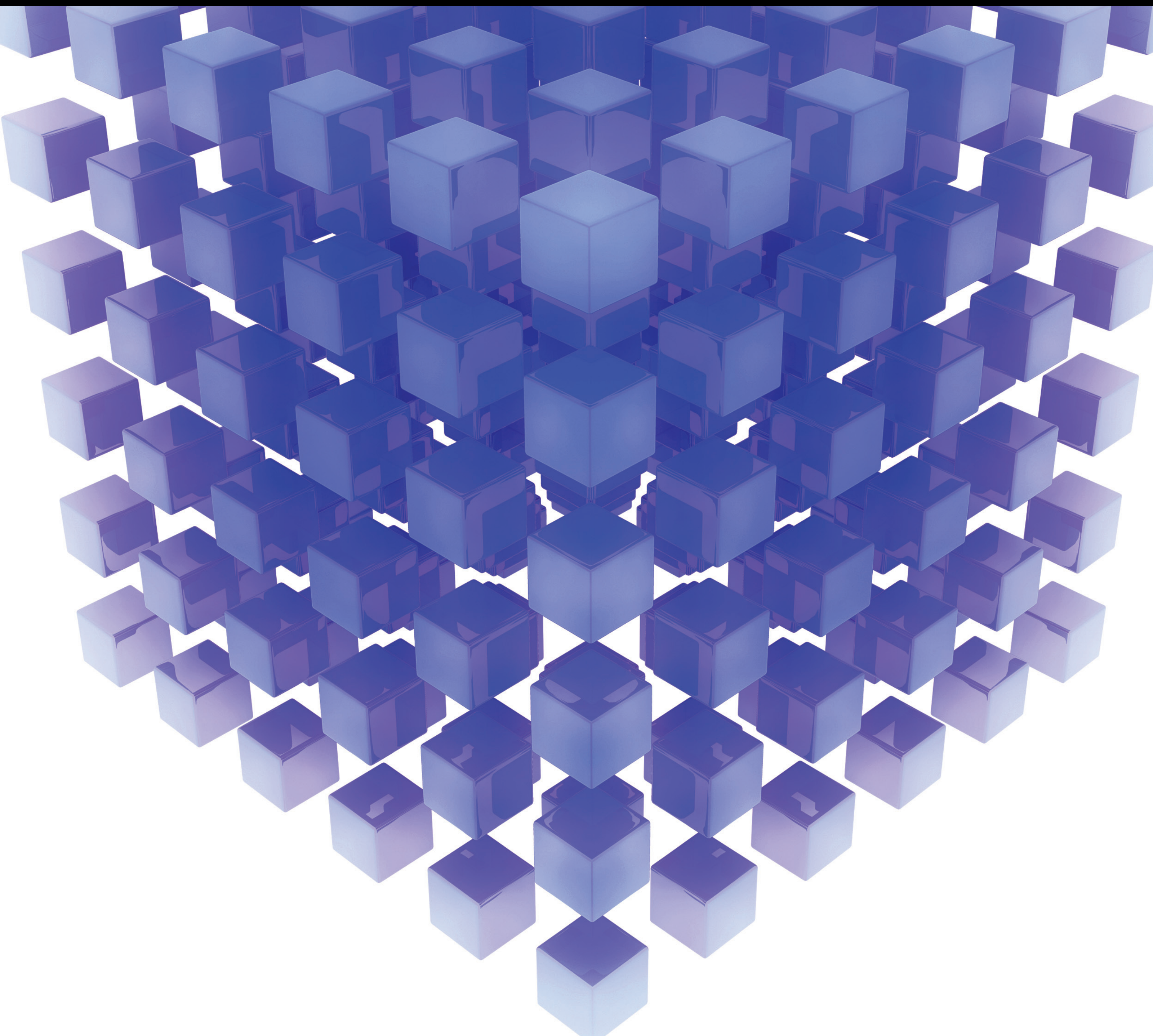
Mathematical Problems in Engineering

Stochastic Systems 2014

Guest Editors: Weihai Zhang, Xuejun Xie, Suiyang Khoo, Guangchen Wang, Wuquan Li, and Ming Gao



Stochastic Systems 2014

Stochastic Systems 2014

Guest Editors: Weihai Zhang, Xuejun Xie, Suiyang Khoo, Guangchen Wang, Wuquan Li, and Ming Gao

Editorial Board

Contents

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 210630, 3 pages
http://dx.doi.org/10.1155/2015/210630

Editorial

Stochastic Systems 2014

Weihai Zhang,[1] Xuejun Xie,[2] Suiyang Khoo,[3] Guangchen Wang,[4] Wuquan Li,[5] and Ming Gao[6]

[1] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China*
[2] *Institute of Automation, Qufu Normal University, Qufu 273165, China*
[3] *School of Engineering, Deakin University, Victoria 3216, Australia*
[4] *School of Control Science and Engineering, Shandong University, Jinan 250061, China*
[5] *School of Mathematics and Information, Ludong University, Yantai 264025, China*
[6] *College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China*

Correspondence should be addressed to Weihai Zhang; w_hzhang@163.com

Received 19 October 2014; Accepted 19 October 2014

Stochastic systems captured by Itô differential equations and stochastic difference equations play a prominent role in modern control theory, which describe the systems disturbed by the randomness in the forms of Brownian motion and white noise. With the development of mathematical finance, network control, biology systems, and multiagent, many challenging stochastic-control problems are springing up, which need to be deeply investigated by means of more advanced theories and tools. To reflect the most recent advances in stochastic systems, we are determined to organize this special issue.

This special issue is focused on the stochastic-control systems and their applications to stability, control, filtering, communication, and fault detection. Topics covered in this issue include (i) stochastic modeling, stability, and stabilization analysis, (ii) stochastic robust/optimal/adaptive control, (iii) stochastic filtering and estimation, (iv) stochastic differential game, and (v) applications of stochastic-control theory to finance, economics, fault detection, and so forth. This special issue has received a total of 82 submitted papers with only 40 papers accepted.

There are 13 manuscripts on the subject "stochastic modeling, stability, and stabilization analysis." In the following, we give a brief summary. The paper entitled "*Discrete-time indefinite stochastic linear quadratic optimal control with second moment constraints*" by W. Zhang and G. Li studies stochastic LQ problem with constraints on the terminal state, where the weighting matrices in the cost functional are allowed to be indefinite. The problem of state-feedback stabilization for a class of stochastic nonlinear systems is investigated by H. Wang et al. in the paper "*Asymptotic stabilization by state feedback for a class of stochastic nonlinear systems with time-varying coefficients*." G. Li and M. Chen investigate the stability and the stabilizability of delayed stochastic systems in "*The stability and stabilization of stochastic delay-time systems*." In "*Integer-valued moving average models with structural changes*," K. Yu et al. present a first order integer-valued moving average model which provides a flexible framework for modeling a wide range of dependence structures. In "*Further results on dynamic additive Hazard rate model*," Z. Zhang and L. Zhang study the dynamic additive hazard rate model and investigate its aging properties for different aging classes. C. Li and J. Duan in "*Impact of correlated noises on additive dynamical systems*" consider Fokker-Planck type equations under the fractional white noise measure. By means of Lyapunov functions, Doob's martingale inequality, and Borel-Cantelli lemma, W. Zhu et al. give some sufficient conditions for the exponential stability in the mean square of a class of stochastic systems in "*Exponential stability of stochastic systems with delay and Poisson jumps*." In "*Stochastic resonance in neuronal network motifs with Ornstein-Uhlenbeck colored noise*," X. Lou considers the effect of the Ornstein-Uhlenbeck colored noise on the stochastic resonance of the feed-forward-loop network motif. In "*Input-to-state stability for a class of switched* *stochastic nonlinear systems by an improved average dwell time method*," the input-to-state stability in the mean property of switched stochastic nonlinear systems is investigated by R. Guo et al. In "*Optimal dividend and capital injection strategies in the Cramér-Lundberg risk model*," Y. Li and G. Liu maximize the discounted dividends payments minus the penalized discounted capital injections. In "*Boundedness of stochastic delay differential systems with impulsive control and impulsive disturbance*," L. Wang et al. derive several sufficient conditions which guarantee the p-moment boundedness of nonlinear impulsive stochastic delay differential systems by using the Lyapunov-Razumikhin method and stochastic analysis techniques. In the paper "*Exponential stability of neutral stochastic functional differential equations with two-time-scale Markovian switching*," J. Hu and Z. Xu develop exponential stability of neutral stochastic equations modeled by a continuous-time Markov chain which has a large state space. Y. Li and Y. Shen discuss the impact of stochastic noise and connection weight matrices uncertainty on global exponential stability of hybrid BAM neural networks with reaction diffusion terms. It is found that the perturbed hybrid BAM neural networks preserve global exponential stability if the intensities of both stochastic noise and the connection weight matrix uncertainty are smaller than the defined upper threshold.

There are 8 contributions closely related to controlled stochastic differential equations. In "*Nonlinear stochastic H_∞ control with Markov jumps and (x, u, v)-dependent noise: finite and infinite horizon cases,*" L. Sheng et al. investigate the H_∞ control problem for nonlinear stochastic Markov jump systems with state, control, and external disturbance-dependent noise. In "*The H_∞ control for bilinear systems with Poisson jumps*," R. Zhang et al. discuss the state feedback H_∞ control problem for bilinear stochastic systems driven by both Brownian motion and Poisson jumps. S. Wang and Z. Wu focus on optimal control derived by forward-backward regime-switching systems with impulse controls in "*Maximum principle for optimal control problems of forward-backward regime-switching systems involving impulse controls*." Maximum principles and verification theorems for optimality are obtained and are used to solve an optimal investment and consumption problem with recursive utility. In "*Mean-field backward stochastic evolution equations in Hilbert spaces and optimal control for BSPDEs*," R. Xu and T. Wu investigate an optimal control problem of backward stochastic partial differential equations. Existence and uniqueness of mild solutions to mean-field backward stochastic evolution equations in Hilbert spaces are proved. In "*Terminal-dependent statistical inference for the FBSDEs models*," Y. Song works out a nonparameter method to estimate parameters of backward stochastic differential equations from noisy data and terminal conditions. In "*Adaptive neural output feedback control of stochastic nonlinear systems with unmodeled dynamics*," X. Xia and T. P. Zhang propose an adaptive neural output feedback control scheme for stochastic systems with unmodeled dynamics and unmeasured states. X. Dai et al. investigate robust stochastic mean-square exponential stabilization and robust H_∞ control for stochastic partial differential time delay systems in "*Robust H_∞ control for linear stochastic partial differential systems with time delay*." Based on the Lyapunov stability theory and stochastic analysis technique, G. Chen et al. establish both delay-independent and delay-dependent dissipativity criteria for nonlinear stochastic delay systems in "*Dissipative delay-feedback control for nonlinear stochastic systems with time-varying delay*."

The subject on stochastic filtering and estimation has occupied 7 contributions. In "*Parallel array bistable stochastic resonance system with independent input and its signal-to-noise ratio improvement*," W. Li et al. discuss the design enhancement of the bistable stochastic resonance performance on sinusoidal signal and Gaussian white noise. A new pruning algorithm for Gaussian mixture PHD filter is proposed by X. Yan in the paper "*Iterative mixture component pruning algorithm for Gaussian mixture PHD filter*," where the pruning algorithm is based on maximizing the posterior probability density of the mixture weights. In "*Covariance-based estimation from multisensor delayed measurements with random parameter matrices and correlated noises*," R. Caballero-Águila et al. address the optimal least-square linear estimation problem for a class of discrete-time multisensor linear stochastic systems subject to randomly delayed measurements with different delay rates. In "*Stochastic signal processing for sound environment system with decibel evaluation and energy observation*," A. Ikuta and H. Orimoto propose a stochastic signal processing method to predict the output response probability distribution of complex sound environment systems. A fusion algorithm based on linear minimum mean-square error estimation is provided by X. Yuan et al. in "*Performance analysis for distributed fusion with different dimensional data*." In "*Two identification methods for dual-rate sampled-data nonlinear output-error systems*," J. Chen and R. Ding present two methods for dual-rate sampled-data nonlinear output-error systems, which can estimate the unknown parameters directly. A particle filter based track-before-detect algorithm is proposed for the monopulse high pulse repetition frequency pulse Doppler radar by F. Cai et al. in "*Dual-channel particle filter based track-before-detect for monopulse radar*."

There are 4 papers that are concerned with stochastic differential games. In "*Linear quadratic nonzero sum differential games with asymmetric information*," D. Chang and H. Xiao consider an LQ nonzero sum stochastic differential game, where the information available to players is asymmetric. In "*Algorithms to solve stochastic H_2/H_∞ control with state-dependent noise*," several algorithms are proposed to solve H_2/H_∞ control problems of stochastic systems by M. Gao et al. X. Chen and Q. Zhu use a maximum principle method to study a partial information nonzero sum differential game of backward stochastic differential equation with jumps in "*Nonzero sum differential game of mean-field BSDEs with jumps under partial information*." Here a feature is that both the game system and the performance functional are of mean-field type. Z. Wu and Q. Zhang's paper "*Backward stochastic H_2/H_∞ control: infinite horizon case*" establishes a necessary and sufficient condition for the existence of H_2/H_∞ control of infinite horizon backward stochastic differential equations.

There are also 8 contributions on applications of stochastic control theory. X. Cao in "*An upper bound of large deviations for capacities*" obtains a type of large deviation principle under the sublinear expectation. Y.-G. Zhang et al. in "*Moving state marine SINS initial alignment based on high degree CKF*" propose a moving state marine initial alignment method for strap-down inertial navigation system. In "*The Gerber-Shiu discounted penalty function of Sparre Andersen risk model with a constant dividend barrier*," Y. Huang and W. Yu construct a new Sparre Andersen risk model with a constant dividend barrier and derive an integrodifferential equation of the Gerber-Shiu discounted penalty function. One paper entitled "*A closed-form solution for robust portfolio selection with worst-case CVaR risk measure*" by L. Tang and A. Ling considers a robust portfolio selection problem with WCCVaR constraint and the corresponding closed-form solution is obtained. In "*Stochastic dominance under the nonlinear expected utilities*," X. Xiao proposes a definition of stochastic dominance under nonlinear expected utilities and gives sufficient conditions on which a random choice X stochastically dominates a random choice Y under the nonlinear expected utilities. In "*Equilibrium model of discrete dynamic supply chain network with random demand and advertisement strategy*," G. Zhang et al. analyze the impact of advertising investment on a discrete dynamic supply chain network which consists of suppliers, manufactures, retailers, and demand markets associated at different tiers under random demand. In "*Research on multiprincipals selecting effective agency mode in the student loan system*," agency modes are discussed by building different principal agent models to solve incentive problems in student loan system. In the paper "*On H_∞ fault estimator design for linear discrete time-varying systems under unreliable communication link*," Y. Li et al. investigate the H_∞ fixed-lag fault estimator design for linear discrete time-varying systems with intermittent measurements, which is described by a Bernoulli distributed random variable.

Acknowledgments

This special issue represents an exciting and insightful snapshot of the current stochastic systems research. As the Lead Guest Editor of this special issue, I would like to express my sincere gratitude to my five coeditors for helping me to undertake this project with a wonderful accomplishment. We are also deeply appreciative of interested authors who have submitted their papers to the special issue.

Weihai Zhang
Xuejun Xie
Suiyang Khoo
Guangchen Wang
Wuquan Li
Ming Gao

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 529862, 13 pages
http://dx.doi.org/10.1155/2015/529862

Research Article

Adaptive Neural Output Feedback Control of Stochastic Nonlinear Systems with Unmodeled Dynamics

Xiaonan Xia and Tianping Zhang

Department of Automation, College of Information Engineering, Yangzhou University, Yangzhou 225127, China

Correspondence should be addressed to Tianping Zhang; tpzhang@yzu.edu.cn

Received 9 April 2014; Revised 17 June 2014; Accepted 4 July 2014

Academic Editor: Ming Gao

An adaptive neural output feedback control scheme is investigated for a class of stochastic nonlinear systems with unmodeled dynamics and unmeasured states. The unmeasured states are estimated by K-filters, and unmodeled dynamics is dealt with by introducing a novel description based on Lyapunov function. The neural networks weight vector used to approximate the black box function is adjusted online. The unknown nonlinear system functions are handled together with some functions resulting from theoretical deduction, and such method effectively reduces the number of adaptive tuning parameters. Using dynamic surface control (DSC) technique, Itô formula, and Chebyshev's inequality, the designed controller can guarantee that all the signals in the closed-loop system are bounded in probability, and the error signals are semiglobally uniformly ultimately bounded in mean square or the sense of four-moment. Simulation results are provided to verify the effectiveness of the proposed approach.

1. Introduction

During the past decades, backstepping in [1] and dynamic surface control (DSC) in [2] have become two most popular methods for adaptive controller design. Many adaptive control schemes based on fuzzy/neural networks have been proposed for uncertain nonlinear systems using backstepping or dynamic surface control method in [3–13]. In the existing literature, three types of uncertainties were commonly considered, which included unknown system functions and parameter uncertainties and unmodeled dynamics. Unmodeled dynamics was dealt with by introducing an available dynamic signal in [3]. In addition, it was handled by a description method of Lyapunov function in [4]. In [4, 5], adaptive tracking control schemes were developed by backstepping and DSC for a class of strict-feedback uncertain nonlinear systems, respectively. In [7–10], adaptive control schemes were presented for a class of pure-feedback nonlinear systems. In [11–13], the adaptive tracking approaches for single-input single-output (SISO) nonlinear systems were extended to uncertain large-scale nonlinear systems.

When system states are assumed to be unmeasurable, output feedback adaptive control based on filters or observers has attracted much attention. In [14], K-filters were firstly proposed, and adaptive output feedback control was developed using K-filters. Inspired by the work in [14], robust adaptive output feedback control schemes were studied for SISO uncertain nonlinear systems in [15, 16]. In [17], combining backstepping technique with small-gain approach, indirect adaptive output feedback fuzzy control was developed. In [18], decentralized adaptive output-feedback control was designed based on high-gain K-filters and dynamic surface control method for a class of uncertain interconnected nonlinear systems.

It is well known that due to the stochastic terms and the extra quadratic variation terms resulting from the Itô differentiation rule, both the structures and the controller design of stochastic systems are commonly more complicated than those of deterministic systems. In the past decade, much effort has focused on the study of adaptive control schemes for uncertain stochastic nonlinear systems and the proof of the control system stability in probability sense. In [19–21], Deng et al. proposed the adaptive control scheme, based on backstepping for stochastic strict feedback or output-feedback nonlinear systems, and introduced a control Lyapunov function formula for stochastic disturbance attenuation earlier. In [22], by employing the stochastic Lyapunov-like theorem, adaptive backstepping state feedback control was developed for a class of stochastic nonlinear systems with unknown backlash-like hysteresis nonlinearities. In [23], the problem of decentralized adaptive output-feedback control was discussed for a class of stochastic nonlinear interconnected systems. In [24, 25], output feedback adaptive fuzzy control approaches were considered using backstepping method for a class of uncertain stochastic nonlinear systems. In [26], by combining stochastic small-gain theorem with backstepping design technique, an adaptive output feedback control scheme was presented for a class of stochastic nonlinear systems with unmodeled dynamics and uncertain nonlinear functions. In [27], a concept of stochastic integral input-to-state stability (SiISS) using Lyapunov function was first introduced, and output feedback control was developed for stochastic nonlinear systems with stochastic inverse dynamics. In [28], two linear output feedback control schemes were studied to make the closed-loop system noise-to-state stable or globally asymptotically stable in probability. In [29], by using the homogeneous domination technique and appropriate Lyapunov functions, an output-feedback stabilizing controller was designed to be globally asymptotically stable in probability. In [30], the small-gain control method was investigated for stochastic nonlinear systems with SiISS inverse dynamics. In [31], based on a reduced-order observer, small-gain type condition on SiISS and stochastic LaSalle theorem, an output feedback controller was developed for stochastic nonlinear systems. In [32], an adaptive output feedback control scheme was investigated by combining K-filters with DSC for a class of stochastic nonlinear systems with dynamic uncertainties and unmeasured states. In [33], adaptive control was developed using the backstepping method for a class of stochastic nonlinear systems with time-varying state delays and unmodeled dynamics.

Motivated by the above-mentioned results [4, 14, 32], in this paper, adaptive neural stochastic output feedback control is developed by combining K-filters with dynamic surface control to guarantee the stability of the closed-loop system. The main contributions of the paper lie in the following.

(i) Adaptive neural output feedback control is developed using K-filters and dynamic surface control for a class of stochastic nonlinear systems with unmodeled dynamics and unmeasured states. The advantage of the design is that once the local system constructed by the filter signals is stabilized, all the signals in the closed-loop system are bounded in probability.

(ii) Unmodeled dynamics is dealt with first by introducing a novel description based on Lyapunove function without using the dynamic signal to handle dynamic uncertainty in [32]. The novel description, which provides an effective method for dealing with unmodeled dynamics in output feedback adaptive controller design, is the development of original idea about handling unmodeled dynamics in [4].

(iii) Utilizing the boundedness of continuous function, the unknown nonlinear system functions are handled together with some functions produced in stability analysis, rather than directly approximated before stability analysis in [6, 8, 9, 11, 12]. Therefore the design effectively reduces the order of filters and the number of adjustable parameters of the whole system, without estimating Ξ in [32].

(iv) Using bounded input bounded output (BIBO) stability and the filter special structure, the stability of the closed-loop system is proved. Therefore, the difficulty, that the transfer function cannot be used in a stochastic system while it was widely used to analyze the boundedness of the K-filters signals in the deterministic systems in [4, 14, 16–18], is solved by the proposed stability analysis approach in this paper.

The rest of the paper is organized as follows. The problem formulation and preliminaries are given in Section 2. The neural filters are designed, and adaptive stochastic output feedback control is developed based on dynamic surface control method. The stability in the closed-loop system in probability sense is analyzed in Section 3. Simulation results are presented to illustrate the effectiveness of the proposed scheme in Section 4. Section 5 contains the conclusions.

2. Problem Statement and Preliminaries

Consider the following uncertain stochastic nonlinear systems with unmodeled dynamics:

$$\begin{aligned}
dz &= q(z,y)\,dt \\
dx_1 &= (x_2 + f_1(y) + \Delta_1(z,y,t))\,dt + g_1^T(y)\,dw \\
&\vdots \\
dx_{\rho-1} &= \left(x_\rho + f_{\rho-1}(y) + \Delta_{\rho-1}(z,y,t)\right)dt \\
&\quad + g_{\rho-1}^T(y)\,dw \\
dx_\rho &= \left(x_{\rho+1} + f_\rho(y) + \Delta_\rho(z,y,t) + b_m\sigma(y)\,u\right)dt \\
&\quad + g_\rho^T(y)\,dw \\
&\vdots \\
dx_{n-1} &= (x_n + f_{n-1}(y) + \Delta_{n-1}(z,y,t) + b_1\sigma(y)\,u)\,dt \\
&\quad + g_{n-1}^T(y)\,dw \\
dx_n &= (f_n(y) + \Delta_n(z,y,t) + b_0\sigma(y)\,u)\,dt + g_n^T(y)\,dw \\
y &= x_1,
\end{aligned} \tag{1}$$

where $x = [x_1, x_2, \ldots, x_n]^T \in R^n$ is the state; $u \in R$ is the input, and $y \in R$ is the output; $\sigma(y) \neq 0$ is a known positive continuous function; $f_i(y)$ is the unknown smooth function; $z \in R^{n_0}$ is the unmodeled dynamics, and $\Delta_i(z,y,t)$ is the unknown smooth nonlinear dynamic disturbance; $b = [b_m, \ldots, b_1, b_0]^T \in R^{m+1}$, $B(s) = b_m s^m + \cdots + b_1 s + b_0$ is a Hurwitz polynomial; $\Delta_i(z,y,t)$ and $q(z,y)$ are the unknown Lipschitz functions; w is an r-dimensional standard Brownian motion defined on the complete probability space (Ω, F, P) with Ω being a sample space, F being a σ field, and P being a probability measure. In this paper, it is assumed that only output y is available for measurement.

The control objective is to design output feedback adaptive control u for system (1) such that the output y follows the specified desired trajectory y_d, and all the signals of the closed-loop system are bounded in probability.

Assumption 1 (see [4]). The unknown nonlinear dynamic disturbances $\Delta_i(z, y, t)$, $i = 1, 2, \ldots, n$, satisfy $|\Delta_i(z, y, t)| \leq \rho_{i1}(|y|) + \rho_{i2}(y)\|z\|$, and $\rho_{i1}(|y|)$ and $\rho_{i2}(y)$ are the unknown nonnegative smooth functions, and $\|\cdot\|$ denotes the Euclidian norm of a vector.

Assumption 2. The system $\dot{z} = q(z, 0, t) - q(0, 0, t)$ is globally exponentially stable when $z = 0$; that is, there exists a Lyapunov function $W(t, z)$ satisfying

$$\begin{gathered} c_1 \|z\|^4 \leq W(z, t) \leq c_2 \|z\|^4, \\ \frac{\partial W}{\partial t}(z, t) + \frac{\partial W}{\partial z}(z, t)\left(q(z, 0, t) - q(0, 0, t)\right) \\ \leq -c_3 \|z\|^4, \\ \left|\frac{\partial W}{\partial z}(z, t)\right| \leq c_4 \|z\|^3, \end{gathered} \tag{2}$$

where c_1, c_2, c_3, c_4 are positive constants, and there exists $c_5 \geq 0$ such that $\|q(0, 0, t)\| \leq c_5$, $\forall t \geq 0$.

Assumption 3. There exists an unknown function ψ_0, and $\psi_0(0) = 0$, such that $\|q(z, y, t) - q(z, 0, t)\| \leq \psi_0(|y|)$ holds.

Assumption 4. The desired trajectory $x_d = [y_d, \dot{y}_d, \ddot{y}_d]^T \in \Omega_d$ is known, where $\Omega_d = \{x_d : y_d^2 + \dot{y}_d^2 + \ddot{y}_d^2 \leq B_0\}$, and B_0 is a known constant.

Assumption 5. There exists a known constant $b_{\max}$ such that the following inequality $0 < |b_m| \leq b_{\max}$ holds.

Remark 6. Assumption 2 is the extension of the description of unmodeled dynamics in [4], and it can effectively deal with unmodeled dynamics in output feedback adaptive controller design. To the best of authors' knowledge, this assumption is first addressed.

Consider the following stochastic nonlinear system:

$$dx = f(t, x)\, dt + h^T(t, x)\, dw, \tag{3}$$

where $x \in R^n$ is the system state, w is an r-dimensional standard Brownian motion, $f : R^+ \times R^n \to R^n$, $h^T : R^+ \times R^n \to R^{n\times r}$ are locally Lipschitz and $f(t, 0)$, $h(t, 0)$ are uniformly ultimately bounded. For any given $V(t, x(t)) \in C^{1,2}$, associated with the stochastic system (3), the infinitesimal generator ℓ is defined as follows:

$$\begin{aligned} \ell V(t, x(t)) = {} & \frac{\partial V(t, x(t))}{\partial t} + \frac{\partial V(t, x(t))}{\partial x^T} f \\ & + \frac{1}{2} \operatorname{tr}\left\{h \frac{\partial^2 V(t, x(t))}{\partial x^T \partial x} h^T\right\}, \end{aligned} \tag{4}$$

where $\operatorname{tr}(A)$ is the trace of a matrix A.

Definition 7 (see [34]). The stochastic process $\{x(t)\}$ is said to be bounded in probability, if $\lim_{c\to\infty} \sup_{0\leq t<\infty} P(|x(t)| > c) = 0$.

Definition 8. The solution $x(t)$ of system (3) is said to be semiglobally uniformly ultimately bounded (SGUUB) in pth moment ($p \geq 1$), if for some compact set $\Omega \subset R^n$ and any initial state $x_0 = x(t_0) \in \Omega$, there exists a constant $\varepsilon > 0$ and a time constant $T = T(\varepsilon, x_0)$ such that $E[\|x(t)\|^p] \leq \varepsilon$ for all $t > t_0 + T$, especially, when $p = 2$, it is usually called SGUUB in mean square.

Lemma 9 (see [32]). *For any stochastic process $\{\xi(t)\}$, if there exists a positive integer p and a positive constant C_0 such that $E|\xi(t)|^p \leq C_0$, $\forall t \geq 0$, then $\{\xi(t)\}$ is bounded in probability.*

Lemma 10 (see [21]). *Consider system (3) and suppose that there exists a C^2 function $V(t, x(t))$: $R^n \times R \to R^+$, two constants $c_1 > 0$, $c_2 \geq 0$, class κ_∞ functions μ_1, μ_2 such that*

$$\begin{gathered} \mu_1(\|x\|) \leq V(t, x) \leq \mu_2(\|x\|), \\ \ell V \leq -c_1 V + c_2 \end{gathered} \tag{5}$$

for all $x \in R^n$ and $t > t_0$. Then, (i) for any initial state $x_0 \in R^n$, there exists a unique strong solution $x(t)$ for system (3); (ii) the solution $x(t)$ of system (3) is bounded in probability; (iii) $E[V(t_0, x)] \leq V(t_0, x_0)e^{-c_1 t} + c_2/c_1$, $\forall t \geq t_0$.

In order to design filters and observer, (1) can be rewritten as follows:

$$\begin{aligned} \dot{z} &= q(z, y), \\ dx &= \left(Ax + f(y) + F^T(y, u)\, b + \Delta\right) dt + g^T(y)\, dw, \\ y &= e_1^T x, \end{aligned} \tag{6}$$

where

$$A = \begin{bmatrix} 0 & I_{n-1} \\ 0 & 0 \end{bmatrix},$$

$$f(y) = \begin{bmatrix} f_1(y) \\ \vdots \\ f_n(y) \end{bmatrix},$$

$$\Delta(z,y,t) = \begin{bmatrix} \Delta_1(z,y,t) \\ \vdots \\ \Delta_n(z,y,t) \end{bmatrix},$$
$$e_1 = [1,0,\ldots,0]^T,$$
$$F^T(y,u) = \begin{bmatrix} 0_{(\rho-1)\times(m+1)} \\ I_{m+1} \end{bmatrix} \sigma(y)u. \tag{7}$$

3. Adaptive Robust Controller Design and Stability Analysis

3.1. Neural Filters and Controller Design. In order to estimate the state x, we introduce the following filters:

$$\begin{aligned} \dot{\xi} &= A_0\xi + Ly, \quad \xi \in R^n, \\ \dot{\Omega}^T &= A_0\Omega^T + F^T(y,u), \quad \Omega^T \in R^{n\times(m+1)}, \end{aligned} \tag{8}$$

where $A_0 = A - Le_1^T$, $L = [l_1,\ldots,l_n]^T$, A_0 is a Hurwitz matrix; that is

$$\begin{aligned} PA_0 + A_0^TP &= -hI, \\ P = P^T &> 0, \end{aligned} \tag{9}$$

where $h > 0$ is a design constant.

Define the state estimate as follows:

$$\hat{x} = \xi + \Omega^T b \tag{10}$$

The observer error is defined as $\varepsilon = x - \hat{x}$. Thus

$$x = \xi + \Omega^T b + \varepsilon, \tag{11}$$

$$d\varepsilon = (A_0\varepsilon + f(y) + \Delta)\,dt + g^T(y)\,dw. \tag{12}$$

Denote the columns of Ω^T as follows:

$$\Omega^T = [v_m,\ldots,v_1,v_0], \quad \Omega^T \in R^{n\times(m+1)}. \tag{13}$$

Inspired by the work in [14], the filters are designed as follows:

$$\begin{aligned} \dot{\xi} &= A_0\xi + Ly, \quad \xi \in R^n, \\ \dot{\lambda} &= A_0\lambda + e_n\sigma(y)u, \quad \lambda \in R^n, \\ \dot{v}_j &= A_0v_j + e_{n-j}\sigma(y)u, \quad v_j \in R^n, \ j = 0,1,\ldots,m. \end{aligned} \tag{14}$$

It is easy to show that

$$A_0^j e_n = e_{n-j}, \quad j = 0,1,\ldots,m, \tag{15}$$

$$v_j = A_0^j\lambda, \quad j = 0,1,\ldots,m, \tag{16}$$

where e_i denotes n dimensional vector with the ith element being one and other elements being all zeros, $i = 1,\ldots,n$.

Let $v_{i,j}$ be the jth element of the vector v_i and λ_l the lth element of the vector λ, respectively. From [14], we know

$$v_{i,j} = [* \ \cdots \ * \ 1] \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{i+j} \end{bmatrix},$$
$$j = 1,\ldots,\rho; \quad i = 0,1,\ldots,m; \ \lambda_l = 0, \ l > n. \tag{17}$$

According to (11), we get

$$\begin{aligned} x_2 &= \xi_2 + \omega^T b + \varepsilon_2 = \xi_2 + [v_{m,2},\ldots,v_{1,2},v_{0,2}]\,b + \varepsilon_2 \\ &= b_m v_{m,2} + \xi_2 + \overline{\omega}^T\overline{b} + \varepsilon_2, \end{aligned} \tag{18}$$

where ω^T denotes the second row of the matrix Ω^T, ξ_2 denotes the second element of the vector ξ, and ε_2 is the second element of ε.

Substituting (18) into (1), it yields

$$\begin{aligned} dy &= \left(b_m v_{m,2} + \xi_2 + \overline{\omega}^T\overline{b} + \varepsilon_2 + f_1(y) + \Delta_1(z,y,t)\right)dt \\ &\quad + g_1^T(y)\,dw, \end{aligned} \tag{19}$$

where

$$\begin{aligned} \omega^T &= [v_{m,2}, v_{m-1,2},\ldots,v_{1,2},v_{0,2}], \\ \overline{\omega}^T &= [v_{m-1,2},\ldots,v_{1,2},v_{0,2}], \\ \overline{b}^T &= [b_{m-1},\ldots,b_1,b_0]. \end{aligned} \tag{20}$$

In view of (19) and (14), the system used to design adaptive output feedback DSC in next section is addressed as follows:

$$\begin{aligned} dy &= \left(b_m v_{m,2} + \xi_2 + \overline{\omega}^T\overline{b} + \varepsilon_2 + f_1(y) + \Delta_1\right)dt \\ &\quad + g_1^T(y)\,dw, \\ \dot{v}_{m,i} &= v_{m,i+1} - l_i v_{m,1}, \quad i = 2,\ldots,\rho-1, \\ \dot{v}_{m,\rho} &= \sigma(y)u + v_{m,\rho+1} - l_\rho v_{m,1}. \end{aligned} \tag{21}$$

3.2. Stochastic Adaptive Dynamic Surface Controller Design. In this subsection, according to (21) and by using dynamic surface control method, we propose an output feedback stochastic adaptive tracking control scheme. Similar to backstepping, the whole design needs ρ steps.

For convenience, some notations are presented below. $\overline{s}_i = [s_1,\ldots,s_i]^T$, $\overline{y}_j = [y_2,\ldots,y_j]^T$, where s_i, y_j will be given in the controller design later, $i = 1,2,\ldots,\rho$, $j = 2,\ldots,\rho$. $y_j = \omega_j - \alpha_{j-1}$, $j = 2,\ldots,\rho$, ω_j is the output of a first-order filter with α_{i-1} as the input, and α_{i-1} is an intermediate control which will be developed for the corresponding $(i-1)$th subsystem.

Define some Lyapunov functions as follows:

$$V_\varepsilon = \varepsilon^T P\varepsilon,$$
$$V_W = \frac{1}{\lambda_0} W(z,t), \quad \lambda_0 > 0, \tag{22}$$
$$V_{s_i} = \frac{1}{4}s_i^4,$$
$$V_{sW\varepsilon} = V_{s_1} + V_\varepsilon + V_W,$$

where $W(z,t)$ is given in Assumption 2.

Using Young's inequality, the infinitesimal generator of V_ε satisfies

$$\begin{aligned}\ell V_\varepsilon &= \varepsilon^T\left(PA_0 + A_0^T P\right)\varepsilon + 2\varepsilon^T Pf(y) + 2\varepsilon^T P\Delta \\ &\quad + \operatorname{tr}\left(g(y)Pg^T(y)\right) \\ &\le \varepsilon^T\left(PA_0 + A_0^T P\right)\varepsilon + 2\varepsilon^T\varepsilon + \|P\|^2\|f(y)\|^2 \\ &\quad + \|P\|^2\|\Delta\|^2 + \operatorname{tr}\left(g(y)Pg^T(y)\right)\end{aligned} \tag{23}$$

According to Assumption 1 and by using Young's inequality, we obtain

$$\begin{aligned}\ell V_\varepsilon &\le -(h-2)\varepsilon^T\varepsilon + \sum_{j=1}^n \|P\|^2 f_j^2(y) \\ &\quad + \sum_{j=1}^n \|P\|^2\left(\rho_{j1}(|y|) + \rho_{j2}(y)\|z\|\right)^2 \\ &\quad + \operatorname{tr}\left(g(y)Pg^T(y)\right) \\ &\le -(h-2)\varepsilon^T\varepsilon + \sum_{j=1}^n \|P\|^2 f_j^2(y) \\ &\quad + \sum_{j=1}^n 2\|P\|^2\left(\rho_{j1}^2(|y|) + \rho_{j2}^2(y)\|z\|^2\right) \\ &\quad + \operatorname{tr}\left(g(y)Pg^T(y)\right) \\ &\le -(h-2)\varepsilon^T\varepsilon + \sum_{j=1}^n \|P\|^2 f_j^2(y) \\ &\quad + \sum_{j=1}^n 2\|P\|^2\rho_{j1}^2(|y|) \\ &\quad + \frac{16n\lambda_0}{c_3}\sum_{j=1}^n \|P\|^4\rho_{j2}^4(y) + \frac{c_3}{16\lambda_0}\|z\|^4 \\ &\quad + \operatorname{tr}\left(g(y)Pg^T(y)\right)\end{aligned} \tag{24}$$

According to Assumptions 2 and 3, using Young's inequality, we get

$$\begin{aligned}\dot V_W &= \frac{1}{\lambda_0}\left(\frac{\partial W}{\partial z}(z,t)\dot z + \frac{\partial W}{\partial t}(z,t)\right) \\ &\le -\frac{5c_3}{8\lambda_0}\|z\|^4 + \frac{16c_4^4}{\lambda_0 c_3^3}\psi_0^4(|y|) + \frac{16c_4^4c_5^4}{\lambda_0 c_3^3}.\end{aligned} \tag{25}$$

Step 1. Let $\omega_1 = y_d$. Define the first dynamic surface as follows:

$$s_1 = x_1 - \omega_1. \tag{26}$$

Using the first equation of (21), we obtain

$$\begin{aligned}ds_1 &= \left(b_m v_{m,2} + \xi_2 + \overline{\omega}^T\overline{b} + \varepsilon_2 + f_1(y) + \Delta_1 - \dot y_d\right)dt \\ &\quad + g_1^T(y)\,dw.\end{aligned} \tag{27}$$

Choose the virtual control law α_1 as follows:

$$\alpha_1 = \frac{\widehat{b}_m}{\widehat{b}_m^2 + \beta}\left(-k_1 s_1 - \overline{\omega}^T\widehat{\overline{b}} - \xi_2 - s_1^3\widehat{\theta}_1^T\psi_1(X)\right), \tag{28}$$

where $\beta > 0$, $k_1 > 0$ are design constants, $\widehat{\theta}_1$, $\widehat{b}_m$, $\widehat{\overline{b}}$ are the estimates of θ_1, b_m, $\overline{b}$ at time t, respectively, and $\widetilde{b}_m = b_m - \widehat{b}_m$, $\widetilde{\theta}_1 = \theta_1 - \widehat{\theta}_1$, $\widetilde{\overline{b}} = \overline{b} - \widehat{\overline{b}}$, θ_1 and $\psi_1(X)$ will be given later. Consider

$$\begin{aligned}\ell\alpha_1 &= \frac{\partial\alpha_1}{\partial y}\left(b_m v_{m,2} + \xi_2 + \overline{\omega}^T\overline{b} + \varepsilon_2 + f_1(y) + \Delta_1(z,y,t)\right) \\ &\quad + \frac{\partial\alpha_1}{\partial\widehat{b}_m}\dot{\widehat{b}}_m + \frac{\partial\alpha_1}{\partial\xi^T}\dot\xi + \frac{\partial\alpha_1}{\partial\widehat{\theta}_1^T}\dot{\widehat{\theta}}_1 + \frac{\partial\alpha_1}{\partial y_d}\dot y_d \\ &\quad + \frac{\partial\alpha_1}{\partial\dot y_d}\ddot y_d + \frac{1}{2}\frac{\partial^2\alpha_1}{\partial y^2}g_1^T(y)g_1(y), \\ d\alpha_1 &= \ell\alpha_1 dt + \frac{\partial\alpha_1}{\partial y}g_1^T(y)\,dw.\end{aligned} \tag{29}$$

Therefore, we have

$$\begin{aligned}&\ell V_{s_1} \\ &\quad = s_1^3\left(b_m v_{m,2} + \xi_2 + \overline{\omega}^T\overline{b} + \varepsilon_2 + f_1(y) + \Delta_1 - \dot y_d\right) \\ &\qquad + \frac{3}{2}s_1^2\|g_1(y)\|^2.\end{aligned} \tag{30}$$

A first-order filter with α_1 as the input is designed as follows:

$$\tau_2\dot\omega_2 + \omega_2 = \alpha_1, \quad \omega_2(0) = \alpha_1(0). \tag{31}$$

Let $y_2 = \omega_2 - \alpha_1$; thus, $\dot{\omega}_2 = -y_2/\tau_2$. Since $v_{m,2} = s_2 + y_2 + \alpha_1$, using Young's inequality, it yields

$$\begin{aligned}\ell V_{s_1} \leq & -\left(k_1 - \frac{3}{2}b_{\max}\right)s_1^4 + \widetilde{b}_m s_1^3\alpha_1 + \frac{b_{\max}}{4}s_2^4 \\ & + \frac{b_{\max}}{4}y_2^4 + s_1^3\overline{\omega}^T\widetilde{\overline{b}} - s_1^3\widehat{\theta}_1^T\psi_1(X) \\ & - \frac{s_1^3\beta}{\widehat{b}_m^2 + \beta}\left(-k_1 s_1 - \overline{\omega}^T\widehat{\overline{b}} - \xi_2 - s_1^3\widehat{\theta}_1^T\psi_1(X)\right) \\ & + s_1^3\varepsilon_2 + s_1^3 f_1(y) + s_1^3\Delta_1 - s_1^3\dot{y}_d \\ & + \frac{3}{2}s_1^2\|g_1(y)\|^2.\end{aligned} \tag{32}$$

From Assumption 1, we obtain

$$\begin{aligned}\left|s_1^3\right||\Delta_1| &\leq \left|s_1^3\right|\rho_{11}(|y|) + \left|s_1^3\right|\rho_{12}(y)\|z\| \\ &\leq \frac{3}{4}s_1^4 + \frac{1}{4}\rho_{11}^4(|y|) + \frac{3}{2}\sqrt[3]{\frac{\lambda_0}{2c_3}}s_1^4\rho_{12}^{4/3}(y) \\ &\quad + \frac{c_3}{16\lambda_0}\|z\|^4.\end{aligned} \tag{33}$$

In view of (24), (25), (32), and (33) and by using Young's inequality, we obtain

$$\begin{aligned}\ell V_{sW\varepsilon} \leq & -(h-2)\varepsilon^T\varepsilon - \left(k_1 - \frac{3}{2}b_{\max} - \frac{3}{4}\right)s_1^4 \\ & - \frac{c_3}{2\lambda_0}\|z\|^4 + \widetilde{b}_m s_1^3\alpha_1 + s_1^3\overline{\omega}^T\widetilde{\overline{b}} + \frac{b_{\max}}{4}s_2^4 \\ & + \frac{b_{\max}}{4}y_2^4 + \left|s_1^3\right|S + s_1^3 H_1(X) + Q(y) + \frac{1}{4}\varepsilon_2^2 \\ & - s_1^3\widehat{\theta}_1^T\psi_1(X) + \frac{16c_4^4c_5^4}{\lambda_0 c_3^3} + 1,\end{aligned} \tag{34}$$

where

$$\begin{aligned}Q(y) &= \sum_{j=1}^n \|P\|^2 f_j^2(y) + \sum_{j=1}^n 2\|P\|^2\rho_{j1}^2(|y|) \\ &\quad + \frac{16n\lambda_0}{c_3}\sum_{j=1}^n \|P\|^4\rho_{j2}^4(y) + \frac{16c_4^4}{\lambda_0 c_3^3}\psi_0^4(|y|) \\ &\quad + \frac{1}{4}\rho_{11}^4(|y|) + \operatorname{tr}\left(g(y)Pg^T(y)\right), \\ H_1(X) &= \frac{3}{2}\sqrt[3]{\frac{\lambda_0}{2c_3}}s_1\rho_{12}^{4/3}(y) - \dot{y}_d + \frac{9}{16}s_1\|g_1(y)\|^4 + s_1^3, \\ X &= \left[s_1, y_d, \dot{y}_d\right]^T \in R^3.\end{aligned} \tag{35}$$

$S(s_1, \widehat{b}_m, \widehat{\overline{b}}, \widehat{\theta}_1, \xi, \overline{\lambda}_{m+2}, y_d)$ is a nonnegative continuous function, and

$$\begin{aligned}&\left|f_1(y) - \frac{\beta}{\widehat{b}_m^2 + \beta}\left(-k_1 s_1 - \overline{\omega}^T\widehat{\overline{b}} - \xi_2 - s_1^3\widehat{\theta}_1^T\psi_1(x)\right)\right| \\ &\quad \leq S\left(s_1, \widehat{b}_m, \widehat{\overline{b}}, \widehat{\theta}_1, \xi, \overline{\lambda}_{m+2}, y_d\right),\end{aligned} \tag{36}$$

where $\overline{\lambda}_{m+2} = [\lambda_1, \ldots, \lambda_{m+2}]^T$.

Let $\Omega_X = \{X \mid \|X\| \leq M_X\} \subset R^3$ be a given compact set with $M_X > 0$ being a design constant, and let $\theta_1^T\phi_1(X)$ be the approximation of the radial basis function neural networks on the compact set Ω_X to $H_1(X)$. Then, we have $H_1(X) = \theta_1^T\psi_1(X) + B_1(X)$, where $B_1(X)$ denotes the approximation error and $\psi_1(X) = [\psi_{11}(X), \ldots, \psi_{1M_1}(X)]^T \in R^{M_1}$ denotes the basis function vector with $\psi_{1j}(X)$ being chosen as the commonly used Gaussian functions, which have the form

$$\psi_{1j}(X) = \exp\left[-\frac{\left\|X - \mu_{1j}\right\|^2}{b_{1j}^2}\right] \tag{37}$$

$j = 1, \ldots, M_1$, and μ_{1j} is the center of the receptive field and b_{1j} is the width of the Gaussian function; θ_1 is an adjustable parameter vector.

According to (34) and by using Young's inequality, it yields

$$\begin{aligned}\ell V_{sW\varepsilon} \leq & -\left(h - \frac{9}{4}\right)\varepsilon^T\varepsilon - \left(k_1 - \frac{3}{2}b_{\max} - \frac{3}{2}\right)s_1^4 - \frac{c_3}{2\lambda_0}\|z\|^4 \\ & + \widetilde{b}_m s_1^3\alpha_1 + s_1^3\overline{\omega}^T\widetilde{\overline{b}} + \frac{b_{\max}}{4}s_2^4 + \frac{b_{\max}}{4}y_2^4 \\ & + \frac{1}{4}S^4 + s_1^3\left(\theta_1^T\psi_1(X) + B_1(X)\right) + Q(y) \\ & - s_1^3\widehat{\theta}_1^T\psi_1(X) + \frac{16c_4^4c_5^4}{\lambda_0 c_3^3} + 1.\end{aligned} \tag{38}$$

There exists a nonnegative continuous function $\kappa(s_1, y_d, \dot{y}_d)$ satisfying

$$\left|B_1(X)\right| \leq \kappa\left(s_1, y_d, \dot{y}_d\right). \tag{39}$$

Using Young's inequality, we have

$$\begin{aligned}\ell V_{sW\varepsilon} \leq & -\left(h - \frac{9}{4}\right)\varepsilon^T\varepsilon - \left(k_1 - \frac{3}{2}b_{\max} - \frac{9}{4}\right)s_1^4 \\ & - \frac{c_3}{2\lambda_0}\|z\|^4 + \widetilde{b}_m s_1^3\alpha_1 + s_1^3\overline{\omega}^T\widetilde{\overline{b}} + \frac{b_{\max}}{4}s_2^4 \\ & + \frac{b_{\max}}{4}y_2^4 + \frac{1}{4}S^4 + Q(y) + s_1^3\widetilde{\theta}_1^T\psi_1(X) \\ & + \frac{1}{4}\kappa^4 + C_0,\end{aligned} \tag{40}$$

where $C_0 = 16c_4^4c_5^4/\lambda_0 c_3^3 + 1$.

Step i ($2 \le i \le \rho - 1$). Define the ith dynamic surface $s_i = v_{m,i} - \omega_i$, thus

$$\dot{s}_i = v_{m,i+1} - l_i v_{m,1} - \dot{\omega}_i. \quad (41)$$

Select the virtual control law α_i as follows:

$$\begin{aligned} \alpha_i &= -k_i s_i + l_i v_{m,1} + \dot{\omega}_i, \\ \ell\alpha_i &= -k_i\left(v_{m,i+1} - l_i v_{m,1} - \dot{\omega}_i\right) + l_i \dot{v}_{m,1} - \frac{\ell y_i}{\tau_i}, \\ d\alpha_i &= \ell\alpha_i dt + \frac{\partial\alpha_i}{\partial y} g_1^T(y)\, dw. \end{aligned} \quad (42)$$

A first-order filter with the input α_i is designed as follows:

$$\tau_{i+1}\dot{\omega}_{i+1} + \omega_{i+1} = \alpha_i, \quad \omega_{i+1}(0) = \alpha_i(0), \quad (43)$$

where $\tau_{i+1} > 0$ is a design constant.

Let $y_{i+1} = \omega_{i+1} - \alpha_i$. Then $\dot{\omega}_{i+1} = -y_{i+1}/\tau_{i+1}$. Noting $v_{m,i+1} = s_{i+1} + y_{i+1} + \alpha_i$, in view of (41) and (42), we obtain

$$\begin{aligned} \ell V_{s_i} &= s_i^3 \dot{s}_i = s_i^3\left(v_{m,i+1} - l_i v_{m,1} - \dot{\omega}_i\right) \\ &= -k_i s_i^4 + s_i^3 s_{i+1} + s_i^3 y_{i+1} \\ &\le -\left(k_i - \frac{3}{2}\right) s_i^4 + \frac{1}{4} s_{i+1}^4 + \frac{1}{4} y_{i+1}^4. \end{aligned} \quad (44)$$

Step ρ. The control law will be determined in this step. Define the ρth dynamic surface as $s_\rho = v_{m,\rho} - \omega_\rho$. The derivative of s_ρ is

$$\dot{s}_\rho = \sigma(y) u + v_{m,\rho+1} - l_\rho v_{m,1} - \dot{\omega}_\rho. \quad (45)$$

Choose the control law as follows:

$$u = \frac{\left(-k_\rho s_\rho - v_{m,\rho+1} + l_\rho v_{m,1} + \dot{\omega}_\rho\right)}{\sigma(y)}. \quad (46)$$

In view of (45) and (46), we have

$$\ell V_{s_\rho} = s_\rho^3 \dot{s}_\rho = -k_\rho s_\rho^4. \quad (47)$$

The parameters $\widehat{\theta}_1$, $\widehat{b}_m$, and $\widehat{\overline{b}}$ are updated as follows:

$$\dot{\widehat{\theta}}_1 = \gamma_1\left(s_1^3 \psi_1(X) - \sigma_1 \widehat{\theta}_1\right), \quad (48)$$

$$\begin{aligned} \dot{\widehat{b}}_m &= \gamma_2\left(s_1^3 \alpha_1 - \sigma_2 \widehat{b}_m\right), \\ \dot{\widehat{\overline{b}}} &= \gamma_3\left(s_1^3 \overline{\omega} - \sigma_3 \widehat{\overline{b}}\right), \end{aligned} \quad (49)$$

where $\gamma_1, \gamma_2, \gamma_3, \sigma_1, \sigma_2, \sigma_3$ are the design constants.

3.3. Stability Analysis of Adaptive Control System. In this subsection, we will discuss the stability analysis of the closed-loop system. Firstly we define some Lyapunov functions and compact sets as follows:

$$\begin{aligned} V_1 &= \frac{1}{2} s_1^4 + 2V_\varepsilon + \frac{1}{\gamma_1}\widetilde{\theta}_1^T \widetilde{\theta}_1 + \frac{1}{\gamma_2}\widetilde{b}_m^2 + \frac{1}{\gamma_3}\widetilde{\overline{b}}^T \widetilde{\overline{b}} \\ &\quad + \frac{c_3}{\lambda_0}\|z\|^4, \\ V_i &= \sum_{j=1}^{i}\frac{1}{2} s_j^4 + 2V_\varepsilon + \sum_{j=2}^{i}\frac{1}{2} y_j^4 + \frac{1}{\gamma_1}\widetilde{\theta}_1^T \widetilde{\theta}_1 + \frac{1}{\gamma_2}\widetilde{b}_m^2 \\ &\quad + \frac{1}{\gamma_3}\widetilde{\overline{b}}^T \widetilde{\overline{b}} + \frac{c_3}{\lambda_0}\|z\|^4, \quad i = 2, \ldots, \rho, \\ \Omega_1 &= \left\{\left(s_1, \varepsilon, \widehat{\theta}_1, \widehat{b}_m, \widehat{\overline{b}}, \|z\|\right) : V_1 \le p\right\} \subset R^{p_1}, \\ \Omega_i &= \left\{\left(\overline{s}_i, \overline{y}_i, \varepsilon, \widehat{\theta}_1, \widehat{b}_m, \widehat{\overline{b}}, \|z\|\right) : V_i \le p\right\} \subset R^{p_i}, \end{aligned} \quad (50)$$

where $i = 2, \ldots, \rho$, $p > 0$ is a design constant, $p_i = 2i + M_1 + n + m + 1$. It is easy to know that $\Omega_1 \times R^{p_\rho - p_1} \supset \Omega_2 \times R^{p_\rho - p_2} \supset \cdots \supset \Omega_{\rho-1} \times R^{p_\rho - p_{\rho-1}} \supset \Omega_\rho$.

According to $y_2 = \omega_2 - \alpha_1$, we obtain

$$\begin{aligned} \ell y_2 &= \dot{\omega}_2 - \ell\alpha_1 = -\frac{y_2}{\tau_2} - \ell\alpha_1, \\ dy_2 &= \ell y_2 dt - \frac{\partial\alpha_1}{\partial y} g_1^T(y)\, dw. \end{aligned} \quad (51)$$

From (4), we obtain

$$\begin{aligned} \ell\left(\frac{1}{4} y_2^4\right) &= y_2^3 \ell y_2 + \frac{3}{2} y_2^2 \left(\frac{\partial\alpha_1}{\partial y}\right)^2 \|g_1(y)\|^2 \\ &= -\frac{y_2^4}{\tau_2} - y_2^3 \ell\alpha_1 + \frac{3}{2} y_2^2 \left(\frac{\partial\alpha_1}{\partial y}\right)^2 \|g_1(y)\|^2. \end{aligned} \quad (52)$$

There exist two nonnegative continuous functions $\eta_2(\overline{s}_2, y_2, \widehat{\theta}_1, \widehat{b}_m, \widehat{\overline{b}}, \xi, \overline{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d)$ and $\zeta_2(\overline{s}_2, y_2, \widehat{\theta}_1, \widehat{b}_m, \widehat{\overline{b}}, \xi, \overline{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d)$ such that

$$\begin{aligned} &\left|\ell y_2 + \frac{y_2}{\tau_2}\right| \\ &\le \eta_2\left(\overline{s}_2, y_2, \widehat{\theta}_1, \widehat{b}_m, \widehat{\overline{b}}, \xi, \overline{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d\right), \end{aligned} \quad (53)$$

$$\begin{aligned} &\frac{3}{2} y_2^2 \left(\frac{\partial\alpha_1}{\partial y}\right)^2 \|g_1(y)\|^2 \\ &\le \zeta_2\left(\overline{s}_2, y_2, \widehat{\theta}_1, \widehat{b}_m, \widehat{\overline{b}}, \xi, \overline{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d\right). \end{aligned} \quad (54)$$

From (53), we have

$$\begin{aligned} y_2^3 \ell y_2 &\leq -\frac{y_2^4}{\tau_2} \\ &\quad + \left|y_2^3\right| \eta_2 \left(\bar{s}_2, y_2, \hat{\theta}_1, \hat{b}_m, \hat{\bar{b}}, \xi, \bar{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d\right) \\ &\leq -\frac{y_2^4}{\tau_2} + \frac{3}{4} y_2^4 + \frac{1}{4} \eta_2^4. \end{aligned} \tag{55}$$

From (52), (54), and (55), we obtain

$$\ell\left(\frac{1}{4} y_2^4\right) \leq -\frac{y_2^4}{\tau_2} + \frac{3}{4} y_2^4 + \frac{1}{4} \eta_2^4 + \zeta_2. \tag{56}$$

The infinitesimal generator of y_{i+1} is

$$\begin{aligned} \ell y_{i+1} &= -\frac{y_{i+1}}{\tau_{i+1}} - \ell \alpha_i, \\ y_{i+1}^3 \ell y_{i+1} &= -\frac{y_{i+1}^4}{\tau_{i+1}} - y_{i+1}^3 \ell \alpha_i, \end{aligned} \tag{57}$$

$$\begin{aligned} \ell\left(\frac{y_{i+1}^4}{4}\right) &= -\frac{y_{i+1}^4}{\tau_{i+1}} - y_{i+1}^3 \ell \alpha_i \\ &\quad + \frac{3}{2} y_{i+1}^2 \left(\frac{\partial \alpha_i}{\partial y}\right)^2 \left\|g_1(y)\right\|^2. \end{aligned} \tag{58}$$

There exist two nonnegative continuous functions $\eta_{i+1}(\bar{s}_{i+1}, \bar{y}_{i+1}, \hat{\theta}_1, \hat{b}_m, \hat{\bar{b}}, \xi, \bar{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d)$ and $\zeta_{i+1}(\bar{s}_{i+1}, \bar{y}_{i+1}, \hat{\theta}_1, \hat{b}_m, \hat{\bar{b}}, \xi, \lambda_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d)$ such that the following inequalities hold:

$$\begin{aligned} &\left|\ell y_{i+1} + \frac{y_{i+1}}{\tau_{i+1}}\right| \\ &\leq \eta_{i+1}\left(\bar{s}_{i+1}, \bar{y}_{i+1}, \hat{\theta}_1, \hat{b}_m, \hat{\bar{b}}, \xi, \bar{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d\right), \end{aligned} \tag{59}$$

$$\begin{aligned} &\frac{3}{2} y_{i+1}^2 \left(\frac{\alpha_i}{\partial y}\right)^2 \left\|g_1(y)\right\|^2 \\ &\leq \zeta_{i+1}\left(\bar{s}_{i+1}, \bar{y}_{i+1}, \hat{\theta}_1, \hat{b}_m, \hat{\bar{b}}, \xi, \bar{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d\right). \end{aligned} \tag{60}$$

From (59), we obtain

$$\begin{aligned} &y_{i+1}^3 \ell y_{i+1} \\ &\leq -\frac{y_{i+1}^4}{\tau_{i+1}} \\ &\quad + \left|y_{i+1}^3\right| \eta_{i+1}\left(\bar{s}_{i+1}, \bar{y}_{i+1}, \hat{\theta}_1, \hat{b}_m, \hat{\bar{b}}, \xi, \bar{\lambda}_{m+2}, \varepsilon_2, y_d, \dot{y}_d, \ddot{y}_d\right) \\ &\leq -\frac{y_{i+1}^4}{\tau_{i+1}} + \frac{3 y_{i+1}^4}{4} + \frac{\eta_{i+1}^4}{4}. \end{aligned} \tag{61}$$

From (58), (60), and (61), we obtain

$$\ell\left(\frac{1}{4} y_{i+1}^4\right) \leq -\frac{y_{i+1}^4}{\tau_{i+1}} + \frac{3}{4} y_{i+1}^4 + \frac{1}{4} \eta_{i+1}^4 + \zeta_{i+1}. \tag{62}$$

The continuous function $S(\cdot)$ on the compact set $\Omega_d \times \Omega_1$ has a maximum $M(p)$, which depends on the constant p, and $\kappa(s_1, y_d, \dot{y}_d)$ on the compact set $\Omega_d \times \Omega_1$ has a maximum $N_0(p)$, $\eta_{i+1}(\cdot)$ and $\zeta_{i+1}(\cdot)$ on the compact set $\Omega_d \times \Omega_{i+1}$ have the maximum $N_{i+1}(p)$ and $C_{i+1}(p)$ when $\xi, \bar{\lambda}_{m+2}$ are bounded.

Theorem 11. *Consider the closed-loop system consisting of the plant (1) under Assumptions 1–5, the controller (46), and the adaptation laws (48) and (49). For any bounded initial conditions, there exist constants k_i, τ_i, h, γ_1, γ_2, γ_3, σ_1, σ_2, σ_3 satisfying $V(0) \leq c$, such that all of the signals in the closed-loop system are bounded in probability, and $s_1, \ldots, s_\rho$, $y_2, \ldots, y_\rho$ are SGUUB in four-moment, $\tilde{\theta}_1$, $\tilde{b}_m$, $\tilde{\bar{b}}$ are SGUUB in mean square, and k_i, τ_i, and h satisfy*

$$\begin{aligned} k_i &\geq \frac{3}{2} b_{\max} + \frac{9}{4} + \frac{1}{4} \alpha_0, \quad i = 1, 2, \ldots, \rho, \\ \frac{1}{\tau_i} &\geq \frac{1}{4} b_{\max} + 1 + \frac{1}{4} \alpha_0, \quad i = 2, \ldots, \rho, \\ h &\geq \frac{9}{4} + \alpha_0 \lambda_{\max}(P), \\ \alpha_0 &= \min\left\{\frac{c_3}{2 c_2}, \gamma_1 \sigma_1, \gamma_2 \sigma_2, \gamma_3 \sigma_3\right\}, \end{aligned} \tag{63}$$

where $c > 0$ is a positive constant; V will be given later in the proof of Theorem 11.

Proof. Choose the following Lyapunov function candidate:

$$\begin{aligned} V &= V_{sW\varepsilon} + \sum_{i=2}^{\rho} V_{s_i} + \frac{1}{4} \sum_{i=2}^{\rho} y_i^4 + \frac{1}{2\gamma_1} \tilde{\theta}_1^T \tilde{\theta}_1 + \frac{1}{2\gamma_2} \tilde{b}_m^2 \\ &\quad + \frac{1}{2\gamma_3} \tilde{\bar{b}}^T \tilde{\bar{b}} \end{aligned} \tag{64}$$

For any given positive constant, if $EV \leq c$, according to Lemma 9, we obtain that $s_1, \ldots, s_\rho$, $y_2, \ldots, y_\rho$, y, $\hat{\theta}_1$, $\hat{b}_m$, $\hat{\bar{b}}$ are bounded in probability. $V_W \leq V_{sW\varepsilon} = V_{s1} + V_W + V_\varepsilon \leq V \leq c$, and, from Assumption 2, we obtain that $(c_1/\lambda_0)\|z\|^4 \leq (1/\lambda_0) W \leq c$; that is, $\|z\|^4 \leq \lambda_0 c / c_1$, so z is bounded in probability.

Furthermore, (64) is rewritten as $V = (1/\lambda_0) W + (1/2) V_\rho - (c_3/2\lambda_0)\|z\|^4$; then $V_\rho = 2V - (2/\lambda_0) W + (c_3/\lambda_0)\|z^4\| \leq (2 + c_3/c_1)c$, and choosing $p = (2 + c_3/c_1)c$, we get $V_\rho \leq p$.

From (14) and (49), we have that $\xi, \bar{\omega}, \alpha_1$ are also bounded in probability. It yields that $v_{m-1,2}, \ldots, v_{0,2}$ are all bounded in probability. Noting $v_{m,2} = s_2 + y_2 + \alpha_1$, we obtain that $v_{m,2}$ is bounded in probability. From (14), we have that $\dot{v}_{0,1} = -l_1 v_{0,1} + v_{0,2}$ and $\dot{v}_{m,1} = -l_1 v_{m,1} + v_{m,2}$. Thus we obtain that $v_{0,1}$, $v_{m,1}$ are also bounded. Furthermore, from (42), we have that α_i $(i = 2, \ldots, \rho - 1)$ are bounded. According to (16) and (17), we obtain

$$\begin{bmatrix} v_{0,1} \\ v_{0,2} \\ v_{1,2} \\ \vdots \\ v_{m-1,2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ * & * & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ * & * & \cdots & * & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_{m+1} \end{bmatrix}. \quad (65)$$

Since $v_{0,1}, v_{0,2}, v_{1,2}, \ldots, v_{m-1,2}$ are bounded in probability, we have that $\lambda_1, \ldots, \lambda_{m+1}$ are all bounded in probability. From (17), we get that λ_{m+2} is also bounded. In view of (40), (44), (47)–(49), and (62), using Young's inequality, we obtain

$$\begin{aligned} \ell V \leq & -\left(h - \frac{9}{4}\right)\varepsilon^T\varepsilon - \sum_{i=1}^{\rho}\left(k_i - \frac{3}{2}b_{\max} - \frac{9}{4}\right)s_i^4 \\ & - \sum_{i=2}^{\rho}\left(\frac{1}{\tau_i} - \frac{1}{4}b_{\max} - 1\right)y_i^4 - \frac{c_3}{2\lambda_0}\|z\|^4 \\ & - \frac{\sigma_1\left\|\tilde{\theta}_1\right\|^2}{2} - \frac{\sigma_2\tilde{b}_m^2}{2} - \frac{\sigma_3\left\|\tilde{\bar{b}}\right\|^2}{2} + Q(y) \\ & + \frac{1}{4}N_0^4(p) + \frac{1}{4}M^4(p) + \frac{1}{4}\sum_{i=2}^{\rho}N_i^4(p) \\ & + \sum_{i=2}^{\rho}C_i(p) + \frac{\sigma_1\left\|\theta_1\right\|^2}{2} + \frac{\sigma_2 b_m^2}{2} + \frac{\sigma_3\left\|\bar{b}\right\|^2}{2} + C_0. \end{aligned} \quad (66)$$

Substituting (63) into (66), we obtain

$$\ell V \leq -\alpha_0 V + \mu_1 + \mu_0, \quad (67)$$

where $\mu_1 = (1/4)N_0^4(p) + (1/4)M^4(p) + (1/4)\sum_{i=2}^{\rho}N_i^4(p) + \sum_{i=2}^{\rho}C_i(p) + \sigma_1\|\theta_1\|^2/2 + \sigma_2 b_m^2/2 + \sigma_3\|\bar{b}\|^2/2 + C_0$. Since $|Q(y)|$ is a nonnegative continuous function, let $|Q(y)| \leq \mu_0$, where $\mu_0 > 0$.

If $\alpha_0 \geq (\mu_1 + \mu_0)/c$, and $EV = c$, then we have $dEV/dt \leq 0$. Thus, if $EV(0) \leq c$, then $EV(t) \leq c$, $\forall t > 0$; that is,

$$\begin{aligned} 0 \leq EV(t) \leq & \frac{\mu_1 + \mu_0}{\alpha_0} + \left[V(0) - \frac{\mu_1 + \mu_0}{\alpha_0}\right]e^{-\alpha_0 t} \\ \leq & V(0). \end{aligned} \quad (68)$$

Similar to the discussion of Theorem 11 in [32], it is easy to know that the conclusion is true. □

Remark 12. This paper differs from [32] in the following several aspects. (1) Unmodeled dynamics is dealt with by introducing a novel description based on Lyapunov function in this paper while the dynamic signal was handled with the help of a dynamic signal in [32]. (2) The unknown nonlinear system functions are handled together with some functions produced in stability analysis, but they were directly approximated before constructing the observer in [32]. Therefore, this brings out a good result that the filter order is reduced. (3) The neural networks weight vector used to approximate the black box function at the first design step is adjusted online in this paper such that much more information of weight vector can be used in adaptive law, whereas only the norm of weight vector acts as adaptive tuning parameter in [32]. (4) Utilizing bounded input bounded output stability and linear equations (65), the stability of the closed-loop system is proved in this paper, which avoids using the transfer function to make stability analysis in [32], which is questionable in probability sense.

Remark 13. The design parameters k_i, τ_i and α_0 determined by (63) in Theorem 11 are only a sufficient condition. They provide a guideline for the designers. From (63), some suggestions are given for the choice of some key design parameters for any given positive constants B_0 and c.

(i) Increasing $\gamma_1, \gamma_2, \gamma_3$ helps to increase α_0, subsequently reduces μ_1/α_0.

(ii) Decreasing $\sigma_1, \sigma_2, \sigma_3$ helps to reduce μ_1 and reduces μ_1/α_0.

(iii) Increasing $k_1, \ldots, k_\rho$ helps to increase α_0 and reduces μ_1/α_0.

In practical applications, to obtain good tracking performance, some experiments need to be done before the valid parameters are given.

4. Simulation Results

To demonstrate the effectiveness of the proposed approach, two numerical examples are given.

Example 1. Consider the following third-order stochastic nonlinear system with unmodeled dynamics:

$$\begin{aligned} \dot{z} &= q(z, y), \\ dx_1 &= \left(x_2 + \frac{y - y^3}{1 + y^2} + 0.5z\right)dt + y\sin\left(y^3\right)dw, \\ dx_2 &= \left(x_3 + \frac{y - y^3}{1 + y^2} + 0.5z + 0.2\left(35 + y^2\right)u\right)dt \\ & \quad + x_1\sin\left(y^3\right)dw, \\ dx_3 &= \left(y^2\tanh(y) - \left(y^2 + 2y\right)\sin y + 0.2\left(35 + y^2\right)u + yz\right)dt \\ & \quad + 0.5y^2 dw, \\ y &= x_1, \end{aligned} \quad (69)$$

where $q(z, y) = -2z + y\sin t + 0.5$, $m = 1$, $\rho = 2$. The desired tracking trajectory is taken as $y_d = 0.5\sin(0.5t)$. Select $W(z, t) = (1/4)z^4$, $c_1 = 1/8$, $c_2 = 1$, $c_3 = 2$, $c_4 = 1$,

$c_5 = 0.5$; then $(\partial W/\partial t)(z,t) + (\partial W/\partial z)(z,t)(q(z,0,t) - q(0,0,t)) = -2z^4$, $|(\partial W/\partial z)(z,t)| = |z|^3$, $|q(0,0,t)| = 0.5$; $\psi_0(|y|) = |y|$, $|q(z,y,t) - q(z,0,t)| = |y\sin t| \le \psi_0(|y|)$, and it satisfies the conditions of Assumptions 2 and 3.

The filters are designed as follows:

$$\begin{aligned}
\dot{\xi}_1 &= -l_1\xi_1 + \xi_2 + l_1 y,\\
\dot{\xi}_2 &= -l_2\xi_1 + \xi_3 + l_2 y,\\
\dot{\xi}_3 &= -l_3\xi_1 + l_3 y,\\
\dot{\lambda}_1 &= -l_1\lambda_1 + \lambda_2,\\
\dot{\lambda}_2 &= -l_2\lambda_1 + \lambda_3,\\
\dot{\lambda}_3 &= -l_3\lambda_1 + \sigma(y)u.
\end{aligned} \tag{70}$$

The adaptation laws are employed as follows:

$$\begin{aligned}
\dot{\widehat{\theta}}_1 &= \gamma_1\left(s_1^3\psi_1(X) - \sigma_1\widehat{\theta}_1\right),\\
\dot{\widehat{b}}_1 &= \gamma_2\left(s_1^3\alpha_1 - \sigma_2\widehat{b}_1\right),\\
\dot{\widehat{b}}_0 &= \gamma_3\left(s_1^3 v_{0,2} - \sigma_3\widehat{b}_0\right),
\end{aligned} \tag{71}$$

where $X = [s_1, y_d, \dot{y}_d]^T$.

The virtual control law α_1 is chosen as follows:

$$\alpha_1 = \frac{\widehat{b}_1}{\widehat{b}_1^2 + \beta}\left(-k_1 s_1 - \xi_2 - v_{0,2}\widehat{b}_0 - s_1^3\widehat{\theta}_1^T\psi_1(X)\right). \tag{72}$$

The control law is employed as follows:

$$u = \frac{(-k_2 s_2 + l_2 v_{0,1} + \dot{\omega}_2)}{\sigma(y)}, \tag{73}$$

where $v_{0,1} = \lambda_1$, $v_{0,2} = \lambda_2$, $\sigma(y) = 35 + y^2$.

In the simulation, $s_1 = y - y_d$, $s_2 = v_{0,2} - \omega_2$, $l_1 = 6$, $l_2 = 11$, $l_3 = 6$, $k_1 = 40$, $k_2 = 50$, $\beta = 0.02$, $\omega_2(0) = 0.1$, $\tau_2 = 0.01$, $\gamma_1 = \gamma_2 = \gamma_3 = 2$, $\sigma_1 = \sigma_2 = \sigma_3 = 0.05$, $x(0) = [0.2, 0, 0]^T$, $\xi(0) = [0,0,0]^T$, $\lambda(0) = [0,0,0]^T$, $\widehat{b}_0(0) = \widehat{b}_1(0) = 1$, $\widehat{\theta}_1(0) = [0.1]^T_{1\times 10}$, $M_1 = 10$. Simulation results are shown in Figures 1, 2, and 3. From Figure 1, it can be seen that fairly good tracking performance is obtained.

Remark 14. According to (69), we know that $b_1 = 0.2$ and $b_{\max} = 1$. From the above selected design parameters and (63), it is easy to see that $c_3/(2c_2) = 1$, $\alpha_0 = 0.1$. The constant h is only used to analyze the stability in the closed-loop system. Therefore, (63) is true for the above selected design parameters $\gamma_1, \gamma_2, \gamma_3, \sigma_1, \sigma_2, \sigma_3, k_1, k_2, \tau_2$.

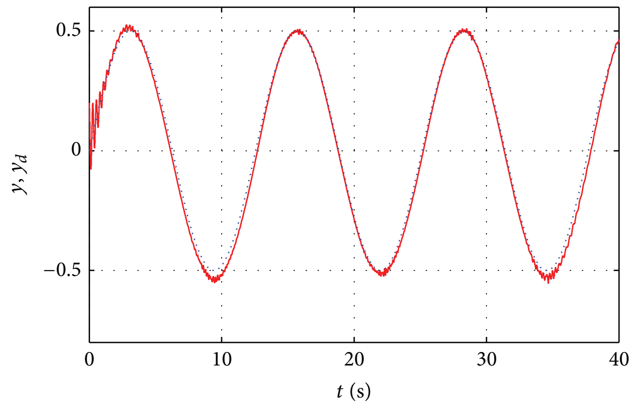


FIGURE 1: Output y (solid line) and desired trajectory y_d (dotted line).

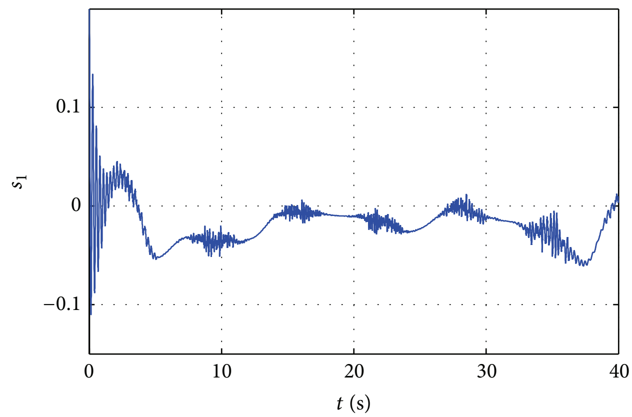


FIGURE 2: Tracking error s_1.

Example 2. To compare the simulation results with [32], consider the following same stochastic nonlinear system with unmodeled dynamics in [32]:

$$\begin{aligned}
\dot{z} &= q(z,y),\\
dx_1 &= \left(x_2 + \frac{x_1 - x_1^3}{1 + x_1^2} + 0.5z\right)dt + x_1\sin\left(x_1^3\right)dw,\\
dx_2 &= \left(x_1^2\tanh(x_1) - \left(x_1^2 + 2x_1\right)\sin x_1\right.\\
&\quad \left. + 0.2\left(0.5 + x_1^2\right)u + x_1 z\right)dt\\
&\quad + 0.5x_1^2\,dw,\\
y &= x_1,
\end{aligned} \tag{74}$$

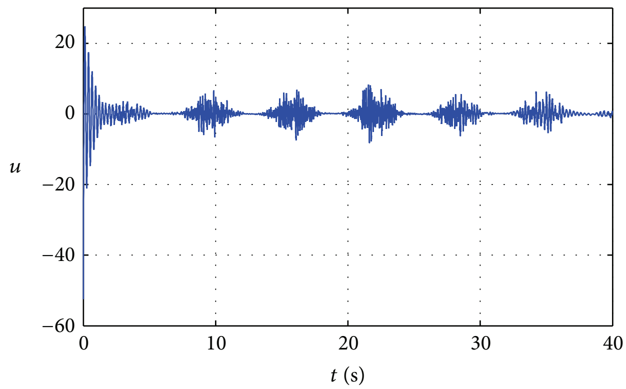


Figure 3: Control signal u.

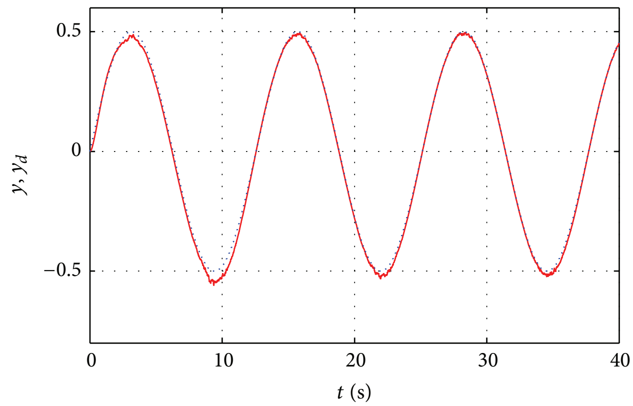


Figure 4: Output y (solid line) and desired trajectory y_d (dotted line).

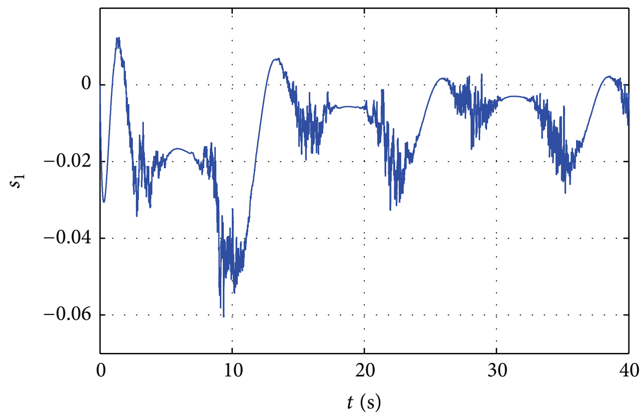


Figure 5: Tracking error s_1.

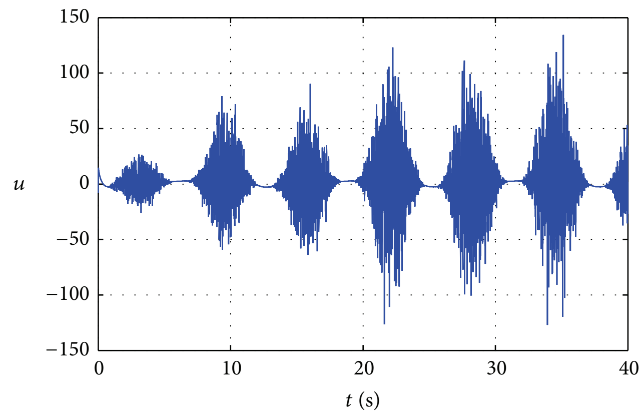


Figure 6: Control signal u.

where $q(z, y) = -2z + y^2$, $m = 0$, $\rho = 2$. The desired tracking trajectory is taken as $y_d = 0.5\sin(0.5t)$. The filters are designed as follows:

$$\begin{aligned}\dot{\xi}_1 &= -l_1\xi_1 + \xi_2 + l_1 y,\\ \dot{\xi}_2 &= -l_2\xi_1 + l_2 y,\\ \dot{\lambda}_1 &= -l_1\lambda_1 + \lambda_2,\\ \dot{\lambda}_2 &= -l_2\lambda_1 + \sigma(y)u.\end{aligned} \tag{75}$$

The adaptation laws are employed as follows:

$$\begin{aligned}\dot{\widehat{\theta}}_1 &= \gamma_1\left(s_1^3\psi_1(X) - \sigma_1\widehat{\theta}_1\right),\\ \dot{\widehat{b}}_0 &= \gamma_3\left(s_1^3 v_{0,2} - \sigma_3\widehat{b}_0\right),\end{aligned} \tag{76}$$

where $X = [s_1, y_d, \dot{y}_d]^T$.

The virtual control law α_1 is chosen as follows:

$$\alpha_1 = \frac{\widehat{b}_0}{\widehat{b}_0^2 + \beta}\left(-k_1 s_1 - \xi_2 - \lambda_2\widehat{b}_0 - s_1^3\widehat{\theta}_1^T\psi_1(X)\right). \tag{77}$$

The control law is employed as follows:

$$u = \frac{\left(-k_2 s_2 + l_2 v_{0,1} + \dot{\omega}_2\right)}{\sigma(y)}, \tag{78}$$

where $v_{0,1} = \lambda_1$, $\sigma(y) = 0.5 + y^2$.

In the simulation, $s_1 = y - y_d$, $s_2 = v_{0,2} - \omega_2$, $l_1 = 5$, $l_2 = 6$, $k_1 = 60$, $k_2 = 60$, $\beta = 0.02$, $\omega_2(0) = 0.1$, $\tau_2 = 0.01$, $\gamma_1 = \gamma_3 = 1.5$, $\sigma_1 = \sigma_3 = 0.05$, $x(0) = [0, 0]^T$, $z(0) = 0$, $\xi(0) = [0, 0]^T$, $\lambda(0) = [0, 0]^T$, $\widehat{b}_0(0) = 1$, $\widehat{\theta}_1(0) = [0.1, 0.1, 0.1, 0.1, 0.1]^T$, $M_1 = 5$. Simulation results are shown in Figures 4–6. If the proposed approach in [32] is utilized, and the design parameters of the adaptive controller are taken, the same values as in [32], the corresponding simulation results are as shown in Figures 7–9.

From Figures 4, 5, 7, and 8, it can be seen that better tracking performance can be obtained than [32]. However, 42 equations need to be solved online using the method in [32] while only 14 equations need to be solved online using the approach in this paper. Moreover, we know that increasing k_1, k_2 helps to improve the tracking precision.

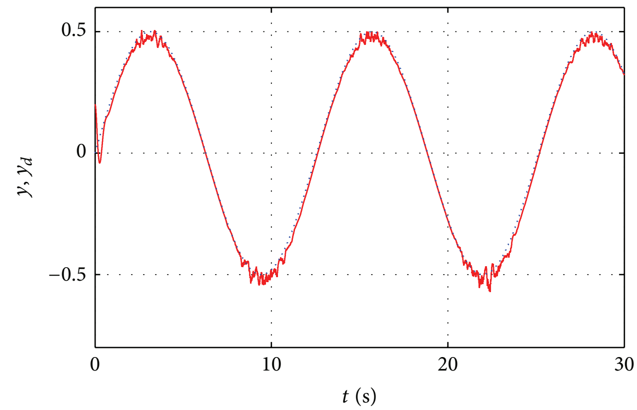


FIGURE 7: Output y (solid line) and desired trajectory y_d (dotted line).

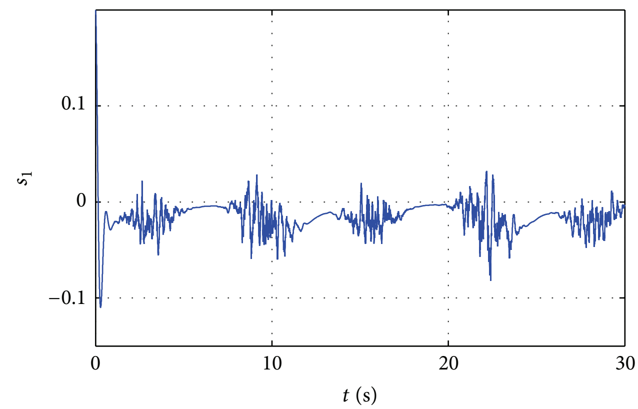


FIGURE 8: Tracking error s_1.

5. Conclusions

Using K-filters and dynamic surface control, an adaptive output feedback neural control scheme has been proposed for a class of stochastic nonlinear systems with unmodeled dynamics. Unmodeled dynamics has been dealt with by introducing the novel description based on Lyapunov function. The unknown nonlinear system functions are handled together with some functions resulting from stability analysis, and the filter order is reduced. The neural network weight vector is adjusted online. Therefore, the more information included in radial basis function can be fully made use of. Using Chebyshev's inequality and Itô formula, the designed controller can guarantee that all the signals in the closed-loop system are bounded in probability and the error signals are semiglobally uniformly ultimately bounded in the sense of four-moment or mean square. Simulation results illustrate the effectiveness of the proposed approach.

Conflict of Interests

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

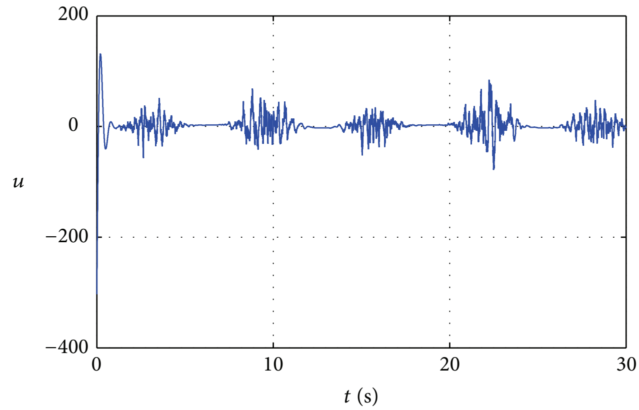


FIGURE 9: Control signal u.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (61174046, 61473249, and 61473250).

References

[1] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Transactions on Automatic Control*, vol. 36, no. 11, pp. 1241–1253, 1991.

[2] D. Swaroop, J. K. Hedrick, P. P. Yip, and J. C. Gerdes, "Dynamic surface control for a class of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1893–1899, 2000.

[3] Z. Jiang and L. Praly, "Design of robust adaptive controllers for nonlinear systems with dynamic uncertainties," *Automatica*, vol. 34, no. 7, pp. 825–840, 1998.

[4] Z. P. Jiang and D. J. Hill, "A robust adaptive backstepping scheme for nonlinear systems with unmodeled dynamics," *IEEE Transactions on Automatic Control*, vol. 44, no. 9, pp. 1705–1711, 1999.

[5] D. Wang and J. Huang, "Neural network-based adaptive dynamic surface control for a class of uncertain nonlinear systems in strict-feedback form," *IEEE Transactions on Neural Networks*, vol. 16, no. 1, pp. 195–202, 2005.

[6] T. S. Li, D. Wang, G. Feng, and S. C. Tong, "A DSC approach to robust adaptive nn tracking control for strict-feedback nonlinear systems," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 40, no. 3, pp. 915–927, 2010.

[7] T. P. Zhang and S. S. Ge, "Adaptive dynamic surface control of nonlinear systems with unknown dead zone in pure feedback form," *Automatica*, vol. 44, no. 7, pp. 1895–1903, 2008.

[8] T. P. Zhang, Q. Zhu, and Y. Q. Yang, "Adaptive neural control of non-affine pure-feedback non-linear systems with input nonlinearity and perturbed uncertainties," *International Journal of Systems Science*, vol. 43, no. 4, pp. 691–706, 2012.

[9] T. Zhang, X. Shi, Q. Zhu, and Y. Yang, "Adaptive neural tracking control of pure-feedback nonlinear systems with unknown gain signs and unmodeled dynamics," *Neurocomputing*, vol. 121, pp. 290–297, 2013.

[10] X. Y. Zhang and Y. Lin, "Adaptive tracking control for a class of pure-feedback non-linear systems including actuator

hysteresis and dynamic uncertainties," *IET Control Theory and Applications*, vol. 5, no. 16, pp. 1868–1880, 2011.

[11] T. Zhang and S. S. Ge, "Adaptive neural network tracking control of MIMO nonlinear systems with unknown dead zones and control directions," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 483–497, 2009.

[12] S. J. Yoo and J. B. Park, "Neural-network-based decentralized adaptive control for a class of large-scale nonlinear systems with unknown time-varying delays," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 39, no. 5, pp. 1316–1323, 2009.

[13] S. J. Yoo, "Decentralised adaptive control of a class of interconnected non-linear systems with unknown time delays and dead-zone inputs," *IET Control Theory & Applications*, vol. 4, no. 11, pp. 2639–2650, 2010.

[14] M. Krstic, I. Kanellakopoulos, and P. V. Kokotocic, *Nonlinear and Adaptive Control Design*, Wiley-Interscience, New York, NY, USA, 1995.

[15] Y. J. Liu and W. Wang, "Adaptive output feedback control of uncertain nonlinear systems based on dynamic surface control technique," *International Journal of Robust and Nonlinear Control*, vol. 22, no. 9, pp. 945–958, 2012.

[16] S. Tong, C. Liu, and Y. Li, "Adaptive fuzzy backstepping output feedback control for strict feedback nonlinear systems with unknown sign of high-frequency gain," *Neurocomputing*, vol. 77, no. 1, pp. 58–70, 2012.

[17] Y. M. Li, S. C. Tong, Y. J. Liu, and T. S. Li, "Adaptive fuzz y robust otput feedback control of nonlinear systems with unknown dead-zones based on a small-gain approach," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 1, pp. 164–176, 2014.

[18] C. Wang and Y. Lin, "Decentralised adaptive dynamic surface control for a class of interconnected non-linear systems," *IET Control Theory & Applications*, vol. 6, no. 9, pp. 1172–1181, 2012.

[19] H. Deng and M. Krstić, "Stochastic nonlinear stabilization—I: a backstepping design," *Systems and Control Letters*, vol. 32, no. 3, pp. 143–150, 1997.

[20] H. Deng and M. Krstić, "Output-feedback stochastic nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 328–333, 1999.

[21] H. Deng, M. Krstić, and R. J. Williams, "Stabilization of stochastic nonlinear systems driven by noise of unknown covariance," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1237–1253, 2001.

[22] H. Fan and W. Xie, "Adaptive state feedback control for a class of stochastic nonlinear systems with unknown backlash-like hysteresis," in *Proceedings of the 2nd IEEE Conference on Industrial Electronics and Applications (ICIEA '07)*, pp. 2797–2801, May 2007.

[23] H. Fan, L. Han, C. Wen, and L. Xu, "Decentralized adaptive output-feedback controller design for stochastic nonlinear interconnected systems," *Automatica*, vol. 48, no. 11, pp. 2866–2873, 2012.

[24] T. Wang, S. Tong, and Y. Li, "Robust adaptive fuzzy control for a class of stochastic nonlinear systems with dynamical uncertainties," *Journal of the Franklin Institute*, vol. 349, no. 10, pp. 3121–3141, 2012.

[25] T. Wang, S. C. Tong, and Y. M. Li, "Robust adaptive fuzzy output feedback control for stochastic nonlinear systems with unknown control direction," *Neurocomputing*, vol. 106, pp. 31–41, 2013.

[26] Z. J. Wu, X. J. Xie, and S. Y. Zhang, "Adaptive backstepping controller design using stochastic small-gain theorem," *Automatica*, vol. 43, no. 4, pp. 608–620, 2007.

[27] X. Yu and X. Xie, "Output feedback regulation of stochastic nonlinear systems with stochastic iISS inverse dynamics," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 304–320, 2010.

[28] S. Liu and J. Zhang, "Output-feedback control of a class of stochastic nonlinear systems with linearly bounded unmeasurable states," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 6, pp. 665–687, 2008.

[29] W. Li, Y. Jing, and S. Zhang, "Output-feedback stabilization for stochastic nonlinear systems whose linearizations are not stabilizable," *Automatica*, vol. 46, no. 4, pp. 752–760, 2010.

[30] X. Yu, X. J. Xie, and N. Duan, "Small-gain control method for stochastic nonlinear systems with stochastic iISS inverse dynamics," *Automatica*, vol. 46, no. 11, pp. 1790–1798, 2010.

[31] C. Zhao and X. Xie, "Output feedback stabilization using small-gain method and reduced-order observer for stochastic nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 523–528, 2013.

[32] T. P. Zhang and X. N. Xia, "Adaptive output feedback tracking control of stochastic nonlinear systems with dynamic uncertainties," *International Journal of Robust and Nonlinear Control*, 2014.

[33] H. T. Gao, T. P. Zhang, and X. N. Xia, "Adaptive neural control of stochastic nonlinear systems with unmodeled dynamics and time-varying state delays," *Journal of the Franklin Institute*, vol. 351, no. 6, pp. 3182–3199, 2014.

[34] R. Z. Khasminskii, *Stochastic Stability of Differential Equations*, Kluwer Academic, Norwell, Mass, USA, 1980.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 707069, 8 pages
http://dx.doi.org/10.1155/2015/707069

Research Article

Boundedness of Stochastic Delay Differential Systems with Impulsive Control and Impulsive Disturbance

Liming Wang,[1] Baoqing Yang,[2] Xiaohua Ding,[1] and Kai-Ning Wu[1]

[1] *Department of Mathematics, Harbin Institute of Technology at Weihai, Weihai 264209, China*
[2] *Control and Simulation Center, Harbin Institute of Technology, Harbin 150001, China*

Correspondence should be addressed to Kai-Ning Wu; kainingwu@163.com

Received 27 February 2014; Accepted 5 August 2014

Academic Editor: Weihai Zhang

This paper considers the p-moment boundedness of nonlinear impulsive stochastic delay differential systems (ISDDSs). Using the Lyapunov-Razumikhin method and stochastic analysis techniques, we obtain sufficient conditions which guarantee the p-moment boundedness of ISDDSs. Two cases are considered, one is that the stochastic delay differential system (SDDS) may not be bounded, and how an impulsive strategy should be taken to make the SDDS be bounded. The other is that the SDDS is bounded, and an impulsive disturbance appears in this SDDS, then what restrictions on the impulsive disturbance should be adopted to maintain the boundedness of the SDDS. Our results provide sufficient criteria for these two cases. At last, two examples are given to illustrate the correctness of our results.

1. Introduction

Boundedness is an important property of a given system; for example, in the population models, the boundedness of a biological population is strongly connected with the persistence and extinction [1]. Another important application is on the stability; the practical stability actually is of a kind of boundedness [2]. Impulsive phenomena widely exist in the real world, and known, impulsive effects can change the properties of a given system; for example, given an unstable system, if a suitable impulsive strategy, including the impulsive strength and impulsive moments, is adopted, this system can be stabilized [3]. It is easy to understand that the impulsive effects can destroy the boundedness of a given system when the impulsive strength is large enough and the impulsive interval is small enough. Time delay is extensive in the engineering and applications and impulsive delay differential systems were considered in lots of papers [3–9]. The boundedness of impulsive delay differential systems has also been paid considerable attentions in the past decades. In [10], the authors presented sufficient conditions for uniform ultimate boundedness by virtue of the Lyapunov functional method. The boundedness of variable impulsive perturbations system was considered in [11] and the eventual boundedness was studied in [12]. Recently, the perturbing Lyapunov function method was also used in the study of boundedness [13].

Stochastic noise is ubiquitous [14–16] and stochastic delay differential systems (SDDSs) have been one of the focuses of scientific research for many years. Many properties of SDDSs have been studied and lots of papers were published; see [17, 18] and the references therein. Being the wide existence of stochastic delay and impulsive effects, it is a natural task to consider the stochastic delay differential systems with impulsive effects. These systems are described by impulsive stochastic delay differential systems (ISDDSs). In the past ten years, the stability of ISDDSs has attracted a lot of researchers, and a great deal of results on the stability of ISDDSs have been reported; see [19–24] and the references therein.

However, little attention has been paid to the boundedness of ISDDSs. In this paper, the boundedness of ISDDSs is considered under two cases. The first case is that the SDDSs may be unbounded, then what kind of impulsive strategy should be taken to make the system be bounded.

The second case is that the SDDSs are bounded, then this system can tolerate what kind of impulsive effect to maintain the boundedness.

In this paper, sufficient conditions are presented to guarantee the boundedness of ISDDSs; these conditions also admit the global existence of solutions for ISDDSs, which usually was a standard assumption in many papers [25–27]. Making use of the Lyapunov-Razumikhin method, we generalize the results of [10] to the stochastic situation. At last, two examples are given to illustrate the correctness of our results.

2. Preliminaries and Model Description

Let $(\Omega, F, \{F_t\}_{t\geqslant 0}, P)$ be a complete probability space with a filtration $\{F_t\}_{t\geqslant 0}$ satisfying the usual conditions (i.e., the filtration contains all P-null sets and is right continuous). Let $\mathbb{R} = (-\infty, +\infty)$, $\mathbb{R}^+ = [0, +\infty)$, and $\mathbb{N} = \{1, 2, \ldots\}$. If A is a vector or a matrix, its transpose is denoted by A^T. Consider $PC(\mathbb{J}; \mathbb{R}^n) = \{\varphi : \mathbb{J} \to \mathbb{R}^n$, $\varphi(s)$ is continuous for all but at most countable points $s \in \mathbb{J}$ and at these points, $\varphi(s^+)$ and $\varphi(s^-)$ exist and $\varphi(s^+) = \varphi(s)\}$, where $\mathbb{J} \subset \mathbb{R}$ is an interval and $\varphi(s^+)$ and $\varphi(s^-)$ denote the right-hand and left-hand limits of the function $\varphi(s)$ at time s, respectively. Consider $PC^{1,2} = \{\varphi(t,x) : \varphi(\cdot, x) \in PC$ and $\varphi(t,x) \in C^{1,2}$ if t is not at the uncontinuous points $s\}$. Let $PC^b_{F_0}([-\tau, 0]; \mathbb{R}^n)(PC^b_{F_t}([-\tau, 0]; \mathbb{R}^n))$ denote the family of all bounded $F_0(F_t)$-measurable, PC-valued random variables. Let $|\cdot|$ be the Euclidean norm in $\mathbb{R}^n$ and $\|\varphi\|_\tau = \sup_{-\tau\leqslant\theta\leqslant 0}|\varphi(t+\theta)|$.

Consider the following nonlinear impulsive stochastic delay differential system:

$$\begin{gathered} dx(t) = f(t, x_t)\,dt + g(t, x_t)\,dB(t), \\ t > t_0, \quad t \neq t_k, \quad k \in \mathbb{N}, \\ x(t_k) = x(t_k^-) + I(t_k, x(t_k^-)), \quad k \in \mathbb{N}, \\ x(t_0 + s) = \varphi(s), \quad s \in [-\tau, 0], \end{gathered} \tag{1}$$

where $x_t(s) = x(t+s)$, $s \in [-\tau, 0]$, $f : \mathbb{R}^+ \times PC([-\tau, 0], \mathbb{R}^n) \to \mathbb{R}^n$, $g : \mathbb{R}^+ \times PC([-\tau, 0], \mathbb{R}^n) \to \mathbb{R}^{n\times m}$, $I : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ and satisfies global Lipschitz condition, τ represents the delay in system (1), impulsive moment t_k satisfies $0 < t_1 < t_2 < \cdots < t_n < \cdots$, and $t_k \to \infty$ as $k \to \infty$. $B(t)$ is an m-dimensenal Brownian motion and $\varphi(s) \in PC^b_{F_0}([-\tau, 0], \mathbb{R}^n)$.

Given a function $V \in PC^{1,2} : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^+$, the operator $\mathscr{L}$ of $V(t,x)$ with respect to system (1) is defined by

$$\begin{aligned} \mathscr{L}V(t,x) &= V_t + V_x f(t, x_t) \\ &\quad + \frac{1}{2}\operatorname{trace}\left[g^T(t, x_t)\, V_{xx}\, g(t, x_t)\right], \end{aligned} \tag{2}$$

where

$$\begin{aligned} V_t &= \frac{\partial V(t,x)}{\partial t}, \\ V_x &= \left(\frac{\partial V(t,x)}{\partial x_1}, \frac{\partial V(t,x)}{\partial x_2}, \ldots, \frac{\partial V(t,x)}{\partial x_n}\right)^T, \\ V_{xx} &= \left(\frac{\partial^2 V(t,x)}{\partial x_i \partial x_j}\right)_{n\times n}. \end{aligned} \tag{3}$$

Definition 1. System (1) is said to be

(1) p-moment bounded if, for every $B_1 > 0$ and $t_0 \in \mathbb{R}_+$, there exists $B_2 = B_2(t_0, B_1)$ such that if $\varphi \in PC^b_{F_0}([-\tau, 0], \mathbb{R}^n)$ with $E\|\varphi\|^p_\tau \leqslant B_1$ and $x = x(t, t_0, \varphi)$ is a solution of (1), then $E|x(t, t_0, \varphi)|^p \leqslant B_2$ for all $t \geqslant t_0$;

(2) p-moment uniformly bounded if the system (1) is p-moment bounded and B_2 is independent of t_0;

(3) p-moment ultimately bounded if the system (1) is p-moment bounded and there exists a positive constant B such that for every $B_3 > 0$ and $t_0 \in \mathbb{R}^+$ there exists some $T = T(t_0, B_3) > 0$; if $\varphi \in PC^b_{F_0}([-\tau, 0], \mathbb{R}^n)$ with $E\|\varphi\|^p_\tau \leqslant B_3$, then $E|x(t, t_0, \varphi)|^p \leqslant B$ for $t \geqslant t_0 + T$;

(4) p-moment uniformly ultimately bounded, if the system (1) is p-moment ultimately bounded and T is independent of t_0.

3. Boundedness with Impulsive Control

In this section, we consider the first case: when the given SDDS may not be bounded, we adopt an impulsive strategy to get the boundedness. The main result is stated as follows.

Theorem 2. *Assume there exist a positive function $V(t,x) \in PC^{1,2}$ and positive constants $\rho, p, a, b, \gamma, \lambda$, where $0 < \lambda < 1$ and $1 - \lambda - \gamma\tau > 0$, such that*

(1) *$a|x|^p \leqslant V(t,x) \leqslant b|x|^p$ for any (t,x);*

(2) *for $t \neq t_k$, any $s \in [-\tau, 0]$, and $\phi(t) \in PC([-\tau, 0], \mathbb{R}^n)$, $\mathscr{L}V(t, \phi(0)) \leqslant \gamma V(t, \phi(0))$ whenever $V(t, \phi(0)) \geqslant \lambda V(t+s, \phi(s))$ and $|\phi(0)|^p \geqslant \rho$;*

(3) *$V(t_k, \phi(0) + I(t_k, \phi(0))) \leqslant \lambda V(t_k^-, \phi(0))$ for all $|\phi(0)|^p \geqslant \rho$;*

(4) *there exists a positive constant $\rho_1 \geqslant \rho$ such that if $|\phi(0)|^p \leqslant \rho$, then $|\phi(0) + I(t_k, \phi(0))|^p \leqslant \rho_1$;*

(5) *$\alpha = \sup_{k\in\mathbb{Z}}\{t_k - t_{k-1}\} < \infty$, $\alpha\gamma < 1 - \lambda$.*

Then the system (1) is p-moment uniformly ultimately bounded.

Proof. We separate the proof into two parts. First, we show the p-moment uniform boundedness and then we give the ultimate uniform boundedness.

Step 1. Let $B_1 > 0$. Without loss of generality, we assume $B_1 \geqslant \rho_1 \geqslant \rho$. Choose $B_2 = B_2(B_1)$ such that $bB_1 < \lambda a B_2$; then we can see $B_2 > B_1$.

Let $E\|\varphi\|_\tau^p < B_1$ and $t_0 \in [t_{l-1}, t_l)$ for some positive integer l. Suppose $x(t) = x(t, t_0, \varphi)$ is a solution of system (1) with initial value φ and its maximal interval of existence is $[t_0 - \tau, t_0 + \beta)$ for some positive constant β. We will show that, for any $t \in [t_0 - \tau, t_0 + \beta)$, $E|x(t)|^p \leqslant B_2$. By the way, if this statement is true, we know that the solution of system (1) is not explored in $[t_0, t_0 + \beta)$, and the global existence of the solution follows.

For the sake of contradiction, suppose $E|x(t)|^p \geqslant B_2$ for some $t \in [t_0, t_0 + \beta)$. Then there exists $\hat{t} = \inf\{t \in [t_0 - \tau, t_0 + \beta) \mid E|x(t)|^p > B_2\}$. Note that $E|x(t)|^p \leqslant E\|\varphi\|_\tau^p \leqslant B_1 < B_2$ for $t \in [t_0 - \tau, t_0]$; we see that $\hat{t} \in (t_0, t_0 + \beta)$ and $E|x(t)|^p \leqslant B_2$ for $t \in [t_0 - \tau, \hat{t})$ and $E|x(\hat{t})|^p \geqslant B_2$.

Write $V(t, x(t)) = V(t)$. For $t \in [t_0 - \tau, t_0]$, we have $EV(t) \leqslant bE|x(t)|^p \leqslant bE\|\varphi\|_\tau^p \leqslant bB_1 < \lambda aB_2 < aB_2$, and $EV(\hat{t}) \geqslant aE|x(\hat{t})|^p \geqslant aB_2$. Define $t^* = \inf\{t \in [t_0, \hat{t}] \mid EV(t) \geqslant aB_2\}$ and then $t^* \in (t_0, \hat{t}]$ and $EV(t) < aB_2$ for $t \in [t_0 - \tau, t^*)$ and $EV(t^*) \geqslant aB_2$.

We claim that $t^* \neq t_k$ for any $k \in \mathbb{N}$ and then $EV(t^*) = aB_2$.

If it is not true, suppose $t^* = t_k$ for some k. If $E|x(t_k^-)|^p \geqslant \rho$, then $aB_2 \leqslant EV(t_k) \leqslant \lambda EV(t_k^-) < \lambda aB_2 < aB_2$, which is a contradiction. If $E|x(t_k^-)|^p < \rho$, then $E|x(t_k)|^p = E|x(t_k^-) + I(t_k, x(t_k^-))|^p < \rho_1 < B_1$. Then $aB_2 < EV(t_k) < bB_1 < \lambda aB_2 < aB_2$, which is a contradiction.

Now we will proceed under two cases.

Case 1. Consider $t_{l-1} \leqslant t_0 < t^* < t_l$.

Let $\bar{t} = \sup\{t \in [t_0, t^*] \mid EV(t) \leqslant \lambda aB_2\}$. Since $EV(t_0) < bB_1 < \lambda aB_2$, $EV(t^*) = aB_2 > \lambda aB_2$, and $EV(t)$ is continuous on $[t_0, t^*]$, then $\bar{t} \in (t_0, t^*)$ and $EV(\bar{t}) = \lambda aB_2$ and, when $t \in [\bar{t}, t^*]$, $EV(t) \geqslant \lambda aB_2$. Hence, for $t \in [\bar{t}, t^*]$ and $s \in [-\tau, 0]$, we have

$$\begin{gathered} \lambda EV(t+s) \leqslant \lambda aB_2 \leqslant EV(t), \\ bB_1 \leqslant \lambda aB_2 \leqslant EV(t) \leqslant bE|x(t)|^p, \end{gathered} \tag{4}$$

and we can get

$$E|x(t)|^p \geqslant B_1 \geqslant \rho. \tag{5}$$

Then, by virtue of condition (2), for $t \in [\bar{t}, t^*]$,

$$\begin{aligned} E\mathscr{L}V(t) &\leqslant \gamma EV(t), \\ EV(t^*) - EV(\bar{t}) &= \int_{\bar{t}}^{t^*} E\mathscr{L}V(s)\,ds \\ &\leqslant \int_{\bar{t}}^{t^*} \gamma EV(s)\,ds < \gamma\alpha aB_2. \end{aligned} \tag{6}$$

However,

$$EV(t^*) - EV(\bar{t}) = aB_2 - \lambda aB_2 = (1-\lambda)aB_2 > \gamma\alpha aB_2, \tag{7}$$

which is contradiction. Then we get, in this case,

$$E|x(t)|^p \leqslant B_2. \tag{8}$$

Case 2. Consider $t_k < t^* < t_{k+1}$ for some $k \geqslant l$.

Note that $EV(t_k) \leqslant \lambda aB_2$. This inequality can be obtained by the following reason: if $E|x(t_k^-)|^p \geqslant \rho$, then $EV(t_k) \leqslant \lambda EV(t_k^-) \leqslant \lambda aB_2$. If $E|x(t_k^-)|^p < \rho$, we get $E|x(t_k)|^p < \rho_1 < B_1$, and then

$$EV(t_k) < bB_1 < \lambda aB_2. \tag{9}$$

Define $\bar{t} = \sup\{t \in [t_k, t^*] \mid EV(t) \leqslant \lambda aB_2\}$, and then $\bar{t} \in [t_k, t^*)$, $EV(\bar{t}) = \lambda aB_2$, and $EV(t) \geqslant \lambda aB_2$ for $t \in [\bar{t}, t^*]$. The same argument as the one in Case 1 yields a contradiction. Therefore, in this case, we have, for any $t \in [t_0 - \tau, \infty)$,

$$E|x(t)|^p \leqslant B_2. \tag{10}$$

Now we get that, under conditions (1) to condition (5), the solutions of (1) are p-moment uniformly bounded. That is, if $E\|\varphi\|_\tau^p \leqslant \rho_1$, there exists a constant $B > 0$, such that $E|x(t, t_0, \varphi)|^p \leqslant B$ for all $t \geqslant t_0 - \tau$, and, from the proof, we have $b\rho_1 < \lambda aB$.

Step 2. Now, let $B_3 > 0$ and assume, without loss of generality, that $B_3 > B$. Then, from the proof of uniform boundedness, there exists some $B_2 = B_2(B_3) > B_3$ for which if $E\|\varphi\|_\tau^p \leqslant B_3$, then $E|x(t)|^p \leqslant B_2$ for $t \geqslant t_0 - \tau$.

Take a constant d satisfying $0 < d \leqslant (1-\lambda-\gamma\tau)aB/(1-\gamma\tau)$; it is easy to verify that $0 < d < (1-\lambda)aB$. Let $N = N(B_3)$ be the smallest positive integer for which $bB_2 < aB + Nd$ and $T = T(B_3) = \alpha + (\tau + \alpha)(N-1)$. Given a solution $x(t) = x(t, t_0, \varphi)$ where $E\|\varphi\|_\tau^p \leqslant B_3$ and $t_0 \in [t_{l-1}, t_l)$, we will show $E|x(t)|^p \leqslant B$ for $t \geqslant t_0 + T$.

Given a constant A satisfying $aB \leqslant A - d \leqslant bB_2$ and $j > l$, we will show that if $EV(t) \leqslant A$ for $t \in [t_j - \tau, t_j)$, then $EV(t) \leqslant A - d$ for $t \geqslant t_j$.

For the sake of contradiction, suppose that there exists some $t \geqslant t_j$ for which $EV(t) > A - d$ and define

$$t^* = \inf\left\{t \geqslant t_j \mid EV(t) > A - d\right\}, \tag{11}$$

and we suppose $t^* \in [t_k, t_{k+1})$ for some $k \in \mathbb{N}$. We can get $EV(t) \leqslant A - d$ for $t \in [t_j - \tau, t^*)$ and $EV(t^*) \geqslant A - d$.

We claim that $EV(t_k) \leqslant \lambda A$. The fact follows that if $E|x(t_k^-)|^p \geqslant \rho$, then $EV(t_k) \leqslant \lambda EV(t_k^-) \leqslant \lambda A$. If $E|x(t_k^-)|^p < \rho$ and we have $E|x(t_k)|^p \leqslant \rho_1$, then $EV(t_k) \leqslant b\rho < bB \leqslant \lambda aB \leqslant \lambda A$.

Now, since $aB \leqslant A$, we have $\lambda A = A - (1-\lambda)A < A - (1-\lambda)aB < A - d$ and $EV(t_k) < A - d$. This implies that $t^* \neq t_k$; that is, $t^* \in (t_k, t_{k+1})$ and $EV(t^*) = A - d$ since $EV(t)$ is continuous at t^*. Also, for $t \in [t_k, t^*]$, we have $EV(t) \leqslant A - d$. Define

$$\bar{t} = \sup\left\{t \in [t_k, t^*] \mid EV(t) \leqslant \lambda(A-d)\right\}. \tag{12}$$

Since $EV(t^*) = A - d > \lambda A > \lambda(A-d)$, we have $\bar{t} \in [t_k, t^*)$ and $EV(\bar{t}) = \lambda(A-d)$ and $EV(t) \geqslant \lambda(A-d)$ for $t \in [\bar{t}, t^*]$. Then, if $t \in [\bar{t}, t^*]$ and $s \in [-\tau, 0]$,

$$\begin{gathered} \lambda EV(t+s) \leqslant \lambda(A-d) < EV(t), \\ bE|x(t)|^p > EV(t) > \lambda(A-d) > \lambda aB > b\rho, \end{gathered} \tag{13}$$

which yields $E|x(t)|^p > \rho$. Then, in light of condition (2),

$$E\mathscr{L}V(t) \leqslant \gamma EV(t). \tag{14}$$

In terms of Itô formula,

$$\begin{aligned} EV(t^*) - EV(\bar{t}) &= \int_{\bar{t}}^{t^*} E\mathscr{L}V(s)\,ds \\ &\leqslant \int_{\bar{t}}^{t^*} \gamma EV(s)\,ds \leqslant \gamma\alpha(A-d). \end{aligned} \tag{15}$$

But

$$EV(t^*) - EV(\bar{t}) = A - d - \lambda(A-d) > \gamma\alpha(A-d), \tag{16}$$

and this contradiction proves that $EV(t) < A - d$ for all $t \geqslant t_j$. □

Now we define a sequence $t_{k^{(i)}} \in \{t_k, k = l, l+1, \ldots\}$, satisfying $t_{k^{(1)}} = t_l$ and $t_{k^{(i)}-1} - \tau \leqslant t_{k^{(i-1)}} \leqslant t_{k^{(i)}} - \tau$, and then we have $t_{k^{(i)}} \leqslant t_{k^{(i)}-1} + \alpha \leqslant t_{k^{(i-1)}} + \tau + \alpha$. By induction, we get $t_{k^{(N)}} \leqslant t_0 + \alpha + (\tau+\alpha)(N-1) = t_0 + T$. We know that when $t \in [t_0 - \tau, t_l)$, that is, $t \in [t_0 - \tau, t_{k^{(1)}})$, $EV(t) \leqslant bB_2$; then by induction we get $EV(t) \leqslant bB_2 - Nd$ for $t \in [t_{k^{(N)}}, \infty)$ and then $EV(t) \leqslant aB$ for $t \in [t_0 + T, \infty)$. Using condition (1), we get that $aE|x(t)|^p \leqslant EV(t) \leqslant aB$; that is,

$$E|x(t)|^p \leqslant B. \tag{17}$$

Remark 3. Condition (2) means the system without impulse may be unbounded. If the impulsive effects satisfy condition (3) to condition (5), then this system can be bounded.

4. Boundedness with Impulsive Disturbance

In this section, we consider the case that the SDDS is bounded, and when the impulsive disturbance appears in the SDDS, then what restrictions should be added to the disturbance to maintain the boundedness. The result is stated as follows.

Theorem 4. *Assume that there exist a positive function $V(t,x)$ and positive constants $a, b, c, p, \lambda_1, \lambda_2, \gamma$, where $1 \leqslant \lambda_1 < \lambda_2$, such that*

(1) *$a|x|^p \leqslant V(t,x) \leqslant b|x|^p$ for any (t,x);*

(2) *for $t \neq t_k$, any $s \in [-\tau, 0]$, and $\phi(s) \in PC([-\tau,0],\mathbb{R}^n)$, $\mathscr{L}V(t,\phi(0)) \leqslant -\gamma V(t,\phi(0))$ whenever $\lambda_2 V(t,\phi(0)) \geqslant V(t+s,\phi(s))$ and $|\phi(0)|^p \geqslant \rho$;*

(3) *$V(t_k, \phi(0) + I(t_k,\phi(0))) \leqslant \lambda_1 V(t_k^-, \phi(0))$ for all $|\phi(0)|^p \geqslant \rho$;*

(4) *there exists a positive constant $\rho_1 \geqslant \rho$ such that if $|\phi(0)|^p \leqslant \rho$, then $|\phi(0) + I(\tau_k,\phi(0))|^p \leqslant \rho_1$;*

(5) *there exist positive constants μ and α, such that $\mu \leqslant t_k - t_{k-1} \leqslant \alpha$ and $\mu\gamma > \lambda_2 - 1$.*

Then, the system (1) is p-moment uniformly ultimately bounded.

Proof. Step 1. Let $B_1 > 0$; without loss of generality, we assume $B_1 \geqslant \rho_1$. Choose $B_2 = B_2(B_1)$, such that $\lambda_2 b B_1 < aB_2$, and then we get $B_2 > B_1$. Let $E\|\varphi\|_\tau^p \leqslant B_1$ and assume $t_0 \in [t_{l-1}, t_l)$; moreover, we assume that (1) has a maximal interval of existence, $[t_0 - \tau, t_0 + \beta)$.

We will prove that $E|x(t)|^p \leqslant B_2$ for $t \in [t_0, t_0+\beta)$. This will show that $\beta = \infty$ and that solutions of (1) are uniformly bounded.

For the sake of contradiction, we suppose that $E|x(t)|^p > B_2$ for some $t \in [t_0, t_0+\beta)$. Let $\hat{t} = \inf\{t \in [t_0, t_0+\beta) \mid E|x(t)|^p > B_2\}$. Note that $E|x(t)|^p \leqslant E\|\varphi\|_\tau^p < B_1 < B_2$ for $t \in [t_0 - \tau, t_0]$, and we get $\hat{t} \in (t_0, t_0+\beta)$, $E|x(t)|^p \leqslant B_2$ for $t \in [t_0 - \tau, \hat{t})$ and $E|x(\hat{t})|^p \geqslant B_2$.

For $t \in [t_0 - \tau, t_0]$, we have $EV(t) \leqslant bE|x(t)|^p \leqslant bE\|\varphi\|_\tau^p \leqslant bB_1$ and then $EV(t) \leqslant \lambda_2 EV(t) \leqslant \lambda_2 bB_1 < aB_2$. Particularly, $EV(t_0) \leqslant \lambda_2 EV(t_0) < aB_2$ and $EV(\hat{t}) \geqslant aE|x(\hat{t})|^p \geqslant aB_2$.

Define $t^* = \inf\{t \in [t_0, \hat{t}] \mid EV(t) \geqslant aB_2\}$ and then $t^* \in (t_0, \hat{t}]$, $EV(t^*) \geqslant aB_2$, and $EV(t) < aB_2$ for $t \in [t_0 - \tau, t^*)$.

Now we will proceed under two cases.

Case 1. Consider $t_{l-1} \leqslant t_0 < t^* < t_l$.

Under this case, we have $EV(t^*) = aB_2$ because of the continuity of $V(t)$ on (t_k, t_{k+1}) and $\lambda_2 EV(t^*) = \lambda_2 aB_2 > aB_2$. Define $\bar{t} = \sup\{t \in [t_0, t^*] \mid \lambda_2 EV(t) \leqslant aB_2\}$ and then $\bar{t} \neq t^*$, $\lambda_2 EV(\bar{t}) = aB_2$, and $\lambda_2 EV(t) \geqslant aB_2$ for $t \in [\bar{t}, t^*]$. Therefor, for any $t \in [\bar{t}, t^*]$ and $s \in [-\tau, 0]$, we have $EV(t+s) \leqslant aB_2 < \lambda_2 EV(t)$ and $\lambda_2 bB_1 < aB_2 < \lambda_2 EV(t)$, which yields $EV(t) > bB_1$, and then we have $E|x(t)|^p > B_1 \geqslant \rho$. Using condition (2), we have, when $t \in [\bar{t}, t^*]$,

$$E\mathscr{L}V(t) \leqslant -\gamma EV(t). \tag{18}$$

By virtue of Itô formula, we have

$$EV(t^*) - EV(\bar{t}) = \int_{\bar{t}}^{t^*} E\mathscr{L}V(s)\,ds \leqslant \int_{\bar{t}}^{t^*} -\gamma EV(s)\,ds \leqslant 0. \tag{19}$$

However,

$$EV(t^*) = aB_2 > \frac{aB_2}{\lambda_2} = EV(\bar{t}). \tag{20}$$

This contradiction gives

$$E|x(t)|^p \leqslant B_2 \quad \text{for } t \in [t_0, t_0+\beta). \tag{21}$$

Case 2. Consider $t_k \leqslant t^* < t_{k+1}$ for some $k \geqslant l$.

We first show $\lambda_2 EV(t_k^-) \leqslant aB_2$. We have two situations to contemplate: $k = l$ and $k > l$.

If $k = l$, we suppose $\lambda_2 EV(t_l^-) > aB_2$. Define $\bar{t} = \sup\{t \in [t_0, t_l) \mid \lambda_2 EV(t) \leqslant aB_2\}$ and then $\bar{t} \in (t_0, t_l)$ and $\lambda_2 EV(\bar{t}) = aB_2$. In light of the definition of $\bar{t}$, we have, for $t \in [\bar{t}, t_l)$ and $s \in [-\tau, 0]$,

$$\lambda_2 EV(t) \geqslant aB_2 \geqslant EV(t+s), \tag{22}$$

and, for $t \in [\bar{t}, t_l)$,

$$E|x(t)|^p \geqslant B_1 \geqslant \rho. \tag{23}$$

By virtue of condition (2), an analogous calculation of $EV(t_l^-) - EV(\bar{t})$ yields $EV(t_l^-) \leqslant EV(\bar{t})$; then we get

$$aB_2 < \lambda_2 EV(t_l^-) \leqslant \lambda_2 EV(\bar{t}) = aB_2. \tag{24}$$

If $k > l$, we suppose $\lambda_2 EV(t_k^-) > aB_2$. We will proceed under two subcases.

Subcase 1. Consider $\lambda_2 EV(t) > aB_2$ for all $t \in [t_{k-1}, t_k)$.

Under this situation, we have $\lambda_2 EV(t) > aB_2 \geqslant EV(t+s)$ and $E|x(t)|^p \geqslant \rho$ for all $t \in [t_{k-1}, t_k)$ and $s \in [-\tau, 0]$. In terms of condition (2), an analogous discussion as done in Case 1 gives

$$\begin{aligned} EV(t_k^-) - EV(t_{k-1}) &= \int_{tk-1}^{t_k^-} E\mathscr{L}V(s)\,ds \\ &\leqslant \int_{tk-1}^{t_k^-} -\gamma EV(s)\,ds \leqslant -\gamma\mu \frac{aB_2}{\lambda_2}. \end{aligned} \tag{25}$$

However, by virtue of condition (5),

$$\begin{aligned} EV(t_k^-) - EV(t_{k-1}) &\geqslant \frac{aB_2}{\lambda_2} - aB_2 = \left(\frac{1}{\lambda_2} - 1\right) aB_2 \\ &> -\gamma\mu \frac{aB_2}{\lambda_2}. \end{aligned} \tag{26}$$

This contradiction implies

$$\lambda_2 EV(t_k^-) \leqslant aB_2 \quad \text{for } t_k \leqslant t^* < t_{k+1},\ k \geqslant l. \tag{27}$$

Subcase 2. Consider $\lambda_2 EV(t) \leqslant aB_2$ for some $t \in [t_{k-1}, t_k)$.

Define $\bar{t} = \sup\{t \in [t_{k-1}, t_k) \mid \lambda_2 EV(t) \leqslant aB_2\}$ and then $\bar{t} \in [t_{k-1}, t_k)$ and $\lambda_2 EV(\bar{t}) = aB_2$. Using the definition of $\bar{t}$, we get, for $t \in [\bar{t}, t_k)$ and $s \in [-\tau, 0]$, $\lambda_2 EV(t) \geqslant aB_2 \geqslant EV(t+s)$. Since $\lambda_2 EV(t) \geqslant aB_2$, using the fact $\rho_1 \geqslant \rho, \lambda_2 bB_1 < aB_2$ and $b|x|^p \geqslant V(t,x)$, we can get $E|x(t)|^p \geqslant \rho$. By virtue of condition (2), we get, for $t \in [\bar{t}, t_k)$,

$$E\mathscr{L}V(t) \leqslant -\gamma EV(t). \tag{28}$$

An analogous discussion as done in the case $k = l$ gives $EV(\bar{t}) \geqslant EV(t_k^-)$. Then we have

$$aB_2 < \lambda_2 EV(t_k^-) \leqslant \lambda_2 EV(\bar{t}) = aB_2. \tag{29}$$

This contradiction gives

$$\lambda_2 EV(t_k^-) \leqslant aB_2 \quad \text{for } t_k \leqslant t^* < t_{k+1},\ k \geqslant l. \tag{30}$$

Now we claim $EV(t_k) < aB_2$. If $E|x(t_k^-)|^p \geqslant \rho$, we get $EV(t_k) \leqslant \lambda_1 EV(t_k^-) < \lambda_2 EV(t_k^-) < aB_2$. If $E|x(t_k^-)|^p < \rho$, we get $EV(t_k) \leqslant b\rho_1 < bB_1 < \lambda_2 bB_1 < aB_2$. That is, the following inequality holds:

$$EV(t_k) < aB_2. \tag{31}$$

Since $EV(t^*) \geqslant aB_2$, we have $t^* \neq t_k$ and $EV(t^*) = aB_2$.

If $\lambda_2 EV(t^*) \geqslant aB_2$ for all $t \in [t_k, t^*]$, then let $\bar{t} = t_k$ and we have $EV(\bar{t}) < aB_2$. Otherwise, let $\bar{t} = \sup\{t \in [t_k, t^*) \mid \lambda_2 EV(t) \leqslant aB_2\}$, and we have $EV(\bar{t}) < \lambda_2 EV(\bar{t}) = aB_2$. Since $EV(t^*) = aB_2$, we get $\bar{t} \in [t_k, t^*)$. Moreover, for $t \in [\bar{t}, t^*]$, we have $\lambda_2 EV(t) \geqslant aB_2 > EV(t+s)$ and, by virtue of $\lambda_2 bB_1 < aB_2 < \lambda_2 EV(t)$, we obtain $EV(t) > bB_1$ and then $E|x(t)|^p > B_1 > \rho$. In terms of condition (2) and Itô formula, we can obtain $EV(\bar{t}) \geqslant EV(t^*)$. But $EV(\bar{t}) < aB_2 = EV(t^*)$, which is a contradiction and yields

$$E|x(t)|^p \leqslant B_2 \quad \text{for } t \in [t_0, t_0 + \beta). \tag{32}$$

Now we get that, under condition (1) to condition (5), the solutions of (1) are p-moment uniformly bounded. Then we know that if $E\|\varphi\|_\tau^p \leqslant \rho_1$, there exists a constant $B > 0$, such that $E|x(t, t_0, \varphi)|^p \leqslant B$ for all $t \geqslant t_0 - \tau$, and, from the above proof, we have $\lambda_2 b\rho_1 < aB$.

Step 2. Now, let $B_3 > 0$ and assume, without loss of generality, that $B_3 > B$. Then, from the proof of uniform boundedness, there exists a constant $B_2 = B_2(B_3) > B_3$ for which if $E\|\varphi\|_\tau^p \leqslant B_3$, then $E|x(t)|^p \leqslant B_2$ for $t \geqslant t_0 - \tau$.

Take a constant d satisfying $0 < d \leqslant \min\{aB - b\rho_1, ((\lambda_2 - \lambda_1)/\lambda_2)aB\}$, $N = \min\{n > ((bB_2 - aB)/d)\}$, and $T = \alpha + (2N-1)(\alpha+\tau)$.

Let $x(t) = x(t, t_0, \varphi)$ be a solution of (1) with $E\|\varphi\|_\tau^p \leqslant B_3$, $t_0 \in [t_{l-1}, t_l)$. We will show $E|x(t)|^p \leqslant B$ for $t \geqslant t_0 + T$.

Given a positive number A satisfying $aB \leqslant A \leqslant bB_2$ and $j \geqslant l$, we will show that if $EV(t) \leqslant A$ for $t \in [t_j - \tau, t_j)$ and $\lambda_2 EV(t_j^-) \leqslant A$, then $EV(t) \leqslant A$ for $t \geqslant t_j$ and $\lambda_2 EV(t_{j+1}^-) \leqslant A$.

For the sake of contradiction, suppose that there exists a constant $t \in [t_j, t_{j+1})$ for which $EV(t) > A$ and define

$$t^* = \inf\left\{t \in \left[t_j, t_{j+1}\right) \mid EV(t) \geqslant A\right\}. \tag{33}$$

Note that $EV(t_j) < A$, and we have that if $E|x(t_j^-)|^p \geqslant \rho$, then $EV(t_k) \leqslant \lambda_1 EV(t_j^-) < \lambda_2 EV(t_j^-) \leqslant A$. If $E|x(t_k^-)|^p < \rho$, we have $EV(t_j) \leqslant b\rho_1 < \lambda_2 b\rho_1 < aB \leqslant A$. Then we get $t^* \neq t_j$, $EV(t^*) = A$, and $EV(t) \leqslant A$ for $t \in (t_j, t_{j+1}]$.

If $\lambda_2 EV(t) > A$ for all $t \in [t_j, t_{j+1})$, we let $\bar{t} = t_j$, and then $EV(\bar{t}) = EV(t_j) < A$. Otherwise, let $\bar{t} = \sup\{t \in [t_j, t^*] \mid \lambda_2 EV(t) \leqslant A\}$, and we get $EV(\bar{t}) \leqslant \lambda_2 EV(\bar{t}) = A$. Since $\lambda_2 EV(t^*) = \lambda_2 A > A$, $\bar{t} \neq t^*$. For $t \in [\bar{t}, t^*]$ and $s \in [-\tau, 0]$, we have $\lambda_2 EV(t) \geqslant A \geqslant EV(t+s)$. Moreover, for $t \in [\bar{t}, t^*]$,

$$\lambda_2 EV(t) \geqslant A \geqslant aB > \lambda_2 b\rho_1, \tag{34}$$

and we get $E|x(t)|^p \geqslant \rho_1 \geqslant \rho$. By virtue of condition (2) and Itô formula, we can get $EV(\bar{t}) \geqslant EV(t^*)$. However, $EV(t^*) = A > EV(\bar{t})$.

Now we have proven $EV(t) \leqslant A$ for $t \in [t_j, t_{j+1})$, and we are on the position to show $\lambda_2 EV(t_{j+1}^-) \leqslant A$. This will follow in the same way as the arguments used in the proof of uniform boundedness, where we show $\lambda_2 EV(t_k^-) \leqslant aB_2$ for the case $k > l$; we just need to replace k by $j+1$ and ab_2 by A.

By induction, we get that if $EV(t) \leqslant A$ for $t \in [t_j - \tau, t_j)$ and $\lambda_2 EV(t_j^-) \leqslant A$, then $EV(t) \leqslant A$ for all $t \geqslant t_j$ and $\lambda_2 EV(t_k^-) \leqslant A$ for $k \geqslant j+1$.

Next, we will show $EV(t) \leqslant A - d$ for $t \in [t_{j+1}, t_{j+2})$, if $EV(t) \leqslant A$ for all $t \geqslant t_j$ and $\lambda_2 EV(t_k^-) \leqslant A$, $k \geqslant j$.

We first show $EV(t_{j+1}) \leqslant A - d$. This can be easily verified under two situations: iIf $E|x(t_{j+1}^-)|^p \leqslant \rho$, we have $EV(t_{j+1}) \leqslant b\rho_1 \leqslant aB - d \leqslant A - d$; if $E|x(t_{j+1}^-)|^p > \rho$, $EV(t_{j+1}) < \lambda_1 EV(t_{j+1}) = (\lambda_1/\lambda_2)\lambda_2 EV(t_{j+1}^-) \leqslant (\lambda_1/\lambda_2)A < A - d$.

In order to verify $EV(t) \leqslant A - d$ for all $t \in [t_{j+1}, t_{j+2})$, suppose that $EV(t) > A - d$ for some $t \in [t_{j+1}, t_{j+2})$. Let $t^* = \inf\{t \in [t_{j+1}, t_{j+2}) \mid EV(t) \geqslant A - d\}$; we know $t^* \neq t_{j+1}$ and then $EV(t^*) = A - d$ and $\lambda_2 EV(t^*) = \lambda_2(A - d) > A$.

If $\lambda_2 EV(t) > A$ for all $t \in [t_{j+1}, t^*]$, let $\bar{t} = t_{j+1}$, $EV(\bar{t}) = EV(t_{j+1}) < A - d$.

If $\lambda_2 EV(t) > A$ for some $t \in (t_{j+1}, t^*]$, let $\bar{t} = \sup\{t \in [t_{j+1}, t^*] \mid \lambda_2 EV(t) \leqslant A\}$ and we know $\bar{t} \neq t^*$, $EV(\bar{t}) = A/\lambda_2$.

For $t \in [\bar{t}, t^*]$ and $s \in [-\tau, 0]$, $\lambda_2 EV(t) \geqslant A > A - d > EV(t + s)$ and $EV(t) \geqslant A/\lambda_2 > aB/\lambda_2 > b\rho_1$, and we get $E|x(t)|^p > \rho_1 \geqslant \rho$. In terms of condition (2) and Itô formula, we can get $EV(t^*) < EV(\bar{t})$. However, $EV(t^*) = A - d > EV(\bar{t})$, which yields

$$EV(t) \leqslant A - d. \tag{35}$$

Applying our results to successive intervals of the form $[t_k, t_{k+1})$ for $k \geqslant j + 1$, we can get $EV(t) \leqslant A - d$ for $t \geqslant t_{j+1}$.

Now we need a fact $\lambda_2 EV(t_{j+2}^-) \leqslant A - d$. This can be verified just as we did in the proof of uniform boundedness, where we show $\lambda_2 EV(t_k^-) \leqslant aB_2$ for the case $k > l$.

Take $t_{k^{(i)}} \in \{t_j, j = l, l + 1, \ldots\}$ satisfying $t_{k^{(i-1)}} + \tau \leqslant t_{k^{(i)}} \leqslant t_{k^{(i-1)}+1} + \tau$. Take $A = bB_2$, when $t \geqslant t_{k^{(2N)}}$, and we get $EV(t) \leqslant bB_2 - Nd < aB$. Since $t_{k^{(2N)}} \leqslant t_{k^{(1)}} + (2N - 1)(\alpha + \tau) \leqslant t_0 + \alpha + (2N - 1)(\alpha + \tau) = t_0 + T$, we have $EV(t) \leqslant aB$ when $t > t_0 + T$. By virtue of condition (1), $E|x(t)|^p \leqslant B$ for $t \geqslant t_0 + T$, which completes the proof. □

Remark 5. Theorem 4 considers that a bounded system without impulse can tolerate what kind of impulsive effects to hold the boundedness. It is not surprising that condition (3) to condition (5) should be satisfied: the interval of impulsive moments (μ) should be large and impulsive strength (λ_1) should be small.

5. Examples

In this section, we present two examples to illustrate our results.

Example 1. Consider the following impulsive stochastic delay differential system:

$$dx(t) = \left(\frac{1}{2}x(t) + \frac{1}{2x(t)}\right)dt + x\left(t - \frac{1}{20}\right)dB(t),$$
$$t > 0, \quad t \neq \frac{k}{10}, \quad k = 1, 2, \ldots, \tag{36}$$
$$x\left(\frac{k}{10}\right) = \frac{\sqrt{2}}{2}x\left(\left(\frac{k}{10}\right)^-\right),$$

where $B(t)$ is a one-dimension Brownian motion.

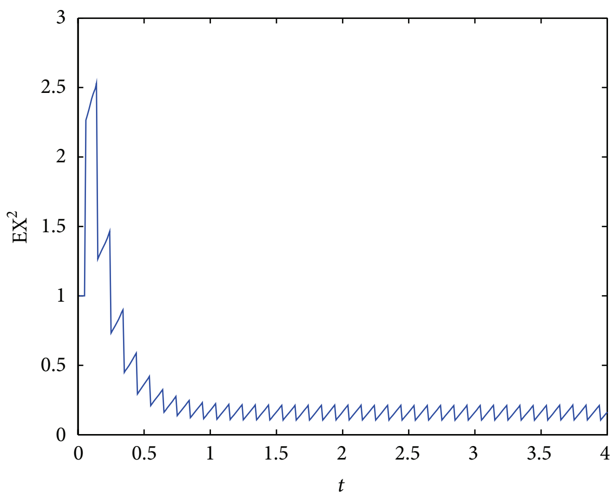


Figure 1: Mean square uniform ultimate boundedness of solution of system (36).

Define $V(t, x) = x^2$; the smoothness requirement is satisfied. Let $a = b = 1$ and $p = 2$; condition (1) of Theorem 2 follows. For any solution $x(t)$ of system (36), we have

$$\begin{aligned}\mathscr{L}V(t, x) &= 2x\left(\frac{1}{2}x(t) + \frac{1}{2x(t)}\right) + x^2\left(t - \frac{1}{20}\right) \\ &= x^2(t) + 1 + x^2\left(t - \frac{1}{20}\right).\end{aligned} \tag{37}$$

Take $\lambda = 1/2$; condition (3) of Theorem 2 is satisfied.

Now let $\rho = 1$; then, when $|x(t)|^2 \geqslant 1$ and $V(t, x) \geqslant \lambda V(t, x(t - \tau))$, that is, $x^2(t) \geqslant (1/2)x^2(t - 1/20)$, we have

$$\mathscr{L}V(t, x) \leqslant x^2(t) + x^2(t) + 2x^2(t) = 4x^2(t) = 4V(t, x). \tag{38}$$

Then let $\gamma = 4$; condition (2) of Theorem 2 is verified.

Condition (4) of Theorem 2 can be verified by taking $\rho_1 = 1$.

Take $\alpha = 1/10$ and then $\alpha\gamma = (1/10) \times 4 = 2/5 < 1/2 = 1 - \lambda$; condition (5) of Theorem 2 is verified.

Therefore, according to Theorem 2, solutions of system (36) are mean square uniformly ultimately bounded. The boundedness can be read from Figure 1, where we take initial condition $x(t) = 1, t \in [-1/20, 0]$.

To see the contribution of impulsive effect on boundedness, we consider the following system:

$$dx(t) = \left(\frac{1}{2}x(t) + \frac{1}{2x(t)}\right)dt + x\left(t - \frac{1}{20}\right)dB(t),$$
$$t > 0, \tag{39}$$

which is the situation of system (36) without impulses. It is easy to be verified that system (39) is unbounded; see Figure 2, where we also take initial condition $x(t) = 1, t \in [-1/20, 0]$.

Now we give another example to illustrate the correctness of Theorem 4.

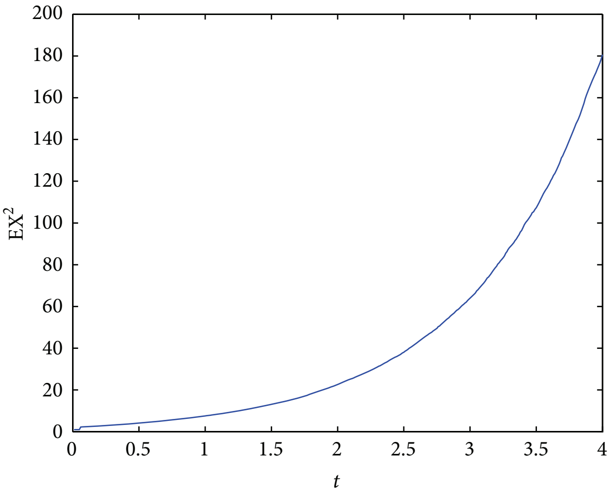


FIGURE 2: Unboundedness of solution of system (39).

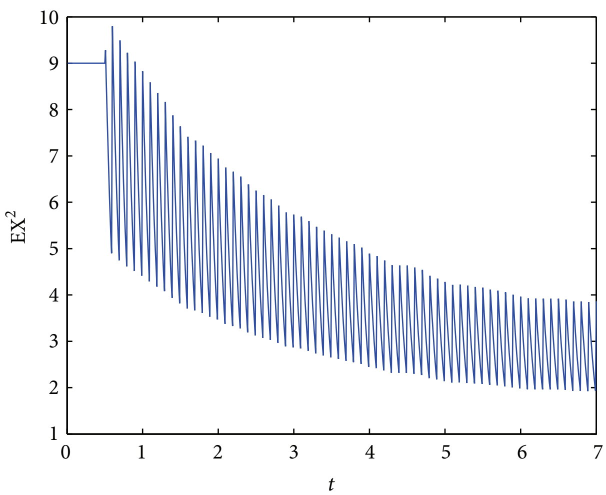


FIGURE 3: Mean square uniform ultimate boundedness of solution of system (40).

Example 2. Consider

$$dx(t) = \left(-4x(t) + \frac{1}{2x(t)}\right)dt + x\left(t - \frac{1}{2}\right)dB(t), \quad t > 0, \quad t \neq 2k, \quad k = 1, 2, \ldots, \quad x(2k) = \sqrt{2}x\left((2k)^{-}\right), \tag{40}$$

where $B(t)$ is a one-dimension Brownian motion.

Define $V(t, x) = x^2$; the smoothness requirement is satisfied. Let $a = b = 1$ and $p = 2$; condition (1) of Theorem 4 follows. For any solution $x(t)$ of system (40), we have

$$\begin{aligned} \mathscr{L}V(t, x) &= 2x\left(-4x(t) + \frac{1}{2x(t)}\right) + x^2\left(t - \frac{1}{2}\right) \\ &= -8x^2(t) + 1 + x^2\left(t - \frac{1}{20}\right). \end{aligned} \tag{41}$$

Take $\lambda_1 = 2$, condition (3) of Theorem 4 is satisfied.

Now let $\rho = 1$ and $\lambda_2 = 3$; then, when $|x(t)|^2 \geqslant 1$ and $V(t, x) \geqslant \lambda_2 V(t, x(t - \tau))$, that is, $3x^2(t) \geqslant x^2(t - 1/2)$, we have

$$\begin{aligned} \mathscr{L}V(t, x) &\leqslant -8x^2(t) + x^2(t) + 3x^2(t) \\ &= -4x^2(t) = -4V(t, x). \end{aligned} \tag{42}$$

Then, let $\gamma = 4$; condition (2) of Theorem 2 is verified.

Condition (4) of Theorem 2 can be verified by taking $\rho_1 = 2$.

Take $\mu = 2$ and then $\mu\gamma = 2 \times 8 = 16 > 3 - 1 = \lambda_2 - 1$ and condition (5) of Theorem 4 is verified.

Therefore, according to Theorem 4, solutions of system (40) are mean square uniformly ultimately bounded. The boundedness can be seen in Figure 3, where we take initial condition $x(t) = 3, t \in [-1/2, 0]$.

We also present the simulation of system (40) without impulsive effects; that is,

$$dx(t) = \left(-4x(t) + \frac{1}{2x(t)}\right)dt + x\left(t - \frac{1}{2}\right)dB(t), \quad t > 0. \tag{43}$$

The property of system (43) can be read from Figure 4, where we take initial condition $x(t) = 3, t \in [-1/2, 0]$.

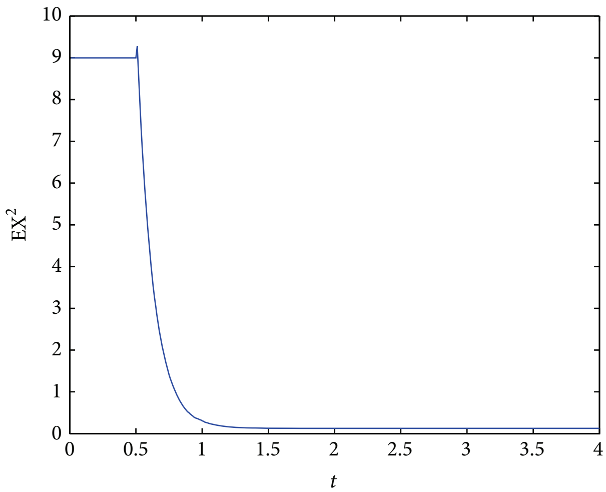


FIGURE 4: Simulation of system (43).

Conflict of Interests

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

Acknowledgments

This work is supported by the foundation under Grant HIT.IBRSEM.A.2014015 and by the National Natural Scientific Foundation of China under Grants 11271101, 61104193, and 61333001.

References

[1] L. Zu, D. Jiang, and D. O'Regan, "Stochastic permanence, stationary distribution and extinction of a single-species nonlinear diffusion system with random perturbation," *Abstract and Applied Analysis*, vol. 2014, Article ID 320460, 14 pages, 2014.

[2] E. Miao, H. Shu, and Y. Che, "Practical Stability in the pth Mean for Itô Stochastic Differential Equations," *Mathematical Problems in Engineering*, vol. 2010, Article ID 380304, 12 pages, 2012.

[3] X. Liu and Q. Wang, "The method of Lyapunov functionals and exponential stability of impulsive systems with time delay," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 66, no. 7, pp. 1465–1484, 2007.

[4] M. de la Sen, "Global stability of polytopic linear time-varying dynamic systems under time-varying point delays and impulsive controls," *Mathematical Problems in Engineering*, vol. 2010, Article ID 693958, 33 pages, 2010.

[5] X. Li, H. Akca, and X. Fu, "Uniform stability of impulsive infinite delay differential equations with applications to systems with integral impulsive conditions," *Applied Mathematics and Computation*, vol. 219, no. 14, pp. 7329–7337, 2013.

[6] Y. Zhang and J. Sun, "Boundedness of the solutions of impulsive differential systems with time-varying delay," *Applied Mathematics and Computation*, vol. 154, no. 1, pp. 279–288, 2004.

[7] L. Berezansky and E. Braverman, "Exponential boundedness of solutions for impulsive delay differential equations," *Applied Mathematics Letters*, vol. 9, no. 6, pp. 91–95, 1996.

[8] I. M. Stamova, "Lyapunov method for boundedness of solutions of nonlinear impulsive functional differential equations," *Applied Mathematics and Computation*, vol. 177, no. 2, pp. 714–719, 2006.

[9] P. Cheng, F. Deng, and L. Wang, "The method of Lyapunov function and exponential stability of impulsive delay systems with delayed impulses," *Mathematical Problems in Engineering*, vol. 2013, Article ID 458047, 7 pages, 2013.

[10] X. Liu and G. Ballinger, "Boundedness for impulsive delay differential equations and applications to population growth models," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 53, no. 7-8, pp. 1041–1062, 2003.

[11] I. M. Stamova, "Boundedness of impulsive functional differential equations with variable impulsive perturbations," *Bulletin of the Australian Mathematical Society*, vol. 77, no. 2, pp. 331–345, 2008.

[12] I. Stamova, "Eventual stability and eventual boundedness for impulsive differential equations with "supremum"," *Mathematical Modelling and Analysis*, vol. 16, no. 2, pp. 304–314, 2011.

[13] A. Li and X. Song, "Stability and boundedness of nonlinear impulsive systems in terms of two measures via perturbing Lyapunov functions," *Journal of Mathematical Analysis and Applications*, vol. 375, no. 1, pp. 276–283, 2011.

[14] B. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[15] W. Zhang, H. Zhang, and B. Chen, "Generalized Lyapunov equation approach to state-dependent stochastic stabilization/detectability criterion," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1630–1642, 2008.

[16] W. Zhang and B. Chen, "H-representation and applications to generalized Lyapunov equations and linear stochastic systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3009–3022, 2012.

[17] F. Wu and P. E. Kloeden, "Mean-square random attractors of stochastic delay differential equations with random delay," *Discrete and Continuous Dynamical Systems B*, vol. 18, no. 6, pp. 1715–1734, 2013.

[18] B. Song, J. H. Park, Z. Wu, and Y. Zhang, "New results on delay-dependent stability analysis for neutral stochastic delay systems," *Journal of the Franklin Institute: Engineering and Applied Mathematics*, vol. 350, no. 4, pp. 840–852, 2013.

[19] X. Li, J. Zou, and E. Zhu, "pth moment exponential stability of impulsive stochastic neural networks with mixed delays," *Mathematical Problems in Engineering*, vol. 2012, Article ID 175934, 20 pages, 2012.

[20] K. Wu and X. Ding, "Stability and stabilization of impulsive stochastic delay differential equations," *Mathematical Problems in Engineering*, vol. 2012, Article ID 176375, 16 pages, 2012.

[21] X. Wu, W. Zhang, and Y. Tang, "pth moment stability of impulsive stochastic delay differential systems with Markovian switching," *Communications in Nonlinear Science and Numerical Simulation*, vol. 18, no. 7, pp. 1870–1879, 2013.

[22] P. Cheng, F. Deng, and F. Yao, "Exponential stability analysis of impulsive stochastic functional differential systems with delayed impulses," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 6, pp. 2104–2114, 2014.

[23] F. Yao, F. Deng, and P. Cheng, "Exponential stability of impulsive stochastic functional differential systems with delayed impulses," *Abstract and Applied Analysis*, vol. 2013, Article ID 548712, 8 pages, 2013.

[24] P. Cheng, F. Deng, and Y. Peng, "Robust exponential stability and delayed-state-feedback stabilization of uncertain impulsive stochastic systems with time-varying delay," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 4740–4752, 2012.

[25] J. Yang, S. Zhong, and W. Luo, "Mean square stability analysis of impulsive stochastic differential equations with delays," *Journal of Computational and Applied Mathematics*, vol. 216, no. 2, pp. 474–483, 2008.

[26] S. Peng and B. Jia, "Some criteria on pth moment stability of impulsive stochastic functional differential equations," *Statistics & Probability Letters*, vol. 80, no. 13-14, pp. 1085–1092, 2010.

[27] 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.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 439537, 16 pages
http://dx.doi.org/10.1155/2015/439537

Research Article

Optimal Dividend and Capital Injection Strategies in the Cramér-Lundberg Risk Model

Yan Li[1] and Guoxin Liu[2]

[1] *School of Insurance and Economics, University of International Business and Economics, Beijing 100029, China*
[2] *Shijiazhuang Tiedao University, Shijiazhuang 050043, China*

Correspondence should be addressed to Yan Li; email.liyan@163.com

Received 12 March 2014; Revised 25 August 2014; Accepted 15 September 2014

Academic Editor: Guangchen Wang

We discuss the optimal dividend and capital injection strategies in the Cramér-Lundberg risk model. The value function $V(x)$ is defined by maximizing the discounted value of the dividend payment minus the penalized discounted capital injection until the time of ruin. It is shown that $V(x)$ can be characterized by the Hamilton-Jacobi-Bellman equation. We find the optimal dividend barrier b, the optimal upper capital injection barrier 0, and the optimal lower capital injection barrier $-z^*$. In the case of exponential claim size especially, we give an explicit procedure to obtain b, $-z^*$, and the value function $V(x)$.

1. Introduction

In the modern theory of risk, people tend to study the cost of postponing or avoiding outright ruin; that is, ruin does not mean the end of the game but only the necessity of raising additional money. So the risk process can continue if there is a suitable injection of surplus.

Borch [1] pointed out that it was a good investment to rescue an insolvent insurance company, provided that its deficit was not too large. He studied this problem for a random walk model and suggested that the company should be rescued only if the deficit was smaller than the expected profits from the rescue operation.

For a diffusion model, Sethi and Taksar [2] considered the problem of finding an optimal financing mix of retained earnings and external equity for maximizing the value of a corporation. They showed that the optimal policy can be characterized in terms of two threshold parameters. Løkka and Zervos [3] studied the same problem with possibility of bankruptcy in a model of Brownian motion with drift. Depending on the relationships between the coefficients, the optimal strategy requires the consideration of two auxiliary suboptimal models. For more references in diffusion model see He and Liang [4, 5], and so forth.

As pointed out by Bäuerle [6], the classical approach is to model the liquid assets or risk reserve process of the insurance company as a piecewise deterministic Markov process (PDMP). However, within this setting the control problem is very hard and many characteristics of the risk process can not be calculated in closed form.

For the Cramér-Lundberg risk model without bankruptcy (i.e., the shareholders will inject capital to cover the deficit whatever serious it is) the optimal dividend problem was studied. See, for example, Dickson and Waters [7], Gerber et al. [8], Kulenko and Schmidli [9], and so forth. This capital injection strategy makes sense for itself; at the same time we notice that the injected capital can be viewed as an investment. Therefore the shareholders should consider the return of it. If the injected amount is small enough to the shareholders to earn positive net profit, they accept to do so and survive the company. Otherwise, they will refuse to inject capital anymore and ruin occurs. So what is the optimal capital injection strategy is worth to be discussed.

In this paper, we will discuss the optimal dividend payment and capital injection strategies in the Cramér-Lundberg risk model. The objective is to maximize the discounted dividends payments minus the penalized discounted capital injections. Through the discussion of the optimal capital injection strategy, we find the maximal deficit which the shareholders can bear. Moreover, from the mathematical point of view we give a rigorous proof that it is optimal to inject capital once the reserves are below 0, that is, the moment ruin occurs (in the previous literature about capital injection strategy, considering discounting, it could not be optimal to inject capital before it is really necessary. Therefore, the shareholders postpone the injection as long as possible and just conjecture that it is optimal to do so when the reserves become 0).

Suppose the reserve process of an insurance company at time t is

$$X_t = x + ct - \sum_{i=1}^{N_t} Y_i, \tag{1}$$

where $x \in R$ is the initial capital, $c > 0$ is the premium rate, $\{N_t, t \geq 0\}$ is a Poisson process with intensity $\lambda > 0$, and $\{Y_i, i \geq 1\}$ is a sequence of strictly positive i.i.d. random variables with the distribution function $F(x)$. In addition, $\{Y_i, i \geq 1\}$ and $\{N_t, t \geq 0\}$ are independent. We assume that $EY_i = \mu < \infty$ and $F(x)$ is continuous. $\{X_t\}$ is on a filtrated probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, where $\{\mathcal{F}_t\}_{t\geq 0}$ is the smallest right-continuous filtration such that $\{X_t\}$ is adapted. Let P_x and E_x denote the probability and the expectation with initial capital x, respectively.

Now we enrich the model with a strategy $\pi = \{(D_t, Z_t)\}$. $\{D_t\}$ and $\{Z_t\}$ denote the aggregate dividends and capital injections paid up to time t, respectively. The strategy π is admissible if

(1) $\{D_t\}$ is càdlàg, increasing and adapted processes with $D_{0-} = 0$;

(2) $\{Z_t\}$ is càglàd, increasing and adapted processes with $Z_0 = 0$.

The reserve turns to

$$X_t^\pi = X_t - D_t + Z_t. \tag{2}$$

Since the strategy π will not assure that the process $\{X_t^\pi\}$ is always larger than 0, ruin is possible. The ruin time is defined by

$$T^\pi = \inf\{t \geq 0, X_{t+}^\pi < 0\}. \tag{3}$$

The value of a strategy π is

$$V^\pi(x) = E_x\left[\int_{0-}^{T^\pi -} e^{-\delta t}\mathrm{d}D_t - \phi\int_0^{T^\pi} e^{-\delta t}\mathrm{d}Z_t\right], \tag{4}$$

where $\delta > 0$ is a discounted factor and $\phi > 1$ is a penalizing factor. The point 0 being included in the integration area is for the reason of taking an immediate dividend $D_0 > 0$ into the value. Our purpose is to maximize $V^\pi(x)$. The *value function* is defined by

$$V(x) = \sup_{\pi\in\Pi} V^\pi(x), \tag{5}$$

where Π denotes the set of all admissible strategies.

The paper is organized as follows. In Section 2, the dividend strategy is constrained by a restricted density. Some properties of the value function $V(x)$ are proved. We show that $V(x)$ can be characterized by the Hamilton-Jacobi-Bellman equation. Moreover, if $V(x)$ is concave, the optimal dividend and capital injection strategies are both barrier strategies. If we remove the constraint on the dividend strategy, the results on $V(x)$ and optimal strategies are extended in Section 3. In the last section, we give an explicit procedure to obtain the optimal dividend barrier b, the optimal lower capital injection barrier $-z^*$, and the value function $V(x)$ when the claim size is exponentially distributed.

2. Dividends with Restricted Densities

In this section, we study this optimization problem under the constraint that the dividends are paid at a dividend rate, which is bounded by a positive constant u_0; that is, $0 \leq U_t \leq u_0 < \infty$. Then $D_t = \int_0^t U_s \mathrm{d}s$ and

$$V^\pi(x) = E_x\left[\int_0^{T^\pi} e^{-\delta t}U_t\mathrm{d}t - \phi\int_0^{T^\pi} e^{-\delta t}\mathrm{d}Z_t\right]. \tag{6}$$

In this section, Π^r denotes the set of all admissible restricted strategies and $\pi = (U_t, Z_t)$. So the value function

$$V(x) = \sup_{\pi\in\Pi^r} V^\pi(x). \tag{7}$$

2.1. The Value Function $V(x)$. $V(x)$ has the following properties.

Lemma 1. *If the capital injection strategy is defined by*

$$Z_t = \max\left\{-\inf_{0\leq s<t}(X_s - D_s), 0\right\}, \tag{8}$$

Then, for $x \in R_+$, the value under any dividend strategy $\{D_t\}$ is bounded from below by $-\phi\lambda\mu/\delta$.

Proof. Under this assumption, ruin time is ∞. The maximal amount of capital injection may be that the shareholders cover all the claims. If we are not considering the dividends, value under such a strategy is the worst one. Using the time of the kth claim T_k is Gamma $\Gamma(\lambda, k)$, so

$$E\left[\sum_{k=1}^{\infty} Y_k e^{-\delta T_k}\right] = \mu\sum_{k=1}^{\infty}\left(\frac{\lambda}{\lambda+\delta}\right)^k = \frac{\lambda\mu}{\delta}. \tag{9}$$

The value is bounded from below by $-\phi\lambda\mu/\delta$. □

Lemma 2. *$V(x)$ is increasing and Lipschitz continuous on $(-\infty, \infty)$. Moreover, $0 \leq V(x) \leq u_0/\delta$ and $\lim_{x\to\infty} V(x) = u_0/\delta$.*

Proof. Obviously, $V(x)$ is increasing. For $x < 0$, if define the strategy π as $Z_t = U_t = 0$, then $V(x) \geq V^\pi(x) = 0$. Because $V(x)$ is increasing, $V(x) \geq 0$ for $x \in R$. If $U_t = u_0$, $Z_t = 0$, then

$$V(x) \leq \int_0^\infty u_0 e^{-\delta t}\mathrm{d}t = \frac{u_0}{\delta}. \tag{10}$$

Consider a strategy $\pi = (U_t, Z_t)$, where $U_t = u_0$ and $Z_t = \max\{-\inf_{0\le s<t}(X_s - D_s), 0\}$. Then $T^\pi = \infty$. Define $\tau_x^\pi = \inf\{t : x + (c - u_0)t - \sum_{i=1}^{N_t} Y_i < 0\}$. Using Lemma 1,

$$\begin{aligned}\int_0^{T^\pi} e^{-\delta t}\mathrm{d}Z_t &= \int_0^{\infty} e^{-\delta t}\mathrm{d}Z_t = \int_{\tau_x^\pi}^{\infty} e^{-\delta t}\mathrm{d}Z_t \\ &= e^{-\delta\tau_x^\pi}\int_0^{\infty} e^{-\delta t}\mathrm{d}Z_{t+\tau_x^\pi} \le e^{-\delta\tau_x^\pi}\phi\frac{\lambda\mu}{\delta}.\end{aligned} \tag{11}$$

When $x \to \infty$, then $\tau_x^\pi \to \infty$ and $P_x(\int_0^\infty e^{-\delta t}\mathrm{d}Z_t > \varepsilon) \to 0$. So we have

$$\begin{aligned}V(x) &\ge V^\pi(x) \\ &\ge E\left[\int_0^{\tau_x^\pi} u_0 e^{-\delta t}\mathrm{d}t - \phi\int_0^{\infty} e^{-\delta t}\mathrm{d}Z_t\right] \longrightarrow \frac{u_0}{\delta}.\end{aligned} \tag{12}$$

Combining with (10), we have $\lim_{x\to\infty}V(x) = u_0/\delta$.

For $x \ge 0$, let $h > 0$ be small. Define

$$\begin{aligned}U_t &= \begin{cases} 0, & \text{if } 0 \le t < h\wedge T_1, \\ \widetilde{U}_{t-h}, & \text{if } t\wedge T_1 \ge h, \\ 0, & \text{if } T_1 < h,\end{cases} \\ Z_t &= \begin{cases} 0, & \text{if } 0 \le t < h\wedge T_1, \\ \widetilde{Z}_{t-h}, & \text{if } t\wedge T_1 \ge h, \\ 0, & \text{if } T_1 < h,\end{cases}\end{aligned} \tag{13}$$

where $\widetilde{\pi} = (\widetilde{U}, \widetilde{Z}) \in \Pi^r$ is for the initial capital $x + ch$. While $P(T_1 \ge h) = e^{-\lambda h}$, then

$$\begin{aligned}V(x) &\ge V^\pi(x) \\ &= E\left[E\left[\int_0^{T^\pi} e^{-\delta t}U_t\mathrm{d}t - \phi\int_0^{T^\pi} e^{-\delta t}\mathrm{d}Z_t \mid T_1\right]\right] \\ &= P(T_1 \ge h) \\ &\quad\times E\left[\int_h^{T^\pi} e^{-\delta t}U_t\mathrm{d}t - \phi\int_h^{T^\pi} e^{-\delta t}\mathrm{d}Z_t \mid T_1 \ge h\right] \\ &= e^{-\lambda h}E\left[E\left[\int_h^{T^\pi} e^{-\delta t}U_t\mathrm{d}t - \phi\int_h^{T^\pi} e^{-\delta t}\mathrm{d}Z_t \mid \mathcal{F}_h\right]\right] \\ &= e^{-\lambda h}E\left[e^{-\delta h}E\left[\int_0^{T^\pi - h} e^{-\delta t}U_{t+h}\mathrm{d}t\right.\right. \\ &\qquad\qquad \left.\left. -\phi\int_0^{T^\pi - h} e^{-\delta t}\mathrm{d}Z_{t+h} \mid \mathcal{F}_h\right]\right] \\ &= e^{-(\lambda+\delta)h}E\left[\int_0^{T^\pi - h} e^{-\delta t}\widetilde{U}_t\mathrm{d}t - \phi\int_0^{T^\pi - h} e^{-\delta t}\mathrm{d}\widetilde{Z}_t\right] \\ &= e^{-(\lambda+\delta)h}V^{\widetilde{\pi}}(x + ch)\end{aligned} \tag{14}$$

and so

$$V(x) \ge \sup_{\widetilde{\pi}\in\Pi^r} e^{-(\lambda+\delta)h}V^{\widetilde{\pi}}(x + ch) = e^{-(\lambda+\delta)h}V(x + ch). \tag{15}$$

From the bounded property of $V(x)$, we have

$$\begin{aligned}0 &\le V(x + ch) - V(x) \le V(x + ch)\left(1 - e^{-(\lambda+\delta)h}\right) \\ &\le V(x + ch)(\lambda + \delta)h \le \frac{u_0}{\delta}(\lambda + \delta)h.\end{aligned} \tag{16}$$

Let the shareholder inject h and follow the optimal strategy afterwards when $x < 0$. So $V(x) \ge V(x + h) - \phi h$; that is,

$$V(x + h) - V(x) \le \phi h. \tag{17}$$

Thus $V(x)$ is Lipschitz continuous on $(-\infty, \infty)$. □

2.2. HJB Equation and the Optimal Strategy. In this section, we will derive the HJB equation satisfied by the value function $V(x)$ and discuss the optimal strategy π^*.

Similar to the discussion in Azcue and Muler [10], the following dynamic programming principle holds:

$$\begin{aligned}V(x) = \sup_\pi E_x\Bigg[&\int_0^{\tau\wedge T^\pi} e^{-\delta t}U_t\mathrm{d}t \\ &-\phi\int_0^{\tau\wedge T^\pi} e^{-\delta t}\mathrm{d}Z_t + e^{-\delta(\tau\wedge T^\pi)}V\left(X^\pi_{\tau\wedge T^\pi}\right)\Bigg]\end{aligned} \tag{18}$$

for $x \in R_+$ and any $\{\mathcal{F}_t\}$-stopping time τ. This principle may serve us to derive the HJB equation.

For $x \ge 0$, $\varepsilon > 0$, and any admissible strategy π, define $\sigma^\pi = \inf\{t \ge 0, X_t^\pi \notin (x - \varepsilon, x + \varepsilon)\}$. Choose ε small enough; then $\sigma^\pi < T^\pi$. Let $\tau^\pi = \sigma^\pi \wedge h$, $h > 0$. So $\tau^\pi \to 0$ a.s. $h \to 0$. Applying Itô formula into $e^{-\delta\tau^\pi}V(X^\pi_{\tau^\pi})$, we have

$$\begin{aligned}&e^{-\delta\tau^\pi}V\left(X^\pi_{\tau^\pi}\right) \\ &\quad= V\left(X_0^\pi\right) \\ &\qquad+ \int_0^{\tau^\pi} e^{-\delta s}\left(c - U_s\right)V'\left(X^\pi_{s-}\right) - \delta e^{-\delta s}V\left(X^\pi_{s-}\right)\mathrm{d}s \\ &\qquad+ \sum_{\substack{0\le s\le\tau^\pi \\ X^\pi_{s-}\ne X^\pi_s}} e^{-\delta s}\left[V\left(X^\pi_s\right) - V\left(X^\pi_{s-}\right)\right] \\ &\qquad+ \sum_{\substack{0\le s<\tau^\pi \\ X^\pi_s\ne X^\pi_{s+}}} e^{-\delta s}\left[V\left(X^\pi_{s+}\right) - V\left(X^\pi_s\right)\right].\end{aligned} \tag{19}$$

If $U_t > c$, $\{X_t^\pi\}$ could become negative before the first claim and so dividends lead to ruin. Considering the early penalty, this dividend strategy with $U_t > c$ at a point where $X_t^\pi = 0$ will not be optimal. So we can assume without restriction that $\{Z_t\}$ only increases when the claim arrives; that is, it is a pure jump process. Thus

$$\sum_{\substack{0\le s<\tau^\pi \\ X^\pi_s\ne X^\pi_{s+}}} e^{-\delta s}\left[V\left(X^\pi_{s+}\right) - V\left(X^\pi_s\right)\right] = \phi\int_0^{\tau^\pi} e^{-\delta s}\mathrm{d}Z_s. \tag{20}$$

When claim arrives, $X^{\pi}_{s-} \neq X^{\pi}_{s}$. Then

$$\begin{aligned} M(\tau^{\pi}) &= M(\sigma^{\pi} \wedge h) \\ &= \sum_{\substack{0\le s\le \tau^{\pi} \\ X^{\pi}_{s-}\neq X^{\pi}_{s}}} e^{-\delta s}\left[V(X^{\pi}_{s}) - V(X^{\pi}_{s-})\right] - \lambda \\ &\quad \times \int_0^{\tau^{\pi}} \int_0^{\infty} e^{-\delta s}\left(V(X^{\pi}_{s-} - y) - V(X^{\pi}_{s-})\right) \mathrm{d}F(y)\,\mathrm{d}s \end{aligned} \tag{21}$$

is a martingale with $M(0) = 0$. So from the dynamic programming principle in (18), we have

$$\begin{aligned} &V(x) \\ &\geq E_x\left[\int_0^{\tau^{\pi}} e^{-\delta s} U_s \mathrm{d}s - \phi \int_0^{\tau^{\pi}} e^{-\delta s}\mathrm{d}Z_s + V(x) \right. \\ &\quad + \int_0^{\tau^{\pi}} e^{-\delta s}\left[(c-U_s)V'(X^{\pi}_{s-}) - \delta V(X^{\pi}_{s-}) \right. \\ &\quad\quad + \lambda \int_0^{\infty} V(X^{\pi}_{s-} - y) \\ &\quad\quad \left.\left. - V(X^{\pi}_{s-})\,\mathrm{d}F(y)\right]\mathrm{d}s + \phi \int_0^{\tau^{\pi}} e^{-\delta s}\mathrm{d}Z_s\right]. \end{aligned} \tag{22}$$

Equivalently

$$\begin{aligned} E_x\left[\int_0^{\tau^{\pi}} e^{-\delta s}\left[(c-U_s)V'(X^{\pi}_{s-}) + U_s \right.\right. \\ + \lambda \int_0^{\infty} V(X^{\pi}_{s-} - y)\,\mathrm{d}F(y) \\ \left.\left. - (\lambda+\delta)V(X^{\pi}_{s-})\right]\mathrm{d}s\right] \leq 0. \end{aligned} \tag{23}$$

Dividing $E\tau^{\pi}$ in (23) and letting $h \to 0$ yield

$$\begin{aligned} &(c-u)V'(x) + u \\ &\quad + \lambda \int_0^{\infty} V(x-y)\,\mathrm{d}F(y) - (\lambda+\delta)V(x) \leq 0. \end{aligned} \tag{24}$$

We have proved that $V(x)$ is increasing, continuous, and nonnegative, so the above inequality can be rewritten as

$$\begin{aligned} &(c-u)V'(x) + u \\ &\quad + \lambda \int_0^{x+z} V(x-y)\,\mathrm{d}F(y) - (\lambda+\delta)V(x) \leq 0 \end{aligned} \tag{25}$$

for $z \in R_+$.

On the other hand, consider a strategy by receiving $\varepsilon > 0$ from the shareholder immediately and following the optimal strategy for the capital $x+\varepsilon$ afterwards; then $V(x) \geq V(x+\varepsilon) - \phi\varepsilon$. Letting $\varepsilon \to 0$, we get

$$V'(x) \leq \phi. \tag{26}$$

A more sophisticated analysis shows that one of the inequalities (25) and (26) is always tight (see Fleming and Soner [11]).

As a result, we get the following HJB equation satisfied by the value function $V(x)$ on $[0,\infty)$:

$$\max\left\{\begin{aligned} &\sup_{\substack{0\le u\le u_0 \\ z\in R_+}}\left\{(c-u)V'(x) + u + \lambda \right. \\ &\quad\quad \left. \times \int_0^{x+z} V(x-y)\,\mathrm{d}F(y) - (\lambda+\delta)V(x)\right\}, \\ &V'(x) - \phi \end{aligned}\right\} = 0. \tag{27}$$

The expressions to be maximized are

$$u\left(1 - V'(x)\right), \qquad \int_0^{x+z} V(x-y)\,\mathrm{d}F(y). \tag{28}$$

First, because $u(1-V'(x))$ is linear in u, $u^*(x)$ maximizing $u(1-V'(x))$ is

$$u^*(x) = \begin{cases} 0, & \text{if } V'(x) > 1, \\ u_0, & \text{if } V'(x) \leq 1. \end{cases} \tag{29}$$

Second, we will maximize $\int_0^{x+z} V(x-y)\mathrm{d}F(y)$. Because $V(x) \geq 0$, we can define $z^* = -\inf\{z; V(z) > 0\}$. If $x < 0$, the shareholders either inject capital to survive the company or default to do so. Ruin occurs in the latter case, while in the former case $V(x)$ will be linear when $x < 0$; that is, $V(x) = V(0) + \phi x$. Thus, from the definition of z^*, we have

$$z^* = \frac{V(0)}{\phi}. \tag{30}$$

In fact, z^* is the maximal deficit that the shareholder should bare. We call $-z^*$ the optimal lower capital injection barrier.

If $V(x)$ is concave on $(0,\infty)$, then there exists an optimal dividend barrier $b := \inf\{x : V'(x) \leq 1\}$ with

$$u^*(x) = \begin{cases} 0, & \text{if } x < b \Longleftrightarrow V'(x) > 1, \\ u_0, & \text{if } x \geq b \Longleftrightarrow V'(x) \leq 1. \end{cases} \tag{31}$$

And also a barrier $a_0 := \sup\{x, V'(x) \geq \phi\}$. If the reserves become less than a_0, according to z^*, the shareholders may take actions between the following two choices.

(a) If the deficit is larger than z^*, they refuse to inject any capital and ruin occurs.

(b) Otherwise, they inject capital and the injected amount should recover the reserves to a_0. If $a_0 < 0$, the injected amount could not survive the company. Therefore, we define the optimal upper capital injection barrier as $a = a_0 \vee 0$.

Recall that in the literature (e.g., Kulenko and Schmidli [9] and He and Liang [4, 5]) concerning the capital injection strategy, considering the discounting, it can not be optimal to inject capital before they really are necessary. Therefore, the shareholders postpone injecting capital as long as possible and just conjecture that it is optimal to do so only when the reserves become 0. In the next proposition, from the mathematical point of view, we will give a rigorous proof of $a = 0$.

Proposition 3. *If $V(x)$ is concave on $(0, \infty)$, the optimal upper capital injection barrier $a = 0$.*

Proof. Under the assumption, a is unique. Suppose $a > 0$. So $V(x) = V(0) + \phi x$ when $x \in [-z^*, a]$. Note that $V'(a) = V'(0) = \phi$. $V(x)$ fulfils the HJB equation (27), so at $x = 0$

$$c\phi + \lambda \int_0^{V(0)/\phi} [V(0) - \phi y] \,\mathrm{d}F(y) - (\lambda + \delta) V(0) \le 0. \quad (32)$$

If we take $V(0) = V(a) - \phi a$ into the left side of (32), the expression turns into

$$\begin{aligned} &c\phi + \lambda \int_0^{V(0)/\phi} [V(a) - \phi a - \phi y] \,\mathrm{d}F(y) \\ &\quad - (\lambda + \delta)[V(a) - \phi a]. \end{aligned} \quad (33)$$

At the optimal upper capital injection barrier a,

$$cV'(a+) + \lambda \int_0^{V(a)/\phi} [V(a) - \phi y] \,\mathrm{d}F(y) - (\lambda + \delta) V(a) = 0. \quad (34)$$

It implies

$$(\lambda + \delta) V(a) = cV'(a+) + \lambda \int_0^{V(a)/\phi} [V(a) - \phi y] \,\mathrm{d}F(y). \quad (35)$$

Pulling (35) into (33), we can rewrite the expression by

$$\begin{aligned} &c\phi + \lambda \int_0^{V(0)/\phi} [V(a) - \phi a - \phi y] \,\mathrm{d}F(y) + (\lambda + \delta)\phi a \\ &\quad - cV'(a+) - \lambda \int_0^{V(a)/\phi} [V(a) - \phi y] \,\mathrm{d}F(y) \\ &= -\lambda \int_0^{V(0)/\phi} \phi a \,\mathrm{d}F(y) - \lambda \int_{V(0)/\phi}^{V(a)/\phi} [V(a) - \phi y] \,\mathrm{d}F(y) \\ &\quad + (\lambda + \delta)\phi a + c(\phi - V'(a+)) \\ &= -\lambda\phi a F\left(\frac{V(0)}{\phi}\right) + \lambda V(a) F\left(\frac{V(0)}{\phi}\right) \\ &\quad - \lambda V(a) F\left(\frac{V(a)}{\phi}\right) + \phi\lambda \int_e^{V(a)/\phi} y \,\mathrm{d}F(y) \\ &\quad + (\lambda + \delta)\phi a + c(\phi - V'(a+)) \\ &= -\lambda\phi a F\left(\frac{V(0)}{\phi}\right) + \lambda V(a) F\left(\frac{V(0)}{\phi}\right) \\ &\quad - \lambda V(a) F\left(\frac{V(a)}{\phi}\right) + c(\phi - V'(a+)) \\ &\quad - \lambda V(0) F\left(\frac{V(0)}{\phi}\right) + \lambda V(a) F\left(\frac{V(a)}{\phi}\right) \\ &\quad - \lambda\phi \int_{V(0)/\phi}^{V(a)/\phi} F(y) \,\mathrm{d}y + (\lambda + \delta)\phi a \\ &= [-\lambda\phi a + \lambda V(a) - \lambda\phi V(0)] F\left(\frac{V(a)}{\phi}\right) + (\lambda + \delta)\phi a \\ &\quad - \lambda\phi \int_{V(0)/\phi}^{V(a)/\phi} F(y) \,\mathrm{d}y + c(\phi - V'(a+)) \\ &= (\lambda + \delta)\phi a - \lambda\phi \int_{V(0)/\phi}^{V(a)/\phi} F(y) \,\mathrm{d}y + c(\phi - V'(a+)). \end{aligned} \quad (36)$$

However

$$\begin{aligned} &-\lambda\phi \int_{V(0)/\phi}^{V(a)/\phi} F(y) \,\mathrm{d}y + (\lambda + \delta)\phi a + c(\phi - V'(a+)) \\ &\quad \ge -\lambda V(a) + \lambda V(0) + (\lambda + \delta)\phi a \\ &\quad = -\lambda(V(a) - V(0) - \phi a) + \delta\phi a \\ &\quad = \delta\phi a > 0, \end{aligned} \quad (37)$$

which is contradictory with (32). So $a > 0$ is impossible and $a = 0$ is proved. □

The above proposition tells us that the moment when deficit occurs is just the time the shareholders consider to inject capital.

Proposition 4. *If $V(x)$ is concave on $(0, \infty)$, it is continuously differentiable on $(0, \infty)$.*

Proof. From the concavity of $V(x)$, (31) is true. When $x \in (0, b)$, from HJB equation (27), and $V(x)$ is Lipschitz continuous, so

$$\begin{aligned} &cV'(x+) - (\lambda + \delta) V(x) + \lambda \int_0^{V(0)/\phi} V(x - y) \,\mathrm{d}F(y) \\ &\quad = cV'(x-) - (\lambda + \delta) V(x) \\ &\quad + \lambda \int_0^{V(0)/\phi} V(x - y) \,\mathrm{d}F(y) = 0. \end{aligned} \quad (38)$$

Thus $V'(x-) = V'(x+)$. Similarly, we can proof $V(x)$ is continuously differentiable on (b, ∞). Now suppose $b > 0$. Note

$$\begin{gathered}(c-u_0)V'(b+)+u_0-(\lambda+\delta)V(b)\\ +\lambda\int_0^{V(0)/\phi}V(b-y)\,\mathrm{d}F(y)=0,\\ cV'(b-)-(\lambda+\delta)V(b)+\lambda\int_0^{V(0)/\phi}V(b-y)\,\mathrm{d}F(y)=0.\end{gathered}\tag{39}$$

So $cV'(b-) = u_0 + (c-u_0)V'(b+)$ or equivalently $c(V'(b-) - V'(b+)) = u_0(1 - V'(b+))$.

If $u_0 < c$, either $V'(b-) = V'(b+) = 1$ or $1 > V'(b-)$. The latter is impossible, so $V(x)$ is continuously differentiable under this case.

If $u_0 \geq c$, the reserve stays at b until the first claim occurs because dividend is a barrier strategy. b is independent of the constant u_0. In fact, because the process does not leave the interval $[0, b]$ and the corresponding strategy is admissible for any $u_0 \geq c$, it must be optimal for any initial value in $[0, b]$. For $x = b$, the expected discounted dividends until the first claim are

$$\lambda\int_0^\infty e^{-\lambda t}\int_0^t ce^{-\delta s}\mathrm{d}s\,\mathrm{d}t=\frac{\lambda c}{\delta}\int_0^\infty\left(1-e^{-\delta t}\right)e^{-\lambda t}\mathrm{d}t=\frac{c}{\lambda+\delta}.\tag{40}$$

The expected discounted dividends after the first claim are

$$\begin{aligned}&\lambda\int_0^\infty e^{-\lambda t}\int_0^b e^{-\delta t}V(b-y)\,\mathrm{d}F(y)\,\mathrm{d}t\\ &\quad+\lambda\int_0^\infty e^{-\lambda t}\int_b^{b+z^*}e^{-\delta t}\left[V(0)+\phi(b-y)\right]\mathrm{d}F(y)\,\mathrm{d}t\\ &=\frac{\lambda}{\lambda+\delta}\left[\int_0^b V(b-y)\,\mathrm{d}F(y)\right.\\ &\qquad\left.+\int_b^{b+z^*}\left[V(0)+\phi(b-y)\right]\mathrm{d}F(y)\right].\end{aligned}\tag{41}$$

Hence, the value at b can be characterized as

$$\begin{aligned}V(b)=\frac{c}{\lambda+\delta}+\frac{\lambda}{\lambda+\delta}&\left[\int_0^b V(b-y)\,\mathrm{d}F(y)\right.\\ &\left.+\int_b^{b+z^*}\left[V(0)+\phi(b-y)\right]\mathrm{d}F(y)\right].\end{aligned}\tag{42}$$

Pulling $V(b)$ into (39), we find $V'(b-) = V'(b+) = 1$. So $V(x)$ is continuously differentiable in this case, too. □

It holds in an interval (T_{i-1}, T_i) between two claims that $\mathrm{d}X_t^\pi = (c - U_t)\mathrm{d}t$. $\Delta Z_{T_i} = Z_{T_i+} - Z_{T_i}$ denotes the injected capital at the ith claim arrivals.

(i) If $X_{T_i-}^\pi - Y_i \geq 0$, then $\Delta Z_{T_i} = 0$;

(ii) If $-z^* < X_{T_i-}^\pi - Y_i < 0$, then the shareholders pay as much that $X_{T_i+}^\pi = X_{T_i-}^\pi - Y_i + \Delta Z_{T_i} = 0$. That is, $\Delta Z_{T_i} = 0 - (X_{T_i-}^\pi - Y_i)$. In this case, the value function fulfils

$$\begin{gathered}V\left(X_{T_i+}^\pi\right)(=V(0))=V\left(X_{T_i-}^\pi-Y_i\right)+\phi\Delta Z_{T_i},\\ \text{if } -z^*<X_{T_i-}^\pi-Y_i<0.\end{gathered}\tag{43}$$

(iii) If $X_{T_i-}^\pi - Y_i \leq -z^*$, then the shareholders would get a negative net profit as long as they cover the deficit (because $V(0) - \phi\Delta Z_{T_i} < 0$). It is unreasonable. Hence, they prefer to "no-injection-no-profit" and refuse to inject capital anymore. In this case, bankruptcy occurs and $T^\pi = T_i$. So

$$\begin{gathered}V\left(X_{T_i+}^\pi\right)=V\left(X_{T_i}^\pi\right)=V\left(X_{T_i-}^\pi-Y_i\right)=0,\\ \text{if } X_{T_i-}^\pi-Y_i\leq -z^*.\end{gathered}\tag{44}$$

Based on the discussion above, when $x < 0$, we can express $V(x)$ by

$$V(x)=\begin{cases}0, & \text{if } x\leq -z^*,\\ V(0)+\phi x, & \text{if } -z^*<x<0.\end{cases}\tag{45}$$

Thus it suffices to consider solutions f to the HJB equation with the properties

$$f(x)=0,\quad \text{if } x\leq -\frac{f(0)}{\phi}.\tag{46}$$

$$f(x)=f(0)+\phi x,\quad \text{if } -\frac{f(0)}{\phi}<x<0.\tag{47}$$

Lemma 5. *Let $f(x)$ be an increasing, bounded, and nonnegative solution to* (27) *with properties* (46) *and* (47). *Then for any admissible strategy $\pi \in \Pi^r$, the process*

$$\begin{aligned}&\left\{f\left(X_{t\wedge T^\pi}^\pi\right)e^{-\delta(t\wedge T^\pi)}-f(x)-\phi\int_0^{t\wedge T^\pi}e^{-\delta s}dZ_s\right.\\ &\quad-\int_0^{t\wedge T^\pi}\left[(c-U_s)f'\left(X_s^\pi\right)-(\lambda+\delta)f\left(X_s^\pi\right)\right.\\ &\qquad\left.\left.+\lambda\int_0^{X_s^\pi+(f(0)/\phi)}f\left(X_s^\pi-y\right)dF(y)\right]e^{-\delta s}ds\right\}\end{aligned}\tag{48}$$

is a martingale.

Proof. First we decompose $f(X_{t\wedge T^\pi}^\pi)e^{-\delta(t\wedge T^\pi)}$

$$\begin{aligned}&f\left(X_{t\wedge T^\pi}^\pi\right)e^{-\delta(t\wedge T^\pi)}\\ &\quad=f\left(X_{0+}^\pi\right)+\sum_{i=1}^{N_{t\wedge T^\pi}}\left[f\left(X_{T_i+}^\pi\right)e^{-\delta T_i}-f\left(X_{T_{i-1}+}^\pi\right)e^{-\delta T_{i-1}}\right]\end{aligned}$$

$$\begin{aligned}
&+ f\left(X^{\pi}_{t\wedge T^{\pi}}\right)e^{-\delta(t\wedge T^{\pi})} - f\left(X^{\pi}_{T_{N_{t\wedge T^{\pi}}+}}\right)e^{-\delta T_{N_{t\wedge T^{\pi}}}}\\
&= f\left(X^{\pi}_{0+}\right) + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\left[f\left(X^{\pi}_{T_i+}\right) - f\left(X^{\pi}_{T_i-}\right)\right]e^{-\delta T_i}\\
&\quad + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\left[f\left(X^{\pi}_{T_i-}\right)e^{-\delta T_i} - f\left(X^{\pi}_{T_{i-1}+}\right)e^{-\delta T_{i-1}}\right]\\
&\quad + f\left(X^{\pi}_{t\wedge T^{\pi}}\right)e^{-\delta(t\wedge T^{\pi})} - f\left(X^{\pi}_{T_{N_{t\wedge T^{\pi}}+}}\right)e^{-\delta T_{N_{t\wedge T^{\pi}}}}\\
&= f(x) + \phi\Delta Z_{T_0} + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\left[f\left(X^{\pi}_{T_i-} - Y_i\right) - f\left(X^{\pi}_{T_i-}\right)\right]e^{-\delta T_i}\\
&\quad + \phi\sum_{i=1}^{N_{t\wedge T^{\pi}}}\Delta Z_{T_i}e^{-\delta T_i} + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\int_{T_{i-1}+}^{T_i-} \mathrm{d}e^{-\delta s} f\left(X^{\pi}_s\right)\\
&\quad + \int_{T_{N_{t\wedge T^{\pi}}}+}^{t\wedge T^{\pi}} \mathrm{d}e^{-\delta s} f\left(X^{\pi}_s\right)\\
&= f(x) + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\left[f\left(X^{\pi}_{T_i-} - Y_i\right) - f\left(X^{\pi}_{T_i-}\right)\right]e^{-\delta T_i}\\
&\quad + \int_{T_{N_{t\wedge T^{\pi}}}+}^{t\wedge T^{\pi}}\left[\left(c - U_s\right)f'\left(X^{\pi}_s\right) - \delta f\left(X^{\pi}_s\right)\right]e^{-\delta s}\mathrm{d}s\\
&\quad + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\int_{T_{i-1}+}^{T_i-}\left[\left(c - U_s\right)f'\left(X^{\pi}_s\right) - \delta f\left(X^{\pi}_s\right)\right]e^{-\delta s}\mathrm{d}s\\
&\quad + \phi\int_0^{t\wedge T^{\pi}} e^{-\delta s}\mathrm{d}Z_s\\
&= f(x) + \sum_{i=1}^{N_{t\wedge T^{\pi}}}\left[f\left(X^{\pi}_{T_i-} - Y_i\right) - f\left(X^{\pi}_{T_i-}\right)\right]e^{-\delta T_i}\\
&\quad + \phi\int_0^{t\wedge T^{\pi}} e^{-\delta s}\mathrm{d}Z_s\\
&\quad + \int_0^{t\wedge T^{\pi}}\left[\left(c - U_s\right)f'\left(X^{\pi}_s\right) - \delta f\left(X^{\pi}_s\right)\right]e^{-\delta s}\mathrm{d}s.
\end{aligned}\tag{49}$$

Then in order to make the process $\{\sum_{i=1}^{N_{t\wedge T^{\pi}}}[f(X^{\pi}_{T_i-} - Y_i) - f(X^{\pi}_{T_i-})]e^{-\delta T_i} - \int_0^{t\wedge T^{\pi}} g(X^{\pi}_s)\mathrm{d}s\}$ become a martingale with the expected value 0, we must find a measurable function g. Since the above expression can be written as

$$\begin{aligned}
&\sum_{i=1}^{N_{t\wedge T^{\pi}}}\left\{\left[f\left(X^{\pi}_{T_i-} - Y_i\right) - f\left(X^{\pi}_{T_i-}\right)\right]e^{-\delta T_i} - \int_{T_{i-1}}^{T_i} g\left(X^{\pi}_s\right)\mathrm{d}s\right\}\\
&\quad - \int_{T_{N_{t\wedge T^{\pi}}}}^{t} g\left(X^{\pi}_s\right)\mathrm{d}s
\end{aligned}\tag{50}$$

it is enough to replace t by $T_1 \wedge t$; that is,

$$\left[f\left(X^{\pi}_{T_1-} - Y_1\right) - f\left(X^{\pi}_{T_1-}\right)\right]e^{-\delta T_1}1_{(T_1\le t)} - \int_0^{t\wedge T_1} g\left(X^{\pi}_s\right)\mathrm{d}s.\tag{51}$$

Because the exponential distribution is lack of memory, we only consider the expected value. g will satisfy

$$\begin{aligned}
&E\left\{\left[f\left(X^{\pi}_{T_1-} - Y_1\right) - f\left(X^{\pi}_{T_1-}\right)\right]e^{-\delta T_1}1_{(T_1\le t)}\right.\\
&\quad\left. - \int_0^{t\wedge T_1} g\left(X^{\pi}_s\right)\mathrm{d}s\right\} = 0.
\end{aligned}\tag{52}$$

The expected values of the first and the second part are

$$\begin{aligned}
&\int_0^t \lambda e^{-\lambda s}e^{-\delta s}\left\{\int_0^{x+\int_0^s(c-U_v)\mathrm{d}v+f(0)/\phi} f\right.\\
&\qquad\times\left(x + \int_0^s\left(c - U_v\right)\mathrm{d}v - y\right)\mathrm{d}F(y)\\
&\qquad\left. - f\left(x + \int_0^s\left(c - U_v\right)\mathrm{d}v\right)\right\}\mathrm{d}s,\\
&\int_0^t \lambda e^{-\lambda s}\int_0^s g\left(x + \int_0^v\left(c - U_w\right)\mathrm{d}w\right)\mathrm{d}v\,\mathrm{d}s\\
&\quad + e^{-\lambda t}\int_0^t g\left(x + \int_0^s\left(c - U_v\right)\mathrm{d}v\right)\mathrm{d}s\\
&= \int_0^t e^{-\lambda s}g\left(x + \int_0^s\left(c - U_v\right)\mathrm{d}v\right)\mathrm{d}s.
\end{aligned}\tag{53}$$

Thus we can choose

$$\begin{aligned}
g\left(X^{\pi}_t\right) &= \lambda e^{-\delta t}\left(\int_0^{X^{\pi}_t+f(0)/\phi} f\left(X^{\pi}_t - y\right)\mathrm{d}F(y) - f\left(X^{\pi}_t\right)\right)\\
&= \lambda e^{-\delta t}\int_0^{X^{\pi}_t+f(0)/\phi} f\left(X^{\pi}_t - y\right)\mathrm{d}F(y) - \lambda e^{-\delta t}f\left(X^{\pi}_t\right).
\end{aligned}\tag{54}$$

So

$$\begin{aligned}
&\left\{\sum_{i=1}^{N_{t\wedge T^{\pi}}}\left[f\left(X^{\pi}_{T_i-} - Y_i\right) - f\left(X^{\pi}_{T_i-}\right)\right]e^{-\delta T_i} - \lambda\right.\\
&\quad\left.\times\int_0^{t\wedge T^{\pi}} e^{-\delta s}\left[\int_0^{X^{\pi}_s+f(0)/\phi} f\left(X^{\pi}_s - y\right)\mathrm{d}F(y) - f\left(X^{\pi}_s\right)\right]\mathrm{d}s\right\}
\end{aligned}\tag{55}$$

and, also, the process

$$\begin{aligned}
&\left\{f\left(X^{\pi}_{t\wedge T^{\pi}}\right)e^{-\delta(t\wedge T^{\pi})} - f(x) - \phi\int_0^{t\wedge T^{\pi}} e^{-\delta s}\mathrm{d}Z_s\right.\\
&\quad - \int_0^{t\wedge T^{\pi}}\left[\left(c - U_s\right)f'\left(X^{\pi}_s\right) - (\lambda + \delta)f\left(X^{\pi}_s\right)\right.\\
&\qquad\left.\left. + \lambda\int_0^{X^{\pi}_s+f(0)/\phi} f\left(X^{\pi}_s - y\right)\mathrm{d}F(y)\right]e^{-\delta s}\mathrm{d}s\right\}
\end{aligned}\tag{56}$$

are $\{\mathscr{F}_t\}$-martingales with expected value 0. □

The following theorem serves as a verification theorem.

Theorem 6. *Let $f(x)$ be an increasing and bounded solution to* (27) *with the properties* (46) *and* (47). *Then* $\lim_{x\to\infty} f(x) = u_0/\delta$ *and* $f(x) = V(x)$ *on* R_+. *The optimal capital injection and dividend barriers are given by* (30) *and* (31).

Proof. Because $f(x)$ is increasing and bounded, we assume $\lim_{x\to\infty} f(x) = f_0$. Then there exists a sequence $x_n \to \infty$ such that $f'(x_n) \to 0$. Let $u_n = u(x_n)$. From the definition of the optimal dividend strategy, we can assume that $u_n = u_0$. As $n \to \infty$, the first term in (27) turns to

$$\begin{aligned} 0 &= (c-u_0) f'(x_n) + u_0 - \delta f(x_n) \\ &\quad + \lambda \left[\int_0^{x_n + f(0)/\phi} f(x_n - y)\, \mathrm{d}F(y) - f(x_n) \right] \\ &\longrightarrow -\delta f_0 + u_0. \end{aligned} \tag{57}$$

Equivalently we have $\lim_{x\to\infty} f(x) = u_0/\delta$.

Let T^* be the ruin time under the strategies (30) and (31) and $V^*(x)$ the corresponding value. From Lemma 5 and the HJB equation (27), we have

$$\begin{aligned} \Bigg\{ & f\left(X_{t\wedge T^*}^{\pi^*}\right) e^{-\delta(t\wedge T^*)} - f(x) \\ & + \int_0^{t\wedge T^*} e^{-\delta s} U_s^* \mathrm{d}s - \phi \int_0^{t\wedge T^*} e^{-\delta s} \mathrm{d}Z_s^* \Bigg\} \end{aligned} \tag{58}$$

is a martingale with expected value 0. Then

$$\begin{aligned} f(x) = E_x \Bigg[& f\left(X_{t\wedge T^*}^{\pi^*}\right) e^{-\delta(t\wedge T^*)} + \int_0^{t\wedge T^*} e^{-\delta s} U_s^* \mathrm{d}s \\ & -\phi \int_0^{t\wedge T^*} e^{-\delta s} \mathrm{d}Z_s^* \Bigg]. \end{aligned} \tag{59}$$

Since f is bounded and from the bounded convergence theorem, as $t \to \infty$, we get that $E[f(X_{t\wedge T^*}^{\pi^*}) e^{-\delta(t\wedge T^*)}] \to 0$. The other terms are monotone, when $t \to \infty$, by interchanging the limit and integration, so we obtain $f(x) = V^*(x) \le V(x)$.

On the other hand, because $f(x)$ is increasing and satisfies (46), $f(x)$ is nonnegative on $(-\infty, \infty)$. For any admissible strategy π, HJB equation (27) gives that

$$\begin{aligned} f(x) &\ge E_x \Bigg[f\left(X_{t\wedge T^\pi}^{\pi}\right) e^{-\delta(t\wedge T^\pi)} + \int_0^{t\wedge T^\pi} e^{-\delta s} U_s \mathrm{d}s \\ &\qquad -\phi \int_0^{t\wedge T^\pi} e^{-\delta s} \mathrm{d}Z_s \Bigg] \\ &\ge E_x \left[\int_0^{t\wedge T^\pi} e^{-\delta s} U_s \mathrm{d}s - \phi \int_0^{t\wedge T^\pi} e^{-\delta s} \mathrm{d}Z_s \right]. \end{aligned} \tag{60}$$

Let $t \to \infty$; then $f(x) \ge V^\pi(x)$, which means $f(x) \ge V(x)$. Thus, $f(x) = V(x)$. □

Based on the discussion above, if $V(x)$ is concave on $(0, \infty)$, it is optimal for the shareholders to take no action as long as the reserve process takes value in $(0, b)$. When the process reaches or exceeds the barrier b, dividends have to be paid at the maximal rate u_0. When the reserve is less than 0, the shareholders should consider either to inject capital to recover the reserve to 0 or default to do so. If the decifit is less than z^*, the shareholders can earn positive net profit. So they inject capital which covers the deficit to survive the company. Otherwise, once the deficit is larger than z^*, the shareholders refuse to do so and ruin occurs.

Remark 7. Diffusion models can be used to approximate the Cramér-Lundberg risk model. During the recent decades, they have been applied to insurance modeling setting extensively. See Radner and Shepp [12], Asmussen and Taksar [13], and Højgaard and Taksar [14, 15], Sethi and Taksar [2], and so forth. Diffusion models have the advantage that some very explicit optimal controls and a smooth value function can be made. Hopefully, these can help to take almost optimal strategies for the original risk model. However, this statement is not trivial.

The optimal dividend and issuance equity strategies (or combined with other strategies) in diffusion risk model had been studied by Løkka and Zervos [3], He and Liang [4, 5], and so forth. In their paper, depending on the relationships between the coefficients, it is optimal for the company either to involve no issuance equity or to involve issuance equity without ruin. In this paper, our conclusion in the Cramér-Lundberg risk model is that the optimal capital injection strategy will depend on the deficit. Once the deficit is large, ruin will still occur. Thus the optimal capital injection strategy looks different for these two models and the diffusion approximations are not effective here.

Discussion on whether the diffusion approximation is true can be found in Maglaras [16] and Bäuerle [6], and so forth.

3. Unrestricted Dividends

In this section, we will discuss the dividend strategy without restriction. Here all increasing, adapted, and càdlàg processes are allowed to be the dividend strategy. Let Π denote the set of all admissible strategies. The value of an admissible strategy π is

$$V^\pi(x) = E\left[\int_{0-}^{T^\pi -} e^{-\delta t} \mathrm{d}D_t - \phi \int_0^{T^\pi} e^{-\delta t} \mathrm{d}Z_t \right]. \tag{61}$$

The value function is $V(x) = \sup_{\pi\in\Pi} V^\pi(x)$.

Lemma 8. *On* $[0, \infty)$, *the function* $V(x)$ *is increasing and local Lipschitz continuous;* $V(x) - V(y) \ge x - y$ *if* $x \ge y$; $0 \le V(x) \le x + c/\delta$.

Proof. For any $\varepsilon > 0$, define a strategy π satisfing $V^\pi(y) \ge V(y) - \varepsilon$. π' is a new strategy for $x \ge y$. $\{Z_t'\}$ in π' is the same as $\{Z_t\}$ in π. While $\{D_t'\}$ is defined as: $x - y$ is paid immediately as dividend and then the strategy $\{D_t\}$ with initial capital y is followed. Therefore, $V(x) \ge x - y + V^\pi(y) \ge x - y + V(y) - \varepsilon$. From the arbitrary property of ε, we have $V(x) - V(y) \ge x - y$. In particulars, $V(x)$ is strictly increasing.

Consider such a strategy π: the initial capital x is paid to the shareholders as dividends immediately and capital injection is forbidden. Then $V(x) \geq V^{\pi}(x) \geq 0$.

To get the upper bound of $V(x)$, we consider a strategy π. $\{D_t\}$ is defined as: if the initial capital is x $(x \geq 0)$, then x is paid immediately and then the dividends are paid at rate c. If we donot take the capital injection into account, then $x + E_x[\int_0^{\infty} e^{-\delta t} c\, dt] = x + c/\delta$ is the upper bound of any admissible strategy π; that is, $V(x) \leq x + c/\delta$.

The local Lipschitz continuity follows by the local boundedness of $V(x)$ as in the proof of Lemma 2. □

3.1. HJB Equation and the Optimal Strategies. Similar to the discussion in Section 2.2, $V(x)$ satisfies the following dynamic programming principle:

$$V(x) = \sup_{\pi} E_x \left[\int_{0-}^{\tau\wedge T^{\pi}} e^{-\delta t} \mathrm{d}D_t - \phi \int_0^{\tau\wedge T^{\pi}} e^{-\delta t} \mathrm{d}Z_t + e^{-\delta(\tau\wedge T^{\pi})} V\left(X^{\pi}_{\tau\wedge T^{\pi}}\right) \right] \tag{62}$$

for $x \in R_+$ and any $\{\mathcal{F}_t\}$-stopping time τ.

For $x \geq 0$, similarly we define τ^{π} as in Section 2.2. Note that $\sigma^{\pi} = T^{\pi}$ is possible here. Applying Itô formula into $e^{-\delta\tau^{\pi}} V(X^{\pi}_{\tau^{\pi}})$, we have

$$\begin{aligned} e^{-\delta\tau^{\pi}} V\left(X^{\pi}_{\tau^{\pi}}\right) = V\left(X^{\pi}_{0-}\right) &+ \int_0^{\tau^{\pi}} e^{-\delta s} cV'\left(X^{\pi}_{s-}\right) - \delta e^{-\delta s} V\left(X^{\pi}_{s-}\right) \mathrm{d}s \\ &+ \sum_{\substack{0\leq s\leq\tau^{\pi} \\ X^{\pi}_{s-}\neq X^{\pi}_s}} e^{-\delta s} \left[V\left(X^{\pi}_s\right) - V\left(X^{\pi}_{s-}\right)\right] \\ &+ \sum_{\substack{0\leq s<\tau^{\pi} \\ X^{\pi}_s\neq X^{\pi}_{s+}}} e^{-\delta s} \left[V\left(X^{\pi}_{s+}\right) - V\left(X^{\pi}_s\right)\right]. \end{aligned} \tag{63}$$

$X^{\pi}_s \neq X^{\pi}_{s+}$ only when capital is injected, so

$$\sum_{\substack{0\leq s<\tau^{\pi} \\ X^{\pi}_s\neq X^{\pi}_{s+}}} e^{-\delta s} \left[V\left(X^{\pi}_{s+}\right) - V\left(X^{\pi}_s\right)\right] = \phi \int_0^{\tau^{\pi}} e^{-\delta s} \mathrm{d}Z_s. \tag{64}$$

When claim arrives or dividend occurs, $X^{\pi}_{s-} \neq X^{\pi}_s$. The jumps caused by claim arrivals lead to

$$\begin{aligned} M\left(\tau^{\pi}\right) &= M\left(\sigma^{\pi} \wedge h\right) \\ &= \sum_{\substack{0\leq s\leq\tau^{\pi} \\ X^{\pi}_{s-}\neq X^{\pi}_s}} e^{-\delta s} \left[V\left(X^{\pi}_s\right) - V\left(X^{\pi}_{s-}\right)\right] \\ &\quad - \lambda \int_0^{\tau^{\pi}} \int_0^{\infty} e^{-\delta s} \left(V\left(X^{\pi}_{s-} - y\right) - V\left(X^{\pi}_{s-}\right)\right) \mathrm{d}F(y)\, \mathrm{d}s \end{aligned} \tag{65}$$

is a martingale with $M(0) = 0$. And the amount of the aggregated jumps caused by dividend are $-\int_{0-}^{\tau^{\pi}} e^{-\delta s} \mathrm{d}D_s$. So from the dynamic programming principle (62), yields

$$\begin{aligned} V(x) &\geq E_x \left[\int_{0-}^{\tau^{\pi}} e^{-\delta s} \mathrm{d}D_s - \phi \int_0^{\tau^{\pi}} e^{-\delta s} \mathrm{d}Z_s + V(x) \right. \\ &\quad + \int_0^{\tau^{\pi}} e^{-\delta s} \left[cV'\left(X^{\pi}_{s-}\right) - \delta V\left(X^{\pi}_{s-}\right) + \lambda \right. \\ &\quad \left. \times \int_0^{\infty} V\left(X^{\pi}_{s-} - y\right) - V\left(X^{\pi}_{s-}\right) \mathrm{d}F(y) \right] \mathrm{d}s \\ &\quad \left. - \int_{0-}^{\tau^{\pi}} e^{-\delta s} \mathrm{d}D_s + \phi \int_0^{\tau^{\pi}} e^{-\delta s} \mathrm{d}Z_s \right]. \end{aligned} \tag{66}$$

Equivalently

$$E_x \left[\int_0^{\tau^{\pi}} e^{-\delta s} \left[cV'\left(X^{\pi}_{s-}\right) + \lambda \int_0^{\infty} V\left(X^{\pi}_{s-} - y\right) \mathrm{d}F(y) - (\lambda+\delta) V\left(X^{\pi}_{s-}\right) \right] \mathrm{d}s \right] \leq 0. \tag{67}$$

If $T^{\pi} = 0$, then $\tau^{\pi} = 0$. Therefore (67) gives no information. If $T^{\pi} > 0$, we can choose ε such that $E\tau^{\pi} > 0$. Dividing $E\tau^{\pi}$ in (67) and letting $h \to 0$, so

$$cV'(x) + \lambda \int_0^{\infty} V(x-y)\, \mathrm{d}F(y) - (\lambda+\delta) V(x) \leq 0. \tag{68}$$

Also we can rewrite the above inequality by

$$cV'(x) + \lambda \int_0^{x+z} V(x-y)\, \mathrm{d}F(y) - (\lambda+\delta) V(x) \leq 0 \tag{69}$$

for $z \in R_+$.

Refering to the proof of (26), we have

$$V'(x) \leq \phi. \tag{70}$$

If the company pays out ε as dividends, then the initial capital reduces from x to $x-\varepsilon$. Using the optimal strategy afterwards, so $V(x) \geq V(x-\varepsilon)+\varepsilon$. Subtracting $V(x-\varepsilon)$ from both sides, dividing by ε, and letting $\varepsilon \to 0$, we get

$$V'(x) \geq 1. \tag{71}$$

One of the inequalities (69), (70), and (71) is always tight (refer to Fleming and Soner [11]).

Thus we derive the HJB equation satisfied by $V(x)$ on $[0,\infty)$

$$\max\left\{ \sup_{z\in R_+} \left\{ cV'(x) + \lambda \int_0^{x+z} V(x-y)\, \mathrm{d}F(y) - (\lambda+\delta) V(x) \right\}, 1 - V'(x), V'(x) - \phi \right\} = 0. \tag{72}$$

To maximize $\int_0^{x+z} V(x-y)\mathrm{d}F(y)$, let us recall the proof of $z^* = V(0)/\phi$ in Section 2.2. We can find that z^* is independent of u_0. So we also have the optimal lower capital injection barrier

$$-z^* = -\frac{V(0)}{\phi}. \tag{73}$$

Hence when $x < 0$, $V(x)$ can be expressed by

$$V(x) = \begin{cases} 0 & \text{if } x \le -z^*, \\ V(0) + \phi x & \text{if } -z^* < x < 0. \end{cases} \tag{74}$$

In Section 2.2, the optimal dividend strategy and the optimal capital injection strategy are both barrier strategies under the assumption that $V(x)$ is concave on $(0,\infty)$. Moreover, the optimal dividend barrier b and the upper optimal capital injection barrier a are both independent of u_0. Here if $V(x)$ is concave on $(0,\infty)$, similar to discussion in Section 2.2, we can define the optimal dividend barrier $b := \inf\{x : V'(x) \le 1\}$ and the optimal upper capital injection barrier $a := \sup\{x : V'(x) \ge \phi\} \vee 0$. And also $V(x)$ is continuously differentiable.

Proposition 9. *If $V(x)$ is concave on $(0,\infty)$, the optimal upper capital injection barrier $a = 0$.*

Proof. The proof is similar as in Proposition 3, so we omit it here. □

Now define a strategy $\pi^1 = (D^1, Z^1)$ as follows:

$$\begin{aligned} D_0^1 &= \max(x-b, 0), \\ D_t^1 &= D_0^1 + \int_0^t c 1_{\{X_s^{\pi^1} = b\}} \mathrm{d}s, \quad \text{for } t > 0, \\ Z_t^1 &= \max\left\{-\inf_{0\le s<t}\left(X_s - D_s^1\right), 0\right\} \quad \text{for } t > 0. \end{aligned} \tag{75}$$

Let $T^* = \inf\{t \ge 0 : X_t^{\pi^1} \le -z^*\}$. Define strategy $\pi^* = (D^*, Z^*)$ by the strategy π^1 stopped at T^*:

$$D_t^* = \begin{cases} D_t^1, & \text{if } t < T^*, \\ D_{T^*-}^1, & \text{if } t \ge T^*, \end{cases} \qquad Z_t^* = \begin{cases} Z_t^1, & \text{if } t < T^*, \\ Z_{T^*}^1, & \text{if } t \ge T^*. \end{cases} \tag{76}$$

Under π^*, if the initial capital $x > b$, $x - b$ will be paid to the shareholders as dividends immediately. When the reserve process takes value in $(0, b)$, insurance company dose not pay dividend and shareholders do not inject capital. When the process reaches the barrier b, the premium income will be paid as dividends. If deficit occurs and it is less than z^*, the shareholders inject capital to recover the reserve process to 0. Otherwise, they refuse to inject any capital and ruin occurs. $X_t^* = X_t - D_t^* + Z_t^*$ is the corresponding reserve process.

Theorem 10. *If $V(x)$ is concave on $(0,\infty)$, the strategy π^* defined in (76) is optimal; that is,*

$$V^{\pi^*}(x) = V(x). \tag{77}$$

Proof. Note that $V'(X_t^*) = V'(b) = 1$ on $\{X_t^* = b\}$. According to (76), the possible increment of $\{Z_t^*\}$ is at the time of claim arrivals. As in Lemma 5,

$$\begin{aligned} &V\left(X_{t\wedge T^*}^*\right) e^{-\delta(t\wedge T^*)} \\ &\quad = V(x) - D_0^1 + \phi \int_0^{t\wedge T^*} e^{-\delta s} \mathrm{d}Z_s^* \\ &\qquad + \sum_{i=1}^{N_{t\wedge T^*}} \left[V\left(X_{T_i-}^* - Y_i\right) - V\left(X_{T_i-}^*\right)\right] e^{-\delta T_i} \\ &\qquad + \sum_{i=1}^{N_{t\wedge T^*}} \left[V\left(X_{T_i-}^*\right) e^{-\delta T_i} - V\left(X_{T_{i-1}+}^*\right) e^{-\delta T_{i-1}}\right] \\ &\qquad + V\left(X_{t\wedge T^*}^*\right) e^{-\delta(t\wedge T^*)} - V\left(X_{T_{N_{t\wedge T^*}}+}^*\right) e^{-\delta T_{N_{t\wedge T^*}}} \\ &\quad = V(x) - D_0^1 + \phi \int_0^{t\wedge T^*} e^{-\delta s} \mathrm{d}Z_s^* \\ &\qquad + \sum_{i=1}^{N_{t\wedge T^*}} \left[V\left(X_{T_i-}^* - Y_i\right) - V\left(X_{T_i-}^*\right)\right] e^{-\delta T_i} \\ &\qquad + \sum_{i=1}^{N_{t\wedge T^*}} \int_{T_{i-1}+}^{T_i-} \left[cV'\left(X_s^*\right) - \delta V\left(X_s^*\right)\right] 1_{\{0<X_s^*<b\}} e^{-\delta s} \mathrm{d}s \\ &\qquad - \sum_{i=1}^{N_{t\wedge T^*}} \int_{T_{i-1}+}^{T_i-} \delta V\left(X_s^*\right) 1_{\{X_s^*=b\}} e^{-\delta s} \mathrm{d}s \\ &\qquad + \int_{T_{N_{t\wedge T^*}}+}^{t\wedge T^*} \left[cV'\left(X_s^*\right) - \delta V\left(X_s^*\right)\right] 1_{\{0<X_s^*<b\}} e^{-\delta s} \mathrm{d}s \\ &\qquad - \int_{T_{N_{t\wedge T^*}}+}^{t\wedge T^*} \delta V\left(X_s^*\right) 1_{\{X_s^*=b\}} e^{-\delta s} \mathrm{d}s. \end{aligned} \tag{78}$$

The process

$$\begin{aligned} &\left\{\sum_{i=1}^{N_{t\wedge T^*}} \left[V\left(X_{T_i-}^* - Y_i\right) - V\left(X_{T_i-}^*\right)\right] e^{-\delta T_i} - \lambda \right. \\ &\quad \left. \times \int_0^{t\wedge T^*} e^{-\delta s} \left[\int_0^{X_s^*+z^*} V\left(X_{s-}^* - y\right) \mathrm{d}F(y) - V\left(X_{s-}^*\right)\right] \mathrm{d}s\right\} \end{aligned} \tag{79}$$

is a martingale with expected value 0. Equivalently,

$$\begin{aligned} &\left\{V\left(X_{t\wedge T^*}^*\right) e^{-\delta(t\wedge T^*)} - V(x) + D_0^1 - \phi \int_0^{t\wedge T^*} e^{-\delta s} \mathrm{d}Z_s^* \right. \\ &\quad - \int_0^{t\wedge T^*} \left[cV'\left(X_s^*\right) + \lambda \int_0^{X_s^*+z^*} V\left(X_s^* - y\right) \mathrm{d}F(y) \right. \\ &\quad \left. - (\lambda+\delta) V\left(X_s^*\right)\right] 1_{\{0<X_s^*<b\}} e^{-\delta s} \mathrm{d}s \end{aligned}$$

$$-\int_0^{t\wedge T^*}\left[\lambda\int_0^{X_s^*+z^*}V\left(X_s^*-y\right)\mathrm{d}F\left(y\right)\right.$$
$$\left.-\left(\lambda+\delta\right)V\left(X_s^*\right)\right]1_{\{X_s^*=b\}}e^{-\delta s}\mathrm{d}s\Bigg\} \tag{80}$$

is a martingale. Because $V(x)$ is concave on $(0,\infty)$, the derivatives of $V(x)$ from left and right exist. Moreover, $F(y)$ is continuous, so $V(x)$ in (72) is continuously differentiable. For $V'(X_s^*) > 1$ on $\{0 < X_s^* < b\}$, the first term on the left-hand side of (72) is 0, thus the integral over $\{0 < X_s^* < b\}$ on the expression above is 0. Furthermore, from $V'(X_s^*) = 1$ on $\{X_s^* = b\}$ and (72), it follows that

$$\lambda\int_0^{X_s^*+z^*}V\left(X_s^*-y\right)\mathrm{d}F\left(y\right)-\left(\lambda+\delta\right)V\left(X_s^*\right)=-c. \tag{81}$$

Taking this expression into (80), we have

$$\Bigg\{V\left(X_{t\wedge T^*}^*\right)e^{-\delta(t\wedge T^*)}-V\left(x\right)+D_0^1$$
$$-\phi\int_0^{t\wedge T^*}e^{-\delta s}\mathrm{d}Z_s^*+\int_0^{t\wedge T^*}c1_{\{X_s^*=b\}}e^{-\delta s}\mathrm{d}s\Bigg\} \tag{82}$$

is a martingale with expected value 0. Then

$$V\left(x\right)=E_x\left[V\left(X_{t\wedge T^*}^*\right)e^{-\delta(t\wedge T^*)}-\phi\int_0^{t\wedge T^*}e^{-\delta s}\mathrm{d}Z_s^*\right.$$
$$\left.+\int_0^{t\wedge T^*}c1_{\{X_s^*=b\}}e^{-\delta s}\mathrm{d}s+D_0^1\right]. \tag{83}$$

Note that

$$E_x\left[V\left(X_{t\wedge T^*}^*\right)e^{-\delta(t\wedge T^*)}\right]=e^{-\delta t}E_x\left[V\left(X_t^*\right)1_{(t\le T^*)}\right]$$
$$\le e^{-\delta t}V\left(b\right). \tag{84}$$

By the bounded convergence theorem,

$$\lim_{t\to\infty}E_x\left[V\left(X_{t\wedge T^*}^*\right)e^{-\delta(t\wedge T^*)}\right]=0. \tag{85}$$

So

$$V\left(x\right)$$
$$=\lim_{t\to\infty}E_x\left[\int_0^{t\wedge T^*}c1_{\{X_s^*=b\}}e^{-\delta s}\mathrm{d}s-\phi\int_0^{t\wedge T^*}e^{-\delta s}\mathrm{d}Z_s^*+D_0^1\right]$$
$$=E_x\left[\int_{0-}^{T^*-}e^{-\delta s}\mathrm{d}D_s^*-\phi\int_0^{T^*}e^{-\delta s}\mathrm{d}Z_s^*\right]=V^*\left(x\right). \tag{86}$$

□

3.2. Characterization of the Solution. How to characterize the solution $V(x)$ among other possible solutions?

Theorem 11. *$V(x)$ is the minimal nonnegative solution to* (72).

Proof. Let f be a nonnegative solution to the HJB equation (72). f is increasing because $f'(x) \ge 1$. $\{X_t^*\}$ is the reserve process under π^*. From Theorem 10

$$\Bigg\{f\left(X_{t\wedge T^*}^*\right)e^{-\delta(t\wedge T^*)}-f\left(x\right)+D_0^1-\phi\int_0^{t\wedge T^*}e^{-\delta s}\mathrm{d}Z_s^*$$
$$-\int_0^{t\wedge T^*}\left[cf'\left(X_s^*\right)+\lambda\int_0^{X_s^*+f(0)/\phi}f\left(X_s^*-y\right)\mathrm{d}F\left(y\right)\right.$$
$$\left.-\left(\lambda+\delta\right)f\left(X_s^*\right)\right]1_{\{0<X_s^*<b\}}e^{-\delta s}\mathrm{d}s$$
$$-\int_0^{t\wedge T^*}\left[\lambda\int_0^{X_s^*+f(0)/\phi}f\left(X_s^*-y\right)\mathrm{d}F\left(y\right)\right.$$
$$\left.-\left(\lambda+\delta\right)f\left(X_s^*\right)\right]1_{\{X_s^*=b\}}e^{-\delta s}\mathrm{d}s\Bigg\} \tag{87}$$

is a martingale with expected value 0. $f(x)$ satisfies (72); then

$$cf'\left(X_s^*\right)+\lambda\int_0^{X_s^*+f(0)/\phi}f\left(X_s^*-y\right)\mathrm{d}F\left(y\right)$$
$$-\left(\lambda+\delta\right)f\left(X_s^*\right)\le 0. \tag{88}$$

Because $f'(x) \ge 1$,

$$\lambda\int_0^{X_s^*+f(0)/\phi}f\left(X_s^*-y\right)\mathrm{d}F\left(y\right)-\left(\lambda+\delta\right)f\left(X_s^*\right)$$
$$\le -cf'\left(X_s^*\right)\le -c. \tag{89}$$

From the non-negative property of $f(x)$, we have

$$f\left(x\right)\ge E_x\left[f\left(X_{t\wedge T^*}^*\right)e^{-\delta(t\wedge T^*)}-\phi\int_0^{t\wedge T^*}e^{-\delta s}\mathrm{d}Z_s^*\right.$$
$$\left.+\int_0^{t\wedge T^*}c1_{\{X_s^*=b\}}e^{-\delta s}\mathrm{d}s+D_0^1\right]$$
$$\ge E_x\left[\int_{0-}^{t\wedge T^*-}e^{-\delta s}\mathrm{d}D_s^*-\phi\int_0^{t\wedge T^*}e^{-\delta s}\mathrm{d}Z_s^*\right]$$
$$=V^*\left(x\right)=V\left(x\right). \tag{90}$$

□

4. Optimal Dividend and Capital Injection Strategies for Exponential Claims

In this section we will consider the case that the claim size is exponentially distributed and the dividend strategy without restriction. Let $F(x) = 1 - e^{-\alpha x}$.

First, we assume that $f(x)$ is an increasing, continuously differentiable and concave solution to the HJB equation (72) on $[0,\infty)$. Define $b=\inf\{x: f'(x)=1\}\vee 0$. On $[0,b]$, $f(x)$ satisfies

$$\begin{aligned} &cf'(x)+\lambda\int_0^x f(x-y)\,\alpha e^{-\alpha y}\mathrm{d}y \\ &\quad+\lambda\int_x^{x+f(0)/\phi}\left[f(0)+\phi(x-y)\right]\alpha e^{-\alpha y}\mathrm{d}y \\ &\quad-(\lambda+\delta)f(x)=0. \end{aligned}\tag{91}$$

Let $z=x-y$ and change (91) into

$$\begin{aligned} &cf'(x)+\lambda e^{-\alpha x}\int_0^x f(z)\,\alpha e^{\alpha z}\mathrm{d}z \\ &\quad+\lambda e^{-\alpha x}\int_{-f(0)/\phi}^0\left[f(0)+\phi z\right]\alpha e^{\alpha z}\mathrm{d}z \\ &\quad-(\lambda+\delta)f(x)=0. \end{aligned}\tag{92}$$

The above expression can be derivative, so it yields

$$\begin{aligned} &cf''(x)-\alpha\lambda e^{-\alpha x}\int_0^x f(z)\,\alpha e^{\alpha z}\mathrm{d}z+\alpha\lambda f(x) \\ &\quad-\alpha\lambda e^{-\alpha x}\int_{-f(0)/\phi}^0\left[f(0)+\phi z\right]\alpha e^{\alpha z}\mathrm{d}z \\ &\quad-(\lambda+\delta)f'(x)=0. \end{aligned}\tag{93}$$

Combining (92) with (93), we get the differentiable equation about $f(x)$

$$cf''(x)+(\alpha c-(\lambda+\delta))f'(x)-\alpha\delta f(x)=0.\tag{94}$$

Its solution is

$$f(x)=C_1e^{v_1x}+C_2e^{v_2x},\tag{95}$$

where v_1, v_2 are the solutions of equation $cv^2+(\alpha c-(\lambda+\delta))v-\alpha\delta=0$; that is,

$$\begin{aligned} v_1&=\frac{\lambda+\delta-\alpha c-\sqrt{(\lambda+\delta-\alpha c)^2+4\alpha c\delta}}{2c}<0, \\ v_2&=\frac{\lambda+\delta-\alpha c+\sqrt{(\lambda+\delta-\alpha c)^2+4\alpha c\delta}}{2c}>0. \end{aligned}\tag{96}$$

When $x\ge b$, we conjecture that

$$f(x)=x-b+f(b),\quad x\ge b.\tag{97}$$

Therefore, from (95) and (97), the suggested solution of HJB equation (72) has the form

$$f(x)=\begin{cases}C_1e^{v_1x}+C_2e^{v_2x} & \text{if } 0\le x\le b,\\ x-b+C_1e^{v_1b}+C_2e^{v_2b} & \text{if } x\ge b,\end{cases}\tag{98}$$

where C_1, C_2, and b are to be determined later.

Lemma 12. *At $x=b$, we have*

$$f''(b)=0,\qquad f(b)=\frac{\alpha c-\lambda-\delta}{\alpha\delta}.\tag{99}$$

Proof. As we have assumed that $f(x)$ satisfies HJB equation (72), when $x>b$, we have

$$c+\lambda\int_0^{x+f(0)/\phi}f(x-y)\,\alpha e^{-\alpha y}\mathrm{d}y-(\lambda+\delta)f(x)\le 0.\tag{100}$$

Note that

$$\begin{aligned} &\int_0^{x+f(0)/\phi}f(x-y)\,\alpha e^{-\alpha y}\mathrm{d}y \\ &\quad=\int_0^{x-b}\left[f(b)+(x-y-b)\right]\alpha e^{-\alpha y}\mathrm{d}y \\ &\qquad+\int_{x-b}^{x+f(0)/\phi}f(x-y)\,\alpha e^{-\alpha y}\mathrm{d}y \\ &\quad=\left[f(b)+(x-b)\right]\left(1-e^{-\alpha(x-b)}\right)-\int_0^{x-b}y\alpha e^{-\alpha y}\mathrm{d}y \\ &\qquad+\int_0^{b+f(0)/\phi}f(b-u)\,\alpha e^{-\alpha(x-b+u)}\mathrm{d}u \\ &\quad=\left[f(b)+(x-b)\right]\left(1-e^{-\alpha(x-b)}\right)+(x-b)\,e^{-\alpha(x-b)} \\ &\qquad-\frac{1}{\alpha}+\frac{1}{\alpha}e^{-\alpha(x-b)} \\ &\qquad+\left(\int_0^{b+f(0)/\phi}f(b-u)\,\alpha e^{-\alpha u}\mathrm{d}u\right)e^{-\alpha(x-b)} \end{aligned}\tag{101}$$

and from the HJB equation (72), when $x=b$,

$$c+\lambda\int_0^{b+f(0)/\phi}f(b-y)\,\alpha e^{-\alpha y}\mathrm{d}y-(\lambda+\delta)V(b)=0.\tag{102}$$

Plugging (101) and (102) into the left side of (100), then we can rewrite the expression by

$$\begin{aligned} &c+\lambda\Big\{\left[f(b)+(x-b)\right]\left(1-e^{-\alpha(x-b)}\right) \\ &\qquad+(x-b)\,e^{-\alpha(x-b)}-\frac{1}{\alpha}+\frac{1}{\alpha}e^{-\alpha(x-b)}\Big\} \\ &\qquad+\left[(\lambda+\delta)f(b)-c\right]e^{-\alpha(x-b)}-(\lambda+\delta)\left(f(b)+(x-b)\right) \\ &\quad=\left(c-\delta f(b)-\frac{\lambda}{\alpha}\right)\left(1-e^{-\alpha(x-b)}\right)-\delta(x-b). \end{aligned}\tag{103}$$

Therefore (100) is established if and only if $f(b)\ge[\alpha c-\lambda-\alpha\delta(x-b)/(1-e^{-\alpha(x-b)})]/(\alpha\delta)$ for all $x>b$. When $x\to b$, $f(b)\ge(\alpha c-\lambda-\delta)/(\alpha\delta)$.

In (94), let $x = b$. From $f(b) \geq (\alpha c - \lambda - \delta)/(\alpha\delta)$, we find that

$$\begin{aligned} 0 &= cf''(b) + (\alpha c - (\lambda + \delta)) - \alpha\delta f(b) \\ &\leq cf''(b) + (\alpha c - (\lambda + \delta)) - \alpha\delta \frac{\alpha c - \lambda - \delta}{\alpha\delta} = cf''(b), \end{aligned} \tag{104}$$

which implies $f''(b) \geq 0$.

On the other hand, because $f(x)$ is concave, we have $f''(x) \leq 0$. Particularly, $f''(b) \leq 0$. Combining the discussion above, $f''(b) = 0$.

Furthermore, taking $f''(b) = 0$ and $f'(b) = 1$ into (94) yields

$$f(b) = \frac{\alpha c - \lambda - \delta}{\alpha\delta}. \tag{105}$$

□

Next we will determine C_1, C_2, and b.

From the expression of $f(x)$ in (95) and $f''(b) = 0$ (it has been proved in Lemma 12), it holds that

$$f''(b) = C_1 v_1^2 e^{v_1 b} + C_2 v_2^2 e^{v_2 b} = 0. \tag{106}$$

The continuously differentiable property of $f(x)$ tells us that

$$f'(b) = C_1 v_1 e^{v_1 b} + C_2 v_2 e^{v_2 b} = 1. \tag{107}$$

Combining the two equations above, we can get the expression of C_1 and C_2:

$$C_1 = \frac{v_2}{(v_2 - v_1) v_1 e^{v_1 b}}, \qquad C_2 = \frac{v_1}{(v_1 - v_2) v_2 e^{v_2 b}}. \tag{108}$$

When $x = 0$, (95) informs us that $f(0) = C_1 + C_2$, $f'(0) = C_1 v_1 + C_2 v_2$. Meanwhile, at $x = 0$ the integral-differential equation (91) implies

$$cf'(0) + \lambda \int_0^{f(0)/\phi} [f(0) - \phi y] \, dF(y) - (\lambda + \delta) f(0) = 0. \tag{109}$$

Together with (108), (109) can be rewritten as

$$\begin{aligned} &\frac{cv_2}{v_2 - v_1} e^{-v_1 b} + \frac{cv_1}{v_1 - v_2} e^{-v_2 b} \\ &\quad - \frac{\lambda\phi}{\alpha} \left(1 - e^{-(\alpha/\phi)[(v_2/(v_2 - v_1)v_1)e^{-v_1 b} + (v_1/(v_1 - v_2)v_2)e^{-v_2 b}]}\right) \\ &\quad - \delta \left[\frac{v_2}{(v_2 - v_1) v_1} e^{-v_1 b} + \frac{v_1}{(v_1 - v_2) v_2} e^{-v_2 b}\right] = 0, \end{aligned} \tag{110}$$

which can be used to calculate b.

Proposition 13. *The solution of* (110) *is unique.* $b = 0$ *if and only if*

$$\lambda + \delta \geq \lambda\phi \left(1 - e^{-(\alpha c - (\lambda + \delta))/\phi\delta}\right). \tag{111}$$

Proof. To analyse the solution of (110), we first define a function

$$\begin{aligned} g(z) := &\frac{cv_2}{v_2 - v_1} e^{-v_1 z} + \frac{cv_1}{v_1 - v_2} e^{-v_2 z} \\ &- \frac{\lambda\phi}{\alpha} \left(1 - e^{-(\alpha/\phi)[(v_2/(v_2 - v_1)v_1)e^{-v_1 z} + (v_1/(v_1 - v_2)v_2)e^{-v_2 z}]}\right) \\ &- \delta \left[\frac{v_2}{(v_2 - v_1) v_1} e^{-v_1 z} + \frac{v_1}{(v_1 - v_2) v_2} e^{-v_2 z}\right], \end{aligned} \tag{112}$$

where $z \geq 0$. In view of $v_1 + v_2 = [(\lambda + \delta) - \alpha c]/c$ and $v_1 v_2 = -\alpha\delta/c$, we find that

$$\begin{aligned} g'(z) = &c \frac{v_1 v_2}{v_1 - v_2} \left(e^{-v_1 z} - e^{-v_2 z}\right) \\ &+ \delta \left(\frac{v_2}{v_2 - v_1} e^{-v_1 z} + \frac{v_1}{v_1 - v_2} e^{-v_2 z}\right) \\ &- \lambda \left(\frac{v_2}{v_1 - v_2} e^{-v_1 z} + \frac{v_1}{v_2 - v_1} e^{-v_2 z}\right) \\ &\cdot e^{-(\alpha/\phi)[(v_2/(v_2 - v_1)v_1)e^{-v_1 z} + (v_1/(v_1 - v_2)v_2)e^{-v_2 z}]} > 0, \end{aligned} \tag{113}$$

which implies $g(z)$ is increasing strictly. So the solution is unique. Consider

$$\begin{aligned} \lim_{z \to \infty} g(z) = \lim_{z \to \infty} \Bigg[&\frac{cv_2}{v_2 - v_1} e^{-v_1 z} \\ &- \frac{\lambda\phi}{\alpha} \left(1 - e^{-(\alpha/\phi)(v_2/(v_2 - v_1)v_1)e^{-v_1 z}}\right) \\ &- \delta \frac{v_2}{(v_2 - v_1) v_1} e^{-v_1 z} \Bigg] = \infty. \end{aligned} \tag{114}$$

Hence $b = 0$ if and only if $g(0) \geq 0$. While

$$\begin{aligned} g(0) &= c - \frac{\lambda\phi}{\alpha} \left(1 - e^{-(\alpha/\phi)((v_1 + v_2)/v_1 v_2)}\right) - \delta \frac{v_1 + v_2}{v_1 v_2} \\ &= \frac{\lambda + \delta}{\alpha} - \frac{\lambda\phi}{\alpha} \left(1 - e^{-(\alpha c - (\lambda + \delta))/\phi\delta}\right), \end{aligned} \tag{115}$$

so the necessary and sufficient condition of $b = 0$ is $\lambda + \delta \geq \lambda\phi(1 - e^{-(\alpha c - (\lambda + \delta))/\phi\delta})$. □

Based on the discussion above, we obtain the expression of $f(x)$. The following proposition will verify the concavity of $f(x)$.

Proposition 14. $f(x)$ *is concave on* $[0, \infty)$.

Proof. When $x \in [0, b)$, from (95) and (108), we have

$$\begin{aligned} f''(x) &= \frac{v_2 v_1}{v_2 - v_1} e^{v_1(x - b)} + \frac{v_1 v_2}{v_1 - v_2} e^{v_2(x - b)} \\ &= \frac{v_2 v_1}{v_2 - v_1} \left(e^{v_1(x - b)} - e^{v_2(x - b)}\right) < 0 \end{aligned} \tag{116}$$

due to the fact $v_1 < 0$ and $v_2 > 0$. What is more, $f''(x) = 0$ for $x \geq b$. Therefore $f''(x) \leq 0$ on $[0, \infty)$. This establishes the concavity of $f(x)$ on $[0, \infty)$. □

Proposition 15. *$f(x)$ is the solution of HJB equation (72) when $x \in [0, \infty)$.*

Proof. From the construction of $f(x)$,

$$f'(x) = 1 \quad \text{for } x \geq b$$
$$c + \lambda \int_0^{x+f(0)/\phi} f(x-y)\,\alpha e^{-\alpha y}\,dy - (\lambda+\delta) f(x) = 0 \quad \text{for } 0 \leq x \leq b \tag{117}$$

are established obviously. We only remain to show that $f(x)$ satisfies

$$f'(x) > 1 \quad \text{for } 0 \leq x < b, \tag{118}$$
$$c + \lambda \int_0^{x+f(0)/\phi} f(x-y)\,\alpha e^{-\alpha y}\,dy - (\lambda+\delta) f(x) < 0 \quad \text{for } x > b, \tag{119}$$
$$f'(x) < \phi \quad \text{for } x > 0. \tag{120}$$

From the concavity of $f(x)$ and $f'(b) = 1$, (118) is true.

Similar to the proof in Lemma 12 we can show (119) is established.

To prove (120), according to the concavity of $f(x)$, we only need to show $f'(0) < \phi$. Let $x = 0$ in (91) and assume that $f'(0) \geq \phi$. We find

$$\begin{aligned} f(0) &= \frac{\alpha c f'(0) - \lambda\phi + \lambda\phi e^{-(\alpha/\phi) f(0)}}{\alpha\delta} \\ &\geq \frac{\alpha c\phi - \lambda\phi + \lambda\phi e^{-(\alpha/\phi) f(0)}}{\alpha\delta} \\ &= \phi\frac{\alpha c - \lambda + \lambda e^{-(\alpha/\phi) f(0)}}{\alpha\delta} \\ &> \phi\frac{\alpha c - \lambda - \delta}{\alpha\delta} = \phi f(b). \end{aligned} \tag{121}$$

The last equality comes from $f(b) = (\alpha c - \lambda - \delta)/\alpha\delta$ which is proved in Lemma 12. While $f(0) > \phi f(b)$ is impossible because $f(x)$ is increasing and $\phi > 1$. This also tells us that $f'(0) < \phi$. Therefore $f'(0) \leq \phi$. And the proof is completed. □

The following theorem gives the optimal value function and optimal strategies when the claim size is exponentially distributed.

Theorem 16. *Suppose $F(x) = 1 - e^{-\alpha x}$. The value function $V(x)$ and the optimal strategy are as follows.*

(1) *If $\lambda + \delta < \lambda\phi(1 - e^{-(\alpha c - (\lambda+\delta))/\phi\delta})$, the value function*

$$V(x) = \begin{cases} 0 & \text{if } x < -z^*, \\ V(0) + \phi x & \text{if } -z^* \leq x < 0, \\ C_1 e^{v_1 x} + C_2 e^{v_2 x} & \text{if } 0 \leq x < b, \\ x - b + C_1 e^{v_1 b} + C_2 e^{v_2 b} & \text{if } x \geq b, \end{cases} \tag{122}$$

where C_1, C_2 are given by (108). The optimal lower capital injection barrier $-z^* = -V(0)/\phi$ and the optimal upper capital injection barrier $a = 0$. The optimal dividend barrier b can be calculated from (110).

(2) *If $\lambda + \delta \geq \lambda\phi(1 - e^{-(\alpha c - (\lambda+\delta))/\phi\delta})$, the value function*

$$V(x) = \begin{cases} 0 & \text{if } x < -z^*, \\ V(0) + \phi x = \left[\phi W\left(\dfrac{\lambda}{\delta} e^{-(\alpha c - \lambda\phi)/\phi\delta}\right) + \dfrac{\alpha c - \lambda\phi}{\delta}\right](\alpha + \phi x)^{-1} & \text{if } -z^* \leq x < 0, \\ x + V(0) = x + \left[\phi W\left(\dfrac{\lambda}{\delta} e^{-(\alpha c - \lambda\phi)/\phi\delta}\right) + \dfrac{\alpha c - \lambda\phi}{\delta}\right](\alpha)^{-1} & \text{if } x \geq 0, \end{cases} \tag{123}$$

where $W(x)$ is Lambert W function which is the solution of $W(x)e^{W(x)} = x$. The optimal lower capital injection barrier $-z^* = -V(0)/\phi$ and the optimal upper capital injection barrier $a = 0$. The optimal dividend barrier $b = 0$.

Proof. (1) For $x \geq 0$, because $f(x)$ is the solution of HJB equation (72) on $[0, \infty)$, from Theorem 11, we know $V(x)$ coincides with $f(x)$ on $[0, \infty)$. Because $V(x)$ is concave on $[0, \infty)$, Proposition 9 and the expression (73) inform us what are the optimal upper and lower capital injection barriers. Under the condition in (1), $b > 0$ by Proposition 13. b can be derived by (110). When $x < 0$, the expression of $V(x)$ has been discussed in (74). Therefore, (122) is established. Figure 1 shows us the sample path of the reserve process under the optimal strategy π^* and Figure 2 is the figure of the value function $V(x)$.

(2) If $\lambda + \delta \geq \lambda\phi(1 - e^{-(\alpha c - (\lambda+\delta))/\phi\delta})$, then $b = 0$ by Proposition 13. $b = 0$ means that under the optimal strategy, the shareholders will act as the insurer: they receive

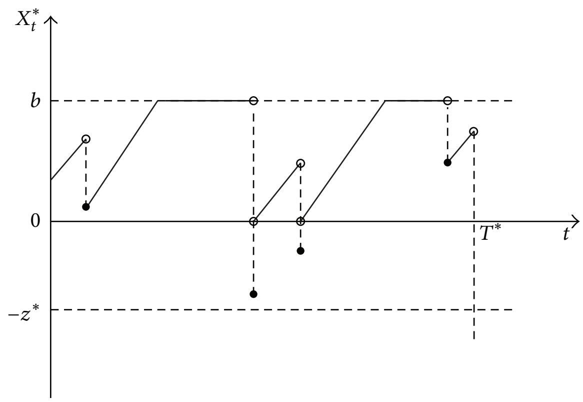


FIGURE 1: The sample path of the reserve process under the optimal strategy π^*.

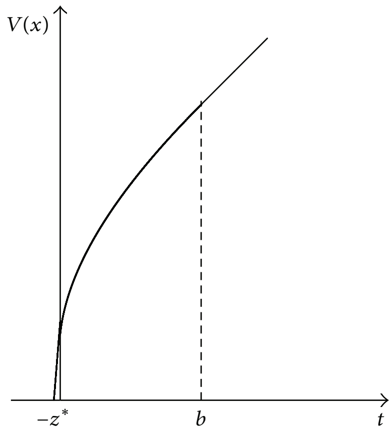


FIGURE 2: The value function $V(x)$.

the premium income and pay each claim in full when it occurs (see Dickson and Waters [7]). $V(0)$ must be recalculated by

$$\begin{aligned} V(0) &= E\left[\int_0^{T_1} ce^{-\delta t}\mathrm{d}t + e^{-\delta T_1}\left(V(0) - \phi Y_1\right) 1_{(Y_1 \le V(0)/\phi)}\right] \\ &= \int_0^{\infty} \lambda e^{-\lambda s}\int_0^{s} ce^{-\delta t}\mathrm{d}t\,\mathrm{d}s \\ &\quad + \int_0^{\infty} \lambda e^{-\lambda s}e^{-\delta s}\int_0^{V(0)/\phi}\left(V(0) - \phi y\right)\alpha e^{\alpha y}\mathrm{d}y\,\mathrm{d}s \\ &= \frac{\alpha c - \lambda\phi}{\alpha\delta} + \frac{\lambda\phi}{\alpha\delta}e^{-\alpha(V(0)/\phi)}. \end{aligned} \tag{124}$$

So $V(0) = [\phi W((\lambda/\delta)e^{-(\alpha c-\lambda\phi)/\phi\delta}) + (\alpha c - \lambda\phi)/\delta]/\alpha$, where $W(x)$ is Lambert W function which is the solution of $W(x)e^{W(x)} = x$. $-z^* = -V(0)/\phi$ and $a = 0$ are same as the discussion in proof of (1). Therefore, $V(x)$ can be expressed by (123). Figure 3 is the figure of the value function $V(x)$.

Note that it is the first time that Lambert W function is used in the risk theory. It simplifies the expression of $V(x)$ when $b = 0$. □

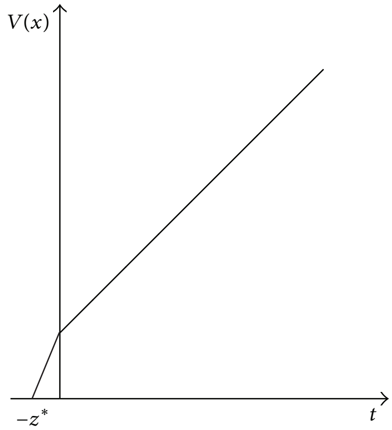


FIGURE 3: The value function $V(x)$.

Conflict of Interests

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

Acknowledgment

This work has been supported by The Natural Science Foundation of China (Grant no. 11471218 and no. 71303047).

References

[1] B. Borch, "The rescue of an insurance company after ruin," *ASTIN Bulletin*, vol. 15, no. 2, pp. 280–292, 1968.

[2] S. P. Sethi and M. I. Taksar, "Optimal financing of a corporation subject to random returns," *Mathematical Finance*, vol. 12, no. 2, pp. 155–172, 2002.

[3] A. Løkka and M. Zervos, "Optimal dividend and issuance of equity policies in the presence of proportional costs," *Insurance: Mathematics & Economics*, vol. 42, no. 3, pp. 954–961, 2008.

[4] L. He and Z. Liang, "Optimal financing and dividend control of the insurance company with proportional reinsurance policy," *Insurance: Mathematics & Economics*, vol. 42, no. 3, pp. 976–983, 2008.

[5] L. He and Z. Liang, "Optimal financing and dividend control of the insurance company with fixed and proportional transaction costs," *Insurance: Mathematics & Economics*, vol. 44, no. 1, pp. 88–94, 2009.

[6] N. Bäuerle, "Approximation of optimal reinsurance and dividend payout policies," *Mathematical Finance*, vol. 14, no. 1, pp. 99–113, 2004.

[7] D. C. Dickson and H. R. Waters, "Some optimal dividends problems," *Astin Bulletin*, vol. 34, no. 1, pp. 47–74, 2004.

[8] H. U. Gerber, E. S. Shiu, and N. Smith, "Maximizing dividends without bankruptcy," *Astin Bulletin*, vol. 36, no. 1, pp. 5–23, 2006.

[9] N. Kulenko and H. Schmidli, "Optimal dividend strategies in a Cramér–Lundberg model with capital injections," *Insurance: Mathematics and Economics*, vol. 43, no. 2, pp. 270–278, 2008.

[10] P. Azcue and N. Muler, "Optimal reinsurance and dividend distribution policies in the Cramér-Lundberg model," *Mathematical Finance*, vol. 15, no. 2, pp. 261–308, 2005.

[11] W. H. Fleming and H. M. Soner, *Controlled Markov Pro cesses and Viscosity Solutions*, vol. 25 of *Applications of Mathematics*, Springer, New York, NY, USA, 1993.

[12] R. Radner and L. Shepp, "Risk vs. profit potential: a model for corporate strategy," *Journal of Economic Dynamics and Control*, vol. 20, no. 8, pp. 1373–1393, 1996.

[13] S. Asmussen and M. Taksar, "Controlled diffusion models for optimal dividend pay-out," *Insurance: Mathematics & Economics*, vol. 20, no. 1, pp. 1–15, 1997.

[14] B. Højgaard and M. Taksar, "Optimal proportional reinsurance policies for diffusion models," *Scandinavian Actuarial Journal*, vol. 1998, no. 2, pp. 166–180, 1998.

[15] B. Højgaard and M. Taksar, "Controlling risk exposure and dividends payout schemes: insurance company example," *Mathematical Finance*, vol. 9, no. 2, pp. 153–182, 1999.

[16] C. Maglaras, "Discrete-review policies for scheduling stochastic networks: trajectory tracking and fluid-scale asymptotic optimality," *The Annals of Applied Probability*, vol. 10, no. 3, pp. 897–929, 2000.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 892304, 13 pages
http://dx.doi.org/10.1155/2015/892304

Research Article

Maximum Principle for Optimal Control Problems of Forward-Backward Regime-Switching Systems Involving Impulse Controls

Shujun Wang and Zhen Wu

School of Mathematics, Shandong University, Jinan, Shandong 250100, China

Correspondence should be addressed to Zhen Wu; wuzhen@sdu.edu.cn

Received 15 April 2014; Accepted 28 August 2014

Academic Editor: Guangchen Wang

This paper is concerned with optimal control problems of forward-backward Markovian regime-switching systems involving impulse controls. Here the Markov chains are continuous-time and finite-state. We derive the stochastic maximum principle for this kind of systems. Besides the Markov chains, the most distinguishing features of our problem are that the control variables consist of regular and impulsive controls, and that the domain of regular control is not necessarily convex. We obtain the necessary and sufficient conditions for optimal controls. Thereafter, we apply the theoretical results to a financial problem and get the optimal consumption strategies.

1. Introduction

Maximum principle was first formulated by Pontryagin et al.'s group [1] in the 1950s and 1960s, which focused on the deterministic control system to maximize the corresponding Hamiltonian instead of the optimization problem. Bismut [2] introduced the linear backward stochastic differential equations (BSDEs) as the adjoint equations, which played a role of milestone in the development of this theory. The general stochastic maximum principle was obtained by Peng in [3] by introducing the second order adjoint equations. Pardoux and Peng also proved the existence and uniqueness of solution for nonlinear BSDEs in [4], which has been extensively used in stochastic control and mathematical finance. Independently, Duffie and Epstein introduced BSDEs under economic background, and in [5] they presented a stochastic recursive utility which was a generalization of the standard additive utility with the instantaneous utility depending not only on the instantaneous consumption rate but also on the future utility. Then El Karoui et al. gave the formulation of recursive utilities from the BSDE point of view. As found by [6], the recursive utility process can be regarded as a solution of BSDE. Peng [7] first introduced the stochastic maximum principle for optimal control problems of forward-backward control system as the control domain is convex. Since BSDEs and forward-backward stochastic differential equations (FBSDEs) are involved in a broad range of applications in mathematical finance, economics, and so on, it is natural to study the control problems involving FBSDEs. To establish the necessary optimality conditions, Pontryagin maximum principle is one fundamental research direction for optimal control problems. Rich literature for stochastic maximum principle has been obtained; see [8–12] and the references therein. Recently, Wu [13] established the general maximum principle for optimal controls of forward-backward stochastic systems in which the control domains were nonconvex and forward diffusion coefficients explicitly depended on control variables.

The applications of regime-switching models in finance and stochastic control also have been researched in recent years. Compared to the traditional system based on the diffusion processes, it is more meaningful from the empirical point of view. Specifically, it modulates the system with a continuous-time finite-state Markov chain with each state representing a regime of the system or a level of economic indicator. Based on the switching diffusion model, much work has been done in the fields of option pricing, portfolio management, risk management, and so on. In [14], Crépey focused on the pricing equations in finance. Crépey and

Matoussi [15] investigated the reflected BSDEs with Markov chains. For the controlled problem with regime-switching model, Donnelly studied the sufficient maximum principle in [16]. Using the results about BSDEs with Markov chains in [14, 15], Tao and Wu [17] derived the maximum principle for the forward-backward regime-switching model. Moreover, in [18] the weak convergence of BSDEs with regime switching was studied. For more results of Markov chains, readers can refer to the references therein.

In addition, stochastic impulse control problems have received considerable research attention due to their wide applications in portfolio optimization problems with transaction costs (see [19, 20]) and optimal strategy of exchange rates between different currencies [21, 22]. Korn [23] also investigated some applications of impulse control in mathematical finance. For a comprehensive survey of theory of impulse controls, one is referred to [24]. Wu and Zhang [25] first studied stochastic optimal control problems of forward-backward systems involving impulse controls, in which they assumed the domain of the regular controls was convex and obtained both the maximum principle and sufficient optimality conditions. Later on, in [26] they considered the forward-backward system in which the domain of regular controls was not necessarily convex and the control variable did not enter the diffusion coefficient.

In this paper, we consider a stochastic control system, in which the control system is described by a forward-backward stochastic differential equation, all the coefficients contain Markov chains, and the control variables consist of regular and impulsive parts. This case is more complicated than [17, 25, 26]. We obtain the stochastic maximum principle by using spike variation on the regular control and convex perturbation on the impulsive one. Applying the maximum principle to a financial investment-consumption model, we also get the optimal consumption processes and analyze the effects on consumption by various economic factors.

The rest of this paper is organized as follows. In Section 2, we give preliminaries and the formulation of our problems. A necessary condition in the form of maximum principle is established in Section 3. Section 4 aims to investigate sufficient optimality conditions. An example in finance is studied in Section 5 to illustrate the applications of our theoretical results and some figures are presented to give more explanations. In the end, Section 6 concludes the novelty of this paper.

2. Preliminaries and Problem Formulation

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{0\le t\le T}, P)$ be a complete filtered probability space equipped with a natural filtration $\mathscr{F}_t$ generated by $\{B_s, \alpha_s;\ 0 \le s \le t\}$, $t \in [0,T]$, where $\{B_t\}_{0\le t\le T}$ is a d-dimensional standard Brownian motion defined on the space, $\{\alpha_t,\ 0 \le t \le T\}$ is a finite-state Markov chain with the state space given by $I = \{1, 2, \dots, k\}$, and $T \ge 0$ is a fixed time horizon. The transition intensities are $\lambda(i,j)$ for $i \ne j$ with $\lambda(i,j)$ nonnegative and bounded. $\lambda(i,i) = -\sum_{j\in I\setminus\{i\}} \lambda(i,j)$. For $p \ge 1$, denote by $S^p(\mathbb{R}^n)$ the set of n-dimensional adapted processes $\{\varphi_t,\ 0 \le t \le T\}$ such that $\mathbb{E}[\sup_{0\le t\le T}|\varphi_t|^p] < +\infty$ and denote by $H^p(\mathbb{R}^n)$ the set of n-dimensional adapted processes $\{\psi_t,\ 0 \le t \le T\}$ such that $\mathbb{E}[(\int_0^T |\psi_t|^2 dt)^{p/2}] < +\infty$.

Define $\mathscr{V}$ as the integer-valued random measure on $([0,T]\times I, \mathscr{B}([0,T])\otimes\mathscr{B}_I)$ which counts the jumps $\mathscr{V}_t(j)$ from α to state j between time 0 and t. The compensator of $\mathscr{V}_t(j)$ is $\mathbf{1}_{\{\alpha_t\ne j\}}\lambda(\alpha_t, j)dt$, which means $d\mathscr{V}_t(j) - \mathbf{1}_{\{\alpha_t\ne j\}}\lambda(\alpha_t, j)dt := d\widetilde{\mathscr{V}}_t(j)$ is a martingale (compensated measure). Then the canonical special semimartingale representation for α is given by

$$d\alpha_t = \sum_{j\in I}\lambda(\alpha_t, j)(j-\alpha_t)\,dt + \sum_{j\in I}(j-\alpha_{t-})\,d\widetilde{\mathscr{V}}_t(j). \tag{1}$$

Define $n_t(j) := \mathbf{1}_{\{\alpha_t\ne j\}}\lambda(\alpha_t, j)$. Denote by $\mathscr{M}_\rho$ the set of measurable functions from $(I, \mathscr{B}_I, \rho)$ to $\mathbb{R}$ endowed with the topology of convergence in measure and $|v|_t := \sum_{j\in I}[v(j)^2 n_t(j)]^{1/2} \in \mathbb{R}_+ \cup \{+\infty\}$ the norm of $\mathscr{M}_\rho$; denote by $H^p_{\mathscr{V}}$ the space of $\widetilde{P}$-measurable functions $V : \Omega\times[0,T]\times I \to \mathbb{R}$ such that $\sum_{j\in I}\mathbb{E}[(\int_0^T V_t(j)^2 n_t(j)dt)^{p/2}] < +\infty$.

Let U be a nonempty subset of $\mathbb{R}^k$ and K nonempty convex subset of $\mathbb{R}^n$. Let $\{\tau_i\}$ be a given sequence of increasing $\mathscr{F}_t$-stopping times such that $\tau_i \uparrow +\infty$ as $i \to +\infty$. Denote by $\mathscr{I}$ the class of right continuous processes $\eta(\cdot) = \sum_{i\ge1}\eta_i\mathbf{1}_{[\tau_i,T]}(\cdot)$ such that each η_i is an $\mathscr{F}_{\tau_i}$-measurable random variable. It's worth noting that, the assumption $\tau_i \uparrow +\infty$ implies that at most finitely many impulses may occur on $[0,T]$. Denote by $\mathscr{U}$ the class of adapted processes $v : [0,T]\times\Omega \to U$ such that $\mathbb{E}[\sup_{0\le t\le T}|v_t|^3] < +\infty$ and denote by $\mathscr{K}$ the class of K-valued impulse processes $\eta \in \mathscr{I}$ such that $\mathbb{E}[(\sum_{i\ge1}|\eta_i|)^3] < +\infty$. $\mathscr{A} := \mathscr{U}\times\mathscr{K}$ is called the admissible control set. For notational simplicity, in what follows we focus on the case where all processes are 1-dimensional.

Now we consider the forward regime-switching systems modulated by continuous-time, finite-state Markov chains involving impulse controls. Let $b : [0,T]\times I\times\mathbb{R} \to \mathbb{R}$, $\sigma : [0,T]\times I\times\mathbb{R} \to \mathbb{R}$, and $C : [0,T] \to \mathbb{R}$ be measurable mappings. Given $x \in \mathbb{R}$ and $\eta(\cdot) \in \mathscr{K}$, the system is formulated by

$$\begin{aligned} dx_t &= b(t,\alpha_t,x_t)\,dt + \sigma(t,\alpha_t,x_t)\,dB_t + C_t d\eta_t, \\ x_0 &= x. \end{aligned} \tag{2}$$

The following result is easily obtained.

Proposition 1. *Assume that b, σ are Lipschitz with respect to x, $b(\cdot,i,0)$, $\sigma(\cdot,i,0) \in H^3(\mathbb{R})$, $\forall i \in I$, and C is a continuous function. Then SDE (2) admits a unique solution $x(\cdot) \in S^3(\mathbb{R})$.*

Given $\zeta \in L^3(\Omega, \mathscr{F}_T, P; \mathbb{R})$ and $\eta(\cdot) \in \mathscr{K}$, consider the following backward regime-switching system modulated by Markov chains α_t involving impulse controls:

$$\begin{aligned} dy_t &= -f(t,\alpha_t,y_t,z_t,W_t(1)n_t(1),\dots,W_t(k)n_t(k))\,dt \\ &\quad + z_t dB_t + \sum_{j\in I} W_t(j)\,d\widetilde{\mathscr{V}}_t(j) - D_t d\eta_t, \\ y_T &= \zeta, \end{aligned} \tag{3}$$

where $f : [0,T] \times I \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$ and $D : [0,T] \to \mathbb{R}$ are measurable mappings and $W : \Omega \times [0,T] \times I \to \mathbb{R}$ is a measurable function such that $\sum_{j\in I} \mathbb{E} [(\int_0^T W_t(j)^2 n_t(j) dt)^{3/2}] < +\infty$.

Proposition 2. *Assume that $f(t,i,y,z,p)$ is Lipschitz with respect to (y,z,p), $f(\cdot,i,0,0,0) \in H^3(\mathbb{R})$, $\forall i \in I$, and D is a continuous function. Then BSDE (3) admits a unique solution $(y(\cdot),z(\cdot),W(\cdot)) \in S^3(\mathbb{R}) \times H^3(\mathbb{R}) \times H^3_{\mathscr{V}}(\mathbb{R})$.*

Proof. Define $A_t := \int_0^t D_s d\eta_s = \sum_{\tau_i \le t} D_{\tau_i}\eta_i$ and $F(t,i,y,z,p) := f(t,i,y-A_t,z,p)$, $\forall i \in I$. It is easy to check that

$$\begin{aligned} &\left|F(t,i,y,z,p) - F(t,i,y',z',p')\right| \\ &\quad \le c_1 \left(\left|y-y'\right| + \left|z-z'\right| + \left|p-p'\right|\right). \end{aligned} \tag{4}$$

Since n_t is uniformly bounded, we have

$$\left|\left(W_t(j) - W_t'(j)\right) n_t(j)\right| \le c_2 \left|W_t - W_t'\right|_t, \quad \forall j \in I. \tag{5}$$

Here c_1, c_2 are positive constants. Then F is Lipschitz with respect to (y,z,W). We also get that $F(\cdot,i,0,0,0) \in H^3(\mathbb{R})$ and $\mathbb{E}|\zeta + A_T|^3 < +\infty$. Hence, the following BSDE

$$\begin{aligned} dY_t = &-F(t,\alpha_t,Y_t,Z_t,M_t(1)n_t(1),\dots,M_t(k)n_t(k))\,dt \\ &+ Z_t dB_t + \sum_{j\in I} M_t(j)\,d\widetilde{\mathscr{V}}_t(j), \\ &Y_T = \zeta + A_T \end{aligned} \tag{6}$$

admits a unique solution $(Y,Z,M) \in S^3(\mathbb{R}) \times H^3(\mathbb{R}) \times H^3_{\mathscr{V}}(\mathbb{R})$ (see [15, 18] for details). Now define $y_t := Y_t - A_t$, $z_t := Z_t$, and $W_t := M_t$. Then it is easy to check that $(y,z,W) \in S^3(\mathbb{R}) \times H^3(\mathbb{R}) \times H^3_{\mathscr{V}}(\mathbb{R})$ solves BSDE (3).

Let (y^1,z^1,W^1) and (y^2,z^2,W^2) be two solutions of (3). Applying Itô's formula to $(y_s^1 - y_s^2)^2$, $t \le s \le T$ and combining Gronwall's inequality, we get the uniqueness of solution. □

Now, we consider the following stochastic control system:

$$\begin{aligned} dx_t &= b(t,\alpha_t,x_t,v_t)\,dt + \sigma(t,\alpha_t,x_t)\,dB_t + C_t d\eta_t, \\ dy_t &= -f(t,\alpha_t,x_t,y_t,z_t,W_t(1)n_t(1),\dots,W_t(k)n_t(k),v_t)\,dt \\ &\quad + z_t dB_t + \sum_{j\in I} W_t(j)\,d\widetilde{\mathscr{V}}_t(j) - D_t d\eta_t, \\ x_0 &= x, \qquad y_T = g(x_T), \end{aligned} \tag{7}$$

where $b : [0,T] \times I \times \mathbb{R} \times U \to \mathbb{R}$, $\sigma : [0,T] \times I \times \mathbb{R} \to \mathbb{R}$, $f : [0,T] \times I \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^k \times U \to \mathbb{R}$, and $g : \mathbb{R} \to \mathbb{R}$ are deterministic measurable functions and $C : [0,T] \to \mathbb{R}$, $D : [0,T] \to \mathbb{R}$ are continuous functions. In what follows $(W_t(1)n_t(1),\dots,W_t(k)n_t(k))$ will be written as $W_t n_t$ for short. The objective is to maximize, over class $\mathscr{A}$, the cost functional

$$\begin{aligned} J(v(\cdot),\eta(\cdot)) = \mathbb{E}\Bigg\{ &\int_0^T h(t,\alpha_t,x_t,y_t,z_t,W_t n_t,v_t)\,dt + \phi(x_T) \\ &+ \gamma(y_0) + \sum_{i\ge 1} l(\tau_i,\eta_i) \Bigg\}, \end{aligned} \tag{8}$$

where $h : [0,T] \times I \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^k \times U \to \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$, $\gamma : \mathbb{R} \to \mathbb{R}$, and $l : [0,T] \times \mathbb{R} \to \mathbb{R}$ are deterministic measurable functions. A control (u,ξ) which solves this problem is called an optimal control.

In what follows, we make the following assumptions.

(H1) b, σ, f, g, h, ϕ, and γ are continuous and continuously differentiable with respect to (x,y,z,p). b, f have linear growth with respect to (x,y,v). l is continuous and continuously differentiable with respect to η.

(H2) The derivatives of b, σ, f, and g are bounded.

(H3) The derivatives of h, ϕ, γ, and l are bounded by $K(1 + |x| + |y| + |z| + |p| + |v|)$, $K(1+|x|)$, $K(1+|y|)$, and $K(1+|\eta|)$, respectively. Moreover, $|h(t,i,0,0,0,0,v)| \le K(1+|v|^3)$ for any (t,v), $i \in I$.

From Propositions 1 and 2, it follows that, under (H1)–(H3), FBSDE (7) admits a unique solution $(x(\cdot),y(\cdot),z(\cdot),W(\cdot)) \in S^3(\mathbb{R}) \times S^3(\mathbb{R}) \times H^3(\mathbb{R}) \times H^3_{\mathscr{V}}(\mathbb{R})$ for any $(v,\eta) \in \mathscr{A}$.

3. Stochastic Maximum Principle

In this section, we will derive the stochastic maximum principle for optimal control problem (7) and (8). We give the necessary conditions for optimal controls.

Let $\xi(\cdot) = \sum_{i\ge 1} \xi_i \mathbf{1}_{[\tau_i,T]}(\cdot)$ and $(u(\cdot),\xi(\cdot)) \in \mathscr{A}$ be an optimal control of this stochastic control problem and let $(x(\cdot),y(\cdot),z(\cdot),W(\cdot))$ be the corresponding trajectory. Now, we introduce the spike variation with respect to $u(\cdot)$ as follows:

$$u^\varepsilon(t) = \begin{cases} v, & \text{if } \tau \le t \le \tau + \varepsilon, \\ u(t), & \text{otherwise}, \end{cases} \tag{9}$$

where $\tau \in [0,T)$ is an arbitrarily fixed time, $\varepsilon > 0$ is a sufficiently small constant, and v is an arbitrary U-valued $\mathscr{F}_\tau$-measurable random variable such that $\mathbb{E}|v|^3 < +\infty$. Let $\eta \in \mathscr{I}$ be such that $\xi + \eta \in \mathscr{K}$. For the reason that domain K is convex, we can check that $\xi^\varepsilon := \xi + \varepsilon\eta$, $0 \le \varepsilon \le 1$, is also an element of $\mathscr{K}$. Let $(x^\varepsilon(\cdot),y^\varepsilon(\cdot),z^\varepsilon(\cdot),W^\varepsilon(\cdot))$ be the trajectory corresponding to $(u^\varepsilon(\cdot),\xi^\varepsilon(\cdot))$. For convenience, we denote $\psi(t) = \psi(t,\alpha_t,x_t,y_t,z_t,W_t n_t,u_t)$, $\psi(u_t^\varepsilon) = \psi(t,\alpha_t,x_t,y_t,z_t,W_t n_t,u_t^\varepsilon)$ for $\psi = b, \sigma, f, h, b_x, b_v, \sigma_x, \sigma_v, f_x, f_y, f_z, f_{w(j)}, f_v, h_x, h_y, h_z, h_{w(j)}, h_v$, where $f_{w(j)} := f_{W(j)n(j)}$, $h_{w(j)} := h_{W(j)n(j)}$.

Introduce the following FBSDE which is called the variational equation:

$$\begin{aligned} dx_t^1 &= \left[b_x(t)x_t^1 + b(u_t^\varepsilon) - b(t)\right]dt + \sigma_x(t)x_t^1 dB_t \\ &\quad + \varepsilon C_t d\eta_t, \\ dy_t^1 &= -\Bigg[f_x(t)x_t^1 + f_y(t)y_t^1 + f_z(t)z_t^1 \\ &\quad + \sum_{j\in I} f_{w(j)}(t)P_t(j)n_t(j) + f(u_t^\varepsilon) - f(t)\Bigg]dt \\ &\quad + z_t^1 dB_t + \sum_{j\in I} P_t(j)\,d\widetilde{\mathscr{V}}_t(j) - \varepsilon D_t d\eta_t, \\ x_0^1 &= 0, \qquad y_T^1 = g_x(x_T)x_T^1. \end{aligned} \tag{10}$$

Obviously, this FBSDE admits a unique solution $(x^1, y^1, z^1, P) \in S^3(\mathbb{R}) \times S^3(\mathbb{R}) \times H^3(\mathbb{R}) \times H^3_{\mathscr{V}}(\mathbb{R})$.

We have the following lemma. In what follows, we denote by c a positive constant which can be different from line to line.

Lemma 3. *Consider*

$$\sup_{0\le t\le T}\mathbb{E}\left|x_t^1\right|^3 \le c\varepsilon^3, \tag{11}$$

$$\begin{aligned} &\sup_{0\le t\le T}\mathbb{E}\left|y_t^1\right|^3 + \mathbb{E}\left[\left(\int_0^T \left|z_t^1\right|^2 dt\right)^{3/2}\right] \\ &\quad + \sum_{j\in I}\mathbb{E}\left[\left(\int_0^T \left|P_t(j)\right|^2 n_t(j)\,dt\right)^{3/2}\right] \le c\varepsilon^3. \end{aligned} \tag{12}$$

Proof. By the boundedness of (b_x, σ_x) and using Hölder's inequality, we have

$$\begin{aligned} &\sup_{0\le t\le r}\mathbb{E}\left|x_t^1\right|^3 \\ &\quad \le c\int_0^r \left[\sup_{0\le s\le t}\mathbb{E}\left|x_s^1\right|^3\right]dt \\ &\quad + c\mathbb{E}\left(\int_0^T \left|b(u_t^\varepsilon) - b(t)\right|dt\right)^3 + c\varepsilon^3\mathbb{E}\left(\int_0^T \left|C_t\right|d\eta_t\right)^3, \end{aligned} \tag{13}$$

$\forall 0 \le r \le T$. Noting the definition of $u^\varepsilon(\cdot)$, we get

$$\begin{aligned} &\mathbb{E}\left(\int_0^T \left|b(u_t^\varepsilon) - b(t)\right|dt\right)^3 \\ &\quad = \mathbb{E}\left(\int_\tau^{\tau+\varepsilon} \left|b(t, \alpha_t, x_t, v) - b(t)\right|dt\right)^3 \\ &\quad \le \varepsilon^2\mathbb{E}\int_\tau^{\tau+\varepsilon} \left|b(t, \alpha_t, x_t, v) - b(t)\right|^3 dt \\ &\quad \le c\varepsilon^3\left(1 + \sup_{0\le t\le T}\mathbb{E}\left[\left|x_t\right|^3 + \left|u_t\right|^3 + |v|^3\right]\right) \\ &\quad \le c\varepsilon^3. \end{aligned} \tag{14}$$

Here we apply Hölder's inequality for $p = 3$, $q = 3/2$, and the growth condition of b in (H1). Since C_t is bounded on $[0, T]$, then (11) is obtained by applying Gronwall's inequality.

By the result of Section 5 in [6] and noting that the predictable covariation of $\widetilde{\mathscr{V}}_t(j)$ is

$$d\left\langle \widetilde{\mathscr{V}}_t(j), \widetilde{\mathscr{V}}_t(j)\right\rangle_t = n_t(j)\,dt, \tag{15}$$

we obtain

$$\begin{aligned} &\sup_{0\le t\le T}\mathbb{E}\left|y_t^1\right|^3 + \mathbb{E}\left[\left(\int_0^T \left|z_t^1\right|^2 dt\right)^{3/2}\right] \\ &\quad + \sum_{j\in I}\mathbb{E}\left[\left(\int_0^T \left|P_t(j)\right|^2 n_t(j)\,dt\right)^{3/2}\right] \\ &\le c\mathbb{E}\left|g_x(x_T)x_T^1\right|^3 + c\mathbb{E}\left(\int_0^T \left|f_x(t)x_t^1 + f(u_t^\varepsilon) - f(t)\right|dt\right)^3 \\ &\quad + c\varepsilon^3\mathbb{E}\left(\int_0^T \left|D_t\right|d\eta_t\right)^3. \end{aligned} \tag{16}$$

On the one hand, since g_x is bounded, by (11), we have

$$\mathbb{E}\left|g_x(x_T)x_T^1\right|^3 \le c\varepsilon^3. \tag{17}$$

On the other hand, since f_x is bounded, using the basic inequality and (11), we have

$$\begin{aligned} &\mathbb{E}\left(\int_0^T \left|f_x(t)x_t^1 + f(u_t^\varepsilon) - f(t)\right|dt\right)^3 \\ &\quad \le c\varepsilon^3 + c\mathbb{E}\left(\int_0^T \left|f(u_t^\varepsilon) - f(t)\right|dt\right)^3. \end{aligned} \tag{18}$$

From the growth condition of f in (H1) and the same technique as above, it follows that

$$\mathbb{E}\left(\int_0^T \left|f(u_t^\varepsilon) - f(t)\right|dt\right)^3 \le c\varepsilon^3. \tag{19}$$

Besides, D_t is bounded on $[0, T]$; then (12) is obtained. The proof is complete. ☐

Denote $\widehat{x}_t = x_t^\varepsilon - x_t - x_t^1$, $\widehat{y}_t = y_t^\varepsilon - y_t - y_t^1$, $\widehat{z}_t = z_t^\varepsilon - z_t - z_t^1$, and $\widehat{W}_t = W_t^\varepsilon - W_t - P_t$, and then we have the following.

Lemma 4. *Consider*

$$\sup_{0\le t\le T}\mathbb{E}\left|\widehat{x}_t\right|^2 \le C_\varepsilon \varepsilon^2, \tag{20}$$

$$\sup_{0\le t\le T}\mathbb{E}\left|\widehat{y}_t\right|^2 + \mathbb{E}\left[\int_0^T \left|\widehat{z}_t\right|^2 dt\right] + \sum_{j\in I}\mathbb{E}\left[\int_0^T \left|\widehat{W}_t(j)\right|^2 n_t(j)\,dt\right] \le C_\varepsilon \varepsilon^2, \tag{21}$$

where $C_\varepsilon \to 0$ *as* $\varepsilon \to 0$.

Proof. It is easy to check that $\widehat{x}$ satisfies

$$d\widehat{x}_t = \left[\Lambda_1(t) + \Lambda_2(t)\right]dt + \left[\Xi_1(t) + \Xi_2(t)\right]dB_t, \qquad \widehat{x}_0 = 0, \tag{22}$$

where

$$\begin{aligned}
\Lambda_1(t) &:= b\left(t,\alpha_t,x_t^\varepsilon,u_t^\varepsilon\right) - b\left(t,\alpha_t,x_t+x_t^1,u_t^\varepsilon\right),\\
\Lambda_2(t) &:= b\left(t,\alpha_t,x_t+x_t^1,u_t^\varepsilon\right) - b\left(u_t^\varepsilon\right) - b_x(t)\,x_t^1,\\
\Xi_1(t) &:= \sigma\left(t,\alpha_t,x_t^\varepsilon\right) - \sigma\left(t,\alpha_t,x_t+x_t^1\right),\\
\Xi_2(t) &:= \sigma\left(t,\alpha_t,x_t+x_t^1\right) - \sigma(t) - \sigma_x(t)\,x_t^1.
\end{aligned} \tag{23}$$

Then we have

$$\sup_{0\le t\le r}\mathbb{E}\left|\widehat{x}_t\right|^2 \le c\mathbb{E}\left(\int_0^r \left|\Lambda_1(t)+\Lambda_2(t)\right|dt\right)^2 + c\mathbb{E}\int_0^r \left|\Xi_1(t)+\Xi_2(t)\right|^2 dt \tag{24}$$

$\forall 0\le r\le T$. Since $\Lambda_1(t) = \int_0^1 b_x(t,\alpha_t,x_t+x_t^1+\lambda\widehat{x}_t,u_t^\varepsilon)d\lambda\widehat{x}_t$, by the boundedness of b_x, we have $|\Lambda_1(t)|\le c|\widehat{x}_t|$. Further we get

$$\mathbb{E}\left(\int_0^r \left|\Lambda_1(t)\right|dt\right)^2 \le c\mathbb{E}\int_0^r \left|\widehat{x}_t\right|^2 dt. \tag{25}$$

On the other hand, since $\Lambda_2(t) = \int_0^1 [b_x(t,\alpha_t,x_t+\lambda x_t^1,u_t^\varepsilon) - b_x(t)]d\lambda x_t^1$, we have

$$\int_0^r \left|\Lambda_2(t)\right|dt \le \int_0^T \left|\Lambda_2(t)\right|dt \le I_1 + I_2, \tag{26}$$

where

$$\begin{aligned}
I_1 &:= \int_\tau^{\tau+\varepsilon}\left|\int_0^1 \left[b_x\left(t,\alpha_t,x_t+\lambda x_t^1,v\right) - b_x(t)\right]d\lambda x_t^1\right|dt,\\
I_2 &:= \int_0^T\left|\int_0^1 \left[b_x\left(t,\alpha_t,x_t+\lambda x_t^1,u_t\right) - b_x(t)\right]d\lambda x_t^1\right|dt.
\end{aligned} \tag{27}$$

Since b_x is bounded, by Lemma 3 we get

$$\begin{aligned}
&\mathbb{E}\left|I_1\right|^2\\
&\le \varepsilon\int_\tau^{\tau+\varepsilon}\mathbb{E}\left[\left|\int_0^1 \left[b_x\left(t,\alpha_t,x_t+\lambda x_t^1,v\right) - b_x(t)\right]d\lambda x_t^1\right|^2\right]dt\\
&\le c\varepsilon^2 \sup_{0\le t\le T}\mathbb{E}\left|x_t^1\right|^2\\
&\le c\varepsilon^4.
\end{aligned} \tag{28}$$

For I_2, by Hölder's inequality, Lemma 3, and the dominated convergence theorem, it follows that

$$\begin{aligned}
&\mathbb{E}\left|I_2\right|^2\\
&\le \mathbb{E}\left\{\int_0^T\left|\int_0^1 \left[b_x\left(t,\alpha_t,x_t+\lambda x_t^1,u_t\right) - b_x(t)\right]d\lambda\right|^2 dt \cdot \int_0^T \left|x_t^1\right|^2 dt\right\}\\
&\le \left\{\mathbb{E}\left(\int_0^T\left|\int_0^1 \left[b_x\left(t,\alpha_t,x_t+\lambda x_t^1,u_t\right) - b_x(t)\right]d\lambda\right|^2 dt\right)^3\right\}^{1/3}\\
&\quad\times\left\{\mathbb{E}\left(\int_0^T \left|x_t^1\right|^2 dt\right)^{3/2}\right\}^{2/3}\\
&\le C_\varepsilon \varepsilon^2.
\end{aligned} \tag{29}$$

Then we get

$$\mathbb{E}\left(\int_0^r \left|\Lambda_2(t)\right|dt\right)^2 \le 2\mathbb{E}\left(\left|I_1\right|^2 + \left|I_2\right|^2\right) \le C_\varepsilon \varepsilon^2 \tag{30}$$

and obtain

$$\mathbb{E}\left(\int_0^r \left|\Lambda_1(t)+\Lambda_2(t)\right|dt\right)^2 \le C_\varepsilon\varepsilon^2 + c\mathbb{E}\int_0^r \left|\widehat{x}_t\right|^2 dt. \tag{31}$$

In the same way, we have

$$\mathbb{E}\int_0^r \left|\Xi_1(t)+\Xi_2(t)\right|^2 dt \le C_\varepsilon\varepsilon^2 + c\mathbb{E}\int_0^r \left|\widehat{x}_t\right|^2 dt. \tag{32}$$

From (24), (31), and (32) it follows that

$$\sup_{0\le t\le r}\mathbb{E}\left|\widehat{x}_t\right|^2 \le C_\varepsilon\varepsilon^2 + c\int_0^r \left[\sup_{0\le s\le t}\mathbb{E}\left|\widehat{x}_s\right|^2\right]dt. \tag{33}$$

Finally, applying Gronwall's inequality implies (20). □

To get estimate (21), for simplicity, we introduce

$$\begin{aligned}
\Theta_t &= \left(t,\alpha_t,x_t+\lambda x_t^1,y_t+\lambda y_t^1,z_t+\lambda z_t^1,\left(W_t+\lambda P_t\right)n_t\right),\\
\Sigma_t &= \left(t,\alpha_t,x_t+x_t^1+\lambda\widehat{x}_t,y_t+y_t^1+\lambda\widehat{y}_t,z_t+z_t^1+\lambda\widehat{z}_t,\right.\\
&\qquad\left.\left(W_t+P_t+\lambda\widehat{W}_t\right)n_t\right).
\end{aligned} \tag{34}$$

It is easy to check that $(\widehat{y}, \widehat{z}, \widehat{W})$ satisfies

$$d\widehat{y}_t = -\left[f_1(t) + f_2(t)\right] dt + \widehat{z}_t dB_t + \sum_{j\in I} \widehat{W}_t(j)\, d\widetilde{\mathscr{V}}_t(j),$$
$$\widehat{y}_T = G_1 + G_2, \tag{35}$$

where

$$\begin{aligned} f_1(t) &:= f\left(t, \alpha_t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, W_t^\varepsilon n_t, u_t^\varepsilon\right) \\ &\quad - f\left(t, \alpha_t, x_t + x_t^1, y_t + y_t^1, z_t + z_t^1, \left(W_t + P_t\right) n_t, u_t^\varepsilon\right), \\ f_2(t) &:= f\left(t, \alpha_t, x_t + x_t^1, y_t + y_t^1, z_t + z_t^1, \left(W_t + P_t\right) n_t, u_t^\varepsilon\right) \\ &\quad - f\left(u_t^\varepsilon\right) - f_x(t)\, x_t^1 - f_y(t)\, y_t^1 - f_z(t)\, z_t^1 \\ &\quad - \sum_{j\in I} f_{w(j)}(t)\, P_t(j)\, n_t(j), \\ G_1 &:= g\left(x_T^\varepsilon\right) - g\left(x_T + x_T^1\right), \\ G_2 &:= g\left(x_T + x_T^1\right) - g\left(x_T\right) - g_x\left(x_T\right) x_T^1. \end{aligned} \tag{36}$$

Similar to the proof above, we have

$$\begin{aligned} f_1(t) &= \int_0^1 f_x\left(\Sigma_t, u_t^\varepsilon\right) d\lambda \widehat{x}_t + \int_0^1 f_y\left(\Sigma_t, u_t^\varepsilon\right) d\lambda \widehat{y}_t \\ &\quad + \int_0^1 f_z\left(\Sigma_t, u_t^\varepsilon\right) d\lambda \widehat{z}_t \\ &\quad + \sum_{j\in I} \int_0^1 f_{w(j)}\left(\Sigma_t, u_t^\varepsilon\right) d\lambda \widehat{W}_t(j)\, n_t(j), \\ f_2(t) &= \int_0^1 \left[f_x\left(\Theta_t, u_t^\varepsilon\right) - f_x(t)\right] d\lambda x_t^1 \\ &\quad + \int_0^1 \left[f_y\left(\Theta_t, u_t^\varepsilon\right) - f_y(t)\right] d\lambda y_t^1 \\ &\quad + \int_0^1 \left[f_z\left(\Theta_t, u_t^\varepsilon\right) - f_z(t)\right] d\lambda z_t^1 \\ &\quad + \sum_{j\in I} \int_0^1 \left[f_{w(j)}\left(\Theta_t, u_t^\varepsilon\right) - f_{w(j)}(t)\right] d\lambda P_t(j)\, n_t(j), \\ G_1 &= \int_0^1 g_x\left(x_T + x_T^1 + \lambda \widehat{x}_T\right) d\lambda \widehat{x}_T, \\ G_2 &= \int_0^1 \left[g_x\left(x_T + \lambda x_T^1\right) - g_x\left(x_T\right)\right] d\lambda x_T^1. \end{aligned} \tag{37}$$

Then for BSDE (35), by the estimates of BSDEs, we obtain

$$\begin{aligned} &\sup_{0\le t\le T} \mathbb{E}\left|\widehat{y}_t\right|^2 + \mathbb{E}\left[\int_0^T \left|\widehat{z}_t\right|^2 dt\right] + \sum_{j\in I} \mathbb{E}\left[\int_0^T \left|\widehat{W}_t(j)\right|^2 n_t(j)\, dt\right] \\ &\le c\mathbb{E}\left[\left|\int_0^1 g_x\left(x_T + x_T^1 + \lambda \widehat{x}_T\right) d\lambda \widehat{x}_T\right|^2\right. \\ &\quad + \left|\int_0^1 \left[g_x\left(x_T + \lambda x_T^1\right) - g_x\left(x_T\right)\right] d\lambda x_T^1\right|^2 \\ &\quad + \left(\int_0^T \left|\int_0^1 f_x\left(\Sigma_t, u_t^\varepsilon\right) d\lambda \widehat{x}_t\right| dt\right)^2 \\ &\quad + \left(\int_0^T \left|\int_0^1 \left[f_x\left(\Theta_t, u_t^\varepsilon\right) - f_x(t)\right] d\lambda x_t^1\right| dt\right)^2 \\ &\quad + \left(\int_0^T \left|\int_0^1 \left[f_y\left(\Theta_t, u_t^\varepsilon\right) - f_y(t)\right] d\lambda y_t^1\right| dt\right)^2 \\ &\quad + \left(\int_0^T \left|\int_0^1 \left[f_z\left(\Theta_t, u_t^\varepsilon\right) - f_z(t)\right] d\lambda z_t^1\right| dt\right)^2 \\ &\quad + \left(\int_0^T \left|\sum_{j\in I} \int_0^1 \left[f_{w(j)}\left(\Theta_t, u_t^\varepsilon\right) - f_{w(j)}(t)\right] d\lambda P_t(j)\right.\right. \\ &\quad \left.\left.\left. \times n_t(j)\right| dt\right)^2\right]. \end{aligned} \tag{38}$$

Applying Hölder's inequality, Cauchy-Schwartz inequality, the dominated convergence theorem, Lemma 3, and (20) and noting the boundedness of n_t, we obtain (21).

Now, we are ready to state the variational inequality.

Lemma 5. *The following variational inequality holds:*

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \left(h_x(t)\, x_t^1 + h_y(t)\, y_t^1 + h_z(t)\, z_t^1\right.\right. \\ &\quad \left.\left. + \sum_{j\in I} h_{w(j)}(t)\, P_t(j)\, n_t(j) + h\left(u_t^\varepsilon\right) - h(t)\right) dt\right] \\ &\quad + \mathbb{E}\left[\phi_x\left(x_T\right) x_T^1 + \gamma_y\left(y_0\right) y_0^1 + \varepsilon \sum_{i\ge 1} l_\xi\left(\tau_i, \xi_i\right) \eta_i\right] \le o(\varepsilon). \end{aligned} \tag{39}$$

Proof. From the optimality of $(u(\cdot), \xi(\cdot))$, we have

$$J\left(u^\varepsilon(\cdot), \xi^\varepsilon(\cdot)\right) - J\left(u(\cdot), \xi(\cdot)\right) \le 0. \tag{40}$$

By Lemmas 3 and 4, we have

$$\begin{aligned}
&\mathbb{E}\left[\phi\left(x_T^{\varepsilon}\right)-\phi\left(x_T\right)\right]\\
&\quad=\mathbb{E}\left[\phi\left(x_T^{\varepsilon}\right)-\phi\left(x_T+x_T^1\right)\right]+\mathbb{E}\left[\phi\left(x_T+x_T^1\right)-\phi\left(x_T\right)\right]\\
&\quad=\mathbb{E}\left[\phi_x\left(x_T\right)x_T^1\right]+o(\varepsilon),\\
&\mathbb{E}\left[\gamma\left(y_0^{\varepsilon}\right)-\gamma\left(y_0\right)\right]\\
&\quad=\mathbb{E}\left[\gamma\left(y_0^{\varepsilon}\right)-\gamma\left(y_0+y_0^1\right)\right]+\mathbb{E}\left[\gamma\left(y_0+y_0^1\right)-\gamma\left(y_0\right)\right]\\
&\quad=\mathbb{E}\left[\gamma_y\left(y_0\right)y_0^1\right]+o(\varepsilon).
\end{aligned} \tag{41}$$

Similarly, we obtain

$$\mathbb{E}\left[\sum_{i\geq 1}l\left(\tau_i,\xi_i^{\varepsilon}\right)-\sum_{i\geq 1}l\left(\tau_i,\xi_i\right)\right]=\varepsilon\mathbb{E}\left[\sum_{i\geq 1}l_{\xi}\left(\tau_i,\xi_i\right)\eta_i\right]+o(\varepsilon). \tag{42}$$

Next, we aim to get the first term of (39). For convenience, we introduce two notations as follows:

$$\begin{aligned}
H_1:=\mathbb{E}\Bigg[\int_0^T\Big(&h\left(t,\alpha_t,x_t^{\varepsilon},y_t^{\varepsilon},z_t^{\varepsilon},W_t^{\varepsilon}n_t,u_t^{\varepsilon}\right)\\
&-h\big(t,\alpha_t,x_t+x_t^1,y_t+y_t^1,z_t\\
&+z_t^1,\left(W_t+P_t\right)n_t,u_t^{\varepsilon}\big)\Big)dt\Bigg],\\
H_2:=\mathbb{E}\Bigg[\int_0^T\Bigg(&h\Bigg(t,\alpha_t,x_t+x_t^1,y_t+y_t^1,\\
&z_t+z_t^1,\left(W_t+P_t\right)n_t,u_t^{\varepsilon}\Bigg)\\
&-h\left(u_t^{\varepsilon}\right)-h_x(t)x_t^1-h_y(t)y_t^1\\
&-h_z(t)z_t^1-\sum_{j\in I}h_{w(j)}(t)P_t(j)n_t(j)\Bigg)dt\Bigg].
\end{aligned} \tag{43}$$

Applying the same technique to the proof of Lemma 4, we obtain

$$H_1\sim H_2=o(\varepsilon). \tag{44}$$

Hence

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T\left(h\left(t,\alpha_t,x_t^{\varepsilon},y_t^{\varepsilon},z_t^{\varepsilon},W_t^{\varepsilon}n_t,u_t^{\varepsilon}\right)-h(t)\right)dt\right]\\
&=\mathbb{E}\Bigg[\int_0^T\Bigg(h_x(t)x_t^1+h_y(t)y_t^1+h_z(t)z_t^1\\
&\qquad+\sum_{j\in I}h_{w(j)}(t)P_t(j)n_t(j)+h\left(u_t^{\varepsilon}\right)-h(t)\Bigg)dt\Bigg]\\
&\quad+o(\varepsilon).
\end{aligned} \tag{45}$$

Thus, variational inequality (39) follows from (41)–(45).

Let us introduce the following adjoint equations:

$$\begin{aligned}
dp_t=&\left[f_y(t)p_t-h_y(t)\right]dt+\left[f_z(t)p_t-h_z(t)\right]dB_t\\
&+\sum_{j\in I}\left[f_{w(j)}(t-)p_{t-}-h_{w(j)}(t-)\right]d\widetilde{\mathscr{V}}_t(j),
\end{aligned} \tag{46}$$

$$p_0=-\gamma_y\left(y_0\right),$$

$$\begin{aligned}
-dq_t=&\left[b_x(t)q_t+\sigma_x(t)k_t-f_x(t)p_t+h_x(t)\right]dt\\
&-k_tdB_t-\sum_{j\in I}M_t(j)d\widetilde{\mathscr{V}}_t(j),
\end{aligned} \tag{47}$$

$$q_T=-g_x\left(x_T\right)p_T+\phi_x\left(x_T\right),$$

where $\varphi_{w(j)}(t-):=\varphi_{w(j)}(t,\alpha_{t-},x_{t-},y_{t-},z_t,W_tn_{t-},u_{t-})$ for $\varphi=f,h$. It is easy to check that SDE (46) admits a unique solution $p(\cdot)\in S^3(\mathbb{R})$. Besides, the generator of BSDE (47) does not contain $M_t(j)$. Therefore, the Lipschitz condition is satisfied obviously. Hence (47) admits a unique solution $(q(\cdot),k(\cdot),M(\cdot))\in S^3(\mathbb{R})\times H^3(\mathbb{R})\times H^3_{\mathscr{V}}(\mathbb{R})$. Now we establish the stochastic maximum principle. □

Theorem 6. *Let assumptions (H1)–(H3) hold. Suppose $(u(\cdot),\xi(\cdot))$ is an optimal control, $(x(\cdot),y(\cdot),z(\cdot),W(\cdot))$ is the corresponding trajectory, and $(p(\cdot),q(\cdot),k(\cdot),M(\cdot))$ is the solution of adjoint equations* (46) *and* (47)*. Then, $\forall v\in U,\eta(\cdot)\in\mathscr{K}$, it holds that*

$$\begin{aligned}
&H\left(t,\alpha_t,x_t,y_t,z_t,W_t,v,p_t,q_t,k_t\right)\\
&\quad-H\left(t,\alpha_t,x_t,y_t,z_t,W_t,u_t,p_t,q_t,k_t\right)\leq 0,\quad a.e.,\ a.s.,
\end{aligned} \tag{48}$$

$$\mathbb{E}\left[\sum_{i\geq 1}\left(l_{\xi}\left(\tau_i,\xi_i\right)+q_{\tau_i}C_{\tau_i}-p_{\tau_i}D_{\tau_i}\right)\left(\eta_i-\xi_i\right)\right]\leq 0, \tag{49}$$

where $H:[0,T]\times I\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathscr{M}_{\rho}\times U\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}$ is the Hamiltonian defined by

$$\begin{aligned}
&H\left(t,\alpha_t,x,y,z,W,v,p,q,k\right)\\
&\quad=-f\left(t,\alpha_t,x,y,z,Wn_t,v\right)p+b\left(t,\alpha_t,x,v\right)q+\sigma\left(t,\alpha_t,x\right)k\\
&\qquad+h\left(t,\alpha_t,x,y,z,Wn_t,v\right),
\end{aligned} \tag{50}$$

where $Wn_t=(W(1)n_t(1),\ldots,W(k)n_t(k))$.

Proof. Applying Itô's formula to $p_t y_t^1 + q_t x_t^1$ and combining with Lemma 5, we obtain

$$\mathbb{E}\left[\int_0^T \left(H\left(t,\alpha_t,x_t,y_t,z_t,W_t,u_t^\varepsilon,p_t,q_t,k_t\right)\right.\right.$$
$$\left.\left.-H\left(t,\alpha_t,x_t,y_t,z_t,W_t,u_t,p_t,q_t,k_t\right)\right)dt\right] \tag{51}$$
$$+\varepsilon\mathbb{E}\left[\sum_{i\geq 1}\left(l_\xi\left(\tau_i,\xi_i\right)+q_{\tau_i}C_{\tau_i}-p_{\tau_i}D_{\tau_i}\right)\theta_i\right]\leq o\left(\varepsilon\right),$$

where $\theta \in \mathcal{I}$ such that $\xi + \theta = \eta \in \mathcal{K}$. Then it follows that

$$\varepsilon^{-1}\mathbb{E}\left[\int_\tau^{\tau+\varepsilon}\left(H\left(t,\alpha_t,x_t,y_t,z_t,W_t,v,p_t,q_t,k_t\right)\right.\right.$$
$$\left.\left.-H\left(t,\alpha_t,x_t,y_t,z_t,W_t,u_t,p_t,q_t,k_t\right)\right)dt\right] \tag{52}$$
$$+\mathbb{E}\left[\sum_{i\geq 1}\left(l_\xi\left(\tau_i,\xi_i\right)+q_{\tau_i}C_{\tau_i}-p_{\tau_i}D_{\tau_i}\right)\theta_i\right]\leq 0.$$

Letting $\varepsilon \to 0$, we obtain

$$\mathbb{E}\left[H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,v,p_\tau,q_\tau,k_\tau\right)\right.$$
$$\left.-H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,u_\tau,p_\tau,q_\tau,k_\tau\right)\right]$$
$$+\mathbb{E}\left[\sum_{i\geq 1}\left(l_\xi\left(\tau_i,\xi_i\right)+q_{\tau_i}C_{\tau_i}-p_{\tau_i}D_{\tau_i}\right)\theta_i\right]\leq 0, \tag{53}$$
$$\text{a.e. } \tau\in[0,T].$$

By choosing $v = u_\tau$ we get (49). Setting $\eta \equiv \xi$, then for any $v \in \mathcal{F}_\tau$ we have

$$\mathbb{E}\left[H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,v,p_\tau,q_\tau,k_\tau\right)\right.$$
$$\left.-H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,u_\tau,p_\tau,q_\tau,k_\tau\right)\right]\leq 0,\quad \text{a.e.} \tag{54}$$

Let $v_\tau = v\mathbf{1}_A + u_\tau\mathbf{1}_{A^c}$ for $A \in \mathcal{F}_\tau$ and $v \in U$. Obviously $v_\tau \in \mathcal{F}_\tau$ and $\mathbb{E}|v_\tau|^3 < +\infty$. Then it follows that for any $A \in \mathcal{F}_\tau$

$$\mathbb{E}\left\{\left[H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,v,p_\tau,q_\tau,k_\tau\right)\right.\right.$$
$$\left.\left.-H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,u_\tau,p_\tau,q_\tau,k_\tau\right)\right]\mathbf{1}_A\right\}\leq 0,\quad \text{a.e.}, \tag{55}$$

which implies

$$\mathbb{E}\left\{\left[H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,v,p_\tau,q_\tau,k_\tau\right)\right.\right.$$
$$\left.\left.-H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,u_\tau,p_\tau,q_\tau,k_\tau\right)\right]\mid\mathcal{F}_\tau\right\}$$
$$=H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,v,p_\tau,q_\tau,k_\tau\right)$$
$$-H\left(\tau,\alpha_\tau,x_\tau,y_\tau,z_\tau,W_\tau,u_\tau,p_\tau,q_\tau,k_\tau\right)\leq 0,\quad \text{a.e., a.s.} \tag{56}$$

The proof is complete. □

4. Sufficient Optimality Conditions

In this section, we add additional assumptions to obtain the sufficient conditions for optimal controls. Let us introduce the following.

(H4) The control domain U is a convex body in $\mathbb{R}$. The measurable functions b, f, and l are locally Lipschitz with respect to v, and their partial derivatives are continuous with respect to (x, y, z, W, v).

Theorem 7. *Let (H1)–(H4) hold. Suppose that the functions $\phi(\cdot)$, $\gamma(\cdot)$, $\eta \to l(t,\eta)$, and $H(t,\alpha_t,\cdot,\cdot,\cdot,\cdot,\cdot,p_t,q_t,k_t)$ are concave and $(p(\cdot),q(\cdot),k(\cdot),M(\cdot))$ is the solution of adjoint equations* (46) *and* (47) *corresponding to control $(u(\cdot),\xi(\cdot)) \in \mathcal{A}$. Moreover, assume that $y_T^{v,\eta}$ is of the special form $y_T^{v,\eta} = K(\alpha_T)x_T^{v,\eta} + \zeta$, $\forall(v,\eta) \in \mathcal{A}$, where K is a deterministic measurable function and $\zeta \in L^3(\Omega,\mathcal{F}_T,P;\mathbb{R})$. Then (u,ξ) is an optimal control if it satisfies* (48) *and* (49).

Proof. Let $(x_t^{v,\eta}, y_t^{v,\eta}, z_t^{v,\eta}, W_t^{v,\eta})$ be the trajectory corresponding to $(v,\eta) \in \mathcal{A}$. By the concavity of ϕ, γ and $\eta \to l(t,\eta)$, we derive

$$J(v,\eta)-J(u,\xi)$$
$$\leq\mathbb{E}\left[\int_0^T\left(h\left(t,\alpha_t,x_t^{v,\eta},y_t^{v,\eta},z_t^{v,\eta},W_t^{v,\eta}n_t,v_t\right)-h\left(t\right)\right)dt\right]$$
$$+\mathbb{E}\left[\phi_x\left(x_T^{u,\xi}\right)\left(x_T^{v,\eta}-x_T^{u,\xi}\right)\right]$$
$$+\mathbb{E}\left[\gamma_y\left(y_0^{u,\xi}\right)\left(y_0^{v,\eta}-y_0^{u,\xi}\right)\right]$$
$$+\mathbb{E}\left[\sum_{i\geq 1}l_\xi\left(\tau_i,\xi_i\right)\left(\eta_i-\xi_i\right)\right]. \tag{57}$$

Define

$$\mathcal{H}^{v,\eta}(t):=H\left(t,\alpha_t,x_t^{v,\eta},y_t^{v,\eta},z_t^{v,\eta},W_t^{v,\eta},v_t,p_t,q_t,k_t\right). \tag{58}$$

Applying Itô's formula to $(x_t^{v,\eta} - x_t^{u,\xi})q_t + (y_t^{v,\eta} - y_t^{u,\xi})p_t$ and noting $q_T = -K(\alpha_T)p_T + \phi_x(x_T^{u,\xi})$, we obtain

$$J(v,\eta)-J(u,\xi)$$
$$\leq\mathbb{E}\left[\int_0^T\left(\mathcal{H}^{v,\eta}(t)-\mathcal{H}^{u,\xi}(t)-\mathcal{H}_x^{u,\xi}(t)\left(x_t^{v,\eta}-x_t^{u,\xi}\right)\right.\right.$$
$$-\mathcal{H}_y^{u,\xi}(t)\left(y_t^{v,\eta}-y_t^{u,\xi}\right)-\mathcal{H}_z^{u,\xi}(t)\left(z_t^{v,\eta}-z_t^{u,\xi}\right)$$
$$\left.\left.-\sum_{j\in I}\mathcal{H}_{w(j)}^{u,\xi}(t)\left(W_t^{v,\eta}(j)-W_t^{u,\xi}(j)\right)n_t(j)\right)dt\right]$$
$$+\mathbb{E}\left[\sum_{i\geq 1}\left(l_\xi\left(\tau_i,\xi_i\right)+q_{\tau_i}C_{\tau_i}-p_{\tau_i}D_{\tau_i}\right)\left(\eta_i-\xi_i\right)\right]$$
$$:=\Gamma_1+\Gamma_2. \tag{59}$$

By (48) and Lemma 2.3 of Chapter 3 in [27], we have

$$0 \in \partial_v \mathscr{H}^{u,\xi}(t). \tag{60}$$

By Lemma 2.4 of Chapter 3 in [27], we further conclude that

$$\left(\mathscr{H}_x^{u,\xi}(t), \mathscr{H}_y^{u,\xi}(t), \mathscr{H}_z^{u,\xi}(t), \mathscr{H}_W^{u,\xi}(t), 0\right) \in \partial_{x,y,z,W,v} \mathscr{H}^{u,\xi}(t). \tag{61}$$

Finally, by the concavity of $H(t, \alpha_t, \cdot, \cdot, \cdot, \cdot, \cdot, p_t, q_t, k_t)$ and (49), we obtain $\Gamma_1 \leq 0, \Gamma_2 \leq 0$. Thus, it follows that $J(v, \eta) - J(u, \xi) \leq 0$. We complete the proof. □

5. Application in Finance

This section is devoted to studying an investment and consumption model under the stochastic recursive utility arising from financial markets, which naturally motivates the study of the problem (7) and (8).

5.1. An Example in Finance. In a financial market, suppose there are two kinds of securities which can be invested: a bond, whose price $S_0(t)$ is given by

$$dS_0(t) = r_t S_0(t)\, dt, \quad S_0(0) > 0, \tag{62}$$

and a stock, whose price is

$$dS_1(t) = S_1(t)\left(\mu_t dt + \sigma_t dB_t\right), \quad S_1(0) > 0. \tag{63}$$

Here, $\{B_t\}$ is the standard Brownian motion and r_t, μ_t, and σ_t are bounded deterministic functions. For the sake of rationality, we assume $\mu_t > r_t, \sigma_t^2 \geq \delta > 0$. Here, δ stands for a positive constant, which ensures that σ_t is nondegenerate. In reality, in order to get stable profit and avoid risk of bankruptcy, many small companies and individual investors usually make a plan at the beginning of a year or a period, in which the weight invested in stock was fixed. Denote by π_t the weight invested in stock which is called the portfolio strategy. It means no matter how much the wealth x_t is, the portfolio strategy π_t is fixed, which is a bounded deterministic function with respect to t. Then the wealth dynamics are given as

$$\begin{aligned} dx_t = {} & \left[r_t x_t + (\mu_t - r_t)\pi_t x_t - c_t\right] dt \\ & + \sigma_t \pi_t x_t dB_t - \theta d\eta_t, \\ & x_0 = x > 0, \end{aligned} \tag{64}$$

where $\theta \geq 0$, $c_t \geq 0$, and $\eta_t = \sum_{i\geq 1} \eta_i \mathbf{1}_{[\tau_i, T]}(t)$. Here, c_t is a continuous consumption process, η_t is a piecewise consumption process, and θ is a weight factor. Not only in the mode of continuous consumption, but also in reality society, one consumes piecewise. Hence our setting of consumption process is practical.

Besides, if the macroeconomic conditions are also taken into account in this model, above model has obvious imperfections because it lacks the flexibility to describe the changing stochastically of investment environment. One can modulate the uncertainty of the economic situation by a continuous-time finite-state Markov chain. Then the wealth is formulated by a switching process as

$$\begin{aligned} dx_t = {} & \left[r(t,\alpha_t) x_t + \left(\mu(t,\alpha_t) - r(t,\alpha_t)\right)\pi(t,\alpha_t) x_t - c_t\right] dt \\ & + \sigma(t,\alpha_t)\pi(t,\alpha_t) x_t dB_t - \theta d\eta_t, \\ & x_0 = x, \qquad \alpha_0 = i. \end{aligned} \tag{65}$$

Let U be a nonempty subset of $\{\mathbb{R}_+ \cup 0\}$ and K a nonempty convex subset of $\{\mathbb{R}_+ \cup 0\}$. Suppose $\{\mathscr{F}_t\}$ is the natural filtration generated by the Brownian motion and the Markov chains, c_t is an $\mathscr{F}_t$-progressively measurable process satisfying

$$c_t \in U, \text{ a.s., a.e.,} \quad \mathbb{E}\int_0^T |c_t|^3 dt < +\infty, \tag{66}$$

$\{\tau_i\}$ is a fixed sequence of increasing $\mathscr{F}_t$-stopping times, and each η_i is an $\mathscr{F}_{\tau_i}$-measurable random variable satisfying

$$\eta_i \in K, \text{ a.s.,} \quad \mathbb{E}\left(\sum_{i\geq 1} |\eta_i|\right)^2 < +\infty. \tag{67}$$

We consider the following stochastic recursive utility, which is described by a BSDE with the Markov chain α_t:

$$\begin{aligned} -dy_t = {} & \left[b(t,\alpha_t) x_t + f(t,\alpha_t) y_t + g(t,\alpha_t) z_t - c_t\right] dt \\ & - z_t dB_t - \sum_{j\in I} W_t(j)\, d\widetilde{\mathscr{V}}_t(j) - \zeta d\eta_t, \\ & y_T = x_T, \end{aligned} \tag{68}$$

where $I = 1, 2, \ldots, k$, $\zeta \geq 0$. The recursive utility is meaningful in economics and theory. Details can be found in Duffie and Epstein [5] and El Karoui et al. [6].

Define the associated utility functional as

$$J(c(\cdot), \eta(\cdot)) = \mathbb{E}\left[\int_0^T L e^{-\beta t} \frac{(c_t)^{1-R}}{1-R} dt + \frac{S}{2}\sum_{i\geq 1}\eta_i^2 + H y_0\right], \tag{69}$$

where L, S, and H are positive constants, β is a discount factor, and $\beta \in (0,1)$ is also called Arrow-Pratt index of risk aversion (see, e.g., Karatzas and Shreve [28]). To get the explicit solution, we also assume $b(t, \alpha_t) \geq 0$. The first and second terms in (69) measure the total utility from $c(\cdot)$ and $\eta(\cdot)$, while the third term characterizes the initial reserve y_0. It is natural to desire to maximize the expected utility functional representing cumulative consumption and the recursive utility y_0, which means to find $(c(\cdot), \eta(\cdot))$ satisfying (66) and (67), respectively, to maximize $J(c(\cdot), \eta(\cdot))$ in (69).

We solve the problem by the maximum principle derived in Section 3. The Hamiltonian corresponding to this model is

$$\begin{aligned} &H(t,\alpha_t,x,y,z,c,p,q,k) \\ &= -p\left[b(t,\alpha_t)x + f(t,\alpha_t)y + g(t,\alpha_t)z - c\right] \\ &\quad + q\left[r(t,\alpha_t)x + (\mu(t,\alpha_t) - r(t,\alpha_t))\pi(t,\alpha_t)x - c\right] \\ &\quad + k\sigma(t,\alpha_t)\pi(t,\alpha_t)x + Le^{-\beta t}\frac{(c_t)^{1-R}}{1-R}, \end{aligned} \tag{70}$$

where (p,q,k,M) is the solution of the following adjoint equations:

$$\begin{gathered} dp_t = f(t,\alpha_t)p_t dt + g(t,\alpha_t)p_t dB_t, \\ p_0 = -H, \end{gathered} \tag{71}$$

$$\begin{aligned} -dq_t = &\left[(r(t,\alpha_t) + (\mu(t,\alpha_t) - r(t,\alpha_t))\pi(t,\alpha_t))q_t \right. \\ &\left. + \sigma(t,\alpha_t)\pi(t,\alpha_t)k_t - b(t,\alpha_t)p_t\right]dt \\ &- k_t dB_t - \sum_{j\in I} M_t(j)\, d\widetilde{\mathscr{V}}_t(j), \\ & q_T = -p_T. \end{aligned} \tag{72}$$

From (71) it is easy to obtain that

$$\begin{aligned} p_t = -H\exp\left\{\int_0^t \left[f(s,\alpha_s) - \frac{1}{2}g^2(s,\alpha_s)\right]ds \right. \\ \left. + \int_0^t g(s,\alpha_s)\,dB_s\right\} < 0. \end{aligned} \tag{73}$$

To solve (72), we introduce the dual process

$$\begin{aligned} d\Lambda_s = &\left[r(s,\alpha_s)(1-\pi(s,\alpha_s)) + \mu(s,\alpha_s)\pi(s,\alpha_s)\right]\Lambda_s ds \\ &+ \sigma(s,\alpha_s)\pi(s,\alpha_s)\Lambda_s dB_s, \\ &\Lambda_t = 1, \quad (s \geq t). \end{aligned} \tag{74}$$

Actually, (74) is solved by

$$\begin{aligned} \Lambda_s = \exp&\left\{\int_t^s \left[r(\tau,\alpha_\tau)(1-\pi(\tau,\alpha_\tau)) + \mu(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau) \right.\right. \\ &\left. - \frac{1}{2}\sigma^2(\tau,\alpha_\tau)\pi^2(\tau,\alpha_\tau)\right]d\tau \\ &\left. + \int_t^s \sigma(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau)\,dB_\tau\right\} > 0. \end{aligned} \tag{75}$$

Applying Itô's formula to $\Lambda_s q_s$ and taking conditional expectation with respect to $\mathscr{F}_t$, we obtain

$$\begin{aligned} q_t &= \mathbb{E}\left[-p_T\Lambda_T - \int_t^T b(s,\alpha_s)p_s\Lambda_s ds \mid \mathscr{F}_t\right] \\ &= H\mathbb{E}\left[\exp\left\{\int_0^T \left[f(\tau,\alpha_\tau) + r(\tau,\alpha_\tau)(1-\pi(\tau,\alpha_\tau)) \right.\right.\right. \\ &\qquad + \mu(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau) - \frac{1}{2}g^2(\tau,\alpha_\tau) \\ &\qquad \left. - \frac{1}{2}\sigma^2(\tau,\alpha_\tau)\pi^2(\tau,\alpha_\tau)\right]d\tau \\ &\qquad - \int_0^t \left[r(\tau,\alpha_\tau)(1-\pi(\tau,\alpha_\tau)) \right. \\ &\qquad + \mu(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau) \\ &\qquad \left. - \frac{1}{2}\sigma^2(\tau,\alpha_\tau)\pi^2(\tau,\alpha_\tau)\right]d\tau \\ &\qquad + \int_0^T \left[g(\tau,\alpha_\tau) + \sigma(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau)\right]dB_\tau \\ &\qquad \left. - \int_0^t \sigma(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau)\,dB_\tau\right\} + \int_t^T b(s,\alpha_s) \\ &\qquad \times \exp\left\{\int_0^s \left[f(\tau,\alpha_\tau) + r(\tau,\alpha_\tau)(1-\pi(\tau,\alpha_\tau)) \right.\right. \\ &\qquad + \mu(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau) - \frac{1}{2}g^2(\tau,\alpha_\tau) \\ &\qquad \left. - \frac{1}{2}\sigma^2(\tau,\alpha_\tau)\pi^2(\tau,\alpha_\tau)\right]d\tau \\ &\qquad - \int_0^t \left[r(\tau,\alpha_\tau)(1-\pi(\tau,\alpha_\tau)) \right. \\ &\qquad + \mu(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau) \\ &\qquad \left. - \frac{1}{2}\sigma^2(\tau,\alpha_\tau)\pi^2(\tau,\alpha_\tau)\right]d\tau \\ &\qquad + \int_0^s \left[g(\tau,\alpha_\tau) + \sigma(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau)\right]dB_\tau \\ &\qquad \left.\left. - \int_0^t \sigma(\tau,\alpha_\tau)\pi(\tau,\alpha_\tau)\,dB_\tau\right\} ds \mid \mathscr{F}_t\right]. \end{aligned} \tag{76}$$

Note that $b(t,\alpha_t) \geq 0$; then we have $q_t > 0$. Thus, by Theorem 6 we get the optimal consumption processes $(c^*(\cdot),\eta^*(\cdot))$ for the regime-switching investment-consumption problem (65)–(69) as follows:

$$\begin{aligned} c_t^* &= \left(\frac{L}{q_t - p_t}\right)^{1/R} e^{-\beta t/R}, \quad \text{a.e., a.s.,} \\ \eta_i^* &= \frac{\theta q_{\tau_i} - \zeta p_{\tau_i}}{S}, \quad \forall i \geq 1, \text{ a.s,} \end{aligned} \tag{77}$$

where (p_t, q_t) is given by (73) and (76), respectively.

5.2. Numerical Simulation. In this part, we calculate the optimal consumption functions explicitly according to (71)–(77) in the case that all coefficients are constants and discuss the relationship between consumption and some financial parameters, which can further illustrate our results obtained in this paper. We only consider the optimal regular consumption process $c^*(\cdot)$ and in this case the Markov chain $\alpha_t \equiv \alpha$ has two states $\{1, -1\}$. Here α_t will not change from 0 to T. Further we fix $[H, \beta, L, R] = [0.1, 0.5, 2, 0.2]$ and $T = 1$ year throughout this part.

5.2.1. The Relationship between $c^*(t)$ and r. As $\alpha = 1$, we set

$$\begin{aligned}&[r1, r2, r3, f(\alpha), g(\alpha), \pi(\alpha), \sigma(\alpha), b(\alpha), \mu] \\ &\quad = [0.02, 0.03, 0.04, 0.1, 0.1, 0.5, 0.2, 0.2, 0.05].\end{aligned} \tag{78}$$

From Figure 1, we find that the higher the risk-free interest rate is, the lower the optimal consumption is. It coincides with the financial behaviors in reality. As the risk-free interest rate r grows higher, the investors can gain more profits via deposit. Consequently, the desire of consumption is declined.

As $\alpha = -1$, we set

$$\begin{aligned}&[r1, r2, r3, f(\alpha), g(\alpha), \pi(\alpha), \sigma(\alpha), b(\alpha), \mu] \\ &\quad = [0.02, 0.03, 0.04, 0.05, 0.05, 0.4, 0.3, 0.15, 0.05].\end{aligned} \tag{79}$$

Figure 2 shows the influence of risk-free interest rate on the optimal consumption function as $\alpha = -1$. Same as Figure 1, when the risk-free interest rate gets higher, the optimal consumption becomes smaller. From Figures 1 and 2, we also find that under different strategies of government's macrocontrol (different α), the optimal consumption has different values and changes trends with respect to t, even for the same risk-free interest rate r. It is natural because α affects some parameters in this model such as f, g, π, σ, and b.

5.2.2. The Relationship between $c^*(t)$ and μ. The following two figures show the relationships between the optimal consumption function and appreciation rate of stock. First, for $\alpha = 1$, we fix

$$\begin{aligned}&[\mu 1, \mu 2, \mu 3, f(\alpha), g(\alpha), \pi(\alpha), \sigma(\alpha), b(\alpha), r] \\ &\quad = [0.05, 0.06, 0.07, 0.1, 0.1, 0.5, 0.2, 0.2, 0.02].\end{aligned} \tag{80}$$

From Figure 3, we can see that the higher the appreciation rate of stock is, the lower the optimal consumption is. It is also reasonable since a higher appreciation rate of stock μ inspires investors to put more money into stock market and thereby reduce the consumption. For $\alpha = -1$, we fix

$$\begin{aligned}&[\mu 1, \mu 2, \mu 3, f(\alpha), g(\alpha), \pi(\alpha), \sigma(\alpha), b(\alpha), r] \\ &\quad = [0.05, 0.06, 0.07, 0.05, 0.05, 0.4, 0.3, 0.15, 0.02].\end{aligned} \tag{81}$$

Figure 4 also presents the same influence of appreciation rate on the optimal consumption function as $\alpha = -1$. In addition, Figures 3 and 4 enhance us to understand that the

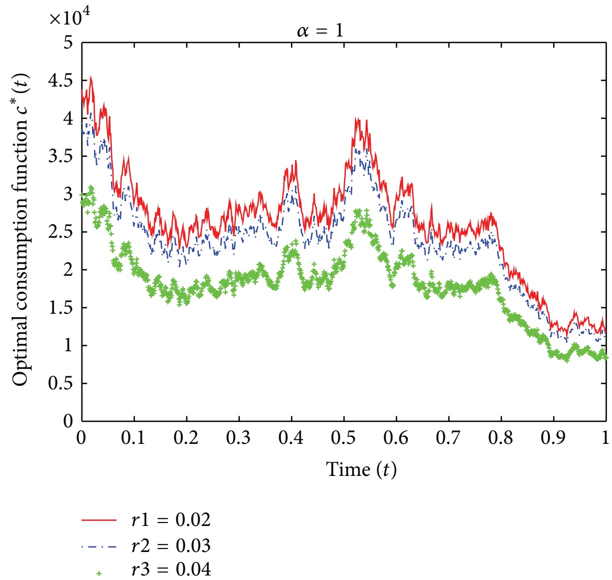


Figure 1: The relationship between $c^*(t)$ and r as $\alpha = 1$.

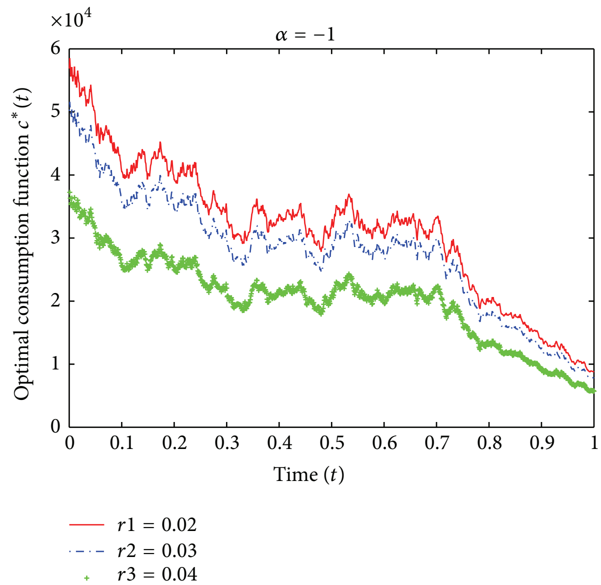


Figure 2: The relationship between $c^*(t)$ and r as $\alpha = -1$.

optimal consumption has different values and changes trends with respect to t for the same appreciation rate μ by considering different strategies of government's macrocontrol.

Based on Figures 1–4, we analyze the relationships between the optimal consumption function and the risk-free interest rate, the appreciation rate of stock, and the government's macrocontrol, which are quite important and applicable in financial problems.

6. Conclusion

In this paper, we consider the optimal control problem of forward-backward Markovian regime-switching systems

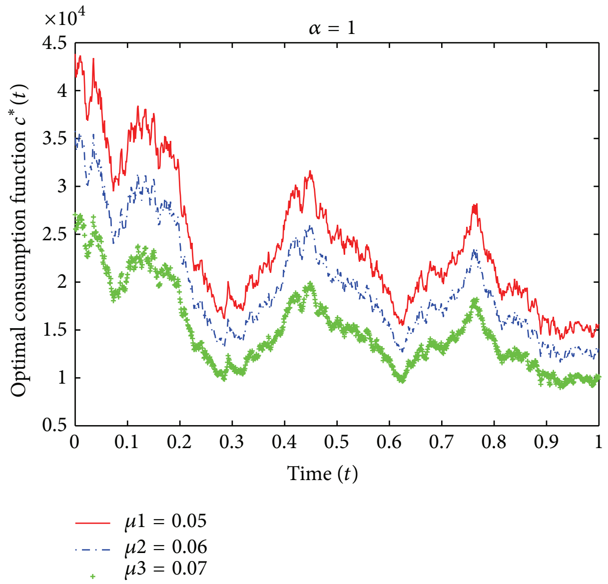


Figure 3: The relationship between $c^*(t)$ and μ as $\alpha = 1$.

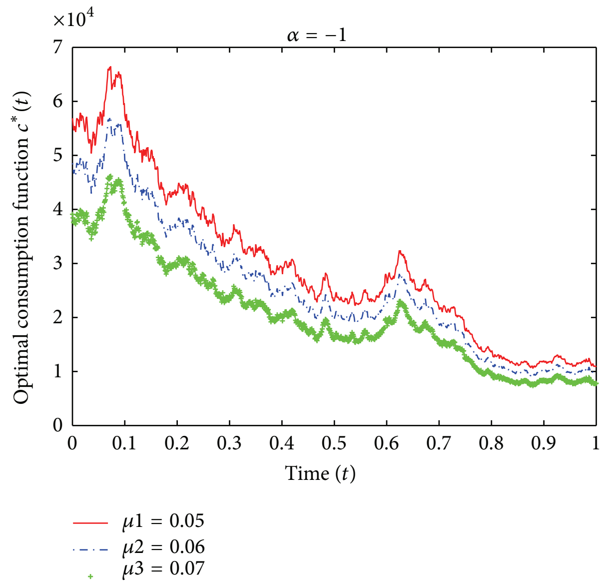


Figure 4: The relationship between $c^*(t)$ and μ as $\alpha = -1$.

involving impulse controls. The control system is described by FBSDEs involving impulse controls and modulated by continuous-time, finite-state Markov chains. Based on both spike and convex variation techniques, we establish the maximum principle and sufficient optimality conditions for optimal controls. Here, the regular control does not enter in the diffusion term of the forward system. In the future, we may focus on the cases that the diffusion coefficient contains controls, fully coupled forward-backward Markovian regime-switching system involving impulse controls, and game problems in this framework. It is worth pointing out that if the domain of regular control is not convex and the control enters in the forward diffusion coefficient, it will be more complicated and bring some difficulties immediately by applying spike variation. Based on the methods and results of [13], we hope to further research for such kind of control problems and investigate more applications in reality.

Conflict of Interests

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grants (11221061 and 61174092), 111 Project (B12023), the Natural Science Fund for Distinguished Young Scholars of China (11125102), and the Science and Technology Project of Shandong Province (2013GRC32201).

References

[1] L. Pontryagin, V. Boltyanskti, R. Gamkrelidze, and E. Mischenko, *The Mathematical Theory of Optimal Control Processes*, John Wiley & Sons, New York, NY, USA, 1962.

[2] J. Bismut, "An introductory approach to duality in optimal stochastic control," *SIAM Review*, vol. 20, no. 1, pp. 62–78, 1978.

[3] S. G. Peng, "A general stochastic maximum principle for optimal control problems," *SIAM Journal on Control and Optimization*, vol. 28, no. 4, pp. 966–979, 1990.

[4] E. Pardoux and S. Peng, "Adapted solution of a backward stochastic differential equation," *Systems and Control Letters*, vol. 14, no. 1, pp. 55–61, 1990.

[5] D. Duffie and L. G. Epstein, "Stochastic differential utility," *Econometrica*, vol. 60, no. 2, pp. 353–394, 1992.

[6] N. El Karoui, S. Peng, and M. C. Quenez, "Backward stochastic differential equations in finance," *Mathematical Finance*, vol. 7, no. 1, pp. 1–71, 1997.

[7] S. Peng, "Backward stochastic differential equations and applications to optimal control," *Applied Mathematics and Optimization*, vol. 27, no. 2, pp. 125–144, 1993.

[8] W. S. Xu, "Stochastic maximum principle for optimal control problem of forward and backward system," *Journal of the Australian Mathematical Society B*, vol. 37, no. 2, pp. 172–185, 1995.

[9] Z. Wu, "Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems," *Systems Science and Mathematical Sciences*, vol. 11, no. 3, pp. 249–259, 1998.

[10] J. Yong, "Optimality variational principle for controlled forward-backward stochastic differential equations with mixed initial-terminal conditions," *SIAM Journal on Control and Optimization*, vol. 48, no. 6, pp. 4119–4156, 2010.

[11] G. Wang and Z. Wu, "The maximum principles for stochastic recursive optimal control problems under partial information," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1230–1242, 2009.

[12] G. Wang and Z. Yu, "A Pontryagin's maximum principle for non-zero sum differential games of BSDEs with applications," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1742–1747, 2010.

[13] Z. Wu, "A general maximum principle for optimal control of forward-backward stochastic systems," *Automatica*, vol. 49, no. 5, pp. 1473–1480, 2013.

[14] S. Crépey, "About the pricing equations in finance," in *Paris-Princeton Lectures on Mathematical Finance 2010*, vol. 2003 of *Lecture Notes in Mathematics*, pp. 63–203, Springer, Berlin, Germany, 2011.

[15] S. Crépey and A. Matoussi, "Reflected and doubly reflected BSDEs with jumps: a priori estimates and comparison," *The Annals of Applied Probability*, vol. 18, no. 5, pp. 2041–2069, 2008.

[16] C. Donnelly, "Sufficient stochastic maximum principle in a regime-switching diffusion model," *Applied Mathematics and Optimization*, vol. 64, no. 2, pp. 155–169, 2011.

[17] R. Tao and Z. Wu, "Maximum principle for optimal control problems of forward-backward regime-switching system and applications," *Systems & Control Letters*, vol. 61, no. 9, pp. 911–917, 2012.

[18] R. Tao, Z. Wu, and Q. Zhang, "BSDEs with regime switching: weak convergence and applications," *Journal of Mathematical Analysis and Applications*, vol. 407, no. 1, pp. 97–111, 2013.

[19] M. H. A. Davis and A. R. Norman, "Portfolio selection with transaction costs," *Mathematics of Operations Research*, vol. 15, no. 4, pp. 676–713, 1990.

[20] B. Øksendal and A. Sulem, "Optimal consumption and portfolio with both fixed and proportional transaction costs," *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1765–1790, 2002.

[21] A. Cadenillas and F. Zapatero, "Classical and impulse stochastic control of the exchange rate using interest rates and reserves," *Mathematical Finance*, vol. 10, no. 2, pp. 141–156, 2000.

[22] M. Jeanblanc-Picqué, "Impulse control method and exchange rate," *Mathematical Finance*, vol. 3, pp. 161–177, 1993.

[23] R. Korn, "Some applications of impulse control in mathematical finance," *Mathematical Methods of Operations Research*, vol. 50, no. 3, pp. 493–518, 1999.

[24] B. M. Miller and E. Y. Rubinovich, *Impulsive Control in Continuous and Discrete-Continuous Systems*, Kluwer Academic, Dordrecht, The Netherlands, 2003.

[25] Z. Wu and F. Zhang, "Stochastic maximum principle for optimal control problems of forward-backward systems involving impulse controls," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1401–1406, 2011.

[26] Z. Wu and F. Zhang, "Maximum principle for stochastic recursive optimal control problems involving impulse controls," *Abstract and Applied Analysis*, vol. 2012, Article ID 709682, 16 pages, 2012.

[27] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, vol. 43 of *Applications of Mathematics*, Springer, New York, NY, USA, 1999.

[28] I. Karatzas and S. E. Shreve, *Methods of Mathematical Finance*, Springer, New York, NY, USA, 1998.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 571572, 8 pages
http://dx.doi.org/10.1155/2014/571572

Research Article

Performance Analysis for Distributed Fusion with Different Dimensional Data

Xianghui Yuan,[1] Zhansheng Duan,[2] and Chongzhao Han[1]

[1] *Ministry of Education Key Laboratory for Intelligent Networks and Network Security (MOE KLINNS), School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*
[2] *Center for Information Engineering Science Research (CIESR), School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*

Correspondence should be addressed to Zhansheng Duan; zsduan@mail.xjtu.edu.cn

Received 24 June 2014; Accepted 28 August 2014; Published 30 September 2014

Academic Editor: Ming Gao

Different sensors or estimators may have different capability to provide data. Some sensors can provide a relatively higher dimensional data, while other sensors can only provide part of them. Some estimators can estimate full dimensional quantity of interest, while others may only estimate part of it due to some constraints. How is such kind of data with different dimensions fused? How do the common part and the uncommon part affect each other during fusion? To answer these questions, a fusion algorithm based on linear minimum mean-square error (LMMSE) estimation is provided in this paper. Then the fusion performance is analyzed, which is the main contribution of this work. The conclusions are as follows. First, the fused common part is not affected by the uncommon part. Second, the fused uncommon part will benefit from the common part through the cross-correlation. Finally, under certain conditions, both the more accurate common part and the stronger correlation can result in more accurate fused uncommon part. The conclusions are all supported by some tracking application examples.

1. Introduction

Estimation of the stochastic system state or parameters has wide applications. For example, in target tracking applications, the evolution of the target state can often be represented by a stochastic dynamic system, where the state transition model is driven by some process noise. The observations of the measurement model are also corrupted by some measurement noise in general. Since the state model and measurement model are both stochastic, the output of the estimators, for example, a Kalman filter, is also stochastic. When there are multiple sensors or estimators, the data fusion techniques are usually used for potential better estimation purpose.

Data fusion is the problem of how to utilize useful information contained in multiple sets of data for the purpose of estimation of an unknown quantity—a parameter or a process [1]. The most common situation is that the data to be fused are of the same dimensions. But, in some cases, the data of different dimensions may need to be fused. The following are some examples to show the different dimensional data fusion in target tracking applications.

Measurement-to-Measurement Fusion. Suppose that we have two radars, A and B. Radar A can sense target 1 and target 2 simultaneously, while radar B can only sense target 1. Then the measurement-to-measurement fusion for such a scenario is a fusion problem with different dimensionalities.

Track-to-Track Fusion. Constant velocity (CV) model based estimator can only provide estimation of position and velocity, while constant acceleration (CA) model based estimator can provide estimation of position, velocity, and acceleration. The fusion of such two estimators is also a fusion problem with different dimensionalities. This is very common in maneuvering target tracking using the interacting multiple model (IMM) algorithm.

Measurement-to-Track Fusion. A CV model based estimator provides the target's state estimation of position and velocity, while a sensor (a radar or GPS) provides the target's position measurement. This is a measurement-to-track fusion problem with different dimensional data.

The reason for such phenomenons is that some sensors or estimators may be subject to some constraints compared to the full dimensional data provider. In the above examples, radar B may have narrower coverage than radar A; the CV model based estimator cannot provide acceleration estimation due to the model itself; the sensor cannot provide target velocity measurement because of its sensing capability.

For such kind of fusion with different dimensional data, how to deal with the uncommon part is a problem which needed to be considered. A simple way is to abandon the uncommon part when fusing. This is quite natural but some useful information will be lost. To fully use all available information, an LMMSE estimator is provided in this work. In fact, if the uncommon part and the common part have some kind of cross-correlation, the correlation will help in fusion.

The relationship between the correlation and the estimator's performance has been discussed in some literatures. For example, Doppler radar's range and range rate measurement errors are often correlated. Reference [2] concluded that negative correlation has the best tracking performance. With more detailed simulation and analysis, [3] concluded that, for steady state estimation, negative correlation has the best tracking performance, positive correlation is not always worse than without correlation. Reference [3] also discussed the coefficient selection strategy for one step state estimation. Reference [4] proposed a fusion algorithm in which local estimates have correlations. Reference [5] analyzed the fusion performance with the correlation for the scalar case. Reference [6–9] also disscussed the fusion algorithm in the existence of correlation. Although these literatures discussed the relationship between correlation and the fusion performance, the fusion performance analysis of the different dimensional data fusion is very rare. To reveal the factors which affect the fusion performance, the performance is analyzed in this paper.

The rest of the paper is organized as follows. Section 2 is the problem formulation part. Fusion algorithm is proposed in Section 3. Performance analysis is given in Section 4, which is the main contribution of this work. Some examples are given in Section 5 and Section 6 is the conclusion.

2. Problem Formulation

In general, filter or model's output can be seen as an estimator. In this work, for the unification of the problem formulation, sensor's measurement is also treated as an "estimator" in which the filter's output is the same as the input, the original measurement.

The following problem is considered. There are two estimators. One can provide the full dimensional estimate of an estimand (the quantity to be estimated), and the other can only provide partial estimate of the estimand. In this paper, the estimators are stochastic, which means estimators are affected by some noises.

Assume X is the estimand, which can be written as $X = [x^T \quad y^T]^T$.

Estimator 1 is as follows:

$$\widehat{X}_1 = \begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} v_1^x \\ v_1^y \end{bmatrix}. \tag{1}$$

Estimator 2 is as follows:

$$\widehat{X}_2 = \hat{x}_2 = x + v_2^x. \tag{2}$$

It can be seen that x is the common part and y is the uncommon part. The dimensions of those vectors are

$$\begin{gathered} \hat{x}_1, \hat{x}_2, x, v_1^x, v_2^x, \widehat{X}_2 \in R^{n\times 1}, \\ \hat{y}_1, y, v_1^y \in R^{m\times 1}, \\ X, \widehat{X}_1 \in R^{(n+m)\times 1}. \end{gathered} \tag{3}$$

The mean, covariance, and cross covariance of the noises are

$$\begin{gathered} E\left[v_1^x\right] = E\left[v_2^x\right] = \mathbf{0}_n, \qquad E\left[v_1^y\right] = \mathbf{0}_m, \\ \operatorname{cov}\left(\begin{bmatrix} v_1^x \\ v_1^y \end{bmatrix}\right) = P_1 = \begin{bmatrix} P_{11}^{xx} & P_{11}^{xy} \\ P_{11}^{yx} & P_{11}^{yy} \end{bmatrix} > 0, \\ \operatorname{cov}\left(v_2^x\right) = P_2 = P_{22}^{xx} > 0, \\ \operatorname{cov}\left(\begin{bmatrix} v_1^x \\ v_1^y \end{bmatrix}, v_2^x\right) = P_{12} = \begin{bmatrix} P_{12}^{xx} \\ P_{12}^{yx} \end{bmatrix}, \end{gathered} \tag{4}$$

where $P_1, P_2 > 0$ means P_1, P_2 are positive definite matrices.

3. Fusion Algorithm with Different Dimensional Data

3.1. Introduction to the LMMSE Estimator. The minimum mean-square error (MMSE) estimation is Bayesian estimation where the expected value of a positive definite cost function is to be minimized. It is a tool which estimates a random variable X in terms of another random variable Z. The solution is the conditional mean $E[X \mid Z]$.

Since the distributional information needed for the evaluation of the conditional mean is not always available, the linear minimum mean-square error (LMMSE) estimator is often used in practice. LMMSE estimator yields the estimate as a linear function of the observation and requires only the first two moments. It is a widely used estimation method.

Consider the vector-valued random variables X and Z, where Z is a measurement of X. The best estimate of X in terms of Z in LMMSE sense [10] is

$$\begin{gathered} \widehat{X} = \overline{X} + P_{XZ}P_{ZZ}^{-1}\left(Z - \overline{Z}\right), \\ P_{XX|Z} = \operatorname{MSE}\left(\widehat{X}\right) = P_{XX} - P_{XZ}P_{ZZ}^{-1}P_{ZX}, \end{gathered} \tag{5}$$

where $\overline{X}$ is the prior mean of X, P_{XX} is the prior covariance matrix of X, $\overline{Z}$ is the prior mean of Z, and P_{ZZ} is the prior covariance matrix of Z. P_{XZ} is the cross covariance matrix between X and Z.

The LMMSE estimator of one random vector in terms of another random vector (the measurement) is such that the estimation error is

(1) zero-mean,
(2) uncorrelated from the measurements.

LMMSE estimator has the following properties.

(1) It is the best estimator (in the MMSE sense) for Gaussian random variables.
(2) It is the best estimator within the class of linear estimators.

LMMSE estimation is essentially known as best linear unbiased estimation (BLUE) [1], which is proved to be identical to the linear weighted least squares (WLS) estimation [11].

3.2. Fusion Algorithm Using the LMMSE Estimation. Since $\widehat{X}_1$ can provide the full estimate of X, $\widehat{X}_1$ can be regarded as the prior information.

The prior information is as follows:

$$\overline{X} = \begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \end{bmatrix},$$
$$P_{XX} = \operatorname{cov}\left(\begin{bmatrix} v_1^x \\ v_1^y \end{bmatrix}\right) = \begin{bmatrix} P_{11}^{xx} & P_{11}^{xy} \\ P_{11}^{yx} & P_{11}^{yy} \end{bmatrix}. \tag{6}$$

Next, $\widehat{X}_2$ is regarded as the measurement. Since $\widehat{X}_1$ is the prior information,

$$\begin{aligned} Z &= \widehat{X}_2 = \hat{x}_2, \\ \overline{Z} &= E\left[\widehat{X}_2 \mid \widehat{X}_1\right] = \hat{x}_1, \\ P_{ZZ} &= E\left[\left(Z - \overline{Z}\right)\left(Z - \overline{Z}\right)'\right], \\ &= E\left[\left(\hat{x}_2 - \hat{x}_1\right)\left(\hat{x}_2 - \hat{x}_1\right)'\right], \\ &= P_{11}^{xx} + P_{22}^{xx} - P_{12}^{xx} - \left(P_{12}^{xx}\right)^T. \end{aligned} \tag{7}$$

The cross covariance between the prior information and the measurement is then

$$\begin{aligned} P_{XZ} &= E\left[\left(X - \overline{X}\right)\left(Z - \overline{Z}\right)'\right] \\ &= E\left[\left(X - \widehat{X}_1\right)\left(\hat{x}_2 - \hat{x}_1\right)'\right] \\ &= E\left[\begin{pmatrix} -v_1^x \\ -v_1^y \end{pmatrix}\left(v_2^x - v_1^x\right)'\right] \\ &= \begin{bmatrix} P_{11}^{xx} - P_{12}^{xx} \\ P_{11}^{yx} - P_{12}^{yx} \end{bmatrix}. \end{aligned} \tag{8}$$

Here it is assumed that $P_{ZZ} > 0$, which means Z or $\hat{x}_2$ can also provide some new information.

The LMMSE fuser for this problem is the following:

$$\begin{aligned} \widehat{X} &= \overline{X} + P_{XZ} P_{ZZ}^{-1}\left(Z - \overline{Z}\right) \\ &= \widehat{X}_1 + P_{XZ} P_{ZZ}^{-1}\left(\hat{x}_2 - \hat{x}_1\right) \\ &= \begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \end{bmatrix} + \begin{bmatrix} P_{11}^{xx} - P_{12}^{xx} \\ P_{11}^{yx} - P_{12}^{yx} \end{bmatrix}\left(P_{11}^{xx} + P_{22}^{xx} - P_{12}^{xx} - \left(P_{12}^{xx}\right)^T\right)^{-1} \\ &\quad \times\left(\hat{x}_2 - \hat{x}_1\right), \end{aligned} \tag{9}$$

$$\begin{aligned} P_{XX|Z} &= P_{XX} - P_{XZ} P_{ZZ}^{-1} P_{ZX} \\ &= \begin{bmatrix} P_{11}^{xx} & P_{11}^{xy} \\ P_{11}^{yx} & P_{11}^{yy} \end{bmatrix} - \begin{bmatrix} P_{11}^{xx} - P_{12}^{xx} \\ P_{11}^{yx} - P_{12}^{yx} \end{bmatrix} \\ &\quad \times\left(P_{11}^{xx} + P_{22}^{xx} - P_{12}^{xx} - \left(P_{12}^{xx}\right)^T\right)^{-1} \\ &\quad \times \begin{bmatrix} P_{11}^{xx} - P_{12}^{xx} \\ P_{11}^{yx} - P_{12}^{yx} \end{bmatrix}^T. \end{aligned} \tag{10}$$

It is the updated covariance $P_{XX|Z}$ which is used for performance analysis. $P_{XX|Z}$ can be rearranged as

$$P_{XX|Z} = \begin{bmatrix} P^{xx} & P^{xy} \\ P^{yx} & P^{yy} \end{bmatrix}, \tag{11}$$

where P^{xx} stands for the updated x part's (common data) covariance matrix:

$$\begin{aligned} P^{xx} &= P_{11}^{xx} - \left(P_{11}^{xx} - P_{12}^{xx}\right)\left(P_{11}^{xx} + P_{22}^{xx} - P_{12}^{xx} - \left(P_{12}^{xx}\right)^T\right)^{-1} \\ &\quad \times\left(P_{11}^{xx} - P_{12}^{xx}\right)^T. \end{aligned} \tag{12}$$

It is the same as the fusion algorithm in [4].

P^{yy} stands for the updated y part's (uncommon data) covariance matrix:

$$\begin{aligned} P^{yy} &= P_{11}^{yy} - \left(P_{11}^{yx} - P_{12}^{yx}\right)\left(P_{11}^{xx} + P_{22}^{xx} - P_{12}^{xx} - \left(P_{12}^{xx}\right)^T\right)^{-1} \\ &\quad \times\left(P_{11}^{yx} - P_{12}^{yx}\right)^T. \end{aligned} \tag{13}$$

It is affected by the x part. The following performance analysis is on the updated uncommon part (y part).

4. Performance Analysis of the Uncommon Part

4.1. The Uncommon Part's Impact on the Fused Common Part. From (12), it is very clear that the fused common part will not be affected by the uncommon part.

4.2. The Cross-Correlation's Impact on the Fused Uncommon Part. From (13), it can be easily seen that the fused uncommon part is affected by the common part.

First, some properties of the positive matrix are introduced. If $A, B \in R^{n\times n}$ are positive definite matrices, then they have the following properties [12].

(I) For $T \in R^{m\times n}$, if $\operatorname{rank}(T) = m$, then $TAT^T > 0$; otherwise $TAT^T \geq 0$.

(II) $A > 0 \Leftrightarrow A^{-1} > 0$.

(III) $A - B > 0 \Leftrightarrow B^{-1} - A^{-1} > 0$.

Before fusion, the covariance matrix of y part is P_{11}^{yy}. After fusion, it becomes P^{yy}. From (13), it can be seen that

$$\begin{aligned} P_{11}^{yy} - P^{yy} &= \left(P_{11}^{yx} - P_{12}^{yx}\right)\left(P_{11}^{xx} + P_{22}^{xx} - P_{12}^{xx} - \left(P_{12}^{xx}\right)^T\right)^{-1} \\ &\quad \times \left(P_{11}^{yx} - P_{12}^{yx}\right)^T \\ &= P^{yx} P_{ZZ}^{-1} \left(P^{yx}\right)^T, \end{aligned} \tag{14}$$

where $P^{yx} = P_{11}^{yx} - P_{12}^{yx}$ and P^{yx} can be regarded as the cross-correlation matrix.

Theorem 1. *If* $\operatorname{rank}(P^{yx}) = m$, *then* $P_{11}^{yy} - P^{yy} > 0$; *otherwise* $P_{11}^{yy} - P^{yy} \geq 0$.

Proof. Because $P_{ZZ} > 0$, from Property (II), it follows that $P_{ZZ}^{-1} > 0$.

The conclusion can then be directly obtained from (14) and Property (I). □

It can be seen from (14) that if $P^{yx} = \mathbf{0}$, $P_{11}^{yy} - P^{yy} = \mathbf{0}$ The following are the conclusions from the above.

(1) If $P^{yx} = \mathbf{0}$, which means there is no cross-correlation between x and y, the fused uncommon part will be the same as the unfused one.

(2) If $\operatorname{rank}(P^{yx}) = m$, which means the cross-correlation is full row rank, the fused uncommon part is definitely better than the unfused one.

If $P^{yx} \neq \mathbf{0}$ and $\operatorname{rank}(P^{yx}) < m$, the following shows which component of y will benefit from the fusion. Assume that

$$P^{yx} = \left[\left(p_1^{yx}\right)^T \ \left(p_2^{yx}\right)^T \ \cdots \ \left(p_m^{yx}\right)^T\right]^T, \tag{15}$$

where p_i^{yx}, $i = 1, \ldots, m$, are row vectors. If only the ith component of y is considered, the following corollaries can be obtained.

Corollary 2. *If* $p_i^{yx} \neq \mathbf{0}$, *then* $P_{11}^{yy}(i,i) - P^{yy}(i,i) > 0$.

Proof. It can be seen from (14) that $P_{11}^{yy}(i,i) - P^{yy}(i,i) = p_i^{yx} P_{ZZ}^{-1} \left(p_i^{yx}\right)^T$.

Thus if $p_i^{yx} \neq \mathbf{0}$, then $\operatorname{rank}(p_i^{yx}) = 1$.

Furthermore, since $P_{ZZ}^{-1} > 0$, from Property (I), it can be seen that $P_{11}^{yy}(i,i) - P^{yy}(i,i) > 0$. □

It can be seen from Corollary 2 that if one certain component of the uncommon part y is cross-correlated with the common part x, then its fused result is better than the unfused one.

Corollary 3. *If* $m = 1$ *and* $P^{yx} \neq \mathbf{0}$, $P_{11}^{yy} - P^{yy} > 0$.

Proof. If $m = 1$, then P^{yx} is a row vector.

If $P^{yx} \neq \mathbf{0}$, then $\operatorname{rank}(P^{yx}) = m = 1$.

From Corollary 2, Corollary 3 can be directly achieved. □

It can be seen from Corollary 3 that if the uncommon part y is a scalar and the cross-correlation exists, the fused result is better than the unfused one.

4.3. The Accuracy of the Independent Common Part's Impact on the Fused Uncommon Part. Assume that estimator $\widehat{X}_2$ can be obtained with different precision. The covariance matrix of higher precision is $P_{22,H}^{xx}$ and the covariance matrix of lower precision is $P_{22,L}^{xx}$. The corresponding fused covariance matrix of y is P_H^{yy} and P_L^{yy}. Assume that $P_{22,L}^{xx} - P_{22,H}^{xx} > 0$. If the two estimators $\widehat{X}_1$ and $\widehat{X}_2$ are independent, which means $P_{12}^{yx} = \mathbf{0}$ and $P_{12}^{xx} = \mathbf{0}$, the following theorem can be obtained.

Theorem 4. *Under the condition that* $\widehat{X}_1$ *and* $\widehat{X}_2$ *are independent, if* $\operatorname{rank}(P_{11}^{yx}) = m$, *then* $P_L^{yy} - P_H^{yy} > 0$; *otherwise* $P_L^{yy} - P_H^{yy} \geq 0$.

Proof. When $\widehat{X}_1$ and $\widehat{X}_2$ are independent, $P^{yx} = P_{11}^{yx}$. From (14), the fusion covariance for y is the following:

$$\begin{aligned} P_H^{yy} &= P_{11}^{yy} - P_{11}^{yx}\left(P_{11}^{xx} + P_{22,H}^{xx}\right)^{-1}\left(P_{11}^{yx}\right)^T, \\ P_L^{yy} &= P_{11}^{yy} - P_{11}^{yx}\left(P_{11}^{xx} + P_{22,L}^{xx}\right)^{-1}\left(P_{11}^{yx}\right)^T. \end{aligned} \tag{16}$$

The difference between the two covariance matrices is

$$P_L^{yy} - P_H^{yy} = P_{11}^{yx}\left(\left(P_{11}^{xx} + P_{22,H}^{xx}\right)^{-1} - \left(P_{11}^{xx} + P_{22,L}^{xx}\right)^{-1}\right)\left(P_{11}^{yx}\right)^T. \tag{17}$$

Since $P_{22,L}^{xx} - P_{22,H}^{xx} > 0$, it thus follows that $P_{11}^{xx} + P_{22,L}^{xx} - (P_{11}^{xx} + P_{22,H}^{xx}) > 0$.

From Property (III),

$$\left(P_{11}^{xx} + P_{22,H}^{xx}\right)^{-1} - \left(P_{11}^{xx} + P_{22,L}^{xx}\right)^{-1} > 0. \tag{18}$$

According to (17) and Property (I), if $\operatorname{rank}(P_{11}^{yx}) = m$, then $P_L^{yy} - P_H^{yy} > 0$; otherwise $P_L^{yy} - P_H^{yy} \geq 0$. □

It can be seen from Theorem 4 that increasing the independent common part's accuracy can improve the fused performance of uncommon part.

The following two corollaries can be easily obtained.

Corollary 5. *Under the condition that* $\widehat{X}_1$ *and* $\widehat{X}_2$ *are independent, if* $p_i^{yx} \neq \mathbf{0}$, *then* $P_L^{yy}(i,i) - P_H^{yy}(i,i) > 0$.

Corollary 6. *Under the condition that $\widehat{X}_1$ and $\widehat{X}_2$ are independent, if $P_{11}^{yx} \neq \mathbf{0}$ and $m = 1$, then $P_L^{yy} - P_H^{yy} > 0$.*

The proof is similar to that of Corollaries 2 and 3 and will be omitted here.

Corollaries 5 and 6 are the supplement of Theorem 4 for the single component case and scalar case, which also mean that increasing the independent common part accuracy can improve the fused result of the uncommon part.

4.4. The Level of Correlation's Impact on the Fused Uncommon Part. Assume that $p_i^{yx}(j)$ is the jth component of vector p_i^{yx} and it is the only nonzero component of p_i^{yx}:

$$\begin{aligned} p_i^{yx} &= \begin{bmatrix} 0 & \cdots & \rho_{i,j}\sigma_{x,j}\sigma_{y,i} & \cdots & 0 \end{bmatrix}^T, \\ p_i^{yx}(j) &= \rho_{i,j}\sigma_{x,j}\sigma_{y,i}, \end{aligned} \tag{19}$$

where $\rho_{i,j}$ is the correlation coefficient.

Theorem 7. *Under the condition that there is only one nonzero component in p_i^{yx}, if the absolute value of the correlation coefficient $|\rho_{i,j}|$ increases, the fused covariance $P^{yy}(i,i)$ will decrease.*

Proof. If there is only one nonzero component in p_i^{yx},

$$\begin{aligned} P^{yy}(i,i) &= P_{11}^{yy}(i,i) - P_{ZZ}^{-1}(j,j)\left(p_i^{yx}(j)\right)^2 \\ &= P_{11}^{yy}(i,i) - \rho_{i,j}^2 P_{ZZ}^{-1}(j,j)\,\sigma_{x,j}^2\sigma_{y,i}^2. \end{aligned} \tag{20}$$

Thus when $|\rho_{i,j}|$ increases, $P^{yy}(i,i)$ will decrease. □

It can be seen from Theorem 7 that under some condition, stronger cross-correlation can result in better fused result.

When $n = 1$, p_i^{yx} is a scalar, and the corresponding correlation coefficient is ρ_i. The following corollary can be obtained.

Corollary 8. *If $n = 1$, when $|\rho_i|$ increases, the fused results $P^{yy}(i,i)$ will decrease.*

The proof is the same as that of Theorem 7.

It can be seen from Corollary 8 that if the common part is a scalar, stronger cross-correlation can lead to better fused result.

5. Illustrative Examples

5.1. The Example for Improving the Fusion Result by the Existence of Cross-Correlation

Example 1. In target tracking applications, constant acceleration (CA) model based estimator can provide position, velocity, and acceleration estimation while constant velocity (CV) model based estimator can only provide position and velocity. The state vector of CA is $[x\ \ \dot{x}\ \ \ddot{x}]'$ and the state vector of CV is $[x\ \ \dot{x}]'$. When fusing the estimates from two models, position and velocity estimates are considered to be the common part and acceleration is considered to be the uncommon part. Assume the two estimators are independent.

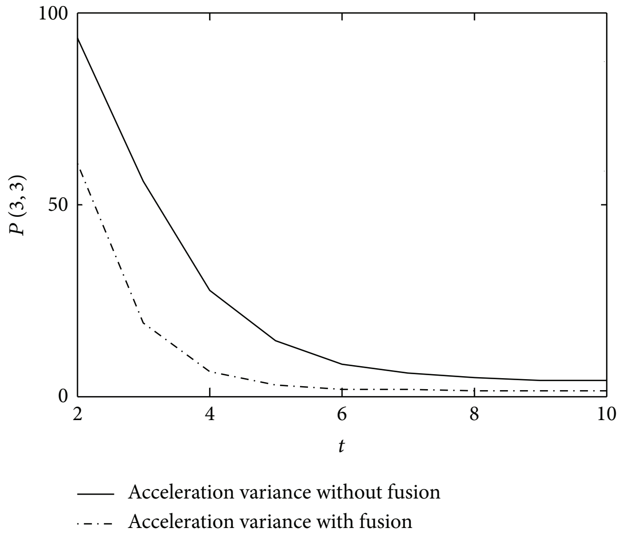


FIGURE 1: Acceleration fusion performance enhancement.

Assume there is a target moving with constant velocity motion. Two estimators are used to estimate the target's state. One estimator uses the CA model and the other one uses the CV model. The two estimators' initial covariance matrices are

$$P_{\text{CA}} = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix}, \qquad P_{\text{CV}} = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}. \tag{21}$$

Assume only the position can be observed by the sensors and the measurement noise variances are both $R = 100$. The sampling interval is $T = 1$. Both estimators' updated state covariance matrices are achieved by the Kalman filter. Because the CV model cannot provide estimation of the acceleration part, there are two ways to achieve the acceleration's estimation. One way is to use the CA model's acceleration estimation directly and the other way is to use the fusion result. Figure 1 shows the acceleration variance of the two ways. Acceleration estimate from the CA model is always correlated with the velocity and position estimates because of the state equation. The fusion results should benefit from the correlation and Figure 1 supports this analysis.

The following are some analyses for one step fusion. Assume the covariance matrices of the two models are

$$P_{\text{CA}} = \begin{bmatrix} 10 & 4 & 6 \\ 4 & 10 & 8 \\ 6 & 8 & 10 \end{bmatrix}, \qquad P_{\text{CV}} = \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix}. \tag{22}$$

The cross covariance vector between the common part and uncommon part is $P^{yx} = [6\ \ 8]$.

Using (10), the fusion result is

$$P = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 6.33 \end{bmatrix}. \tag{23}$$

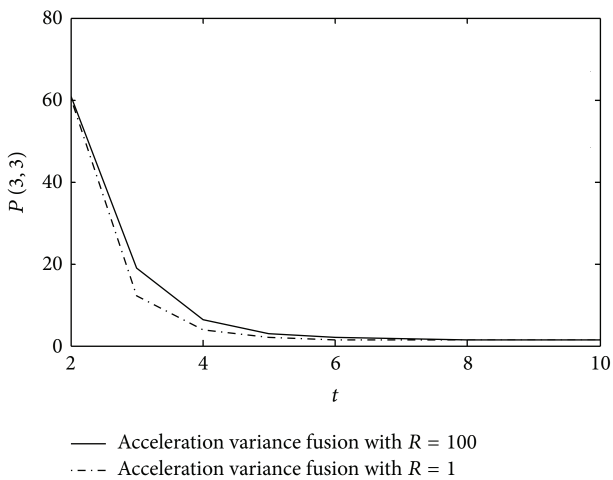


FIGURE 2: Acceleration fusion performance enhancement with more accurate estimator.

If there is no cross-correlation between acceleration and the other part,

$$P_{CA} = \begin{bmatrix} 10 & 4 & 0 \\ 4 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix}, \qquad P_{CV} = \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix}, \tag{24}$$

the fusion result

$$P = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 10 \end{bmatrix}. \tag{25}$$

It can be seen that without cross-correlation, the performance of the uncommon part cannot be improved.

However, with correlation, we have $6.33 < 10$, which means that the existence of cross-correlation can help improve the fusion result.

5.2. *The Examples for Increasing the Accuracy of the Independent Common Part to Improve the Fusion Result*

Example 2. The simulation setting is the same as in Example 1, which is a CA-CV fusion problem. In Example 2, CA model is the same as in Example 1, CV model's measurement is more accurate than in Example 1, and the measurement noise variance is $R = 1$.

Figure 2 shows the fusion results using two different CV estimators. It is known that more accurate measurement can lead to more accurate estimation. So the CV estimator in Example 2 is more accurate than the CV estimator in Example 1. Figure 2 supports the conclusion that more accurate independent common part estimator can lead to more accurate uncommon part's fusion result.

The following are some more analyses compared with Example 1. Here the covariance matrices of the two models are assumed to be

$$P_{CA} = \begin{bmatrix} 10 & 4 & 6 \\ 4 & 10 & 8 \\ 6 & 8 & 10 \end{bmatrix}, \qquad P_{CV} = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix}, \tag{26}$$

and the fusion result is

$$P = \begin{bmatrix} 0.91 & 0.36 & 0.55 \\ 0.36 & 0.91 & 0.73 \\ 0.55 & 0.73 & 3.33 \end{bmatrix}. \tag{27}$$

In Example 1, $P(3,3) = 6.33$. Here $P(3,3) = 3.33$.

Since $3.33 < 6.33$, it can be easily seen that more accurate common part estimation can lead to better fusion result.

Example 3. There are two radars which observe the same target. One is a Doppler radar, which can provide range and range rate measurements. The other is a regular radar, which can only provide range measurement. Doppler radar's range and range rate measurement errors are sometimes correlated. The two radars' measurement errors are independent of each other. The state vectors are $[r\ \dot{r}]'$ and r, respectively. The corresponding covariance matrices are

$$P_1 = \begin{bmatrix} \sigma_{r1}^2 & \rho\sigma_{r1}\sigma_{\dot{r}} \\ \rho\sigma_{r1}\sigma_{\dot{r}} & \sigma_{\dot{r}}^2 \end{bmatrix}, \qquad P_2 = \sigma_{r2}^2. \tag{28}$$

After fusion,

$$P_{\dot{r}} = \sigma_{\dot{r}}^2 \left(1 - \frac{\rho^2\sigma_{r1}^2}{\sigma_{r1}^2 + \sigma_{r2}^2} \right). \tag{29}$$

When σ_{r2}^2 decreases, $P_{\dot{r}}$ will also decrease.

When $\sigma_{r2}^2 \to 0$, $P_{\dot{r}} \to \sigma_{\dot{r}}^2(1-\rho^2)$.

Let

$$P_1 = \begin{bmatrix} 10 & 5 \\ 5 & 10 \end{bmatrix}. \tag{30}$$

When $P_2 = 10$, the covariance after fusion is

$$P = \begin{bmatrix} 5 & 2.5 \\ 2.5 & 8.75 \end{bmatrix}. \tag{31}$$

When $P_2 = 1$, the covariance after fusion is

$$P = \begin{bmatrix} 0.91 & 0.45 \\ 0.45 & 7.73 \end{bmatrix}. \tag{32}$$

Since $7.73 < 8.75$, it can be easily seen that more accurate common part estimation can lead to more accurate fusion result.

Figure 3 shows $P_{\dot{r}}$ as a function of σ_{r2}^2, which changes from 0 to 10. From the figure, it can be clearly seen that when improving the regular radar's range accuracy, the range rate accuracy will be improved.

5.3. *The Example for the Stronger Correlation to Improve the Fusion Result*

Example 4. The simulation setting is the same as in Example 3. The correlation coefficient is a variable. From (29), it can be seen that the bigger the $|\rho|$, the smaller the $P_{\dot{r}}$, which means stronger correlation can lead to better fusion result.

When $|\rho| \to 1$, $P_{\dot{r}} \to \sigma_{\dot{r}}^2\sigma_{r1}^2/(\sigma_{r1}^2 + \sigma_{r2}^2)$.

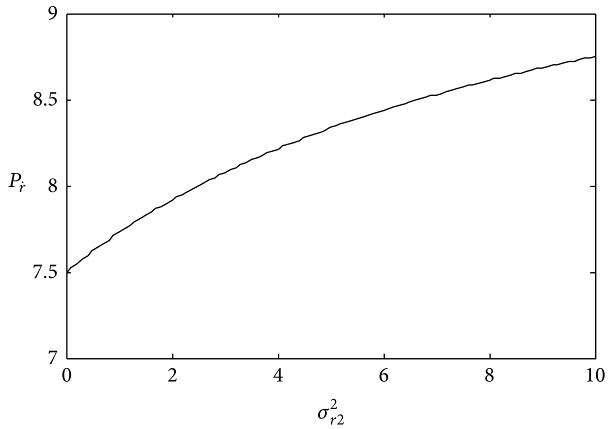


FIGURE 3: The relationship between regular radar's range accuracy and fused range rate accuracy.

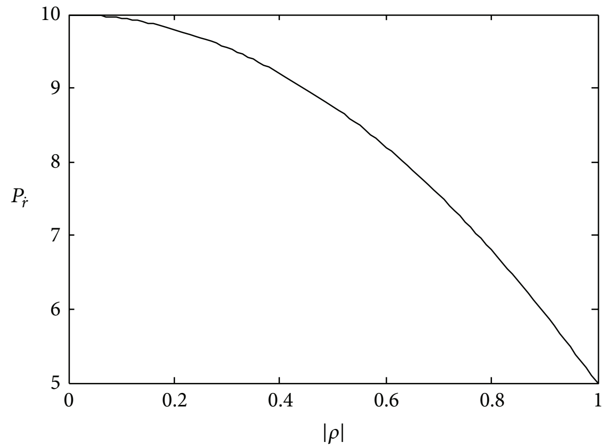


FIGURE 4: $P_{\dot{r}}$ as a function of $|\rho|$.

Let

$$P_1 = \begin{bmatrix} 10 & 5 \\ 5 & 10 \end{bmatrix}, \qquad P_2 = 10; \tag{33}$$

then

$$P = \begin{bmatrix} 5 & 2.5 \\ 2.5 & 8.75 \end{bmatrix}. \tag{34}$$

Let

$$P_1 = \begin{bmatrix} 10 & 9 \\ 9 & 10 \end{bmatrix}, \qquad P_2 = 10; \tag{35}$$

then

$$P = \begin{bmatrix} 5 & 4.5 \\ 4.5 & 5.95 \end{bmatrix}. \tag{36}$$

Since $5.95 < 8.75$, it can be easily seen that stronger correlation can lead to better fusion result.

Figure 4 shows $P_{\dot{r}}$ as a function of $|\rho|$, which changes from 0 to 1.

It can be seen that the stronger the correlation, the better the fused result.

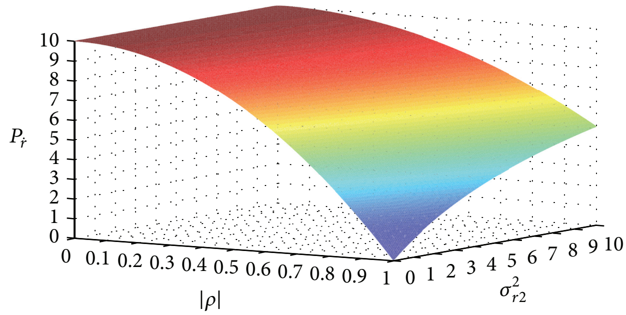


FIGURE 5: $P_{\dot{r}}$ as a function of σ^2_{r2} and $|\rho|$.

Example 5. Examples 3 and 4 are combined together. The range accuracy and correlation coefficient are changing simultaneously. From (29), when $|\rho|$ increases and σ^2_{r2} decreases, $P_{\dot{r}}$ will decrease.

And if $|\rho| \to 1$ and $\sigma^2_{r2} \to 0$, $P_{\dot{r}} \to 0$.

Figure 5 shows $P_{\dot{r}}$ as a function of σ^2_{r2} and $|\rho|$.

Figure 5 supports the conclusion that fusion result benefits from stronger correlation and more accurate common part.

6. Conclusion

Some sensors or estimators can provide higher dimensional measurement or estimation. But due to some constraints, other sensors or estimators can only provide partial measurement or estimation. To fuse such kind of data with different dimensions, a fusion algorithm based on LMMSE estimation is provided. To reveal the relationship between the common part and the uncommon part, the fusion performance is analyzed and the following four conclusions are obtained. (1) The fused common part is not affected by the uncommon part. (2) The fused uncommon part benefits from the common part through the cross-correlation. (3) The more accurate independent common part will result in better performance of the fused uncommon part. (4) In some cases, stronger cross-correlation will result in better performance of the fused uncommon part. The above conclusions are all supported by some target tracking examples.

Conflict of Interests

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

Acknowledgments

This research work was supported by the National Key Fundamental Research & Development Programs (973) of China (2013CB329405), Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61221063), Natural Science Foundation of China (61203221, 61174138, and 61473217), Ph.D. Programs Foundation of

Ministry of Education of China (20100201120036), and Fundamental Research Funds for the Central University.

References

[1] X. R. Li, Y. Zhu, J. Wang, and C. Han, "Optimal linear estimation fusion: part .I. Unified fusion rules," *IEEE Transactions on Information Theory*, vol. 49, no. 9, pp. 2192–2323, 2003.

[2] Y. Bar-Shalom, "Negative correlation and optimal tracking with doppler measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 3, pp. 1117–1120, 2001.

[3] X. H. Yuan, C. Z. Han, and Z. S. Duan, "Performance analysis and correlation selection with Doppler measurements," in *Proceedings of the 12th International Conference on Information Fusion*, pp. 373–379, Seattle, Wash, USA, July 2009.

[4] Y. Bar-Shalom and L. Campo, "The effect of the common process noise on the two-sensor fused-track covariance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 6, pp. 803–805, 1986.

[5] X. H. Yuan, C. Z. Han, and F. Lian, "Fusion performance analysis with the correlation," in *Proceedings of the 13th Conference on Information Fusion*, Edinburgh, UK, July 2010.

[6] X. R. Li and P. Zhang, "Optimal linear estimation fusion—part III: cross-correlation of local estimation errors," in *Proceedings of the 4th International Conference of Information Fusion*, pp. WeB1.11–WeB1.18, Montreal, Canada, 2001.

[7] S. J. Julier and J. K. Uhlmann, "Non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the American Control Conference*, pp. 2369–2373, June 1997.

[8] Y. C. Eldar, A. Beck, and M. Teboulle, "A minmax Chebyshev estimator for bounded error estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1388–1397, 2008.

[9] Y. Wang and X. R. Li, "Distributed estimation fusion with unavailable cross-correlation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 259–278, 2012.

[10] Y. Bar-shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*, Wiley, New York, NY, USA, 2001.

[11] X. R. Li, K. Zhang, J. Zhao, and Y. Zhu, "Optimal linear estimation fusion, part V: relationships," in *Proceeding of the 5th International Conference on Information Fusion (FUSION '02)*, vol. 1, pp. 497–504, Annapolis, Md, USA, July 2002.

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, chapter 7, Cambridge University Press, 1985.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 678976, 6 pages
http://dx.doi.org/10.1155/2014/678976

Research Article

Impact of Correlated Noises on Additive Dynamical Systems

Chujin Li[1] and Jinqiao Duan[1,2]

[1] *School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China*
[2] *Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA*

Correspondence should be addressed to Chujin Li; lichujin@hust.edu.cn

Received 14 April 2014; Accepted 4 August 2014; Published 3 September 2014

Academic Editor: Wuquan Li

Impact of correlated noises on dynamical systems is investigated by considering Fokker-Planck type equations under the fractional white noise measure, which correspond to stochastic differential equations driven by fractional Brownian motions with the Hurst parameter $H > 1/2$. Firstly, by constructing the fractional white noise framework, one small noise limit theorem is proved, which provides an estimate for the deviation of random solution orbits from the corresponding deterministic orbits. Secondly, numerical experiments are conducted to examine the probability density evolutions of two special dynamical systems, as the Hurst parameter H varies. Certain behaviors of the probability density functions are observed.

1. Introduction

Dynamical systems arising from financial, biological, physical, or geophysical sciences are often subject to random influences. These random influences may be modeled by various stochastic processes, such as Brownian motions, Lévy motions, or fractional Brownian motions. A fractional Brownian motion B_t^H, $t \geq 0$, in a probability space $(\Omega, \mathcal{F}, P)$, with Hurst parameter $H \in (0,1)$, is a continuous-time Gaussian process with mean zero, starting at zero and having the following correlation function:

$$\mathbb{E}\left[B_s^H B_t^H\right] = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right). \tag{1}$$

In particular, when $H = 1/2$ it is just the standard Brownian motion. The time derivative of a fractional Brownian motion, dB_t^H/dt, as a generalized stochastic process, has nonvanishing correlation [1, 2] and it is thus called a correlated noise or colored noise. In the special case of $H = 1/2$, this noise is uncorrelated and thus is called white noise [3]. Correlated noises appear in the modeling of some geophysical systems [4–6].

For systematic discussions about fractional Brownian motions and their stochastic calculus, we refer to [7–12] and the references therein. Fractional Brownian motions have stationary increments and are Hölder continuous with exponent less than H, but they are no longer semimartingales, even no longer Markovian. They possess some other significant properties such as long range dependence and self-similarity which result in wide applications in fields such as hydrology, telecommunications, and mathematical finance. During the last decade or so, several reasonable stochastic integrations with respect to fractional Brownian motions were developed. See, for example, Lin [13], Duncan et al. [14], Decreusefond and Üstunel [15], and the references mentioned therein. Stochastic differential equations (SDEs) driven by fractional Brownian motions also have been attracting more attention recently [1, 10, 16–18].

In this paper, we consider the following scalar stochastic differential equation (SDE):

$$dX_t = b\left(X_t\right)dt + \varepsilon dB_t^H, \qquad X_0 = x, \tag{2}$$

where the drift $b(\cdot)$ is a Lipschitz continuous function on R, $\varepsilon > 0$ is the noise intensity, B_t^H is a fractional Brownian motion with $H > 1/2$, and the initial state value ξ is assumed to be independent of the natural filtration of B_t^H. Since this system has a unique solution [17, 19], here we intend to understand some impact of correlated noises on this additive dynamical system as the Hurst parameter H varies.

This paper is organized as follows. In Section 2, we set up a fractional white noise analysis framework which makes correlated noises as functionals of standard white noises and prove a small noise limit theorem which implies the stochastic continuity of the system with respect to noise intensity. In Section 3, we show that the probability density function of X_t satisfies a Fokker-Planck type partial differential equation with respect to the fractional white noise measure. Then, we implement numerical experiments to examine the probability density evolutions as the Hurst parameter H varies. As to one linear system and one double-well system, certain behaviors of the probability density functions are observed.

2. Analysis Framework and Small Noise Limit

2.1. Analysis Framework. White noise framework is one natural and flexible stochastic analysis thoughtway, and fractional white noise analysis takes correlated noise as functionals of standard white noise. This approach has shown to be very effective in investigating distributions and path properties of stochastic processes. In the following, we describe the fractional white noise analysis framework.

Let $\mathcal{S}(R)$ be the Schwartz space of rapidly decreasing smooth functions on R and $\mathcal{S}'(R)$ the space of tempered distributions. And denote by $\langle\cdot,\cdot\rangle$ the dual pairing on $\mathcal{S}'(R)\times\mathcal{S}(R)$. For $1/2 < H < 1$, define

$$\varphi(s,t) = H(2H-1)|s-t|^{2H-2}, \quad s,t\in R;$$
$$c_H^2 = \frac{H(2H-1)}{B(H-1/2, 2-2H)}, \tag{3}$$

where $B(\cdot,\cdot)$ is beta function; $K_{\pm}(t) = c_H t_{\pm}^{H-3/2}$, $t_+ = t\vee 0$, $t_- = -(t\wedge 0)$.

Lemma 1. *For $f\in\mathcal{S}(R)$, let*

$$\Gamma_\varphi f(u) = (K_- * f)(u) = c_H\int_u^\infty (s-u)^{H-3/2} f(s)\,ds,$$
$$\Gamma_\varphi^* f(t) = (K_+ * f)(t) = c_H\int_{-\infty}^t (t-u)^{H-3/2} f(u)\,du. \tag{4}$$

Then, for $f,g\in\mathcal{S}(R)$,

$$(\Gamma_\varphi f, g)_{L^2(R)} = (f, \Gamma_\varphi^* g)_{L^2(R)}; \tag{5}$$

that is, Γ_φ^* is the dual map of Γ_φ.

Now we can only prove the linear map Γ_φ is continuous from $\mathcal{S}(R)$ to $L^2(R)$. Since Γ_φ is not continuous from $\mathcal{S}(R)$ to $\mathcal{S}(R)$ (even not a proper operator in $\mathcal{S}(R)$), we could not obtain a dual map from $\mathcal{S}'(R)$ to $\mathcal{S}'(R)$ by duality. By using Itô's regularization theorem, we construct a unique $\mathcal{S}'(R)$-valued random variable $T:\mathcal{S}'(R)\to\mathcal{S}'(R)$ such that

$$\langle T\omega,\xi\rangle = \langle\omega,\Gamma_\varphi\xi\rangle\ \mu-\text{a.e. }\omega, \tag{6}$$

which extends the map Γ_φ^* in view of (5).

Theorem 2. *Let $\mu_\varphi = \mu\circ T^{-1}$ be the image measure of μ induced by the map T. Then, for any $\xi\in\mathcal{S}(R)$, the distribution of $\langle\cdot,\xi\rangle$ under μ_φ is the same as $\langle\cdot,\Gamma_\varphi\xi\rangle$ under μ. In particular,*

$$B_t^H \equiv \langle\omega,\Gamma_\varphi\mathbf{1}_{[0,t]}\rangle, \quad t\geq 0 \tag{7}$$

is a fractional Brownian motion with Hurst constant H. Moreover,

$$B_t^H = c_H\left(H-\frac{1}{2}\right)^{-1}\int_{-\infty}^t\left[(t-u)^{H-1/2} - u_-^{H-1/2}\right]dB_u, \tag{8}$$

where $B_t(\omega)\equiv\langle\omega,\mathbf{1}_{[0,t]}\rangle$ is the standard Brownian motion. (See proof in [20].)

Let $\{\mathcal{F}_t, t\in R_+\}$ and $\{\mathcal{F}_t^H, t\in R_+\}$ be the filtrations generated by $\{B_t\}$ and $\{B_t^H\}$, respectively. Then, in view of (8), we have

(1) $\mathcal{F}_t\supset T^{-1}(\mathcal{F}_t^H)$, for all $t\in R_+$;

(2) for any $f\in L^\infty(\mu_\varphi)$, $\mathbb{E}_\mu[T_* f\mid\mathcal{F}_t] = T_*\mathbb{E}_{\mu_\varphi}[f\mid\mathcal{F}_t^H]$ a.s. $[\mu]$, where $(T_* f)(\omega) := f(T\omega)$. So, the filtrated probability space $(\mathcal{S}'(R),\mathcal{F}_t,\mu)$ is the extension of $(\mathcal{S}'(R),\mathcal{F}_t^H,\mu_\varphi)$. Thus the stochastic analysis with respect to measure μ_φ could be reduced to the standard white noise framework naturally. Therefore, we choose the standard white noise measure μ as the reference measure rather than μ_φ, and this treatment is more useful and more convenient for applications. For more details, we refer to [20] and the reference therein.

2.2. Small Noise Limit. Now, we consider the SDE (2) in fractional white noise framework

$$dX_t = b(X_t)\,dt + \varepsilon dB_t^H, \qquad X_0 = x. \tag{9}$$

And to investigate the impact of noise on deterministic dynamical system

$$\frac{d}{dt}x(t) = b(x(t)), \qquad x(0) = x, \tag{10}$$

which is solvable on any finite time interval $[0,T]$. We have the following result.

Theorem 3. *The solution X_t of (2) converges in probability to the solution $x(t)$ of (10) uniformly on any finite time interval $[0,T]$.*

Proof. Firstly, we rewrite the equation as

$$X_t - x(t) = \int_0^t\left[b(X_s) - b(x(s))\right]ds + \varepsilon B_t^H. \tag{11}$$

Then, by assuming the Lipschitz condition on $b(x)$ with Lipschitz constant $K > 0$, it follows from the Gronwall inequality that

$$\sup_{0\leq t\leq T}|X_t - x(t)| \leq \varepsilon e^{KT}\sup_{0\leq t\leq T}|B_t^H|. \tag{12}$$

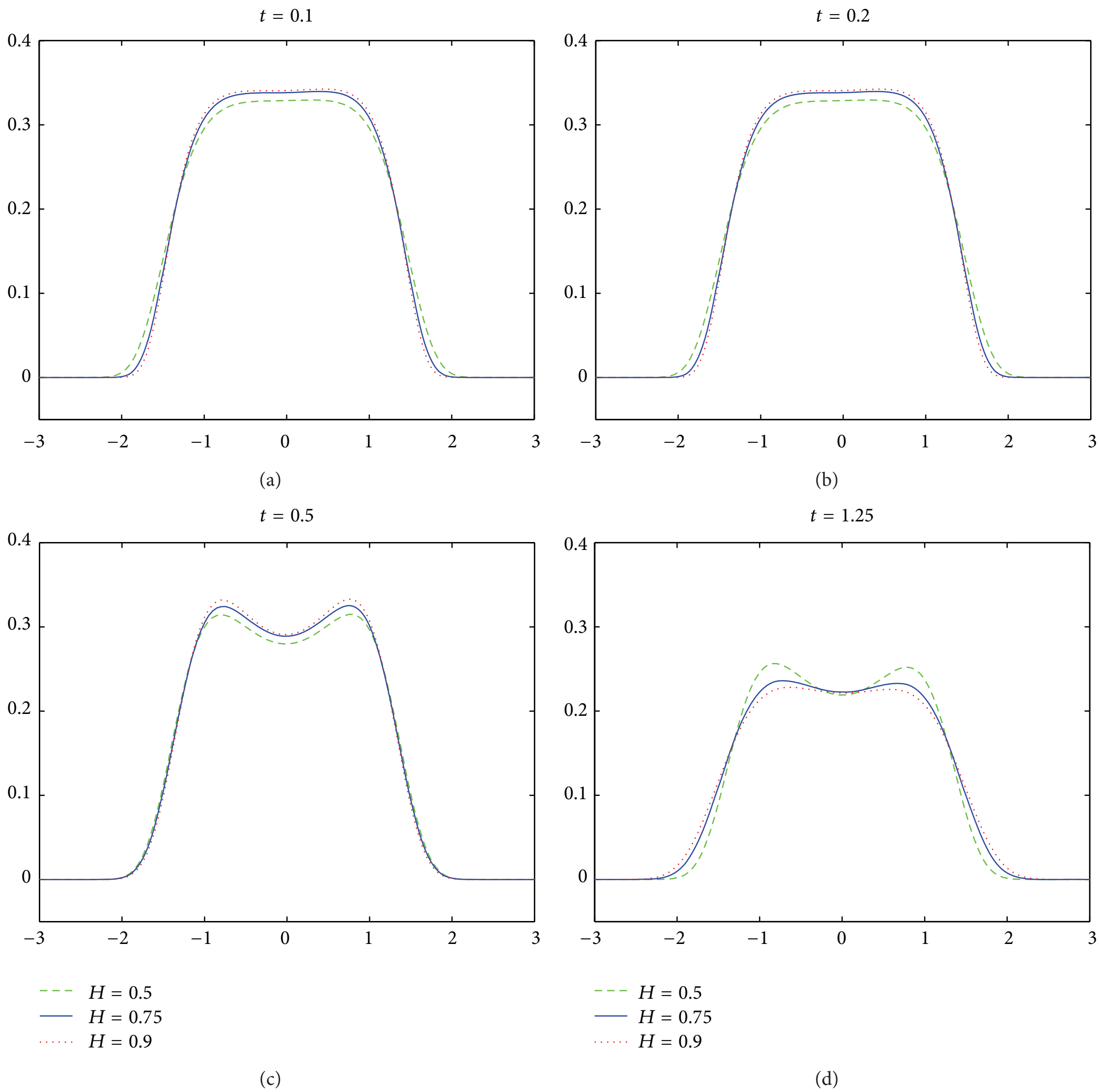


FIGURE 1: Plot of $p(x,t)$ with $b(x) = x - x^3$, at $t = 0.1, 0.2, 0.5, 1.25$.

Hence, for any small enough $\delta > 0$, we have

$$\begin{aligned} P\left\{\sup_{0\le t\le T}\left|X_t - x(t)\right| > \delta\right\} &\le P\left\{\sup_{0\le t\le T}\left|B_t^H\right| > \frac{\delta}{\varepsilon}e^{-KT}\right\} \\ &\le \frac{\varepsilon e^{KT}}{\delta}\mathbb{E}\sup_{0\le t\le T}\left|B_t^H\right| \\ &\le \frac{\varepsilon e^{KT}T^H}{\delta}\mathbb{E}\left|B_1^H\right|, \end{aligned} \tag{13}$$

which completes the proof when $\varepsilon \to 0$. In the final step, we have used the self-similarity of the fractional Brownian motion

$$\mathbb{E}\sup_{0\le t\le T}\left|B_t^H\right| \le T^H\mathbb{E}\left|B_1^H\right|. \tag{14}$$

□

This theorem provides an estimate for the deviation of random solution orbits from the corresponding deterministic orbits. Note that the expectation $\mathbb{E}$ in the above theorem corresponds to the fractional white noise measure. And, henceforth, we take all expectations $\mathbb{E}$ with respect to the fractional white noise measure (i.e., for simplicity, we omit the subscript μ mentioned above).

3. Probability Density Evolution

For SDE, such as (2), the probability density function of the solution X_t carries significant dynamical information. This is considered here by examining a fractional Fokker-Planck type equation. The key step in the derivation of this Fokker-Planck type equation is the application of Ito's formula for SDEs driven by fractional Brownian motion, under fractional

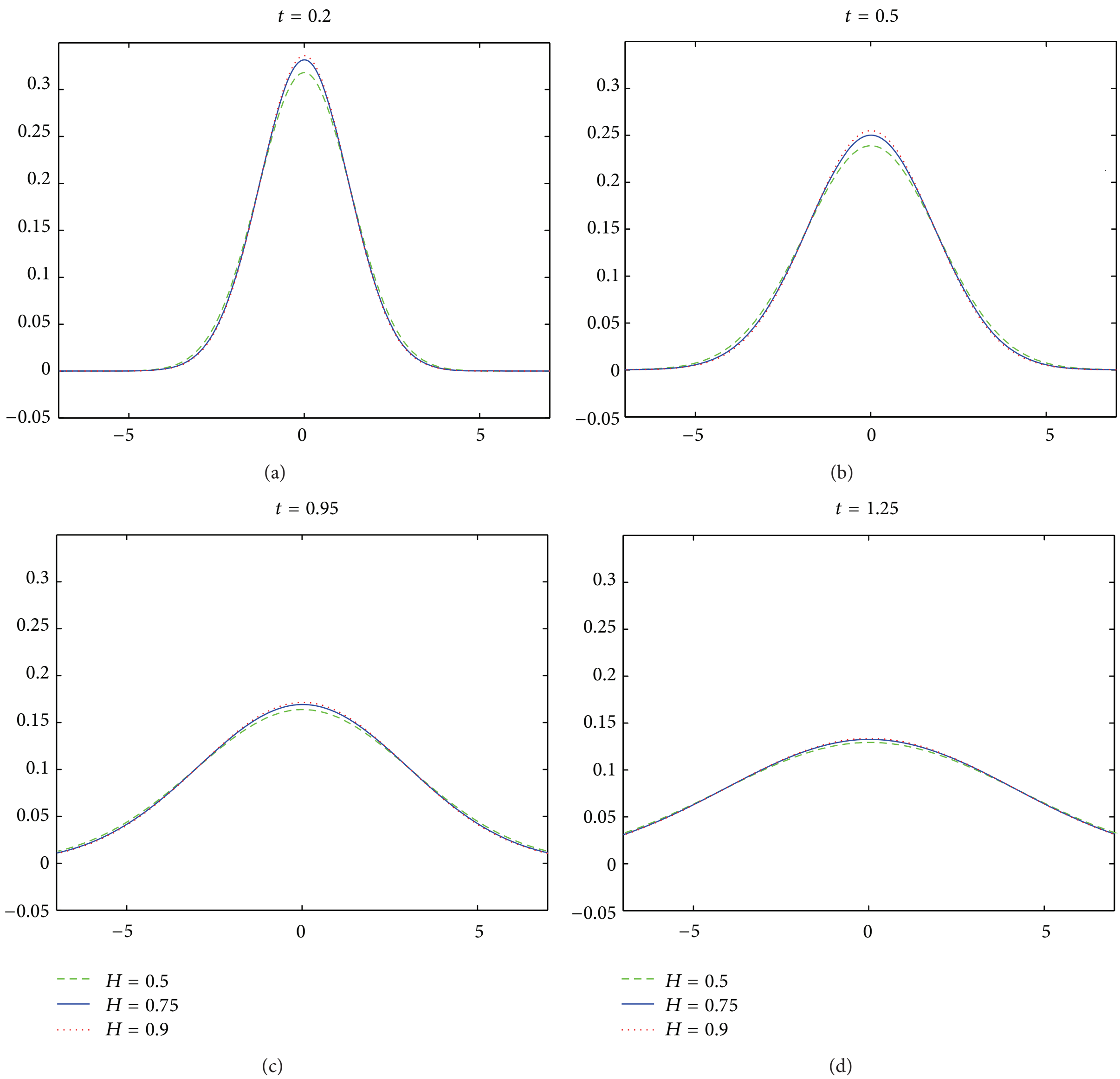


FIGURE 2: Plot of $p(x,t)$ with $b(x) = x$: $t = 0.2$, $t = 0.5$, $t = 0.95$, and $t = 1.25$.

white noise analysis framework [1, 10, 16, 20, 21]. We sketch the derivation here.

By Ito's formula [10], Theorem 6.3.6, for a second order differentiable function $h(\cdot)$ with compact support, we have

$$dh(X_t) = \left[b(X_t)\frac{\partial h}{\partial x}(X_t) + Ht^{2H-1}\varepsilon^2\frac{\partial^2 h}{\partial x^2}(X_t)\right] dt + \varepsilon\frac{\partial h}{\partial x}(X_t)\, dB_t^H. \tag{15}$$

Taking expectations on both sides yields

$$\mathbb{E}\left[\frac{dh(X_t)}{dt}\right] = \mathbb{E}\left[b(X_t)\frac{\partial h}{\partial x}(X_t)\right] + Ht^{2H-1}\varepsilon^2\mathbb{E}\left[\frac{\partial^2 h}{\partial x^2}(X_t)\right]. \tag{16}$$

Let $p = p(x,t)$ be the probability density function of the solution X_t of the system (2). Recall that $\mathbb{E}[h(X_t)] = \int_{\mathbb{R}} h(x)p(x,t)dx$; by integration by parts and $p = 0$ at $x = \pm\infty$, we obtain

$$\int_{\mathbb{R}} h(x)\left[\frac{\partial p}{\partial t} + \frac{b(x)\,p}{\partial x} - \varepsilon^2 Ht^{2H-1}\frac{\partial^2 p}{\partial x^2}\right] dx = 0; \tag{17}$$

that is,

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial\left[b(x)\,p(x,t)\right]}{\partial x} + \varepsilon^2 Ht^{2H-1}\frac{\partial^2 p(x,t)}{\partial x^2}. \tag{18}$$

In the following, we numerically simulate this partial differential equation for two special cases: $b(x) = x - x^3$ and $b(x) = x$, with finite noise intensity (for simplicity we take $\varepsilon = 1$). Through these two special cases, we expect to illustrate the impact of correlated noises on additive dynamical systems as the Hurst parameter H varies.

Here, we perform the popular Crank-Nicolson scheme in Matlab for (17) with zero boundary values;, the grid size is 0.05, total grid points are 801, and the time step size is 0.01. And the initial probability density function is taken to be standard normal; that is, $p(x,0) = (1/\sqrt{2\pi})e^{-x^2/2}$.

Since the system is tridiagonal, we could solve it using Thomas Algorithm efficiently. Moreover, for other initial conditions and other drift coefficients, for instance, the initial uniform distribution or $b(x) = x - x^2$, this method also applies smoothly.

3.1. Numerical Simulation: $b(x) = x - x^3$. We first simulate the dynamical evolutions of the probability density function $p(x,t)$ for the corresponding stochastic differential equation (2) with the double-well drift $b(x) = x - x^3$, for various values of $H > 1/2$. The double-well dynamics is a rich and typical model for understanding numerous physical or geophysical systems [22, 23], focusing on the maxima (minima), symmetry, kurtosis, and so forth.

As observed in Figure 1, the probability density function $p(x,t)$ evolves from the unimodal (one peak) to the flat top and then to the bimodal (two peaks) shape for various Hurst parameter values H, as time t increases. Simultaneously, the effect of Hurst parameter H on the dynamics is significant. As H value increases, the plateau for $p(x,t)$ becomes lower when time exceeds $t = 0.5$.

3.2. Numerical Simulation: $b(x) = x$. Now, for comparison we investigate the dynamical evolutions of the probability density function $p(x,t)$ of the corresponding stochastic differential equation (2) with the linear drift $b(x) = x$, which is a rich toy example for understanding dynamical systems.

Also as observed in Figure 2, at given time instants, $p(x,t)$'s peak becomes higher as H increases. This illustrates the significant and distinguishing influence of Hurst parameter H on the dynamics when time t evolves. The bigger H makes the solution X_t of (2) has more centralized value, but the long time effect shows that the values of the solution X_t distribute more scatteredly.

Conflict of Interests

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

Acknowledgment

Chujin Li acknowledges the support of Self-Renovation Project HUST2013QN171.

References

[1] C. Bender, "An Itô formula for generalized functionals of a fractional Brownian motion with arbitrary Hurst parameter," *Stochastic Processes and their Applications*, vol. 104, no. 1, pp. 81–106, 2003.

[2] J. Duan, C. Li, and X. Wang, "Modeling colored noise by fractional Brownian motion," *Interdisciplinary Mathematical Sciences*, vol. 8, pp. 119–130, 2009.

[3] L. Arnold, *Stochastic Differential Equations*, John Wiley & Sons, New York, NY, USA, 1974.

[4] A. Du and J. Duan, "A stochastic approach for parameterizing unresolved scales in a system with memory," *Journal of Algorithms & Computational Technology*, vol. 3, no. 3, pp. 393–405, 2009.

[5] J. Duan, "Stochastic modeling of unresolved scales in complex systems," *Frontiers of Mathematics in China*, vol. 4, no. 3, pp. 425–436, 2009.

[6] B. Chen and J. Duan, "Stochastic quantification of missing mechanisms in dynamical systems," *Interdisciplinary Mathematical Sciences*, vol. 8, pp. 67–76, 2009.

[7] D. Nualart, "Stochastic calculus with respect to the fractional Brownian motion and applications," *Contemporary Mathematics*, vol. 336, pp. 3–39, 2003.

[8] E. Alòs and D. Nualart, "Stochastic integration with respect to the fractional Brownian motion," *Stochastics and Stochastics Reports*, vol. 75, no. 3, pp. 129–152, 2003.

[9] A. Chronopoulou and F. Viens, "Hurst index estimation for self-similar processes with long-memory," *Interdisciplinary Mathematical Sciences*, vol. 8, pp. 91–118, 2009.

[10] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang, *Stochastic Calculus for Fractional Brownian Motion and Applications*, Springer, New York, NY, USA, 2008.

[11] Y. S. Mishura, *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, Springer, New York, NY, USA, 2008.

[12] C. A. Tudor and F. G. Viens, "Variations and estimators for self-similarity parameters via Malliavin calculus," *The Annals of Probability*, vol. 37, no. 6, pp. 2093–2134, 2009.

[13] S. J. Lin, "Stochastic analysis of fractional Brownian motions," *Stochastics and Stochastics Reports*, vol. 55, no. 1-2, pp. 121–140, 1995.

[14] T. E. Duncan, Y. Hu, and B. Pasik-Duncan, "Stochastic calculus for fractional Brownian motion. I. Theory," *SIAM Journal on Control and Optimization*, vol. 38, no. 2, pp. 582–612, 2000.

[15] L. Decreusefond and A. S. Üstunel, "Stochastic analysis of the fractional Brownian motion," *Potential Analysis*, vol. 10, no. 2, pp. 177–214, 1999.

[16] F. Baudoin and L. Coutin, "Operators associated with a stochastic differential equation driven by fractional Brownian motions," *Stochastic Processes and Their Applications*, vol. 117, no. 5, pp. 550–574, 2007.

[17] M. A. Diop and Y. Ouknine, "A linear stochastic differential equation driven by a fractional Brownian motion with Hurst parameter > 1/2," *Statistics & Probability Letters*, vol. 81, no. 8, pp. 1013–1020, 2011.

[18] B. Saussereau, "A stability result for stochastic differential equations driven by fractional Brownian motions," *International Journal of Stochastic Analysis*, vol. 2012, Article ID 281474, 13 pages, 2012.

[19] L. Denis, M. Erraoui, and Y. Ouknine, "Existence and uniqueness for solutions of one dimensional SDE's driven by an additive fractional noise," *Stochastics and Stochastics Reports*, vol. 76, no. 5, pp. 409–427, 2004.

[20] Z. Huang and C. Li, "On fractional stable processes and sheets: white noise approach," *Journal of Mathematical Analysis and Applications*, vol. 325, no. 1, pp. 624–635, 2007.

[21] G. Ünal, "Fokker-Planck-Kolmogorov equation for fBM: derivation and analytical solution," in *Proceedings of the 12th Regional Conference*, pp. 53–60, Islamabad, Pakistan, 2006.

[22] D. Farrelly and J. E. Howard, "Double-well dynamics of two ions in the Paul and Penning traps," *Physical Review A*, vol. 49, no. 2, pp. 1494–1497, 1994.

[23] E. Kierig, U. Schnorrberger, A. Schietinger, J. Tomkovic, and M. K. Oberthaler, "Single-particle tunneling in strongly driven double-well potentials," *Physical Review Letters*, vol. 100, no. 19, Article ID 190405, 2008.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 713738, 6 pages
http://dx.doi.org/10.1155/2014/713738

Research Article

Stochastic Dominance under the Nonlinear Expected Utilities

Xinling Xiao[1,2]

[1] *Shandong University Qilu Securities Institute for Financial Studies and School of Mathematics, Jinan 250100, China*
[2] *School of Mathematical Sciences, Shandong Normal University, Jinan 250014, China*

Correspondence should be addressed to Xinling Xiao; xinlingxiao@mail.sdu.edu.cn

Received 26 March 2014; Revised 28 May 2014; Accepted 19 June 2014; Published 31 August 2014

Academic Editor: Guangchen Wang

In 1947, von Neumann and Morgenstern introduced the well-known expected utility and the related axiomatic system (see von Neumann and Morgenstern (1953)). It is widely used in economics, for example, financial economics. But the well-known Allais paradox (see Allais (1979)) shows that the linear expected utility has some limitations sometimes. Because of this, Peng proposed a concept of nonlinear expected utility (see Peng (2005)). In this paper we propose a concept of stochastic dominance under the nonlinear expected utilities. We give sufficient conditions on which a random choice X stochastically dominates a random choice Y under the nonlinear expected utilities. We also provide sufficient conditions on which a random choice X strictly stochastically dominates a random choice Y under the sublinear expected utilities.

1. Introduction

In [1], von Neumann and Morgenstern introduced the well-known expected utility and the related axiomatic system. It is widely used in economics, for example, financial economics. They exhibited four relatively modest axioms of "rationality" such that any agent satisfying the axioms has a utility function. They claimed that $U(\cdot)$ can be characterized by $U(X) = E[u(X)]$. That is to say they proved that an agent is (VNM-) rational if and only if there exists a real-valued function $u(\cdot)$ defined on possible outcomes such that every preference of the agent is characterized by maximizing the expected value of $u(\cdot)$, which can then be defined as the agent's VNM-utility. Here $u(\cdot) : R \to R$ is a continuous and strictly increasing function, and $E[\cdot]$ is the linear expectation in some probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

However, some real world utilities cannot be represented by this expected utility. A famous counterexample is the well-known Allais paradox (see [2]). Allais paradox shows that linear expected utility has some limitations sometimes.

In [3], Peng developed nonlinear expectation and sublinear expectation theory. G-expectation is a kind of special sublinear expectation. More details can be found in [4–7]. In [8–11], G-expectation is used in financial economics.

In [12], Peng developed a nonlinear type of von Neumann-Morgenstern representation theorem to utilities. He proved that there exists a nonlinear expected utility $U(\cdot)$, such that an agent Â prefers a random choice X than Y which is formulated by $U(X) > U(Y)$.

But nonlinear expected utility can only describe an agent's preference; how to describe a group of agents' preference? In this paper we consider the question raised upward; to this end we define a corresponding concept of stochastic dominance under the nonlinear expected utilities.

The rest of this paper is organized as follows. In Section 2, we give some basic notions and results of nonlinear expectations and nonlinear expected utilities. In Section 3, we give the main results and the proofs.

2. Nonlinear Expectations and Nonlinear Expected Utilities

In this section we shall give some results of nonlinear expectations and nonlinear expected utilities.

2.1. Nonlinear Expectations. We present some preliminaries in the theory of nonlinear expectations and sublinear expectations. The following definitions and properties can be found in [3].

Let Ω be a given set and let $\mathscr{H}$ be a linear space of real valued functions defined on Ω satisfying the following: if $X_i \in \mathscr{H}$, $i = 1, 2, \ldots, n$, then $\varphi(X_1, X_2, \ldots, X_n) \in \mathscr{H}$, for all $\varphi \in \mathscr{C}_{l,\mathrm{lip}}(R)$, where $\mathscr{C}_{l,\mathrm{lip}}(R)$ is the space of all real continuous functions defined on R such that

$$\left|\varphi(\mathbf{x}) - \varphi(\mathbf{y})\right| \le \mathscr{C}\left(1 + \mathbf{x}^k + \mathbf{y}^k\right)\left|\mathbf{x} - \mathbf{y}\right|, \quad \forall \mathbf{x}, \mathbf{y} \in R, \ k \text{ depends on } \varphi. \tag{1}$$

Definition 1. $\mathbb{E} : \mathscr{H} \to R$ is said to be a nonlinear expectation defined on $\mathscr{H}$ if it satisfies the following.

(i) Monotonicity:

$$\mathbb{E}(X) \ge \mathbb{E}(Y), \quad \text{if } X \ge Y. \tag{2}$$

(ii) Constant Preserving:

$$\mathbb{E}(c) = c, \quad \text{for } c \in R. \tag{3}$$

A nonlinear expectation is called sublinear expectation if it also satisfies the following.

(iii) Subadditivity: for each $X, Y \in \mathscr{H}$,

$$\mathbb{E}(X + Y) \le \mathbb{E}(X) + \mathbb{E}(Y). \tag{4}$$

(iv) Positive homogeneity:

$$\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X), \quad \text{for } \lambda \ge 0. \tag{5}$$

The triple $(\Omega, \mathscr{H}, \mathbb{E})$ is called nonlinear expectation space and sublinear expectation space correspondingly.

Definition 2. Let $\mathbb{E}_1$ and $\mathbb{E}_2$ be two nonlinear expectations defined on $(\Omega, \mathscr{H})$; $\mathbb{E}_1$ is said to be dominated by $\mathbb{E}_2$ if

$$\mathbb{E}_1(X) - \mathbb{E}_1(Y) \le \mathbb{E}_2(X - Y), \quad \text{for } X, Y \in \mathscr{H}. \tag{6}$$

Remark 3. From (iii), a sublinear expectation is dominated by itself. In many situations, (iii) is also called the property of self-domination. It is easy to conclude that in a sublinear expectation space $(\Omega, \mathscr{H}, \mathbb{E})$, $-\mathbb{E}(-X) \le \mathbb{E}(X)$, for $X \in \mathscr{H}$. If $-\mathbb{E}(-X) = \mathbb{E}(X)$, we say X has no mean uncertainty.

Theorem 4 (Represent theorem). *Let $\mathbb{E}$ be a functional defined on a linear space $\mathscr{H}$ satisfying subadditivity and positive homogeneity. Then there exists a family of linear functionals defined on $\mathscr{H}$ such that*

$$\mathbb{E}(X) = \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}(X), \quad \text{for } X \in \mathscr{H} \tag{7}$$

and, for each $X \in \mathscr{H}$, there exists $\mathbb{P}_X \in \mathscr{P}$ such that $\mathbb{E}(X) = \mathbb{E}_{\mathbb{P}_X}(X)$.

Furthermore, if $\mathbb{E}$ is a sublinear expectation, then the corresponding $\mathbb{E}_{\mathbb{P}_X}$ is a linear expectation.

According to the represent theorem, if $\mathbb{E}$ is a sublinear expectation, we have

$$\mathbb{E}(X) = \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}(X), \quad \text{for } X \in \mathscr{H}. \tag{8}$$

Suppose $(\Omega, \mathscr{F})$ is a measurable space, for such $\mathscr{P}$, we can define an upper probability

$$\mathbb{V}(A) = \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{P}(A), \quad A \in \mathscr{F} \tag{9}$$

and a lower probability

$$\nu(A) = \inf_{\mathbb{P} \in \mathscr{P}} \mathbb{P}(A), \quad A \in \mathscr{F}. \tag{10}$$

Obviously $\mathbb{V}$ and ν are conjugated to each other; that is,

$$\mathbb{V}(A) + \nu(A^c) = 1, \tag{11}$$

where A^c is the complementary set of A.

Definition 5. A set A is polar if $\mathbb{V}(A) = 0$. A property holds quasisurely $(q.s)$. if it holds outside a polar set.

Definition 6. Let X be a given random variable on a nonlinear expectation space $(\Omega, \mathscr{H}, \mathbb{E})$. One defines a functional on $\mathscr{C}_{l,\mathrm{lip}}(R)$ by

$$\mathbb{F}_X[\varphi] := \mathbb{E}[\varphi(X)] : \varphi \in \mathscr{C}_{l,\mathrm{lip}}(R) \longrightarrow R. \tag{12}$$

$\mathbb{F}_X$ is called the distribution of X under $\mathbb{E}$.

Definition 7. Let X_1 and X_2 be two random variables defined on nonlinear expectation spaces $(\Omega_1, \mathscr{H}_1, \mathbb{E}_1)$ and $(\Omega_2, \mathscr{H}_2, \mathbb{E}_2)$, respectively. They are called identically distributed, denoted by $X_1 \overset{d}{=} X_2$, if

$$\mathbb{E}_1[\varphi(X_1)] = \mathbb{E}_2[\varphi(X_2)], \quad \text{for } \varphi \in \mathscr{C}_{l,\mathrm{lip}}(R). \tag{13}$$

It is clear that $X_1 \overset{d}{=} X_2$ if and only if their distributions coincide. One says that the distribution of X_1 is stronger than that of X_2 if

$$\mathbb{E}_1[\varphi(X_1)] \ge \mathbb{E}_2[\varphi(X_2)], \quad \text{for each } \varphi \in \mathscr{C}_{l,\mathrm{lip}}(R). \tag{14}$$

2.2. Nonlinear Expected Utilities. The following definitions and properties can be found in [12]. Let $\mathbb{E}$ be a self-dominated nonlinear expectation defined on $\mathscr{H}$. Define a quasinorm $\|X\|_{*\infty} := \inf_{\omega \in \Omega}\{c \in R;\ c \ge |X| \text{ in } \mathscr{H}\}$. A utility functional of an agent $\hat{A}$ is a real functional $U : \mathscr{H} \to R$. This functional satisfies the following obvious axioms:

(u1) monotonicity: if $X \ge Y$ in $\mathscr{H}$, then $\mathbb{U}(X) \ge \mathbb{U}(Y)$, and if $X \ge Y$ and $\|X - Y\|_* > 0$, then $\mathbb{U}(X) > \mathbb{U}(Y)$;

(u2) continuity: if $\|X_i - X\|_{*\infty} \to 0$, then $\mathbb{U}(X_i) \to \mathbb{U}(X)$.

Then we have the following nonlinear expected utility theorem which generalized the well-known von Neuman-Morgenstern's axiom on expected utility.

Proposition 8. *Let $\mathbb{E}[\cdot]$ be a strictly monotonic expectation satisfying (i) and (ii) in Definition 1. One assumes that $\mathbb{E}[\cdot]$ is continuous in $\mathscr{H}$ and let $u(\cdot)$ be a continuous and strictly increasing function $u(\cdot) : R \rightarrow R$. Then the functional $U(\cdot)$ defined by*

$$U(X) := \mathbb{E}[u(X)] \tag{15}$$

is a utility functional satisfying (u1) and (u2).

Conversely, for each given utility $U(\cdot)$ satisfying (u1) and (u2), there exist a strict monotonic nonlinear expectation $\mathbb{E}[\cdot]$ and a continuous and strictly increasing function $u(\cdot) : R \rightarrow R$ such that (15) holds.

3. Stochastic Dominance under the Nonlinear Expected Utilities

Using nonlinear expected utility to determine the advantages between two random choices is only for a single economic actor. Here comes a problem: can we raise the same question to a group of economic actors? If we still discuss it by using nonlinear expected utility, this means asking the same question to a class of expected utility functions.

In mathematics, it can form such a problem: suppose $\mathscr{H}$ is a collection of random variables. $\mathscr{U}$ is a class of strictly increasing and continuously differentiable functions, which represents the collection of all the utility functions of an investor group. Define a partial ordering $\succeq$ in $\mathscr{H}$: for any $X, Y \in \mathscr{H}$,

$$X \succeq Y \Longleftrightarrow \forall u \in \mathscr{U}, \quad \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]. \tag{16}$$

Here $\mathbb{E}[\cdot]$ is nonlinear expectation, X, Y can be regarded as two risky securities, and $u \in \mathscr{U}$ can be an investor's expected utility function. Thus, this definition of the partial ordering means that all members of the investor group think the former is better than the latter. Here it is important to note that, in general, this is a partially ordering rather than a complete ordering. That is to say, for some pairs of risky securities, neither one stochastically dominates the other, and yet they cannot be said to be equal. At the same time, it is just investor group's preferences characterized by the expected utilities functions theory.

Definition 9. The above-mentioned partial ordering is called stochastic dominance under the nonlinear expected utility.

Remark 10. Stochastic dominance is a form of stochastic ordering. The term is used in decision theory and decision analysis to refer to situations where one random choice can be ranked as superior to another. It is based on preferences regarding outcomes. In linear expected utility, there are first-order stochastic dominance and second-order stochastic dominance and so on. For more results, see [13–16].

Definition 11. In Definition 9, if for any $X, Y \in \mathscr{H}$,

$$X \succ Y \Longleftrightarrow \forall u \in \mathscr{U}, \quad \mathbb{E}[u(X)] > \mathbb{E}[u(Y)], \tag{17}$$

then the partial ordering is called strictly stochastic dominance under the nonlinear expected utility.

Next, we give the main results of this paper.

Theorem 12. *Let $(\Omega, \mathscr{H}, \mathbb{E})$ be a nonlinear expectation space, $X, Y \in \mathscr{H}$, and $\mathscr{U}$ a class of strictly increasing and continuously differentiable functions. If any of the following conditions is satisfied:*

(1) $X \geq Y$,

(2) the distribution of X is stronger than Y, namely,

$$\mathbb{E}[\varphi(X)] \geq \mathbb{E}[\varphi(Y)], \quad \text{for each } \varphi \in \mathscr{C}_{l,\mathrm{lip}}(R), \tag{18}$$

then X stochastically dominates Y, that is,

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)], \quad \forall u \in \mathscr{U}. \tag{19}$$

Proof. (1) If $X \geq Y$, then

$$\forall u \in \mathscr{U}, \quad \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)] \tag{20}$$

is easily concluded by the fact that $u \in \mathscr{U}$ is strictly increasing and the monotonicity of $\mathbb{E}[\cdot]$.

(2) First, notice that an everywhere differentiable function $g(\cdot)$ which is a lipschitz continuous with $k = \sup |g'(x)|$ is equivalent to the fact that $g(\cdot)$ has bounded first derivative. In particular, any continuously differentiable function is locally lipschitz, as continuous functions are locally bounded so its gradient is locally bounded as well. It means that for all $u \in \mathscr{U}, u \in \mathscr{C}_{l,\mathrm{lip}}(R)$. So if the distribution of X is stronger than that of Y, we have

$$\mathbb{E}[\varphi(X)] \geq \mathbb{E}[\varphi(Y)], \quad \text{for each } \varphi \in \mathscr{C}_{l,\mathrm{lip}}(R). \tag{21}$$

Because for all $u \in \mathscr{U}, u \in \mathscr{C}_{l,\mathrm{lip}}(R)$; then we can have

$$\forall u \in \mathscr{U}, \quad \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]. \tag{22}$$

So the result holds by Definition 9. □

Remark 13. (α) The above conclusion (2) gives sufficient condition on which a random choice X stochastically dominates a random choice Y under the nonlinear expected utilities. It is a general extension of the first-order stochastic dominance under the linear expectation utility.

(β) The above conclusion (1) is very intuitive. Next, we give an example which is not intuitive.

Example 14. Suppose

$$\Omega = \{\omega_1, \omega_2\}, \tag{23}$$

and we have two probabilities $\{2/3, 1/3\}$ and $\{2/5, 3/5\}$ denoted by $\mathbb{P}$ and $\mathbb{Q}$, respectively, where

$$\begin{aligned} \mathbb{P}(\{\omega_1\}) &= \frac{2}{3}, \\ \mathbb{P}(\{\omega_2\}) &= \frac{1}{3}, \\ \mathbb{Q}(\{\omega_1\}) &= \frac{2}{5}, \\ \mathbb{Q}(\{\omega_2\}) &= \frac{3}{5}. \end{aligned} \tag{24}$$

We assume $\mathscr{H}$ is a collection of random variables, and $\mathscr{U}$ is a class of strictly increasing and continuously differentiable functions. Take the nonlinear expectation utility like the following:

$$\begin{aligned}\mathbb{E}[u(\xi)] = \frac{3}{4}\max\{E_{\mathbb{P}}[u(\xi)], E_{\mathbb{Q}}[u(\xi)]\} \\ + \frac{1}{4}\min\{E_{\mathbb{P}}[u(\xi)], E_{\mathbb{Q}}[u(\xi)]\}, \\ u \in \mathscr{U}, \quad \xi \in \mathscr{H}.\end{aligned} \tag{25}$$

We set

$$\begin{aligned}X(\omega_1) &= 1, \\ X(\omega_2) &= 0, \\ Y(\omega_1) &= 0, \\ Y(\omega_2) &= 1.\end{aligned} \tag{26}$$

For all $u \in \mathscr{U}$, we can calculate the following results:

$$\begin{aligned}E_{\mathbb{P}}[u(X)] &= \frac{2}{3}u(1) + \frac{1}{3}u(0), \\ E_{\mathbb{Q}}[u(X)] &= \frac{2}{5}u(1) + \frac{3}{5}u(0), \\ \mathbb{E}[u(X)] &= \frac{6}{15}u(0) + \frac{9}{15}u(1), \\ E_{\mathbb{P}}[u(Y)] &= \frac{2}{3}u(0) + \frac{1}{3}u(1), \\ E_{\mathbb{Q}}[u(X)] &= \frac{2}{5}u(0) + \frac{3}{5}u(1), \\ \mathbb{E}[u(Y)] &= \frac{7}{15}u(0) + \frac{8}{15}u(1).\end{aligned} \tag{27}$$

Since $u \in \mathscr{U}$ is strictly increasing, then $u(1) > u(0)$; so

$$\begin{aligned}\mathbb{E}[u(X)] - \mathbb{E}[u(Y)] = \frac{1}{15}u(1) - \frac{1}{15}u(0) > 0, \\ u \in \mathscr{U}.\end{aligned} \tag{28}$$

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)], \quad u \in \mathscr{U}.$$

Hence we can say that X strictly stochastic dominates Y.

It is easy to see that neither $X \geq Y$ nor $Y \geq X$ in a whole. We can check that X strictly stochastic dominates Y. This is not the intuitive way; this implies that stochastic dominance by the nonlinear expected utilities is meaningful.

When $\mathbb{E}[\cdot]$ is a sublinear expectation, Theorem 12 is still valid. Furthermore, we can also have the following theorem.

Theorem 15. *Let $(\Omega, \mathscr{H}, \mathbb{E})$ be a sublinear expectation space, $X, Y \in \mathscr{H}$, and $\mathscr{U}$ a class of strictly increasing and continuously differentiable functions. If $X \geq Y$ q.s., then X stochastically dominates Y, that is,*

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)], \quad \forall u \in \mathscr{U}. \tag{29}$$

Proof. We claim that if $X \geq Y$ *q.s.*, then $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

This is because we can get $\mathbb{V}(X < Y) = 0$ by $X \geq Y$ *q.s.*, which means $\mathbb{V}(X < Y) = 0$, $\nu(X \geq Y) = 1$; namely,

$$\nu(X \geq Y) = \inf_{\mathbb{P} \in \mathscr{P}} \mathbb{P}(X \geq Y) = 1. \tag{30}$$

Then we can get

$$\forall \mathbb{P} \in \mathscr{P}, \quad E_{\mathbb{P}}[X] \geq E_{\mathbb{P}}[Y], \tag{31}$$

so

$$\sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}(X) \geq \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}(Y). \tag{32}$$

According to the represent theorem, we have $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

Since $X \geq Y$ *q.s.* and $u \in \mathscr{U}$ is strictly increasing, $u(X) \geq u(Y)$ *q.s.* is available by the same procedure as above. So we can obtain

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)], \quad \forall u \in \mathscr{U} \tag{33}$$

by the above conclusion. □

Next, we shall give a lemma first, then present a strictly stochastic dominance result under sublinear expectations.

Lemma 16. *Let $(\Omega, \mathscr{H}, \mathbb{E})$ be a sublinear expectation space, $Y \in \mathscr{H}$, and $\mathscr{U}$ a class of strictly increasing and continuously differentiable functions. If $u(Y)$ has no mean uncertainty, that is,*

$$\mathbb{E}[u(Y)] = -\mathbb{E}[-u(Y)], \tag{34}$$

then

$$\mathbb{E}[u(Y)] = E_{\mathbb{P}}[u(Y)], \quad \forall \mathbb{P} \in \mathscr{P}. \tag{35}$$

Proof. Since

$$\begin{aligned}\mathbb{E}[u(Y)] &= \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}[u(Y)], \\ -\mathbb{E}[-u(Y)] &= -\sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}[-u(Y)] = -\sup_{\mathbb{P} \in \mathscr{P}} (-\mathbb{E}_{\mathbb{P}}[u(Y)]) \\ &= -\left(-\inf_{\mathbb{P} \in \mathscr{P}} (\mathbb{E}_{\mathbb{P}}[u(Y)])\right) = \inf_{\mathbb{P} \in \mathscr{P}} (\mathbb{E}_{\mathbb{P}}[u(Y)]),\end{aligned} \tag{36}$$

we can get

$$\sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}[u(Y)] = \inf_{\mathbb{P} \in \mathscr{P}} (\mathbb{E}_{\mathbb{P}}[u(Y)]). \tag{37}$$

Then

$$\begin{aligned}\mathbb{E}[u(Y)] &= \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{E}_{\mathbb{P}}[u(Y)] = \inf_{\mathbb{P} \in \mathscr{P}} (\mathbb{E}_{\mathbb{P}}[u(Y)]) \\ &= E_{\mathbb{P}}[u(Y)], \quad \forall \mathbb{P} \in \mathscr{P}.\end{aligned} \tag{38}$$

□

Theorem 17. *Let $(\Omega, \mathscr{H}, \mathbb{E})$ be a sublinear expectation space, $X, Y \in \mathscr{H}$, and $\mathscr{U}$ a class of strictly increasing and continuously differentiable functions. If $X \geq Y$ q.s., $\mathbb{V}(X > Y) > 0$, and $\mathbb{E}[u(Y)] = -\mathbb{E}[-u(Y)]$, for all $u \in \mathscr{U}$, that is, $u(Y)$ has no mean uncertainty, then X strictly stochastically dominates Y, that is,*

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)], \quad \forall u \in \mathscr{U}. \tag{39}$$

Proof. Since $X \geq Y$ *q.s.* and $\mathbb{V}(X > Y) > 0$, there exists $\mathbb{P} \in \mathscr{P}$, such that

$$\begin{aligned} \mathbb{P}(X \geq Y) &= 1, \\ \mathbb{P}(X > Y) &> 0. \end{aligned} \tag{40}$$

Therefore

$$E_{\mathbb{P}}[X] > E_{\mathbb{P}}[Y]. \tag{41}$$

Since

$$\mathbb{E}[u(Y)] = -\mathbb{E}[-u(Y)], \quad \forall u \in \mathscr{U}, \tag{42}$$

according to Lemma 16, we have

$$\mathbb{E}[u(X)] \geq E_{\mathbb{P}}[u(X)] > E_{\mathbb{P}}[u(Y)] = \mathbb{E}[u(Y)], \quad \forall u \in \mathscr{U}. \tag{43}$$

□

Corollary 18. *Let $(\Omega, \mathscr{H}, \mathbb{E})$ be a sublinear expectation space, $X, Y \in \mathscr{H}$, and $X \geq Y$ q.s. Assume $u(x) = kx$, $k > 0$. If $\mathbb{V}(X > Y) > 0$ and Y has no mean uncertainty, that is, $\mathbb{E}[Y] = -\mathbb{E}[-Y]$, then*

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)]. \tag{44}$$

Proof. If $u(x) = kx$, $k > 0$, and Y has mean certainty, it is easy to verify that $u(Y)$ has mean certainty; that is, $\mathbb{E}[u(Y)] = -\mathbb{E}[-u(Y)]$. Then the consequence attains immediately by Theorem 17. □

Remark 19. This corollary gives sufficient condition for the result that X strictly stochastically dominates Y to the risk-neutral group.

Corollary 20. *Let $(\Omega, \mathscr{H}, \mathbb{E})$ be a sublinear expectation space, $X, Y \in \mathscr{H}$, and $\mathscr{U}$ a class of strictly increasing and continuously differentiable functions. If $X \geq Y$ q.s., $v(X \leq Y) < 1$, and $\mathbb{E}[u(Y)] = -\mathbb{E}[-u(Y)]$, for all $u \in \mathscr{U}$, that is, $u(Y)$ has no mean uncertainty, then X strictly stochastically dominates Y, that is,*

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)], \quad \forall u \in \mathscr{U}. \tag{45}$$

Proof. By using the relationship between $\mathbb{V}$ and v, the consequence attains immediately by Theorem 17. □

Next, we give an example to apply Theorem 17.

Example 21. Suppose there is an outcome $\Omega = [0, 1]$, which indicates the market conditions. $\mathscr{F}$ is the σ-algebra of Borel sets on Ω and $|\cdot|$ is the Lebesgue measure on $[0, 1]$.

We have two prior probabilities denoted by $\mathbb{P}$ and $\mathbb{Q}$, respectively, where

$$\begin{aligned} \mathbb{P}(A) &= \frac{1}{2}\left|A \cap \left[0, \frac{1}{3}\right)\right| + \left|A \cap \left[\frac{1}{3}, \frac{2}{3}\right)\right| + \frac{3}{2}\left|A \cap \left[\frac{2}{3}, 1\right]\right|, \\ &\qquad A \in \mathscr{F}, \\ \mathbb{Q}(A) &= \frac{1}{4}\left|A \cap \left[0, \frac{1}{3}\right)\right| + \frac{5}{4}\left|A \cap \left[\frac{1}{3}, \frac{2}{3}\right)\right| + \frac{3}{2}\left|A \cap \left[\frac{2}{3}, 1\right]\right|, \\ &\qquad A \in \mathscr{F}. \end{aligned} \tag{46}$$

We assume $\mathscr{H}$ is a collection of random variables, which represents risky securities and $\mathscr{U}$ is a class of strictly increasing and continuously differentiable functions, which represents the collection of all the utility functions of an investor group. Take the sublinear expectation utility as follows:

$$\begin{aligned} \mathbb{E}[u(\xi)] = \max\{E_{\mathbb{P}}[u(\xi)], E_{\mathbb{Q}}[u(\xi)]\}, \\ u \in \mathscr{U}, \quad \xi \in \mathscr{H}. \end{aligned} \tag{47}$$

There are two risky securities

$$\begin{aligned} X(\omega) &= \begin{cases} -10, & \omega = 0, \\ 0, & \omega \in \left(0, \frac{1}{3}\right), \\ 1, & \omega \in \left[\frac{1}{3}, \frac{2}{3}\right), \\ 10, & \omega \in \left[\frac{2}{3}, 1\right], \end{cases} \\ Y(\omega) &= \begin{cases} 0, & \omega \in \left[0, \frac{2}{3}\right), \\ 10, & \omega \in \left[\frac{2}{3}, 1\right]. \end{cases} \end{aligned} \tag{48}$$

It is clear that above conditions guarantee Theorem 17; therefore we have that X strictly stochastically dominates Y, that is,

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)], \quad u \in \mathscr{U}. \tag{49}$$

This means that all members of the investor group think the former is better than the latter.

In fact, for all $u \in \mathscr{U}$, we can calculate the following results:

$$\begin{aligned} E_{\mathbb{P}}[u(X)] &= \frac{1}{6}u(0) + \frac{1}{3}u(1) + \frac{1}{2}u(10), \\ E_{\mathbb{Q}}[u(X)] &= \frac{1}{12}u(0) + \frac{5}{12}u(1) + \frac{1}{2}u(10); \end{aligned} \tag{50}$$

then

$$\mathbb{E}[u(X)] = \frac{1}{12}u(0) + \frac{5}{12}u(1) + \frac{1}{2}u(10);$$

$$E_{\mathbb{P}}[u(Y)] = \frac{1}{2}u(0) + \frac{1}{2}u(10), \quad (51)$$

$$E_{\mathbb{Q}}[u(Y)] = \frac{1}{2}u(0) + \frac{1}{2}u(10);$$

then

$$\mathbb{E}[u(Y)] = \frac{1}{2}u(0) + \frac{1}{2}u(10). \quad (52)$$

Since $u \in \mathcal{U}$ is strictly increasing, then $u(1) > u(0)$; so

$$\mathbb{E}[u(X)] - \mathbb{E}[u(Y)] = \frac{5}{12}(u(1) - u(0)) > 0, \quad u \in \mathcal{U}. \quad (53)$$

Then

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)], \quad u \in \mathcal{U}; \quad (54)$$

that is, X strictly stochastically dominates Y.

4. Conclusion

In this paper, we study stochastic dominance under the nonlinear expected utilities. First we attain sufficient conditions on which a random choice X stochastically dominates a random choice Y under the nonlinear expected utilities; then we attain sufficient conditions on which strictly stochastic dominance of a random choice X over a random choice Y under the sublinear expected utilities; finally we give sufficient condition for strictly stochastic dominance of X over Y under the sublinear expected utilities to the risk-neutral group.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Xiao's researches are supported by the NNSF project (11301309) of China, NSF project (ZR2011AQ007) of Shandong Province of China, Project of Shandong Province Higher Educational Science and Technology Program (J11LA08), Project of Shandong Provincial Scientific Research Foundation for Excellent Young Scientists (BS2013SF003), and the Postdoctoral Science Foundation of China (2013M540540).

References

[1] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1953.

[2] M. Allais, "La paychologie de I'home rationnel devant le risqu: critique des postulats et axioms del'cole Americaine," *Econometrica*, vol. 21, no. 1953, pp. 503–546, 1979, translated and reprinted in Allais and Hagen.

[3] S. Peng, "Nonlinear expectations and stochastic calculus under uncertainty," http://arxiv.org/abs/1002.4546.

[4] X. Li, "On the strict comparison theorem for G-expectations," http://arxiv.org/abs/1002.1765.

[5] M. Hu, S. Ji, S. Peng, and Y. Song, "Comparison theorem, Feynman–Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion," *Stochastic Processes and Their Applications*, vol. 124, no. 2, pp. 1170–1195, 2014.

[6] M. Hu, S. Ji, S. Peng, and Y. Song, "Backward stochastic differential equations driven by G-Brownian motion," *Stochastic Processes and Their Applications*, vol. 124, no. 1, pp. 759–784, 2014.

[7] M. Hu, S. Ji, and S. Yang, "A stochastic recursive optimal control problem under the G-expectation framework," *Applied Mathematics & Optimization*, 2014.

[8] S. Peng, "G-Brownian motion and dynamic risk measure under volatility uncertainty," http://arxiv.org/abs/0711.2834.

[9] L. G. Epstein and S. Ji, "Ambiguous volatility, possibility and utility in continuous time," *Journal of Mathematical Economics*, vol. 50, pp. 269–282, 2014.

[10] L. G. Epstein and S. Ji, "Ambiguous volatility and asset pricing in continuous time," *Review of Financial Studies*, vol. 26, no. 7, pp. 1740–1786, 2013.

[11] X. Cao, "An upper bound of large deviations for capacities," *Mathematical Problems in Engineering*, vol. 2014, Article ID 516291, 6 pages, 2014.

[12] S. Peng, "Nonlinear expectations and nonlinear Markov chains," *Chinese Annals of Mathematics B*, vol. 26, no. 2, pp. 159–184, 2005.

[13] V. S. Bawa, "Optimal rules for ordering uncertain prospects," *Journal of Financial Economics*, vol. 2, no. 1, pp. 95–121, 1975.

[14] D. Dentcheva and A. Ruszczyski, "Optimization with stochastic dominance constraints," *SIAM Journal on Optimization*, vol. 14, no. 2, pp. 548–566, 2003.

[15] D. Dentcheva and A. Ruszczyński, "Semi-infinite probabilistic optimization: first-order stochastic dominance constraints," *Optimization*, vol. 53, no. 5-6, pp. 583–601, 2004.

[16] E. Jonathan and Jr. Ingersoll, *Theory of Financial Decision Making*, Rowman and Little field Publishers, 1987.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 437843, 14 pages
http://dx.doi.org/10.1155/2014/437843

Research Article

Parallel Array Bistable Stochastic Resonance System with Independent Input and Its Signal-to-Noise Ratio Improvement

Wei Li, Hanzhi Lu, and Yanyan Zuo

Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, China

Correspondence should be addressed to Wei Li; li.wei@hitsz.edu.cn

Received 6 March 2014; Revised 10 July 2014; Accepted 12 July 2014; Published 27 August 2014

Academic Editor: Wuquan Li

We study the design enhancement of the bistable stochastic resonance (SR) performance on sinusoidal signal and Gaussian white noise. The bistable system is known to show an SR property; however the performance improvement is limited. Our work presents two main contributions: first, we proposed a parallel array bistable system with independent components and averaged output; second, we give a deduction of the output signal-to-noise ratio (SNR) for this system to show the performance. Our examples show the enhancement of the system and how different parameters influence the performance of the proposed parallel array.

1. Introduction

Stochastic resonance has attracted considerable attention over the past decades. SR is defined as a phenomenon that is manifest in nonlinear systems whereby generally feeble input information (such as a weak signal) can be amplified and optimized by the assistance of noise.

The physical mechanism of SR has been known since the initial work by Benzi et al. at the beginning of the 1980s [1–3] and received much attention by the physical community in the following years. SR has been observed in a large variety of systems, including bistable ring lasers and semiconductor devices. The first discussed and most developed SR mechanism was the bistable system. Since it has a precise mathematical expression and can be interpreted visually, the bistable system draws much attention of the researchers.

SR can be envisioned as a particular problem of signal extraction from background noise. It is quite natural that a number of authors tried to characterize SR within the formalism of data analysis, most notably by introducing the notion of SNR [4–6]. The focus of our present work is on bistable system and its SNR improvement. SNR is a very important quantity, since it influences information, detection, estimation, and many other measures [7].

The early study of SR system focused on nature nonlinear system to analyse its properties [1–3]. Later, the benefit of the system was known, and researchers started to design new systems to meet the need in engineering to enhance the performance of the system. Many works dealt with SR in engineering such as signal estimation and detection [8–10]. A good way of designing the SR system is using array since array can enhance system performance which is widely studied [11–14]. The array for SR systems can be designed either in coupled way [12, 15–18] or in uncoupled way [13, 14]. For the coupled array, the processing in each component is complicated due to coupling with other components. Uncoupled parallel array has been widely studied in SR system due to its simplicity, such as superthreshold system. For bistable system, the work in [19] gives a brief introduction of a type of array enhancement for the sinusoidal signal in bistable array with a similar structure as superthreshold system. In [20] the theory for this type of array is demonstrated. However, even with the uncoupled components, the performance of the system still has room to be improved. Since, in these types of array, the components are not independent of each other, the independence in statistics is an importance feature to the best performance.

This paper is in fact inspired by traditional parallel system, proposes a new parallel array with independent sensors, and focuses on the output SNR performance. This is different from traditional parallel SR system since traditional system uses one receiving sensor and parallel array processing components so that input for each component is not independent in statistics. And it is also different from traditional array signal processing [7] since we do not need to consider the shape of the array. To simplify the analysis, we limit our study to two-state bistable system driven by sinusoidal signal and Gaussian white noise and assume some identical independent settings in every bistable component. To analyse the performance of this array theoretically, we give a complete proof on output SNR and experiments to demonstrate the parameter influences.

This paper is organized as follows. The framework of two-state model of bistable system is described in Section 2. Section 3 deals with the case that a new structure of the parallel array is assigned to bistable system and the output SNR of this system is deduced. Section 4 is devoted to instances of the proposed system whose performance is indeed enhanced by adding noise. And the influence of the parameter on the system is also analysed in this section. Finally, in Section 5, we summarize the following.

Notation. $E(\cdot)$ stands for ensemble average, upper dot $\dot{a}$ denotes a time derivative of a, $A'(b)$ represents the derivative of A with respect to b, $\delta(\cdot)$ is Dirac delta function, and $f * g$ represents the convolution of f and g.

2. Two-State Model of Bistable Systems

We consider the overdamped motion of a Brownian particle in a bistable potential in the presence of noise and periodic forcing [21, 22]. The system can be presented by Fokker-Planck equation. Consider

$$\dot{x}(t) = -U'(x) + A_0 \cos(\Omega t) + \xi(t), \tag{1}$$

where x is the position of Brownian particle, $U(x)$ denotes the reflection-symmetric quartic potential,

$$U(x) = -\frac{a}{2}x^2 + \frac{b}{4}x^4, \tag{2}$$

$\xi(t)$ denotes a zero-mean, Gaussian white noise with variance $2D$, and $A_0\cos(\Omega t)$ is periodic forcing. The potential $U(x)$ is bistable with minima located at $\pm x_m$, with $x_m = (a/b)^{1/2}$. The height of the potential barrier between the minima is given by $\Delta U = a^2/4b$.

To simplify the problem in this paper we discuss two-state model [23, 24] that epitomizes the class of symmetric bistable systems introduced. Such a discrete model under certain restrictions renders an accurate representation of most continuous bistable systems. Let us consider a symmetric unperturbed system that switches between two discrete states $\pm m_x$. We define $n_{\pm}(t)$ to be the probabilities that the system occupies either state $\pm$ at time t; that is, $x(t) = \pm x_m$. Then

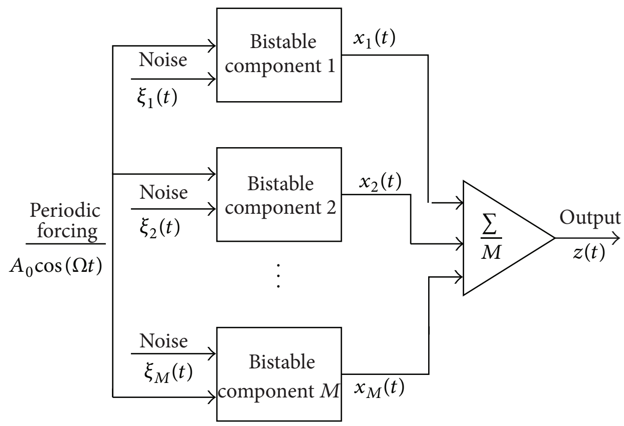


Figure 1: The parallel bistable array with M independent components.

the power spectral density of this symmetric bistable system commonly reported in the literature [21] is

$$\begin{aligned} S(\omega) = {} & \left[1 - \frac{1}{2}\left(\frac{A_0 x_m}{D}\right)^2 \frac{4r_k^2}{4r_k^2 + \Omega^2}\right] \times \frac{4r_k x_m^2}{4r_k^2 + \omega^2} \\ & + \frac{\pi}{2}\left(\frac{A_0 x_m}{D}\right)^2 \times \frac{4r_k x_m^2}{4r_k^2 + \Omega^2}\left[\delta(\omega - \Omega) + \delta(\omega + \Omega)\right], \end{aligned} \tag{3}$$

in which Kramers rate

$$r_k = \frac{1}{\sqrt{2}\pi} \exp\left(-\frac{\Delta U}{D}\right). \tag{4}$$

It is rate of transitions between the neighboring potential wells caused by the fluctuational forces.

Since the noise in the output of the system is no longer Gaussian, the definition mean^2/variance for SNR is not suitable in this system. Here we adopt the definition for input and output SNR according to [21] as follow:

$$\mathrm{SNR} = \frac{2\left[\lim_{\Delta\omega \to 0} \int_{\Omega-\Delta\omega}^{\Omega+\Delta\omega} S(\omega)\, d\omega\right]}{S_N(\Omega)}. \tag{5}$$

For the weak signal ($A_0 x_m \ll \Delta U$), we can omit high order items; then the output SNR for this symmetric bistable system is approximately

$$\mathrm{SNR} = \pi\left(\frac{A_0 x_m}{D}\right)^2 r_k. \tag{6}$$

3. Parallel Bistable Array with Independent Components

3.1. Proposed Array and Its Output SNR. In this section, we discuss the parallel array bistable system and its SR performance.

We consider the parallel array with M components in Figure 1. Each component has a receiving sensor and a processing property as described in the last section. We suppose that the receiving time difference $\Delta t \ll T_\Omega = 2\pi/\Omega$ but there still is a considerable distance between every two different components. This can be set by a suitable Ω. Then the input periodic signal can be taken as $A_0\cos(\Omega t)$ for all the components. And the noise $\xi_i(t)$ is independent and identically distributed (IID) for each input, and the output is $x_i(t)$ for the ith component; then all the outputs are averaged and the response of the array is given as

$$z(t) = \frac{\sum_{i=1}^{M} x_i(t)}{M}. \tag{7}$$

In the following of this section, we present the main results with respect to the parallel array bistable system. Two theorems form the SR performance analysis on output SNR. We utilize four lemmas for proving the theorems. The proofs of all the theorems and lemmas of this section are relegated to appendices.

Theorem 1. *For the parallel bistable system with M components, the output SNR is*

$$\mathrm{SNR}(M) = \frac{\pi(A_0x_m/D)^2 r_k}{-(1/2)(A_0x_m/D)^2\left(4r_k^2/\left(4r_k^2+\Omega^2\right)\right) - H(M)}, \tag{8}$$

where

$$H(M) = \begin{cases} F(M), & \text{if } M \text{ is even} \\ G(M), & \text{if } M \text{ is odd,} \end{cases} \tag{9}$$

in which

$$\begin{aligned} F(M) &= \left(\frac{1}{2}\right)^M \sum_{k=0}^{M/2-1} \frac{2k-M}{M}\binom{M}{k} \\ &\quad\times\left\{\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}B^{M-i-j} f(M-i-j)\right. \\ &\quad\left.\times\left[(-1)^{M-k-j}+(-1)^{k-i}\right]\right\}, \\ G(M) &= \left(\frac{1}{2}\right)^M \sum_{k=0}^{(M-1)/2} \frac{2k-M}{M}\binom{M}{k} \\ &\quad\times\left\{\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}B^{M-i-j} f(M-i-j)\right. \\ &\quad\left.\times\left[(-1)^{M-i-j}+(-1)^{k-i}\right]\right\}. \end{aligned} \tag{10}$$

In the above equations,

$$\begin{aligned} B &= \frac{2r_k(A_0x_m/D)}{\sqrt{\left(4r_k^2+\Omega^2\right)}}, \\ f(x) &= \frac{\Gamma((x+1)/2)}{\sqrt{\pi}\,\Gamma(x/2+1)}, \end{aligned} \tag{11}$$

and Γ is gamma function and defined as

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt. \tag{12}$$

The theory is based on the following lemmas.

Lemma 2. *The pdf of $z(t)$ is*

$$p_z(z,t \mid z_0,t_0) = \sum_{m=0}^{M}\binom{M}{m} n_+^m(t)\, n_-^{M-m}(t)\,\delta\left(z - \frac{2m-M}{M}x_m\right). \tag{13}$$

Lemma 3. *If M is even, the autocorrelation function of $z(t)$ is*

$$\begin{aligned} R_z(t+\tau,t) &= x_m^2\left\{\left[-\exp(-2r_k|\tau|)\kappa^2(t) + \kappa(t+\tau)\right]\kappa(t)\right. \\ &\quad - \left(\frac{1}{2}\right)^M \exp(-2r_k|\tau|) \\ &\quad\times\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k} \\ &\quad\times\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}\times B^{M-i-j} \\ &\quad\times\cos(\Omega t-\varphi)^{M-i-j} \\ &\quad\left.\times\left[(-1)^{M-k-j}+(-1)^{k-i}\right]\right\}, \end{aligned} \tag{14}$$

where

$$\kappa(t) = B\cos(\Omega t-\varphi). \tag{15}$$

Lemma 4. *If M is odd, the autocorrelation function of $z(t)$ is*

$$\begin{aligned}
&R_z(t+\tau,t)\\
&= x_m^2 \Bigg\{ \left[-\exp\left(-2r_k|\tau|\right)\kappa^2(t) + \kappa(t+\tau) \right]\kappa(t)\\
&\quad - \left(\frac{1}{2}\right)^M \exp\left(-2r_k|\tau|\right)\\
&\quad \times \sum_{k=0}^{(M-1)/2} \frac{2k-M}{M}\binom{M}{k}\\
&\quad \times \sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j} \times B^{M-i-j}\\
&\quad \times \cos\left(\Omega t-\varphi\right)^{M-i-j}\\
&\quad \times \left[(-1)^{M-i-j} + (-1)^{k-i}\right] \Bigg\}.
\end{aligned} \tag{16}$$

Lemma 5. *The power spectral density of the output of the parallel array bistable system with M components is*

$$\begin{aligned}
S(\omega) &= -\left[\frac{1}{2}\left(\frac{A_0x_m}{D}\right)^2 \frac{4r_k^2}{4r_k^2+\Omega^2} + H(M)\right]\frac{4r_kx_m^2}{4r_k^2+\omega^2}\\
&\quad + \frac{\pi}{2}\left(\frac{A_0x_m}{D}\right)^2 \frac{4r_k^2x_m^2}{4r_k^2+\omega^2}\\
&\quad \times \left[\delta(\omega-\Omega)+\delta(\omega+\Omega)\right].
\end{aligned} \tag{17}$$

Theorem 6. *For the weak signal ($A_0x_m \ll \Delta U$), the output SNR for this parallel bistable system with two components is approximately*

$$SNR = 2\pi\left(\frac{A_0x_m}{D}\right)^2 r_k. \tag{18}$$

3.2. Remark. We conclude this section with three remarks. Our first remark is about the simplified noise. The noise $\xi_i(t)$ in practical problems is the sum of two parts in each component. The first part is the receiving noise buried in the receiving signal, and the second part is the tuning noise. Here we suppose that the receiving noise and the tuning noise are independent Gaussian white noise. And the variance of the tuning noise can be set by us. Then we can simplify the noise in each component as Gaussian white noise with $2D$ variance.

Our second remark is to point out that the proposed array is different from the traditional SR array [20] due to its independent sensors, and it is also different from traditional array signal processing [7] since the shape of the array does not affect the performance. We not only proposed an M-array system for bistable SR, but also provided a rigorous proof for the output SNR which is nontrivial as evidenced. And the results also divide M into odd and even situations.

Our third remark is that Theorem 1 can arrive to (6) by setting $M = 1$ and Theorem 6 by setting $M = 2$. In fact, Figure 1 shows that if $M = 1$, the system without array becomes the conventional single bistable SR system. The equivalence is also shown in Figure 2 on the same other parameters.

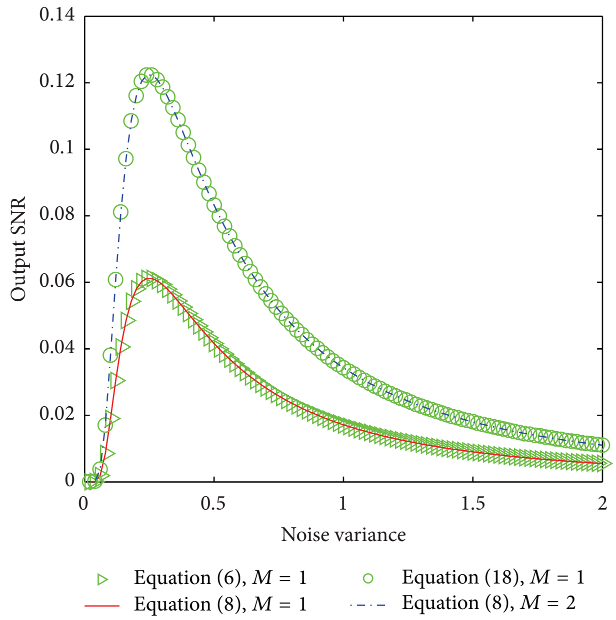


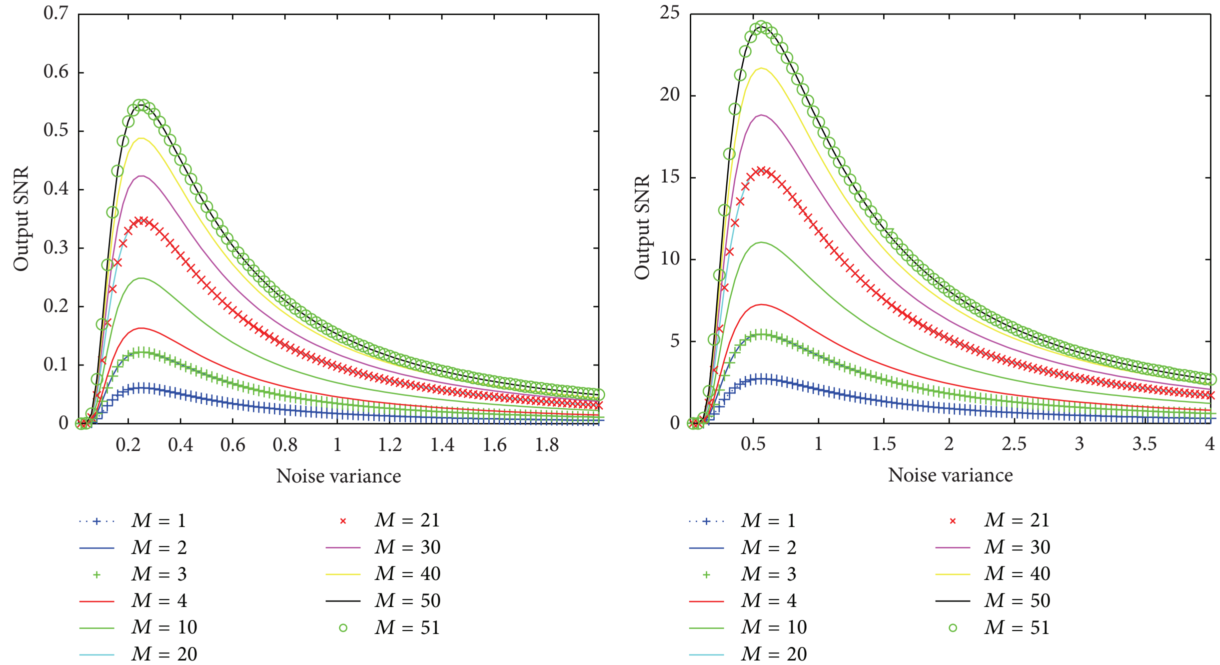
Figure 2: Output SNR as a function of noise variance with $A_0 = 0.1$, $a = 1$, $b = 1$, and $f = 100$.

4. Simulation Results

We now provide examples to illustrate the properties of our proposed bistable parallel array system.

4.1. SR Effect and the Influence of M. For illustration of the possibility of an SR in the output SNR, we consider two different systems based on the theory of (8) for the proposed array: case A: $a = 1$, $b = 1$, $A = 0.1$, and $f = 100$ in Figure 3(a); case B: $a = 1.5$, $b = 1$, $A = 1$, and $f = 10$ in Figure 3(b). Though the two systems are set by different parameters, they both display evolutions of the resulting output SNR of (8), as a function of the noise variance, in some typical conditions. Since D cannot be zero in (8), all the curves start from a small amount D close to zero. As noise power increases, the SR peak rises, shifts to higher noises, and then subsides. This result shows us that if the input noise variance is smaller than the peak point, the tuning noise can be added to improve the output SNR performance. For increasing M, the efficiency of the array and the maximum output SNR increase. This demonstrates that the array of nonlinear devices can play the role of an SNR amplifier, in definite conditions.

At $M = 2$, SR effect gets more pronounced than at $M = 1$ which is the traditional SR bistable system. The output SNR is twice the output of the traditional bistable system according to Theorem 6. As M increases, SR effect gets more pronounced. However if M is even ($M \geq 2$), the output



(a) (b)

FIGURE 3: SNR curve changes as noise power. (a) $a = 1$, $b = 1$, $A = 0.1$, and $f = 100$. (b) $a = 1.5$, $b = 1$, $A = 1$, and $f = 10$.

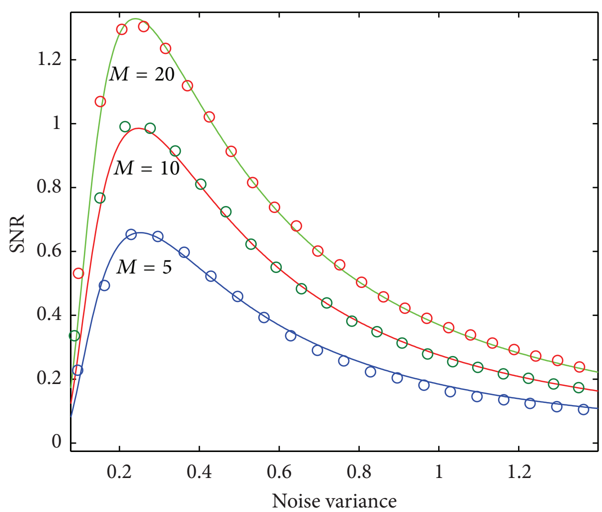


FIGURE 4: Output SNR with $a = 1$, $b = 1$, $A = 0.02$, and $f = 0.01$. The solid lines are from the theory of (8). The sets of discrete points (∘) are from Monte Carlo simulations.

SNR between M and $M+1$ array is very approximate leaving very small difference. This is because, for an even M, we have $(G(M+1) - F(M))/(F(M+2) - G(M+1)) \approx 0$. Then $\mathrm{SNR}(M+1)/\mathrm{SNR}(M) \approx 1$. And the increment grows smaller even if M is only odd or even with increasing M.

The results of Figure 3 reveal that the characteristic behaviors that identify the array bistable SR are precisely exhibited by the evolutions of the SNR. However for equal M, SNR displays different evolutions in the two figures in Figure 3. This is caused by the other parameters of the system. We will show the influence of these parameters in the following examples.

We also offered a validation by a Monte Carlo simulation of the proposed system in Figure 4 by setting $a = 1$, $b = 1$, $A = 0.02$, and $f = 0.01$. The results coincided with (8).

4.2. SR Effect and the Influence of Signal Amplitude. We consider, in Figure 5, the transmission by the array of a sinusoidal wave $s(t) = A_0 \cos(2\pi t/Ts)$ buried in noise based on the theory of (8). The values of the amplitude A_0 determine how the input $s(t)$ is seen by the array. We choose a parallel array with $a = 1$, $b = 1$, $M = 38$, and $f = 100$. Figure 5 shows various evolutions of the SNR at the output of the array, for different values of the constant A_0. For the value of A_0 tested in Figure 5, the performance for the periodic input $s(t)$ is always SR. With increasing A_0, as the level of noise variance is increased, the output SNR experiences nonmonotonic evolutions. In this experiment, we also set A_0 to be very large numbers and very small numbers under $A_0 > 0$. The outputs of the system all perform the SR effect. And as A_0 grows, the effect gets more enhanced. We relate these results to the phenomenon of the proposed parallel array SR, by which nonlinear transmission or processing of signals with arbitrary amplitude can be improved by adding noises in arrays.

4.3. SR Effect and the Influence of Signal Frequency. In this example, we consider $A_0 = 0.1$, $a = 1.5$, $b = 1$, and $M = 33$ array system based on the theory of (8). Let $f = \Omega/2\pi$. We choose different f to see the influence of the signal frequency.

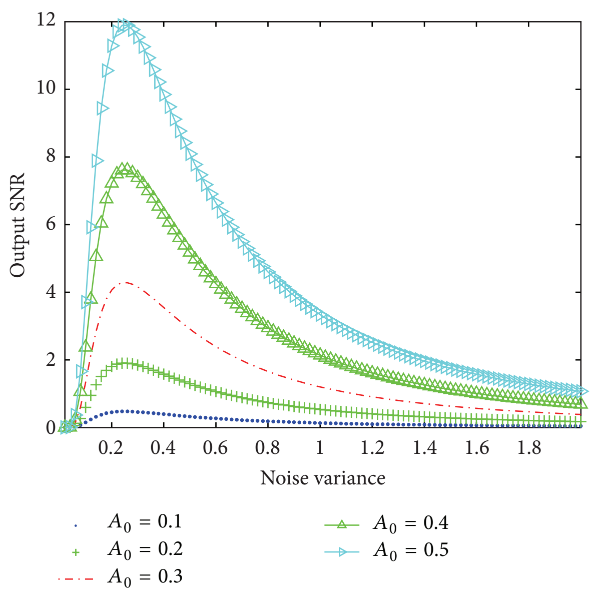


FIGURE 5: Output SNR as a function of noise variance with $a = 1$, $b = 1$, $M = 38$, and $f = 100$.

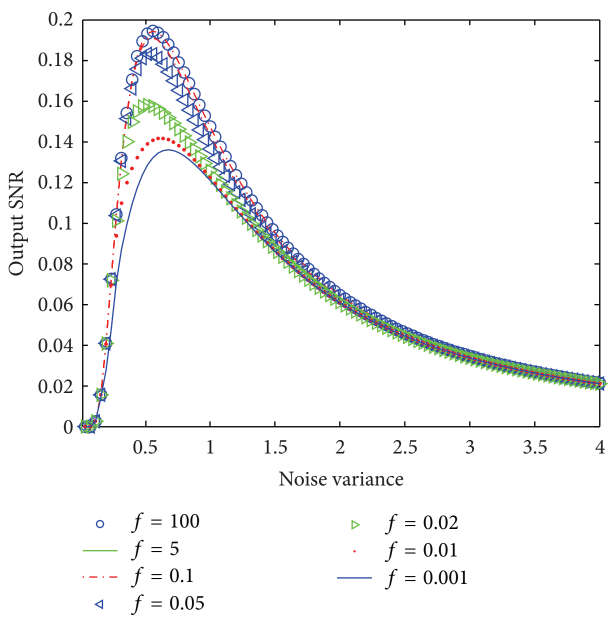


FIGURE 6: Output SNR as a function of noise variance with $A_0 = 0.1$, $a = 1.5$, $b = 1$, and $M = 33$.

The output SNR versus noise variance of the two systems is given in Figure 6. We observe that the output SNR grows from near zero point to the maximum point and then goes down with different frequency tested in this example. The noise variances of maximum output SNR points in this case are slightly different. With the growing of frequency the SR effect becomes more enhanced. However, when f is big enough, the growth of f does not affect the output SNR. No matter how we increase the frequency, the system stays the same at the extremal SNR. This property also helps us to choose a suitable signal frequency under our hypothesis $\Delta t \ll 1/f$. In our experiment when $f < 0.2$ the system loses the SR effect. This phenomenon also shows that the system does not have the SR effect when the signal is DC signal, since we can take the periodic signal as DC signal, if the frequency of the signal is extremely low.

4.4. SR Effect and the Influence of System Parameters. Figure 7 shows various evolutions of the SNR at the output of the array, for different values of system parameters a and b based on the theory of (8). In Figure 7(a) to observe the influence of a, we set $A = 0.1$, $b = 1.5$, $f = 10$, and $M = 33$. As we can see from the figure, the smaller the parameter a is, the stronger the SR effect is. When a becomes big enough, the system loses the SR effect. In fact, in this condition the output SNR is nearly zero.

In Figure 7(b) to observe the influence of b, we set $A = 0.1$, $a = 1.5$, $f = 10$, and $M = 33$, the same as Figure 7(a) except parameters a and b. The result of the output SNR versus noise variance indicates that the bigger the parameter b, the stronger the SR effect. This is the opposite to the influence of a, because b has a positive effect on reflection-symmetric quartic potential, while a has a negative one. And comparing the two figures, the influence of a outweighs that of b, and this is obvious due to Kramers rate.

4.5. Input-Output SNR Gain. We still adopt the definition for input and output SNR in (5). Then input SNR for the sinusoidal signal and a zero-mean, Gaussian white noise $\mathrm{SNR}_{\mathrm{in}} = \pi A^2/D$. The input-output SNR gain is defined below. Consider

$$G_{\mathrm{SNR}} = \frac{\mathrm{SNR}_{\mathrm{out}}}{\mathrm{SNR}_{\mathrm{in}}}. \tag{19}$$

In this experiment, we let $A_0 = 0.1$, $a = 1$, $b = 1$, and $f = 1$. In Figure 8, the G_{SNR} grows first and then decreases with increased noise variance. And the result in Figure 8 shows the array system outweighs the signal bistable system on SR effect and the G_{SNR} can exceed unity for $M > 3$ in this experiment. It means that the array can improve the signal-to-noise ratio (SNR) by noise incoherently. The improvement is measured by the array gain. For $M = 50$, the maximum of G_{SNR} is 2.95. Thus, this SR array with independent sensors provides a preferable strategy for processing periodic signals to the array without independent sensors which exceeds unity much less [19].

5. Conclusion

In this work, we study the design of structure of bistable system aimed at enhancing the SR effect to improve the performance, driven by sinusoidal signal and Gaussian white noise. We first proposed a parallel array bistable system with M independent components and averaged output. We further deduced the output signal-to-noise ratio (SNR) for this parallel array system to analyse the performance of this SR system. Our examples not only show the proposed system reserves the SR property, but also give an analysis of different

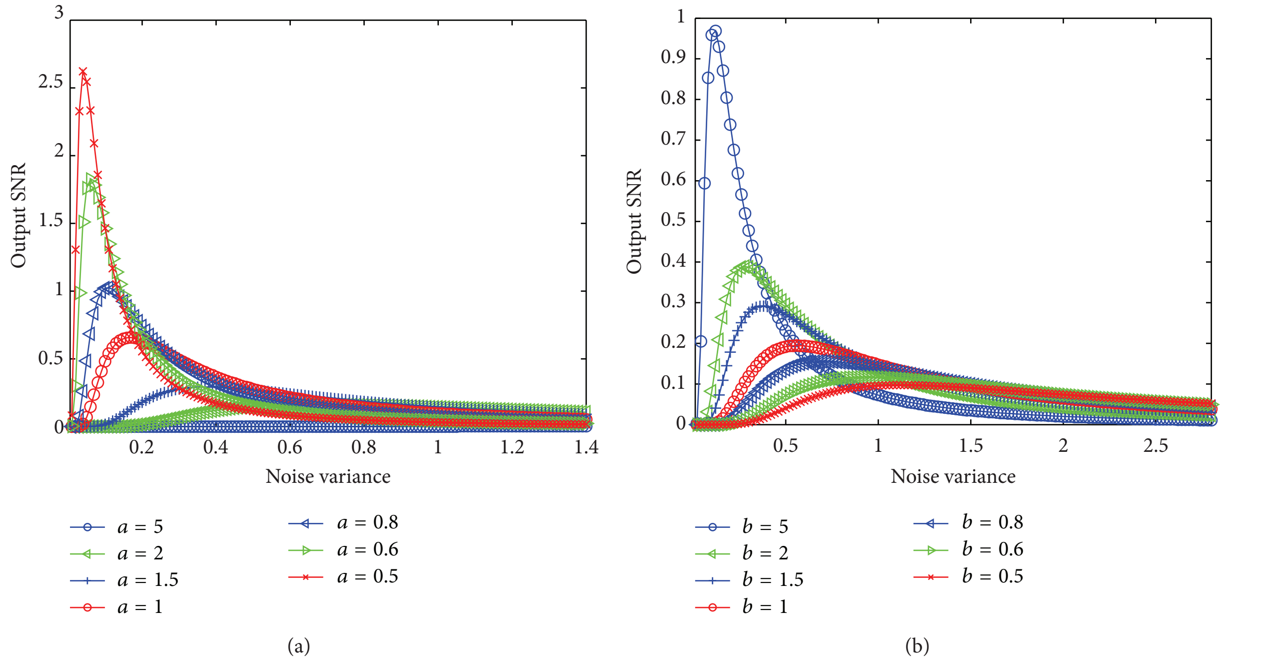


FIGURE 7: Output SNR as a function of noise variance. (a) $A = 0.1$, $b = 1.5$, $f = 10$, and $M = 33$. (b) $A = 0.1$, $a = 1.5$, $f = 10$, and $M = 33$.

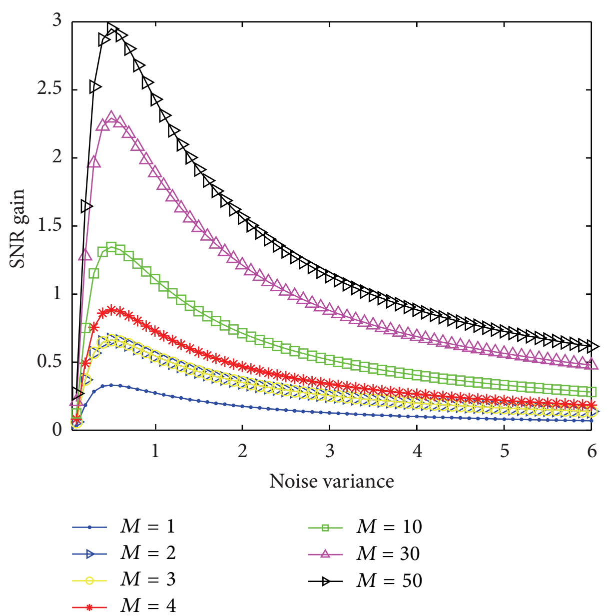


FIGURE 8: Output SNR gain as a function of noise variance, with $A_0 = 0.1$, $a = 1$, $b = 1$, and $f = 1$.

parameter influences on the performance of the proposed parallel array, indicating a promising application in array signal processing.

Appendices

A. Proofs of Lemmas 2–5 and Theorem 1

Proof of Lemma 2. From the structure of the parallel array bistable system, the output of the system is

$$z(t) = \frac{\sum_{i=1}^{M} x_i(t)}{M}. \tag{A.1}$$

For the ith bistable component, the pdf of $x_i(t)$ is

$$p_x(x_i, t \mid x_0, t_0) = n_+(t)\,\delta(x_i - x_m) + n_-(t)\,\delta(x_i + x_m). \tag{A.2}$$

Since $x_i(t)$ is independent in statistics for $i = 1, 2, \ldots, M$, the pdf of the sum of the outputs of M components is the convolution of each component. Then due to property of delta function, we simplify the result of the convolution and obtain the following pdf of $z(t)$. Consider

$$p_z(z, t \mid z_0, t_0) = \sum_{m=0}^{M} \binom{M}{m} n_+^m(t)\, n_-^{M-m}(t)\, \delta\left(z - \frac{2m - M}{M} x_m\right). \tag{A.3}$$

□

Proof of Lemma 3. Based on Lemma 2 and the general definition of autocorrelation function, we can deduce the autocorrelation function of $z(t)$,

$$R_z(t+\tau, t) = \iint_{-\infty}^{+\infty} z_1 z_2 p_z(z_1, t+\tau \mid z_2, t) \times p_z(z_2, t \mid z_0, t_0)\, dz_1 dz_2$$

$$= \iint_{-\infty}^{+\infty} z_1 z_2 \times \sum_{m=0}^{M} \binom{M}{m} n_+^m(t+\tau \mid z_2, t)$$

$$\times n_-^{M-m}(t+\tau\mid z_2,t)$$
$$\times\delta\left(z_1-\frac{2m-M}{M}x_m\right)$$
$$\times\sum_{k=0}^{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)$$
$$\times n_-^{M-k}(t\mid z_0,t_0)$$
$$\times\delta\left(z_2-\frac{2k-M}{M}x_m\right)dz_1dz_2$$
$$=x_m^2\sum_{m=0}^{M}\frac{2m-M}{M}\binom{M}{m}$$
$$\times n_+^m\left(t+\tau\mid\frac{2k-M}{M}x_m,t\right)$$
$$\times n_-^{M-m}\left(t+\tau\mid\frac{2k-M}{M}x_m,t\right)$$
$$\times\sum_{k=0}^{M}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)$$
$$\times n_-^{M-k}(t\mid z_0,t_0).\tag{A.4}$$

The last step follows from the property of delta function.

Since

$$\sum_{i=0}^{N}\binom{N}{i}\frac{2i-N}{N}x^iy^{N-i}=(x+y)^{N-1}(x-y),\tag{A.5}$$

and making use of the normalization condition $n_+(t)+n_-(t)=1$, (A.4) becomes

$$R_z(t+\tau,t)$$
$$=x_m^2\left[n_+\left(t+\tau\mid\frac{2k-M}{M}x_m,t\right)\right.$$
$$\left.-n_-\left(t+\tau\mid\frac{2k-M}{M}x_m,t\right)\right]$$
$$\times\sum_{k=0}^{M}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)n_-^{M-k}(t\mid z_0,t_0)$$
$$=x_m^2\left[2n_+\left(t+\tau\mid\frac{2k-M}{M}x_m,t\right)-1\right]$$
$$\times\sum_{k=0}^{M}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)n_-^{M-k}(t\mid z_0,t_0).\tag{A.6}$$

If M is even, the range $(0,M)$ of k can be divided into $(0,M/2-1)$ and $(M/2,M)$. If $k=M/2$, it is obvious that $R_z(t+\tau,t)=0$. Then

$$R_z(t+\tau,t)$$
$$=x_m^2\left\{[2n_+(t+\tau\mid -,t)-1]\right.$$
$$\times\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)$$
$$\times n_-^{M-k}(t\mid z_0,t_0)$$
$$+[2n_+(t+\tau\mid +,t)-1]$$
$$\times\sum_{k=M/2+1}^{M}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)$$
$$\left.\times n_-^{M-k}(t\mid z_0,t_0)\right\}.\tag{A.7}$$

From [21], we have

$$n_+(t+\tau\mid -,t)$$
$$=\frac{1}{2}\{\exp(-2r_k|\tau|)[-1-\kappa(t)]+1+\kappa(t+\tau)\},$$
$$n_+(t+\tau\mid +,t)\tag{A.8}$$
$$=\frac{1}{2}\{\exp(-2r_k|\tau|)[1-\kappa(t)]+1+\kappa(t+\tau)\},$$

where

$$\kappa(t)=B\cos(\Omega t-\varphi),$$
$$B=\frac{2r_k(A_0x_m/D)}{\sqrt{(4r_k^2+\Omega^2)}}.\tag{A.9}$$

$n_+(t+\tau\mid -,t)$ and $n_+(t+\tau\mid +,t)$ are $\alpha<0$ and $\alpha>0$ in $n_+(t+\tau\mid\alpha,t)$, respectively [21].

It greatly simplifies in the stationary limit $t_0\to-\infty$,

$$\lim_{t_0\to-\infty}R_z(t+\tau,t)$$
$$=x_m^2\left\{\{\exp(-2r_k|\tau|)[-1-\kappa(t)]+\kappa(t+\tau)\}\right.$$
$$\times\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)$$
$$\times n_-^{M-k}(t\mid z_0,t_0)$$
$$+\{\exp(-2r_k|\tau|)[1-\kappa(t)]+\kappa(t+\tau)\}$$
$$\times\sum_{k=M/2+1}^{M}\frac{2k-M}{M}\binom{M}{k}n_+^k(t\mid z_0,t_0)$$

$$\times n_-^{M-k}(t \mid z_0, t_0)\Bigg\}$$

$$= x_m^2\Bigg\{\left[-\exp(-2r_k|\tau|)\kappa^2(t)+\kappa(t+\tau)\right]\kappa(t) - \exp(-2r_k|\tau|) \times \sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}n_+^k(t \mid z_0, t_0) \times n_-^{M-k}(t \mid z_0, t_0) + \exp(-2r_k|\tau|) \times \sum_{k=M/2+1}^{M}\frac{2k-M}{M}\binom{M}{k}n_+^k(t \mid z_0, t_0) \times n_-^{M-k}(t \mid z_0, t_0)\Bigg\}. \tag{A.10}$$

Take

$$\lim_{t_0\to-\infty} n_+(t \mid z_0, t_0) = \frac{1}{2}[1+\kappa(t)], \quad \lim_{t_0\to-\infty} n_-(t \mid z_0, t_0) = \frac{1}{2}[1-\kappa(t)] \tag{A.11}$$

into (A.10) to obtain

$$\lim_{t_0\to-\infty} R_z(t+\tau, t) = x_m^2\Bigg\{\left[-\exp(-2r_k|\tau|)\kappa^2(t)+\kappa(t+\tau)\right]\kappa(t) - \exp(-2r_k|\tau|) \times \sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}\left\{\frac{1}{2}[1+\kappa(t)]\right\}^k \times \left\{\frac{1}{2}[1-\kappa(t)]\right\}^{M-k} + \exp(-2r_k|\tau|) \times \sum_{k=M/2+1}^{M}\frac{2k-M}{M}\binom{M}{k} \times \left\{\frac{1}{2}[1+\kappa(t)]\right\}^k \times\left\{\frac{1}{2}[1-\kappa(t)]\right\}^{M-k}\Bigg\}$$

$$= x_m^2\Bigg\{\left[-\exp(-2r_k|\tau|)\kappa^2(t)+\kappa(t+\tau)\right]\kappa(t) - \left(\frac{1}{2}\right)^M\exp(-2r_k|\tau|) \times \Bigg\{\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}[1+\kappa(t)]^k \times [1-\kappa(t)]^{M-k} - \sum_{k=M/2+1}^{M}\frac{2k-M}{M}\binom{M}{k} \times [1+\kappa(t)]^k \times [1-\kappa(t)]^{M-k}\Bigg\}\Bigg\}. \tag{A.12}$$

The last part can be reexpressed as

$$\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}[1+\kappa(t)]^k[1-\kappa(t)]^{M-k} - \sum_{k=0}^{M/2-1}\frac{M-2k}{M}\binom{M}{M-k}[1+\kappa(t)]^{M-k}[1-\kappa(t)]^k$$
$$= \sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k} \times \sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}\kappa(t)^{M-i-j} \times \left[(-1)^{M-k-j}+(-1)^{k-i}\right]$$
$$= \sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k} \times \sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}B^{M-i-j} \times \cos(\Omega t-\varphi)^{M-i-j} \times \left[(-1)^{M-k-j}+(-1)^{k-i}\right]. \tag{A.13}$$

Then we have

$$\begin{aligned}
&R_z(t+\tau,t)\\
&= x_m^2\Bigg\{\left[-\exp(-2r_k|\tau|)\kappa^2(t)+\kappa(t+\tau)\right]\kappa(t)\\
&\quad-\left(\frac{1}{2}\right)^M\exp(-2r_k|\tau|)\\
&\quad\times\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}\\
&\quad\times\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}B^{M-i-j}\\
&\quad\times\cos(\Omega t-\varphi)^{M-i-j}\\
&\quad\times\left[(-1)^{M-k-j}+(-1)^{k-i}\right]\Bigg\}.
\end{aligned}\tag{A.14}$$

□

Proof of Lemma 4. If M is odd, we can prove Lemma 4 in a similar manner as Lemma 3. After some mathematical manipulations, we obtain the following:

$$\begin{aligned}
&R_z(t+\tau,t)\\
&= x_m^2\Bigg\{\left[-\exp(-2r_k|\tau|)\kappa^2(t)+\kappa(t+\tau)\right]\kappa(t)\\
&\quad-\left(\frac{1}{2}\right)^M\exp(-2r_k|\tau|)\\
&\quad\times\sum_{k=0}^{M-1/2}\frac{2k-M}{M}\binom{M}{k}\\
&\quad\times\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}B^{M-i-j}\\
&\quad\times\cos(\Omega t-\varphi)^{M-i-j}\\
&\quad\times\left[(-1)^{M-i-j}+(-1)^{k-i}\right]\Bigg\}.
\end{aligned}\tag{A.15}$$

□

Proof of Lemma 5. It is obvious that the autocorrelation function depends on both times $t+\tau$ and t. However, in real experiments t represents the time set for the trigger in the data acquisition procedure. Typically, the averages implied by the definition of the autocorrelation function are taken over many sampling records of the signal $x(t)$, triggered at a large number of times t within one period of the forcing T_Ω. Hence, the corresponding phases of the input signal, $\theta=\Omega t+\psi$, are uniformly distributed between 0 and 2π. This corresponds to averaging autocorrelation function as with respect to t uniformly over an entire forcing period, whence if M is even,

$$\begin{aligned}
&R_z(\tau)\\
&=\frac{1}{T_\Omega}\int_0^{T_\Omega}R_z(t+\tau,t)\,dt\\
&=\frac{1}{T_\Omega}\int_0^{T_\Omega}x_m^2\\
&\qquad\times\Bigg\{\left[-\exp(-2r_k|\tau|)\kappa^2(t)+\kappa(t+\tau)\right]\kappa(t)\\
&\qquad\quad-\left(\frac{1}{2}\right)^M\exp(-2r_k|\tau|)\\
&\qquad\quad\times\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\binom{M}{k}\\
&\qquad\quad\times\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}B^{M-i-j}\\
&\qquad\quad\times\cos(\Omega t-\varphi)^{M-i-j}\\
&\qquad\quad\times\left[(-1)^{M-k-j}+(-1)^{k-i}\right]\Bigg\}dt\\
&=x_m^2\Bigg\{-\frac{1}{2}\exp(-2r_k|\tau|)B^2\\
&\qquad+\frac{1}{2}B^2\cos(\Omega\tau)-\left(\frac{1}{2}\right)^M\exp(-2r_k|\tau|)\\
&\qquad\times\sum_{k=0}^{M/2-1}\frac{2k-M}{M}\\
&\qquad\times\binom{M}{k}\sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j}\\
&\qquad\times B^{M-i-j}f(M-i-j)\\
&\qquad\times\left[(-1)^{M-k-j}+(-1)^{k-i}\right]\Bigg\}
\end{aligned}\tag{A.16}$$

in which

$$\begin{aligned}
f(x)&=\frac{\Gamma((x+1)/2)}{\sqrt{\pi}\Gamma(x/2+1)},\\
\Gamma(x)&=\int_0^\infty t^{x-1}e^{-t}dt.
\end{aligned}\tag{A.17}$$

Then

$$R_z(\tau) = x_m^2 \left\{ -\exp(-2r_k|\tau|)\frac{1}{2}B^2 + \frac{1}{2}B^2\cos(\Omega\tau) - \exp(-2r_k|\tau|)F(M) \right\}, \quad (A.18)$$

where

$$F(M) = \left(\frac{1}{2}\right)^M \sum_{k=0}^{M/2-1} \frac{2k-M}{M}\binom{M}{k} \times \sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j} \times B^{M-i-j} f(M-i-j) \times \left[(-1)^{M-k-j} + (-1)^{k-i}\right]. \quad (A.19)$$

Using Fourier transform of (A.18), we obtain the power spectrum density under the condition that M is even. Consider

$$\begin{aligned} S(\omega) &= \frac{4r_k x_m^2}{4r_k^2+\omega^2}\left(-\frac{1}{2}B^2\right) + \frac{\pi}{2}x_m^2 B^2 \left[\delta(\omega-\Omega)+\delta(\omega+\Omega)\right] - \frac{4r_k x_m^2}{4r_k^2+\omega^2}F(M) \\ &= -\left[\frac{1}{2}\left(\frac{A_0 x_m}{D}\right)^2 \frac{4r_k^2}{4r_k^2+\Omega^2} + F(M)\right]\frac{4r_k x_m^2}{4r_k^2+\omega^2} + \frac{\pi}{2}\left(\frac{A_0 x_m}{D}\right)^2 \frac{4r_k^2 x_m^2}{4r_k^2+\Omega^2}\left[\delta(\omega-\Omega)+\delta(\omega+\Omega)\right]. \end{aligned} \quad (A.20)$$

If M is odd, the method is similar. Consider

$$\begin{aligned} R_z(\tau) &= \frac{1}{T_\Omega}\int_0^{T_\Omega} R_z(t+\tau,t)\,dt \\ &= \frac{1}{T_\Omega} \times \int_0^{T_\Omega} x_m^2 \times \Bigg\{ \left[-\exp(-2r_k|\tau|)\kappa^2(t) + \kappa(t+\tau)\right]\kappa(t) \\ &\quad - \left(\frac{1}{2}\right)^M \exp(-2r_k|\tau|) \times \sum_{k=0}^{(M-1)/2}\frac{2k-M}{M}\binom{M}{k} \times \sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j} \\ &\quad \times B^{M-i-j}\cos(\Omega t-\varphi)^{M-i-j} \times \left[(-1)^{M-i-j}+(-1)^{k-i}\right]\Bigg\}\,dt \\ &= x_m^2\Bigg\{-\frac{1}{2}\exp(-2r_k|\tau|)\frac{1}{2}B^2 + \frac{1}{2}B^2\cos(\Omega\tau) - \left(\frac{1}{2}\right)^M \exp(-2r_k|\tau|) \\ &\quad \times \sum_{k=0}^{(M-1)/2}\frac{2k-M}{M} \times \binom{M}{k}\sum_{i=0}^{k}\sum_{j=0}^{M-\mathrm{k}}\binom{k}{i}\binom{M-k}{j} \times B^{M-i-j}f(M-i-j) \times \left[(-1)^{M-i-j}+(-1)^{k-i}\right]\Bigg\}, \end{aligned} \quad (A.21)$$

in which $f(x)$ and $\Gamma(x)$ have the same definition in (A.17). Then

$$R_z(\tau) = x_m^2\left\{-\frac{1}{2}\exp(-2r_k|\tau|)B^2 + \frac{1}{2}B^2\cos(\Omega\tau) - \exp(-2r_k|\tau|)G(M)\right\}, \quad (A.22)$$

in which

$$G(M) = \left(\frac{1}{2}\right)^M \sum_{k=0}^{(M-1)/2}\frac{2k-M}{M}\binom{M}{k} \times \sum_{i=0}^{k}\sum_{j=0}^{M-k}\binom{k}{i}\binom{M-k}{j} \times B^{M-i-j}f(M-i-j) \times \left[(-1)^{M-i-j}+(-1)^{k-i}\right]. \quad (A.23)$$

The Fourier transform of power spectrum density is

$$\begin{aligned} S(\omega) &= \frac{4r_k x_m^2}{4r_k^2+\omega^2}\left(-\frac{1}{2}B^2\right) \\ &\quad + \frac{\pi}{2}x_m^2 B^2\left[\delta(\omega-\Omega)+\delta(\omega+\Omega)\right] \\ &\quad - \frac{4r_k x_m^2}{4r_k^2+\omega^2}G(M) \\ &= -\left[\frac{1}{2}\left(\frac{A_0 x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}+G(M)\right]\frac{4r_k x_m^2}{4r_k^2+\omega^2} \\ &\quad + \frac{\pi}{2}\left(\frac{A_0 x_m}{D}\right)^2\frac{4r_k^2 x_m^2}{4r_k^2+\Omega^2} \\ &\quad \times\left[\delta(\omega-\Omega)+\delta(\omega+\Omega)\right]. \end{aligned} \tag{A.24}$$

In conclusion, the power spectrum density of the output of the system is

$$\begin{aligned} &S(\omega) \\ &= -\left[\frac{1}{2}\left(\frac{A_0 x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}+H(M)\right]\frac{4r_k x_m^2}{4r_k^2+\omega^2} \\ &\quad + \frac{\pi}{2}\left(\frac{A_0 x_m}{D}\right)^2\frac{4r_k^2 x_m^2}{4r_k^2+\omega^2} \\ &\quad \times\left[\delta(\omega-\Omega)+\delta(\omega+\Omega)\right], \end{aligned} \tag{A.25}$$

in which

$$H(M) = \begin{cases} F(M), & M \text{ is even number} \\ G(M), & M \text{ is odd number.} \end{cases} \tag{A.26}$$

□

Proof of Theorem 1. In (A.25), we can easily separate an exponentially decaying branch due to randomness and a periodically oscillating tail driven by the periodic input signal. And as a matter of fact, power spectrum density of noise $S_N(\omega)$ is the product of the Lorentzian curve obtained with no input signal $A_0 = 0$ and a factor that depends on the forcing amplitude A_0, but it is smaller than unity. Then the first part of power spectrum density is caused by noise. Consider

$$S_N(\omega) = -\left[\frac{1}{2}\left(\frac{A_0 x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}+H(M)\right]\frac{4r_k x_m^2}{4r_k^2+\omega^2}, \tag{A.27}$$

based on the definition of output SNR in (5).

Then for the parallel bistable system with M components, the output SNR following the definition in (5) is

$$\text{SNR} = \frac{\pi(A_0 x_m/D)^2 r_k}{-(1/2)(A_0 x_m/D)^2\left(4r_k^2/\left(4r_k^2+\Omega^2\right)\right)-H(M)}. \tag{A.28}$$

□

B. Proof of Theorem 6

Proof of Theorem 6. The proof of Theorem 6 is similar to that of Theorem 1; an outline is provided as follows. The output in this system is

$$z(t) = \frac{x_1(t)+x_2(t)}{2}. \tag{B.1}$$

The pdfs of x_i for $i = 1, 2$ are

$$p_x(x_i, t \mid x_0, t_0) = n_+(t)\,\delta(x_i - x_m) + n_-(t)\,\delta(x_i + x_m). \tag{B.2}$$

Then the pdf of $z(t)$ is

$$p_z(z, t \mid x_0, t_0) = 2\left[p_x(2z, t \mid x_0, t_0) * p_x(2z, t \mid x_0, t_0)\right]. \tag{B.3}$$

Then

$$\begin{aligned} p_z(z, t \mid x_0, t_0) &= n_+(t)^2\delta(z - x_m) \\ &\quad + 2n_+(t)\,n_-(t)\,\delta(z) + n_-(t)^2\delta(z + x_m). \end{aligned} \tag{B.4}$$

According to the general definition of autocorrelation function, the autocorrelation function of $z(t)$ is

$$\begin{aligned} &R_z(t+\tau, t) \\ &\quad = E\left[z(t+\tau)\,z(t)\right] \\ &\quad = \iint_{-\infty}^{+\infty} z_1 z_2\, p_z(z_1, t+\tau \mid z_2, t) \\ &\qquad\qquad \times p_z(z_2, t \mid z_0, t_0)\, dz_1 dz_2 \\ &\quad = \iint_{-\infty}^{+\infty} z_1 z_2 \\ &\qquad\qquad \times\left[n_+(t+\tau \mid z_2, t)^2\delta(z_1 - x_m)\right. \\ &\qquad\qquad + 2n_+(t+\tau \mid z_2, t)\, n_-(t+\tau \mid z_2, t)\,\delta(z_1) \\ &\qquad\qquad \left. + n_-(t+\tau \mid z_2, t)^2\delta(z_1 + x_m)\right] \\ &\qquad\qquad \times\left[n_+(t \mid z_0, t_0)^2\delta(z_2 - x_m)\right. \\ &\qquad\qquad + 2n_+(t \mid z_0, t_0)\, n_-(t \mid z_0, t_0)\,\delta(z_2) \\ &\qquad\qquad \left. + n_-(t \mid z_0, t_0)^2\delta(z_2 + x_m)\right] dz_1 dz_2. \end{aligned} \tag{B.5}$$

Due to the property of delta function, we obtain

$$\begin{aligned}
&R_z(t+\tau,t)\\
&= x_m^2 n_+(t+\tau \mid x_m,t)^2 n_+(t \mid z_0,t_0)^2\\
&\quad - x_m^2 n_+(t+\tau \mid -x_m,t)^2 n_-(t \mid z_0,t_0)^2\\
&\quad - x_m^2 n_-(t+\tau \mid x_m,t)^2 n_+(t \mid z_0,t_0)^2\\
&\quad + x_m^2 n_-(t+\tau \mid -x_m,t)^2 n_-(t \mid z_0,t_0)^2\\
&= x_m^2 \{n_+(t \mid z_0,t_0)^2\\
&\qquad \times [2n_+(t+\tau \mid x_m,t) - 2n_+(t+\tau \mid -x_m,t)]\\
&\qquad + [1-2n_+(t+\tau \mid -x_m,t)]\\
&\qquad \times [1-2n_+(t \mid z_0,t_0)]\}.
\end{aligned} \tag{B.6}$$

Simplify it in the stationary limit $t_0 \to -\infty$,

$$\begin{aligned}
&\lim_{t_0\to-\infty} R_z(t+\tau,t)\\
&= x_m^2 \exp(-2r_k|\tau|)\left[\frac{1-\kappa(t)^2}{2}\right] + x_m^2\kappa(t+\tau)\kappa(t).
\end{aligned} \tag{B.7}$$

And to obtain the average autocorrelation function

$$\begin{aligned}
R_z(\tau) &= \left(\frac{1}{T_\Omega}\right)\int_0^{T_\Omega} R_z(t+\tau,t)\,dt\\
&= \frac{x_m^2}{2}\exp(-2r_k|\tau|)\left[1-\frac{1}{2}\left(\frac{A_0x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}\right]\\
&\quad + \frac{x_m^2}{2}\left(\frac{A_0x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}\cos(\Omega\tau).
\end{aligned} \tag{B.8}$$

The output power spectrum density is as follows:

$$\begin{aligned}
S(\omega) &= \frac{1}{2}\left[1-\frac{1}{2}\left(\frac{A_0x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}\right]\frac{4r_kx_m^2}{4r_k^2+\omega^2}\\
&\quad + \frac{\pi}{2}\left(\frac{A_0x_m}{D}\right)^2\frac{4r_k^2x_m^2}{4r_k^2+\Omega^2}[\delta(\omega-\Omega)+\delta(\omega+\Omega)].
\end{aligned} \tag{B.9}$$

The first part is due to the noise, and then the power spectrum density of noise is

$$S_N(\omega) = \frac{1}{2}\left[1-\frac{1}{2}\left(\frac{A_0x_m}{D}\right)^2\frac{4r_k^2}{4r_k^2+\Omega^2}\right]\frac{4r_kx_m^2}{4r_k^2+\omega^2}. \tag{B.10}$$

The system output SNR is

$$\mathrm{SNR} = 2\pi\left(\frac{A_0x_m}{D}\right)^2 r_k + O\left(A_0^4\right). \tag{B.11}$$

Omitting the high order items due to the weak signal, the SNR following the definition in (5) becomes

$$\mathrm{SNR} = 2\pi\left(\frac{A_0x_m}{D}\right)^2 r_k. \tag{B.12}$$

□

Conflict of Interests

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China, Grant no. 61102157, and the National Basic Research Program of China (973 Program), Grant no. 2013CB329003.

References

[1] R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *Journal of Physics A*, vol. 14, no. 11, pp. L453–L457, 1981.

[2] R. Benzi, G. Parisi, and A. Sutera, "Stochastic resonance in climatic change," *Tellus*, vol. 34, pp. 10–16, 1982.

[3] R. Benzi, G. Parisi, and A. Sutera, "A theory of s tochastic resonance in climatic change," *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 565–578, 1983.

[4] A. A. Saha and G. V. Anand, "Design of detectors based on stochastic resonance," *Signal Processing*, vol. 83, no. 6, pp. 1193–1212, 2003.

[5] S. M. Saberali and H. Amindavar, "Analytical investigation of stochastic resonance phenomenon in linear detectors and comparison to maximum likelihood detector with application in co-channel interferer channels," *IET Signal Processing*, vol. 7, no. 2, pp. 128–133, 2013.

[6] D. He, "Chaotic stochastic resonance energy detection fusion used in cooperative spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 2, pp. 620–627, 2013.

[7] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*, Wiley-Interscience, 2002.

[8] T. Okano, A. Kitagawa, and K. Miyakawa, "Array-enhanced coherence resonance and phase synchronization in a two-dimensional array of excitable chemical oscillators," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 76, no. 4, Article ID 046201, 6 pages, 2007.

[9] Y. Tang, W. Zou, J. Lu, and J. Kurths, "Stochastic resonance in an ensemble of bistable systems under stable distribution noises and nonhomogeneous coupling," *Physical Review E*, vol. 85, no. 4, Article ID 046207, 6 pages, 2012.

[10] D. Rousseau and F. Chapeau-Blondeau, "Constructive role of noise in signal detection from parallel arrays of quantizers," *Signal Processing*, vol. 85, no. 3, pp. 571–580, 2005.

[11] F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Noise-enhanced SNR gain in parallel array of bistable oscillators," *Electronics Letters*, vol. 42, no. 17, pp. 1008–1009, 2006.

[12] F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Stochastic resonance in a parallel array of nonlinear dynamical elements," *Physics Letters A*, vol. 372, no. 13, pp. 2159–2166, 2008.

[13] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Reviews of Modern Physics*, vol. 70, no. 1, pp. 223–287, 1998.

[14] T. Wellens, V. Shatokhin, and A. Buchleitner, "Stochastic resonance," *Reports on Progress in Physics*, vol. 67, no. 1, pp. 45–105, 2004.

[15] B. McNamara and K. Wiesenfeld, "Theory of stochastic resonance," *Physical Review A*, vol. 39, no. 9, pp. 4854–4869, 1989.

[16] P. Jung, U. Behn, E. Pantazelou, and F. Moss, "Collective response in globally coupled bistable systems," *Physical Review A*, vol. 46, no. 4, pp. R1709–R1912, 1992.

[17] M. Locher, D. Cigna, E. R. Hunt et al., "Stochastic resonance in coupled nonlinear dynamic elements," *Chaos*, vol. 8, no. 3, pp. 604–615, 1998.

[18] M. Locher, N. Chatterjee, F. Marchesoni, W. L. Ditto, and E. R. Hunt, "Noise sustained propagation: local versus global noise," *Physical Review E*, vol. 61, no. 5, pp. 4954–4961, 2000.

[19] F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Noise-enhanced snr gain in parallel array of bistable oscillators," *Electronics Letters*, vol. 42, no. 17, pp. 1008–1009, 2006.

[20] F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Stochastic resonance in a parallel array of nonlinear dynamical elements," *Physics Letters A*, vol. 372, pp. 2159–2166, 2008.

[21] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Reviews of Modern Physics*, vol. 70, no. 1, pp. 223–287, 1998.

[22] T. Wellens, V. Shatokhin, and A. Buchleitner, "Stochastic resonance," *Reports on Progress in Physics*, vol. 67, pp. 45–105, 2004.

[23] C. Nicolis, "Stochastic aspects of climatic transitions response to a periodic forcing," *Tellus*, vol. 34, no. 1, pp. 1–9, 1982.

[24] B. McNamara and K. Wiesenfeld, "Theory of stochastic resonance," *Physical Review A*, vol. 39, pp. 4854–4869, 1989.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 539768, 14 pages
http://dx.doi.org/10.1155/2014/539768

Research Article

Equilibrium Model of Discrete Dynamic Supply Chain Network with Random Demand and Advertisement Strategy

Guitao Zhang,[1] Qingmei Sui,[2] Jinsong Hu,[1] Yongguang Zhong,[1] and Hao Sun[1]

[1] *School of Management Science and Engineering, Qingdao University, Qingdao 266071, China*
[2] *School of Control Science and Engineering, Shandong University, Jinan 250062, China*

Correspondence should be addressed to Jinsong Hu; hujinsong@qdu.edu.cn

Received 25 April 2014; Accepted 17 July 2014; Published 26 August 2014

Academic Editor: Ming Gao

The advertisement can increase the consumers demand; therefore it is one of the most important marketing strategies in the operations management of enterprises. This paper aims to analyze the impact of advertising investment on a discrete dynamic supply chain network which consists of suppliers, manufactures, retailers, and demand markets associated at different tiers under random demand. The impact of advertising investment will last several planning periods besides the current period due to delay effect. Based on noncooperative game theory, variational inequality, and Lagrange dual theory, the optimal economic behaviors of the suppliers, the manufactures, the retailers, and the consumers in the demand markets are modeled. In turn, the supply chain network equilibrium model is proposed and computed by modified project contraction algorithm with fixed step. The effectiveness of the model is illustrated by numerical examples, and managerial insights are obtained through the analysis of advertising investment in multiple periods and advertising delay effect among different periods.

1. Introduction

In the 1980s, the interest in supply chain and supply chain management increased tremendously. Supply chain management, which incorporates the raw materials supplying, production and distribution in the demand markets in the end [1], is a hot topic in the academic world as well as the business community. There is abundance of research available on the supply chain management. We refer the readers to the work of [2] to achieve a comprehensive review on the supply chain topic.

These researches mainly focus on the stringy supply chain or a single manufacturer. In fact, the supply chain is a network which consists of suppliers, manufacturers, retailers, and demand markets [3]. Thus, there is limited contribution in the previous literature that addresses the competition between the players with the same function, such as various manufacturers making the homogenous products, and the complexity resulting from so many actors in the supply chain network system. By the concept of equilibrium, Nagurney et al. [4] explore in the general supply chain network setting. Other researchers expand the work of Nagurney et al. [4]. In particular, Dong et al. [5] study the supply chain network equilibrium with stochastic market demand which need get the density function or distribution function of random demand from history data.

In practice, demand uncertainties arise from the complexity and the evolvement over time of supply chain network which is actually a dynamic system [6]. The dynamics of our world results in the changing of network construct; thus we can discrete the fixed time into several planning periods, and in one planning period, the parameters in the network are stable, whereas in different periods, there are some changes such as the raw materials price fluctuation or the demand parameter transformation in the markets. In this paper, we model the discrete dynamic supply chain network equilibrium.

Moreover, in order to promote the product, firms usually use some marketing strategies such as advertising. Advertising is a common marketing activity and is widely used by enterprises. Local advertising, which focuses on the local market, is mainly accomplished by the cooperation between manufacturer and retailer [7]. Since the retailer is closer and familiar with the consumers, she may have an efficient local advertising channel, and the manufacturer may provide the retailer a part of money for local advertising purpose. Warner Brothers, a maker of corsets, issued the first co-op agreement in 1903 [7]. From then on, the use of co-op advertising spreads to other industrials such as grocery stores and fashion, and the automobile is the most common user of cooperative advertising today.

The advocating of advertisement could make consumers learn about the characters and related knowledge of the products provided by manufacturers and retailers, so more consumers will purchase this product, which result in the total market share increasing. If we consider the advertisement strategy in a dynamic decision context, then the relationship among different periods must be taken into consideration. For example, the advertising investment in the current period also has some effects in the next periods, and this effect will reduce over time. This paper incorporates the co-op advertising investment strategy in discrete dynamic decision-making environment, and the investment will be shared by manufacturers and retailers; the sharing ratio is determined by negotiation between the two tire players. As we see in the numerical examples, it is interesting to note that the value of ratio does not impact the equilibrium results. Since the advertising strategy is an option that is underutilized, enterprises are unsure about the economic performance of advertising investment.

To mitigate the ambiguity about advertising investment for decision makers, in the paper, we model the role of advertising investment in a supply chain network over time. Similar to literatures of supply chain network, we assume the players in the same tier such as all manufacturers compete in a noncooperative fashion and the players in different tiers such as manufacturers and retailers must cooperate in order to agree with each other in transaction price and amount. In the network, decision makers including manufacturers and retailers need to decide on the appreciation level of advertisement investment so that they sell more products to demand markets to maximize the profit. To simplify problem studied, we will illustrate this point through numerical examples and consider the investment levels as a constant instead of a decision variable.

This paper is organized as follows. Section 2 gives assumptions and notations. In Section 3, we model the optimal behaviors of various players in supply chain network. In turn, we establish the equilibrium model of the whole network. Section 4 provides solution algorithm for the model established, and in Section 5, we illustrate the effective and managerial insights by numerical examples. Finally, in Section 6, we conclude the paper.

2. Literature Review

Over the past decades, in the context of supply chain, advertising strategy has grown up and becomes an important research topic in operations research and management area. Cooperative advertising generally has five different meanings [23]. In our research, we employ the first one that is vertical cooperative advertising which is also the most common comprehension. The manufacturers offer to share a certain percentage of the downstream retailers' advertising expenditures [24]. We also refer the readers to the work of [23] and the literature therein to get a general review about advertising. Based on the time dependence of parameters and decision variables, Lei et al. [25] and Xiao et al. [26] propose various multiperiod models to illustrate the impact of advertising investment on supply chain, whereas Chen [27], He et al. [28], Tsao and Sheen [29], and Xiao et al. [26] pick up the topic of stochastic environment associated with advertisement. Using game theoretic methods and from two main parts, simple marketing channels and a more complex structure, Jogensen and Zaccour [30] survey the literature on cooperative advertising in marketing channels (supply chains). Considering corporate social responsibility, Zhang et al. [31] examine the effectiveness of an advertising initiative in a leader-follower supply chain with one manufacturer and one retailer. Lambertini [32] characterizes an optimal two-part tariff specified as a linear function of the upstream firm's advertising effort, performing this task both in the static and in the dynamic games. It is necessary to point out that these researches mainly pay attention to the simple supply chain or a single firm but do not consider the complexity and the mutual impacts among firms in the supply chain network.

Besides the research of Dong et al. [5], Nagurney et al. [8], Nagurney and Toyasaki [9], Wu et al. [10], Hammond and Beullens [11], Yang et al. [12], Masoumi et al. [13], and Yu and Nagurney [15], Toyasaki et al. [16] study the supply chain network equilibrium problems from various perspectives and different supply chain networks. Qiang et al. [14] establish a closed-loop supply chain network model considering the competition, distribution channel investment, and demand uncertainties. The literatures mentioned above deal with static supply chain or static supply chain network equilibrium problems.

Recently, a few authors explore supply chain network equilibrium problems in dynamic setting. For example, Cruz and Wakolbinger [17] develop a framework for the analysis of the optimal levels of corporate social responsibility (CSR) activities in a multiperiod supply chain network consisting of manufacturers, retailers, and consumers and describe the problem of carbon emissions. Daniele [18] considers a supply chain network model with three tiers of decision makers (manufacturers, retailers, and consumers) in the case when prices and shipments are evolving on time. Cruz and Liu [19] analyze the effects of levels of social relationship on a multiperiod supply chain network with multiple decision makers associated at different tiers. Hamdouch [20] establishes a three-tier equilibrium model with capacity constraints and retailers' purchase strategy from a multiperiod perspective. Liu and Cruz [21] provide an analytical framework to investigate how financial risks affect the values of interconnected supply chain firms from a network perspective and how financial risks affect the supply chain firms' profitability and the cash and credit transactions. Feng et al. [22] develop a closed-loop supply chain super network model in which

Table 1: Literature sources for network equilibrium.

No.	Authors	Static/dynamic	Demand characteristic	Considering factor
1	Nagurney et al. (2002) [4]	Static	Deterministic	No
2	Dong et al. (2004) [5]	Static	Random uncertainty	No
3	Nagurney et al. (2005) [8]	Static	Random uncertainty	B2B transaction, risk
4	Nagurney and Toyasaki (2005) [9]	Static	Deterministic	No
5	Wu et al. (2006) [10]	Static	Deterministic	Pollution tax
6	Hammond and Beullens (2007) [11]	Static	Deterministic	Collection
7	Yang et al. (2009) [12]	Static	Deterministic	Collection
8	Masoumi et al. (2012) [13]	Static	Deterministic	No
9	Qiang et al. (2013) [14]	Static	Random uncertainty	Channel investment
10	Yu and Nagurney (2013) [15]	Static	Deterministic	No
11	Toyasaki et al. (2014) [16]	Static	Deterministic	No
12	Cruz and Wakolbinger (2008) [17]	Discrete dynamic	Deterministic	Corporate social responsibility
13	Daniele (2010) [18]	Continuous dynamic	Deterministic	No
14	Cruz and Liu (2011) [19]	Discrete dynamic	Deterministic	Social relationship
15	Hamdouch (2011) [20]	Discrete dynamic	Deterministic	Purchase strategy
16	Liu and Cruz (2012) [21]	Discrete dynamic	Deterministic	Corporate financial risks, trade credits
17	Feng et al. (2014) [22]	Continuous dynamic	Deterministic	Channel investment

the demand is seasonal and the manufacturers invest the reverse distribution channel for advocating consumers to return more end-of-life products.

The metamorphosis of supply chain network equilibrium literature of recent years is reviewed in Table 1. From Table 1 and literature survey, it is clearly evident that there is no research on discrete dynamic supply chain network equilibrium with advertising strategy and demand uncertainties.

In this paper, our model captures the planning process and the change of costs and demands and highlights the performance of advertising with delay effect, and moreover, this model expresses the uncertainties popularly existing in practice.

3. Model Assumptions and Notations

3.1. Model Assumptions. We consider a supply chain network consisting of S suppliers, M manufacturers, N retailers, and K demand markets and let s denote a typical supplier, m a typical manufacturer, n a typical retailer, and k a typical demand market; a retailer is matching a demand market; that is, one retailer only deals with the demand of one demand market. All actors in the same tire compete in a noncooperative fashion. Figure 1 illustrates the simple supply chain network with 2 suppliers, 2 manufacturers, 2 retailers, and 2 demand markets in 2 periods. $s_1(1)$ denotes the first supplier in the first period, and $s_2(1)$ denotes the second supplier in the first period; the other notations can be explained in the same way. The real lines between two adjacent tiers denote the related transaction activities, and the dash lines between 2 periods denote inventory transferring from the former period to the latter period.

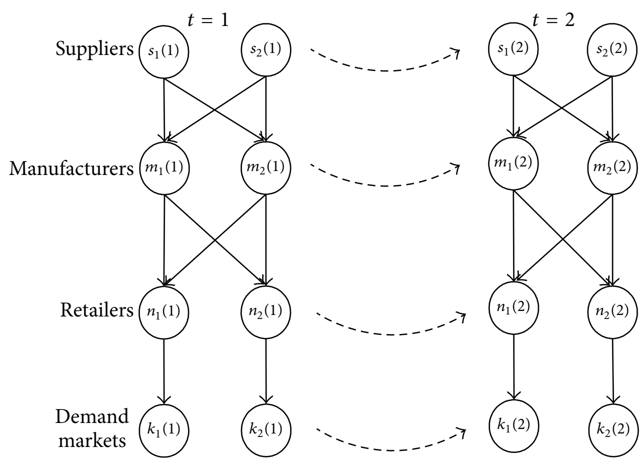


Figure 1: An illustration of 2-period supply chain network.

In order to explicate the problem studied, we give the following assumptions:

(1) All vectors are column vectors;

(2) The equilibrium solution or the optimal value of a decision variable is denoted by "$*$";

(3) The advertising investment is a constant and shared between the pairs of manufacturer and retailer;

(4) All cost functions and transaction functions are continuous convex and differentiable;

(5) All players in the network are risk neutral.

Table 2: Basic parameters in the closed-loop supply chain network figure.

Notation	Definition
β^r	Raw material conversion rate
t	A typical period, $t = 1, 2, \ldots, T$
s	A typical supplier, $s = 1, 2, \ldots, S$
m	A typical manufacturer, $m = 1, 2, \ldots, M$
n	A typical retailer, $n = 1, 2, \ldots, N$
k	A typical demand market, $k = 1, 2, \ldots, K$
$\lambda_n^-(>0)$	The unit cost of product shortage of retailer n
$\lambda_n^+(>0)$	The unit cost of product excess of retailer n
ϕ_{mn}	Advertising investment ratio shared by manufacturer m

Table 3: Transactions and production variables associated with various players in the network.

Notation	Definition
$q_{sm}^r(t)$	The raw material transaction volume from supplier s to manufacturer m at period t; group all of these variables into a column vector $Q^1 \in R_+^{SMT}$
$q_s^r(t)$	The total raw material volume provided by supplier s to all manufacturers at period t; group all of these variables into a column vector $q_1^r \in R_+^{ST}$
$q_m^r(t)$	The total raw material volume of manufacturer m used to produce at period t; group all of these variables into a column vector $q^r \in R_+^{MT}$
$q_{mn}(t)$	The product transaction volume from manufacturer m to retailer n at period t; group all of these variables into a column vector $Q^2 \in R_+^{MNT}$
$\rho_{mn}(t)$	The transaction price charged by manufacturer m for retailer n at period t
$\rho_n(t)$	Price charged by retailer n to the product in his outlet for corresponding demand market k at period t
$\phi_n(x; \rho_n(t))$	The density function of random variable x
$\Phi_n(x; \rho_n(t))$	The distribution function of random variable x
$I_m(t)$	The inventory of manufacturer m at period t; group all of $I_m(t)$ into a column vector $I \in R_+^{MT}$

3.2. Variables and Notations. The variables and notations are defined as in Tables 2 and 3, and the production functions and transaction functions are defined as in Table 4.

4. Discrete Dynamic Supply Chain Network Equilibrium Model

4.1. The Optimal Behavior and Equilibrium Condition of Suppliers. In each period, supplier s provides raw material to various manufacturers at the beginning of every period and makes decision associated with trade and production volume of raw material to maximize the profit in the entire planning horizon. Using the notations defined previously, the profit maximum criterion for supplier s can be described as

$$\pi_s = \max \left\{ \sum_{t=1}^{T} \sum_{m=1}^{M} \rho_{sm}^r(t)\, q_{sm}^r(t) - \sum_{t=1}^{T} \sum_{m=1}^{M} c_{sm}^r(t) - \sum_{t=1}^{T} f_s^r\left(q_1^r(t)\right) \right\}, \quad (1)$$

$$\text{s.t.} \quad \sum_{m=1}^{M} q_{sm}^r(t) \le q_s^r(t), \quad (2)$$

$$\left(q_{sm}^r(t), q_s^r(t)\right) \in R_+^{(M+1)T}, \quad \forall s. \quad (3)$$

Equation (2) expresses that production output of raw material cannot be lower than total volume of the raw material transaction between the supplier s and the various manufacturers.

In this paper, we assume that all the suppliers compete in a noncooperative fashion. Therefore, we can simultaneously express the equilibrium condition of the suppliers as the variational inequality, determining $(q_1^{r*}, Q^{1*}, \eta_s^*) \in \Omega^S$, such that

$$\begin{aligned} &\sum_{t=1}^{T} \sum_{s=1}^{S} \left[\frac{\partial f_s^r\left(q_s^{r*}(t)\right)}{\partial q_s^r(t)} - \eta_s^*(t) \right] \times \left[q_s^r(t) - q_s^{r*}(t) \right] \\ &\quad + \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{m=1}^{M} \left[\frac{\partial c_{sm}^{r*}(t)}{\partial q_{sm}^r(t)} - \rho_{sm}^{r*}(t) + \eta_s^*(t) \right] \times \left[q_{sm}^r(t) - q_{sm}^{r*}(t) \right] \\ &\quad + \sum_{t=1}^{T} \sum_{s=1}^{S} \left[q_s^{r*}(t) - \sum_{m=1}^{M} q_{sm}^{r*}(t) \right] \times \left[\eta_s(t) - \eta_s^*(t) \right] \ge 0 \\ &\qquad \forall \left(q_1^r, Q^1, \eta_s\right) \in \Omega^S, \end{aligned} \quad (4)$$

where $\Omega^S = R_+^{ST+SMT+ST}$.

Table 4: Functions associated with various players in the network.

Notations	Definition
$f_s^r(t) = f_s^r(q_1^r(t))$	The raw material production cost function of supplier s at period t
$c_{sm}^r(t) = c_{sm}^r(q_{sm}^r(t))$	The transaction cost function between supplier s and manufacturer m at period t
$f_m^M(t) = f_m^M(\beta_r, q^r(t))$	The production cost function using raw materials of manufacturer m at period t
$c_n(t)$	The exhibition and disposal cost at retailer n at period t
$c_{mn}(t) = c_{mn}(q_{mn}(t))$	The transaction cost function between manufacturer m and retailer n at period t
$H_m(t) = H_m(I_m(t))$	The inventory cost function at manufacturer m
f_t^{t+i}	The delay effect factor of advertising investment at period t on the period $t+i$
$d_k(\rho_k(t), I_A^{mn}(t))$	The demand function associated with demand market k

In (4), $\eta_s(t)$ is the Lagrange multiplier corresponding to constraint (2) and $\eta_s \in R_+^{ST}$ is the column vector with the elements of $\eta_s(t)$.

Based on the equivalence of variational inequality and complement problem, from the second term of (4), we get

$$\rho_{sm}^{r*}(t) = \frac{\partial c_{sm}^{r*}(t)}{\partial q_{sm}^r(t)} + \eta_s^*(t). \tag{5}$$

From the 1st term of (4), in the equilibrium state, $\eta_s^*(t) = \partial f_s^r(q_s^{r*}(t))/\partial q_s^r(t)$; that is, $\eta_s^*(t)$ is equal to the marginal production cost. Therefore, (5) shows that the transaction price between suppliers and manufacturers is equal to the sum of marginal transaction cost and marginal production cost.

4.2. The Optimal Behavior and Equilibrium Condition of Manufacturers. The manufacturers purchase the raw materials from various suppliers to make products and sell the new products to retailers at every period and in the same time manage inventory between periods according to the market conditions. The manufacturer m seeks to maximize her profit that can be described as follows:

$$\begin{aligned}\pi_m = \max \Bigg\{ &\sum_{t=1}^{T}\sum_{n=1}^{N}\rho_{mn}(t)\,q_{mn}(t) - \sum_{t=1}^{T}\sum_{s=1}^{S}\rho_{sm}(t)\,q_{sm}(t) \\ &- \sum_{t=1}^{T} f_m^M(t) - \sum_{t=1}^{T}\sum_{n=1}^{N} c_{mn}(t) - \sum_{t=1}^{T} H_m(t) \\ &- \sum_{t=1}^{T}\sum_{n=1}^{N}\phi_{mn} I_A^{mn}(t)\,q_{mn}(t)\Bigg\},\end{aligned} \tag{6}$$

$$\text{s.t.}\quad I_m(t-1) + \beta_r q_m^r(t) = I_m(t) + \sum_{n=1}^{N} q_{mn}(t), \tag{7}$$

$$q_m^r(t) \le \sum_{s=1}^{S} q_{sm}^r(t). \tag{8}$$

Equation (7) expresses the flow conservation; the sum of production volume from raw materials in t period and the transferring inventory from $t-1$ period is equal to the sum of the transaction volume with all retailers and the transferring inventory to next period, and assume the corresponding Lagrange multiplier is $\lambda_m(t)$; $\lambda \in R^{MT}$ is the column vector with the elements of $\lambda_m(t)$. Equation (8) shows that the raw materials amount obtained in manufacturer m is not higher than that various suppliers sent to her; similarly, assume the corresponding Lagrange multiplier is $\gamma_m(t)$and $\gamma \in R_+^{MT}$ is the column vector with the elements of $\gamma_m(t)$.

The profit maximum object of all manufacturers can be described as a variational inequality, determining $(q^{r*}, Q^{1*}, Q^{2*}, I^*, \gamma^*, \lambda^*) \in \Omega^M$, such that

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{m=1}^{M}\left[\frac{\partial f_m^{M*}(t)}{\partial q_m^r(t)} - \beta_r\lambda_m^*(t) + \gamma_m^*(t)\right]\times\left[q_m^r(t) - q_m^{r*}(t)\right] \\ &+ \sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{s=1}^{S}\left[\rho_{sm}^*(t) - \gamma_m^*(t)\right]\times\left[q_{sm}(t) - q_{sm}^*(t)\right] \\ &+ \sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[\frac{\partial c_{mn}^*(t)}{\partial q_{mn}(t)} + \lambda_m^*(t) - \rho_{mn}^*(t) + \phi_{mn} I_A^{mn}(t)\right] \\ &\qquad\times\left[q_{mn}(t) - q_{mn}^*(t)\right] \\ &+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[\frac{\partial H_m^*(t)}{\partial I_m(t)} + \lambda_m^*(t) - \lambda_m^*(t+1)\right] \\ &\qquad\times\left[I_m(t) - I_m^*(t)\right] \\ &+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[\sum_{s=1}^{S} q_{sm}^{r*}(t) - q_m^{r*}(t)\right]\times\left[\gamma_m(t) - \gamma_m^*(t)\right] \\ &+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[I_m^*(t-1) + \beta_r q_m^{r*}(t) - I_m^*(t) - \sum_{n=1}^{N} q_{mn}^*(t)\right] \\ &\qquad\times\left[\lambda_m(t) - \lambda_m^*(t)\right] \ge 0 \\ &\qquad\forall\left(q^r, Q^1, Q^2, I, \gamma, \lambda\right) \in \Omega^M,\end{aligned} \tag{9}$$

where $\Omega^M = R_+^{MT+SMT+MNT+2MT} \times R^{MT}$.

From the third term of (9), the transaction price can be written as when the network is in equilibrium:

$$\rho_{mn}^*(t) = \frac{\partial c_{mn}^*(t)}{\partial q_{mn}(t)} + \lambda_m^*(t) + \phi_{mn} I_A^{mn}(t). \tag{10}$$

From the 2nd term of (9), in the equilibrium state, we get $\rho_{sm}^*(t) = \gamma_m^*(t)$; then from the 1st term, we get $\lambda_m^*(t) =$

$(1/\beta_r)[\partial f_m^{M*}(t)/\partial q_m^r(t) + \gamma_m^*(t)] = (1/\beta_r)[\partial f_m^{M*}(t)/\partial q_m^r(t) + \rho_{sm}^*(t)]$. Equation (10) shows that in the equilibrium state, the transaction price between manufacturers and retailers is equal to the sum of marginal transaction cost between manufacturers and retailers, the Lagrange multiplier corresponding to constraint (7), and the advertisement investment amount shared by manufacturer m.

4.3. The Optimal Behavior and Equilibrium Condition of Retailers. The retailers need to decide to purchase how many products from manufacturers and sell to consumers in corresponding demand markets in a certain price.

Due to $\hat{d}_n(I_A^{mn}(t), \rho_n(t))$ denoting the random demand of retailer outlet n, the demand depends on the advertising investment and the trade price; it is obvious that the more advertising investment paid by manufacturers and retailers is, the larger consumer demand is, whereas the increase of price charged by retailers will lower the product demand. For a given product transaction price $\rho_n(t)$ at period t, according to the notations illustrated in Table 3, $\Phi_n(x; \rho_n(t)) = \int_0^x \phi_n(x; \rho_n(t))dx$. Let $s_n(t)$ denote the wholesale amount from manufacturers and $s_n(t) = \sum_{m=1}^M q_{mn}(t)$; group all $s_n(t)$ in period t into a column vector $s(t) \in R_+^N$, and group all $s_n(t)$ into a column vector $s_n \in R_+^{NT}$. In order to express the competition among retailers, we assume that the exhibition function and disposal cost function at retailer n $c_n(t) = c_n(s(t))$ are related with all retailers.

For retailer n, if given $s_n(t)$, it is similar as in Dong et al. [5] and Nagurney et al. [4], the expected sales quantity, expected shortage quantity, and expected exceed quantity can be expressed as

$$\begin{aligned}
&S_n(s_n(t), I_A^{mn}(t), \rho_n(t)) \\
&\quad = E\left[\min\left\{\hat{d}_n(I_A^{mn}(t), \rho_n(t)), s_n(t)\right\}\right] \\
&\quad = s_n(t) - \int_0^{s_n(t)} (s_n(t) - x)\, d\Phi_n \\
&\qquad \times (x, I_A^{mn}(t), \rho_n(t)), \\
&H_n(s_n(t), I_A^{mn}(t), \rho_n(t)) \\
&\quad = E\left[\max\left\{0, s_n(t) - \hat{d}_n(I_A^{mn}(t), \rho_n(t))\right\}\right] \\
&\quad = \int_0^{s_n} (s_n - x)\, d\Phi_n(x, I_A^{mn}(t), \rho_n(t)), \\
&Q_n(s_n(t), I_A^{mn}(t), \rho_n(t)) \\
&\quad = E\left[\max\left\{0, \hat{d}_n(I_A^{mn}(t), \rho_n(t)) - s_n(t)\right\}\right] \\
&\quad = \int_{s_n(t)}^{+\infty} (x - s_n(t))\, d\Phi_j(x, I_A^{mn}(t), \rho_n(t)).
\end{aligned} \tag{11}$$

From (10), we can easily obtain

$$\begin{aligned}
\frac{\partial S_n(s_n(t), I_A^{mn}(t), \rho_n(t))}{\partial s_n(t)} &= 1 - \Phi_n(s_n(t), I_A^{mn}(t), \rho_n(t)), \\
\frac{\partial H_n(s_n(t), I_A^{mn}(t), \rho_n(t))}{\partial s_n(t)} &= \Phi_n(s_n(t), I_A^{mn}(t), \rho_n(t)), \\
\frac{\partial Q_n(s_n(t), I_A^{mn}(t), \rho_n(t))}{\partial s_n(t)} &= \Phi_n(s_n(t), I_A^{mn}(t), \rho_n(t)) - 1.
\end{aligned} \tag{12}$$

For retailer n, the maximum expected profit model can be expressed as

$$\begin{aligned}
\pi_n = \max \Bigg\{ &\sum_{t=1}^T \rho_n(t) S_n(s_n(t), I_A^{mn}(t), \rho_n(t)) \\
&- \lambda_n^+ \sum_{t=1}^T H_n(s_n(t), I_A^{mn}(t), \rho_n(t)) \\
&- \lambda_n^- \sum_{t=1}^T Q_n(s_n(t), I_A^{mn}(t), \rho_n(t)) \\
&- \sum_{t=1}^T c_n(s(t)) - \sum_{t=1}^T \sum_{m=1}^M \rho_{mn}(t) q_{mn}(t) - (1 - \phi_{mn}) \\
&\times \sum_{t=1}^T \sum_{m=1=1}^M I_A^{mn}(t) q_{mn}(t) \Bigg\},
\end{aligned} \tag{13}$$

$$\text{s.t.} \quad s_n(t) = \sum_{m=1}^M q_{mn}(t). \tag{14}$$

Using (11) and (13) can be rewritten as

$$\begin{aligned}
\pi_n = \max \Bigg\{ &\sum_{t=1}^T \rho_n(t) s_n(t) - \sum_{t=1}^T (\rho_n(t) + \lambda_n^+) \\
&\times \int_0^{s_n(t)} (s_n(t) - x)\, d\Phi_n(x, \rho_n(t)) - \lambda_n^- \\
&\times \sum_{t=1}^T \int_{s_n}^{+\infty} (x - s_n(t))\, d\Phi_n(x, \rho_n(t)) \\
&- \sum_{t=1}^T c_n(s(t)) - \sum_{t=1}^T \sum_{m=1}^M \rho_{mn}(t) q_{mn}(t) - (1 - \phi_{mn}) \\
&\times \sum_{t=1}^T \sum_{m=1}^M I_A^{mn}(t) q_{mn}(t) \Bigg\}.
\end{aligned} \tag{15}$$

All retailers compete in a noncooperation fashion; using (12), their equilibrium conditions can be described as a variational inequality, determining $(s_n^*, Q^{2*}, \theta^*) \in \Omega^N$, such that

$$\sum_{t=1}^{T}\sum_{n=1}^{N}\Bigg[(\rho_n^*(t)+\lambda_n^+ +\lambda_n^-)\,\Phi_n(s_n^*(t),\rho_n^*(t)) - \rho_n^*(t)-\lambda_n^- + \frac{\partial c_n(s^*(t))}{\partial s_n(t)}\Bigg]\times[s_n(t)-s_n^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{n=1}^{N}[(1-\phi_{mn})I_A^{mn}(t)+\rho_{mn}^*(t)] \times[q_{mn}(t)-q_{mn}^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{n=1}^{N}\left[\sum_{m=1}^{M}q_{mn}^*(t)-s_n^*(t)\right]\times[\theta_n(t)-\theta_n^*(t)]\geq 0$$
$$\forall\left(s_n, Q^2,\theta\right)\in\Omega^N, \tag{16}$$

where $\Omega^N = R_+^{NT+MNT}\times R^{NT}$.

In (16), $\theta_n(t)$ is the Lagrange multiplier corresponding to constraint (14) and $\theta_n(t)\in R^{NT}$ is the column vector with the elements of $\theta_n(t)$. The transaction price $\rho_n^*(t)$ is a decision variable which can be obtained from the computing results.

4.4. The Optimal Behavior and Equilibrium Condition of Demand Markets. For the supply chain network, given a fixed advertising investment, the consumers of demand markets buy the products under a price charged by the retailers and it is similar as in Dong et al. [5] and Nagurney et al. [4]

$$d_n(I_A^{mn}(t),\rho_n^*(t))\begin{cases}= s_n^*(t), & \rho_n^*(t)>0\\ < s_n^*(t), & \rho_n^*(t)=0.\end{cases} \tag{17}$$

The consumers' optimal behaviors and equilibrium conditions can be described as a variational inequality, determining $\rho_n^*(t)\in\Omega^K$, such that

$$[s_n^*(t)-d_n(I_A^{mn}(t),\rho_n^*(t))]\times[\rho_n(t)-\rho_n^*(t)]\geq 0 \qquad \forall\rho_n(t)\in\Omega^K, \tag{18}$$

where $\Omega^K = R_+^{KT}$.

4.5. The Equilibrium Condition of the Supply Chain Network. Each player in the supply chain network selects the optimal strategy in every period and seeks to maximize the profit in the entire planning horizon on the basis of the other players making optimal decisions. Thus, the network will experience a strategy selecting process and carry out Nash equilibrium in the end. In particular, the product transaction amount and price between the adjacent tires must be equal to that the players want to purchase or sell at every period, and the manufacturers and retailers also need to make decisions about the advertising investment to enhance the expected sales to maximize their profits. So, the whole network equilibrium condition is the sum of (4), (9), (15), and (18). We sum up these equations and obtain the following theorem.

Theorem 1. *A strategy pattern $(q_1^{r*}, q^{r*}, Q^{1*}, Q^{2*}, I^*, s^*, \rho^*, \eta_s^*, \gamma^*, \lambda^*)\in\Omega$ of the discrete dynamic supply chain network can be called an equilibrium pattern if and only if it satisfies the following inequality, determining $(q_1^{r*}, q^{r*}, Q^{1*}, Q^{2*}, I^*, s_n^*, \rho^*, \eta_s^*, \gamma^*, \lambda^*, \theta^*)\in\Omega$, such that*

$$\sum_{t=1}^{T}\sum_{s=1}^{S}\left[\frac{\partial f_s^r(q_s^{r*}(t))}{\partial q_s^r(t)}-\eta_s^*(t)\right]\times[q_s^r(t)-q_s^{r*}(t)]$$
$$+\sum_{t=1}^{T}\sum_{m=1}^{M}\left[\frac{\partial f_m^{M*}(t)}{\partial q_m^r(t)}-\beta_r\lambda_m^*(t)+\gamma_m^*(t)\right]\times[q_m^r(t)-q_m^{r*}(t)]$$
$$+\sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{m=1}^{M}\left[\frac{\partial c_{sm}^{r*}(t)}{\partial q_{sm}^r(t)}+\eta_s^*(t)-\gamma_m^*(t)\right]\times[q_{sm}^r(t)-q_{sm}^{r*}(t)]$$
$$+\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[\frac{\partial c_{mn}^*(t)}{\partial q_{mn}(t)}+\lambda_m^*(t)-\theta_n^*(t)+I_A^{mn*}(t)\right]\times[q_{mn}(t)-q_{mn}^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{m=1}^{M}\left[\frac{\partial H_m^*(t)}{\partial I_m(t)}+\lambda_m^*(t)-\lambda_m^*(t+1)\right]\times[I_m(t)-I_m^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{n=1}^{N}\Bigg[(\rho_n^*(t)+\lambda_n^+ +\lambda_n^-)\,\Phi_n(s_n^*(t),\rho_n^*(t)) - \rho_n^*(t)-\lambda_n^- +\theta_n^*(t)+\frac{\partial c_n(s^*(t))}{\partial s_n(t)}\Bigg]\times[s_n(t)-s_n^*(t)]$$
$$+[s_n^*(t)-d_n(I_A^{mn*}(t),\rho_n^*(t))]\times[\rho_n(t)-\rho_n^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{s=1}^{S}\left[q_s^{r*}(t)-\sum_{m=1}^{M}q_{sm}^{r*}(t)\right]\times[\eta_s(t)-\eta_s^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{m=1}^{M}\left[\sum_{s=1}^{S}q_{sm}^{r*}(t)-q_m^{r*}(t)\right]\times[\gamma_m(t)-\gamma_m^*(t)]$$
$$+\sum_{t=1}^{T}\sum_{m=1}^{M}\Bigg[I_m^*(t-1)+\beta_r q_m^{r*}(t) - I_m^*(t)-\sum_{n=1}^{N}q_{mn}^*(t)\Bigg]$$

$$\times \left[\lambda_m(t) - \lambda_m^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{n=1}^{N}\left[\sum_{m=1}^{M} q_{mn}^*(t) - s_n^*(t)\right] \times \left[\theta_n(t) - \theta_n^*(t)\right] \ge 0$$

$$\forall \left(q_1^r, q^r, Q^1, Q^2, I, s, \rho, \eta_s, \gamma, \lambda, \theta\right) \in \Omega$$

$$+ \sum_{t=1}^{T}\sum_{n=1}^{N}\left[\sum_{m=1}^{M} q_{mn}^*(t) - s_n^*(t)\right] \times \left[\theta_n(t) - \theta_n^*(t)\right] \ge 0$$

$$\forall \left(q_1^r, q^r, Q^1, Q^2, I, s_n, \rho, \eta_s, \gamma, \lambda, \theta\right) \in \Omega, \tag{19}$$

where $\Omega = \Omega^S \times \Omega^M \times \Omega^N \times \Omega^K$.

Proof. Let us sum up (4), (9), (15), and (18); we get the total inequality, determining $(q_1^{r*}, q^{r*}, Q^{1*}, Q^{2*}, I^*, s_n^*, \rho^*, \eta_s^*, \gamma^*, \lambda^*, \theta^*) \in \Omega$, such that

$$\sum_{t=1}^{T}\sum_{s=1}^{S}\left[\frac{\partial f_s^r\left(q_s^{r*}(t)\right)}{\partial q_s^r(t)} - \eta_s^*(t)\right]$$

$$\times \left[q_s^r(t) - q_s^{r*}(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[\frac{\partial f_m^{M*}(t)}{\partial q_m^r(t)} - \beta_r \lambda_m^*(t) + \gamma_m^*(t)\right]$$

$$\times \left[q_m^r(t) - q_m^{r*}(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{m=1}^{M}\left[\frac{\partial c_{sm}^{r*}(t)}{\partial q_{sm}^r(t)} + \eta_s^*(t) - \gamma_m^*(t)\right.$$

$$\left. + \rho_{sm}^*(t) - \rho_{sm}^*(t)\right]$$

$$\times \left[q_{sm}^r(t) - q_{sm}^{r*}(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[\frac{\partial c_{mn}^*(t)}{\partial q_{mn}(t)} + \lambda_m^*(t) - \theta_n^*(t)\right.$$

$$+ \left(1 - \phi_{mn}\right) I_A^{mn*}(t) + \phi_{mn} I_A^{mn*}(t)$$

$$\left. - \rho_{mn}^*(t) + \rho_{mn}^*(t)\right]$$

$$\times \left[q_{mn}(t) - q_{mn}^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[\frac{\partial H_m^*(t)}{\partial I_m(t)} + \lambda_m^*(t) - \lambda_m^*(t+1)\right]$$

$$\times \left[I_m(t) - I_m^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{n=1}^{N}\left[\left(\rho_n^*(t) + \lambda_n^+ + \lambda_n^-\right)\Phi_n\left(s_n^*(t), \rho_n^*(t)\right)\right.$$

$$\left. - \rho_n^*(t) - \lambda_n^- + \theta_n^*(t) + \frac{\partial c_n\left(s^*(t)\right)}{\partial s_n(t)}\right]$$

$$\times \left[s_n(t) - s_n^*(t)\right]$$

$$+ \left[s_n^*(t) - d_n\left(I_A^{mn*}(t), \rho_n^*(t)\right)\right]$$

$$\times \left[\rho_n(t) - \rho_n^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{s=1}^{S}\left[q_s^{r*}(t) - \sum_{m=1}^{M} q_{sm}^{r*}(t)\right]$$

$$\times \left[\eta_s(t) - \eta_s^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[\sum_{s=1}^{S} q_{sm}^{r*}(t) - q_m^{r*}(t)\right]$$

$$\times \left[\gamma_m(t) - \gamma_m^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left[I_m^*(t-1) + \beta_r q_m^{r*}(t)\right.$$

$$\left. - I_m^*(t) - \sum_{n=1}^{N} q_{mn}^*(t)\right]$$

$$\times \left[\lambda_m(t) - \lambda_m^*(t)\right]$$

$$+ \sum_{t=1}^{T}\sum_{n=1}^{N}\left[\sum_{m=1}^{M} q_{mn}^*(t) - s_n^*(t)\right] \times \left[\theta_n(t) - \theta_n^*(t)\right] \ge 0$$

$$\forall \left(q_1^r, q^r, Q^1, Q^2, I, s_n, \rho, \eta_s, \gamma, \lambda, \theta\right) \in \Omega. \tag{20}$$

We simplify the 3rd and 4th terms in (20) and obtain (19). From (19), we note that the share ratio of advertising investment between manufacturers and retailers does not impact the network equilibrium results; therefore, determining the share ratio will be up to the power of two kinds of players in their bargain. □

5. Numerical Examples

In this section, we will provide some numerical examples to illustrate the efficiency of the previous equilibrium model and analyze the relevant parameters. To solve the model, there are several algorithms to choose, such as logarithmic-quadratic proximal prediction-correction method [33], modified contraction project method [34], smoothing Newton algorithm [35], and others, to name a few. In this paper, we employ the modified contraction project method to solve the variational inequality (19) for its simple steps and obtain the decision variables and Lagrange multiplexer simultaneously. Set the related parameters as follows: the initial value of decision variables and Lagrange multipliers is set to 1 and

Table 5: Cost functions for computational study.

Notation	Definition
$f_s(q_1^r(t)) = tq_s^r(t)^2 + q_s^r(t) + 1$	Cost function of producing raw materials for supplier s at period t
$c_{sm}(t) = q_{sm}(t)^2 + 1.5q_{sm}(t) + 1$	Transaction cost function undertaken by supplier s related to supply chain sm at period t
$f_m^M(t) = t(\beta_r q_m^r(t))^2 + 3\beta_r q_m^r(t) + 2$	Production cost function for manufacturer m at period t
$H_m(t) = tI_m(t)$	Inventory cost for manufacturer m at period t
$c_{mn}(t) = 5.5q_{mn}(t)^2 + 3q_{mn}(t) + 2$	Transaction cost function undertaken by manufacturer m related to supply chain mn at period t
$c_n(t) = 0.25\left(\sum_{m=1}^{2} q_{mn}(t)\right)^2$	Disposal costs at retailer n at period t

Table 6: Equilibrium results with delay effect of advertising investment.

Variables $t = 1, 2, 3$	$I_A^{mn}(1) = 0.15$ $I_A^{mn}(2) = 0$ $I_A^{mn}(3) = 0$	$I_A^{mn}(1) = 0$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0$	$I_A^{mn}(1) = 0$ $I_A^{mn}(2) = 0$ $I_A^{mn}(3) = 0.15$	$I_A^{mn}(1) = 0.15$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0$	$I_A^{mn}(1) = 0$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0.15$	$I_A^{mn}(1) = 0.15$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0.15$
$q_{sm}^r(t)$	0.7282	0.7259	0.7214	0.7407	0.7340	0.7488
$s = 1, 2$	0.4443	0.4431	0.4408	0.451	0.4474	0.4552
$m = 1, 2$	0.3821	0.3813	0.3797	0.3867	0.3842	0.3896
$q_s^r(t), q_m^r(t)$	1.4563	1.4519	1.4429	1.4814	1.4679	1.4976
$s = 1, 2$	0.8886	0.8863	0.8815	0.9019	0.8948	0.9105
$m = 1, 2$	0.7643	0.7627	0.7594	0.7733	0.7685	0.7791
$q_{mn}(t)$	0.5357	0.5174	0.5182	0.5336	0.5161	0.5322
$m = 1, 2$	0.5228	0.5313	0.5132	0.5399	0.5300	0.5386
$n = 1, 2$	0.4961	0.5017	0.5106	0.5048	0.5195	0.5228
$I_m(t)$	0.3849	0.4170	0.4066	0.4142	0.4358	0.4331
$m = 1, 2$	0.2280	0.2407	0.2617	0.2364	0.2706	0.2664
	0	0	0	0	0	0
π_s	3.3263	3.2928	3.2254	3.5170	3.4143	3.6415
π_m	14.3234	14.2429	14.0845	14.7572	14.5163	15.0348
π_n	27.2004	27.0716	26.7721	28.1674	27.7355	28.8402

the convergence criterion, for example, the absolute value of difference of decision variables and Lagrange multipliers between two steps is lower than or equal to 10^{-8}. We assume $\phi_{mn} = 0.4$, $\beta_r = 1$, $\lambda_n^- = 1$, $\lambda_n^+ = 1$, $I_A^{mn}(t) = 0.15$, $f_t^t = 0.2$, $f_t^{t+1} = 0.1$, and $f_t^{t+2} = 0.05$. The related cost functions and parameters are set as listed in Table 5. It is assumed that the random demands follow uniform distribution in, $d_k(\rho_k(t), I_A^{mn}(t)) \sim [0, b_k(t)/\rho_k(t)]$, $b_k(1) = 90(1 + I_A^{mn}(1))^{f_1^1}$, $b_k(2) = 93\prod_{i=0}^{1}(1 + I_A^{mn}(2-i))^{f_{2-i}^2}$, and $b_k(3) = 96\prod_{i=0}^{2}(1 + I_A^{mn}(3-i))^{f_{3-i}^3}$, for $k = 1, 2$, $m = 1, 2$, $n = 1, 2$, and $t = 1, 2, 3$.

This paper focuses on the analysis of the following four aspects: (1) the equilibrium results of advertising investment with delay effect and the results listed as in Table 6; (2) the equilibrium results of advertising investment with no delay effect, that is, $f_t^t = f_t^{t+1} = f_t^{t+2} = 0$, and the results listed as in Table 7; (3) the equilibrium results with one manufacturer advertising investment and the results listed as in Table 8; and (4) the profits of various players with the 1st period advertising investment increasing with/without delay effect, which is illustrated as in Figure 2.

From the first three columns in Table 6, we can find that in the case the advertising delay effect exists, the production volumes, the transaction volumes, and all the players' profits are the highest when the manufacturers and the retailers make advertisements in the 1st period, and then is the 2nd period, the lowest is the 3rd period.

From the latter three columns in Table 6, it can be seen that when advertising is in the 1st and 2nd periods, all the players' profits are higher than that in the 2nd and 3rd periods and lower than that in all the three periods, which implies that the earlier the advertisement is made, the higher profits the players can obtain.

We now turn to analyze the inventory between adjacent periods which describes the characteristic of discrete dynamic supply chain network. The manufacturers can adjust the inventory to maximize the profits in the whole planning horizon. Compare the 1st three columns of Table 6, because of the demand increasing in the 2nd period as a result of advertising, the inventory transfer from the 1st period to 2nd period increases; on the other hand, due to the delay effect, the advertising in the 2nd period also has much influence in

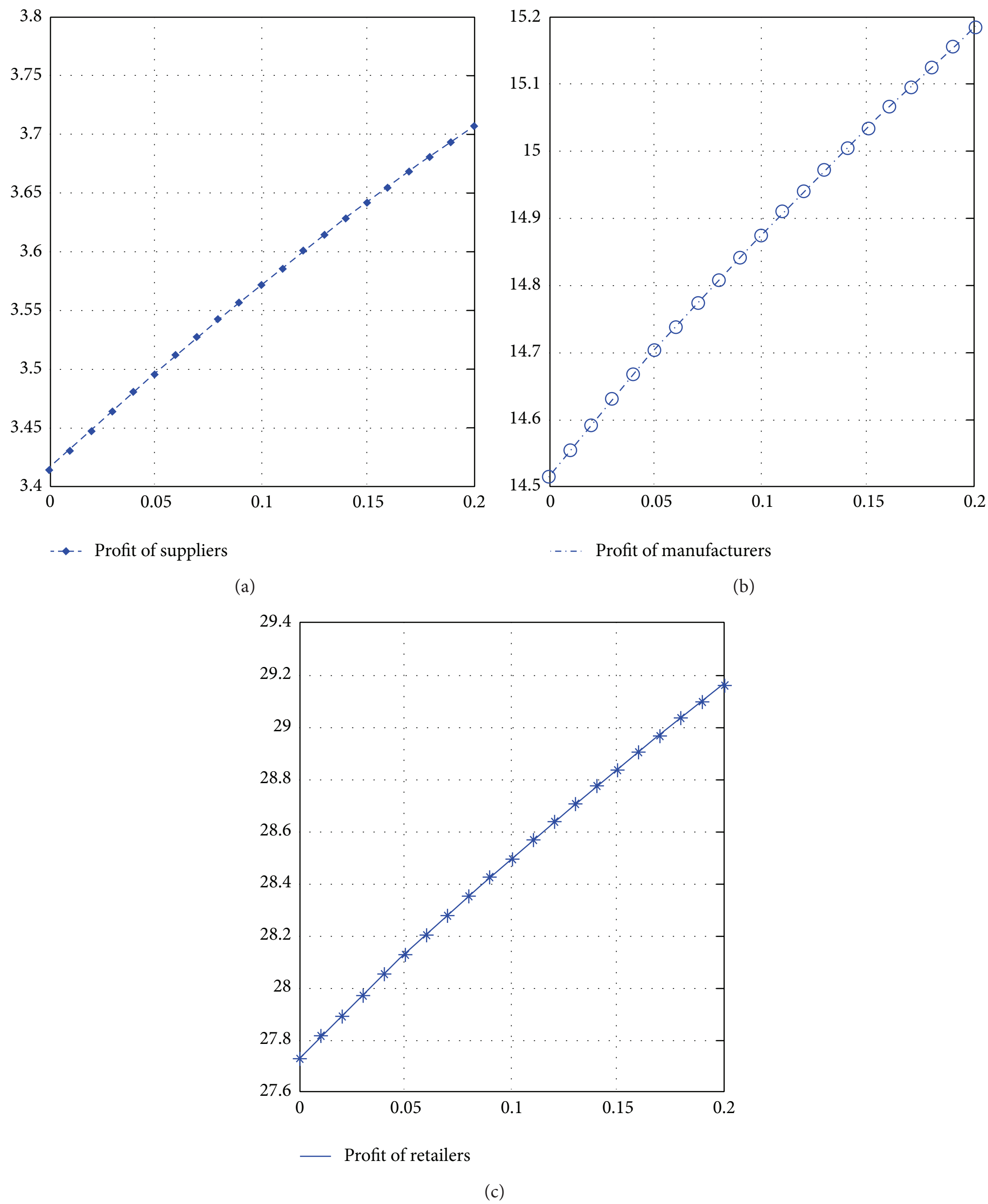


Figure 2: Players' profits in the discrete dynamic supply chain network with delay effect.

the 3rd period; thus the inventory from the 2nd period to the 3rd period also increases. Due to the increasing of the demand in the 3rd period, the inventory from the 2nd period to the 3rd period increases obviously. The latter 3 columns can be analyzed in the same way.

From the latter three columns in Table 7, we can find that, in the absence of delay effect, the manufacturers' profits when making advertisement in all the three periods are lower than that only in the 1st and 2nd periods instead. It illustrates that, in some cases, the increased profit through advertisement is less than its investment volume, so at this time, it is meaningless and should not be the manufacturer's optimal strategy. On the other hand, the retailers' profits remain unchanged.

In Table 7, it is interesting that the volume of $q^r_{sm}(t)$ $q^r_s(t)$ and $q^r_m(t)$ is almost identical and the profits of all players in these cases are similar too in the first three columns and the 4th and 5th, respectively.

Comparing Table 6 with Table 7, we can see that because we only consider three periods, the 3rd period is the last one;

Table 7: Equilibrium results without delay effect of advertising investment.

Variables $t = 1, 2, 3$	$I_A^{mn}(1) = 0.15$ $I_A^{mn}(2) = 0$ $I_A^{mn}(3) = 0$	$I_A^{mn}(1) = 0$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0$	$I_A^{mn}(1) = 0$ $I_A^{mn}(2) = 0$ $I_A^{mn}(3) = 0.15$	$I_A^{mn}(1) = 0.15$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0$	$I_A^{mn}(1) = 0$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0.15$	$I_A^{mn}(1) = 0.15$ $I_A^{mn}(2) = 0.15$ $I_A^{mn}(3) = 0.15$
$q_{sm}^r(t)$ $s = 1, 2$ $m = 1, 2$	0.7214	0.7215	0.7214	0.7294	0.7294	0.7373
	0.4408	0.4408	0.4408	0.4450	0.4450	0.4491
	0.3797	0.3797	0.3797	0.3826	0.3826	0.3854
$q_s^r(t), q_m^r(t)$ $s = 1, 2$ $m = 1, 2$	1.4429	1.4430	1.4429	1.4588	1.4588	1.4746
	0.8815	0.8816	0.8815	0.8900	0.8900	0.8983
	0.7594	0.7595	0.7594	0.7652	0.7652	0.7708
$q_{mn}(t)$ $m = 1, 2$ $n = 1, 2$	0.5368	0.5182	0.5182	0.5355	0.5168	0.5342
	0.5132	0.5320	0.5132	0.5307	0.5307	0.5295
	0.4919	0.4919	0.5106	0.4907	0.5094	0.5082
$I_m(t)$ $m = 1, 2$	0.3692	0.4067	0.4066	0.3878	0.4251	0.4062
	0.2243	0.2243	0.2617	0.2163	0.2536	0.2456
	0	0	0	0	0	0
π_s	3.2254	3.2263	3.2254	3.3450	3.3450	3.4646
π_m	14.0953	14.0955	14.0845	14.3703	14.3597	14.6342
π_n	26.7327	26.7527	26.7721	27.3596	27.3991	28.0059

Table 8: Equilibrium results with only advertising investment of manufacturer 1.

Variables $t = 1, 2, 3$	$I_A^{1n}(1) = 0.15$ $I_A^{1n}(2) = 0$ $I_A^{1n}(3) = 0$	$I_A^{1n}(1) = 0$ $I_A^{1n}(2) = 0.15$ $I_A^{1n}(3) = 0$	$I_A^{1n}(1) = 0$ $I_A^{1n}(2) = 0$ $I_A^{1n}(3) = 0.15$	$I_A^{1n}(1) = 0.15$ $I_A^{1n}(2) = 0.15$ $I_A^{1n}(3) = 0$	$I_A^{1n}(1) = 0$ $I_A^{1n}(2) = 0.15$ $I_A^{1n}(3) = 0.15$	$I_A^{1n}(1) = 0.15$ $I_A^{1n}(2) = 0.15$ $I_A^{1n}(3) = 0.15$
$q_{1n}(t)$ $n = 1, 2$	0.5236	0.5207	0.5211	0.5248	0.5222	0.5263
	0.5212	0.5189	0.5161	0.5257	0.5205	0.5273
	0.4970	0.4999	0.4978	0.5038	0.5048	0.5087
$I_1(t)$	0.3955	0.4954	0.3934	0.4064	0.4043	0.4118
	0.2345	0.2413	0.2389	0.2435	0.2480	0.2503
	0	0	0	0	0	0
π_1	14.0075	13.9653	13.8845	14.1539	14.0303	14.2186
$q_{2n}(t)$ $n = 1, 2$	0.5327	0.5161	0.5165	0.5293	0.5131	0.5263
	0.5166	0.5280	0.5115	0.5302	0.5250	0.5273
	0.4924	0.4954	0.5069	0.4947	0.5093	0.5087
$I_2(t)$	0.3773	0.4081	0.4025	0.3973	0.4225	0.4118
	0.2255	0.2322	0.2571	0.2253	0.2571	0.2503
	0	0	0	0	0	0
π_2	14.1659	14.1223	14.0353	14.4704	14.3393	14.6873

therefore, the 3rd column in the two Tables has no difference with or without delay effect. When considering delay effect, the transaction volume in the 1st period is lower than that without delay effect except the 3rd column, whereas in the next 2 periods, the former is higher than the latter.

From Table 8, it can be obviously seen that when only manufacturer 1 makes advisements, the quantity of selling products to retailers is higher than that of manufacturer 2, but his profit is lower; in the same time, the higher the advertising investment of manufacturer 1 is, the bigger the profits of the two manufacturers are, and the bigger the profit difference between the two manufacturers is. Therefore, we can draw a conclusion when multiple firms engage in homogeneous products; one firm's advertising activity also has a positive effect on the other firms, which makes the so-called "Free-Rider Phenomenon" emerge and when the advertising investment is bigger, this phenomenon is more obvious.

Figures 2 and 3 illustrate the impacts of the advertising investment in the first period when the advertising in the next

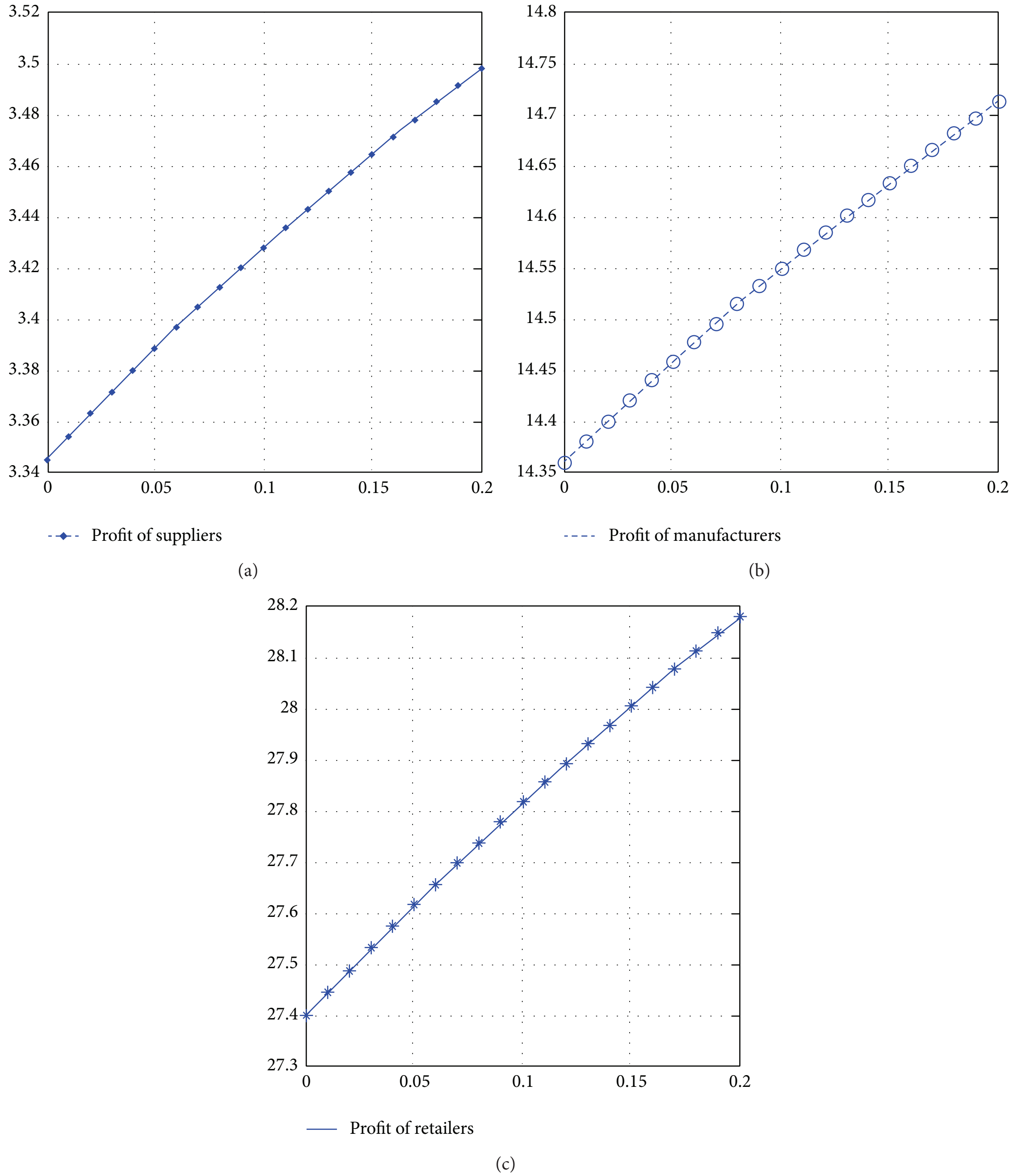


Figure 3: Players' profits in the discrete dynamic supply chain network without delay effect.

2 periods is fixed. The profits of all players in the supply chain network are higher with advertising than that without advertising and increase depending on the advertising investment volume, whereas the increasing margin is smaller and smaller.

From Figures 2 and 3, we also note that the profit differences of all actors are becoming bigger and bigger when the advertising investment in the 1st period increases. For example, when $I_A^{mn}(1) = 0$, the profit difference of manufacturers is $14.5163 - 14.360 = 0.1563$; when $I_A^{mn}(1) = 0.2$, the profit difference is $14.5163 - 14.360 = 0.1563$; $15.1832 - 14.7126 = 0.4706$. The profits of suppliers and retailers can be computed in the same way and have similar trends.

6. Conclusions

In the discrete dynamic decision making environment, this paper proposes a supply chain network model with demand uncertainties. The manufacturers purchase the raw materials from suppliers and sell products to consumers in demand markets by way of retailers; in the same time, the manufacturers and retailers use the advertising strategy to increase the demand of products, and the advertising investment has delay effect in the next periods. Using variational inequality theory, complement theory, and Lagrange duality theory, we formulate the profit functions and optimal behaviors of various players in the network and in turn compute the equilibrium results by modified projection and contraction algorithm. In the numerical examples, we illustrate the effectiveness of our model and analyze the impact of different advertising strategies on the equilibrium results.

From the numerical examples, we obtain the following conclusions: (1) when considering the delay effects, the earlier the advertising investment is made, the more profits the enterprises can obtain, and the whole supply chain network will benefit from the advertising strategy; (2) when not considering the delay effect, the advertising strategy is not always beneficial for the enterprises; if the investment is higher than the profit resulting from the strategy, the extra investment is harmful to enterprise; (3) if there are only part of the enterprises that make advertising activities, it is likely that the so-called "Free-Rider Phenomenon" emerges; (4) when advertising investment increases, the profit difference will magnify with delay effect than that without the effect. The managerial insights obtained in this paper may give insights to the decision makers in the enterprises and theorists in the supply chain management.

Future research may be in the following directions: as a common policy for promoting products, advertising strategy investment must have the cap constraints because of the limitation of funds.

Conflict of Interests

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

Acknowledgments

This paper is supported by the National Natural Science Foundation, China (nos. 71371102 and 71273149) and International (regional) Cooperation and Exchange, China (no. 71311120090).

References

[1] M. C. Cooper, M. L. Douglas, and D. P. Janus, "Supply chain management: more than a new name for logistics," *International Journal of Logistics Management*, vol. 8, no. 1, pp. 1–14, 1997.

[2] M. Christopher, *Logistics and Supply Chain Management*, Pearson UK, 2012.

[3] B. M. Beamon, "Supply chain design and analysis: models and methods," *International Journal of Production Economics*, vol. 55, no. 3, pp. 281–294, 1998.

[4] A. Nagurney, J. Dong, and D. Zhang, "A supply chain network equilibrium model," *Transportation Research E*, vol. 38, no. 5, pp. 281–303, 2002.

[5] J. Dong, D. Zhang, and A. Nagurney, "A supply chain network equilibrium model with random demands," *European Journal of Operational Research*, vol. 156, no. 1, pp. 194–212, 2004.

[6] H. Sarimveis, P. Patrinos, C. D. Tarantilis, and C. T. Kiranoudis, "Dynamic modeling and control of supply chain systems: a review," *Computers and Operations Research*, vol. 35, no. 11, pp. 3530–3561, 2008.

[7] J. Yue, J. Austin, Z. Huang, and B. Chen, "Pricing and advertisement in a manufacturer-retailer supply chain," *European Journal of Operational Research*, vol. 231, no. 2, pp. 492–502, 2013.

[8] A. Nagurney, J. Cruz, J. Dong, and D. Zhang, "Supply chain networks, electronic commerce, and supply side and demand side risk," *European Journal of Operational Research*, vol. 164, no. 1, pp. 120–142, 2005.

[9] A. Nagurney and F. Toyasaki, "Reverse supply chain management and electronic waste recycling: a multitiered network equilibrium framework for e-cycling," *Transportation Research E*, vol. 41, no. 1, pp. 1–28, 2005.

[10] K. Wu, A. Nagurney, Z. Liu, and J. K. Stranlund, "Modeling generator power plant portfolios and pollution taxes in electric power supply chain networks: a transportation network equilibrium transformation," *Transportation Research D*, vol. 11, no. 3, pp. 171–190, 2006.

[11] D. Hammond and P. Beullens, "Closed-loop supply chain network equilibrium under legislation," *European Journal of Operational Research*, vol. 183, no. 2, pp. 895–908, 2007.

[12] G. Yang, Z. Wang, and X. Li, "The optimization of the closed-loop supply chain network," *Transportation Research E*, vol. 45, no. 1, pp. 16–28, 2009.

[13] A. H. Masoumi, M. Yu, and A. Nagurney, "A supply chain generalized network oligopoly model for pharmaceuticals under brand differentiation and perishability," *Transportation Research E*, vol. 48, no. 4, pp. 762–780, 2012.

[14] Q. Qiang, K. Ke, T. Anderson, and J. Dong, "The closed-loop supply chain network with competition, distribution channel investment, and uncertainties," *Omega*, vol. 41, no. 2, pp. 186–194, 2013.

[15] M. Yu and A. Nagurney, "Competitive food supply chain networks with application to fresh produce," *European Journal of Operational Research*, vol. 224, no. 2, pp. 273–282, 2013.

[16] F. Toyasaki, P. Daniele, and T. Wakolbinger, "A variational inequality formulation of equilibrium models for end-of-life products with nonlinear constraints," *European Journal of Operational Research*, vol. 236, no. 1, pp. 340–350, 2014.

[17] J. M. Cruz and T. Wakolbinger, "Multiperiod effects of corporate social responsibility on supply chain networks, transaction costs, emissions, and risk," *International Journal of Production Economics*, vol. 116, no. 1, pp. 61–74, 2008.

[18] P. Daniele, "Evolutionary variational inequalities and applications to complex dynamic multi-level models," *Transportation Research E*, vol. 46, no. 6, pp. 855–880, 2010.

[19] J. M. Cruz and Z. Liu, "Modeling and analysis of the multiperiod effects of social relationship on supply chain networks," *European Journal of Operational Research*, vol. 214, no. 1, pp. 39–52, 2011.

[20] Y. Hamdouch, "Multi-period supply chain network equilibrium with capacity constraints and purchasing strategies," *Transportation Research C*, vol. 19, no. 5, pp. 803–820, 2011.

[21] Z. Liu and J. M. Cruz, "Supply chain networks with corporate financial risks and trade credits under economic uncertainty," *International Journal of Production Economics*, vol. 137, no. 1, pp. 55–67, 2012.

[22] Z. Feng, Z. Wang, and Y. Chen, "The equilibrium of closed-loop supply chain super network with time-dependent parameters," *Transportation Research Part E*, vol. 64, pp. 1–11, 2014.

[23] G. Aust and U. Buscher, "Cooperative advertising models in supply chain management: a review," *European Journal of Operational Research*, vol. 234, no. 1, pp. 1–14, 2014.

[24] M. Bergen and G. John, "Understanding cooperative advertising participation rates in conventional channels," *Journal of Marketing Research*, vol. 34, no. 3, pp. 357–369, 1997.

[25] M. Lei, S. Sun, and D. Yang, "A study of the joint advertising channels," *Journal of Service Science and Management*, vol. 2, no. 4, pp. 418–426, 2009.

[26] T. Xiao, X. Yan, and J. Zhao, "Coordination of a supply chain with advertising investment and allowing the second ordering," *Technology and Investment*, vol. 1, no. 3, pp. 191–200, 2010.

[27] T. Chen, "Coordinating the ordering and advertising policies for a single-period commodity in a two-level supply chain," *Computers and Industrial Engineering*, vol. 61, no. 4, pp. 1268–1274, 2011.

[28] X. He, A. Krishnamoorthy, A. Prasad, and S. P. Sethi, "Retail competition and cooperative advertising," *Operations Research Letters*, vol. 39, no. 1, pp. 11–16, 2011.

[29] Y. Tsao and G. Sheen, "Effects of promotion cost sharing policy with the sales learning curve on supply chain coordination," *Computers and Operations Research*, vol. 39, no. 8, pp. 1872–1878, 2012.

[30] S. Jogensen and G. Zaccour, "A survey of game-theoretic models of cooperative advertising," *European Journal of Operational Research*, vol. 237, no. 1, pp. 1–14, 2014.

[31] W. Zhang, S. Li, D. Zhang, and W. Hou, "On the impact of advertising initiatives in supply chains," *European Journal of Operational Research*, vol. 234, no. 1, pp. 99–107, 2014.

[32] L. Lambertini, "Coordinating static and dynamic supply chains with advertising through two-part tariffs," *Automatica*, vol. 50, no. 2, pp. 565–569, 2014.

[33] B. He, Y. Xu, and X. Yuan, "A logarithmic-quadratic proximal prediction-correction method for structured monotone variational inequalities," *Computational Optimization and Applications*, vol. 35, no. 1, pp. 19–46, 2006.

[34] G. M. Korpelevich, "The extra gradient method for finding saddle points and other problems," *Matekon*, vol. 13, pp. 35–49, 1977.

[35] L. Zhang and Y. Zhou, "A new approach to supply chain network equilibrium models," *Computers & Operations Research*, vol. 63, no. 1, pp. 82–88, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 750279, 10 pages
http://dx.doi.org/10.1155/2014/750279

Research Article

Dual-Channel Particle Filter Based Track-Before-Detect for Monopulse Radar

Fei Cai, Hongqi Fan, and Qiang Fu

ATR Key Laboratory, National University of Defense Technology, Changsha, Hunan 410073, China

Correspondence should be addressed to Fei Cai; caifei.nudt@gmail.com

Received 28 February 2014; Accepted 1 August 2014; Published 18 August 2014

Academic Editor: Guangchen Wang

A particle filter based track-before-detect (PF-TBD) algorithm is proposed for the monopulse high pulse repetition frequency (PRF) pulse Doppler radar. The actual measurement model is adopted, in which the range is highly ambiguous and the sum and difference channels exist in parallel. A quantization method is used to approximate the point spread function to reduce the computation load. The detection decisions of the PF-TBD are fed to a binary integrator to further improve the detection performance. Simulation results show that the proposed algorithm can detect and track the low SNR target efficiently. The detection performance is improved significantly for both the single frame and the multiframe detection compared with the classical detector. A performance comparison with the PF-TBD using sum channel only is also supplied.

1. Introduction

The developments of stealthy military aircraft and cruise missiles recently have emphasized the need for detection and tracking of low signal-to-noise ratio (SNR) targets. This need is especially urgent for a radar seeker because of its limited battery capacity and antenna size. High pulse repetition frequency (PRF) pulse Doppler is generally used in a radar seeker at early detection stage, which allows thermal noise-limited detection of targets with high radial velocities [1]. Noncoherent or binary integration is often used after the coherent processing to improve the detection performance. But the radar data rate and the unknown target motion have limited the coherent processing interval (CPI) and noncoherent/binary times. The azimuth and elevation are measured by monopulse generally, which is a widely used technique to provide accurate angle measurements in the tracking radar. A monopulse system for estimating one angle typically consists of two identical antennas, either separated by some distance (phase monopulse) or at the same phase center but with a squint angle (amplitude monopulse), whose outputs are summed up to produce a sum channel Σ and are subtracted to yield the difference channel Δ as shown in Figure 1. The angular information θ is contained in the monopulse ratio $\gamma(\theta) = \Delta(\theta)/\Sigma(\theta)$ providing the function $\theta \rightarrow \gamma(\theta)$ is reversible. Poor monopulse estimation performance under low SNR has also deteriorated the guidance performance.

Track-before-detect (TBD) is a simultaneous detection and tracking paradigm that uses unthresholded data or thresholded data with significantly lower thresholds than those used in conventional detectors and integrates them over time according to the target dynamic model to improve the sensitivity to low SNR targets. Typical TBD is implemented as a batch algorithm using the Hough transform [2] or dynamic programming [3]. The Hough transform TBD is suitable only for linear trajectories. The dynamic programming TBD is studied more for the radar application and is applied in pulse Doppler radars in [4–6]. Particle filter based TBD (PF-TBD) was introduced by [7] and extended by [8–10]. Compared to the typical methods, it is recursive and does not require discretization of the state space.

For simplicity, most researches on PF-TBD are based on grayscale-image-like measurements (e.g., [8, 10]). Boers and Driessen [9, 11] have studied PF-TBD on search radar measurements. A Rao-Blackwellised PF-TBD is proposed for over-the-horizon radar in [12]. Multisensor PF-TBD is studied for MIMO radar [13]. There is no open literature addressing PF-TBD on monopulse radar to the best of

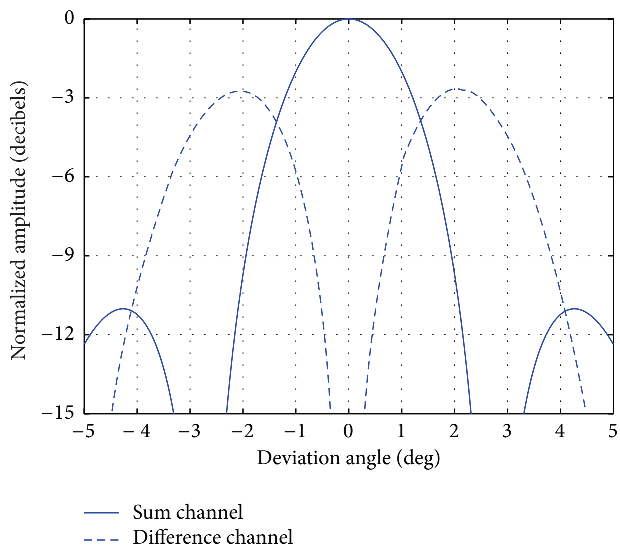


Figure 1: Amplitude of the sum and difference channels at different deviation angles.

our knowledge. In monopulse radar systems, the sum and difference channels exist in parallel as Figure 1 has shown. A PF-TBD algorithm similar to [12] can be applied by using only the output of the sum channel as the measurements. The target Doppler and intensity are estimated by it and then the bearing and azimuth are estimated by classical monopulse methods (e.g., ML method proposed by [14]). But from Figure 1 we can see that amplitude of the difference channel is comparable to that of the sum channel when the target is not at the beam center, which often occurred in the target searching stage. So fusion of the sum and difference channels using Bayesian theory in the PF-TBD algorithm is possible to improve the detection performance as well as the monopulse estimation performance.

In this work, the target and measurement models of the monopulse high PRF pulse Doppler radar are constructed. Based on them, we derive a PF-TBD algorithm which can effectively detect and track the low SNR target. Its detection performance is compared with the classical detector, which shows that more than 7 dB gain in SNR can be attained. A quantization method of approximating the point spread function is proposed to reduce the computation load of the PF-TBD. Binary integration of the PF-TBD's detection result is proposed to further improve the detection performance, which is shown to be very effective and not limited by the target maneuver.

The rest of the work is organized as follows. In Section 2 the target and sensor models are formulated. The recursive Bayesian TBD filter for this application is described in Section 3 and its PF implementation procedure is derived in Section 4. Two simulated examples are presented in Section 5, in which the detection and estimation performances of the proposed algorithm are evaluated in comparison with the classical method and the sum-only PF-TBD. Conclusions and future work are drawn in the last section.

2. Target and Measurement Models

2.1. Target Model. The high PRF can measure Doppler unambiguously, but it is highly ambiguous in range, which precludes the pulse delay ranging. The range information is not a must for a radar seeker, however, since the proportional navigation is commonly adopted. As a result, only the target Doppler is involved in the target state vector in this paper. The target azimuth and elevation are measured by monopulse. For the sake of brevity, only one difference channel (azimuth difference or elevation difference) is considered. Moreover, the unknown target echo amplitude is also incorporated to implement the PF-TBD algorithm. The target state vector is then defined as

$$\mathbf{x}_k = \left[f_d^k, A_\Sigma^k, \gamma_k \right]^T, \tag{1}$$

where f_d^k, A_Σ^k, and γ_k denote the Doppler frequency, echo amplitude of the sum channel, and monopulse ratio of the target in frame k, respectively. The Doppler frequency $f_d = 2v_r/\lambda$, where λ is the wavelength and v_r is the radial velocity.

Although the dynamic model can be as general as $\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_{k-1})$ for a particle implementation, where $\mathbf{v}_{k-1}$ is the process noise sequence, for simplicity we model the target motion relative to the radar as the nearly constant velocity model with a white acceleration noise $v_k^{(1)}$. The target echo amplitude and monopulse ratio are modeled as random walk processes with process noises $v_k^{(2)}$ and $v_k^{(3)}$, respectively. The process noises $v_k^{(1)}$, $v_k^{(2)}$, and $v_k^{(3)}$ are mutually independent, zero mean white noise with variances $\sigma_{(1)}^2$, $\sigma_{(2)}^2$, and $\sigma_{(3)}^2$, respectively. Thus, the system dynamic equation is

$$\mathbf{x}_k = \mathbf{x}_{k-1} + T \cdot \mathbf{v}_{k-1}, \tag{2}$$

where T is the CPI and $\mathbf{v}_{k-1} = [v_{k-1}^{(1)}, v_{k-1}^{(2)}, v_{k-1}^{(3)}]^T$. This target model accommodates not only target maneuver but also fluctuations of the target intensity and the monopulse ratio.

Target existence variable E_k is modeled as a two-state Markov chain and $E_k \in \{0, 1\}$. Here 0 denotes the event that the target is absent, while 1 denotes the opposite [15]. Furthermore, we define the transitional probabilities of target "birth" (P_b) and "death" (P_d) as

$$\begin{aligned} P_b &\triangleq P\left\{E_k = 1 \mid E_{k-1} = 0\right\} \\ P_d &\triangleq P\left\{E_k = 0 E_{k-1} = 1\right\}. \end{aligned} \tag{3}$$

Thus the transitional probability matrix Π is given by

$$\Pi = \begin{bmatrix} 1 - P_d & P_b \\ P_d & 1 - P_b \end{bmatrix}. \tag{4}$$

2.2. Measurement Model. We assume that the target is located in the clutter-free region; thus the clutter is not considered in the signal model. When the target is present, the received signal sequences at the video stage of the sum and difference channels in frame k are denoted as $s^k_\Sigma(n)$ and $s^k_\Delta(n)$ and given by

$$s^k_\Sigma(n) = A^k_\Sigma \exp\left\{j2\pi\left(f^k_d nT_r + \phi_k\right)\right\} + n^k_\Sigma(n) \quad (5)$$

$$s^k_\Delta(n) = \gamma_k A^k_\Sigma \exp\left\{j2\pi\left(f^k_d nT_r + \phi_k\right)\right\} + n^k_\Delta(n) \quad (6)$$

$$\triangleq A^k_\Delta \exp\left\{j2\pi\left(f^k_d nT_r + \phi_k\right)\right\} + n^k_\Delta(n), \quad (7)$$

respectively, where A^k_Δ is the amplitude of the difference channel, ϕ_k is some arbitrary phase, T_r is the pulse repetition interval (PRI), and $n = 0, 1, \ldots, N-1$ is index of the sample in an CPI. The background thermal noises $n^k_\Sigma(n)$ and $n^k_\Delta(n)$ are mutually independent, zero mean, and temporally white complex Gaussian processes with the same variance. The Doppler frequency f^k_d is assumed to be constant within an CPI.

The coherent integrations of the sum and difference echoes are done via fast Fourier transform (FFT) independently. To reduce peak side-lobe levels, the signal sequences are windowed before the FFT. The result of the coherent integration is given by

$$y^k_{\Sigma/\Delta}(l) = \sum_{n=0}^{N_f-1} s^k_{\Sigma/\Delta}(n)\, w_n \exp\left\{-j2\pi\left(\frac{nl}{N_f}\right)\right\}, \quad (8)$$

where the subscript Σ/Δ denotes sum channel Σ or difference channel Δ for simplification, N_f is the next power of two that is greater than or equal to N, $s_{\Sigma/\Delta}(n) = 0$ for $n > N-1$ (also known as zero padding), w_n is the windowing function, and $l = 0, 1, \ldots, N_f - 1$ is the index of the frequency bin. The signal's unknown phase component is useless, so the magnitude of the spectrum in each frequency bin forms the set of measurements in frame k. Then the measurement can be modeled as

$$z^k_{\Sigma/\Delta}(l) = \begin{cases} \left|A^k_{\Sigma/\Delta} B_k\left(f^k_d, l\right) + u^k_{\Sigma/\Delta}(l)\right| & E_k = 1 \\ \left|u^k_{\Sigma/\Delta}(l)\right| & E_k = 0, \end{cases} \quad (9)$$

where $|\cdot|$ is the complex modulus, $B_k(f^k_d, l) = \exp\{j2\pi\phi_k\} \sum_{n=0}^{N-1} w_n \exp\{j2\pi f^k_d nT_r\} \exp\{-j2\pi(nl/N_f)\}$, and $u^k_\Sigma(l)$ and $u^k_\Delta(l)$ are the background noises of the sum and difference channels, respectively, after the coherent integration. Because of linearity of the FFT, $u^k_\Sigma(l)$ and $u^k_\Delta(l)$ are also zero mean i.i.d. complex Gaussian noise processes. We assume that they both have a variance $2\sigma^2_u$.

As has been stated that not all the frequency bins of the FFT result are of interest, only bins in clutter-free region constitute the set of measurements at frame k; that is, $\mathbf{z}^k_{\Sigma/\Delta} = \{z^k_{\Sigma/\Delta}(l_c : (N_f - l_c - 1))\}$, where $l_c = \mathrm{ceil}(2v_M/(\lambda\delta_f))$, v_M is the horizontal velocity of the missile, and $\delta_f = 1/(T_r N_f)$ is the Doppler bin size.

Following the model described above, the likelihood in each frequency bin when the target is present has a Ricean distribution

$$\begin{aligned} p\left(z^k_{\Sigma/\Delta}(l) \mid \mathbf{x}_k, E_k = 1\right) &= \frac{z^k_{\Sigma/\Delta}(l)}{\sigma^2_u} I_0\left(\frac{A^k_{\Sigma/\Delta} \mid B_k\left(f^k_d, l\right) \mid z^k_{\Sigma/\Delta}(l)}{\sigma^2_u}\right) \\ &\quad \times \exp\left\{-\frac{z^k_{\Sigma/\Delta}(l)^2 + \left(A^k_{\Sigma/\Delta}\right)^2 \left|B_k\left(f^k_d, l\right)\right|^2}{2\sigma^2_u}\right\}, \end{aligned} \quad (10)$$

where $I_0(\cdot)$ is the modified Bessel function of order zero. The likelihood when the target is absent has a Rayleigh distribution

$$p\left(z^k_{\Sigma/\Delta}(l) \mid E_k = 0\right) = \frac{z^k_{\Sigma/\Delta}(l)}{\sigma^2_u} \exp\left\{-\frac{z^k_{\Sigma/\Delta}(l)^2}{2\sigma^2_u}\right\}. \quad (11)$$

Because of the windowing before the FFT, the target (if present) power will spread into the bins in the vicinity of its location. Let $C(\mathbf{x}_k)$ denote the bins affected by the target (i.e., the target's effect on the other bins is negligible); then the likelihood function of the whole measurement set when the target is present can be approximated as follows:

$$\begin{aligned} p\left(\mathbf{z}^k_{\Sigma/\Delta} \mid \mathbf{x}_k, E_k = 1\right) &\approx \prod_{l \in C(\mathbf{x}_k)} p\left(z^k_{\Sigma/\Delta}(l) \mid \mathbf{x}_k, E_k = 1\right) \\ &\quad \times \prod_{l \notin C(\mathbf{x}_k)} p\left(z^k_{\Sigma/\Delta}(l) \mid E_k = 0\right) \end{aligned} \quad (12)$$

and the likelihood function when the target is absent is

$$p\left(\mathbf{z}^k_{\Sigma/\Delta} \mid E_k = 0\right) = \prod_{l=l_c}^{N_f - l_c - 1} p\left(z^k_{\Sigma/\Delta}(l) \mid E_k = 0\right). \quad (13)$$

We denote the set of complete measurements up to frame k as $\mathbf{Z}_k = \{\mathbf{z}^i_\Sigma, \mathbf{z}^i_\Delta, i = 1, \ldots, k\}$.

It is computational complex to calculate the $|B_k(f^k_d, l)|$ for bins in $C(\mathbf{x}_k)$ in real time applications. The contribution of $\mathbf{x}_k$ to bin l in $C(\mathbf{x}_k)$ (i.e., point spread function) is generally approximated by a Gaussian-like function (e.g., [7, 8] for optical sensor). Using the Gaussian approximation method, the point spread function $h(\mathbf{x}_k, l)$ is

$$\begin{aligned} h(\mathbf{x}_k, l) &= A^k_{\Sigma/\Delta} \left|B_k\left(f^k_d, l\right)\right| \\ &\approx A^k_{\Sigma/\Delta} G \exp\left\{-\frac{\left(l\delta_f - f^k_d\right)^2}{2\beta^2}\right\}, \end{aligned} \quad (14)$$

where $G = \sum_{n=0}^{N-1} w_n$ is the coherent integration gain and β is a parameter to be designed to better approximate the amount of blurring introduced by the FFT windowing functions. But this approximation is valid only within a limited range as Figure 2 has shown. To solve this

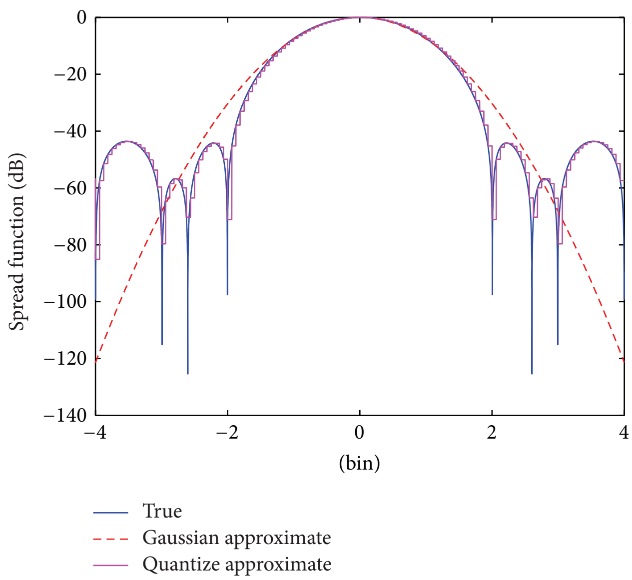


FIGURE 2: Comparison of different point spread function approximation methods. Hamming window is used, $N = 1024$, $N_{app} = 64$.

problem, we present an approximation approach which is calculation-free and more precise. Note that $|B_k(f_d^k, l)| = |\sum_{n=0}^{N-1} w_n \exp\{j2\pi n T_r(f_d^k - l\delta_f)\}|$ can be expressed as a function $g_{\mathbf{w}}(x)$ with a parameter $\mathbf{w} = \{w_n\}$ and a variable $x = |f_d^k - l\delta_f|$. Because the windowing function $\mathbf{w}$ can be taken as known a priori, we can quantize $g_{\mathbf{w}}(x)$ into a number of points (e.g., $g_{\mathbf{w}}(k\delta_f/N_{app})$, $k = 0, 1, \ldots, N_{app} - 1$ for $x \in [0, \delta_f)$, where N_{app} is the number of points each bin is quantized into, and we can store them as a look-up table in the read-only memory (ROM). In real time operations, the value of the quantized point nearest to the true point is read from the ROM and used; that is,

$$h(\mathbf{x}_k, l) \approx A_{\Sigma/\Delta}^k \left| B_k \left(\left\lfloor \left| l\delta_f - f_d^k \right| \frac{N_{app}}{\delta_f} + 0.5 \right\rfloor \frac{\delta_f}{N_{app}}, l \right) \right|, \tag{15}$$

where $\lfloor \cdot \rfloor$ is the floor function and $\lfloor x + 0.5 \rfloor$ rounds x to the nearest integer. The result of this approximation is also presented in Figure 2.

3. Recursive Bayesian Filtering Procedure

The posterior probability of target existence $P_E^k \triangleq P\{E_k = 1 \mid \mathbf{Z}_k\}$ and $\mathbf{x}_k$ are estimated recursively by a Bayesian method as follows. Given the joint posterior PDF at frame $k-1$, $p(\mathbf{x}_{k-1}, E_{k-1} \mid \mathbf{Z}_{k-1})$ and the latest measurement $\mathbf{Z}_k$, the goal is to construct the joint posterior PDF at frame k, $p(\mathbf{x}_k, E_k \mid \mathbf{Z}_k)$. P_E^k and $\mathbf{x}_k$ are then estimated using $p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_k)$.

Prediction. Prediction of E_k is given by

$$\begin{bmatrix} P\{E_k = 1 \mid \mathbf{Z}_{k-1}\} \\ P\{E_k = 0 \mid \mathbf{Z}_{k-1}\} \end{bmatrix} = \Pi \begin{bmatrix} P\{E_{k-1} = 1 \mid \mathbf{Z}_{k-1}\} \\ P\{E_{k-1} = 0 \mid \mathbf{Z}_{k-1}\} \end{bmatrix}. \tag{16}$$

If $E_k = 0$, $\mathbf{x}_k$ is undefined and no prediction of it is needed. If $E_k = 1$, the prediction step of $\mathbf{x}_k$ can be expressed as

$$\begin{aligned} p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_{k-1}) &= \int p(\mathbf{x}_k, E_k = 1 \mid \mathbf{x}_{k-1}, E_{k-1} = 1, \mathbf{Z}_{k-1}) \\ &\quad \cdot p(\mathbf{x}_{k-1}, E_{k-1} = 1 \mid \mathbf{Z}_{k-1})\, d\mathbf{x}_{k-1} \\ &\quad + \int p(\mathbf{x}_k, E_k = 1 \mid \mathbf{x}_{k-1}, E_{k-1} = 0, \mathbf{Z}_{k-1}) \\ &\quad \cdot p(\mathbf{x}_{k-1}, E_{k-1} = 0 \mid \mathbf{Z}_{k-1})\, d\mathbf{x}_{k-1}, \end{aligned} \tag{17}$$

where

$$\begin{aligned} &p(\mathbf{x}_k, E_k = 1 \mid \mathbf{x}_{k-1}, E_{k-1} = 1, \mathbf{Z}_{k-1}) \\ &\quad = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, E_k = 1, E_{k-1} = 1)\, P\{E_k = 1 \mid E_{k-1} = 1\} \\ &\quad = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, E_k = 1, E_{k-1} = 1)(1 - P_d), \\ &p(\mathbf{x}_k, E_k = 1 \mid \mathbf{x}_{k-1}, E_{k-1} = 0, \mathbf{Z}_{k-1}) \\ &\quad = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, E_k = 1, E_{k-1} = 0)\, P\{E_k = 1 \mid E_{k-1} = 0\} \\ &\quad = p_b(\mathbf{x}_k)\, P_b. \end{aligned} \tag{18}$$

The transitional density $p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, E_k = 1, E_{k-1} = 1)$ is defined by the target dynamic model (2). The PDF $p_b(\mathbf{x}_k)$ denotes the initial target density on its appearance.

Update. The update equation using Bayes' rule is given by

$$\begin{aligned} &p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_k) \\ &\quad = \frac{p(\mathbf{z}_\Sigma^k, \mathbf{z}_\Delta^k \mid \mathbf{x}_k, E_k = 1)\, p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_{k-1})}{p(\mathbf{z}_\Sigma^k, \mathbf{z}_\Delta^k \mid \mathbf{Z}_{k-1})}, \end{aligned} \tag{19}$$

where the prediction density $p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_{k-1})$ is given by (17), the normalizing constant in the denominator is $p(\mathbf{z}_\Sigma^k, \mathbf{z}_\Delta^k \mid \mathbf{Z}_{k-1}) = \int p(\mathbf{z}_\Sigma^k, \mathbf{z}_\Delta^k \mid \mathbf{x}, E_k = 1) p(\mathbf{x}, E_k = 1 \mid \mathbf{Z}_{k-1}) d\mathbf{x}$, and the likelihood function $p(\mathbf{z}_\Sigma^k, \mathbf{z}_\Delta^k \mid \mathbf{x}_k, E_k = 1)$ is

$$\begin{aligned} &p(\mathbf{z}_\Sigma^k, \mathbf{z}_\Delta^k \mid \mathbf{x}_k, E_k = 1) \\ &\quad = p(\mathbf{z}_\Sigma^k \mid \mathbf{x}_k, E_k = 1)\, p(\mathbf{z}_\Delta^k \mid \mathbf{x}_k, E_k = 1), \end{aligned} \tag{20}$$

where the likelihood function $p(\mathbf{z}_{\Sigma/\Delta}^k \mid \mathbf{x}_k, E_k = 1)$ is given by (12).

Estimate. P_E^k is estimated by taking marginal of $p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_k)$ as follows:

$$\widehat{P}_E^k = \int p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_k)\, d\mathbf{x}_k. \tag{21}$$

Using expected a posterior (EAP) estimator, $\mathbf{x}_k$ is estimated by

$$\widehat{\mathbf{x}}_k = \frac{\int \mathbf{x}_k p(\mathbf{x}_k, E_k = 1 \mid \mathbf{Z}_k)\, d\mathbf{x}_k}{\widehat{P}_E^k}. \tag{22}$$

4. Particle Filter Implementation

To implement the recursive Bayesian filtering procedure, a SIR particle filter based TBD algorithm described in [8] is adopted with some modifications. As the particle filter tends to suffer from a progressive degeneration as the sequence evolves, an MCMC step referred to as resample-move in [16] is employed after importance resampling, which adds diversity to the particles without altering the underlying distribution [10]. A Metropolis-Hasting resample-move method is used as described in [10, 17]. Taking move of the γ, for example, a proposal distribution $q_m(\gamma_k' \mid \gamma_k)$ is defined, from which a sample is drawn for each particle after resampling. A monopulse ratio γ_k' is obtained conditioned on the old monopulse ratio γ_k while keeping the other two states unchanged. Under the assumption that the proposal is symmetric, $q_m(\gamma_k' \mid \gamma_k) = q_m(\gamma_k \mid \gamma_k')$, the new particle is accepted or rejected on a test, formed by a ratio of likelihoods

$$T_{\gamma',\gamma} = \frac{L\left(\mathbf{z}_k \mid f_d^k, A_\Sigma^k, \gamma_k'\right)}{L\left(\mathbf{z}_k \mid f_d^k, A_\Sigma^k, \gamma_k\right)}. \tag{23}$$

If $T_{\gamma',\gamma} > 1$, then the new particle, with monopulse ratio γ', is kept. Otherwise the new particle is kept in preference to the old particle only if $U < T_{\gamma',\gamma}$, where U is a uniform random number between 0 and 1. The move operation is used twice in this application, firstly to the amplitude A_Σ^k and then to the monopulse ratio γ_k. Truncated Gaussian distributions with different variances and means at A_Σ^k and γ_k, respectively, are used as the proposal distributions.

A detailed description of the TBD algorithm is given as follows.

Initialization. Set $k = 0$ and generate N_s samples $\{E_0^i\}_{i=1}^{N_s}$ from $P_E^0 = P(E_0 = 1)$. If $E_0^i = 1$, generate $\mathbf{x}_0^i$ from the birth density $q_b(\mathbf{x}_0 \mid \mathbf{z}_0)$, or else, $\mathbf{x}_0^i$ is undefined.

Then, given $[\{\mathbf{x}_{k-1}^i\}_{i=1}^{N_s}, \mathbf{z}_k]$ at each frame k, go from Steps 1 to 5.

Step 1 (prediction). Generate $\{E_k^i\}_{i=1}^{N_s}$ on the basis of $\{E_{k-1}^i\}_{i=1}^{N_s}$ and Π. If $E_k^i = 0$, $\mathbf{x}_k^i$ is undefined. If $E_{k-1}^i = 1$ and $E_k^i = 1$, predict $\mathbf{x}_k^i$ according to (2). For the new born particles, that is, those with $E_{k-1}^i = 0$ and $E_k^i = 1$, generate $\mathbf{x}_k^i$ from the birth density $q_b(\mathbf{x}_k \mid \mathbf{z}_k)$.

Step 2 (update). In the SIR filter, the prior PDF $p(z_k \mid x_{k-1}^i)$ is chosen to be the important density and, thus, unnormalized weights are proportional to the likelihood functions. Consequently, using the likelihood ratios as unnormalized weights will have no effect on the performance of the SIR filter. Thus the importance weights are calculated by the following [7]:

$$\widetilde{w}_k^i = \begin{cases} \prod\limits_{l \in C_i(\mathbf{x}_k)} L\left(z_\Sigma^k(l), z_\Delta^k(l) \mid \mathbf{x}_k^i\right) & \text{if } E_k^i = 1 \\ 1 & \text{if } E_k^i = 0. \end{cases} \tag{24}$$

We simplify the likelihood $L(z_\Sigma^k(l), z_\Delta^k(l) \mid \mathbf{x}_k^i)$ as follows:

$$\begin{aligned} L\left(z_\Sigma^k(l), z_\Delta^k(l) \mid \mathbf{x}_k^i\right) &\triangleq \frac{p\left(z_\Sigma^k(l), z_\Delta^k(l) \mid \mathbf{x}_k^i, E_k = 1\right)}{p\left(z_\Sigma^k(l), z_\Delta^k(l) \mid E_k = 0\right)} \\ &= \frac{p\left(z_\Sigma^k(l) \mid \mathbf{x}_k^i, E_k = 1\right) p\left(z_\Delta^k(l) \mid \mathbf{x}_k^i, E_k = 1\right)}{p\left(z_\Sigma^k(l) \mid E_k = 0\right) p\left(z_\Delta^k(l) \mid E_k = 0\right)} \\ &= L\left(z_\Sigma^k(l) \mid \mathbf{x}_k^i\right) L\left(z_\Delta^k(l) \mid \mathbf{x}_k^i\right). \end{aligned} \tag{25}$$

From (10) and (11), $L(z_{\Sigma/\Delta}^k(l) \mid \mathbf{x}_k^i)$ can be simplified as

$$\begin{aligned} L\left(z_{\Sigma/\Delta}^k(l) \mid \mathbf{x}_k^i\right) &= \frac{p\left(z_{\Sigma/\Delta}^k(l) \mid \mathbf{x}_k^i, E_k = 1\right)}{p\left(z_{\Sigma/\Delta}^k(l) \mid E_k = 0\right)} \\ &= I_0\left(\frac{\left|A_{\Sigma/\Delta}^{(k,i)} B_k\left(f_d^{(k,i)}, l\right)\right| z_{\Sigma/\Delta}^k(l)}{\sigma_u^2}\right) \\ &\quad \times \exp\left\{-\frac{\left|A_{\Sigma/\Delta}^{(k,i)} B_k\left(f_d^{(k,i)}, l\right)\right|^2}{2\sigma_u^2}\right\}, \end{aligned} \tag{26}$$

where $A_\Delta^{(k,i)}$ is calculated by $A_\Delta^{(k,i)} = \gamma_k^i A_\Sigma^{(k,i)}$. Then get the normalized weights $\{w_k^i\}_{i=1}^{N_s}$ by $w_k^i = \widetilde{w}_k^i / \sum_{i=1}^{N_s} \widetilde{w}_k^i$.

Step 3 (resample). Generate a new set of samples $[\{E_k^i, x_k^i\}_{i=1}^{N_s}]$ from $[\{E_k^i, x_k^i, w_k^i\}_{i=1}^{N_s}]$ and replace them using systematic resampling algorithm [18]. The weights of the new samples are not required since they are all equal to $1/N_s$.

Step 4 (MCMC move). Generate a new set of samples from $[\{x_k^i\}_{i=1}^{N_s}]$ and replace them by move of A_Σ^k using the Metropolis-Hasting method described above; do this again by move of γ_k. Note that this operation only changes the particles with $E_k^i = 1$.

Step 5 (state estimation). Estimate the posterior probability of target existence P_E^k by

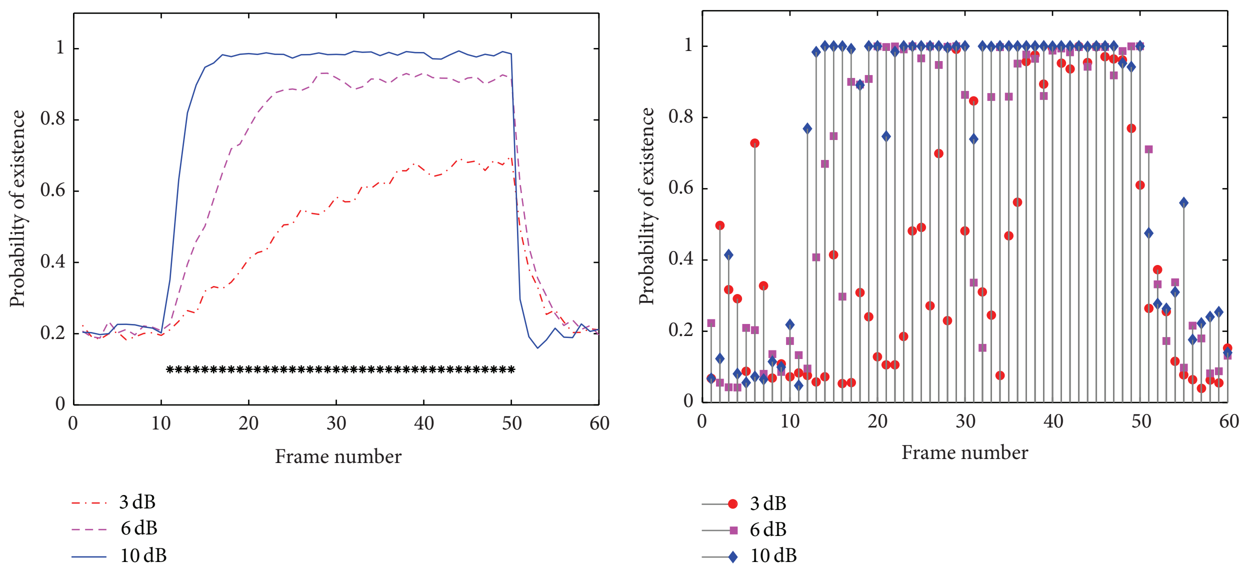
$$\widehat{P}_E^k = \frac{\sum_{i=1}^{N_s} E_k^i}{N_s}. \tag{27}$$

If $\widehat{P}_E^k$ exceeds a certain threshold $\mathrm{Th} \in (0, 1)$, target presence is declared, and then the target state is estimated by

$$\widehat{\mathbf{x}}_{k|k} = \frac{\sum_{i=1}^{N_s} \mathbf{x}_k^i E_k^i}{\sum_{i=1}^{N_s} E_k^i}. \tag{28}$$

To be more specific, some application issues are discussed as follows.

If there is no additional information, the birth density should be a uniform density over the surveillance region. For



(a) Averaged by 200 runs

(b) Single run

Figure 3: Probability of target existence under different SNRs, asterisk signs ($*$) at the bottom indicate the presence of the target.

example, for Doppler component, f_d^k, uniform samples are drawn from bins in the measurements which have amplitudes that exceed a predefined threshold. For echo amplitude A_k, the birth density is uniform over $[A_{\min}, A_{\max}]$, where $A_{\min}$ and $A_{\max}$ are expected minimum and maximum intensity levels, respectively. For monopulse ratio, γ_k, we assume that the target only exists within the half-power beamwidth, and from Figure 1 we can get that γ takes value within $[0, 0.8]$; thus, we choose its birth density to be uniform within $[0, 0.8]$. If other information is available (e.g., angle, range, or Doppler information supplied by the carrier aircraft, which usually has a normal law of error distribution and can be easily sampled as $q_b(\mathbf{x}_k \mid \mathbf{z}_k)$), the information should be used rather than the uniform one to improve the performance.

The bins in $C(\mathbf{x}_k)$ should be selected carefully, one practical choice is $C(\mathbf{x}_k) = \{i_0 - p, \ldots, i_0 - 1, i_0, i_0 + 1, \ldots, i_0 + p\}$, where i_0 is the bin nearest to the predicted $\mathbf{x}_k^i$ and p is a design parameter. Bins near the true Doppler position have comparatively higher amplitudes and can be beneficial to the performance, while the others will, on the contrary, deteriorate the performance because the signal amplitudes there are too low. As can be seen from Figure 2, the spread function for the points that are one bin away from the true position is below -20 dB; thus we choose $p = 1$ in this application.

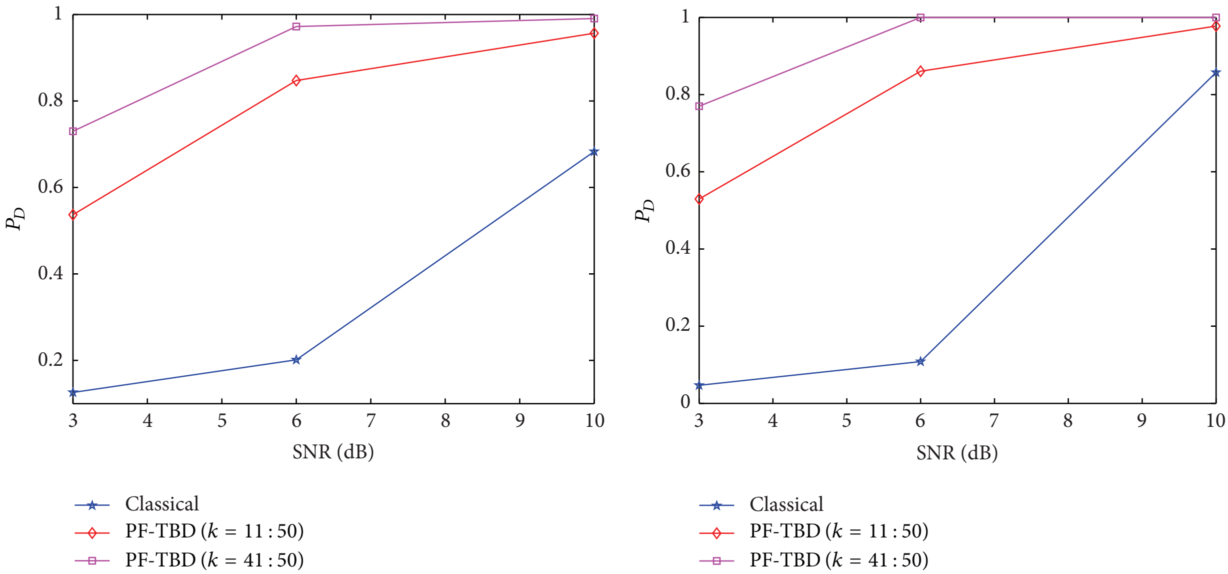
5. Experiments

5.1. Experiment 1: Stationary Scenario. The radar parameters are set as follows: the wavelength is $\lambda = 3$ cm, the PRI is $T_r = 4\,\mu$s, and the number of pulses per CPI is $N = 1000$. Hamming FFT windowing function is used. The target SNR represents the envelope of the target return compared to that of just noise. The SNR is measured after the entire coherent process (losses caused by windowing and straddle effect are considered). The initial relative velocity between target and radar is 1900 m/s. The initial monopulse ratio is 0.2. There are 368 bins in the clutter-free region. The initial amplitudes for 3, 6, and 10 dB are 0.87, 1.23, and 1.95, respectively. The levels of process noise used in the target model are $\sigma_{(1)}^2 = 0.01 \cdot \delta_f$, $\sigma_{(2)}^2 = 0.001$, and $\sigma_{(3)}^2 = 0.01$ (the SNR varies only marginally). The target is born at frame 11 and disappeared at frame 51.

The particle filter parameters are set as follows: the level of the process noise is perfectly matched to the simulated data, the probabilities of target "birth" P_b and "death" P_d are both set as 0.05, the initial target existence probability is $P_E^0 = 0.1$, the threshold $T_1 = 0.32$, and each bin of the point spread function is quantized into $N_{\text{app}} = 64$ points. The birth density $q_b(\mathbf{x}_0 \mid \mathbf{z}_0)$ is selected as follows: $A_\Sigma^0 \sim U(0.5, 3)$, $\gamma_0 \sim U(0, 0.8)$, and f_d^0 uniformly distributed in the clutter-free region. The variances of the proposal distributions in the MCMC move for A_Σ and γ are 0.04 and 0.01, respectively. $p = 1$ and 4000 particles are used.

Figure 3 shows the estimation result of the existence probability $\widehat{P}_E^k$; asterisk signs ($*$) at the bottom of the figure indicate the presence of the target. It can be seen that it is possible to detect target under an SNR as low as 3 dB. Setting the threshold Th $= 0.6$, for example, we can see that the target can be detected after several frames' accumulations. From Figure 3(b) we can see that the false alarms are isolated. Thus a binary integrator can be used to mitigate them and at the same time keep the successful detections, which are continuous after $\widehat{P}_E^k$ becomes stable.

Now we evaluate the detection performance of the PF-TBD algorithm in the detection terminologies. We estimate



(a) Single frame (b) Binary integration

Figure 4: Probability of detection. For single frame detection, $P_{\text{FA}} = 0.1$. For binary integration (3-out-of-5), the P_{FA} of classical detector is 0.02, while that of PF-TBD is 0 (no false alarm occurs in the 200 runs).

the probability of false alarm P_{FA} using frames 1 to 10 of the 200 Monte Carlo runs, where no target is present. More explicitly,

$$P_{\text{FA}} = \frac{1}{200 \times 10} \sum_{m=1}^{200} \sum_{k=1}^{10} \left(\widehat{P}_E^{k,m} \underset{0}{\overset{1}{\gtrless}} \text{Th} \right), \tag{29}$$

where m is the index of each Monte Carlo run. Similarly, P_D is computed when the target is present. To see performance in the stable region as well as in the whole region, we estimate P_D using frames 41 to 50 and frames 11 to 50, respectively. For example, P_D using frames 41 to 50 is

$$P_D = \frac{1}{200 \times 10} \sum_{m=1}^{200} \sum_{k=41}^{50} \left(\widehat{P}_E^{k,m} \underset{0}{\overset{1}{\gtrless}} \text{Th} \right). \tag{30}$$

For comparison, the classical detector is applied to the same data. Because the PF-TBD algorithm makes one decision in each frame, for a fairly comparison, the classical detector declares a detection once any bin in the clutter-free region exceeds the threshold Th'. Setting $P_{\text{FA}} = 0.1$ for both the PF-TBD and the classical detector (correspondingly, probability of false alarm for the classical detector in each single bin is 2.86×10^{-4} and the threshold for the PF-TBD is $\text{Th} = 0.45$), the P_D performances of them are shown in Figure 4(a). It can be observed that the P_D of PF-TBD in the stable region at 3 dB is better than that of the classical detector at 10 dB. Thus an SNR gain of up to 7 dB is obtained.

Taking results of Figure 4(a) as the primary detection results, we apply the 3-out-of-5 binary integration strategy to both the PF-TBD and the classical detector. Once 3 or more frames of consecutive 5 frames pass the primary detection, a secondary detection is declared. The resulting P_{FA} of the classical detector is 0.02, while that of the PF-TBD is 0 (no false alarm occurs in the 200 runs), which has proved that the binary integration after the PF-TBD performs well at false alarm mitigation. The P_D in binary integration is defined as the quotient of the number of secondary detections that have past the 3-out-of-5 logic divided by the total number of secondary detections. The P_D results are shown in Figure 4(b). We can see that the P_D improvement over the classical detector is more compared with the single frame detection even under lower P_{FA}.

Remark 1. As the number of Monte Carlo runs is comparatively small, these results are not intended to provide a performance assessment. More precise results can be attained by performing a large number of Monte Carlo simulations. Compared with the classical target detection problem, it seems more reasonable to define an index to describe the delay before the $\widehat{P}_E^k$ becomes stable and then evaluate the detection and estimation performances in the stable region.

5.2. Experiment 2: Maneuvering Target. Now we consider a real scenario on a 2D plane. As Figure 5 has shown, the missile performs a straight motion with its antenna direction 1 degree deviated off the south to the east side. After 10 noise only frames, the target enters the main beam of the seeker radar and performs a 2 s evasive maneuver. The trajectory of the target is generated by the simulation software JSBSim (http://jsbsim.sourceforge.net/). The target's velocity is about 280 m/s and its normal acceleration during the maneuver is 6 g. The missile's velocity is 1200 m/s and its monopulse sum and difference beam patterns are the same as those shown in Figure 1. The echo amplitude is inversely proportional to the square of the range between missile and target (the eclipsing

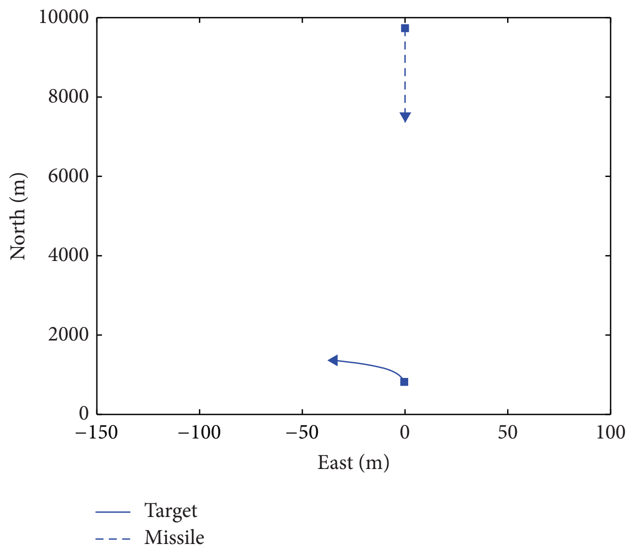


Figure 5: Missile and target trajectories. The "□" and "△" denote start and end of the trajectory, respectively.

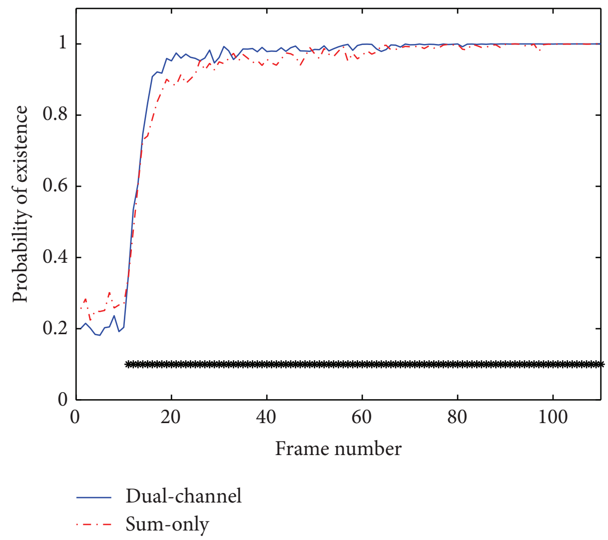


Figure 6: Probability of target existence (averaged by 100 runs), asterisk signs (∗) at the bottom indicate the presence of the target.

effect and the target fluctuation are not considered). The radar parameters are the same as those in Experiment 1 except that $N = 5000$; thus, the CPI is 20 ms and there are 100 target presented frames. Because the number of bins in the clutter-free region is too large, only 200 bins (bins from 3100 to 3300) containing the target are used. The initial SNR is 6 dB. The levels of process noise used in the particle filter are set as $\sigma^2_{(1)} = 5 \cdot \delta_f$, $\sigma^2_{(2)} = 0.05$, and $\sigma^2_{(3)} = 0.05$. The birth density $q_b(\mathbf{x}_0 \mid \mathbf{z}_0)$ is $A^0_{\Sigma} \sim U(1,4)$, $\gamma_0 \sim U(0.79, 0.8)$, and f^0_d uniformly distributed in the 200 bins. The other parameters of the particle filter are the same as Experiment 1. For comparison, the PF-TBD algorithm using sum channel only is also developed and tested using the same data. The sum-only PF-TBD is obtained through omitting the γ_k in the state vector and the filtering process. To distinguish them, the filter proposed in this paper is referred to as the dual-channel PF-TBD.

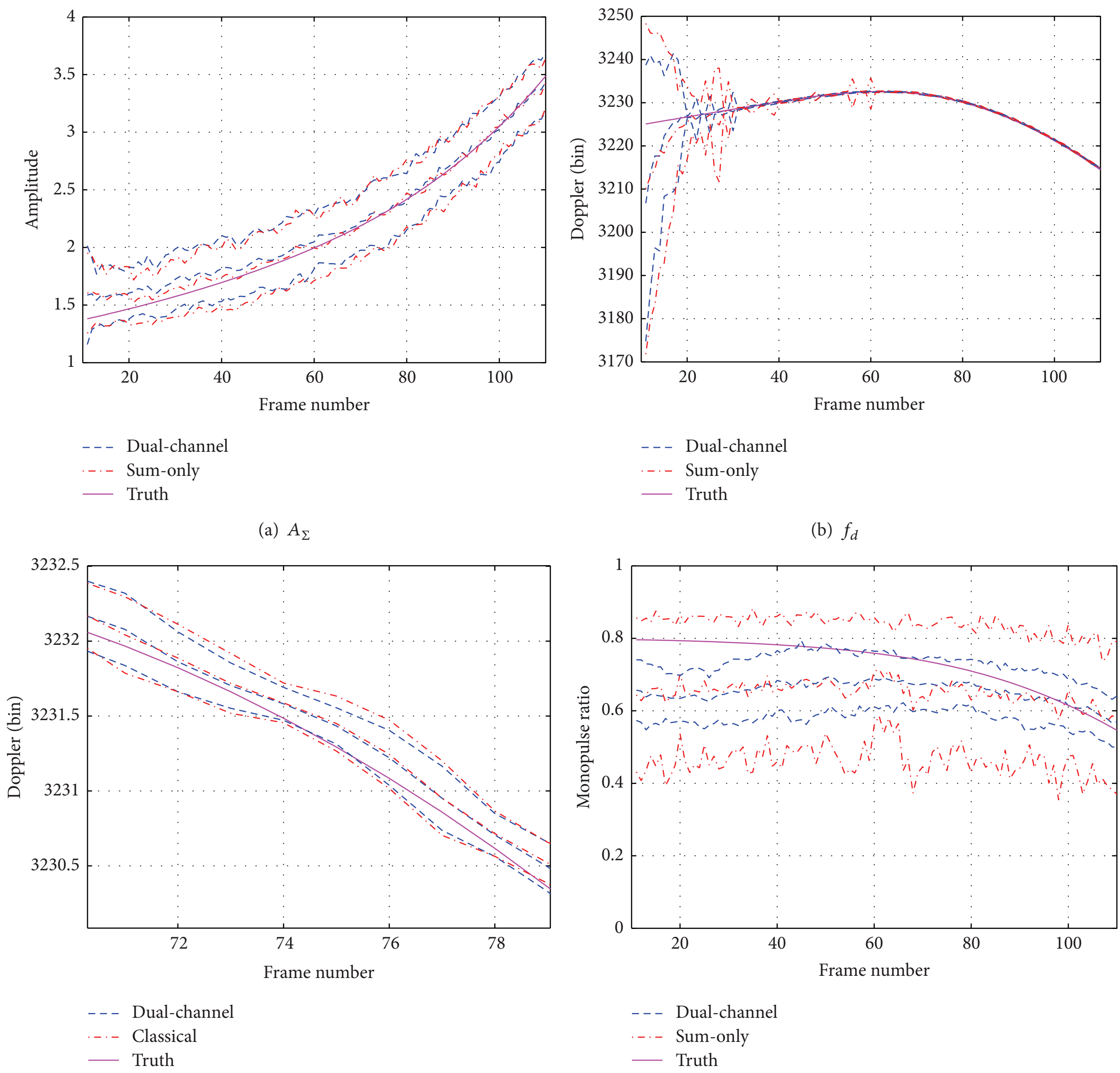
In Figure 6, the estimated probabilities of existence prove the effectiveness of the two filters in target detection. Note that the sum-only filter results in worse $\widehat{P}^k_E$ when the target is both absent and present, which means that its detection performance is worse than that of the dual-channel one. This is because the dual-channel PF-TBD benefits from the difference channel whose amplitude is high near the half-power point.

Figures 7(a)–7(c) present the state estimation results of the two filters. We can see that both of the two filters can successfully track in target maneuvering. The dual-channel filter has better Doppler estimation performance. Note that the target Doppler can travel across half the bin size per frame; the binary integration of the classical detector will fail while that of the PF-TBD is unaffected. As the sum-only filter does not output the monopulse estimation result, the monopulse estimation performance of the dual-channel PF-TBD is compared with the classical single frame monopulse estimation method as shown in Figure 7(d). To use the same a priori knowledge, the result of the classical method is constrained to be within (0, 0.8) and that is why its estimation result is biased. The classical method assumes index of the bin which contains the target is known while the PF-TBD does not use this information. In spite of this, the monopulse estimation performance of the dual-channel PF-TBD is better.

Remark 2. This example shows that the detection performance can be improved by using the difference channel when the target is near beam edge. When the target is at the beam center, however, the difference channel amplitude is approximately zero as can be seen from Figure 1. Then the detection performance may be deteriorated instead compared with the sum-only PF-TBD. In fact, through simulation we have found that when $\gamma > 0.1$, detection performance of the dual-channel PF-TBD is better. In practical application, the two methods should be selected according to the scenario (e.g., whether there is precise angular targeting information), and the estimation performance should also be taken into account.

6. Conclusions and Future Work

Using PF-TBD in monopulse high PRF pulse Doppler radar to improve detection and estimation performances under low SNR is addressed in this paper. The target and measurement models are analyzed and defined for this application. Based on them, a PF-TBD algorithm with resample-move operations is developed. Extensive simulations have shown that the proposed algorithm can improve both the detection and estimation performances compared with the classical and sum-only methods. To further improve the detection performance, binary integration after the PF-TBD is proposed. Simulation



(a) A_Σ (b) f_d (c) Detail of f_d (d) γ

FIGURE 7: Estimation results of A_Σ, f_d, and γ. The thick dashed lines show the mean value over 100 Monte Carlo runs. The thin dashed lines are mean ± one standard deviation.

result shows that it can effectively mitigate the false alarms in the PF-TBD detection result.

As a byproduct of the PF-TBD algorithm, the estimated amplitude can be used to predict range eclipsing and to estimate the SNR. Application of the PF-TBD requires exact knowledge of the thermal noise power, which can be estimated on-the-fly before the PF-TBD is enabled. For seekers incorporating multispectral sensors, targeting information (e.g., angular information of the target, probability of existence of target in the main beam) from other sensors like the infrared sensor or the passive radar can be fused easily as Section 4 has stated.

Conflict of Interests

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

References

[1] M. I. Skolnik, *Radar Handbook*, McGraw-Hill, New York, NY, USA, 3rd edition, 2008.

[2] B. D. Carlson, E. D. Evans, and S. L. Wilson, "Search radar detection and track with the hough transform I: system concept," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 1, pp. 102–108, 1994.

[3] Y. Barniv, "Dynamic programming solution for detecting dim moving targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 21, no. 1, pp. 144–156, 1985.

[4] J. David, R. Kramer, and W. S. Reid, "Track-before-detect processing for a range-ambiguous radar," in *Proceedings of the IEEE National Radar Conference*, pp. 113–116, April 1993.

[5] W. Wallace, "The use of track-before-detect in pulse-doppler radar," in *RADAR 2002*, pp. 315–319, October 2002.

[6] S. Buzzi, M. Lops, and L. Venturino, "Track-before-detect procedures for early detection of moving target from airborne radars," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 3, pp. 937–954, 2005.

[7] D. Salmond and H. Birch, "A particle filter for track-before-detect,," in *Proceedings of the American Control Conference*, Arlington, Va, USA, 2001.

[8] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter-particle Filter for Tracking Applications*, Artech House, Boston, MA, USA, 2004.

[9] Y. Boers and J. Driessen, "Multitarget particle filter track before detect application," *IEE Proceedings Radar, Sonar and Navigation*, vol. 151, no. 6, pp. 351–357, 2004.

[10] M. Rutten, N. Gordon, and S. Maskell, "Recursive track-before-detect with target amplitude fluctuations," *IEE Proceedings Radar, Sonar and Navigatio*, vol. 152, no. 5, pp. 345–352, 2005.

[11] Y. Boers and J. Driessen, "Particle filter based detection for tracking," in *Proceedings of the American Control Conference*, pp. 4393–4397, Arlington, Va, USA, June 2001.

[12] H.-T. Su, T.-P. Wu, H.-W. Liu, and Z. Bao, "Rao-blackwellised particle filter based track-before-detect algorithm," *IET Signal Processing*, vol. 2, no. 2, pp. 169–176, 2008.

[13] B. K. Habtemariam, R. Tharmarasa, and T. Kirubarajan, "Multitarget track before detect with MIMO radars," in *Proceedings of the IEEE Aerospace Conference*, pp. 1–9, IEEE, Big Sky, Mont, USA, March 2010.

[14] E. Mosca, "Angle estimation in amplitude comparison monopulse systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 5, no. 2, pp. 205–212, 1969.

[15] D. Musicki, R. Evans, and S. Stankovic, "Integrated probabilistic data association," *IEEE Transactions on Automatic Control*, vol. 39, no. 6, pp. 1237–1241, 1994.

[16] W. R. Gilks and C. Berzuini, "Following a moving target—Monte Carlo inference for dynamic Bayesian models," *Journal of the Royal Statistical Society B*, vol. 63, no. 1, pp. 127–146, 2001.

[17] C. Berzuini and W. Gilks, "Resample-move filtering with cross-model jumps," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Feitas, and N. Gordon, Eds., Stat. Eng. Inf. Sci., chapter 6, pp. 117–138, Springer, New York, NY, USA, 2001.

[18] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 205967, 9 pages
http://dx.doi.org/10.1155/2014/205967

Research Article

Algorithms to Solve Stochastic H_2/H_∞ Control with State-Dependent Noise

Ming Gao,[1,2] Changdi Fu,[2] and Weihai Zhang[2]

[1] *College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China*
[2] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Ming Gao; yuming_gao@163.com

Received 24 April 2014; Accepted 22 June 2014; Published 17 August 2014

Academic Editor: Wuquan Li

This paper is concerned with the algorithms which solve H_2/H_∞ control problems of stochastic systems with state-dependent noise. Firstly, the algorithms for the finite and infinite horizon H_2/H_∞ control of discrete-time stochastic systems are reviewed and studied. Secondly, two algorithms are proposed for the finite and infinite horizon H_2/H_∞ control of continuous-time stochastic systems, respectively. Finally, several numerical examples are presented to show the effectiveness of the algorithms.

1. Introduction

Mixed H_2/H_∞ control is an important robust control method and has been extensively investigated by many researchers [1–4]. Compared with the sole H_∞ control, the mixed H_2/H_∞ control is more attractive in engineering practice [4], since the former is a worst-case design which tends to be conservative while the latter minimizes the average performance with a guaranteed worst-case performance. Recently, stochastic H_2/H_∞ control for continuous- and discrete-time systems with multiplicative noise has become a popular topic and has attracted a lot of attention [5–7]. In [5], the finite and infinite horizon H_2/H_∞ control problems were discussed for continuous-time stochastic systems with state-dependent noise. The finite and infinite horizon H_2/H_∞ control problems were solved for discrete-time stochastic systems with state and disturbance dependent noise by [6] and [7], respectively. Moreover, mixed H_2/H_∞ control was widely studied for stochastic systems with Markov jumps and multiplicative noise [8–11] due to their powerful modeling ability in many fields [12, 13].

Generally, the existence of a H_2/H_∞ controller is equivalent to the solvability of several coupled matrix-valued equations. However, it is difficult to solve these coupled matrix-valued equations analytically. Several numerical algorithms have appeared in dealing with deterministic and stochastic H_2/H_∞ control. In [1], the finite horizon H_2/H_∞ controller for continuous-time deterministic systems was obtained by using the Runge-Kutta integration procedure. In [14], an exact solution to the suboptimal deterministic H_2/H_∞ control problem was studied via convex optimization. Two iterative algorithms were proposed for finite and infinite horizon H_2/H_∞ control of discrete-time stochastic systems in [6] and [7], respectively. In [15], an iterative algorithm was proposed to solve a kind of stochastic algebraic Riccati equation in LQ zero-sum game problems.

However, most of these algorithms were concerned with the H_2/H_∞ control for discrete-time systems. Up to now, the algorithm for stochastic H_2/H_∞ control of continuous-time systems has received little research attention. This is because the coupled matrix-valued equations for the continuous-time H_2/H_∞ control cannot be solved by recursive algorithms as in the discrete-time case. In this paper, we will study the algorithms to solve H_2/H_∞ control problems for stochastic systems with state-dependent noise. Firstly, the algorithms for finite and infinite horizon H_2/H_∞ control of discrete-time stochastic systems are reviewed. An iterative algorithm is presented to solve the infinite horizon H_2/H_∞ control of discrete-time time-varying stochastic systems. For continuous-time stochastic systems, two algorithms are proposed for the finite and infinite horizon H_2/H_∞ control, respectively. Some numerical examples are presented to illustrate the developed algorithms.

For conveniences, we make use of the following notations throughout this paper: $\mathscr{R}^n$: n-dimensional Euclidean space; $\mathcal{S}^n$: the set of all $n\times n$ symmetric matrices; $A>0$ $(A\geq 0)$: A is a positive definite (positive semidefinite) symmetric matrix; A': the transpose of a matrix A; I: the identity matrix; $\mathrm{Tr}[A]$: the trace of matrix A; $E(x)$: the mathematical expectation of x.

2. Preliminaries

In this section, we will present some preliminary results for stochastic H_2/H_∞ control, including the finite horizon case for discrete-time time-varying systems, the infinite horizon case for discrete-time time-invariant systems, the finite horizon case for continuous-time time-varying systems, and the infinite horizon case for continuous-time time-invariant systems.

Consider the following discrete-time time-varying stochastic system with state-dependent noise:

$$\begin{gathered} x_{k+1}=A_kx_k+B_ku_k+C_kv_k+D_kx_kw_k,\\ z_k=\begin{bmatrix}F_kx_k\\ u_k\end{bmatrix},\quad x_0\in\mathscr{R}^n,\ k=0,1,\ldots,T, \end{gathered}\tag{1}$$

where $x_k\in\mathscr{R}^n$, $u_k\in\mathscr{R}^{n_u}$, $v_k\in\mathscr{R}^{n_v}$, and $y_k\in\mathscr{R}^{n_y}$ are, respectively, the system state, control, disturbance signal, and output; $\{w_k\}_{k\geq 0}\in\mathscr{R}$ is a sequence of independent white noise defined on the filtered probability space $(\Omega,\mathscr{F},\mathscr{F}_k,\mathscr{P})$ with $E(w_k)=0$ and $E(w_kw_s)=\delta_{ks}$, where δ_{ks} is a Kronecker function defined by $\delta_{ks}=0$ for $k\neq s$ while $\delta_{ks}=1$ for $k=s$. x_0 is assumed to be deterministic for simplicity purposes. A_k, B_k, C_k, D_k, and F_k are matrix-valued continuous functions of appropriate dimensions.

Lemma 1 (see [6]). *For given $\gamma>0$, the finite horizon H_2/H_∞ control for system* (1) *has solutions (u^*,v^*) as $u^*(k,x_k)=K_{2k}x_k$, and $v^*(k,x_k)=K_{1k}x_k$, with K_{1k} and K_{2k} being continuous matrix-valued functions, if and only if the following coupled difference matrix-valued equations*

$$\begin{gathered} -P_{1k}+(A_k+B_kK_{2k})'P_{1,k+1}(A_k+B_kK_{2k})+D_kP_{1,k+1}D_k\\ -F_k'F_k-K_{2k}'K_{2k}-\left[(A_k+B_kK_{2k})'P_{1,k+1}C_k\right]\Gamma_1(P_{1,k+1})^{-1}\\ \times\left[(A_k+B_kK_{2k})'P_{1,k+1}C_k\right]'=0,\\ P_{1,T+1}=0,\\ \Gamma_1(P_{1,k+1})=\gamma^2I+C_k'P_{1,k+1}C_k>0, \end{gathered}\tag{2}$$

$$K_{1k}=-\Gamma_1(P_{1,k+1})^{-1}\left[(A_k+B_kK_{2k})'P_{1,k+1}C_k\right]',\tag{3}$$

$$\begin{gathered} -P_{2k}+(A_k+C_kK_{1k})'P_{2,k+1}(A_k+C_kK_{1k})+D_kP_{2,k+1}D_k\\ +F_k'F_k-\left[(A_k+C_kK_{1k})'P_{2,k+1}B_k\right]\Gamma_2(P_{2,k+1})^{-1}\\ \times\left[(A_k+C_kK_{1k})'P_{2,k+1}B_k\right]'=0,\\ P_{2,T+1}=0, \end{gathered}\tag{4}$$

$$K_{2k}=-\Gamma_2(P_{2,k+1})^{-1}\left[(A_k+C_kK_{1k})'P_{2,k+1}B_k\right]'\tag{5}$$

with $\Gamma_2(P_{2,k+1})=I+B_k'P_{2,k+1}B_k$, admit a bounded solution $(P_{1k}\leq 0,K_{1k};P_{2k}\geq 0,K_{2k})$, $k=0,1,\ldots,T$.

Consider the following discrete-time time-invariant stochastic systems with state-dependent noise:

$$\begin{gathered} x_{k+1}=Ax_k+Bu_k+Cv_k+Dx_kw_k,\\ z_k=\begin{bmatrix}Fx_k\\ u_k\end{bmatrix},\quad x_0\in\mathscr{R}^n,\ k=0,1,2,\ldots. \end{gathered}\tag{6}$$

Briefly, system (6) can be denoted by $(A,B,C;D\mid F)$, and similar notations will be used in the following section.

Lemma 2 (see [7]). *Suppose that $(A;D\mid F)$ and $(A+BK_1;D\mid F)$ are exactly observable. For given $\gamma>0$, the infinite horizon H_2/H_∞ control for system* (6) *has solutions (u^*,v^*) as $u_k^*=K_2x_k$, and $v_k^*=K_1x_k$, if and only if the following coupled algebraic matrix-valued equations*

$$\begin{gathered} -P_1+(A+BK_2)'P_1(A+BK_2)+DP_1D-F'F-K_2'K_2\\ -\left[(A+BK_2)'P_1C\right]\Gamma_1(P_1)^{-1}\left[(A+BK_2)'P_1C\right]'=0,\\ \Gamma_1(P_1)=\gamma^2I+C'P_1C>0, \end{gathered}\tag{7}$$

$$K_1=-\Gamma_1(P_1)^{-1}\left[(A+BK_2)'P_1C\right]',\tag{8}$$

$$\begin{gathered} -P_2+(A+CK_1)'P_2(A+CK_1)+DP_2D+F'F\\ -\left[(A+CK_1)'P_2B\right]\Gamma_2(P_2)^{-1}\left[(A+CK_1)'P_2B\right]'=0, \end{gathered}\tag{9}$$

$$K_2=-\Gamma_2(P_2)^{-1}\left[(A+CK_1)'P_2B\right]'\tag{10}$$

with $\Gamma_2(P_2)=I+B'P_2B$, have a solution $(P_1<0,K_1;P_2>0,K_2)$.

Consider the following continuous-time time-varying stochastic system with state-dependent noise:

$$\begin{gathered} dx(t)=[A(t)x(t)+B(t)u(t)+C(t)v(t)]\,dt\\ +D(t)x(t)\,dw(t),\\ z(t)=\begin{bmatrix}F(t)x(t)\\ u(t)\end{bmatrix},\quad x(0)=x_0\in\mathscr{R}^n,\ t\in[0,T], \end{gathered}\tag{11}$$

where $x(t)\in\mathscr{R}^n$, $u(t)\in\mathscr{R}^{n_u}$, $v(t)\in\mathscr{R}^{n_v}$, and $y(t)\in\mathscr{R}^{n_y}$ are, respectively, the system state, control, disturbance signal, and output. $w(t)$ is a standard one-dimensional Wiener process defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{P})$ with $\mathcal{F}_t = \sigma\{w(s) : 0 \leq s \leq t\}$. x_0 is assumed to be deterministic for simplicity purposes. $A(t)$, $B(t)$, $C(t)$, $D(t)$, and $F(t)$ are matrix-valued continuous functions of suitable dimensions.

Lemma 3 (see [5]). *For given $\gamma > 0$, the finite horizon H_2/H_∞ control for system* (11) *has solutions (u^*, v^*) as $u^*(t, x_t) = K_2(t)x(t)$, and $v^*(t, x_t) = K_1(t)x(t)$, with $K_1(t)$ and $K_2(t)$ being continuous matrix-valued functions, if and only if the following coupled differential matrix-valued equations*

$$\begin{aligned}
-\dot{P}_1(t) &= P_1(t)A(t) + A'(t)P_1(t) + D'(t)P_1(t)D(t) \\
&\quad - F'(t)F(t) - [P_1(t) \;\; P_2(t)] \\
&\quad \times \begin{bmatrix} \gamma^{-2}C(t)C'(t) & B(t)B'(t) \\ B(t)B'(t) & B(t)B'(t) \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix}, \\
P_1(T) &= 0,
\end{aligned} \tag{12}$$

$$\begin{aligned}
-\dot{P}_2(t) &= P_2(t)A(t) + A'(t)P_2(t) + D'(t)P_2(t)D(t) \\
&\quad + F'(t)F(t) - [P_1(t) \;\; P_2(t)] \\
&\quad \times \begin{bmatrix} 0 & \gamma^{-2}C(t)C'(t) \\ \gamma^{-2}C(t)C'(t) & B(t)B'(t) \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix} \\
P_2(T) &= 0
\end{aligned} \tag{13}$$

have a bounded solution $(P_1(t) \leq 0, K_1(t); P_2(t) \geq 0, K_2(t))$, with $K_1(t) = -\gamma^{-2}C(t)P_1(t)$ and $K_2(t) = -B(t)P_2(t)$, $t \in [0, T]$.

Consider the following continuous-time time-invariant stochastic system with state-dependent noise:

$$\begin{aligned}
dx(t) &= [Ax(t) + Bu(t) + Cv(t)]\,dt + Dx(t)\,dw(t), \\
z(t) &= \begin{bmatrix} Fx(t) \\ u(t) \end{bmatrix}, \quad x(0) = x_0 \in \mathcal{R}^n, \; t \in [0, \infty).
\end{aligned} \tag{14}$$

Lemma 4 (see [5]). *Suppose that $(A; D \mid F)$ and $(A + BK_1; D \mid F)$ are exactly observable. For given $\gamma > 0$, the infinite horizon H_2/H_∞ control for system* (14) *has solutions (u^*, v^*) as $u^*(t) = K_2x(t)$ and $v^*(t) = K_1x(t)$, if and only if the following coupled algebraic matrix-valued equations*

$$P_1A + A'P_1 + D'P_1D - F'F - [P_1 \;\; P_2] \begin{bmatrix} \gamma^{-2}CC' & BB' \\ BB' & BB' \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = 0, \tag{15}$$

$$P_2A + A'P_2 + D'P_2D + F'F - [P_1 \;\; P_2] \begin{bmatrix} 0 & \gamma^{-2}CC' \\ \gamma^{-2}CC' & BB' \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = 0 \tag{16}$$

have a solution $(P_1 \leq 0, K_1; P_2 \geq 0, K_2)$, with $K_1 = -\gamma^{-2}CP_1$, $K_2 = -BP_2$.

3. Discrete-Time Case

In [6, 7], Zhang et al. provided the recursive algorithms to solve the coupled matrix-valued equations in Lemmas 1 and 2, respectively. Based on those results, this paper will present an algorithm to solve the infinite horizon H_2/H_∞ control of discrete-time time-varying stochastic systems.

The following algorithm can be used to solve the coupled difference matrix-valued equations (2)–(5) in Lemma 1 [6].

Algorithm 5. Consider the following.

(i) Set $k = T$, then $\Gamma_1(P_{1,T+1})$ and $\Gamma_2(P_{2,T+1})$ can be computed according to the final conditions $P_{1,T+1} = 0$ and $P_{2,T+1} = 0$.

(ii) Solve the matrix recursions (3) and (5), then K_{1T} and K_{2T} are derived.

(iii) Substituting the obtained K_{2T} and K_{1T} into the matrix recursions (2) and (4), respectively, then $P_{1T} \leq 0$ and $P_{2T} \geq 0$ are available.

(iv) Repeat the above procedures; $(P_{1k}, K_{1k}; P_{2k}, K_{2k})$ can be computed for $k = T - 1, T - 2, \ldots, 0$, recursively.

In Algorithm 5, the priori condition $\Gamma_1(P_{1k}) > 0$ should be checked to guarantee it to proceed backward. Otherwise, the algorithm has to stop. It is noted that $\Gamma_1(P_{1,T+1})$ and $\Gamma_2(P_{2,T+1})$ can be computed first, provided that the final conditions $P_{1,T+1} = 0$ and $P_{2,T+1} = 0$ are known.

The following algorithm can be used to solve the coupled algebraic matrix-valued equations (7)–(10) in Lemma 2 [7].

Algorithm 6. Consider the following.

(i) Establish difference equations (2)–(5) corresponding to algebraic equations (7)–(10).

(ii) Give a large T. By means of Algorithm 5, the difference equations (2)–(5) can be solved and $(P_{1k}, K_{1k}; P_{2k}, K_{2k})$ $k = T, T - 1, T - 2, \ldots, 0$ can be derived.

(iii) If the sequences $(P_{1k}, K_{1k}; P_{2k}, K_{2k})k = T, T - 1, T - 2, \ldots, 0$ are convergent, then (7)–(10) have solutions $(P_1, K_1; P_2, K_2) = (P_{10}, K_{10}; P_{20}, K_{20})$. Otherwise, the problem is unsolvable.

In [10], a necessary and sufficient condition for the infinite horizon H_2/H_∞ control problem of discrete-time time-varying stochastic systems with Markov jumps was derived in terms of four coupled discrete-time Riccati equations. However, the Riccati equations in [10] were solved by trial and error and cannot be extended to the complicated case. The condition for the infinite horizon H_2/H_∞ control of time-varying stochastic system $(A_k, B_k, C_k; D_k \mid F_k)$ (or system (1)) is as follows.

Lemma 7. *For systems $(A_k, B_k, C_k; D_k \mid F_k)$, assume that $(A_k; D_k \mid F_k)$ and $(A_k + B_kK_{1k}; D_k \mid F_k)$ are stochastically detectable. The infinite horizon H_2/H_∞ control problem has solutions $u^*(k, x_k) = K_{2k}x_k$, and $v^*(k, x_k) = K_{1k}x_k$, with K_{1k}*

and K_{2k} being continuous matrix-valued functions, if and only if the following coupled difference matrix-valued equations

$$\begin{aligned} & -P_{1k} + (A_k + B_k K_{2k})' P_{1,k+1} (A_k + B_k K_{2k}) \\ & \quad + D_k P_{1,k+1} D_k - F_k' F_k - K_{2k}' K_{2k} \\ & \quad - \left[(A_k + B_k K_{2k})' P_{1,k+1} C_k \right] \Gamma_1 (P_{1,k+1})^{-1} \\ & \quad \times \left[(A_k + B_k K_{2k})' P_{1,k+1} C_k \right]' = 0, \\ & \quad \Gamma_1 (P_{1,k+1}) = \gamma^2 I + C_k' P_{1,k+1} C_k > 0, \end{aligned} \tag{17}$$

$$K_{1k} = -\Gamma_1 (P_{1,k+1})^{-1} \left[(A_k + B_k K_{2k})' P_{1,k+1} C_k \right]', \tag{18}$$

$$\begin{aligned} & -P_{2k} + (A_k + C_k K_{1k})' P_{2,k+1} (A_k + C_k K_{1k}) + D_k P_{2,k+1} D_k \\ & \quad + F_k' F_k - \left[(A_k + C_k K_{1k})' P_{2,k+1} B_k \right] \Gamma_2 (P_{2,k+1})^{-1} \\ & \quad \times \left[(A_k + C_k K_{1k})' P_{2,k+1} B_k \right]' = 0, \end{aligned} \tag{19}$$

$$K_{2k} = -\Gamma_2 (P_{2,k+1})^{-1} \left[(A_k + C_k K_{1k})' P_{2,k+1} B_k \right]' \tag{20}$$

with $\Gamma_2(P_{2,k+1}) = I + B_k' P_{2,k+1} B_k$, *admit a bounded solution* $(P_{1k} \leq 0, K_{1k}; P_{2k} \geq 0, K_{2k})$, $k = 0, 1, 2, \ldots$.

Proof. This is a direct corollary of Theorem 2 in [10] and the proof is omitted. □

In this paper, the essential difference between Lemmas 1 and 7 is that k is finite in the former while it is infinite in the later. Based on Algorithm 6, the coupled matrix-valued equations (17)–(20) can be solved by the following recursive algorithm.

Algorithm 8. Consider the following.

(i) Given $k = k_1$, (17)–(20) reduce to time-invariant matrix-valued equations.

(ii) Compute the solution of these time-invariant matrix-valued equations by using Algorithm 6.

(iii) Set $k_1 = k_1 + 1$ and go to step 1.

It is difficult for Algorithm 8 to compute all the solutions as $k \to \infty$ for general time-varying system. However, it is easy to verify that the solutions of (17)–(20) are also periodic for periodic systems. Hence, Algorithm 8 is suitable for the periodic case, which will be shown by Example 1.

4. Continuous-Time Case

In contrast to the discrete-time case, it is more difficult to deal with the continuous-time stochastic H_2/H_∞ control in Lemmas 3 and 4. In this study, the Runge-Kutta integration procedure and the convex optimization approach are applied to solve the coupled matrix-valued equations in Lemmas 3 and 4, respectively.

In Lemma 3, the coupled differential matrix-valued equations (12) and (13) can be viewed as a set of backward differential equations with known terminal conditions, which can be solved by the Runge-Kutta integration procedure [1]. The following algorithm can be used to solve (12) and (13) in Lemma 3.

Algorithm 9. Consider the following.

(i) Rewrite (12) and (13) as a set of equations with $(n(n+1)/2) \times 2$ time-varying backward differential equations with known terminal conditions.

(ii) Solve this set of equations by using the Runge-Kutta integration procedure.

(iii) If the solutions of the set of equations are convergent, then the finite horizon H_2/H_∞ control problem is solvable. Otherwise, the problem is unsolvable.

Next, we will study the algorithm for the solution of coupled algebraic matrix-valued equations (15) and (16) in Lemma 4. In the scalar case, the curves represented by (15) and (16) can be plotted in a (P_1, P_2)-plane, and the intersections of these curves, if they exist, are the solutions of (15) and (16). Moreover, the intersection in the second quadrant is the solution that we need, which will be shown in Example 2.

In the high-dimensional case, a suboptimal H_2/H_∞ controller design algorithm for Lemma 4 was obtained in [8] by solving a convex optimization problem. However, this algorithm was developed under the assumption $P_1 = -P_2$ which was very conservative. Rewrite (15) and (16) as

$$\begin{aligned} \Theta_1 &= P_1 A + A' P_1 + D' P_1 D - F' F - P_2 B B' P_2 \\ & \quad - \gamma^{-2} P_1 C C' P_1 - P_1 B B' P_2 - P_2 B B' P_1 = 0, \\ \Theta_2 &= P_2 A + A' P_2 + D' P_2 D + F' F - P_2 B B' P_2 \\ & \quad - \gamma^{-2} P_1 C C' P_2 - \gamma^{-2} P_2 C C' P_1 = 0. \end{aligned} \tag{21}$$

Substituting $P_1 = -P_2$ into (21) yields

$$\begin{aligned} & P_2 A + A' P_2 + D' P_2 D + F' F - P_2 B B' P_2 + \gamma^{-2} P_2 C C' P_2 = 0, \\ & P_2 A + A' P_2 + D' P_2 D + F' F - P_2 B B' P_2 + 2\gamma^{-2} P_2 C C' P_2 = 0. \end{aligned} \tag{22}$$

From the above, it can be seen that one matrix P_2 cannot satisfy two different equations simultaneously expect in some very special cases.

In this paper, we try to present another convex optimization algorithm to solve (15) and (16). By Theorem 10 of [16], $(P_1, P_2) \in \mathcal{S}^n \times \mathcal{S}^n$ is the optimal solution to

$$\max_{\text{s.t. } \Theta_1 \geq 0, \Theta_2 \geq 0, P_1 \leq 0, P_2 \geq 0} \sum_{i=1}^{N} \operatorname{Tr}\left[P_1 + P_2 \right]. \tag{23}$$

Since

$$-P_1BB'P_2 - P_2BB'P_1 \geq -P_1BB'P_1 - P_2BB'P_2,$$
$$-\gamma^{-2}P_1CC'P_2 - \gamma^{-2}P_2CC'P_1 \geq -\gamma^{-2}P_1CC'P_1 - \gamma^{-2}P_2CC'P_2, \quad (24)$$

we have $\Theta_1 \geq 0$ and $\Theta_2 \geq 0$ if

$$\begin{aligned}\overline{\Theta}_1 &= P_1A + A'P_1 + D'P_1D - F'F - 2P_2BB'P_2 \\ &\quad - P_1\left(BB' + \gamma^{-2}CC'\right)P_1 \\ &\geq 0, \\ \overline{\Theta}_2 &= P_2A + A'P_2 + D'P_2D + F'F - \gamma^{-2}P_1CC'P_1 \\ &\quad - P_2\left(BB' + \gamma^{-2}CC'\right)P_2 \\ &\geq 0,\end{aligned} \quad (25)$$

respectively. According to Schur's complement lemma, $\overline{\Theta}_1 \geq 0$ and $\overline{\Theta}_2 \geq 0$ are, respectively, equivalent to

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & P_2B \\ \Sigma_{12}' & I & 0 \\ B'P_2 & 0 & \frac{1}{2}I \end{bmatrix} \geq 0,$$
$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & P_1C \\ \Omega_{12}' & I & 0 \\ C'P_1 & 0 & \gamma^2 I \end{bmatrix} \geq 0, \quad (26)$$

with

$$\begin{aligned}\Sigma_{11} &= P_1A + A'P_1 + D'P_1D - F'F, \\ \Sigma_{12} &= P_1\left(BB' + \gamma^{-2}CC'\right)^{1/2}, \\ \Omega_{11} &= P_2A + A'P_2 + D'P_2D + F'F, \\ \Omega_{12} &= P_2\left(BB' + \gamma^{-2}CC'\right)^{1/2}.\end{aligned} \quad (27)$$

Since (26) are linear matrix inequalities (LMIs), a suboptimal solution to coupled matrix-valued equations (21) may be derived by solving the following convex optimization problem:

$$\max_{\text{s.t. } \Sigma\geq 0,\, \Omega\geq 0,\, P_1\leq 0,\, P_2\geq 0} \sum_{i=1}^{N} \operatorname{Tr}\left[P_1 + P_2\right]. \quad (28)$$

Moreover, the infinite horizon H_2/H_∞ control problem of system (14) has a pair of solutions:

$$\begin{aligned}u^*(t) &= K_2x(t) = -B'P_2x(t), \\ v^*(t) &= K_1x(t) = -\gamma^{-2}C'P_1x(t).\end{aligned} \quad (29)$$

Summarizing the above, the following algorithm can be used to solve (15) and (16) in Lemma 4.

Algorithm 10. Consider the following.

(i) Establish LMIs (26) corresponding to algebraic equations (15) and (16) in Lemma 4.

(ii) If the convex optimization problem (28) is solvable, then P_1 and P_2 can be derived. Moreover, $K_1 = -\gamma^{-2}C'P_1$ and $K_2 = -B'P_2$ can be computed. Otherwise, (15) and (16) in Lemma 4 are unsolvable.

Remark 11. Note that, in Algorithm 10, conditions (26) are given in terms of linear matrix inequalities; therefore, by using the Matlab LMI-Toolbox, it is straightforward to check the feasibility of the convex optimization problem (28) without tuning any parameters. In fact, Algorithm 10 is also a suboptimal algorithm, and the conservatism comes from the inequality transforms (24).

Remark 12. In this paper, we consider the H_2/H_∞ control for stochastic systems with only state-dependent noise. As discussed in [17, 18], for most natural phenomena described by Itô stochastic systems, not only state but also control input or external disturbance maybe corrupted by noise. Therefore, it is necessary to study stochastic systems with state, control, and disturbance-dependent noise which makes the conditions for H_2/H_∞ control more complicated. Searching for the numerical solutions for these conditions deserves further study.

5. Numerical Examples

In this section, several numerical examples will be provided to illustrate the effectiveness of Algorithms 8–10.

Example 1. Consider the infinite horizon H_2/H_∞ control for two-dimensional periodic stochastic systems (1) with the following parameters:

$$A_k = \begin{bmatrix} 0.2 & 0 \\ -1 & 0.2 * (-1)^k \end{bmatrix}, \qquad B_k = \begin{bmatrix} 0.5 \\ (-1)^k \end{bmatrix},$$
$$C_k = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \qquad D_k = \begin{bmatrix} 0.5 * (-1)^k & 0 \\ 0 & -0.2 \end{bmatrix}, \quad (30)$$

$F_k = \begin{bmatrix} 0 & 1 \end{bmatrix}$. Apparently, the period of this system is $\tau = 2$. By setting $\gamma = 1.8$ and applying Algorithm 8, the evolutions of P_{1k}, P_{2k}, K_{1k}, K_{2k}, $k = 1, 2$ are illustrated in Figures 1 and 2, respectively, which clearly show the convergence of the algorithm.

Example 2. Consider the finite horizon H_2/H_∞ control for the following one-dimensional stochastic system:

$$dx(t) = [2x(t) + 3u(t) + v(t)]\,dt + x(t)\,dw(t),$$
$$z(t) = \begin{bmatrix} 3x(t) \\ u(t) \end{bmatrix}, \quad t \in [0, 2]. \quad (31)$$

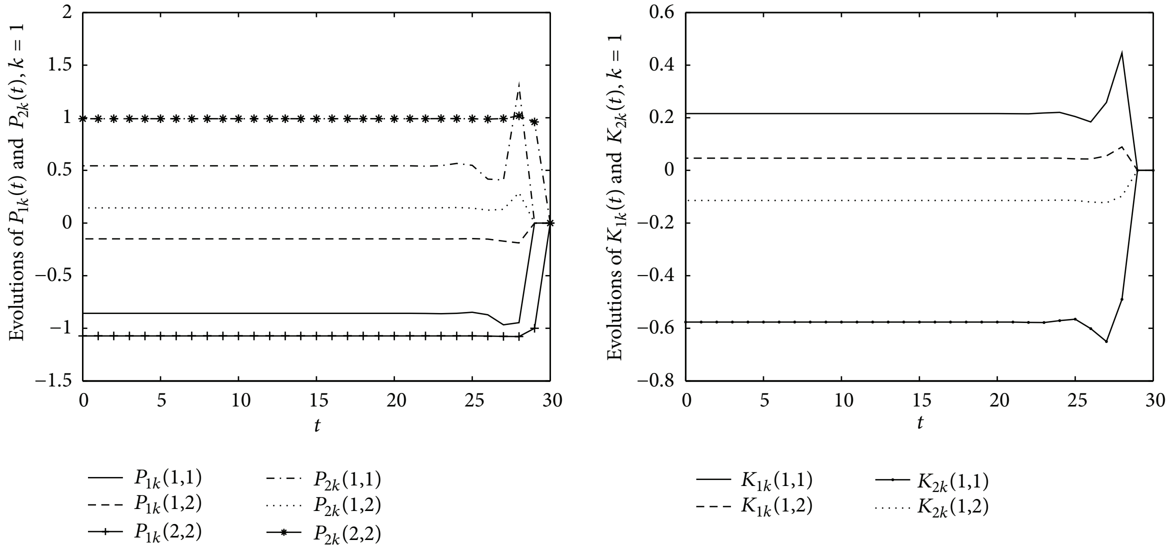


FIGURE 1: Convergence of P_{1k}, P_{2k}, K_{1k}, and K_{2k} in Example 1, $k = 1$.

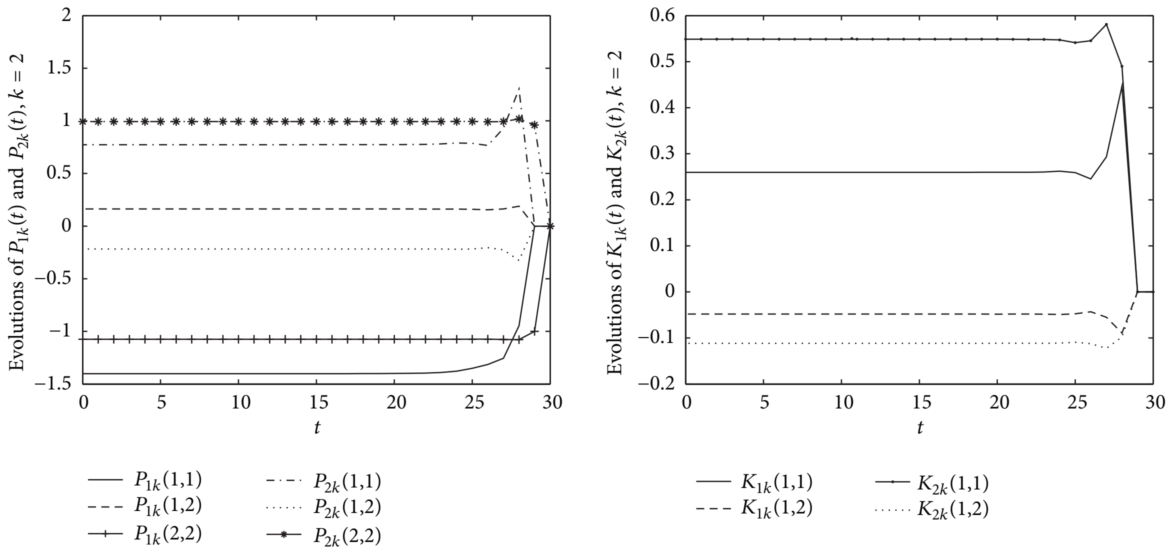


FIGURE 2: Convergence of P_{1k}, P_{2k}, K_{1k}, and K_{2k} in Example 1, $k = 2$.

According to Algorithm 9, coupled differential equations (12) and (13) can be viewed as the following set of equations with known terminal conditions:

$$\begin{gathered} -\dot{P}_1(t) = 5P_1(t) - 9 - \gamma^{-2}P_1^2(t) - 9P_2^2(t) - 18P_1(t)P_2(t), \\ P_1(2) = 0, \\ -\dot{P}_2(t) = 5P_2(t) + 9 - 9P_2^2(t) - 2\gamma^{-2}P_1(t)P_2(t), \\ P_2(2) = 0. \end{gathered} \tag{32}$$

Setting $\gamma = 0.4$ and using the Runge-Kutta integration procedure, the evolutions of $P_1(t)$, $P_2(t)$, $K_1(t)$, $K_2(t)$ are given in Figure 3, which clearly show the convergence of the solutions of (12) and (13).

On the other hand, for one-dimensional time-invariant system (31), the infinite horizon H_2/H_∞ control can be solved by searching for the intersection in the second quadrant of the curves represented by (15) and (16). From Figure 4, it can be found that the solution of (15) and (16) is $P_1 = -7.9777$, $P_2 = 11.7207$, which coincides with the $P_1(0)$ and $P_2(0)$ in Figure 3. Therefore, algebraic matrix-valued equations (15) and (16) can be solved by computing the initial conditions of the corresponding differential matrix-valued equations (12) and (13), which will be called "initial condition method" in the following analysis.

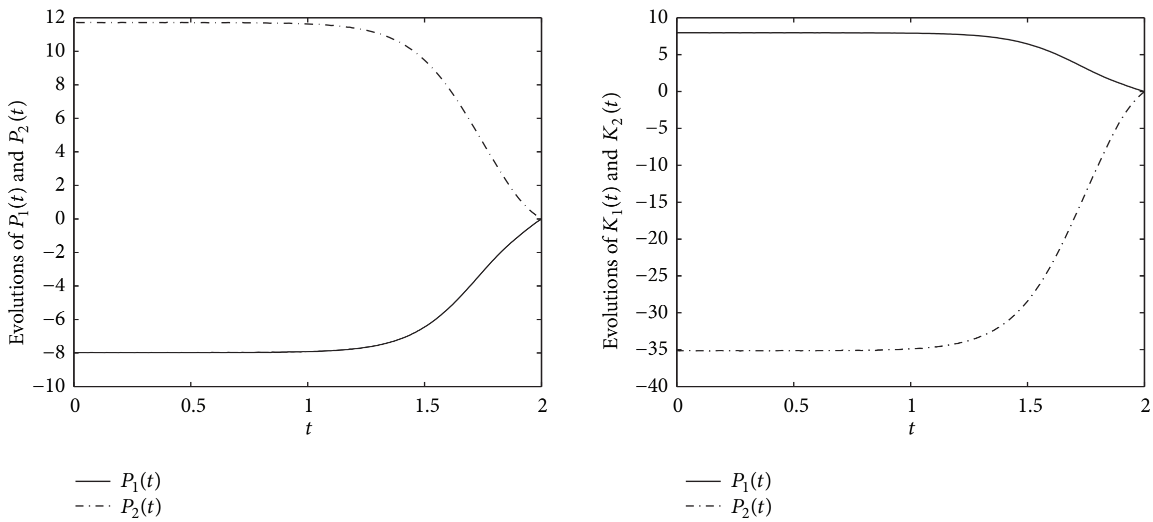


Figure 3: Convergence of $P_1(t)$, $P_2(t)$, $K_1(t)$, and $K_2(t)$ in Example 2.

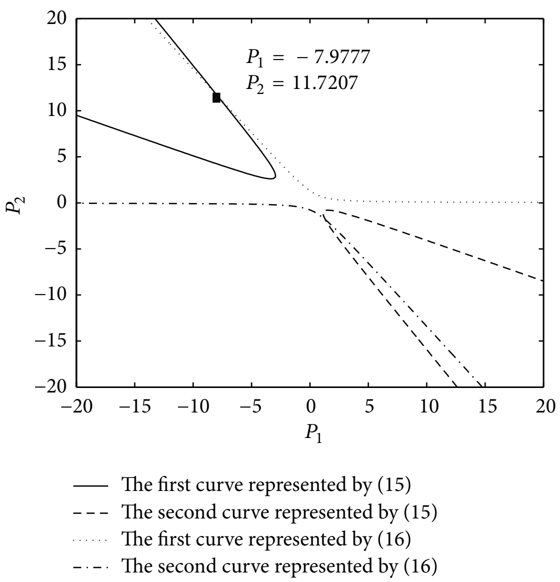


Figure 4: The intersection in the second quadrant of curves represented by (15) and (16) in Example 2.

Example 3. Consider the finite horizon H_2/H_∞ control for two-dimensional time-varying stochastic systems (11) with the following parameters:

$$A(t) = \begin{bmatrix} -1 & 0.5\cos(t) \\ 0 & -1 \end{bmatrix}, \qquad B(t) = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix},$$
$$C(t) = \begin{bmatrix} 1 \\ 0.2 \end{bmatrix}, \qquad D(t) = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}, \qquad F(t) = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}'. \tag{33}$$

In this case, let $P_1(t) = \begin{bmatrix} P_1(11)(t) & P_1(12)(t) \\ P_1(12)(t) & P_1(22)(t) \end{bmatrix}$, $P_2(t) = \begin{bmatrix} P_2(11)(t) & P_2(12)(t) \\ P_2(12)(t) & P_2(22)(t) \end{bmatrix}$, and (12) and (13) correspond to a set of equations with 6 differential equations.

Set $\gamma = 1$, $T = 12$. By applying Algorithm 9, the evolutions of $P_1(t)$, $P_2(t)$, $K_1(t)$, $K_2(t)$ are shown in Figure 5.

Example 4. Consider the infinite horizon H_2/H_∞ control for three-dimensional stochastic systems (14) with the following parameters:

$$A = \begin{bmatrix} -2 & 0 & -0.2 \\ 0.3 & -2 & 0.5 \\ 0.5 & 0.1 & -1.5 \end{bmatrix}, \qquad B = \begin{bmatrix} 1.2 \\ 0.4 \\ 0.8 \end{bmatrix}, \qquad C = \begin{bmatrix} 0.6 \\ 0.5 \\ 1.3 \end{bmatrix},$$
$$D = \begin{bmatrix} 1 & 0.2 & 0.3 \\ 0 & -0.3 & 0.2 \\ 0.5 & 0.6 & 0.8 \end{bmatrix}, \qquad F = \begin{bmatrix} 0.1 \\ 0.3 \\ 0.2 \end{bmatrix}'. \tag{34}$$

According to Algorithm 10, by solving the convex optimization problem (28), we have the following solutions to (15) and (16):

$$P_1 = \begin{bmatrix} -0.0567 & -0.0320 & -0.0525 \\ -0.0320 & -0.0320 & -0.0370 \\ -0.0525 & -0.0370 & -0.0535 \end{bmatrix},$$
$$P_2 = \begin{bmatrix} 0.0196 & 0.0142 & 0.0204 \\ 0.0142 & 0.0123 & 0.0164 \\ 0.0204 & 0.0164 & 0.0225 \end{bmatrix}, \tag{35}$$
$$K_1 = \begin{bmatrix} 0.0296 & 0.0208 & 0.0299 \end{bmatrix},$$
$$K_2 = \begin{bmatrix} -0.0455 & -0.0351 & -0.0490 \end{bmatrix}.$$

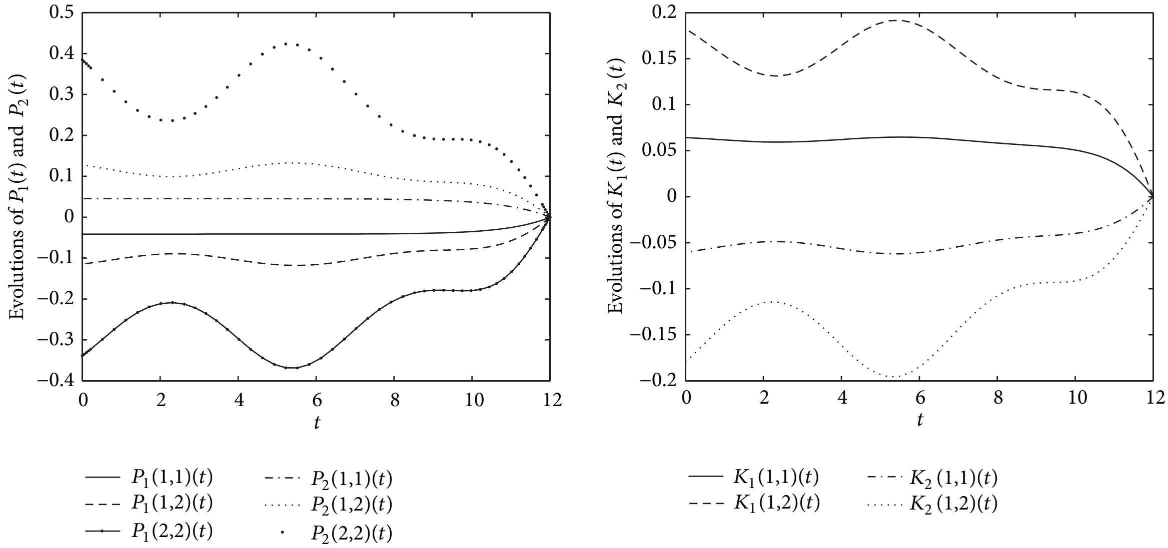


FIGURE 5: Convergence of $P_1(t)$, $P_2(t)$, $K_1(t)$, and $K_2(t)$ in Example 3.

Example 5. Consider the infinite horizon H_2/H_∞ control for two-dimensional stochastic systems (14) with the following parameters:

$$A = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 \\ 0.2 \end{bmatrix},$$
$$D = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}, \qquad F = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}'. \tag{36}$$

In this example, (15) and (16) will be solved by two different methods, that is, the initial condition method and Algorithm 10. Set $\gamma = 2$ and $T = 20$. By using Algorithm 9, the convergence of the solutions to (12) and (13) is shown in Figure 6, and the initial conditions of $P_1(t)$ and $P_2(t)$ are as follows:

$$P_1(0) = \begin{bmatrix} -0.0376 & -0.1267 \\ -0.1267 & -0.5100 \end{bmatrix},$$
$$P_2(0) = \begin{bmatrix} 0.0384 & 0.1307 \\ 0.1307 & 0.5319 \end{bmatrix}. \tag{37}$$

On the other hand, according to Algorithm 10, we have the following solutions to (15) and (16):

$$P_1 = \begin{bmatrix} -0.0524 & -0.2200 \\ -0.2200 & -1.1381 \end{bmatrix},$$
$$P_2 = \begin{bmatrix} 0.0081 & 0.0307 \\ 0.0307 & 0.1159 \end{bmatrix}. \tag{38}$$

Remark 13. Substituting the solutions from initial condition method and those from Algorithm 10 into (15) and (16), it can be found that the former has a higher accuracy than the later. Moreover, the initial condition method is less conservative than Algorithm 10 in some cases. For instance, the infinite horizon H_2/H_∞ control of system (31) can be solved by initial condition method (see Example 2), while there is no optimization solution by using Algorithm 10. However, Algorithm 10 has more advantages in the high-dimensional case than initial condition method. For example, it is difficult to deal with the problem in Example 4 for initial condition method, since it needs to solve a set of equations with 12 differential equations. Therefore, each method has its own advantage and proper scope.

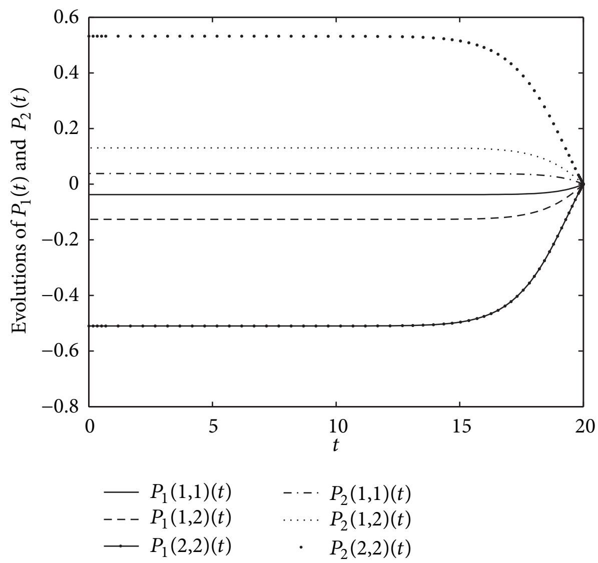


FIGURE 6: Convergence of $P_1(t)$ and $P_2(t)$ in Example 5.

6. Conclusions

In this paper, we have studied the algorithms for H_2/H_∞ control problems of stochastic systems with state-dependent noise. For the finite and infinite horizon stochastic H_2/H_∞ control problems, algorithms in the discrete-time case have been reviewed and studied, and algorithms in the continuous-time case have been developed. The validity of the obtained algorithms has been verified by numerical examples. This subject yields many interesting and challenging topics. For example, how can we design numerical algorithms to solve the H_2/H_∞ control problems of stochastic systems with state, control, and disturbance-dependent noise? This issue deserves further research.

Conflict of Interests

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (nos. 61203053 and 61174078), China Postdoctoral Science Foundation (no. 2013M531635), Special Funds for Postdoctoral Innovative Projects of Shandong Province (no. 201203096), Research Fund for the Taishan Scholar Project of Shandong Province of China, and SDUST Research Fund (no. 2011KYTD105).

References

[1] D. J. N. Limebeer, B. D. O. Anderson, and B. Hendel, "A nash game approach to mixed H_2/H_∞ control," *IEEE Transactions on Automatic Control*, vol. 39, no. 1, pp. 69–82, 1994.

[2] R. Muradore and G. Picci, "Mixed H_2/H_∞ control: the discrete-time case," *Systems & Control Letters*, vol. 54, no. 1, pp. 1–13, 2005.

[3] O. L. V. Costa and R. P. Marques, "Mixed H_2/H_∞ control of discrete-time Markovian jump linear systems," *IEEE Transactions on Automatic Control*, vol. 43, pp. 95–100, 1998.

[4] C. Du, L. Xie, J. N. Teoh, and G. Guo, "An improved mixed H_2/H_∞ control design for hard disk drives," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 5, pp. 832–839, 2005.

[5] B.-S. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[6] W. Zhang, Y. Huang, and H. Zhang, "Stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise," *Automatica*, vol. 43, no. 3, pp. 513–521, 2007.

[7] W. Zhang, Y. Huang, and L. Xie, "Infinite horizon stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise," *Automatica*, vol. 44, no. 9, pp. 2306–2316, 2008.

[8] Y. Huang, W. Zhang, and G. Feng, "Infinite horizon H_2/H_∞ control for stochastic systems with Markovian jumps," *Automatica*, vol. 44, no. 3, pp. 857–863, 2008.

[9] T. Hou, W. Zhang, and H. Ma, "Finite horizon H_2/H_∞ control for discrete-time stochastic systems with Markovian jumps and multiplicative noise," *IEEE Transactions on Automatic Control*, vol. 55, no. 5, pp. 1185–1191, 2010.

[10] H. Ma, W. Zhang, and T. Hou, "Infinite horizon H_2/H_∞ control for discrete-time time-varying Markov jump systems with multiplicative noise," *Automatica*, vol. 48, no. 7, pp. 1447–1454, 2012.

[11] L. Sheng, W. Zhang, and M. Gao, "Relationship between Nash equilibrium strategies and H_2/H_∞ control of stochastic Markov jump systems with multiplicative noise," *IEEE Transactions on Automatic Control*, 2014.

[12] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, UK, 2006.

[13] V. Dragan, T. Morozan, and A. M. Stoica, *Mathematical Methods in Robust Control of Discrete-Time Linear Stochastic Systems*, Springer, New York, NY, USA, 2010.

[14] M. Sznaier and H. Rotstein, "An exact solution to general 4-blocks discrete-time mixed H_2/H_∞ problems via convex optimization," in *Proceedings of the American Control Conference*, pp. 2251–2256, July 1994.

[15] Y. Feng and B. D. O. Anderson, "An iterative algorithm to solve state-perturbed stochastic algebraic Riccati equations in LQ zero-sum games," *Systems & Control Letters*, vol. 59, no. 1, pp. 50–56, 2010.

[16] M. A. Rami and X. Y. Zhou, "Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1131–1143, 2000.

[17] W. Zhang and G. Feng, "Nonlinear stochastic H_2/H_∞ control with (x,u,v)-dependent noise: infinite horizon case," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1323–1328, 2008.

[18] W. Zhang, B. Chen, H. Tang, L. Sheng, and M. Gao, "Some remarks on general nonlinear stochastic H_∞ control with state, control, and disturbance-dependent noise," *IEEE Transactions on Automatic Control*, vol. 59, no. 1, pp. 237–242, 2014.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 903821, 10 pages
http://dx.doi.org/10.1155/2014/903821

Research Article

Exponential Stability of Stochastic Systems with Delay and Poisson Jumps

Wenli Zhu, Jiexiang Huang, and Zhao Zhao

School of Economic Mathematics, Southwestern University of Finance and Economics, Chengdu 611130, China

Correspondence should be addressed to Wenli Zhu; zhuwl@swufe.edu.cn

Received 6 April 2014; Revised 3 June 2014; Accepted 21 June 2014; Published 5 August 2014

Academic Editor: Guangchen Wang

This paper focuses on the model of a class of nonlinear stochastic delay systems with Poisson jumps based on Lyapunov stability theory, stochastic analysis, and inequality technique. The existence and uniqueness of the adapted solution to such systems are proved by applying the fixed point theorem. By constructing a Lyapunov function and using Doob's martingale inequality and Borel-Cantelli lemma, sufficient conditions are given to establish the exponential stability in the mean square of such systems, and we prove that the exponentially stable in the mean square of such systems implies the almost surely exponentially stable. The obtained results show that if stochastic systems is exponentially stable and the time delay is sufficiently small, then the corresponding stochastic delay systems with Poisson jumps will remain exponentially stable, and time delay upper limit is solved by using the obtained results when the system is exponentially stable, and they are more easily verified and applied in practice.

1. Introduction

In nature, physics, society, engineering, and so on we always meet two kinds of functions with respect to time: one is deterministic and another is random. Stochastic differential equations (SDEs for short) were first initiated and developed by K. Itô [1]. Today they have become a very powerful tool applied to mathematics, physics, biology, finance, and so forth.

Currently, the study of analysis and synthesis of stochastic time delay systems, described by stochastic delayed differential equations (SDDE for short), is a popular topic in the field of control theory [2–8]. Delays in the dynamics can represent memory or inertia in the financial system [9]. Because the existence of time delay is the main reason about bringing instability and deteriorating the control performance, the study on time delay systems stability and control has important theoretical and practical values. Furthermore, it often happens in real lives that a stochastic system jumps from a "normal state" or "good state" to a "bad state," and the strength of system is random. For this class of systems, it is natural and necessary to include a jump term in them. The effect of Poisson jumps should be taken into account when studying the stability of SDEs [10–16]. Therefore, except stochastic and delay effects, Poisson jumps' effects is likely to exist widely in variety of evolution processes in which states are changed abruptly at some moments of time, including such fields as finance, economy, medicine, electronics, and so forth. Then, it is natural to consider the effect of Poisson jumps when studying the stability of SDDEs.

So far, these topics have received a lot of attention and there are so many references about them. For instance, [2–8] established some stability criteria of the stochastic systems with delay by using Lyapunov function method or Razumikhin technique or inequality technique and so on. By using the fixed point theory and Borel-Cantelli lemma, Guo and Zhu [13] studied that the solution to a class of stochastic Volterra-Levin equations with Poisson jumps is not only existent and unique but also pth moment exponentially stable. By constructing a novel Lyapunov-Krasovskii functional and using some new approaches and techniques, Zhu and Cao [14] focused on the exponential stability for a class of Markovian jump impulsive stochastic Cohen-Grossberg neural networks with mixed time delays and got several novel sufficient conditions. By applying a Lyapunov-Krasovskii functional, the stochastic analysis theory, and LMI approach, Zhu and Cao [15] investigated a class of stochastic neural networks with both Markovian jump parameters and mixed time delays and derived some novel sufficient conditions. In [16], Zhu proposed several good sufficient conditions under which he proved the asymptotic stability in the pth moment and almost sure stability of the SDEs with Lévy noise. Based on fixed point theory, Chen et al. [17] proved that the mild solution to a class of impulsive SPDEs with delays and Poisson jumps is not only existent and unique but also pth moment exponentially stable.

Delay and Poisson jumps always coexist in real dynamic systems. Thus, it is reasonable to consider them together, leading us to investigate SDDEs with Poisson jumps. However, the delayed response gives us more difficulties to deal with the delayed stochastic control problems, not only for the infinite-dimensional problem, but also for the absence of Itô's formula to deal with the delayed part of the trajectory. So the stochastic controlled delay systems are more complicated. Because Lévy processes are not continuous, but their sample paths are right-continuous and have a number of random jump discontinuities occurring at random times, on each finite time interval. Since Lévy noise has more advantages than the standard Gausian noise despite its increased mathematical complexity, it is very interesting and challenging to study SDDEs with Lévy noise. There is little literature focusing on a certain class of this system, [14–17], that discussed the exponential stability of the trivial solution for this system, but these stable conditions only ensure the exponential stability of the respective solution and do not give a bound for the time delay δ, and Chen et al. pointed out that it is impossible to analyze the stability of mild solutions to SDDEs by Lyapunov method.

The main objective of this paper is to fill this gap. We investigate not only the exponential stability in the mean square but also the almost surely exponential stability for a class of SDDE with Poisson jumps based on Lyapunov stability theory, Itô formula, stochastic analysis, and inequality technique. We first consider the existence and uniqueness of the adapted solution by employing fixed point theorem. Next, some sufficient conditions of exponential stability and corollaries for SDDE with Poisson jumps are obtained by using Lyapunov function. By utilizing Doob's martingale inequality and Borel-Cantelli lemma, it is shown that the exponentially stable in the mean square of SDDE with Poisson jumps implies the almost surely exponentially stable. Our results generalize and improve some recent results (for instance [5–8, 14–17]). In particular, our results show that if SDE is exponentially stable and the time delay is sufficiently small, then the corresponding SDDE with Poisson jumps will remain exponentially stable. Moreover, when the system is exponentially stable, the time delay upper limit is solved by using our results which are more easily verified and applied in practice. Our approach in the current paper is different from the above [14–17]. Finally, we present a simple example to illustrate the effectiveness of our stable results.

The rest of this paper is organized as follows. In Section 2, we give the preliminary results about SDDE with Poisson jumps. Main results and proofs for SDDE with Poisson jumps are provided in Section 3. Section 4 presents a simple example to illustrate our stable results. Section 5 lists some concluding remarks.

2. Preliminaries

Throughout this paper and unless specified, we let $B(t) = B(t,\omega)$ be an m-dimensional motion and $\widetilde{N}(dt,dz) = N(dt,dz) - \nu(dz)dt$ which is the l-dimensional compensated jump measure of $\eta(\cdot)$an independent compensated Poisson random measure on a filtered probability space $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\le t\le T},P)$. $N(dt,dz)$ is the l-dimensional jump measure and $\nu(dz)$ is the Lévy measure of l-dimensional Lévy process $\eta(\cdot)$ and $T>0$.

We denote the notation $|\cdot|$ for the Euclidean norm. If A is a vector or matrix, its transpose is denoted by A^T. If A is a square matrix, the trace of A is denoted by $\operatorname{tr}(A)$ and then the operator norm of A is denoted by $\|A\|$; that is, $\|A\| = \sqrt{\operatorname{tr}(A^T A)}$. We also use the notation: $L^2_{\mathcal{F}}([s,r];\mathbf{R}^n) = \{\phi(t) : \{\phi(t), s\le t\le r\}$ which is $\mathbf{R}^n$-valued adapted stochastic processes s.t. $\int_s^r E|\phi(t)|^2 dt < \infty\}$.

Suppose $X(t)\in\mathbf{R}^n$ is an Itô-Lévy process of the form

$$\begin{aligned} dX(t) = {} & b(t, X(t), Y(t), \omega)\,dt \\ & + \sigma(t, X(t), Y(t), \omega)\,dB(t) \\ & + \int_{\mathbf{R}_0^n} \gamma(t, X(t^-), Y(t^-), z, \omega)\,\widetilde{N}(dt,dz); \\ & \qquad t\in[0,T], \end{aligned} \tag{1}$$

$$X(t) = \xi(t); \quad t\in[-\delta,0], \tag{2}$$

where $Y(t) = X(t-\delta)$, $\mathbf{R}_0^n := \mathbf{R}^n/\{0\}$, and $\delta>0$. Here $b : [0,T]\times\mathbf{R}^n\times\mathbf{R}^n\times\Omega \to \mathbf{R}^n$, $\sigma : [0,T]\times\mathbf{R}^n\times\mathbf{R}^n\times\Omega \to \mathbf{R}^{n\times m}$, and $\gamma : [0,T]\times\mathbf{R}^n\times\mathbf{R}^n\times\mathbf{R}_0^n\times\Omega \to \mathbf{R}^{n\times l}$, are given functions such that for all t, $b(t,x,y,\cdot)$, $\sigma(t,x,y,\cdot)$, and $\gamma(t,x,y,z,\cdot)$ are $\mathcal{F}_t$-measurable for all $x\in\mathbf{R}^n$, $y\in\mathbf{R}^n$ and $z\in\mathbf{R}_0^n$. In the following, we suppress the ω, for notational simplicity. The initial date $X(t) = \xi(t)$ is satisfied with $\xi := \{\xi(s) : -\delta\le s\le 0\}\in L^p_{\mathcal{F}_0}([-\delta,0];\mathbf{R}^n)$.

Now let us present an existence and uniqueness result for (1)-(2). First we let the maps $b(t,x,y,\cdot)$, $\sigma(t,x,y,\cdot)$, and $\gamma(t,x,y,z,\cdot)$ satisfy the following conditions.

(H2.1) At most linear growth: there exists a constant $C_1>0$ such that

$$\begin{aligned} & |b(t,x,y)|^2 + \|\sigma(t,x,y)\|^2 \\ & \quad + \int_{\mathbf{R}_0} \sum_{k=1}^{l} \left|\gamma^{(k)}(t,x,y,z_k)\right|^2 \nu_k(dz_k) \\ & \le C_1\left(1+|x|^2+|y|^2\right) \end{aligned} \tag{3}$$

for all $x\in\mathbf{R}^n$, $y\in\mathbf{R}^n$, where $\gamma^{(k)}\in\mathbf{R}^n$ is column number k of the $n\times l$ matrix $\gamma = [\gamma_{ik}]$ and $\gamma_i^{(k)} = \gamma_{ik}$ is the coordinate number i of $\gamma^{(k)}$, and $\gamma^{(k)}(t,x,y,z) = \gamma^{(k)}(t,x,y,z_k)$; $z = (z_1,\ldots,z_l)\in\mathbf{R}^l$.

(H2.2) Lipschitz continuity: there exists a constant $C_2 > 0$ such that

$$\left|b(t,x,y)-b(t,x',y')\right|^2+\left\|\sigma(t,x,y)-\sigma(t,x',y')\right\|^2 + \int_{\mathbf{R}_0}\sum_{k=1}^{l}\left|\gamma^{(k)}(t,x,y,z_k)-\gamma^{(k)}(t,x',y',z_k)\right|^2\nu_k(dz_k) \leq C_2\left(\left|x-x'\right|^2+\left|y-y'\right|^2\right) \tag{4}$$

for all $x\in\mathbf{R}^n$, $y\in\mathbf{R}^n$.

Lemma 1 (see [18]). *For any real matrices $\zeta_1,\zeta_2\in R^n$ and a constant $\theta>0$, the following matrix inequality holds:*

$$2\zeta_1^T\zeta_2 \leq \theta\zeta_1^T\zeta_1+\frac{1}{\theta}\zeta_2^T\zeta_2. \tag{5}$$

Theorem 2. *Let (H2.1) and (H2.2) hold. Then for any $\xi(t)\in L^2_{\mathcal{F}_0}([-\delta,0];\mathbf{R}^n)$, (1)-(2) have a unique adapted solution $X(t;\xi)$ such that*

$$E\left[\left|X(t;\xi)\right|^2\right]<\infty \tag{6}$$

for all t. When $b(t,0,0,0)=\sigma(t,0,0,0)\equiv 0$, it is easy to see that (1)-(2) have a trivial solution $X(t;0)=0$.

We present the proof of Theorem 2 which is left in Appendix.

To develop our theories and results, we need to introduce the following concepts. For stochastic system, exponential stability in mean square and almost surely exponential stability are generally used [7].

Definition 3. The trivial solution of (1)-(2) is said to be pth moment exponentially stable. If there exists a positive constant ε such that

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left(E\left|X(t;\xi)\right|^p\right)\leq-\varepsilon \tag{7}$$

for any $\xi\in L^p_{\mathcal{F}_0}([-\delta,0];\mathbf{R}^n)$.

Particularly, $p=2$; it is called mean square exponentially stable.

Definition 4. The trivial solution of (1)-(2) is said to be almost surely exponentially stable. If there exists a positive constant η such that

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left|X(t;\xi)\right|\leq-\eta \quad \text{a.s.} \tag{8}$$

for any $\xi\in L^p_{\mathcal{F}_0}([-\delta,0];\mathbf{R}^n)$.

3. Main Results

For simplicity, in what follows we write $X(t;\xi)=X(t)$.

We make the following assumptions for the coefficients of (1)-(2).

In the study of mean square exponential stability, it is often to use a quadratic function as the Lyapunov function; that is, $V(t,x)=x^TGx$, where G is a symmetric positive definite $n\times n$ matrix.

Theorem 5. *Let (H2.1)-(H2.2) hold; then the trivial solution of (1)-(2) is exponentially stable in the mean square. Assume that there exists a symmetric positive definite $n\times n$ matrices G and a constant $\lambda>0$ such that*

$$2X^TGb(t,X,X)\leq-\lambda|X|^2,\quad \forall(t,x,y)\in[0,T]\times R^n\times R^n, \tag{9}$$

$$\lambda>4\|G\|C_1+2\|G\|\sqrt{6C_1C_2\delta(\delta+2)}. \tag{10}$$

In order to prove Theorem 5, we need two lemmas, proofs of which are left in Appendix.

Lemma 6. *Fix the initial data $\xi(t)$ arbitrarily. Then,*

$$\int_\delta^t e^{\varepsilon s}\int_{s-\delta}^s E|X(\tau)|^2d\tau\,ds\leq\delta e^{\varepsilon\delta}\int_0^t e^{\varepsilon\tau}E|X(\tau)|^2d\tau, \tag{11}$$

$$\int_0^t e^{\varepsilon s}E|X(s-\delta)|^2ds\leq c_{11}e^{\varepsilon\delta}+e^{\varepsilon\delta}\int_0^t e^{\varepsilon s}E|X(s)|^2ds \tag{12}$$

for any $t\geq\delta$, where c_{11} is a constant larger than $\int_{-\delta}^0 E|\xi(s)|^2ds$.

Lemma 7. *Let (H2.1) and (H2.2) hold. Fix the initial data $\xi(t)$ arbitrarily; then,*

$$\int_\delta^t e^{\varepsilon s}\int_{s-\delta}^s E|X(\tau-\delta)|^2d\tau\,ds \leq \delta c_{11}e^{2\varepsilon\delta}+\delta e^{2\varepsilon\delta}\int_0^t e^{\varepsilon s}E|X(s)|^2ds, \tag{13}$$

$$\begin{aligned}\int_0^t e^{\varepsilon s}E|X(s)-X(s-\delta)|^2ds &\leq c_{22} + 3C_1(\delta+2)\delta e^{\varepsilon\delta}\int_0^t e^{\varepsilon s}E|X(s)|^2ds \\ &\quad + 3C_1(\delta+2)\left(c_{11}\delta e^{2\varepsilon\delta}+\delta e^{2\varepsilon\delta}\int_0^t e^{\varepsilon s}E|X(s)|^2ds\right)\end{aligned} \tag{14}$$

for any $t\geq\delta$, where

$$c_{22}\geq\int_0^\delta e^{\varepsilon s}E|X(s)-X(s-\delta)|^2ds + \frac{3\delta C_1(\delta+2)}{\varepsilon}\left(e^{\varepsilon T}-e^{\varepsilon\delta}\right). \tag{15}$$

Based on Lemmas 6 and 7 above, we now carry out a proof for Theorem 5.

Proof of Theorem 5. Fix the initial data $\xi(t)$ arbitrarily. Applying Itô's formula to $X^T GX$, we have

$$\begin{aligned} d\left(X^T GX\right) = {} & 2X^T Gb(t,X,Y)\,dt \\ & + 2X^T G\sigma(t,X,Y)\,dB(t) \\ & + \operatorname{tr}\left[\sigma^T(t,X,Y)\,G\sigma(t,X,Y)\right]dt \\ & + \sum_{k=1}^{l}\int_{|z_k|<R}\gamma^{(k)}(t,X,Y,z_k) \\ & \qquad \times G\gamma^{(k)}(t,X,Y,z_k)\,\nu_k(dz_k)\,dt \\ & + \sum_{k=1}^{l}\int_{\mathbf{R}_0}\left\{\gamma^{(k)}(t,X,Y,z_k)\,G\gamma^{(k)}(t,X,Y,z_k)\right. \\ & \qquad + X^T(t^-)\,G\gamma^{(k)}(t,X,Y,z_k) \\ & \qquad \left. + \gamma^{(k)}(t,X,Y,z_k)\,GX(t^-)\right\} \\ & \times \widetilde{N}_k(dt,dz_k). \end{aligned} \tag{16}$$

Applying Itô's formula to $e^{\varepsilon t}X^T GX$ and taking the expectation, we have

$$\begin{aligned} E\left(e^{\varepsilon t}X^T GX\right) \le {} & E\left(\xi^T(0)\,G\xi(0)\right) \\ & + \varepsilon E\int_0^t e^{\varepsilon s}X^T GX\,ds \\ & + E\int_0^t e^{\varepsilon s}2X^T Gb(s,X,Y)\,ds \\ & + E\int_0^t e^{\varepsilon s}\operatorname{tr}\left[\sigma^T(s,X,Y)\,G\sigma(s,X,Y)\right]ds \\ & + E\int_0^t e^{\varepsilon s}\left[\sum_{k=1}^{l}\int_{|z_k|<R}\gamma^{(k)}(s,X,Y,z_k)\right. \\ & \qquad \times G\gamma^{(k)}(s,X,Y,z_k) \\ & \qquad \left. \times \nu_k(dz_k)\right]ds \\ := {} & E\left(\xi^T(0)\,G\xi(0)\right) \\ & + \varepsilon E\int_0^t e^{\varepsilon s}X^T GX\,ds + I_1 + I_2 + I_3, \end{aligned} \tag{17}$$

where

$$\begin{aligned} I_1 &= E\int_0^t e^{\varepsilon s}2X^T Gb(s,X,Y)\,ds, \\ I_2 &= E\int_0^t e^{\varepsilon s}\operatorname{tr}\left[\sigma^T(s,X,Y)\,G\sigma(s,X,Y)\right]ds, \\ I_3 &= E\int_0^t e^{\varepsilon s}\left[\sum_{k=1}^{l}\int_{|z_k|<R}\gamma^{(k)}(t,X,Y,z_k)\right. \\ & \qquad \left. \times G\gamma^{(k)}(t,X,Y,z_k)\,\nu_k(dz_k)\right]ds. \end{aligned} \tag{18}$$

Combining Lemma 1 and (9) as well as (H2.2), we can estimate I_1 as follows:

$$\begin{aligned} I_1 &= E\int_0^t e^{\varepsilon s}\left\{2X^T Gb(s,X,X)\right. \\ & \qquad \left. +2X^T G\left[b(s,X,Y)-b(s,X,X)\right]\right\}ds \\ &\le E\int_0^t e^{\varepsilon s}\left\{-\lambda|X|^2+\theta|X|^2\right. \\ & \qquad \left. +\left(\frac{\|G\|^2}{\theta}\right)\cdot C_2|X-Y|^2\right\}ds, \end{aligned} \tag{19}$$

where $\theta > 0$ is a constant.

By (H2.1), $I_2 + I_3$ of (17) yields

$$\begin{aligned} I_2 + I_3 \le {} & 2C_1\|G\|\left(e^{\varepsilon T}-1\right) \\ & + 2C_1\|G\|\int_0^t e^{\varepsilon s}\left(E|X|^2+E|Y|^2\right)ds. \end{aligned} \tag{20}$$

Substituting the above two into (17) and using Lemmas 6 and 7, we get an estimate of $E(e^{\varepsilon t}X^T GX)$ as follows:

$$\begin{aligned} & E\left(e^{\varepsilon t}X^T GX\right) \\ & \le E\left(\xi^T(0)\,G\xi(0)\right) + 2C_1\|G\|\,e^{\varepsilon T} \\ & \quad - \left[\lambda-\theta-\|G\|\left(2C_1+\varepsilon\right)\right]\int_0^t e^{\varepsilon s}E|X|^2 ds \\ & \quad + 2C_1\|G\|\left(c_{11}e^{\varepsilon\delta}+e^{\varepsilon\delta}\int_0^t e^{\varepsilon s}E|X|^2 ds\right) \\ & \quad + \left(\frac{\|G\|^2}{\theta}\right)\cdot C_2\left[c_{22}+3C_1(\delta+2)\,\delta e^{\varepsilon\delta}\right. \\ & \qquad \times\int_0^t e^{\varepsilon s}E|X|^2 ds + 3C_1(\delta+2) \\ & \qquad \left.\times\left(c_{11}\delta e^{2\varepsilon\delta}+\delta e^{2\varepsilon\delta}\int_0^t e^{\varepsilon s}E|X|^2 ds\right)\right] \\ & = c_{33}. \end{aligned} \tag{21}$$

for $t \ge \delta$, where

$$\begin{aligned} c_{33} = {} & E\left(\xi^T(0)\,G\xi(0)\right) + 2C_1\|G\|\,e^{\varepsilon T} + 2C_1c_{11}\|G\|\,e^{\varepsilon\delta} \\ & + \frac{\left(\|G\|\sqrt{C_2}\right)^2}{\theta}\cdot\left[c_{22}+3C_1c_{11}\delta(\delta+2)\,e^{2\varepsilon\delta}\right). \end{aligned} \tag{22}$$

For small enough $\varepsilon > 0$, we derive

$$\begin{aligned} &\|G\|(2C_1+\varepsilon)+\theta+2C_1\|G\|e^{\varepsilon\delta} \\ &\quad +3\frac{\left(\|G\|\sqrt{C_1C_2(\delta+2)\delta}\right)^2}{\theta}\cdot\left(e^{\varepsilon\delta}+e^{2\varepsilon\delta}\right) \\ &\geq 4\|G\|C_1+\theta \\ &\quad +\frac{\left(\|G\|\sqrt{6C_1C_2\delta(\delta+2)}\right)^2}{\theta}. \end{aligned} \tag{23}$$

If (10) holds, then we can choose $\varepsilon > 0$ small enough such that

$$\begin{aligned} \lambda &= \|G\|(2C_1+\varepsilon)+\theta+2C_1\|G\|e^{\varepsilon\delta} \\ &\quad +3\frac{\left(\|G\|\sqrt{C_1C_2(\delta+2)\delta}\right)^2}{\theta}\cdot\left(e^{\varepsilon\delta}+e^{2\varepsilon\delta}\right). \end{aligned} \tag{24}$$

Since G is positive definite,

$$X^TGX \geq \lambda_{\min}(G)|X|^2, \tag{25}$$

where $\lambda_{\min}(G) > 0$ is the smallest eigenvalue of G.

Then,

$$E\left(e^{\varepsilon t}X^TGX\right) \geq E\left(e^{\varepsilon t}\lambda_{\min}(G)|X|^2\right). \tag{26}$$

It then follows from (21) that

$$\begin{aligned} \frac{1}{t}\ln\left(E|X(t)|^2\right) &\leq \frac{1}{t}\ln\left(\left[\frac{c_{44}}{\lambda_{\min}(G)}\right]e^{-\varepsilon t}\right) \\ &= -\varepsilon+\frac{1}{t}\ln\left[\frac{c_{44}}{\lambda_{\min}(G)}\right]. \end{aligned} \tag{27}$$

This easily yields

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left(E|X(t)|^2\right) \leq -\varepsilon. \tag{28}$$

Then (1)-(2) is exponentially stable in the mean square. □

Theorem 8. *Let $\varepsilon > 0$, under the same assumption as Theorem 5. If inequality* (28) *holds, then,*

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left(|X(t)|\right) \leq -\frac{\varepsilon}{2} \quad a.s. \tag{29}$$

Proof. Let $\varepsilon > 0$, under the same assumption as Theorem 5. It follows from (27) that

$$\frac{1}{t}\ln\left(E|X(t)|^2\right) \leq -\varepsilon+\frac{1}{t}\cdot\ln M \tag{30}$$

for all $t \geq \delta$. Here $M = c_{44}/\lambda_{\min}(G)$. Then, for $t \in [k\delta,(k+1)\delta]$, $k = 2,3,\ldots$, we have

$$E\left(\sup_{k\delta\leq t\leq(k+1)\delta}E|X(t)|^2\right) \leq Me^{-\varepsilon k\delta}. \tag{31}$$

Let $\varepsilon_0 \in (0,\varepsilon)$ be arbitrary. By Doob's martingale inequality. It follows from (31) that

$$P\left(\omega : \sup_{k\delta\leq t\leq(k+1)\delta}|X(t)| > e^{-(\varepsilon-\varepsilon_0)k\delta/2}\right) \leq c_{33}e^{-\varepsilon_0 k\delta}. \tag{32}$$

Thus, it follows from the Borel-Cantelli lemma that, for almost all $\omega \in \Omega$, there exists $k_0(\omega)$, and $k \geq k_0(\omega)$,

$$P\left(\omega : \sup_{k\delta\leq t\leq(k+1)\delta}|X(t)| \leq e^{-(\varepsilon-\varepsilon_0)k\delta/2}\right) = 1. \tag{33}$$

Since ε_0 is arbitrary, we must have

$$\limsup_{t\to\infty}\frac{1}{t}\ln|X(t)| \leq -\frac{\varepsilon}{2} \quad \text{a.s.} \tag{34}$$

□

Remark 9. The exponentially stable in the mean square of (1)-(2) implies the almost surely exponentially stable. In general, Theorem 8 is still true for pth moment exponential stable.

Let us single out three important special cases.

Case 1. If $\sigma = 0$ and $N = 0$ (no jumps), then (1)-(2) reduces to ODE with delay

$$\begin{aligned} dX(t) &= b(t,X,Y)\,dt, \quad t \in [0,T], \\ X(t) &= \xi(t), \quad t \in [-\delta,0]. \end{aligned} \tag{35}$$

Applying Theorem 5 to (35), we obtain the following useful result.

Corollary 10. *Let (H2.1)-(H2.2) hold; then the trivial solution of* (35) *is exponentially stable in the mean square. Assume that there exists a symmetric positive definite $n\times n$ matrices G and a constant $\lambda > 0$ such that*

$$2X^TGb(t,X,X) \leq -\lambda|X|^2, \quad \forall(t,x,x) \in [0,T]\times R^n\times R^n, \tag{36}$$

$$\lambda > 2\|G\|\delta\sqrt{2C_1C_2}. \tag{37}$$

Case 2. If $N = 0$ (no jumps), then (1)-(2) reduces to SDE with delay

$$\begin{aligned} dX(t) &= b(t,X,Y)\,dt+\sigma(t,X,Y)\,dB(t), \quad t \in [0,T], \\ X(t) &= \xi(t), \quad t \in [-\delta,0]. \end{aligned} \tag{38}$$

Applying Theorem 5 to (38), we obtain the following useful result.

Corollary 11. *Let (H2.1)-(H2.2) hold; then the trivial solution of* (38) *is exponentially stable in the mean square. Assume that there exists symmetric positive definite $n\times n$ matrices G and a constant $\lambda > 0$ such that*

$$2X^TGb(t,X,X) \leq -\lambda|X|^2, \quad \forall(t,x,x) \in [0,T]\times R^n\times R^n, \tag{39}$$

$$\lambda > 2\|G\|C_1+4\|G\|\sqrt{C_1C_2\delta(\delta+1)}. \tag{40}$$

Remark 12. The bound for the time delay δ when (1)-(2) is exponentially stable which follows from (10), the bound for the corresponding deterministic case follows from (37), and the bound for the corresponding stochastic case follows from (40).

Case 3. If the time delay $\delta = 0$, then (1)-(2) reduces to the nondelay SDE with jumps

$$dX(t) = b(t, X, X)\,dt + \sigma(t, X, X)\,dB(t) + \int_{\mathbf{R}_0^n} \gamma(t, X, X, z, \omega)\,\widetilde{N}(dt, dz); \quad t \in [0, T]. \tag{41}$$

One of the powerful techniques employed in the study of the stability problem is the method of the Lyapunov functions or functional [19]. However, it is generally much more difficult to construct the Lyapunov functionals in the case of delay than the Lyapunov functions in the case of nondelay. Therefore another useful technique has been developed, that is, to compare the stochastic differential delay equations with the corresponding nondelay equations. To explain, let us look at a SDE (1) with delay and jumps?

$$dX(t) = b(t, X, Y)\,dt + \sigma(t, X, Y)\,dB(t) + \int_{\mathbf{R}_0^n} \gamma(t, X(t^-), Y(t^-), z)\,\widetilde{N}(dt, dz); \quad t \in [0, T]. \tag{42}$$

Equation (1) can be rewritten as

$$\begin{aligned} dX(t) &= b(t, X, X)\,dt + \sigma(t, X, X)\,dB(t) \\ &\quad + \int_{\mathbf{R}_0^n} \gamma(t, X(t^-), X(t^-), z)\,\widetilde{N}(dt, dz) \\ &\quad - [b(t, X, X) - b(t, X, Y)]\,dt \\ &\quad - [\sigma(t, X, X) - \sigma(t, X, Y)]\,dB(t) \\ &\quad - \int_{\mathbf{R}_0^n} [\gamma(t, X(t^-), X(t^-), z) \\ &\quad - \gamma(t, X(t^-), Y(t^-), z)]\,\widetilde{N}(dt, dz), \end{aligned} \tag{43}$$

and regard it as the perturbed system of the corresponding nondelay SDE (41). Obviously, if the time delay δ is sufficiently small then the perturbation term,

$$\begin{aligned} &[b(t, X, X) - b(t, X, Y)]\,dt \\ &\quad + [\sigma(t, X, X) - \sigma(t, X, Y)]\,dB(t) \\ &\quad + \int_{\mathbf{R}_0^n} [\gamma(t, X(t^-), X(t^-), z) \\ &\quad - \gamma(t, X(t^-), Y(t^-), z)]\,\widetilde{N}(dt, dz), \end{aligned} \tag{44}$$

could be so small that the perturbed equation (1) would behave in a similar way as (41) asymptotically. Applying Theorem 5 and Remark 12 in [20], we derive (1) which will remain exponentially stable.

Corollary 13. *If the nondelay equation* (41) *is exponentially stable and the time delay δ is sufficiently small, then the corresponding delay equation* (1) *will remain exponentially stable.*

4. Example

Let us now present a simple example to illustrate our results, which can help us find the time delay upper limit.

Example 1. For simplicity of presentation, let us consider a simple one-dimensional (i.e., $n = m = l = 1$, thus, the indices i and j in Theorem 5 will be omitted below) delay equation with jumps

$$dX(t) = b(t, X, Y)\,dt + \sigma(t, X, Y)\,dB(t) + \int_{\mathbf{R}_0} \gamma(t, X(t^-), Y(t^-), z)\,\widetilde{N}(dt, dz); \quad t \in [0, T], \tag{45}$$

where $B(t)$ is one-dimensional Brownian motion. Constants $T > 0$ and $\delta > 0$ is a given finite time delay. For convenience, let us choose $G = 1$ in this one-dimensional case. Hence (9) is satisfied with $\lambda \geq 2$.

Moreover, we let $b(t, x, y) = -y$, $\sigma(t, x, y) = 0.5x - 0.5y$ and $\gamma(t, x, y, z) = z$, $N(dt, dz) = dN(t)$, $\nu(dz) = \kappa f(z)dz$, where $dN(t)$ is a Poisson process with jump intensity κ, $f(z)$ is log-normal density: $f(z) = (1/\sqrt{2\pi}\omega z)e^{-(\ln z - \mu)^2/2\omega^2}$ with $E[z] = e^{\mu + \omega^2/2}$ and $D[z] = (e^{\omega^2} - 1)e^{2\mu + \omega^2}$, μ, is mean of jump z and ω is the variance of jump z and $x \in R$, $y \in R$. Then $\widetilde{N}(dt, dz) = dN(t) - \kappa f(z)dz\,dt$. Here we let $\mu = -0.9$, $\omega = 0.45$, and $\kappa = 0.1$. Then

$$\int_{\mathbf{R}} |z|^2 \nu(dz) = \kappa E\left[z^2\right] = \kappa e^{2\mu + 2\omega^2} = 0.025. \tag{46}$$

One can write (45) as the following stochastic differential delay equation with jumps:

$$\begin{aligned} dX(t) &= -X(t - \delta)\,dt \\ &\quad + [0.5X(t) - 0.5X(t - \delta)]\,dB(t) \\ &\quad + \int_{\mathbf{R}_0} z\widetilde{N}(dt, dz). \end{aligned} \tag{47}$$

It is easy to see that hypotheses (H2.1)-(H2.1) are satisfied with $C_1 = 1.5$, $C_2 = 1.5$. On the other hand, it is easy to see that condition (9) is satisfied with $\lambda = 8$ and (10) becomes $\delta < 0.037$.

Therefore, by Theorems 5 and 8, we can conclude that (47) is both mean square and almost surely exponentially stable provided $\delta < 0.037$.

Particularly, $\kappa = 0$; then (45) reduces to SDE with delay

$$dX(t) = -X(t - \delta)\,dt + [0.5X(t) - 0.5X(t - \delta)]\,dB(t). \tag{48}$$

It is easy to see that hypotheses (H2.1)-(H2.2) are satisfied with $C_1 = 1.5$ and $C_2 = 1.5$. On the other hand, it is easy to see that condition (9) is satisfied with $\lambda = 8$ and (40) becomes $\delta < 0.395$.

Therefore, by Corollary 11 and Theorem 8, we can conclude that (48) is both mean square and almost surely exponentially stable provided $\delta < 0.395$.

Moreover, setting $b(t, x, y) = -y$, $\sigma(t, x, y) = 0$, and $\kappa = 0$, then (45) becomes

$$dX(t) = -X(t-\delta)\,dt. \tag{49}$$

It is easy to see that hypotheses (H2.1)-(H2.2) are satisfied with $C_1 = 1$ and $C_2 = 1$. On the other hand, it is easy to see that condition (9) is satisfied with $\lambda = 8$ and (37) becomes $\delta < 2\sqrt{2}$.

Therefore, by Corollary 10 and Theorem 8, we can conclude that (49) is both mean square and almost surely exponentially stable provided $\delta < 2\sqrt{2}$.

Remark 14. Figure 1 gives the simulation results of Example 1 when $\sigma \neq 0$, $\kappa \neq 0$, and $\delta < 0.037$. The parameter values used in the calculations are $\|G\| = 1$, $\lambda = 8$, $C_1 = 1.5$, $C_2 = 1.5$, and $\delta = 0.03$. Figure 2 gives the simulation results of Example 1 when $\sigma \neq 0$, $\kappa = 0$, and $\delta < 0.395$. The parameter values used in the calculations are $\|G\| = 1$, $\lambda = 8$, $C_1 = 1.5$, $C_2 = 1.5$, and $\delta = 0.3$. Figure 3 gives the simulation results of Example 1 when $\sigma = 0$, $\kappa = 0$, and $\delta < 2\sqrt{2}$. The parameter values used in the calculations are $\|G\| = 1$, $\lambda = 8$, $C_1 = 1$, $C_2 = 1$, and $\delta = 1$.

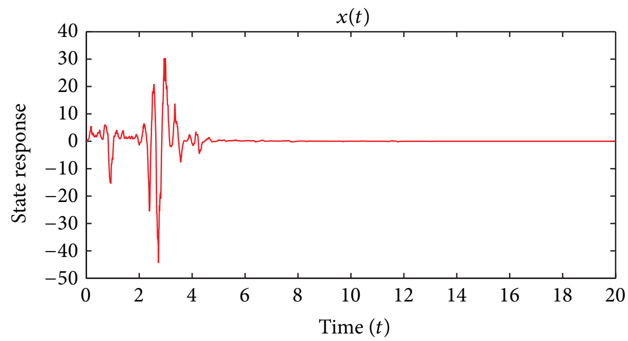


FIGURE 1: The simulation results of Example 1 when $\sigma \neq 0$, $\kappa \neq 0$, and $\delta = 0.03$.

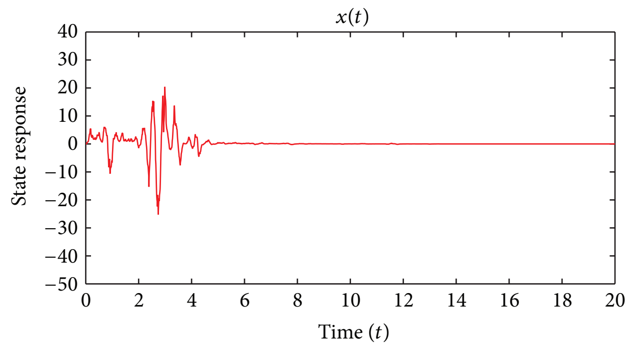


FIGURE 2: The simulation results of Example 1 when $\sigma \neq 0$, $\kappa = 0$, and $\delta = 0.3$.

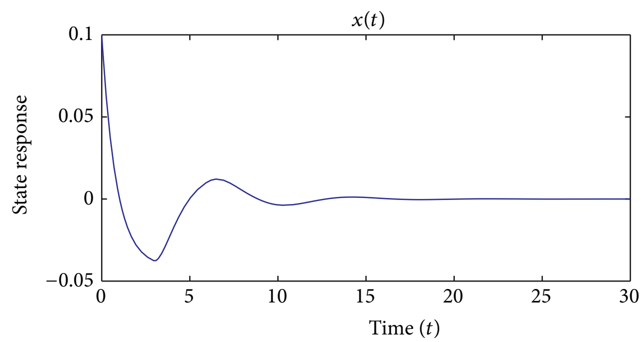


FIGURE 3: The simulation results of Example 1 when $\sigma = 0$, $N = 0$, $\delta = 1$.

5. Concluding Remarks

In this paper, we investigate not only the exponential stability in the mean square but also the almost surely exponential stability for a class of SDDE with Poisson jumps based on Lyapunov stability theory, Itô formula, stochastic analysis, and inequality technique. We first consider the existence and uniqueness of the adapted solution by employing fixed point theorem. Next, some sufficient conditions of exponential stability and corollaries for SDDE with Poisson jumps are obtained by using Lyapunov function. By utilizing Doob's martingale inequality and Borel-Cantelli lemma, we find that the exponentially stable in the mean square of SDDE with Poisson jumps implies the almost surely exponentially stable. Our results generalize and improve some recent results ([5–8, 14–17]). In particular, our results show that if SDE is exponentially stable and the time delay is sufficiently small, then the corresponding SDDE with Poisson jumps will remain exponentially stable. Moreover, when the system is exponentially stable, the time delay upper limit is solved by using our results which are more easily verified and applied in practice. Our approach in the current paper is different from the above [14–17]. Finally, we present a simple example to illustrate the effectiveness of our stable results. Another challenging problem is to study a class of SDEs with variable delays and Poisson jumps. We hope to study these problems in forthcoming papers.

Appendix

We now present proof of Theorem 2.

Proof of Theorem 2. Let us define a norm in Banach space $L^2_{\mathcal{F}}([-\delta, T]; \mathbf{R}^n)$ as follows:

$$|\chi(\cdot)|_\vartheta = \left(E\left[\int_{-\delta}^{T} e^{-\vartheta s}|\chi(s)|^2 ds\right]\right)^{1/2}, \quad \vartheta > 0. \tag{A.1}$$

Clearly it is equivalent to the original norm of $L^2_{\mathcal{F}}([-\delta, T]; \mathbf{R}^n)$. We consider

$$\begin{aligned} X(t) = \xi(0) &+ \int_0^t b(s, \chi, y_\chi)\, ds \\ &+ \int_0^t \sigma(s, \chi, y_\chi)\, dB(s) \\ &+ \int_{\mathbf{R}_0^n} \gamma(s, X(s^-), Y(s^-), z)\, \widetilde{N}(ds, dz); \\ &\qquad\qquad t \in [0, T] \\ X(t) = \xi(t), &\quad t \in [-\delta, 0], \end{aligned} \tag{A.2}$$

where $y_\chi = \chi(t-\delta)$. Define a mapping $\mathbb{T} : L^2_{\mathcal{F}}([-\delta, T]; \mathbf{R}^n) \to L^2_{\mathcal{F}}([-\delta, T]; \mathbf{R}^n)$ such that $\mathbb{T}(\chi(\cdot)) = X(\cdot)$. We desire to prove that $\mathbb{T}$ is a contraction mapping under the norm $|\chi(\cdot)|_\vartheta$. For arbitrary $\chi(\cdot), \chi'(\cdot) \in L^2_{\mathcal{F}}([-\delta, T]; \mathbf{R}^n)$, set $\mathbb{T}(\chi(\cdot)) = X(\cdot)$, $\mathbb{T}(\chi'(\cdot)) = \chi'(\cdot)$, and $\widehat{\chi}(\cdot) = \chi(\cdot) - \chi'(\cdot)$, $\widehat{X}(\cdot) = X(\cdot) - X'(\cdot)$. Then, $\widehat{X}(\cdot)$ satisfies

$$\begin{aligned} \widehat{X}(\cdot) = &\int_0^t \left[b(s, \chi, y_\chi) - b(s, \chi', y'_\chi) \right] ds \\ &+ \int_0^t \left[\sigma(s, \chi, y_\chi) - \sigma(s, \chi', y'_\chi) \right] dB(s) \\ &+ \int_{\mathbf{R}_0^n} \left[\gamma(s, \chi(s^-), y_\chi(s^-), z) \right. \\ &\qquad \left. - \gamma(s, \chi'(s^-), y'_\chi(s^-), z) \right] \widetilde{N}(ds, dz); \\ &\qquad\qquad t \geq 0, \\ \widehat{X}(\cdot) = 0, &\quad t \in [-\delta, 0]. \end{aligned} \tag{A.3}$$

Applying Itô's formula to $e^{-\vartheta t} |\widehat{X}(t)|^2$ and taking the expectation, we have

$$\begin{aligned} &\vartheta E \int_0^T e^{-\vartheta t} \left| \widehat{X}(t) \right|^2 dt \\ &= 2E \int_0^T e^{-\vartheta t} \widehat{X}(t) \left| b(t, \chi, y_\chi) - b(t, \chi', y'_\chi) \right| dt \\ &\quad + E \int_0^T e^{-\vartheta t} \left| \sigma(s, \chi, y_\chi) - \sigma(s, \chi', y'_\chi) \right|^2 dt \\ &\quad + E \int_0^T \sum_{k=1}^{l} \int_{|z_k|<R} e^{-\vartheta t} \left| \gamma^{(k)}(t, \chi(t^-), y_\chi(t^-), z_k) \right. \\ &\qquad\qquad \left. - \gamma^{(k)}(t, \chi'(t^-), y'_\chi(t^-), z_k) \right|^2 \\ &\quad \times \nu_k(dz_k)\, dt. \end{aligned} \tag{A.4}$$

Lemma 1 yields

$$\begin{aligned} &\vartheta E \int_0^T e^{-\vartheta t} \left| \widehat{X}(t) \right|^2 dt \\ &\leq E \int_0^T e^{-\vartheta t} \left(\left| \widehat{X}(t) \right|^2 + \left| b(t, \chi, y_\chi) - b(t, \chi', y'_\chi) \right|^2 \right) dt \\ &\quad + E \int_0^T e^{-\vartheta t} \left| \sigma(s, \chi, y_\chi) - \sigma(s, \chi', y'_\chi) \right|^2 dt \\ &\quad + E \int_0^T \sum_{k=1}^{l} \int_{|z_k|<R} e^{-\vartheta t} \left| \gamma^{(k)}(t, \chi(t^-), y_\chi(t^-), z_k) \right. \\ &\qquad\qquad \left. - \gamma^{(k)}(t, \chi'(t^-), y'_\chi(t^-), z_k) \right|^2 \\ &\quad \times \nu_k(dz_k)\, dt. \end{aligned} \tag{A.5}$$

Then by (H2.2), we obtain

$$\begin{aligned} &(\vartheta - 1) E \int_0^T e^{-\vartheta t} \left| \widehat{X}(t) \right|^2 dt \\ &\leq 3C_2 \cdot E \left[\int_0^T e^{-\vartheta t} \left| \widehat{\chi}(t) \right|^2 dt \right. \\ &\qquad \left. + \int_0^T e^{-\vartheta t} \left| \widehat{y}_\chi(t) \right|^2 dt \right], \end{aligned} \tag{A.6}$$

where

$$\begin{aligned} &\int_0^T e^{-\vartheta t} \left| \widehat{y}_\chi(t) \right|^2 dt \underline{\underline{\tau = t - \delta}} \\ &= e^{-\vartheta \delta} \int_{-\delta}^{T-\delta} e^{-\vartheta \tau} \left| \widehat{\chi}(\tau) \right|^2 d\tau \leq \int_{-\delta}^{T-\delta} e^{-\vartheta \tau} \left| \widehat{\chi}(\tau) \right|^2 d\tau \\ &\leq \int_{-\delta}^{T} e^{-\vartheta \tau} \left| \widehat{\chi}(\tau) \right|^2 d\tau. \end{aligned} \tag{A.7}$$

Then,

$$(\vartheta - 1) E \int_0^T e^{-\vartheta t} \left| \widehat{X}(t) \right|^2 dt \leq 6C_2 \cdot E \int_{-\delta}^T e^{-\vartheta t} \left| \widehat{\chi}(t) \right|^2 dt. \tag{A.8}$$

Let $\vartheta = 12C_2 + 1$, then the above yields

$$E \int_{-\delta}^T e^{-\vartheta t} \left| \widehat{X}(t) \right|^2 dt \leq \frac{1}{2} E \int_{-\delta}^T e^{-\vartheta t} \left| \widehat{\chi}(t) \right|^2 dt. \tag{A.9}$$

That is,

$$\left| \widehat{X}(\cdot) \right|_\vartheta = \frac{1}{\sqrt{2}} \left| \widehat{\chi}(\cdot) \right|_\vartheta. \tag{A.10}$$

This implies that $\mathbb{T}$ is a strict contraction mapping. Then it follows from the fixed point theorem that (1)-(2) has a unique solution in $L^2_{\mathcal{F}}([-\delta, T]; \mathbf{R}^n)$. Since b and σ satisfy (H2.1) and (H2.2), we can easily derive that $E[|X(t;\xi)|^2] < \infty$, and $x(t;\xi)$ is continuous with respect to $t \in [0, T]$. Furthermore, by $b(t, 0, 0, 0) = \sigma(t, 0, 0, 0) \equiv 0$, (1)-(2) have a trivial solution $X(t; 0) = 0$. □

Proof of Lemma 6. For any $t \geq \delta$, we easily get

$$
\begin{aligned}
&\int_{\delta}^{t} e^{\varepsilon s} \int_{s-\delta}^{s} E|X(\tau)|^2 d\tau\, ds \\
&\quad = \int_{0}^{t} E|X(\tau)|^2 \left(\int_{\tau \vee \delta}^{(\tau+\delta)\wedge t} e^{\varepsilon s} ds \right) d\tau \\
&\quad \leq \delta e^{\varepsilon\delta} \int_{0}^{t} e^{\varepsilon\tau} \cdot E|X(\tau)|^2 d\tau. \\
&\int_{0}^{t} e^{\varepsilon s} E|X(s-\delta)|^2 ds \\
&\quad \leq e^{\varepsilon\delta} \int_{0}^{\delta} E|X(s-\delta)|^2 ds \\
&\quad\quad + e^{\varepsilon\delta} \int_{\delta}^{t} e^{\varepsilon(s-\delta)} E|X(s-\delta)|^2 ds \underline{\underline{\tau = s-\delta}} e^{\varepsilon\delta} \\
&\quad\quad \times \int_{-\delta}^{0} E|\xi(\tau)|^2 d\tau \\
&\quad\quad + e^{\varepsilon\delta} \int_{0}^{t-\delta} e^{\varepsilon\tau} E|X(\tau)|^2 d\tau \\
&\quad \leq c_{11} e^{\varepsilon\delta} + e^{\varepsilon\delta} \int_{0}^{t} e^{\varepsilon\tau} E|X(\tau)|^2 d\tau
\end{aligned}
\tag{A.11}
$$

for any $t \geq \delta$, where $c_{11} \geq \int_{-\delta}^{0} E|\xi(\tau)|^2 d\tau$. □

Proof of Lemma 7. Similar to (11), for any $t \geq \delta$, we have

$$
\begin{aligned}
&\int_{\delta}^{t} e^{\varepsilon s} \int_{s-\delta}^{s} E|X(\tau-\delta)|^2 d\tau\, ds \\
&\quad \leq \delta e^{\varepsilon\delta} \int_{0}^{t} e^{\varepsilon\tau} E|X(\tau-\delta)|^2 d\tau.
\end{aligned}
\tag{A.12}
$$

Substituting (12) into the above inequality yields

$$
\begin{aligned}
&\int_{\delta}^{t} e^{\varepsilon s} \int_{s-\delta}^{s} E|X(\tau-\delta)|^2 d\tau\, ds \\
&\quad < \delta c_{11} e^{2\varepsilon\delta} + \delta e^{2\varepsilon\delta} \int_{0}^{t} e^{\varepsilon s} E|X(s)|^2 ds.
\end{aligned}
\tag{A.13}
$$

The relation (13) in Lemma 7 is then proved.

On the other hand, for $s \geq \delta$, we have

$$
\begin{aligned}
X(s) - X(s-\delta) = &\int_{s-\delta}^{s} b(t, X, Y)\, dt \\
&+ \int_{s-\delta}^{s} \sigma(t, X, Y)\, dB(t) \\
&+ \int_{s-\delta}^{s} \int_{\mathbf{R}^n} \gamma(t, X, Y, z)\, \widetilde{N}(dt, dz).
\end{aligned}
\tag{A.14}
$$

By (H2.1), we get

$$
\begin{aligned}
E|X-Y|^2 \leq\, &3\delta E \int_{s-\delta}^{s} |b(t, X, Y)|^2 dt \\
&+ 3E \int_{s-\delta}^{s} \operatorname{tr}\left[\sigma^T(t, X, Y)\, \sigma(t, X, Y)\right] dt \\
&+ 3E \int_{s-\delta}^{s} \int_{\mathbf{R}_0} \sum_{k=1}^{l} \left|\gamma^{(k)}(t, X, Y, z_k)\right|^2 \nu_k(dz_k)\, dt \\
= \,&3\delta C_1(\delta+2) + 3C_1(\delta+2) \\
&\times \int_{s-\delta}^{s} \left(E|X|^2 + E|Y|^2\right) dt.
\end{aligned}
\tag{A.15}
$$

Similar to (12), for $t \geq \delta$, we have

$$
\begin{aligned}
&\int_{0}^{t} e^{\varepsilon s} E|X-Y|^2 ds \\
&\quad = \int_{0}^{\delta} e^{\varepsilon s} E|X(s) - X(s-\delta)|^2 ds \\
&\quad\quad + \int_{\delta}^{t} e^{\varepsilon s} E|X(s) - X(s-\delta)|^2 ds \\
&\quad \leq c_{22} + 3C_1(\delta+2) \int_{\delta}^{t} e^{\varepsilon s} \int_{s-\delta}^{s} \left(E|X|^2 + E|Y|^2\right) dt\, ds,
\end{aligned}
\tag{A.16}
$$

where

$$
\begin{aligned}
c_{22} \geq &\int_{0}^{\delta} e^{\varepsilon s} E|X(s) - X(s-\delta)|^2 ds \\
&+ \frac{3\delta C_1(\delta+2)}{\varepsilon} \left(e^{\varepsilon T} - e^{\varepsilon\delta}\right).
\end{aligned}
\tag{A.17}
$$

Substituting (11) and (13) into (A.16), for $t \geq \delta$, we get

$$
\begin{aligned}
&\int_{0}^{t} e^{\varepsilon s} E|X(s) - X(s-\delta)|^2 ds \\
&\quad \leq c_{22} + 3C_1(\delta+2)\, \delta e^{\varepsilon\delta} \int_{0}^{t} e^{\varepsilon s} E|X(s)|^2 ds \\
&\quad\quad + 3C_1(\delta+2) \left(c_{11} \delta e^{2\varepsilon\delta} + \delta e^{2\varepsilon\delta} \int_{0}^{t} e^{\varepsilon s} E|X(s)|^2 ds \right).
\end{aligned}
\tag{A.18}
$$

□

Conflict of Interests

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

Acknowledgment

This work is supported by the Fundamental Research Funds for the Central Universities (JBK140205).

References

[1] K. Itô, "Differential equations determining Markov processes," *Zenkoku Shijo Sugaku Danwakai*, vol. 244, no. 1077, pp. 1352–1400, 1942.

[2] S. Cong, "On exponential stability conditions of linear neutral stochastic differential systems with time-varying delay," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 11, pp. 1265–1276, 2013.

[3] Q. Zhu and X. Li, "Exponential and almost sure exponential stability of stochastic fuzzy delayed Cohen-Grossberg neural networks," *Fuzzy Sets and Systems*, vol. 203, pp. 74–94, 2012.

[4] Y. Sun and J. Cao, "pth moment exponential stability of stochastic recurrent neural networks with time-varying delays," *Nonlinear Analysis: Real World Applications*, vol. 8, no. 4, pp. 1171–1185, 2007.

[5] Q. Zhu and B. Song, "Exponential stability of impulsive nonlinear stochastic differential equations with mixed delays," *Nonlinear Analysis: Real World Applications*, vol. 12, no. 5, pp. 2851–2860, 2011.

[6] X. Mao and A. Shah, "Exponential stability of stochastic differential delay equations," *Stochastics and Stochastics Reports*, vol. 60, no. 1-2, pp. 135–153, 1997.

[7] W. Zhu, X. Ruan, Y. Qin, and J. Zhuang, "Exponential stability of stochastic nonlinear dynamical price system with delay," *Mathematical Problems in Engineering*, vol. 2013, Article ID 168169, 9 pages, 2013.

[8] W. Zhu and J. Hu, "Stability analysis of stochastic delayed cellular neural networks by LMI approach," *Chaos, Solitons and Fractals*, vol. 29, no. 1, pp. 171–174, 2006.

[9] B. Øksendal and A. Sulem, "A maximum principle for optimal control of stochastic systems with delay, with applications to finance," in *Optimal Control and Partial Differential Equations*, J. M. Menaldi, E. Rofman, and A. Sulem, Eds., pp. 64–79, ISO Press, Amsterdam, The Netherlands, 2000.

[10] B. Øksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Universitext, Springer, Berlin, Germany, 2nd edition, 2007.

[11] F. E. Benth and T. Meyer-Brandis, "The density process of the minimal entropy martingale measure in a stochastic volatility model with jumps," *Finance and Stochastics*, vol. 9, no. 4, pp. 563–575, 2005.

[12] D. Bates, "The crash of '87: Was it expected? The evidence from options markets," *Journal of Finance*, vol. 46, no. 3, pp. 1009–1044, 1991.

[13] L. Guo and Q. Zhu, "Stability analysis for stochastic Volterra-Levin equations with Poisson jumps: fixed point approach," *Journal of Mathematical Physics*, vol. 52, no. 4, Article ID 042702, 2011.

[14] Q. Zhu and J. Cao, "Robust exponential stability of markovian jump impulsive stochastic Cohen-Grossberg neural networks with mixed time delays," *IEEE Transactions on Neural Networks*, vol. 21, no. 8, pp. 1314–1325, 2010.

[15] Q. Zhu and J. Cao, "Exponential stability of stochastic neural networks with both Markovian jump parameters and mixed time delays," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 41, no. 2, pp. 341–353, 2011.

[16] Q. Zhu, "Asymptotic stability in the pth moment for stochastic differential equations with Lévy noise," *Journal of Mathematical Analysis and Applications*, vol. 416, no. 1, pp. 126–142, 2014.

[17] L. Chen, M. Chen, and L. Guo, "Exponential stability of impulsive stochastic delay partial differential equations with Poisson jumps," *Journal of Ningbo University*, vol. 25, no. 3, pp. 30–34, 2012.

[18] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of *SIAM Studies in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1994.

[19] V. B. Kolmanovskii and V. R. Nosov, *Stability of Functional Differential Equations*, Academic Press, 1986.

[20] Y. Hu and F. Wu, "Stochastic Kolmogorov-type population dynamics with infinite distributed delays," *Acta Applicandae Mathematica*, vol. 110, no. 3, pp. 1407–1428, 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 262314, 11 pages
http://dx.doi.org/10.1155/2014/262314

Research Article

Linear Quadratic Nonzero Sum Differential Games with Asymmetric Information

Dejian Chang[1,2] and Hua Xiao[3]

[1] Qilu Securities Institute for Financial Studies, Shandong University, Jinan 250100, China
[2] School of Mathematics, Shandong University, Jinan 250100, China
[3] School of Mathematics and Statistics, Shandong University, Weihai 264209, China

Correspondence should be addressed to Hua Xiao; xiao_hua@sdu.edu.cn

Received 3 March 2014; Accepted 7 July 2014; Published 5 August 2014

Academic Editor: Ming Gao

This paper studies a linear quadratic nonzero sum differential game problem with asymmetric information. Compared with the existing literature, a distinct feature is that the information available to players is asymmetric. Nash equilibrium points are obtained for several classes of asymmetric information by stochastic maximum principle and technique of completion square. The systems of some Riccati equations and forward-backward stochastic filtering equations are introduced and the existence and uniqueness of the solutions are proved. Finally, the unique Nash equilibrium point for each class of asymmetric information is represented in a feedback form of the optimal filtering of the state, through the solutions of the Riccati equations.

1. Introduction

Throughout this article, we denote by R^k the k-dimensional Euclidean space, $R^{k\times l}$ the collection of $k \times l$ matrices. The superscript $*$ denotes the transpose of vectors or matrices. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t), P)$ be a complete filtered probability space in which $\mathscr{F}_t$ denotes a natural filtration generated by a three dimensional standard Brownian motion $(W_1(t), W_2(t), W_3(t))$, $\mathscr{F} = \mathscr{F}_T$, and $T > 0$ be a fixed time horizon. For a given Euclidean space, we denote by $\langle \cdot, \cdot \rangle$ (resp., $|\cdot|$) the scalar product (resp., norm). We also denote by $\mathscr{L}^2_{\mathscr{F}_t}(0,T;S)$ the space of all S-valued, $\mathscr{F}_t$-adapted and square integrable processes, by $\mathscr{L}^2_{\mathscr{F}_T}(\Omega;S)$ the space of all S-valued, $\mathscr{F}_T$-measurable and square integrable random variables, by $\mathscr{L}^2(0,T;S)$ the space of all S-valued functions satisfying $\int_0^T |f(t)|^2 dt < \infty$, and by $f(t)^2$ the square of $f(t)$. For the sake of simplicity, we set

$$
\begin{gathered}
\mathscr{F}_t^j = \sigma\left\{W_j(s), 0 \le s \le t\right\} \quad (j = 1,2,3),\\
\mathscr{F}_t^{1,2} = \sigma\left\{W_1(s), W_2(s), 0 \le s \le t\right\},\\
\mathscr{F}_t^{2,3} = \sigma\left\{W_2(s), W_3(s), 0 \le s \le t\right\},\\
\widehat{h}(t) = \mathbb{E}\left(h(t) \mid \mathscr{F}_t^{1,2}\right),\\
\widetilde{h}(t) = \mathbb{E}\left(h(t) \mid \mathscr{F}_t^{2}\right), \qquad \check{h}(t) = \mathbb{E}\left(h(t) \mid \mathscr{F}_t^{3}\right),\\
\bar{h}(t) = \mathbb{E}\left(h(t) \mid \mathscr{F}_t^{2,3}\right), \qquad \dot{h}(t) = \frac{dh(t)}{dt}.
\end{gathered} \tag{1}
$$

This work is interested in linear quadratic (LQ, for short) non-zero sum differential game with asymmetric information. For simplicity, we only study the case of two players. Let us now begin to specify the problem. Consider the following one-dimensional stochastic differential equation (SDE, for short)

$$
\begin{aligned}
dx^{v_1,v_2}(t) = {} & \left[a(t)x^{v_1,v_2}(t) + b_1(t)v_1(t)\right.\\
& \left. + b_2(t)v_2(t) + c(t)\right]dt + g_1(t)dW_1(t)
\end{aligned}
$$

$$\begin{aligned} &+ \left[e(t)\, x^{v_1,v_2}(t) + g_2(t)\right] dW_2(t) \\ &+ g_3(t)\, dW_3(t), \\ x^{v_1,v_2}(0) &= x_0, \end{aligned} \tag{2}$$

and cost functionals of the form

$$J_i\left(v_1(\cdot), v_2(\cdot)\right) = \frac{1}{2}\mathbb{E}\left[\int_0^T \left(l_i(t)\, x^{v_1,v_2}(t)^2 + m_i(t)\, v_i(t)^2\right) dt + r_i x^{v_1,v_2}(T)^2\right] \quad (i = 1, 2). \tag{3}$$

Here $a, b_1, b_2, c, e, g_1, g_2$ and g_3 are bounded and deterministic functions in t, l_1 and l_2 are bounded, nonnegative and deterministic functions in t, m_1 and m_2 are bounded, positive and deterministic functions in t, and r_1 and r_2 are two nonnegative constants. Hereinafter, we omit all dependence on time variable t of all processes or deterministic functions if there is no risk of ambiguity from the context for the notational simplicity; $v_1(\cdot)$ and $v_2(\cdot)$ are the control processes of Player 1 and Player 2, respectively. We always use the subscript 1 (resp., the subscript 2) to characterize the control variable corresponding to Player 1 (resp., Player 2) and use the notation x^{v_1,v_2} to denote the dependence of the state on the control variable (v_1, v_2).

Let $\mathcal{F}_t$ denote the full information up to time t and $\mathcal{G}_t^i \subseteq \mathcal{F}_t$ be a given sub-filtration which represents the information available to Player i $(i = 1, 2)$ at time $t \in [0, T]$. If $\mathcal{G}_t^i \subseteq \mathcal{F}_t$ and $\mathcal{G}_t^i \neq \mathcal{F}_t$, we call the available information *partial* or *incomplete* for Player i. If $\mathcal{G}_t^1 \neq \mathcal{G}_t^2$, we call the available information *asymmetric* for Player 1 and Player 2. Now we introduce the admissible control set

$$\mathcal{U}_i = \left\{v_i(\cdot) \in \mathcal{L}^2_{\mathcal{G}_t^i}(0, T; R) \mid v_i(t) \in U_i, t \in [0, T]\right\}, \tag{4}$$

where $\mathcal{G}^i = \mathcal{G}_T^i$ and U_i are nonempty convex subsets of R $(i = 1, 2)$. Each element of $\mathcal{U}_i$ is called an open-loop admissible control for Player i $(i = 1, 2)$. And $\mathcal{U}_1 \times \mathcal{U}_2$ is said to be the set of open-loop admissible controls for the players.

Suppose each player hopes to minimize her/his cost functional $J_i(v_1(\cdot), v_2(\cdot))$ by selecting a suitable admissible control $v_i(\cdot)$ $(i = 1, 2)$. In this study, the problem is, under the setting of asymmetric information, to look for $(u_1(\cdot), u_2(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2$ which is called the Nash equilibrium point of the game, such that

$$\begin{aligned} J_1\left(u_1(\cdot), u_2(\cdot)\right) &= \min_{v_1(\cdot)\in\mathcal{U}_1} J_1\left(v_1(\cdot), u_2(\cdot)\right), \\ J_2\left(u_1(\cdot), u_2(\cdot)\right) &= \min_{v_2(\cdot)\in\mathcal{U}_2} J_1\left(u_1(\cdot), v_2(\cdot)\right). \end{aligned} \tag{5}$$

We call the problem above an LQ non-zero sum differential game with asymmetric information. For simplicity, we denote it by *Problem (LQ NZSDG)*.

The LQ problems constitute an extremely important class of optimal control or differential game problems, since they can model many problems in applications, and also reasonably approximate nonlinear control or game problems. On the other hand, there also exist so called partial and asymmetric information problems in real world. For example, investors only partially know the information from security market (see [1, 2]); in many situations, "insider trading" maybe exist, which means that the insider has access to material and non-public information about the security and the available information is asymmetric between the insider and the common trader (see, e.g., [3, 4]); the principal faces information asymmetric and risk with regards to whether the agent has effectively completed a contract, when a principal hires an agent to perform specific duties (see, e.g., [5, 6]). For more information about LQ control or game problems, the interested readers may refer the following partial list of the works including [7–13] with complete information, and [14] with partial information, and the references therein.

It is very important and meaningful to find explicit Nash equilibrium points for differential game problems. When the available information is partial or asymmetric, we need to derive the corresponding optimal filtering of the states and adjoint variables which will be used to represent the Nash equilibrium points. It is very difficult to obtain the equations satisfied by the optimal filtering when the available information is asymmetric for Player 1 and Player 2. Up till now, it seems that there has been no literature about LQ differential games with asymmetric information $\mathcal{G}_t^1$ and $\mathcal{G}_t^2$. However, in case where $\mathcal{G}_t^i$ $(i = 1, 2)$ are chosen as certain special forms, we can still derive the filtering equations and then obtain the explicit form of the Nash equilibrium point. In the sequel, we will study *Problem (LQ NZSDG)* under the following four classes of asymmetric information:

(i) $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^{2,3}$; that is, the two players possess the common partial information $\mathcal{F}_t^2$;

(ii) $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$; that is, Player 1 possesses more information than Player 2;

(iii) $\mathcal{G}_t^1 = \mathcal{F}_t$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$; that is, Player 1 possesses the full information and Player 2 possesses the partial informaion;

(iv) $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^3$; that is, the two players possess the mutually independent information.

In Section 3, we will point out that some other cases similar to (i)–(iv) can be also solved by the same idea and method. To our knowledge, this paper is the first try to study LQ nonzero sum differential games in the setting of the asymmetric information.

The rest of this paper is organized as follows. In Section 2, we introduce some preliminaries which will be used to derive the forward-backward filtering equations and prove the corresponding existence and uniqueness of the solutions. In Section 3, we obtain the unique explicit Nash equilibrium point for each class of asymmetric information above. We also introduce some Riccati equations and represent the unique Nash equilibrium point in a feedback form of the optimal filtering of the state with respect to the corresponding asymmetric information, through the solutions of the Riccati equations. Some conclusions are given in Section 4.

2. Preliminary Results

In this section, we are going to introduce two lemmas, which will be often used later. First, we present existence and uniqueness for the solutions of the forward-backward stochastic differential equation (FBSDE, for short), whose dynamics is described by

$$\begin{aligned} dx &= b(t,x,y)\,dt + \sigma(t,x,y)\,dW, \\ -dy &= f(t,x,y,z)\,dt - z\,dW, \\ x(0) &= x_0, \qquad y(T) = \varphi(x(T)). \end{aligned} \tag{6}$$

Here $x(\cdot)$ satisfies an (forward) SDE, $(y(\cdot), z(\cdot))$ satisfies a backward stochastic differential equation, $W(\cdot)$ is a d-dimensional standard Brownian motion, (x, y, z) takes value in $R^n \times R^n \times R^{n\times d}$, and b, σ, f, and φ are the mappings with suitable sizes.

We introduce the notations

$$u = (x,y,z)^*, \qquad A(t,u) = (-f, b, \sigma)^*(t,u), \tag{7}$$

and make the following assumption.

(H_1) $A(t,u)$ and φ are uniformly Lipschitz continuous with respect to their variables; for each x, $\varphi(x)$ is in $\mathscr{L}^2_{\mathscr{F}_T}(\Omega; R^n)$; for every $(\omega, t) \in \Omega \times [0,T]$, $b(\omega,t,0,0) \in \mathscr{L}^2_{\mathscr{F}}(0,T;R^n)$, $\sigma(\omega,t,0,0) \in \mathscr{L}^2_{\mathscr{F}}(0,T;R^{n\times d})$, and $f(\omega,t,0,0,0) \in \mathscr{L}^2_{\mathscr{F}}(0,T;R^n)$.

We also make the following assumption.

(H_2) The functions $A(t,u)$ and φ satisfy the monotonic conditions:

$$\begin{aligned} &\langle A(t,u_1) - A(t,u_2), u_1 - u_2\rangle \\ &\qquad \le -\kappa_1 |x_1 - x_2|^2 - \kappa_2 |y_1 - y_2|^2, \\ &\langle \varphi(x_1) - \varphi(x_2), x_1 - x_2\rangle \ge \kappa_3 |x_1 - x_2|^2, \\ &\forall u_1 - u_2 = (x_1 - x_2, y_1 - y_2, z_1 - z_2), \end{aligned} \tag{8}$$

where κ_1, κ_1, and κ_3 are given nonnegative constants satisfying $\kappa_1 + \kappa_2 > 0$, $\kappa_2 + \kappa_3 > 0$.

Then we have the following lemma, which is a direct deduction of Theorem 1 in Wu and Yu [11] with no random jumps.

Lemma 1. *If the assumptions* (H_1) *and* (H_2) *hold, then* (6) *has a unique triple* $(x(\cdot), y(\cdot), z(\cdot)) \in \mathscr{L}^2_{\mathscr{F}_t}(0,T;R^{n+n+n\times d})$.

Remark 2. If we assume $\sigma \equiv 0$ and all functions are deterministic, then (6) is reduced to a forward-backward ordinary differential equation (ODE, for short):

$$\begin{aligned} dx &= b(t,x,y)\,dt, \\ -dy &= f(t,x,y)\,dt, \\ x(0) &= x_0, \qquad y(T) = \varphi(x(T)). \end{aligned} \tag{9}$$

We define the notation $u = (x,y)^*$, $G(t,u) = (-f, b)^*(t,u)$. If b, f, φ, and G satisfy the assumptions (H_1) and (H_2) with $\mathscr{L}^2_{\mathscr{F}_t}(0,T;S)$ replaced by $\mathscr{L}^2(0,T;S)$ and φ is uniformly bounded, then (9) has a unique solution $(x,y) \in \mathscr{L}^2(0,T;\mathbf{R}^{n+n})$.

The following lemma is from the monograph by Chung [15] (see the example, Section 9.2).

Lemma 3. *If $\mathscr{F}_1$, $\mathscr{F}_2$, and $\mathscr{F}_3$ are three σ-algebras, and $\mathscr{F}_1 \vee \mathscr{F}_2$ is independent of $\mathscr{F}_3$, then, for any integrable random variable $X \in \mathscr{F}_1$, we have $\mathbb{E}[X \mid \mathscr{F}_2 \vee \mathscr{F}_3] = \mathbb{E}[X \mid \mathscr{F}_2]$.*

3. Nash Equilibrium Point

In this section, we will derive the explicit form of the Nash equilibrium point for *Problem (LQ NZSDG)*, applying stochastic maximum principle for partial information optimal control problem and the technique of complete square. Further, we also introduce the Riccati equations and represent the Nash equilibrium point as a feedback of the optimal filters $\widehat{x}$, $\check{x}$, and $\overline{x}$, through the solutions to the Riccati equations.

We first introduce two LQ stochastic control problems with two pieces of general asymmetric information $\mathscr{G}_t^1$ and $\mathscr{G}_t^2$ which is closely related to *problem (LQ NZSDG)*.

Problem (LQSC1):

$$\begin{aligned} &\min \{ J_1^{u_2}(v_1(\cdot)) \mid v_1(\cdot) \in \mathscr{U}_1 \}, \\ &J_1^{u_2}(v_1(\cdot)) = \frac{1}{2}\mathbb{E}\left[\int_0^T \left(l_1 (x^{v_1})^2 + m_1 (v_1)^2\right) dt + r_1 x^{v_1}(T)^2\right], \end{aligned} \tag{10}$$

subject to

$$\begin{aligned} dx^{v_1} = {} & [a x^{v_1} + b_1 v_1(t) + b_2 u_2 + c]\,dt \\ & + g_1 dW_1 + [e x^{v_1} + g_2]\,dW_2 + g_3 dW_3, \\ x^{v_1}(0) = {} & x_0. \end{aligned} \tag{11}$$

Problem (LQSC2):

$$\begin{aligned} &\min \{ J_2^{u_1}(v_2(\cdot)) \mid v_2(\cdot) \in \mathscr{U}_2 \}, \\ &J_2^{u_1}(v_2(\cdot)) = \frac{1}{2}\mathbb{E}\left[\int_0^T \left(l_2 (x^{v_2})^2 + m_2 (v_2)^2\right) dt + r_2 x^{v_2}(T)^2\right], \end{aligned} \tag{12}$$

subject to

$$\begin{aligned} dx^{v_2} = {} & \left[ax^{v_2} + b_1 u_1(t) + b_2 v_2 + c\right] dt \\ & + g_1 dW_1 + \left[ex^{v_2} + g_2\right] dW_2 + g_3 dW_3, \\ x^{v_2}(0) = {} & x_0. \end{aligned} \tag{13}$$

We can check that $x^{u_1} = x^{u_2} = x^{u_1,u_2} \equiv x$, $J_1^{u_2}(u_1(\cdot)) = J_1(u_1(\cdot), u_2(\cdot))$, and $J_2^{u_1}(u_2(\cdot)) = J_2(u_1(\cdot), u_2(\cdot))$ hold. If (u_1, u_2) is a Nash equilibrium point, then, from the definition of Nash equilibrium point (see (5)), we can conclude that u_1 (resp., u_2) is an optimal control for *Problem (LQSC1)* (resp., *Problem (LQSC2)*). Appealing to the stochastic maximum principle under partial information (see [16], Remark 2.1 with the drift coefficient of the observation equation being zero and convex control domain, or [17], Theorem 3.1 with nonrandom jumps), we can derive the following necessary conditions of the optimal controls for *Problem (LQSC1)* and *Problem (LQSC2)*.

Lemma 4. *If u_1 (resp., u_2) is an optimal control for Problem (LQSC1) (resp., Problem (LQSC2)), then we have*

$$\begin{aligned} u_1(t) &= -m_1^{-1}(t) b_1(t) \mathbb{E}\left(y_1(t) \mid \mathcal{G}_t^1\right), \\ u_2(t) &= -m_2^{-1}(t) b_2(t) \mathbb{E}\left(y_2(t) \mid \mathcal{G}_t^2\right), \end{aligned} \tag{14}$$

where $(x, (y_1, z_{11}, z_{12}, z_{13}), (y_2, z_{21}, z_{22}, z_{23}))$ is a solution to the following FBSDE:

$$\begin{aligned} dx = {} & \left[ax - b_1^2 m_1^{-1} \mathbb{E}\left(y_1(t) \mid \mathcal{G}_t^1\right) - b_2^2 m_2^{-1} \mathbb{E}\left(y_2(t) \mid \mathcal{G}_t^2\right) \right. \\ & \left. + c\right] dt + g_1 dW_1 + \left[ex + g_2\right] dW_2 + g_3 dW_3, \end{aligned} \tag{15a}$$

$$-dy_1 = \left[ay_1 + ez_{12} + l_1 x\right] dt - z_{11} dW_1 - z_{12} dW_2 - z_{13} dW_3, \tag{15b}$$

$$-dy_2 = \left[ay_2 + ez_{22} + l_2 x\right] dt - z_{21} dW_1 - z_{22} dW_2 - z_{23} dW_3, \tag{15c}$$

$$x(0) = x_0, \qquad y_1(T) = r_1 x(T), \qquad y_2(T) = r_2 x(T). \tag{15d}$$

It is obvious that $(u_1, u_2) \in \mathcal{U}_1 \times \mathcal{U}_2$ is a candidate Nash equilibrium point for *Problem (LQ NZSDG)*. We will prove (u_1, u_2) is exactly a Nash equilibrium point in the sequel.

Lemma 5. *(u_1, u_2) in (21) is indeed a Nash equilibrium point for Problem (LQ NZSDG).*

Proof. For any $v_1(\cdot) \in \mathcal{U}_1$, we have

$$\begin{aligned} & J_1(v_1(\cdot), u_2(\cdot)) - J_1(u_1(\cdot), u_2(\cdot)) \\ & \quad = \frac{1}{2}\mathbb{E}\int_0^T \left[l_1 \left(x^{v_1,u_2} - x\right)^2 + m_1 \left(v_1 - u_1\right)^2\right] dt \\ & \quad + \frac{1}{2}\mathbb{E}\left[r_1 \left(x^{v_1,u_2}(T) - x(T)\right)^2\right] + \Theta, \end{aligned} \tag{16}$$

where

$$\begin{aligned} \Theta = {} & \mathbb{E}\int_0^T \left[l_1 x \left(x^{v_1,u_2} - x\right) + m_1 u_1 \left(v_1 - u_1\right)\right] dt \\ & + \mathbb{E}\left[r_1 x(T) \left(x^{v_1,u_2}(T) - x(T)\right)\right]. \end{aligned} \tag{17}$$

We apply Itô's formula to $y_1(x^{v_1,u_2} - x)$ and get

$$\begin{aligned} \Theta &= \mathbb{E}\int_0^T \left(m_1(t) u_1(t) + b_1(t) y_1(t)\right)\left(v_1(t) - u_1(t)\right) dt \\ &= \mathbb{E}\int_0^T \mathbb{E}\left[\left(m_1(t) u_1(t) + b_1(t) y_1(t)\right) \right. \\ & \qquad \left. \times \left(v_1(t) - u_1(t)\right) \mid \mathcal{G}_t^1\right] dt \\ &= \mathbb{E}\int_0^T \left(m_1(t) u_1(t) + b_1(t) \mathbb{E}\left(y_1(t) \mid \mathcal{G}_t^1\right)\right) \\ & \qquad \times \left(v_1(t) - u_1(t)\right) dt = 0. \end{aligned} \tag{18}$$

Then, because l_1 and r_1 are nonnegative, and m_1 is positive, we have

$$J_1(v_1(\cdot), u_2(\cdot)) - J_1(u_1(\cdot), u_2(\cdot)) \geq 0. \tag{19}$$

Similarly, for any $v_2(\cdot) \in \mathcal{U}_2$, we also have

$$J_2(u_1(\cdot), v_2(\cdot)) - J_2(u_1(\cdot), u_2(\cdot)) \geq 0. \tag{20}$$

Therefore, we can conclude that (u_1, u_2) in (14) is a Nash equilibrium point for *Problem (LQ NZSDG)* indeed. □

Combining Lemmas 4 and 5, we obtain the following theorem.

Theorem 6. *(u_1, u_2) is a Nash equilibrium point for Problem (LQ NZSDG) if and only if (u_1, u_2) has the form denoted by* (14) *and $(x, (y_1, z_{11}, z_{12}, z_{13}), (y_2, z_{21}, z_{22}, z_{23}))$ satisfies FBSDE* (15a)–(15d).

Remark 7. If (15a)–(15d) has a unique solution, then *Problem (LQ NZSDG)* has a unique Nash equilibrium point. If (15a)–(15d) have many solutions, then *Problem (LQ NZSDG)* may have many Nash equilibrium points. If (22a)–(22d) have no solution, *Problem (LQ NZSDG)* has no Nash equilibrium point. The existence and uniqueness of the Nash equilibrium point for *Problem (LQ NZSDG)* are equivalent to the existence and uniqueness of (15a)–(15d).

Note that, under the two pieces of general asymmetric information $\mathcal{G}_t^1$ and $\mathcal{G}_t^2$, the optimal filtering $\mathbb{E}(y_i(t) \mid \mathcal{G}_t^i)$ $(i = 1, 2)$ is very abstract which leads to the difficulty in finding the filtering equations satisfied by $\mathbb{E}(y_i(t) \mid \mathcal{G}_t^i)$ $(i = 1, 2)$. In the following, we begin to study *Problem (LQ NZSDG)* under several classes of particular asymmetric information. Though the chosen observable information is a bit special, the mathematical deductions are still highly complicated, and the derived results are interesting and meaningful.

3.1. Case 1: $\mathscr{G}_t^1 = \mathscr{F}_t^{1,2}$ *and* $\mathscr{G}_t^2 = \mathscr{F}_t^{2,3}$. In this case, from the notations defined by (1), we have $\mathbb{E}(y_1(t) \mid \mathscr{G}_t^1) = \widehat{y}_1(t)$ and $\mathbb{E}(y_2(t) \mid \mathscr{G}_t^2) = \overline{y}_2(t)$. Hereinafter, we simply call $\widehat{y}_1$ and $\overline{y}_2$ the optimal filters of y_1 and y_2, respectively, if there is no ambiguity from the notations and context. Then Theorem 6 can be rewritten as follows.

Theorem 8. *(u_1, u_2) is a Nash equilibrium point for Problem (LQ NZSDG) if and only if (u_1, u_2) has the following form:*

$$\begin{aligned} u_1(t) &= -m_1^{-1}(t)\, b_1(t)\, \widehat{y}_1(t), \\ u_2(t) &= -m_2^{-1}(t)\, b_2(t)\, \overline{y}_2(t), \end{aligned} \tag{21}$$

where $(x, (y_1, z_{11}, z_{12}, z_{13}), (y_2, z_{21}, z_{22}, z_{23}))$ is a solution to the following FBSDE:

$$\begin{aligned} dx = {} & \left[ax - b_1^2 m_1^{-1} \widehat{y}_1 - b_2^2 m_2^{-1} \overline{y}_2 + c\right] dt \\ & + g_1 dW_1 + [ex + g_2]\, dW_2 + g_3 dW_3, \end{aligned} \tag{22a}$$

$$-dy_1 = [ay_1 + ez_{12} + l_1 x]\, dt - z_{11} dW_1 - z_{12} dW_2 - z_{13} dW_3, \tag{22b}$$

$$-dy_2 = [ay_2 + ez_{22} + l_2 x]\, dt - z_{21} dW_1 - z_{22} dW_2 - z_{23} dW_3, \tag{22c}$$

$$x(0) = x_0, \qquad y_1(T) = r_1 x(T), \qquad y_2(T) = r_2 x(T). \tag{22d}$$

We can see that (22a)–(22d) is a very complicated FBSDE. First, (forward) SDE (22a) is one dimensional and the combination of BSDEs (22b) and (22c) is two dimensional, which is more intricate than the case of forward SDE and BSDE with the same dimension. Second, the drift terms and terminal conditions in (22b) and (22c) contain x. Finally, the drift term in (22a) contains the optimal filter $\widehat{y}_1$ (resp., $\overline{y}_2$) of y_1 (resp., y_2) with respect to $\mathscr{F}_t^{1,2}$ (resp., $\mathscr{F}_t^{2,3}$), whose dynamics has not been known.

Now it is the position to seek the dynamics of $\widehat{y}_1(t)$ and $\overline{y}_2(t)$ which will be used to construct the analytical representation of the Nash equilibrium point. Applying Lemma 5.4 in Xiong [18] and Lemma 3, we obtain the optimal filters of x and y_1 in (22a) and (22b) with respect to $\mathscr{F}_t^{1,2}$ for Player 1 which satisfies

$$\begin{aligned} d\widehat{x} = {} & \left[a\widehat{x} - b_1^2 m_1^{-1} \widehat{y}_1 - b_2^2 m_2^{-1} \widetilde{y}_2 + c\right] dt \\ & + g_1 dW_1 + [e\widehat{x} + g_2]\, dW_2, \end{aligned} \tag{23a}$$

$$-d\widehat{y}_1 = [a\widehat{y}_1 + e\widehat{z}_{12} + l_1 \widehat{x}]\, dt - \widehat{z}_{11} dW_1 - \widehat{z}_{12} dW_2, \tag{23b}$$

$$\widehat{x}(0) = x_0, \qquad \widehat{y}_1(T) = r_1 \widehat{x}(T). \tag{23c}$$

Similarly, we can obtain the optimal filters of x and y_2 in (22a) and (22c) with respect to $\mathscr{F}_t^{2,3}$ for Player 2 which satisfies

$$\begin{aligned} d\overline{x} = {} & \left[a\overline{x} - b_1^2 m_1^{-1} \widetilde{y}_1 - b_2^2 m_2^{-1} \overline{y}_2 + c\right] dt \\ & + [e\overline{x} + g_2]\, dW_2 + g_3 dW_3, \end{aligned} \tag{24a}$$

$$-d\overline{y}_2 = [a\overline{y}_2 + e\overline{z}_{22} + l_2 \overline{x}]\, dt - \overline{z}_{22} dW_2 - \overline{z}_{23} dW_3, \tag{24b}$$

$$\overline{x}(0) = x_0, \qquad \overline{y}_2(T) = r_2 \overline{x}(T). \tag{24c}$$

Note that (23a) and (24a) involve the optimal filter $\widetilde{y}_i$ of y_i with respect to $\mathscr{F}_t^2$; that is, $\widetilde{y}_i(t) = \mathbb{E}(y_i(t) \mid \mathscr{F}_t^2)$ $(i = 1, 2)$. We can derive that $\widetilde{y}_2$ and $\widetilde{y}_1$ together with the optimal filter $\widetilde{x}$ of x satisfy

$$d\widetilde{x} = \left[a\widetilde{x} - b_1^2 m_1^{-1} \widetilde{y}_1 - b_2^2 m_2^{-1} \widetilde{y}_2 + c\right] dt + [e\widetilde{x} + g_2]\, dW_2, \tag{25a}$$

$$-d\widetilde{y}_1 = [a\widetilde{y}_1 + e\widetilde{z}_{12} + l_1 \widetilde{x}]\, dt - \widetilde{z}_{12} dW_2, \tag{25b}$$

$$-d\widetilde{y}_2 = [a\widetilde{y}_2 + e\widetilde{z}_{22} + l_2 \widetilde{x}]\, dt - \widetilde{z}_{22} dW_2, \tag{25c}$$

$$\widetilde{x}(0) = x_0, \qquad \widetilde{y}_1(T) = r_1 \widetilde{x}(T), \qquad \widetilde{y}_2(T) = r_2 \widetilde{x}(T). \tag{25d}$$

Note that (23a)–(25d) are coupled forward-backward stochastic filtering equations. It is remarkable that the filtering equations are essentially different from the classical ones of SDEs, and the main reason is that BSDEs are included in the equations. To our best knowledge, this class of filtering equations is originally found by Huang et al. [19] when they studied the partial information control problems of backward stochastic systems. This class of filtering equations is later also discussed when some authors investigated the optimal control or differential games of partial informatio in BSDEs or FBSDEs (see [20–26]).

We introduce an assumption:

(H_3) $b_1^2(t) m_1^{-1}(t) = b_2^2(t) m_2^{-1}(t)$, $t \in [0, T]$,

which is needed in the following lemmas and theorems.

Lemma 9. *Under the assumption (H_3), (25a)–(25d) have a unique solution $(\widetilde{x}, (\widetilde{y}_1, \widetilde{z}_{12}), (\widetilde{y}_2, \widetilde{z}_{22})) \in \mathscr{L}^2_{\mathscr{F}_t^2}(0, T; R^5)$.*

Proof. We first introduce another FBSDE:

$$\begin{aligned} dn &= \left(an - b_1^2 m_1^{-1} p + c\right) dt + (en + g_2)\, dW_2, \\ -dp &= \left(ap + eq + (l_1 + l_2)\, n\right) dt - q\, dW_2, \\ n(0) &= x_0, \qquad p(T) = (r_1 + r_2)\, n(T). \end{aligned} \tag{26}$$

If $(\widetilde{x}, (\widetilde{y}_1, \widetilde{z}_{12}), (\widetilde{y}_2, \widetilde{z}_{22}))$ is a solution to (25a)–(25d), then (n, p, q) is a solution to (26), where

$$n = \widetilde{x}, \qquad p = \widetilde{y}_1 + \widetilde{y}_2, \qquad q = \widetilde{z}_{12} + \widetilde{z}_{22}. \tag{27}$$

On the other hand, if (n, p, q) is a solution to (26), we introduce the following BSDE:

$$\begin{aligned}
-dp_1 &= [ap_1 + eq_{12} + l_1 n]\,dt - q_{12}dW_2,\\
-dp_2 &= [ap_2 + eq_{22} + l_2 n]\,dt - q_{22}dW_2,\\
p_1(T) &= r_1 n(T), \qquad p_2(T) = r_2 n(T).
\end{aligned}\tag{28}$$

From the existence and uniqueness of BSDE (see [27]), (28) has a unique solution $(p_1, q_{12}, p_2, q_{22})$ with $p_1 + p_2 = p$, $q_{12} + q_{22} = q$. Further, we can check that $(n, (p_1, q_{12}), (p_2, q_{22}))$ is a solution to (25a)–(25d). In other words, the existence and uniqueness of (25a)–(25d) are equivalent to those of (26). It is easy to check that (26) satisfies the assumptions (H_1) and (H_2). From Lemma 1, it has a unique solution (n, p, q). So do (25a)–(25d). □

We observe that (25a)–(25d) are independent of (23a)–(23c) and (24a)–(24c). We can first solve (25a)–(25d) and derive the unique solution $(\widetilde{x}, (\widetilde{y}_1, \widetilde{z}_{12}), (\widetilde{y}_2, \widetilde{z}_{22}))$. Then we plug $\widetilde{y}_2$ (resp., $\widetilde{y}_1$) into (23a)–(23c) (resp., (24a)–(24c)). From Lemma 1, we have the following lemma.

Lemma 10. *If the assumption* (H_3) *holds, there exists a unique solution* $(\widehat{x}, \widehat{y}_1, \widehat{z}_{11}, \widehat{z}_{12}) \in \mathscr{L}^2_{\mathscr{F}_t^{1,2}}(0, T; R^4)$ *(resp.,* $(\overline{x}, \overline{y}_2, \overline{z}_{22}, \overline{z}_{23}) \in \mathscr{L}^2_{\mathscr{F}_t^{2,3}}(0, T; R^4)$*) to* (23a)–(23c) *(resp.,* (24a)–(24c)*).*

After we obtain the unique solutions $\widehat{y}_1$ and $\overline{y}_2$ by solving (23a)–(23c) and (24a)–(24c), respectively, from the existence and uniqueness of solutions of SDEs, we conclude that (22a) has a unique solution x. Further, (22b) and (22c) also have unique solutions $(y_1, z_{11}, z_{12}, z_{13})$ and $(y_2, z_{21}, z_{22}, z_{23})$, respectively. Then we can say that (22a)–(22d) have a unique solution, which implies the following theorem.

Theorem 11. *Under the assumption* (H_3)*, Problem (LQ NZSDG) has a unique Nash equilibrium point denoted by*

$$\begin{aligned}
u_1(t) &= -m_1^{-1}(t)\, b_1(t)\, \widehat{y}_1(t),\\
u_2(t) &= -m_2^{-1}(t)\, b_2(t)\, \overline{y}_2(t),
\end{aligned}\tag{29}$$

where $\widehat{y}_1$ *and* $\overline{y}_2$ *are uniquely determined by the systems of* (23a)–(25d)*.*

In the following, the Riccati equations are introduced, and the Nash equilibrium point is represented in a feedback of the optimal filters $\widehat{x}, \widetilde{x}$, and $\overline{x}$. Hereinafter, we suppose the assumption (H_3) always holds.

Set

$$\widetilde{y}_i = \alpha_i \widetilde{x} + \beta_i \quad (i = 1, 2), \tag{30}$$

where α_i and β_i are undetermined deterministic functions on $[0, T]$ satisfying $\alpha_i(T) = r_i$ and $\beta_i(T) = 0$.

Applying Itô's formula to $\widetilde{y}_1$ in (30), it yields

$$\begin{aligned}
d\widetilde{y}_1 = {}& \left[\left(\dot{\alpha}_1 - b_1^2 m_1^{-1}\alpha_1^2 - b_2^2 m_2^{-1}\alpha_1\alpha_2 + a\alpha_1\right)\widetilde{x}\right.\\
&\left. + \left(\dot{\beta}_1 - b_1^2 m_1^{-1}\alpha_1\beta_1 - b_2^2 m_2^{-1}\alpha_1\beta_2 + \alpha_1 c\right)\right] dt\\
& + \alpha_1\left(e\widetilde{x} + g_2\right) dW_2,
\end{aligned}\tag{31}$$

which implies

$$\widetilde{z}_{12} = \alpha_1\left(e\widetilde{x} + g_2\right). \tag{32}$$

Substituting (30) and (32) into (25b) and comparing the coefficients between (25b) and (31), we have

$$\dot{\alpha}_1 - b_1^2 m_1^{-1}\alpha_1^2 + \left(2a + e^2\right)\alpha_1 - b_2^2 m_2^{-1}\alpha_1\alpha_2 + l_1 = 0, \tag{33a}$$

$$\dot{\beta}_1 + \left(a - b_1^2 m_1^{-1}\alpha_1\right)\beta_1 - b_2^2 m_2^{-1}\alpha_1\beta_2 + \left(c + eg_2\right)\alpha_1 = 0. \tag{33b}$$

Applying Itô's formula to $\widetilde{y}_2$ in (30), it yields

$$\begin{aligned}
d\widetilde{y}_2 = {}& \left[\left(\dot{\alpha}_2 - b_2^2 m_2^{-1}\alpha_2^2 - b_1^2 m_1^{-1}\alpha_1\alpha_2 + a\alpha_2\right)\widetilde{x}\right.\\
&\left. + \left(\dot{\beta}_2 - b_2^2 m_2^{-1}\alpha_2\beta_2 - b_1^2 m_1^{-1}\alpha_2\beta_1 + \alpha_2 c\right)\right] dt\\
& + \alpha_2\left(e\widetilde{x} + g_2\right) dW_2,
\end{aligned}\tag{34}$$

which implies

$$\widetilde{z}_{22} = \alpha_2\left(e\widetilde{x} + g_2\right). \tag{35}$$

Substituting (30) and (35) into (25c) and comparing the coefficients between (25c) and (34), we have

$$\dot{\alpha}_2 - b_2^2 m_2^{-1}\alpha_2^2 + \left(2a + e^2\right)\alpha_2 - b_1^2 m_1^{-1}\alpha_1\alpha_2 + l_2 = 0, \tag{36a}$$

$$\dot{\beta}_2 + \left(a - b_2^2 m_2^{-1}\alpha_2\right)\beta_2 - b_1^2 m_1^{-1}\alpha_2\beta_1 + \left(c + eg_2\right)\alpha_2 = 0. \tag{36b}$$

Let $\alpha = \alpha_1 + \alpha_2$. From (H_3), we have

$$\begin{gathered}
\dot{\alpha} - b_1^2 m_1^{-1}\alpha^2 + \left(2a + e^2\right)\alpha + l_1 + l_2 = 0\\
\text{on } [0, T), \quad \alpha(T) = r_1 + r_2.
\end{gathered}\tag{37}$$

Since (37) is a standard Riccati equation, it has a unique solution $\alpha(\cdot)$. Introduce two auxiliary equations:

$$\begin{gathered}
\dot{\overline{\alpha}}_1 + \left[\left(2a + e^2\right) - b_1^2 m_1^{-1}\alpha\right]\overline{\alpha}_1 + l_1 = 0\\
\text{on } [0, T), \quad \dot{\overline{\alpha}}_1(T) = r_1,
\end{gathered}\tag{38}$$

$$\begin{gathered}
\dot{\overline{\alpha}}_2 + \left[\left(2a + e^2\right) - b_2^2 m_2^{-1}\alpha\right]\overline{\alpha}_2 + l_2 = 0\\
\text{on } [0, T), \quad \dot{\overline{\alpha}}_2(T) = r_2,
\end{gathered}\tag{39}$$

where α is the solution to (37). Obviously, ODEs (38) and (39) have unique solutions $\overline{\alpha}_1$ and $\overline{\alpha}_2$, respectively. In addition, we can check that α_1 and α_2 in (33a) and (36a) are also the solutions to (38) and (39), respectively. From the uniqueness of solutions to (38) and (39), it follows that

$$\overline{\alpha}_1 = \alpha_1, \qquad \overline{\alpha}_2 = \alpha_2, \tag{40}$$

which implies in turn that (33a) and (36a) have the unique solutions to α_1 and α_2.

Let $\beta = \beta_1 + \beta_2$, and then we have

$$\begin{aligned} &\dot{\beta} + \left(a - b_1^2 m_1^{-1}\alpha\right)\beta + \left(c + eg_2\right)\alpha = 0 \\ &\qquad \text{on } [0,T), \quad \beta(T) = 0, \end{aligned} \tag{41}$$

where α is the solution to (37). Note that ODE (41) has a unique solution β. Introduce two another auxiliary equations:

$$\begin{aligned} &\dot{\overline{\beta}}_1 + a\overline{\beta}_1 - b_2^2 m_2^{-1}\alpha_1\beta + \left(c + eg_2\right)\alpha_1 = 0 \\ &\qquad \text{on } [0,T), \quad \overline{\beta}_1(T) = 0, \\ &\dot{\overline{\beta}}_2 + a\overline{\beta}_2 - b_1^2 m_1^{-1}\alpha_2\beta + \left(c + eg_2\right)\alpha_2 = 0 \\ &\qquad \text{on } [0,T), \quad \overline{\beta}_2(T) = 0, \end{aligned} \tag{42}$$

where α_1, α_2, and β are the solutions to (38), (39), and (41), respectively. Similarly, we can prove that (33b) and (36b) also have unique solutions β_1 and β_2 satisfying

$$\overline{\beta}_1 = \beta_1, \qquad \overline{\beta}_2 = \beta_2. \tag{43}$$

Based on the arguments above, we can derive the analytical expressions for α_1, α_2, β_1, β_2, α, and β. Then (25a) can be rewritten as

$$\begin{aligned} d\widetilde{x} &= \left[\left(a - b_1^2 m_1^{-1}\alpha\right)\widetilde{x} - b_1^2 m_1^{-1}\beta + c\right]dt + \left[e\widetilde{x} + g_2\right]dW_2, \\ \widetilde{x}(0) &= x_0, \end{aligned} \tag{44}$$

which has a unique solution

$$\begin{aligned} \widetilde{x}(t) = \Gamma_0^t x_0 + \int_0^t \Gamma_s^t \Big[\big(& c(s) - b_1^2(s)\, m_1^{-1}(s)\, \beta(s) \\ & - e(s)\, g_2(s) \big) ds + g_2(s)\, dW_2(s) \Big], \end{aligned} \tag{45}$$

with $\Gamma_s^t = \exp\{\int_s^t [a(r) - b_1^2(r) m_1^{-1}(r)\alpha(r) - (1/2)e^2(r)]dr + \int_s^t e(r)dW_2(r)\}$.

From the uniqueness of α_1, α_2, β_1, β_2, and $\widetilde{x}$, it follows that $\widetilde{y}_i$ in (30) has a unique analytical expression.

Substituting $\widetilde{y}_2$ in (30) into (23a)–(23c), we have

$$\begin{aligned} d\widehat{x} = \Big[& a\widehat{x} - b_1^2 m_1^{-1}\widehat{y}_1 - b_2^2 m_2^{-1}\alpha_2\widetilde{x} - b_2^2 m_2^{-1}\beta_2 + c \Big] dt \\ & + g_1 dW_1 + \left[e\widehat{x} + g_2\right] dW_2, \end{aligned} \tag{46a}$$

$$-d\widehat{y}_1 = \left[a\widehat{y}_1 + e\widehat{z}_{12} + l_1\widehat{x}\right]dt - \widehat{z}_{11}dW_1 - \widehat{z}_{12}dW_2, \tag{46b}$$

$$\widehat{x}(0) = x_0, \quad \widehat{y}_1(T) = r_1\widehat{x}(T). \tag{46c}$$

Set

$$\widehat{y}_1 = \gamma_1\widehat{x} + \gamma_2\widetilde{x} + \gamma_3 \tag{47}$$

with $\gamma_1(T) = r_1, \gamma_2(T) = \gamma_3(T) = 0$. Applying Itô's formula to $\widehat{y}_1$ in (47), we have

$$\begin{aligned} d\widehat{y}_1 = \Big[& \left(\dot{\gamma}_1 - b_1^2 m_1^{-1}\gamma_1^2 + a\gamma_1\right)\widehat{x} \\ & + \Big(\dot{\gamma}_2 + \left(a - b_1^2 m_1^{-1}\alpha \; - b_1^2 m_1^{-1}\gamma_1\right)\gamma_2 \\ & \quad - b_2^2 m_2^{-1}\alpha_2\gamma_1\Big)\widetilde{x} \\ & + \dot{\gamma}_3 - b_1^2 m_1^{-1}\gamma_1\gamma_3 + \left(c - b_2^2 m_2^{-1}\beta_2\right)\gamma_1 \\ & + \left(c - b_1^2 m_1^{-1}\beta\right)\gamma_2 \Big] dt \\ & + \gamma_1 g_1 dW_1 + \left[\gamma_1\left(e\widehat{x} + g_2\right) + \gamma_2\left(e\widetilde{x} + g_2\right)\right] dW_2, \end{aligned} \tag{48}$$

which implies

$$\widehat{z}_{11} = \gamma_1 g_1, \qquad \widehat{z}_{12} = \gamma_1\left(e\widehat{x} + g_2\right) + \gamma_2\left(e\widetilde{x} + g_2\right). \tag{49}$$

Substituting (47) and (49) into (46b) and comparing the drift and diffusion coefficients with (48), we conclude that

$$\dot{\gamma}_1 - b_1^2 m_1^{-1}\gamma_1^2 + \left(2a + e^2\right)\gamma_1 + l_1 = 0, \tag{50a}$$

$$\dot{\gamma}_2 + \left(2a + e^2 - b_1^2 m_1^{-1}\alpha - b_1^2 m_1^{-1}\gamma_1\right)\gamma_2 - b_2^2 m_2^{-1}\alpha_2\gamma_1 = 0, \tag{50b}$$

$$\begin{aligned} & \dot{\gamma}_3 + \left(a - b_1^2 m_1^{-1}\gamma_1\right)\gamma_3 + \left(c - b_2^2 m_2^{-1}\beta_2 + eg_2\right)\gamma_1 \\ & \quad + \left(c - b_1^2 m_1^{-1}\beta + eg_2\right)\gamma_2 = 0. \end{aligned} \tag{50c}$$

It is clear that there exists a unique solution $(\gamma_1, \gamma_2, \gamma_3)$ to (50a)–(50c). We denote

$$f_1 \triangleq -\left(b_2^2 m_2^{-1}\alpha_2 + b_1^2 m_1^{-1}\gamma_2\right)\widetilde{x} - b_1^2 m_1^{-1}\gamma_3 - b_2^2 m_2^{-1}\beta_2 + c, \tag{51}$$

and then, in terms of (47), (46a) can be rewritten as

$$\begin{aligned} d\widehat{x} &= \left[\left(a - b_1^2 m_1^{-1}\gamma_1\right)\widehat{x} + f_1\right]dt + g_1 dW_1 + \left[e\widehat{x} + g_2\right]dW_2, \\ \widehat{x}(0) &= x_0, \end{aligned} \tag{52}$$

which has a unique solution:

$$\begin{aligned} \widehat{x}(t) = \Upsilon_0^t x_0 + \int_0^t \Upsilon_s^t \big[& \left(f_1(s) - e(s)\, g_2(s)\right) ds \\ & + g_1(s)\, dW_1(s) + g_2(s)\, dW_2(s) \big], \end{aligned} \tag{53}$$

with $\Upsilon_s^t = \exp\{\int_s^t [a(r) - b_1^2(r) m_1^{-1}(r)\gamma_1(r) - (1/2)e^2(r)]dr + \int_s^t e(r)dW_2(r)\}$.

Substituting $\tilde{y}_1$ in (30) into (24a)–(24c), we have

$$\begin{aligned} d\overline{x} = &\left[a\overline{x} - b_1^2 m_1^{-1}\alpha_1 \tilde{x} - b_2^2 m_2^{-1} \overline{y}_2 - b_1^2 m_1^{-1}\beta_1 + c\right] dt \\ &+ \left[e\overline{x} + g_2\right] dW_2 + g_3 dW_3, \end{aligned} \tag{54a}$$

$$-d\overline{y}_2 = \left[a\overline{y}_2 + e\overline{z}_{22} + l_2\overline{x}\right] dt - \overline{z}_{22} dW_2 - \overline{z}_{23} dW_3, \tag{54b}$$

$$\overline{x}(0) = x_0, \qquad \overline{y}_2(T) = r_2 \overline{x}(T). \tag{54c}$$

Set

$$\overline{y}_2 = \tau_1 \overline{x} + \tau_2 \tilde{x} + \tau_3, \tag{55}$$

with $\tau_1(T) = r_2, \tau_2(T) = \tau_3(T) = 0$. In the similar manner, we can deduce that (τ_1, τ_2, τ_3) satisfies

$$\dot{\tau}_1 - b_2^2 m_2^{-1} \tau_1^2 + \left(2a + e^2\right)\tau_1 + l_2 = 0, \tag{56a}$$

$$\dot{\tau}_2 + \left(2a + e^2 - b_1^2 m_1^{-1}\alpha - b_2^2 m_2^{-1}\tau_1\right)\tau_2 - b_1^2 m_1^{-1}\alpha_1 \tau_1 = 0, \tag{56b}$$

$$\begin{aligned} &\dot{\tau}_3 + \left(a - b_2^2 m_2^{-1}\tau_1\right)\tau_3 + \left(c - b_1^2 m_1^{-1}\beta_1 + e g_2\right)\tau_1 \\ &\quad + \left(c - b_1^2 m_1^{-1}\beta + e g_2\right)\tau_2 = 0, \end{aligned} \tag{56c}$$

which has a unique solution (τ_1, τ_2, τ_3). We denote

$$f_2 \triangleq -\left(b_2^2 m_2^{-1}\tau_2 + b_1^2 m_1^{-1}\alpha_1\right)\tilde{x} - b_1^2 m_1^{-1}\beta_1 - b_2^2 m_2^{-1}\tau_3 + c, \tag{57}$$

and then, in terms of (55), (54a) can be rewritten as

$$\begin{aligned} &d\overline{x} = \left[\left(a - b_2^2 m_2^{-1}\tau_1\right)\overline{x} + f_2\right] dt + \left[e\overline{x} + g_2\right] dW_2 + g_3 dW_3, \\ &\overline{x}(0) = x_0, \end{aligned} \tag{58}$$

which has a unique solution

$$\begin{aligned} \overline{x}(t) = \Psi_0^t x_0 + \int_0^t \Psi_s^t &\left[\left(f_2(s) - e(s) g_2(s)\right) ds \right. \\ &\left. + g_2(s)\, dW_2(s) + g_3(s)\, dW_3(s)\right], \end{aligned} \tag{59}$$

with $\Psi_s^t = \exp\{\int_s^t [a(r) - b_2^2(r) m_2^{-1}(r)\tau_1(r) - (1/2)e^2(r)]dr + \int_s^t e(r) dW_2(r)\}$.

Based on the arguments above, we derive the Nash equilibrium point which is represented in the feedback of the optimal filters $\hat{x}, \tilde{x}$, and $\overline{x}$ of the state x. Then Theorem 11 can be rewritten as follows.

Theorem 12. *Under the assumption* (H_3), *Problem (LQ NZSDG) has a unique Nash equilibrium point denoted by*

$$\begin{aligned} u_1(t) &= -m_1^{-1}(t)\, b_1(t) \left(\gamma_1(t)\, \hat{x}(t) + \gamma_2(t)\, \tilde{x}(t) + \gamma_3(t)\right), \\ u_2(t) &= -m_2^{-1}(t)\, b_2(t) \left(\tau_1(t)\, \overline{x}(t) + \tau_2(t)\, \tilde{x}(t) + \tau_3(t)\right), \end{aligned} \tag{60}$$

where $\tilde{x}, \hat{x}$, *and* $\overline{x}$ *are as shown in* (45), (53), *and* (59), *respectively, and* γ_i *and* τ_i $(i = 1, 2, 3)$ *are uniquely determined by the systems of* (50a)–(50c) *and* (56a)–(56c), *respectively.*

Remark 13. We introduce another assumption:

$(\mathrm{H}_3)'$ $b_i^2 m_i^{-1}$ $(i = 1, 2)$ are independent of t.

We can check that when the assumption (H_3) is replaced by $(\mathrm{H}_3)'$, the foregoing lemmas and theorems still hold.

3.2. Case 2: $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ *and* $\mathcal{G}_t^2 = \mathcal{F}_t^2$. In this case, we have $\mathbb{E}(y_1(t) \mid \mathcal{G}_t^1) = \hat{y}_1(t)$ and $\mathbb{E}(y_2(t) \mid \mathcal{G}_t^2) = \tilde{y}_2(t)$. Applying the similar methods shown in Section 3.1, we can obtain the following theorem.

Theorem 14. (u_1, u_2) *is a Nash equilibrium point for Problem (LQ NZSDG) if and only if*

$$\begin{aligned} u_1(t) &= -m_1^{-1}(t)\, b_1(t)\, \hat{y}_1(t), \\ u_2(t) &= -m_2^{-1}(t)\, b_2(t)\, \tilde{y}_2(t), \end{aligned} \tag{61}$$

where $(x, (y_1, z_{11}, z_{12}, z_{13}), (y_2, z_{21}, z_{22}, z_{23}))$ *is a solution of the following FBSDE:*

$$\begin{aligned} dx &= \left[ax - b_1^2 m_1^{-1}\hat{y}_1 - b_2^2 m_2^{-1}\tilde{y}_2 + c\right] dt \\ &\quad + g_1 dW_1 + \left[ex + g_2\right] dW_2 + g_3 dW_3, \\ -dy_1 &= \left[a y_1 + e z_{12} + l_1 x\right] dt \\ &\quad - z_{11} dW_1 - z_{12} dW_2 - z_{13} dW_3, \\ -dy_2 &= \left[a y_2 + e z_{22} + l_2 x\right] dt \\ &\quad - z_{21} dW_1 - z_{22} dW_2 - z_{23} dW_3, \\ x(0) &= x_0, \qquad y_1(T) = r_1 x(T), \qquad y_2(T) = r_2 x(T). \end{aligned} \tag{62}$$

Under the assumption (H_3), *we can check that the filtering equations* (23a)–(23c), (25a)–(25d), *and the linear relations* (30) *and* (47) *still hold, and the systems of equations* (33a), (33b), (36a), (36b), *and* (50a)–(50c) *are still uniquely solvable. Then we have the following theorem.*

Theorem 15. *If* (H_3) *holds, then Problem (LQ NZSDG) has a unique Nash equilibrium point denoted by*

$$\begin{aligned} u_1(t) &= -\, m_1^{-1}(t)\, b_1(t) \left(\gamma_1(t)\, \hat{x}(t) + \gamma_2(t)\, \tilde{x}(t) + \gamma_3(t)\right), \\ u_2(t) &= -\, m_2^{-1}(t)\, b_2(t) \left(\alpha_2(t)\, \tilde{x}(t) + \beta_2(t)\right), \end{aligned} \tag{63}$$

where $\tilde{x}$ *and* $\hat{x}$ *are shown in* (45) *and* (53), *respectively.*

Remark 16. In the cases similar to Case 2, such as $\mathcal{G}_t^1 = \mathcal{F}_t^{2,3}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$, $\mathcal{G}_t^1 = \mathcal{F}_t^{1,3}$, and $\mathcal{G}_t^2 = \mathcal{F}_t^1$, the corresponding results can be easily derived.

3.3. Case 3: $\mathcal{G}_t^1 = \mathcal{F}_t$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$.

In this case, we have $\mathbb{E}(y_1(t) \mid \mathcal{G}_t^1) = y_1(t)$ and $\mathbb{E}(y_2(t) \mid \mathcal{G}_t^2) = \tilde{y}_2(t)$. Then we have the following theorem.

Theorem 17. *(u_1, u_2) is a Nash equilibrium point for Problem (LQ NZSDG) if and only if*

$$u_1(t) = -m_1^{-1}(t) b_1(t) y_1(t),$$
$$u_2(t) = -m_2^{-1}(t) b_2(t) \tilde{y}_2(t), \tag{64}$$

where $(x, (y_1, z_{11}, z_{12}, z_{13}), (y_2, z_{21}, z_{22}, z_{23}))$ is a solution to the following FBSDE:

$$dx = \left[ax - b_1^2 m_1^{-1} y_1 - b_2^2 m_2^{-1} \tilde{y}_2 + c\right] dt + g_1 dW_1 + \left[ex + g_2\right] dW_2 + g_3 dW_3, \tag{65a}$$

$$-dy_1 = \left[ay_1 + ez_{12} + l_1 x\right] dt - z_{11} dW_1 - z_{12} dW_2 - z_{13} dW_3, \tag{65b}$$

$$-dy_2 = \left[ay_2 + ez_{22} + l_2 x\right] dt - z_{21} dW_1 - z_{22} dW_2 - z_{23} dW_3, \tag{65c}$$

$$x(0) = x_0, \qquad y_1(T) = r_1 x(T), \qquad y_2(T) = r_2 x(T). \tag{65d}$$

Under the assumption (H_3), we can check that $(\tilde{x}, (\tilde{y}_1, \tilde{z}_{12}), (\tilde{y}_2, \tilde{z}_{22}))$ still satisfies the filtering equations (25a)–(25d). *From* Section 3.1, *we know that $\tilde{x}$ is shown as* (45) *and $\tilde{y}_i$ is uniquely represented by* (30). *Then* (65a) *can be rewritten as*

$$dx = \left[ax - b_1^2 m_1^{-1} y_1 - b_2^2 m_2^{-1} \alpha_2 \tilde{x} - b_2^2 m_2^{-1} \beta_2 + c\right] dt + g_1 dW_1 + \left[ex + g_2\right] dW_2 + g_3 dW_3, \tag{66a}$$

$$-dy_1 = \left[ay_1 + ez_{12} + l_1 x\right] dt - z_{11} dW_1 - z_{12} dW_2 - z_{13} dW_3, \tag{66b}$$

$$x(0) = x_0, \qquad y_1(T) = r_1 x(T). \tag{66c}$$

From Lemma 1, *we can say that* (66a)–(66c) *has a unique solution $(x, y_1, z_{11}, z_{12}, z_{13})$. Further, the relation between y_1 and $(x, \tilde{x})$ is as follows:*

$$y_1 = \gamma_1 x + \gamma_2 \tilde{x} + \gamma_3, \tag{67}$$

where γ_i $(i = 1, 2, 3)$ is the solution to (50a)–(50c), *and*

$$x(t) = \Upsilon_0^t x_0 + \int_0^t \Upsilon_s^t \left[\left(f_1(s) - e(s) g_2(s)\right) ds + \sum_{i=1}^{3} g_i(s) dW_i(s) \right], \tag{68}$$

with $\Upsilon_s^t = \exp\{\int_s^t [a(r) - b_1^2(r) m_1^{-1}(r) \gamma_1(r) - (1/2) e^2(r)] dr + \int_s^t e(r) dW_2(r)\}$ and f_1 defined by (51). *Then we have the following theorem.*

Theorem 18. *Under the assumptions (H_3), Problem (LQ NZSDG) has a unique Nash equilibrium point denoted by*

$$u_1(t) = -m_1^{-1}(t) b_1(t) \left(\gamma_1(t) x(t) + \gamma_2(t) \tilde{x}(t) + \gamma_3(t)\right),$$
$$u_2(t) = -m_2^{-1}(t) b_2(t) \left(\alpha_2(t) \tilde{x}(t) + \beta_2(t)\right), \tag{69}$$

where $\tilde{x}$ and x are shown as (45) *and* (68), *respectively.*

Remark 19. In the cases similar to Case 3, such as $\mathcal{G}_t^1 = \mathcal{F}_t$ and $\mathcal{G}_t^2 = \mathcal{F}_t^1$, $\mathcal{G}_t^1 = \mathcal{F}_t$ and $\mathcal{G}_t^2 = \mathcal{F}_t^3$, the corresponding results can be easily derived.

3.4. Case 4: $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^3$.

In this case, we have $\mathbb{E}(y_1(t) \mid \mathcal{G}_t^1) = \hat{y}_1(t)$ and $\mathbb{E}(y_2(t) \mid \mathcal{G}_t^2) = \check{y}_2(t)$. Throughout this subsection, we make an additional assumption on (2):

(H_4) $e(t) = 0, t \in [0, T]$.

Similar to Sections 3.2 and 3.3, we directly present the following theorem.

Theorem 20. *(u_1, u_2) is a Nash equilibrium point for Problem (LQ NZSDG) if and only if*

$$u_1(t) = -m_1^{-1}(t) b_1(t) \hat{y}_1(t),$$
$$u_2(t) = -m_2^{-1}(t) b_2(t) \check{y}_2(t), \tag{70}$$

where $(x, (y_1, z_{11}, z_{12}, z_{13}), (y_2, z_{21}, z_{22}, z_{23}))$ is a solution to the following FBSDE:

$$dx = \left[ax - b_1^2 m_1^{-1} \hat{y}_1 - b_2^2 m_2^{-1} \check{y}_2 + c\right] dt + g_1 dW_1 + g_2 dW_2 + g_3 dW_3, \tag{71a}$$

$$-dy_1 = \left[ay_1 + l_1 x\right] dt - z_{11} dW_1 - z_{12} dW_2 - z_{13} dW_3, \tag{71b}$$

$$-dy_2 = \left[ay_2 + l_2 x\right] dt - z_{21} dW_1 - z_{22} dW_2 - z_{23} dW_3, \tag{71c}$$

$$x(0) = x_0, \qquad y_1(T) = r_1 x(T), \qquad y_2(T) = r_2 x(T). \tag{71d}$$

Using the similar method shown in Section 3.1, *we obtain the optimal filters of x and y_1 in* (71a) *and* (71b) *with respect to $\mathcal{F}_t^{1,2}$ which satisfies*

$$d\hat{x} = \left[a\hat{x} - b_1^2 m_1^{-1} \hat{y}_1 - b_2^2 m_2^{-1} \mathbb{E} y_2 + c\right] dt + g_1 dW_1 + g_2 dW_2, \tag{72a}$$

$$-d\hat{y}_1 = \left[a\hat{y}_1 + l_1 \hat{x}\right] dt - \hat{z}_{11} dW_1 - \hat{z}_{12} dW_2, \tag{72b}$$

$$\hat{x}(0) = x_0, \qquad \hat{y}_1(T) = r_1 \hat{x}(T). \tag{72c}$$

Here we denote by $\mathbb{E}\eta$ the mathematical expectation $\mathbb{E}(\eta(t))$ of the variable $\eta(t)$ and omit t for simplicity. Similarly, we can *obtain the optimal filters of x and y_2 in* (71a) *and* (71c) *with respect to $\mathcal{F}_t^3$ which satisfy*

$$d\check{x} = \left[a\check{x} - b_1^2 m_1^{-1}\mathbb{E}y_1 - b_2^2 m_2^{-1}\check{y}_2 + c\right]dt + g_3 dW_3, \quad (73a)$$

$$-d\check{y}_2 = \left[a\check{y}_2 + l_2\check{x}\right]dt - \check{z}_{23}dW_3, \quad (73b)$$

$$\check{x}(0) = x_0, \qquad \check{y}_2(T) = r_2\check{x}(T). \quad (73c)$$

In addition, $\mathbb{E}y_1$ and $\mathbb{E}y_2$ together with $\mathbb{E}x$ satisfy

$$\dot{\mathbb{E}}x = a\mathbb{E}x - b_1^2 m_1^{-1}\mathbb{E}y_1 - b_2^2 m_2^{-1}\mathbb{E}y_2 + c, \quad (74a)$$

$$-\dot{\mathbb{E}}y_1 = a\mathbb{E}y_1 + l_1\mathbb{E}x, \quad (74b)$$

$$-\dot{\mathbb{E}}y_2 = a\mathbb{E}y_2 + l_2\mathbb{E}x, \quad (74c)$$

$$\begin{gathered}\mathbb{E}x(0) = x_0, \qquad \mathbb{E}y_1(T) = r_1\mathbb{E}x(T), \\ \mathbb{E}y_2(T) = r_2\mathbb{E}x(T),\end{gathered} \quad (74d)$$

where $\dot{\mathbb{E}}\eta$ denotes $d\mathbb{E}(\eta(t))/dt$ for $\eta = x, y_1, y_2$.

It is clear that (74a)–(74d) are a forward-backward ODE independent of (72a)–(73c). Using the similar method shown in Lemma 9 and Remark 2, we conclude that (74a)– (74d) have a unique solution ($\mathbb{E}x$, $\mathbb{E}y_1$, and $\mathbb{E}y_2$). Plugging the solutions $\mathbb{E}y_2$ and $\mathbb{E}y_1$ into (72a)–(72c) and (73a)–(73c), respectively, and applying Lemma 1, we conclude that (72a)–(73c) have the unique solutions $(\hat{x}, \hat{y}_1, \hat{z}_{11}, \hat{z}_{12})$ and $(\check{x}, \check{y}_2, \check{z}_{23})$, respectively. Then we derive the more explicit representation of the Nash equilibrium point in (70) as follows.

Theorem 21. *Under the assumptions* (H_3) *and* (H_4)*, Problem (LQ NZSDG) has a unique Nash equilibrium point denoted by*

$$\begin{aligned} u_1(t) &= -m_1^{-1}(t)\,b_1(t)\,\hat{y}_1(t), \\ u_2(t) &= -m_2^{-1}(t)\,b_2(t)\,\check{y}_2(t), \end{aligned} \quad (75)$$

where $\hat{y}_1$ and $\check{y}_2$ are uniquely determined by the systems of (72a)–(74d).

In the sequel, we only present the results and omit the deduction procedures, because the method and technique are parallel to those in Section 3.1.

The relation between $\mathbb{E}y_i$ and $\mathbb{E}x$ is as follows:

$$\mathbb{E}y_i = \alpha_i\mathbb{E}x + \beta_i \quad (i = 1, 2), \quad (76)$$

where α_i, β_i, α, and β are the unique solutions to the systems of (33a), (33b), (36a), (36b), (37), and (41) with $e(\cdot)$ replaced by 0, respectively, and

$$\mathbb{E}x(t) = \Gamma_0^t x_0 + \int_0^t \Gamma_s^t\left[\left(c(s) - b_1^2(s)\,m_1^{-1}(s)\,\beta(s)\right)\right]ds \quad (77)$$

with $\Gamma_s^t = \exp\{\int_s^t [a(r) - b_1^2(r)m_1^{-1}(r)\alpha(r)]dr\}$.

The relation between $\hat{y}_1$ and $(\hat{x}, \mathbb{E}x)$ is as follows:

$$\hat{y}_1 = \gamma_1\hat{x} + \gamma_2\mathbb{E}x + \gamma_3, \quad (78)$$

where γ_i $(i = 1, 2, 3)$ is the solution to (50a)–(50c) with $e(\cdot)$ replaced by 0, and

$$\begin{aligned}\hat{x}(t) = \Upsilon_0^t x_0 + \int_0^t \Upsilon_s^t\big[f_1(s)\,ds + g_1(s)\,dW_1(s) \\ + g_2(s)\,dW_2(s)\big],\end{aligned} \quad (79)$$

with

$$\Upsilon_s^t = \exp\left\{\int_s^t \left[a(r) - b_1^2(r)\,m_1^{-1}(r)\,\gamma_1(r)\right]dr\right\},$$

$$f_1 = -\left(b_2^2 m_2^{-1}\alpha_2 + b_1^2 m_1^{-1}\gamma_2\right)\mathbb{E}x - b_1^2 m_1^{-1}\gamma_3 - b_2^2 m_2^{-1}\beta_2 + c. \quad (80)$$

The relation between $\check{y}_2$ and $(\check{x}, \mathbb{E}x)$ is as follows:

$$\overline{y}_2 = \tau_1\check{x} + \tau_2\mathbb{E}x + \tau_3, \quad (81)$$

where τ_i $(i = 1, 2, 3)$ is the unique solution to (56a)–(56c) with $e(\cdot)$ replaced by 0, and

$$\check{x}(t) = \Psi_0^t x_0 + \int_0^t \Psi_s^t\left[f_2(s)\,ds + g_3(s)\,dW_3(s)\right], \quad (82)$$

with

$$\Psi_s^t = \exp\left\{\int_s^t \left[a(r) - b_2^2(r)\,m_2^{-1}(r)\,\tau_1(r)\right]dr\right\},$$

$$f_2 = -\left(b_2^2 m_2^{-1}\tau_2 + b_1^2 m_1^{-1}\alpha_1\right)\mathbb{E}x - b_1^2 m_1^{-1}\beta_1 - b_2^2 m_2^{-1}\tau_3 + c. \quad (83)$$

Then Theorem 21 can be rewritten as follows.

Theorem 22. *Under the assumption* (H_3) *and* (H_4)*, Problem (LQ NZSDG) has a unique Nash equilibrium point denoted by*

$$\begin{aligned} u_1(t) &= -m_1^{-1}(t)\,b_1(t)\left(\gamma_1(t)\,\hat{x}(t) + \gamma_2(t)\,\mathbb{E}x(t) + \gamma_3(t)\right), \\ u_2(t) &= -m_2^{-1}(t)\,b_2(t)\left(\tau_1(t)\,\check{x}(t) + \tau_2(t)\,\mathbb{E}x(t) + \tau_3(t)\right), \end{aligned} \quad (84)$$

where $\mathbb{E}x$, $\hat{x}$, and $\check{x}$ are shown in (77), (79), *and* (82)*, respectively, and γ_i and τ_i $(i = 1, 2, 3)$ are uniquely determined by the systems of* (50a)–(50c) *and* (56a)–(56c) *with $e(\cdot)$ replaced by 0, respectively.*

Remark 23. In the cases similar to Case 4, such as $\mathcal{G}_t^1 = \mathcal{F}_t^{1,3}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$, $\mathcal{G}_t^1 = \mathcal{F}_t^{2,3}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^1$, the corresponding results can be easily derived.

4. Conclusion Remark

In this paper, we investigate LQ nonzero sum differential game problem where the information available to players is asymmetric. We discuss the game problem under the four classes of cases: (i) $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^{2,3}$; (ii) $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$; (iii) $\mathcal{G}_t^1 = \mathcal{F}_t$ and $\mathcal{G}_t^2 = \mathcal{F}_t^2$; (iv) $\mathcal{G}_t^1 = \mathcal{F}_t^{1,2}$ and $\mathcal{G}_t^2 = \mathcal{F}_t^3$. Some forward-backward stochastic filtering equations with respect to the asymmetric information $\mathscr{G}_t^1$ and $\mathscr{G}_t^2$ are introduced and the existence and uniqueness of the solutions are proved. Finally, the corresponding unique Nash equilibrium point is represented in a feedback form of the optimal filtering of the state, through the solutions of some Riccati equations.

Conflict of Interests

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

Acknolwedgments

Hua Xiao acknowledges the financial support from the National Nature Science Foundation of China (11201263, 11371228), the Program for New Century Excellent Talents in University of China (NCET-12-0338), and the Postdoctor Foundation of China (2013M540540) and the Nature Science Foundation of Shandong Province (ZR2012AQ004, BS2011SF010). Dejian Chang acknowledges the financial support from the National Nature Science Foundation of China (10921101, 61174092) and the National Nature Science Foundation for Distinguished Young Scholars of China (11125102).

References

[1] G. Gennotte, "Optimal portfolio choice under incomplete information," *Journal of Finance*, vol. 41, pp. 733–746, 1986.

[2] J. Xiong and X. Y. Zhou, "Mean-variance portfolio selection under partial information," *SIAM Journal on Control and Optimization*, vol. 46, no. 1, pp. 156–175, 2007.

[3] F. Biagini and B. Øksendal, "Minimal variance hedging for insider trading," *International Journal of Theoretical and Applied Finance*, vol. 9, no. 8, pp. 1351–1375, 2006.

[4] L. Campi, "Some results on quadratic hedging with insider trading," *Stochastics*, vol. 77, no. 4, pp. 327–348, 2005.

[5] H. M. Müller, "The first-best sharing rule in the continuous-time principal-agent problem with exponential utility," *Journal of Economic Theory*, vol. 79, no. 2, pp. 276–280, 1998.

[6] H. M. Müller, "Asymptotic efficiency in dynamic principal-agent problems," *Journal of Economic Theory*, vol. 91, no. 2, pp. 292–301, 2000.

[7] S. Hamadène, "Nonzero sum linear-quadratic stochastic differential games and backward-forward equations," *Stochastic Analysis and Applications*, vol. 17, no. 1, pp. 117–130, 1999.

[8] Q. Meng, "General linear quadratic optimal stochastic contr ol problem driven by a Brownian motion and a poisson random martingale measure with random coefficients," *Stochastic Analysis and Applications*, vol. 32, no. 1, pp. 88–109, 2014.

[9] L. Mou and J. Yong, "Two-person zero-sum linear quadratic stochastic differential games by a Hilbert space method," *Journal of Industrial and Management Optimization*, vol. 2, no. 1, pp. 95–117, 2006.

[10] S. Tang, "General linear quadratic optimal stochastic control problems with random coefficients: linear stochastic hamilton systems and backward stochastic riccati equations," *SIAM Journal on Control and Optimization*, vol. 42, no. 1, pp. 53–75, 2003.

[11] Z. Wu and Z.-Y. Yu, "Linear quadratic nonzero-sum differential games with random jumps," *Applied Mathematics and Mechanics*, vol. 26, no. 8, pp. 1034–1039, 2005.

[12] J. Yong, "A leader-follower stochastic linear quadratic differential game," *SIAM Journal on Control and Optimization*, vol. 41, no. 4, pp. 1015–1041, 2002.

[13] Z. Yu, "Linear-quadratic optimal control and nonzero-sum differential game of forward-backward stochastic system," *Asian Journal of Control*, vol. 14, no. 1, pp. 173–185, 2012.

[14] Y. Hu and B. Øksendal, "Partial information linear quadratic control for jump diffusions," *SIAM Journal on Control and Optimization*, vol. 47, no. 4, pp. 1744–1761, 2008.

[15] K. L. Chung, *A Course in Probability Theory*, Academic Press, San Diego, Calif, USA, 3rd edition, 2001.

[16] S. Tang, "The maximum principle for partially observed optimal control of stochastic differential equations," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1596–1617, 1998.

[17] F. Baghery and B. Øksendal, "A maximum principle for stochastic control with partial information," *Stochastic Analysis and Applications*, vol. 25, no. 3, pp. 705–717, 2007.

[18] J. Xiong, *An Introduction to Stochastic Filtering Theory*, Oxford University Press, Oxford, UK, 2008.

[19] J. Huang, G. Wang, and J. Xiong, "A maximum principle for partial information backward stochastic control problems with applications," *SIAM Journal on Control and Optimization*, vol. 48, no. 4, pp. 2106–2117, 2009.

[20] E. Hui and H. Xiao, "Differential games of partial information forwardbackward doubly stochastic differential equations and applications," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 20, no. 1, pp. 78–94, 2014.

[21] G. Wang, Z. Wu, and J. Xiong, "Maximum principles for forward-backward stochastic control systems with correlated state and observation noises," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 491–524, 2013.

[22] G. Wang and Z. Yu, "A partial information non-zero sum differential game of backward stochastic differential equations with applications," *Automatica A: Journal of IFAC*, vol. 48, no. 2, pp. 342–352, 2012.

[23] G. Wang, C. Zhang, and W. Zhang, "Stochastic maximum principle for mean-field type optimal control under partial information," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 522–528, 2014.

[24] H. Xiao, "The maximum principle for partially observed optimal control of forward-backward stochastic systems with random jumps," *Journal of Systems Science and Complexity*, vol. 24, no. 6, pp. 1083–1099, 2011.

[25] H. Xiao and G. Wang, "The filtering equations of forward-backward stochastic systems with random jumps and applications to partial information stochastic optimal control," *Stochastic Analysis and Applications*, vol. 28, no. 6, pp. 1003–1019, 2010.

[26] H. Xiao and G. Wang, "A necessary condition for optimal control of initial coupled forward-backward stochastic differential equations with partial information," *Journal of Applied Mathematics and Computing*, vol. 37, no. 1-2, pp. 347–359, 2011.

[27] E. Pardoux and S. G. Peng, "Adapted solution of a backward stochastic differential equation," *Systems & Control Letters*, vol. 14, no. 1, pp. 55–61, 1990.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 208612, 8 pages
http://dx.doi.org/10.1155/2014/208612

Research Article

Stochastic Signal Processing for Sound Environment System with Decibel Evaluation and Energy Observation

Akira Ikuta and Hisako Orimoto

Department of Management Information Systems, Prefectural University of Hiroshima, Hiroshima 734-8558, Japan

Correspondence should be addressed to Akira Ikuta; ikuta@pu-hiroshima.ac.jp

Received 21 April 2014; Accepted 3 June 2014; Published 24 July 2014

Academic Editor: Suiyang Khoo

In real sound environment system, a specific signal shows various types of probability distribution, and the observation data are usually contaminated by external noise (e.g., background noise) of non-Gaussian distribution type. Furthermore, there potentially exist various nonlinear correlations in addition to the linear correlation between input and output time series. Consequently, often the system input and output relationship in the real phenomenon cannot be represented by a simple model using only the linear correlation and lower order statistics. In this study, complex sound environment systems difficult to analyze by using usual structural method are considered. By introducing an estimation method of the system parameters reflecting correlation information for conditional probability distribution under existence of the external noise, a prediction method of output response probability for sound environment systems is theoretically proposed in a suitable form for the additive property of energy variable and the evaluation in decibel scale. The effectiveness of the proposed stochastic signal processing method is experimentally confirmed by applying it to the observed data in sound environment systems.

1. Introduction

A specific signal in real sound environment system usually exhibits multifarious and complex characteristics such as non-Gaussian distribution and nonlinear property relating to natural, social, or human factors. Furthermore, the observation data usually are contaminated by external noise (e.g., background noise) with complex statistical properties. In this situation, in order to evaluate the sound environment system, precise estimation of the system characteristics of the sound environment is required by considering the contaminated observation data.

Furthermore, the internal physical mechanism of the real sound environment system is often difficult to recognize analytically, and it contains unknown structural characteristics. In our previous study, it was found that complex sound environment systems are difficult to analyze by using usual structural methods based on the physical mechanism [1]. Therefore, a nonlinear system model was derived in the expansion series form reflecting various types of correlation information from the lower order to the higher order between input and output variables [2]. The conditional probability density function contains the linear and nonlinear correlations in the expansion coefficients and these correlations play an important role as the statistical information for the input and output relationship of sound environment system.

On the other hand, in considering the relationship between the evaluation from top-down viewpoint and the countermeasure from bottom-up viewpoint in the sound environment system, noise evaluation quantities in decibel scale like L_x ($(100 - x)$ percentile level) and L_{Aeq} (averaged energy on decibel scale) and some countermeasure methods in energy scale are widely used. Since there is a certain scale transform between decibel and energy variables, a unified general consideration without losing their mutual relationship has to be derived.

In this study, a general type of complex sound environment systems is considered. A stochastic signal processing method for predicting the output response probability distribution in decibel scale based on the input observations is proposed for complex sound environment systems.

More specifically, an expansion expression of the conditional probability distribution in decibel scale is adopted as the system characteristics. Next, a method to estimate the system parameters reflecting several orders of correlation information between the input and output variables is derived by considering the additive property of energy variables under existence of external noise. Furthermore, a prediction method for the output probability distribution in decibel scale is also considered.

The effectiveness of the proposed theory is confirmed experimentally by applying it to real data of a sound insulation system and the road traffic noise environment measured around a national road in Hiroshima city.

2. Evaluation of Sound Environment System under Existence of External Noise

2.1. Statistical Model for Sound Environment System. Let X and Y be the sound pressure levels of input and output signals for a complex sound environment system. The probability distribution of output Y has to be predicted on the basis of the observed data of the input level X, because noise evaluation quantities connected with probability distribution are widely used. All the information on linear and/or nonlinear correlations between X and Y is included in the conditional probability density function $P(Y \mid X)$ [2].

In order to find explicitly the various correlation properties between X and Y, let us expand the joint probability density function $P(X, Y)$ into an orthogonal polynomial series [3], as follows:

$$P(X,Y) = P_0(X) P_0(Y) \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} A_{rs} \varphi_r^{(1)}(X) \varphi_s^{(2)}(Y),$$
$$A_{rs} \equiv \left\langle \varphi_r^{(1)}(X) \varphi_s^{(2)}(Y) \right\rangle, \tag{1}$$

where $\langle\rangle$ denotes the averaging operation. $P_0(X)$ and $P_0(Y)$ can be chosen arbitrarily as the probability density functions describing the dominant parts of the actual fluctuation pattern. Two functions $\varphi_r^{(1)}(X)$ and $\varphi_s^{(2)}(Y)$ are orthogonal polynomials with the weighting functions $P_0(X)$ and $P_0(Y)$. The information on the various types of linear and/or nonlinear correlations between X and Y is reflected hierarchically in each expansion coefficient A_{rs}. In this section, the Gaussian distribution suitable for the random variables in decibel scale is adopted as $P_0(X)$ and $P_0(Y)$

$$P_0(X) = N\left(X; \mu_X, \sigma_X^2\right),$$
$$P_0(Y) = N\left(Y; \mu_Y, \sigma_Y^2\right) \tag{2}$$

with

$$N(x; \mu, \sigma) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2},$$
$$\mu_X \equiv \langle X \rangle, \qquad \sigma_X^2 \equiv \left\langle (X - \mu_X)^2 \right\rangle, \tag{3}$$
$$\mu_Y \equiv \langle Y \rangle, \qquad \sigma_Y^2 \equiv \left\langle (Y - \mu_Y)^2 \right\rangle.$$

Thus, orthogonal polynomials $\varphi_r^{(1)}(X)$ and $\varphi_s^{(2)}(Y)$ are given by the Hermite polynomial [3]:

$$\varphi_r^{(1)}(X) = \frac{1}{\sqrt{r!}} H_r\left(\frac{X - \mu_X}{\sigma_X}\right),$$
$$\varphi_s^{(2)}(Y) = \frac{1}{\sqrt{s!}} H_s\left(\frac{Y - \mu_Y}{\sigma_Y}\right). \tag{4}$$

Substituting (1) into the definition of the conditional probability, $P(Y \mid X)$ can be expressed in an expansion series form as follows:

$$P(Y \mid X) = \frac{P(X,Y)}{P(X)}$$
$$= \frac{P_0(Y) \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} A_{rs} \varphi_r^{(1)}(X) \varphi_s^{(2)}(Y)}{\sum_{r=0}^{\infty} A_{r0} \varphi_r^{(1)}(X)}. \tag{5}$$

2.2. Estimation of Correlation Information Based on Energy Observation. In the measurement of the sound environment, the effects by external noise (e.g., background noise) are inevitable. Then, based on the additive property of energy variable, the observed sound intensity z_k at a discrete time k is expressed as

$$z_k = y_k + v_k, \tag{6}$$

where y_k and v_k are sound intensities of the output signal for the sound environment system and external noise. We assume that the statistics of the external noise are known. In this section, an estimation method for the expansion coefficients A_{rs} in (1), reflecting the correlation information between X and Y, is derived on the basis of the observed data z_k. There are relationships between energy variables x_k, y_k and decibel variables X_k, Y_k for the input and output signals, as

$$X_k = 10\log_{10}\frac{x_k}{E_0}\ [\text{dB}], \qquad Y_k = 10\log_{10}\frac{y_k}{E_0}\ [\text{dB}],$$
$$E_0 = 10^{-12}\ \left[\text{W/m}^2\right]. \tag{7}$$

Therefore, from (7), the following relationships can be obtained

$$x_k = e^{(X_k - K)/C}, \qquad y_k = e^{(Y_k - K)/C} \tag{8}$$

with

$$C \equiv \frac{10}{\ln 10}, \qquad K \equiv -C \ln E_0. \tag{9}$$

Next, considering the expansion coefficients A_{rs} as unknown parameter vector $\mathbf{a}$,

$$\mathbf{a} \equiv (a_1, a_2, a_3, \ldots) \equiv (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots),$$
$$\mathbf{a}_{(s)} \equiv (A_{0s}, A_{1s}, A_{2s} \ldots), \quad (s = 1, 2, \ldots), \tag{10}$$

the simple dynamical model,

$$\mathbf{a}_{k+1} = \mathbf{a}_k,$$
$$\left(\mathbf{a}_k \equiv (a_{1,k}, a_{2,k}, a_{3,k}, \ldots) \equiv (\mathbf{a}_{(1),k}, \mathbf{a}_{(2),k}, \ldots)\right), \tag{11}$$

is naturally introduced for the successive estimation of the parameter.

In order to derive the estimation algorithm of the parameter, attention is focused on Bayes' theorem for the conditional probability distribution:

$$P\left(\mathbf{a}_k \mid Z_k\right) = \frac{P\left(\mathbf{a}_k, z_k \mid Z_{k-1}\right)}{P\left(z_k \mid Z_{k-1}\right)}, \tag{12}$$

where $Z_k \equiv \{z_1, z_2, \ldots, z_k\}$ is a set of observation data up to time k. Based on (12), using the similar calculation process to the previously reported paper [4], the estimate for an arbitrary polynomial function $f_{\mathbf{M}}(\mathbf{a}_k)$ of $\mathbf{a}_k$ with $\mathbf{M}$th order can be derived as follows (cf. Appendix):

$$\widehat{f}_{\mathbf{M}}\left(\mathbf{a}_k\right) \equiv \left\langle f_{\mathbf{M}}\left(\mathbf{a}_k\right) \mid Z_k\right\rangle = \frac{\sum_{\mathbf{m}=\mathbf{0}}^{\mathbf{M}} \sum_{n=0}^{\infty} B_{\mathbf{m}n} C_{\mathbf{M}\mathbf{m}} \theta_n^{(2)}\left(z_k\right)}{\sum_{n=0}^{\infty} B_{\mathbf{0}n} \theta_n^{(2)}\left(z_k\right)}, \tag{13}$$

with

$$B_{\mathbf{m}n} \equiv \left\langle \theta_{\mathbf{m}}^{(1)}\left(\mathbf{a}_k\right) \theta_n^{(2)}\left(z_k\right) \mid Z_{k-1}\right\rangle, \quad \left(\mathbf{m} \equiv \left(m_1, m_2, \ldots\right)\right). \tag{14}$$

Two functions $\theta_{\mathbf{m}}^{(1)}(\mathbf{a}_k)$ and $\theta_n^{(2)}(z_k)$ are orthonormal polynomials with the weighting functions $P_0(\mathbf{a}_k \mid Z_{k-1})$ and $P_0(z_k \mid Z_{k-1})$. Furthermore, $C_{\mathbf{Mm}}$ is the coefficient when the function $f_{\mathbf{M}}(\mathbf{a}_k)$ is expanded as

$$f_{\mathbf{M}}\left(\mathbf{a}_k\right) = \sum_{\mathbf{m}=\mathbf{0}}^{\mathbf{M}} C_{\mathbf{M}\mathbf{m}} \theta_{\mathbf{m}}^{(1)}\left(\mathbf{a}_k\right). \tag{15}$$

As the concrete expression on the fundamental probability function for the parameter $\mathbf{a}_k$ fluctuating in both positive and negative range, a standard Gaussian distribution is adopted. Furthermore, a gamma distribution is adopted as the probability function for the sound intensity z_k:

$$\begin{aligned} P_0\left(\mathbf{a}_k \mid Z_{k-1}\right) &= \prod_i N\left(a_{i,k}; a_{i,k}^*, \Gamma_{i,k}\right), \\ P_0\left(z_k \mid Z_{k-1}\right) &= P_\Gamma\left(z_k; m_k^*, s_k^*\right), \end{aligned} \tag{16}$$

with

$$\begin{gathered} P_\Gamma(x; m, s) \equiv \frac{x^{m-1}}{\Gamma(m) s^m} e^{-x/s}, \\ a_{i,k}^* \equiv \left\langle a_{i,k} \mid Z_{k-1}\right\rangle, \qquad \Gamma_{i,k} \equiv \left\langle \left(a_{i,k} - a_{i,k}^*\right)^2 \mid Z_{k-1}\right\rangle, \\ m_k^* \equiv \frac{z_k^{*2}}{\Omega_k}, \qquad s_k^* \equiv \frac{\Omega_k}{z_k^*}, \\ z_k^* \equiv \left\langle z_k \mid Z_{k-1}\right\rangle, \qquad \Omega_k \equiv \left\langle \left(z_k - z_k^*\right)^2 \mid Z_{k-1}\right\rangle. \end{gathered} \tag{17}$$

Therefore, the orthogonal polynomials with the weighting functions of (16) are given by Hermite polynomial and Laguerre polynomial [3]:

$$\begin{aligned} \theta_{\mathbf{m}}^{(1)}\left(\mathbf{a}_k\right) &= \prod_i \frac{1}{\sqrt{m_i!}} H_{m_i}\left(\frac{a_{i,k} - a_{i,k}^*}{\sqrt{\Gamma_{i,k}}}\right), \\ \theta_n^{(2)}\left(z_k\right) &= \sqrt{\frac{\Gamma\left(m_k^*\right) n!}{\Gamma\left(m_k^* + n\right)}} L_n^{(m_k^*-1)}\left(\frac{z_k}{s_k^*}\right). \end{aligned} \tag{18}$$

By considering (6) and independence of y_k and v_k, two parameters z_k^* and Ω_k in (17) can be given by

$$\begin{aligned} z_k^* &= y_k^* + \left\langle v_k \right\rangle, \qquad \left(y_k^* \equiv \left\langle y_k \mid Z_{k-1}\right\rangle\right), \\ \Omega_k &= \left\langle \left(y_k - y_k^*\right)^2 \mid Z_{k-1}\right\rangle + \left\langle \left(v_k - \left\langle v_k \right\rangle\right)^2 \right\rangle. \end{aligned} \tag{19}$$

Considering (5) and the property of conditional expectation, the first terms of the right sides in the above equations are expressed as follows:

$$\begin{aligned} \left\langle y_k \mid Z_{k-1}\right\rangle &= \left\langle \left\langle y_k \mid x_k, Z_{k-1}\right\rangle \mid Z_{k-1}\right\rangle \\ &= \left\langle \int_0^\infty y_k P\left(y_k \mid x_k\right) dy_k \mid Z_{k-1}\right\rangle, \\ \left\langle \left(y_k - y_k^*\right)^2 \mid Z_{k-1}\right\rangle & \\ &= \left\langle \int_0^\infty y_k^2 P\left(y_k \mid x_k\right) dy_k \mid Z_{k-1}\right\rangle - y_k^{*2}. \end{aligned} \tag{20}$$

The integrals in (20) can be calculated by using the relationship between energy and decibel variables in (8) and expansion expression in (1), as follows:

$$\begin{aligned} I_l &\equiv \int_0^\infty y_k^l P\left(y_k \mid x_k\right) dy_k \\ &= \int_{-\infty}^\infty e^{l(Y_k - K)/C} P\left(Y_k \mid X_k\right) dY_k \\ &= \left(\int_{-\infty}^\infty e^{l(Y_k - K)/C} P_0\left(Y_k\right) \right. \\ &\qquad \left. \times \sum_{r=0}^\infty \sum_{s=0}^\infty A_{rs} \varphi_r^{(1)}\left(X_k\right) \frac{1}{\sqrt{s!}} H_s\left(\frac{Y_k - \mu_Y}{\sigma_Y}\right) dY_k \right) \\ &\quad \times \left(\sum_{r=0}^\infty A_{r0} \varphi_r^{(1)}\left(X_k\right)\right)^{-1} \end{aligned}$$

$$= \exp\left\{\frac{\left(C\mu_Y + l\sigma_Y^2\right)^2}{2\sigma_Y^2 C^2} - \frac{\left(C\mu_Y^2 + 2l\sigma_Y^2 K\right)}{2\sigma_Y^2 C}\right\}$$
$$\times \left(\int_{-\infty}^{\infty} N\left(Y_k; \xi_k, \sigma_Y^2\right)\right.$$
$$\left.\times \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} A_{rs}\varphi_r^{(1)}\left(X_k\right)\frac{1}{\sqrt{s!}}H_s\left(\frac{Y_k-\mu_Y}{\sigma_Y}\right)dY_k\right)$$
$$\times \left(\sum_{r=0}^{\infty} A_{r0}\varphi_r^{(1)}\left(X_k\right)\right)^{-1}, \tag{21}$$

where

$$\xi_k \equiv \frac{1}{C}\left(C\mu_Y + l\sigma_Y^2\right), \tag{22}$$

and $d_{s0}^{(l)}$ are coefficients satisfying the following relationship:

$$H_s\left(\frac{Y_k-\mu_Y}{\sigma_Y}\right) = \sum_{j=0}^{s} d_{sj}^{(l)} H_j\left(\frac{Y_k-\xi_k}{\sigma_Y}\right). \tag{23}$$

By using the orthonormal condition of Hermite polynomial,

$$\int_{-\infty}^{\infty} N\left(Y_k; \xi_k, \sigma_Y^2\right)\frac{1}{\sqrt{i!}}H_i\left(\frac{Y_k-\xi_k}{\sigma_Y}\right)\frac{1}{\sqrt{j!}}H_j\left(\frac{Y_k-\xi_k}{\sigma_Y}\right)dY_k$$
$$= \delta_{ij}, \tag{24}$$

the function I_l in (21) can be calculated as

$$I_l = \exp\left\{\frac{\left(C\mu_Y + l\sigma_Y^2\right)^2}{2\sigma_Y^2 C^2} - \frac{C\mu_Y^2 + 2l\sigma_Y^2 K}{2\sigma_Y^2 C}\right\}$$
$$\times \frac{\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} A_{rs}\varphi_r^{(1)}\left(X_k\right)\left(d_{s0}^{(l)}/\sqrt{s!}\right)}{\sum_{r=0}^{\infty} A_{r0}\varphi_r^{(1)}\left(X_k\right)}. \tag{25}$$

Therefore, (20) can be expressed as follows:

$$\left\langle y_k \mid Z_{k-1}\right\rangle = \exp\left\{\frac{\left(C\mu_Y + \sigma_Y^2\right)^2}{2\sigma_Y^2 C^2} - \frac{C\mu_Y^2 + 2\sigma_Y^2 K}{2\sigma_Y^2 C}\right\}$$
$$\times \frac{\sum_{s=0}^{\infty} \mathbf{A}_{(s),k}\Phi\left(X_k\right)\left(d_{s0}^{(1)}/\sqrt{s!}\right)}{\sum_{r=0}^{\infty} A_{r0}\varphi_r^{(1)}\left(X_k\right)},$$
$$\left\langle \left(y_k - y_k^*\right)^2 \mid Z_{k-1}\right\rangle \tag{26}$$
$$= \exp\left\{\frac{\left(C\mu_Y + 2\sigma_Y^2\right)^2}{2\sigma_Y^2 C^2} - \frac{C\mu_Y^2 + 4\sigma_Y^2 K}{2\sigma_Y^2 C}\right\}$$
$$\times \frac{\sum_{s=0}^{\infty} \mathbf{A}_{(s),k}\Phi\left(X_k\right)\left(d_{s0}^{(2)}/\sqrt{s!}\right)}{\sum_{r=0}^{\infty} A_{r0}\varphi_r^{(1)}\left(X_k\right)} - y_k^*,$$

with

$$\Phi\left(X_k\right) \equiv \left(\varphi_0^{(1)}\left(X_k\right), \varphi_1^{(1)}\left(X_k\right), \ldots\right)^t,$$
$$\mathbf{A}_{(s),k} \equiv \mathbf{a}_{(s),k}^*, \quad (s = 1, 2, \ldots),$$
$$\mathbf{A}_{(0),k} \equiv \left(A_{00}, A_{10}, A_{20}, \ldots\right),$$
$$\mathbf{a}_{(s),k}^* \equiv \left\langle \mathbf{a}_{(s),k} \mid Z_{k-1}\right\rangle, \tag{27}$$

where t denotes the transpose of a matrix. Furthermore, using the definition of Laguerre polynomial and (1), the expansion coefficient $B_{\mathbf{m}n}$ can be calculated as follows:

$$B_{\mathbf{m}n} = \sqrt{\frac{\Gamma\left(m_k^*\right)n!}{\Gamma\left(m_k^*+n\right)}}\sum_{s_1=0}^{n}(-1)^{s_1}\binom{n}{s_1}\frac{1}{n!}\frac{\Gamma\left(m_k^*+n\right)}{\Gamma\left(m_k^*+s_1\right)}\left(\frac{1}{s_k^*}\right)^{s_1}$$
$$\times \sum_{s_2=0}^{s_1}\binom{s_1}{s_2}\exp\left\{\frac{\left(C\mu_Y + s_2\sigma_Y^2\right)^2}{2\sigma_Y^2 C^2} - \frac{C\mu_Y^2 + 2s_2\sigma_Y^2 K}{2\sigma_Y^2 C}\right\}$$
$$\times \left\langle v_k^{s_1-s_2}\right\rangle$$
$$\times \frac{\sum_{s=0}^{\infty}\left\langle \theta_{\mathbf{m}}^{(1)}\left(\mathbf{a}_k\right)\mathbf{a}_{(s),k} \mid Z_{k-1}\right\rangle \Phi\left(X_k\right)\left(1/\sqrt{s!}\right)d_{s0}^{(s_2)}}{\sum_{r=0}^{\infty} A_{r0}\varphi_r^{(1)}\left(X_k\right)}. \tag{28}$$

From (19) and (26)–(28), it can be found that the parameters z_k^*, Ω_k and the expansion coefficient $B_{\mathbf{m}n}$ are given by the predictions of unknown parameter $\mathbf{a}_k$, the statistics of the external noise v_k, and the input observations X_k.

By considering (11), the prediction to perform the recurrence estimation can be given for an arbitrary polynomial function $g_{\mathbf{N}}(\mathbf{a}_{k+1})$ with $\mathbf{N}$th order of $\mathbf{a}_{k+1}$ can be expressed as

$$g_{\mathbf{N}}^*\left(\mathbf{a}_{k+1}\right) \equiv \left\langle g_{\mathbf{N}}\left(\mathbf{a}_{k+1}\right) \mid Z_k\right\rangle = \left\langle g_{\mathbf{N}}\left(\mathbf{a}_k\right) \mid Z_k\right\rangle = \widehat{g}_{\mathbf{N}}\left(\mathbf{a}_k\right). \tag{29}$$

2.3. Prediction of Output Probability Distribution for Sound Environment System. Because the conditional probability density function $P(Y \mid X)$ can be considered as an invariant system characteristic, reflecting mainly the proper correlation relationship between the two sound pressure levels X and Y in the sound environment system, the output probability distribution $P_s(Y)$ in decibel scale can be predicted, as $P_s(Y) = \langle P(Y \mid X)\rangle_X$. Thus, based on (1) and using the estimated parameter $\widehat{\mathbf{a}}_k \equiv (\widehat{A}_{11}, \widehat{A}_{12}, \ldots)$, the output probability density function $P_s(Y)$ can be predicted from the observed input data X, as follows:

$$P_s(Y) = P_0(Y)\sum_{s=0}^{\infty}\left\langle \frac{\sum_{r=0}^{\infty}\widehat{A}_{rs}\varphi_r^{(1)}(X)}{\sum_{r=0}^{\infty}A_{r0}\varphi_r^{(1)}(X)}\right\rangle_X \varphi_s^{(2)}(Y). \tag{30}$$

3. Application to Real Sound Environment System

3.1. Application to Sound Insulation System. In order to confirm the effectiveness of the proposed method, it was

Table 1: Statistics of the input, output signals and the background noise in sound insulation system.

	Mean [W/m^2]	Standard deviation [W/m^2]
Input signal	1.3981×10^{-4}	8.2881×10^{-5}
Output signal	1.0633×10^{-6}	5.0934×10^{-6}
Background noise	1.0633×10^{-6}	4.4550×10^{-6}

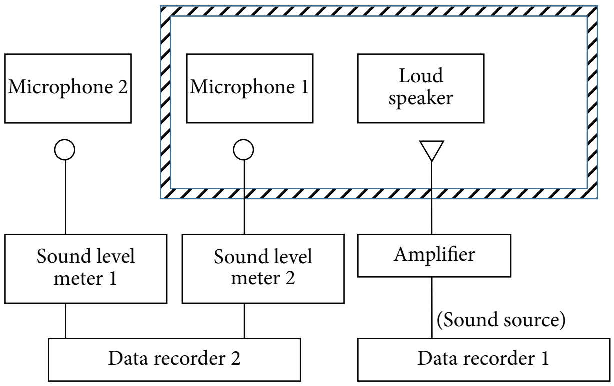


Figure 1: A schematic drawing of the experimental setup in sound insulation system.

applied to real data observed in a sound environment system. Acoustic signals observed by two microphones in indoors and outdoors for a house were adopted as input and output data for the sound insulation system. The schematic drawing of the sound environment system is shown in Figure 1. The rock music was selected as an input signal by considering the aggravation of "Karaoke" noise pollution problem, and white noise was adopted as a background noise. The statistics of the input, output signals and the background noise are shown in Table 1. The input and output fluctuation data simultaneously measured with every sampling interval 1 s. Based on the 500 data, the expansion coefficients A_{rs} in (1) were estimated on the basis of the input signal X_k and observation z_k under existence of the background noise.

Based on the estimated expansion coefficients, the output response probability distribution excited by an arbitrary input signal was predicted. The 200 sampled data following the data used for the evaluation of expansion coefficients were adopted for predicting the output response probability distribution. Figure 2 shows the comparison between theoretically predicted curves and experimentally sampled points on the output probability distribution. The cumulative distributions of sound level are shown in this figure. The "theoretical curves" in this figure were obtained by predicting the probability density function of the output level Y based on the observed data of the input level X by use of the theoretical expression. In this figure, "1st–3rd approximations" considered the expansion coefficients $\widehat{A}_{11}$, $\widehat{A}_{12}$, $\widehat{A}_{21}$, and $\widehat{A}_{22}$ successively in (30). The "experimental values" represent the frequency distributions obtained directly from the observed data of the output level Y.

For comparison, the prediction results of the output probability distribution by introducing the standard regression

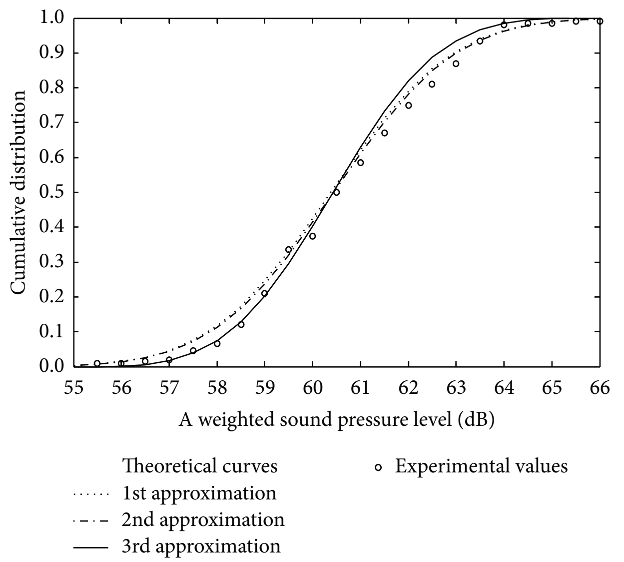


Figure 2: Comparison between experimentally sampled values and theoretically predicted curves by the proposed method on the output probability distribution for the sound insulation system.

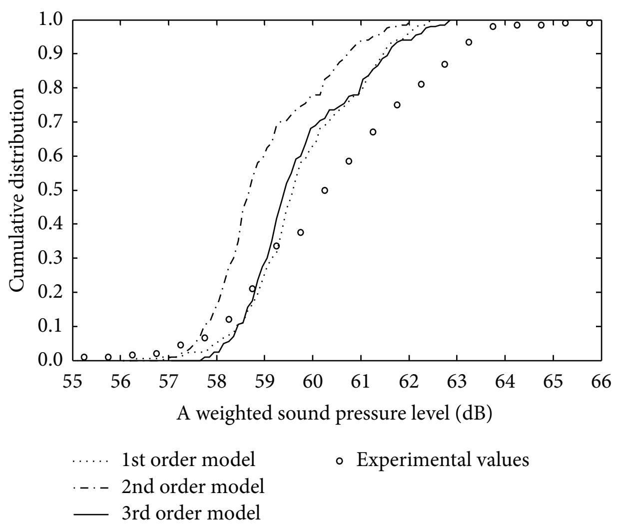


Figure 3: Comparison between experimentally sampled values and theoretically predicted curves by the standard regression models on the output probability distribution for the sound insulation system.

models described by the following equations are shown in Figure 3:

$$\widehat{Y} = a_1 + b_1 X \quad \text{(1st order model)},$$
$$\widehat{Y} = a_2 + b_2 X + c_2 X^2 \quad \text{(2nd order model)}, \tag{31}$$
$$\widehat{Y} = a_3 + b_3 X + c_3 X^2 + d_3 X^3 \quad \text{(3rd order model)}.$$

Table 2: Comparison between the experimental values and theoretically predicted values for several noise evaluation quantities in dB evaluated from Figures 2 and 3.

Noise evaluation quantities	L_5	L_{10}	L_{50}	L_{90}	L_{95}
Experimental values	63.6	63.1	60.5	58.2	57.6
Theoretical curve (1st approximation)	63.7	63.0	60.4	57.7	57.1
Theoretical curve (2nd approximation)	63.7	63.0	60.5	57.7	57.1
Theoretical curve (3rd approximation)	63.2	62.5	60.5	58.2	57.7
Regression model (1st order model)	61.9	61.5	59.6	58.4	57.9
Regression model (2nd order model)	61.2	60.7	58.7	57.7	57.5
Regression model (3rd order model)	62.1	61.6	59.4	58.4	58.1

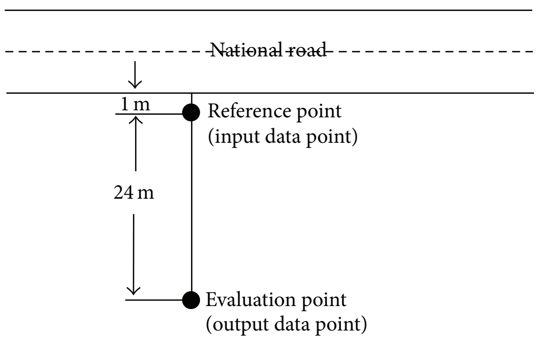


Figure 4: A schematic drawing of the experiment in road traffic noise environment near a national road.

The regression coefficients a_i, b_i, c_i, and d_i in (31) were determined by applying the extended Kalman filter [5] after introducing the following observation equation:

$$z_k = 10^{Y_k/10-12} + v_k. \tag{32}$$

The theoretically predicted curves based on the proposed method show better agreement with the experimentally sampled values than the results by applying the extended Kalman filter based on the standard regression models.

From the cumulative distributions in Figures 2 and 3, noise evaluation quantities L_x ((100 − x) percentile level) can be evaluated. Several noise evaluation quantities obtained from these figures are shown in Table 2. It is obvious that the proposed method provides a more accurate prediction than the results based on the standard regression models.

3.2. Application to Road Traffic Noise Environment. The effectiveness of the proposed method was confirmed experimentally by applying it to real road traffic noise data observed in the complicated sound environment near a national road.

Table 3: Statistics of the input, output signals and the background noise in road traffic noise environment.

	Mean [W/m^2]	Standard deviation [W/m^2]
Input signal	2.7352×10^{-5}	3.3987×10^{-5}
Output signal	1.7185×10^{-6}	2.0577×10^{-6}
Background noise	1.7185×10^{-6}	2.1789×10^{-6}

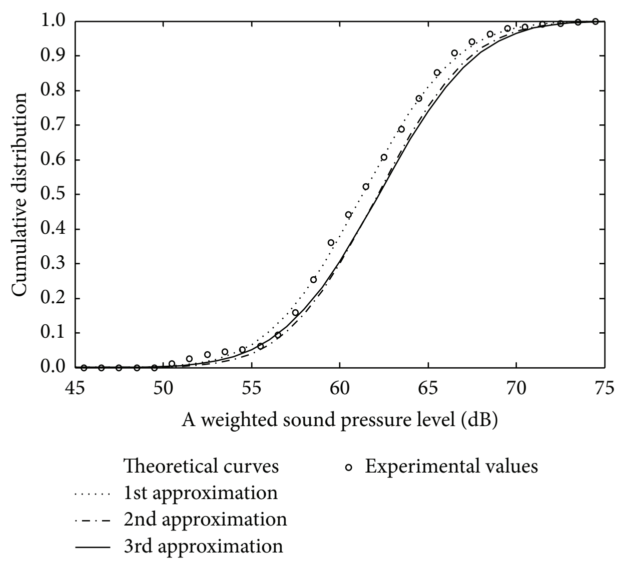


Figure 5: Comparison between experimentally sampled values and theoretically predicted curves by the proposed method on the probability distribution at the evaluation point for the road traffic noise environment near a national road.

In order to evaluate the sound environment around the main line, the sound level at an evaluation point has to be predicted on the basis of the observation at a reference point. After regarding the sound levels at a reference point and an evaluation point as system input X and output Y, respectively, the probability distribution in decibel scale at the evaluation point connected with several evaluation quantities of the sound environment was predicted on the basis of the observation at the reference point. The reference point and the evaluation point were chosen at the positions being 1 m and 25 m apart from one side of the road as shown in Figure 4. The statistics of the input, output signals and the background noise are shown in Table 3. By applying the proposed method, the probability density function of the sound level at the evaluation point was predicted on the basis of the observation at the reference point. Road traffic noise was measured by the use of the sound level meter at every 0.2 s. Through the same procedure in Section 3.1, the expansion coefficients A_{rs} in (1) were first estimated.

Based on the estimates of the expansion coefficients, the probability density function of road traffic noise at the evaluation point was predicted by measuring the road traffic noise data at the reference point. The predicted results are shown in Figures 5 and 6. Several noise evaluation quantities evaluated from Figures 5 and 6 are shown in Table 4. From these results, it can be seen that the theoretically predicted

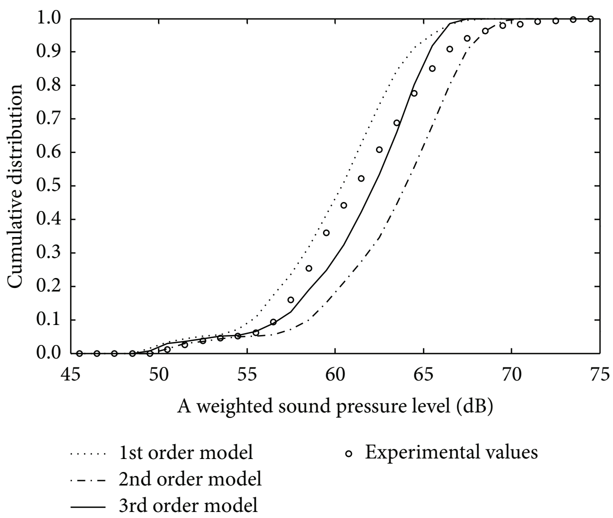


FIGURE 6: Comparison between experimentally sampled values and theoretically predicted curves by the standard regression models on the probability distribution at the evaluation point for the road traffic noise environment near a national road.

TABLE 4: Comparison between the experimental values and theoretically predicted values for several noise evaluation quantities in dB evaluated from Figures 5 and 6.

Noise evaluation quantities	L_5	L_{10}	L_{50}	L_{90}	L_{95}
Experimental values	68.0	66.3	61.3	56.6	54.5
Theoretical curve (1st approximation)	68.1	66.7	61.3	55.8	54.5
Theoretical curve (2nd approximation)	68.9	67.4	62.2	56.9	55.4
Theoretical curve (3rd approximation)	69.3	67.7	62.2	56.5	55.0
Regression model (1st order model)	65.6	64.4	60.4	55.2	52.8
Regression model (2nd order model)	68.4	67.5	64.1	58.4	55.0
Regression model (3rd order model)	66.0	65.4	62.2	56.7	53.7

values by use of the proposed method show good agreement with the experimental values, as compared with the results applying the extended Kalman filter for the standard regression models.

4. Conclusions

In this paper, an evaluation method of complex sound environment systems under existence of an external noise has been proposed. More specifically, by paying attention to the energy variables satisfying the additive property of the specific signal and the external noise, a method for estimating the correlation information between the input and output variables has been theoretically derived on the basis of the observations contaminated by the external noise. Furthermore, a prediction method of the output probability distribution in decibel scale has been derived based on the observations of the input level. The proposed prediction method has been realized by introducing a sound environment model of the conditional probability type in decibel scale. The proposed method has then been applied to the estimation and prediction of a real sound insulation system and road traffic noise environment, and it has been experimentally verified that good results have been achieved with this method.

The proposed stochastic signal processing method is quite different from the traditional standard approach. However, it is still at its early stage of study, and there are a number of practical problems to be explored in the future, starting from the result of the basic study in this paper. Some of the problems are the following.

(i) The proposed method should be applied to real prediction problems of output probability distribution for many other sound environment systems, and its practical usefulness should be verified in each real situation.

(ii) The theory should be extended to further practical cases with multi-input and multioutput systems.

(iii) An optimal number of expansion terms in the proposed stochastic signal processing method of expansion expression type should be found.

Appendix

Derivation of the Estimate

The conditional joint probability density function of the parameter $\mathbf{a}_k$ and the observation z_k can be generally expanded in a statistical orthogonal expansion series:

$$P(\mathbf{a}_k, z_k \mid Z_{k-1}) = P_0(\mathbf{a}_k \mid Z_{k-1}) P_0(z_k \mid Z_{k-1}) \times \sum_{\mathbf{m}=\mathbf{0}}^{\infty} \sum_{n=0}^{\infty} B_{\mathbf{m}n} \theta_{\mathbf{m}}^{(1)}(\mathbf{a}_k) \theta_n^{(2)}(z_k). \quad \text{(A.1)}$$

After substituting (A.1) into (12), taking the conditional expectation of the function $f_{\mathbf{M}}(\mathbf{a}_k)$, and using the orthonormal condition for the function $\theta_{\mathbf{m}}^{(1)}(\mathbf{a}_k)$, (13) can be derived.

Conflict of Interests

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

References

[1] A. Ikuta, H. Masuike, and M. Ohta, "A digital filter for stochastic systems with unknown structure and its application to psychological evaluation of sound environment," *IEICE Transactions on Information and Systems*, vol. 88, no. 7, pp. 1519–1525, 2005.

[2] M. Ohta and A. Ikuta, "Acoustic signal processing for generalized regression analysis with reduced information loss based on data observed with an amplitude limitation," *Acustica*, vol. 81, no. 2, pp. 129–135, 1995.

[3] M. Ohta and T. Koizumi, "General statistical treatment of the response of a nonlinear rectifying device to a stationary random input," *IEEE Transactions on Information Theory*, vol. 14, no. 4, pp. 595–598, 1968.

[4] M. Ohta and H. Yamada, "New methodological trials of dynamical state estimation for the noise and vibration environmental system—establishment of general theory and its application to urban noise problems," *Acustica*, vol. 55, no. 4, pp. 199–212, 1984.

[5] H. J. Kushner, "Approximations to optimal nonlinear filters," *IEEE Transactions on Automatic Control*, vol. AC-12, no. 5, pp. 546–556, 1967.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 752831, 5 pages
http://dx.doi.org/10.1155/2014/752831

Research Article

Further Results on Dynamic Additive Hazard Rate Model

Zhengcheng Zhang[1] and Limin Zhang[2]

[1] *School of Mathematics and Physics, Lanzhou Jiaotong University, P.O. Box 405, Anning West Road, Anning District, Lanzhou, Gansu 730070, China*
[2] *Department of Applied Mathematics, College of Basic Science and Information Engineering, Yunnan Agricultural University, Kunming, Yunnan 650201, China*

Correspondence should be addressed to Zhengcheng Zhang; zhzhcheng004@163.com

Received 15 April 2014; Revised 22 June 2014; Accepted 7 July 2014; Published 21 July 2014

Academic Editor: Wuquan Li

In the past, the proportional and additive hazard rate models have been investigated in the works. Nanda and Das (2011) introduced and studied the dynamic proportional (reversed) hazard rate model. In this paper we study the dynamic additive hazard rate model, and investigate its aging properties for different aging classes. The closure of the model under some stochastic orders has also been investigated. Some examples are also given to illustrate different aging properties and stochastic comparisons of the model.

1. Introduction

It is common practice in statistical analysis that covariates are often introduced to account for factors that increase the heterogeneity of a population. When the effect of a factor under study has a multiplicative (or additive) effect on the baseline hazard function, we have a proportional (or an additive) hazard model. The latter category of model is preferred in any situation. For example, in tumorigenicity cases, where the dose effect on tumor risk is of interest, the excess risk becomes an important factor. Clinical trials that seek the effectiveness of treatments often experience lag times of treatment effectiveness after which treatment procedures will be in full effect.

In reliability and survival analysis, devices or systems always operate in a changing environment. The conditions under which systems operate can be harsher or gentler in modeling lifetime of the devices or systems. The most known Cox [1] model is that the changing conditions are assumed to act multiplicatively on the baseline hazard rate. This model has been widely used in many experiments where the time to systems' failure depends on a group of covariates, which may be regarded as different treatments, operating conditions, heterogeneous environments, and so forth. P. L. Gupta and R. C. Gupta [2] studied the relation between the conditional and unconditional failure rates in mixtures when the distributions in the mixture follow the proportional hazard rate. For further research, one may see Cox and Oakes [3], Kumar and Westberg [4], Dupuy [5], Lau [6], Zhao and Zhou [7], X. Li and Z. Li [8], and Yu [9].

R. C. Gupta and R. D. Gupta [10] proposed and studied the proportional reversed hazard model to analyze failure time data. For more details on this model, see Gupta and Wu [11], X. Li and Z. Li [12], and so forth.

Recently, Nanda and Das [13] introduced the dynamic proportional hazard rate (DPHR) model and the dynamic proportional reversed hazard rate (DPRHR) model and studied their properties for different aging classes. The closure of the models under different stochastic orders has also been studied.

Aranda-Ordaz [14] first dealt with an additive hazard model

$$h(t \mid Z(s), s \le t) = \beta' z(t) + h_0(t), \quad \text{for } t \ge 0, \qquad (1)$$

where $h_0(t)$ is a baseline hazard rate and a time-dependent covariate vector Z, representing the changes in the operating conditions, and β is a vector of parameters. For more details, one may see Cox and Oakes [3], Thomas [15], Breslow and Day [16], Finkelstein and Esaulova [17], Lim and Zhang [18], and so forth.

Assume that X and Y are the lifetimes of two systems with corresponding hazard rate functions $h_X(t)$ and $h_Y(t)$ for $t \geq 0$. Let $c(t) = \beta' z(t)$; the model (with time-dependent covariates) in (1) would reduce to the form

$$h_Y(t) = c(t) + h_X(t), \quad \forall t \geq 0, \tag{2}$$

which is named as dynamic additive hazard rate (DAHR) model.

Sometimes the hazard rate functions of X and Y may not be additive over the whole interval $[0, \infty)$, but they may be additive differently from different intervals. Specifically, they may be related as

$$h_Y(t) = c_i + h_X(t), \quad t_{i-1} \leq t \leq t_i \tag{3}$$

for $i = 1, 2, \ldots$, and $t_0 = 0$, where c_i $(i = 1, 2, \ldots)$ are some constants. When the intervals $[t_{i-1}, t_i)(i = 1, 2, \ldots)$ become smaller and smaller, a model as in (2) will be naturally obtained.

In order to guarantee that $h_Y(t)$ is a hazard rate function of a nonnegative random variable Y, the following lemma is given.

Lemma 1. *Assume that $c(t)$ and $h_X(t)$ are defined before. Then, for $t \geq 0$, $h_Y(t) = c(t) + h_X(t)$ is a hazard rate function if and only if the following conditions hold:*

(i) $c(t) + h_X(t) \geq 0$, *for all* $t \geq 0$;

(ii) $\int_0^{\infty} (c(t) + h_X(t))dt = \infty$;

(iii) *if* $\int_0^{t_0} h_X(t)dt = \infty$, *then*

$$\int_0^{t_0} (c(t) + h_X(t))\, dt = \infty, \tag{4}$$

for some $t_0 < \infty$.

In Section 2 of the paper, we discuss some aging properties of the DAHR model. In Section 3, the closure of DAHR model under different stochastic orderings is studied. Some examples are given to illustrate the results concerned in Sections 2 and 3.

Throughout the paper, assume that all random variables under consideration have 0 as the common left end point of their supports, and the terms increasing and decreasing stand for monotone nondecreasing and monotone nonincreasing, respectively.

2. Aging Properties of DAHR Model

At first we introduce some concepts of aging notions that will be useful in the section. Recall that a random variable X is said to be (a) increasing in failure rate (IFR) [decreasing in failure rate (DFR)] if $h_X(t)$ is increasing [decreasing] in $t \geq 0$; (b) increasing in failure rate in average (IFRA) [decreasing in failure rate in average (DFRA)] if $\int_0^t h_X(u)du/t$ is increasing [decreasing] in $t \geq 0$; (c) new better than used (NBU) [new worse than used (NWU)] if $\overline{F}(x+t) \leq [\geq] \overline{F}(t)\overline{F}(x)$, for all $t, x \geq 0$; (d) new better than used in failure rate (NBUFR) [new worse than used in failure rate (NWUFR)] if $h_X(t) \geq (\leq)h_X(0)$, for all $t \geq 0$; (e) new better than used in failure rate average (NBAFR) [new worse than used in failure rate average (NWAFR)] if $\int_0^t h_X(u)du/t \geq [\leq] h_X(0)$, for all $t \geq 0$. For more discussions on properties of aging notions, readers may refer to Barlow and Proschan [19], Müller and Styan [20], and so forth.

In the following we give some aging closure properties between the random variables X and Y under some conditions of $c(t)$. Some results are obvious and hence their proofs are omitted.

Proposition 2. *If the random variable X is IFR (DFR) and, for $t \geq 0$, $c(t)$ is increasing (decreasing), then the random variable Y is IFR (DFR).*

In the following, we give two examples related to this proposition. Example 3 is an application of the proposition. Example 4 indicates that the condition of $c(t)$ is sufficient but not a necessary one.

Example 3. Let X be a random variable having Weibull distribution with hazard rate function $h_X(t) = 2t$, $t \geq 0$. Take $c(t) = t$ for $t \geq 0$. It is obvious that $c(t)$ satisfies all the conditions of Lemma 1. Obviously, if X is IFR and $c(t)$ is increasing in t, hence Y is IFR.

Example 4. Let X be a random variable having Weibull distribution with hazard rate function $h_X(t) = 2t$, $t \geq 0$. Let $c(t) = (2 + t^2)/(1 + t)$ for $t \geq 0$. It can be verified that $h_X(t) + c(t)$ is increasing in $t \geq 0$, and hence Y is IFR. However, $c(t)$ is decreasing in $t \in [0, \sqrt{3} - 1)$ but increasing in $t \in [\sqrt{3} - 1, +\infty)$.

Proposition 5. *If the random variable X is IFRA (DFRA) and $c(t)$ is increasing (decreasing) in $t \geq 0$, then the random variable Y is IFRA (DFRA).*

Proof. For $t \geq 0$, let

$$q(t) = \frac{\int_0^t h_Y(x)\, dx}{t} = \frac{\int_0^t (c(x) + h_X(x))\, dx}{t}. \tag{5}$$

Note that X is IFRA (DFRA) and $c(t)$ is increasing (decreasing) implying that

$$\begin{aligned} q'(t) &= \frac{c(t) + h_X(t)}{t} - \frac{\int_0^t (c(x) + h_X(x))\, dx}{t^2} \\ &= \frac{\int_0^t (c(t) - c(x))\, dx}{t^2} + \frac{t h_X(t) - \int_0^t h_X(x)\, dx}{t^2} \\ &\geq 0 \ (\leq 0). \end{aligned} \tag{6}$$

Hence the desired result follows directly. □

Example 3 can be regarded as an application of the above proposition. Example 6 below indicates that the condition of

$c(t)$ is sufficient but not a necessary one for the monotone property of Y.

Example 6. Let X be a random variable having Weibull distribution with hazard rate function $h_X(t) = 2t$, $t \geq 0$. Take $c(t) = -t$ for $t \geq 0$. It is obvious that $c(t)$ satisfies all the conditions of Lemma 1. Obviously, X is IFRA and Y is IFRA. However, $c(t)$ is decreasing in $t \geq 0$.

Proposition 7. *If the random variable X is NBU (NWU) and $c(t)$ is increasing (decreasing) in $t \geq 0$, then the random variable Y is NBU (NWU).*

Proof. We only give the proof for the case of NBU. In order to prove that Y is NBU, it is sufficient to prove that, for all $t \geq 0$ and $x \geq 0$,

$$e^{-\int_0^{x+t}(c(u)+h_X(u))du} \leq e^{-\int_0^{x}(c(u)+h_X(u))du} \times e^{-\int_0^{t}(c(u)+h_X(u))du}. \tag{7}$$

It is equivalent to

$$e^{-\int_t^{x+t}(c(u)+h_X(u))du} \leq e^{-\int_0^{x}(c(u)+h_X(u))du}. \tag{8}$$

That is,

$$e^{-\int_0^{x}(c(u+t)+h_X(u+t))du} \leq e^{-\int_0^{x}(c(u)+h_X(u))du}. \tag{9}$$

Note that X is NBU which implies that

$$e^{-\int_0^{x+t}h_X(u)du} \leq e^{-\int_0^{x}h_X(u)du} \cdot e^{-\int_0^{t}h_X(u)du}. \tag{10}$$

That is,

$$e^{-\int_0^{x}h_X(u+t)du} \leq e^{-\int_0^{x}h_X(u)du}. \tag{11}$$

From the fact that $c(t)$ is increasing and (11), (9) holds, and hence the desired result follows. □

Example 3 is an application of the above proposition. The following example indicates that the condition of $c(t)$ is sufficient but not a necessary one for the NBU property of Y.

Example 8. Assume that X is a random variable having exponential distribution with mean 1/2. It is clear that X is NBU. Let $c(t) = (1+t)/(1+t^2)$ for $t \geq 0$. By some computations, we have

$$\begin{aligned} a(t,x) &= \int_0^x (c(t+u) - c(u) + h_X(t+u) - h_X(u))\, du \\ &= \arctan(t+x) + \frac{1}{2}\ln\left(1+(t+x)^2\right) - \arctan x \\ &\quad + \frac{1}{2}\ln\left(1+x^2\right) + 2xt. \end{aligned} \tag{12}$$

It can be verified that $a(t,x)$ is nonnegative for $t, x \geq 0$ (see also Figure 1). From (9), we conclude that Y is NBU. However, it is easily obtained that $c(t)$ is increasing in $[0, \sqrt{2}-1)$ but decreasing in $(\sqrt{2}-1, +\infty)$.

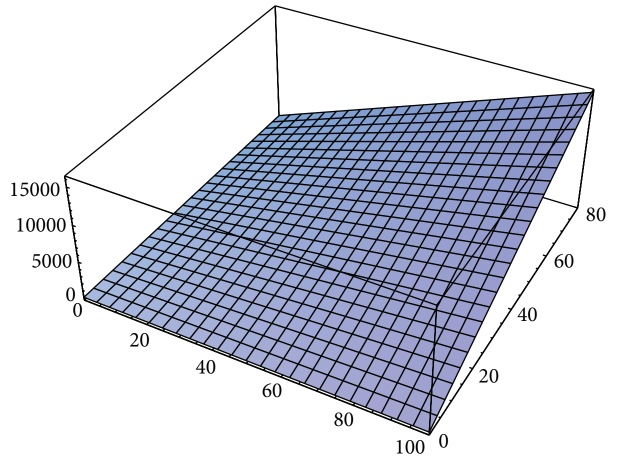


FIGURE 1: Plot of the $a(t,x)$ for $(x,t) \in [0,100] \times [0,80]$.

Proposition 9. *If the random variable X is NBUFR (NWUFR) and $c(t) \geq 0\,(\leq 0)$ for $t \geq 0$, then the random variable Y is NBUFR (NWUFR).*

Proposition 10. *If the random variable X is NBAFR (NWAFR) and $\int_0^t c(u)du \geq (\leq) tc(0)$ for $t \geq 0$, then the random variable Y is NBAFR (NWAFR).*

Proof. We only give the proof for the case of NBAFR. It is noted that Y is NBAFR which is equivalent to that, for all $t \geq 0$, $(\int_0^t (c(u) + h_X(u))du)/t = (\int_0^t h_Y(u)du)/t \geq h_Y(0) = c(0) + h_X(0)$. Note that X is NBAFR if and only if $\int_0^t h_X(u)du/t \geq h_X(0)$. Hence the desired result follows from the condition $\int_0^t c(u)du \geq tc(0)$. □

Remark 11. Example 3 is an application of Propositions 9 and 10. Example 6 can be regarded as a counterexample, which shows that the condition $c(t) \geq 0$ is a sufficient but not a necessary one in Propositions 9 and 10.

3. Stochastic Comparisons of DAHR Model

Firstly let us recall the concepts of some stochastic orders that are closely related to the main results in this section. A random variable X is said to be larger than another random variable Y in (a) aging intensity ordering (denoted by $X \geq_{ai} Y$), if

$$\frac{h_X(t)}{\int_0^t h_X(u)\,du} \leq \frac{h_Y(t)}{\int_0^t h_Y(u)\,du}, \tag{13}$$

for all $t \geq 0$; (b) usual stochastic order (denoted by $X \leq_{st} Y$) if $\overline{F}_X(t) \leq \overline{F}_Y(t)$, for all $t \geq 0$; (c) hazard rate order (denoted by $X \leq_{hr} Y$) if $h_X(t) \geq h_Y(t)$, for all $t \geq 0$; (d) up hazard rate order (denoted by $X \leq_{hr\uparrow} Y$) if $X - t \leq_{hr} Y$, for all $t \geq 0$; (e) down hazard rate order (denoted by $X \leq_{hr\downarrow} Y$) if $X \leq_{hr} [Y - t \mid Y > t]$, for all $t \geq 0$. For more details about stochastic orders, please refer to Shaked and Shanthikumar [21].

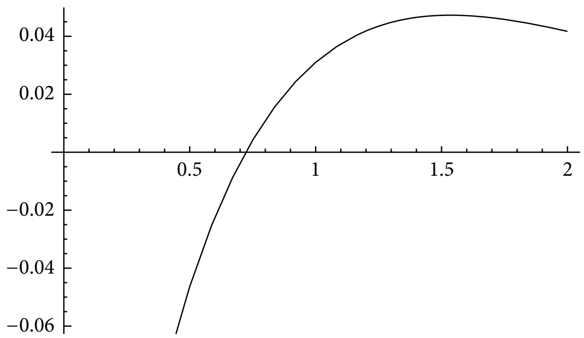


Figure 2: Plot of the $a(t)$ for $t \in [0, 2]$.

In the following we give some sufficient (and necessary) conditions of stochastic ordering between random variables X and Y. Some results are obvious and hence their proofs are omitted.

Proposition 12. *Suppose X and Y are two nonnegative random variables satisfying* (2). *Then, $X \geq_{ai} (\leq_{ai}) Y$ if $c(t)/h_X(t)$ is increasing (decreasing) in $t \geq 0$.*

Proof. Note that $X \geq_{ai} Y$ if and only if, for all $t \geq 0$,

$$\frac{h_X(t)}{\int_0^t h_X(u)\,du} \leq \frac{c(t) + h_X(t)}{\int_0^t (c(u) + h_X(u))\,du}. \tag{14}$$

It is equivalent to that $\int_0^t (c(t)h_X(u) - c(u)h_X(t))du \geq 0$. It holds if $c(t)/h_X(t)$ is increasing in $t \geq 0$. The proof of the parenthetical statement is similar. □

The following example indicates that the condition of the monotone property of the $c(t)/h_X(t)$ is sufficient but not a necessary one for the aging intensity ordering between X and Y.

Example 13. Assume that X is a random variable having exponential distribution with mean 1/2. Let $c(t) = (1+t^2)/(1+t)$ for $t \geq 0$. By some computations, we have

$$\begin{aligned} a(t) &= \int_0^t \left(c(t)\,h_X(u) - c(u)\,h_X(t)\right) du \\ &= \frac{2t\left(1+t^2\right)}{1+t} - t^2 + 2t + 4\ln(1+t). \end{aligned} \tag{15}$$

It can be verified that $a'(t) \geq 0$ for $t \geq 0$, and hence $a(t)$ is increasing in $t \geq 0$ (see also Figure 2). Note that $a(0) = 0$. Thus $a(t) \geq 0$, for all $t \geq 0$, and hence $X \geq_{ai} Y$. However, it is easily obtained that $c(t)/h_X(t)$ is decreasing in $[0, \sqrt{2} - 1)$ but increasing in $(\sqrt{2} - 1, +\infty)$.

Proposition 14. *Suppose X and Y are two nonnegative random variables satisfying* (2). *Then, $X \geq_{st} (\leq_{st}) Y$ if and only if $\int_0^t c(u)du \geq (\leq) 0$, for all $t \geq 0$.*

The following corollary follows immediately from the proposition above.

Corollary 15. *If $c(t) \geq (\leq) 0$, for all $t \geq 0$, then $X \geq_{st} (\leq_{st}) Y$.*

Proposition 16. *Suppose X and Y are two nonnegative random variables satisfying* (2). *Then, $X \geq_{hr} (\leq_{hr}) Y$ if and only if $c(t) \geq (\leq) 0$, for all $t \geq 0$.*

Proposition 17. *Suppose that X and Y are two nonnegative random variables satisfying* (2). *Then, $X \leq_{hr\uparrow} (\geq_{hr\uparrow}) Y$ if and only if $h_X(y+t) - h_X(t) - c(t) \geq (\leq) 0$, for all $y \geq 0$ and $t \geq 0$.*

Proof. Note that $X \leq_{hr\uparrow} Y$ if and only if

$$\frac{\exp\left[-\int_0^x h_Y(u)\,du\right]}{\exp\left[-\int_0^{x+t} h_X(u)\,du\right]} \tag{16}$$

is increasing in x, for all $t \geq 0$. It is equivalent to the fact that

$$\exp\left[\int_0^{x+t} h_X(u) - \int_0^x h_Y(u)\,du\right] \tag{17}$$

is increasing in x, which is equivalent to the fact that its derivative is nonnegative; that is, $h_X(y+t) - h_X(t) - c(t) \geq 0$, for all $y \geq 0$ and $t \geq 0$. It follows from the condition. The proof of the parenthetical statement is similar. □

Proposition 18. *Suppose that X and Y are two nonnegative continuous random variables satisfying* (2). *Then, $X \leq_{hr\downarrow} (\geq_{hr\downarrow}) Y$ if and only if $h_X(y) - h_X(t+y) - c(t+y) \geq (\leq) 0$, for all $y \geq 0$ and $t \geq 0$.*

Its proof is similar to that of Proposition 17 and hence is omitted.

Conflict of Interests

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

Acknowledgment

The research was supported by the National Natural Science Foundation of China (71361020).

References

[1] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society*, vol. 34, pp. 187–220, 1972.

[2] P. L. Gupta and R. C. Gupta, "Ageing characteristics of the Weibull mixtures," *Probability in the Engineering and Informational Sciences*, vol. 10, no. 4, pp. 591–600, 1996.

[3] D. R. Cox and D. Oakes, *Analysis of Survival Data*, Chapman & Hall, New York, NY, USA, 1984.

[4] D. Kumar and U. Westberg, "Maintenance scheduling under age replacement policy using proportional hazards model and TTT-plotting," *European Journal of Operational Research*, vol. 99, no. 3, pp. 507–515, 1997.

[5] J. Dupuy, "The proportional hazards model with covariate measurement error," *Journal of Statistical Planning and Inference*, vol. 135, no. 2, pp. 260–275, 2005.

[6] J. W. Lau, "Bayesian semi-parametric modeling for mixed proportional hazard models with right censoring," *Statistics & Probability Letters*, vol. 76, no. 7, pp. 719–728, 2006.

[7] X. Zhao and X. Zhou, "Proportional hazards models for survival data with long-term survivors," *Statistics & Probability Letters*, vol. 76, no. 15, pp. 1685–1693, 2006.

[8] R. Li and L. Chambless, "Test for additive interaction in proportional hazards models," *Annals of Epidemiology*, vol. 17, no. 3, pp. 227–236, 2007.

[9] Q. Yu, "A note on the proportional hazards model with discontinuous data," *Statistics & Probability Letters*, vol. 77, no. 7, pp. 735–739, 2007.

[10] R. C. Gupta and R. D. Gupta, "Proportional reversed hazard rate model and its applications," *Journal of Statistical Planning and Inference*, vol. 137, no. 11, pp. 3525–3536, 2007.

[11] R. C. Gupta and H. Wu, "Analyzing survival data by proportional reversed hazard model," *International Journal of Reliability and Applications*, vol. 2, pp. 1–26, 2001.

[12] X. Li and Z. Li, "A mixture model of proportional reversed hazard rate," *Communications in Statistics: Theory and Methods*, vol. 37, no. 18-20, pp. 2953–2963, 2008.

[13] A. K. Nanda and S. Das, "Dynamic proportional hazard rate and reversed hazard rate models," *Journal of Statistical Planning and Inference*, vol. 141, no. 6, pp. 2108–2119, 2011.

[14] F. J. Aranda-Ordaz, "An extension of the proportional-hazards model for grouped data," *Biometrics*, vol. 39, no. 1, pp. 109–117, 1983.

[15] D. C. Thomas, "Use of auxiliary information in fitting nonproportional hazards models," in *Modern Statistical Methods in Chronic Disease Epidemiolgy*, E. D. S. Moolgavkar and R. L. Prentice, Eds., pp. 197–210, Wiley, New York, NY, USA, 1986.

[16] N. E. Breslow and N. E. Day, *Statistical Methods in Cancer Research—vol 1, The Design and Analysis of Case-Control Studies*, IARC, Lyon, France, 1987.

[17] M. S. Finkelstein and V. Esaulova, "On an inverse problem in mixture failure rates modelling," *Applied Stochastic Models in Business and Industry*, vol. 17, no. 2, pp. 221–229, 2001.

[18] H. J. Lim and X. Zhang, "Semi-parametric additive risk models: application to injury duration study," *Accident Analysis and Prevention*, vol. 41, no. 2, pp. 211–216, 2009.

[19] R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, New York, NY, USA, 1981.

[20] A. Müller and D. Styan, *Comparison Methods for Stochastic Models and Risks*, John Wlley & Sons, Westsusse, UK, 2002.

[21] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and Their Applications*, Academic Press, San Diego, Calif, USA, 1994.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 231592, 5 pages
http://dx.doi.org/10.1155/2014/231592

Research Article

Integer-Valued Moving Average Models with Structural Changes

Kaizhi Yu,[1] Hong Zou,[2] and Daimin Shi[1]

[1] *Statistics School, Southwestern University of Finance and Economics, Chengdu 611130, China*
[2] *School of Economics, Southwestern University of Finance and Economics, Chengdu 611130, China*

Correspondence should be addressed to Kaizhi Yu; kaizhiyu.swufe@gmail.com

Received 2 March 2014; Revised 22 June 2014; Accepted 7 July 2014; Published 21 July 2014

Academic Editor: Wuquan Li

It is frequent to encounter integer-valued time series which are small in value and show a trend having relatively large fluctuation. To handle such a matter, we present a new first order integer-valued moving average model process with structural changes. The models provide a flexible framework for modelling a wide range of dependence structures. Some statistical properties of the process are discussed and moment estimation is also given. Simulations are provided to give additional insight into the finite sample behaviour of the estimators.

1. Introduction

Integer-valued time series occur in many situations, often as counts of events in consecutive points of time, for example, the number of births at a hospital in successive months, the number of road accidents in a city in successive months, and big numbers even for frequently traded stocks. Integer-valued time series represent an important class of discrete-valued time series models. Because of the broad field of potential applications, a number of time series models for counts have been proposed in literature. McKenzie [1] introduced the first order integer-valued autoregressive, INAR(1), model. The statistical properties of the INAR(1) are discussed in McKenzie [2], Al-Osh and Alzaid [3]. The model is further generalized to a pth-order autoregression, INAR(p), by Alzaid and Al-Osh [4] and Du and Li [5]. The qth-order integer-valued moving average model, INMA(q), was introduced by Al-Osh and Alzaid [6] and in a slightly different form by McKenzie [7]. Ferland et al. [8] proposed an integer-valued GARCH model to study overdispersed counts, and Fokianos and Fried [9], Weiß [10], and Zhu and Wang [11–13] made further studies. Györfi et al. [14] proposed a nonstationary inhomogeneous INAR(1) process, where the autoregressive type coefficient slowly converges to one. Bakouch and Ristić [15] introduced a new stationary integer-valued autoregressive process of the first order with zero truncated Poisson marginal distribution. Kachour and Yao [16] introduced a class of autoregressive models for integer-valued time series using the rounding operator. Kim and Park [17] proposed an extension of integer-valued autoregressive INAR models by using a signed version of the thinning operator. Zheng et al. [18] proposed a first order random coefficient integer-valued autoregressive model and got its ergodicity, moments, and autocovariance functions of the process. Gomes and Canto e Castro [19] presented a random coefficient autoregressive process for count data based on a generalized thinning operator. Existence and weak stationarity conditions for these models were established. A simple bivariate integer-valued time series model with positively correlated geometric marginals based on the negative binomial thinning mechanism was presented by Ristić et al. [20], and some properties of the model are also considered. Pedeli and Karlis [21] considered a bivariate INAR(1) (BINAR(1)) process where cross correlation is introduced through the use of copulas for the specification of the joint distribution of the innovations.

Structural changes in economic data frequently correspond to instabilities in the real world. However, most work in this area has been concentrated on models without structural changes. It seems that the integer-valued autoregressive moving average (INARMA) model with break point has not attracted too much attention. For instance, a new method for modelling the dynamics of rain sampled by a tipping bucket rain gauge was proposed by Thyregod et al. [22].

The models take the autocorrelation and discrete nature of the data into account. First order, second order, and threshold models are presented together with methods to estimate the parameters of each model. Monteiro et al. [23] introduced a class of self-exciting threshold integer-valued autoregressive models driven by independent Poisson-distributed random variables. Basic probabilistic and statistical properties of this class of models were discussed. Moreover, parameter estimation was also addressed. Hudecová [24] suggested a procedure for testing a change in the autoregressive models for binary time series. The test statistic is a maximum of normalized sums of estimated residuals from the model, and thus it is sensitive to any change which leads to a change in the unconditional success probability. Structural change is a statement about parameters, which only have meaning in the context of a model. In our discussion, we will focus on structural change in the simple count data model, the first order integer-valued moving average model, whose coefficient varies with the value of innovation. One of the leading reasons is that piecewise linear functions can offer a relatively simple approximation to the complex nonlinear dynamics.

The rest of this paper is divided into four sections. In Section 2, we give the definition and basic properties of the new INMA(1) model with structural changes. Section 3 discusses the estimation of the unknown parameters. We test the accuracy of the estimation via simulations in Section 4. Section 5 includes some concluding remarks.

2. Definition and Basic Properties

Definition 1. Let $\{X_t\}$ be a process with state space $\mathbb{N}_0$; let $0 < \alpha_i < 1$, $i = 1, \ldots, m$, and τ_i, $i = 1, \ldots, m-1$, be positive integers. The process $\{X_t\}$ is said to be first order integer-valued moving average model with structural change (INMASC(1)) if X_t satisfies the following equation:

$$X_t = \begin{cases} \alpha_1 \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \tau_0 \le \varepsilon_{t-1} \le \tau_1 \\ \alpha_2 \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \tau_1 < \varepsilon_{t-1} \le \tau_2 \\ \vdots & \\ \alpha_{m-1} \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \tau_{m-2} < \varepsilon_{t-1} \le \tau_{m-1} \\ \alpha_m \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \tau_{m-1} < \varepsilon_{t-1} < \tau_m, \end{cases} \tag{1}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed Poisson random variables with mean λ and $\tau_0 := 0$, $\tau_m := \infty$.

The aim of this section is to provide expressions for the moments and stationary of INMASC(1) model. For this purpose, we introduce the following notations:

$$\begin{gathered} p_i := P(\tau_{i-1} < \varepsilon_t \le \tau_i), \qquad u_i := E(\tau_{i-1} < \varepsilon_{t-1} \le \tau_i), \\ \sigma_i^2 := \operatorname{Var}(\tau_{i-1} < \varepsilon_{t-1} \le \tau_i), \qquad q_i := 1 - p_i, \\ I_{t-1,i} := \begin{cases} 1, & \text{if } \tau_{i-1} < \varepsilon_{t-1} \le \tau_i \\ 0, & \text{otherwise,} \end{cases} \qquad i = 1, \ldots, m. \end{gathered} \tag{2}$$

Theorem 2. *The numerical characteristics of $\{X_t\}$ are as follows:*

$$\begin{aligned} &(i)\ \mu_X := E(X_t) = \sum_{i=1}^m p_i \alpha_i u_i + \lambda, \\ &(ii)\ \sigma_X^2 := \operatorname{Var}(X_t) \\ &\quad = \sum_{i=1}^m p_i \alpha_i \left[\alpha_i \left(u_i^2 + \sigma_i^2\right) + (1-\alpha_i) u_i\right] \\ &\qquad - \left(\sum_{i=1}^m p_i \alpha_i u_i\right)^2 + \lambda, \\ &(iii)\ \gamma_X(k) := \operatorname{cov}(X_t, X_{t-k}) \\ &\quad = \begin{cases} \sum_{i=1}^m p_i \alpha_i \left(u_i^2 + \sigma_i^2 - \lambda u_i\right), & k = 1 \\ 0, & k \ge 2. \end{cases} \end{aligned} \tag{3}$$

Proof. (i) It is easy to get the mean and variance of X_t by using the law of iterated expectations:

$$\begin{aligned} E(X_t) &= E\left[I_{t-1,1}(\alpha_1 \circ \varepsilon_{t-1}) + \cdots + I_{t-1,m}(\alpha_m \circ \varepsilon_{t-1}) + \varepsilon_t\right] \\ &= E\{E[I_{t-1,1}(\alpha_1 \circ \varepsilon_{t-1}) \\ &\qquad + \cdots + I_{t-1,m}(\alpha_m \circ \varepsilon_{t-1}) \mid \varepsilon_{t-1}]\} + E(\varepsilon_t) \\ &= \alpha_1 E(I_{t-1,1}\varepsilon_{t-1}) + \cdots + \alpha_m E(I_{t-1,m}\varepsilon_{t-1}) + E(\varepsilon_t) \\ &= \sum_{i=1}^m p_i \alpha_i u_i + \lambda. \end{aligned} \tag{4}$$

(ii) Moreover,

$$\begin{aligned} &\operatorname{Var}(X_t) \\ &= \operatorname{Var}\left(I_{t-1,1}(\alpha_1 \circ \varepsilon_{t-1}) + \cdots + I_{t-1,m}(\alpha_m \circ \varepsilon_{t-1}) + \varepsilon_t\right) \\ &= \sum_{i=1}^m \operatorname{Var}\left(I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1})\right) + \operatorname{Var}(\varepsilon_t) \\ &\quad + 2\sum_{i<j} \operatorname{cov}\left(I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1}), I_{t-1,j}(\alpha_j \circ \varepsilon_{t-1})\right) \\ &= \sum_{i=1}^m \{\operatorname{Var}(E[I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1}) \mid \varepsilon_{t-1}]) \\ &\qquad + E(\operatorname{Var}[I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1}) \mid \varepsilon_{t-1}])\} + \lambda \\ &\quad + 2\sum_{i<j} \left\{E\left[I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1}) I_{t-1,j}(\alpha_j \circ \varepsilon_{t-1})\right]\right. \\ &\qquad \left. - 2E\left[I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1})\right] E\left[I_{t-1,j}(\alpha_j \circ \varepsilon_{t-1})\right]\right\} \\ &= \sum_{i=1}^m \left[\alpha_i^2 \operatorname{Var}(I_{t-1,i}\varepsilon_{t-1}) + \alpha_i (1-\alpha_i) E(I_{t-1,i}\varepsilon_{t-1})\right] + \lambda \\ &\quad - 2\sum_{i<j} E\{E[I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1}) \mid \varepsilon_{t-1}]\} \\ &\qquad \times E\left\{E\left[I_{t-1,j}(\alpha_j \circ \varepsilon_{t-1}) \mid \varepsilon_{t-1}\right]\right\} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{m} \left\{ \left[\alpha_i^2 E\left(I_{t-1}^2 \varepsilon_{t-1}^2\right) - E^2\left(I_{t-1,i}\varepsilon_{t-1}\right) \right] \right. \\
&\qquad \left. + p_i \alpha_i \left(1-\alpha_i\right) u_i \right\} + \lambda \\
&\qquad - 2\sum_{i<j} \alpha_i \alpha_j E\left(I_{t-1,i}\varepsilon_{t-1}\right) E\left(I_{t-1,j}\varepsilon_{t-1}\right) \\
&= \sum_{i=1}^{m} \left\{ \left[p_i \alpha_i^2 \left(u_i^2 + \sigma_i^2\right) - p_i^2 \alpha_i^2 u_i^2 \right] + p_i \alpha_i \left(1-\alpha_i\right) u_i \right\} + \lambda \\
&\qquad - 2\sum_{i<j} p_i \alpha_i u_i p_j \alpha_j u_j \\
&= \sum_{i=1}^{m} p_i \alpha_i \left[\alpha_i \left(u_i^2 + \sigma_i^2\right) + \left(1-\alpha_i\right) u_i \right] - \left(\sum_{i=1}^{m} p_i \alpha_i u_i \right)^2 + \lambda.
\end{aligned} \tag{5}
$$

(iii) Note the correlation between $\alpha_i \circ \varepsilon_{t-1}$ and ε_{t-1}; we have

$$
\begin{aligned}
&\operatorname{cov}\left(X_t, X_{t-1}\right) \\
&= \operatorname{cov}\left(\sum_{i=1}^{m} I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right) + \varepsilon_t, \sum_{j=1}^{m} I_{t-2,j}\left(\alpha_j \circ \varepsilon_{t-2}\right) + \varepsilon_{t-1} \right) \\
&= \sum_{i=1}^{m} \operatorname{cov}\left(I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right) + \varepsilon_t, \varepsilon_{t-1}\right) \\
&= \sum_{i=1}^{m} \left\{ E\left[I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right)\varepsilon_{t-1}\right] - E\left[I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right)\right] E\left(\varepsilon_{t-1}\right) \right\} \\
&= \sum_{i=1}^{m} \left\{ E\left[E\left(I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right)\varepsilon_{t-1} \mid \varepsilon_{t-1}\right)\right] \right. \\
&\qquad \left. - E\left[E\left(I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right) \mid \varepsilon_{t-1}\right)\right] E\left(\varepsilon_{t-1}\right) \right\} \\
&= \sum_{i=1}^{m} \alpha_i \left[E\left(I_{t-1,i}\varepsilon_{t-1}^2\right) - \lambda \alpha_i E\left(I_{t-1,i}\varepsilon_{t-1}\right) \right] \\
&= \sum_{i=1}^{m} p_i \alpha_i \left(u_i^2 + \sigma_i^2 - \lambda u_i\right).
\end{aligned} \tag{6}
$$

□

Theorem 3. *Let X_t be the process defined by the equation in* (1)*; then the $\{X_t\}$ is a covariance stationary process.*

Proof. Both the unconditional mean and the unconditional variance of the $\{X_t\}$ are finite constant. And the autocovariance function does not change with time. Thus $\{X_t\}$ is a stationary process. □

Theorem 4. *Suppose $\{X_t\}$ is INMASC(1) process. Then*

(i) $\sqrt{T}(\overline{X} - \mu_X) \xrightarrow{L} N(0, \sigma_X^2 + 2\gamma_X(1))$;

(ii) $E(X_t^k \mid I_{t-1,i} = 1) < \infty$, $k = 1, 2, 3, i = 1, \ldots, m$.

Proof. (i) From definition and Theorem 2, we have that $(X_1, \ldots, X_i)$ and $(X_j, X_{j+1}, \ldots)$ are independent whenever $j - i > 1$. According to Theorem 9.1 of DasGupta [25], the process $\{X_t\}$ is a stationary 1-dependent sequence. Therefore we can complete the proof.

(ii) For $k = 1$, it follows that

$$
\begin{aligned}
E\left(X_t\right) &\le \max\left\{E\left[I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right) + \varepsilon_t\right],\ i = 1, \ldots, m\right\} \\
&\le \max\left\{E\left(\alpha_i \circ \varepsilon_{t-1}\right) + E\left(\varepsilon_t\right),\ i = 1, \ldots, m\right\} \\
&\le \lambda\left(\alpha_{\max} + 1\right) < \infty, \alpha_{\max} = \max\left(\alpha_1, \ldots, \alpha_m\right).
\end{aligned} \tag{7}
$$

For $k = 2$,

$$
\begin{aligned}
E\left(X_t^2\right) &\le \max\left\{E[I_{t-1,i}(\alpha_i \circ \varepsilon_{t-1}) + \varepsilon_t]^2,\ i = 1, \ldots, m\right\} \\
&= \max\left\{E[I_{t-i,1}\left(\alpha_i \circ \varepsilon_{t-1}\right)]^2 + E\left(\varepsilon_t^2\right) \right. \\
&\qquad \left. + 2E\left[I_{t-i,1}\left(\alpha_i \circ \varepsilon_{t-1}\right)\varepsilon_t\right],\ i = 1, \ldots, m\right\} \\
&\le \max\left\{E[(\alpha_i \circ \varepsilon_{t-1})]^2 + E\left(\varepsilon_t^2\right) \right. \\
&\qquad \left. + 2E\left[\left(\alpha_i \circ \varepsilon_{t-1}\right)\varepsilon_t\right],\ i = 1, \ldots, m\right\} \\
&= \max\left\{\left[\left(\lambda + \lambda^2\right)\alpha_i^2 + \lambda\alpha_i\left(1 - \alpha_i\right)\right] \right. \\
&\qquad \left. + \left(\lambda + \lambda^2\right) + \lambda^2\alpha_i,\ i = 1, \ldots, m\right\} \\
&\le 2\left(\lambda + \lambda^2\right)\alpha_{\max} + 0.25\lambda + \lambda^2\alpha_{\max} < \infty.
\end{aligned} \tag{8}
$$

For $k = 3$,

$$
\begin{aligned}
E\left(X_t^3\right) &\le \max\left\{E[I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right) + \varepsilon_t]^3,\ i = 1, \ldots, m\right\} \\
&= \max\left\{E[I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right)]^3 + E\left(\varepsilon_t^3\right) \right. \\
&\qquad + 3E\left[I_{t-1,i}^2(\alpha_i \circ \varepsilon_{t-1})^2\varepsilon_t\right] \\
&\qquad \left. + 3E\left[I_{t-1,i}\left(\alpha_i \circ \varepsilon_{t-1}\right)\varepsilon_t^2\right],\ i = 1, \ldots, m\right\} \\
&\le \max\left\{E(\alpha_i \circ \varepsilon_{t-1})^3 + E\left(\varepsilon_t^3\right) + 3E\left[(\alpha_i \circ \varepsilon_{t-1})^2\varepsilon_t\right] \right. \\
&\qquad \left. + 3E\left[\left(\alpha_i \circ \varepsilon_{t-1}\right)\varepsilon_t^2\right],\ i = 1, \ldots, m\right\} \\
&\le \max\left\{\left[\alpha_i^3\tau_1 + 3\alpha_i^2\left(1 - \alpha_i\right)\tau_2 \right.\right. \\
&\qquad \left. + \left(\alpha_i - 3\alpha_i^2\left(1 - \alpha_i\right) - \alpha_i^3\right)\lambda\right] \\
&\qquad + \tau_1 + 3\left\{\left[\alpha_i^2\tau_2 + \alpha_i\left(1 - \alpha_i\right)\lambda\right]\lambda\right\} \\
&\qquad \left. + 3\lambda\alpha_i\tau_2,\ i = 1, \ldots, m\right\} \\
&\le \lambda\alpha_{\max}\left[\alpha_{\max}^2\left(\tau_1 - 1 - 3\lambda\right) \right. \\
&\qquad \left. + 3\tau_2\left(\alpha_{\max} + 1\right) + 3\lambda + 1\right] + \tau_1 < \infty,
\end{aligned} \tag{9}
$$

where $\tau_1 := \lambda^3 + 3\lambda^2 + \lambda$, $\tau_2 := \lambda^2 + \lambda$, and $\alpha_{\max} = \max(\alpha_1, \ldots, \alpha_m)$. Then note that $E(X_t^k) < \infty$ implies $E(X_t^k \mid I_{t-1,i} = 1) < \infty$ for $k = 1, 2, 3$, $i = 1, 2, \ldots, m$. □

Theorem 5. *Let $\{X_t\}$ be a INMASC(1) process according to Definition 1. Let $\overline{X}$ be the sample mean of $\{X_t\}$; then the stochastic process $\{X_t\}$ is ergodic in the mean.*

Proof. Since $\gamma_X(k) \to 0, k \to \infty$.

From Theorem 7.1.1 in Brockwell and Davis [26], we get

$$\operatorname{Var}\left(\overline{X}_T\right) = E\left(\overline{X}_T - \mu_X\right)^2 \longrightarrow 0. \tag{10}$$

Then $\overline{X}_T$ converges in probability to μ_X. Therefore, the process $\{X_t\}$ is ergodic in the mean. □

Theorem 6. *Suppose $\{X_t\}$ is a INMASC(1) process; then*

$$P\left(\left|\widehat{\gamma}_X(k) - \gamma_X(k)\right| \geq \varepsilon\right) \xrightarrow{P} 0, \tag{11}$$

where $\widehat{\gamma}_X(k) := (1/T)\sum_{t=1}^{T-k}(X_{t+k} - \overline{X}_T)(X_t - \overline{X}_T)$.

The proof of Theorem 6 is similar to Theorem 4 given in Yu et al. [27]. It is easy to verify; we skip the details.

3. Estimation of Parameters

In this paper, we consider one method, namely, moment estimation. An advantage of the method is that it is simple and often produces good results. The estimation problem of INMASC(1) parameters is complex. In fact, for the INMASC(1) processes, the conditional distribution of the X_t given ε_{t-1} is the convolution of the distribution of the arrival process ε_t and one thinning operation $\alpha_i \circ \varepsilon_{t-1}$. On the other hand, there are too many unknown parameters of the model, such as λ, α_i, p_i, u_i, and σ_i^2, $i = 1, \ldots, m$, whereas the number of moment conditions is small.

Therefore we cannot estimate all the parameters unless additional assumptions are made. Then, we assume that the number of break point m is two and assume that the value of break point τ_i, $i = 1, \ldots, m$, and the mean of innovation λ are also known. Thus, here we estimate INMASC(1) model with two break points. Under these assumptions, all the parameters λ, p_i, u_i, and σ_i^2, $i = 1, 2, 3$, are known. We only need to estimate the autoregressive coefficients α_1, α_2, and α_3. Using the sample mean and sample covariance function, we can get the moment estimators via solving the following equations:

$$\begin{aligned}\widehat{\gamma}(0) &= \sum_{i=1}^{3} p_i \alpha_i \left[\alpha_i \left(u_i^2 + \sigma_i^2\right) + \left(1 - \alpha_i\right) u_i\right] \\ &\quad - \left(\sum_{i=1}^{3} p_i \alpha_i u_i\right)^2 + \lambda \\ \widehat{\gamma}(1) &= \sum_{i=1}^{3} p_i \alpha_i \left(u_i^2 + \sigma_i^2 - \lambda u_i\right) \\ \overline{X} &= \sum_{i=1}^{3} p_i \alpha_i u_i + \lambda.\end{aligned} \tag{12}$$

Table 1: Bias and mean square error for models A, B, and C.

Model	Parameter	Sample size 50	200	500	
A	α_1	0.0267	0.0097	0.0034	Bias
		(0.3948)	(0.0753)	(0.0453)	MSE
	α_2	0.0645	0.0115	0.0025	Bias
		(0.4731)	(0.1314)	(0.0376)	MSE
	α_3	0.0417	0.0083	0.0046	Bias
		(0.2908)	(0.0811)	(0.0342)	MSE
B	α_1	0.0335	0.0127	0.0036	Bias
		(0.4623)	(0.1803)	(0.0745)	MSE
	α_2	0.0297	0.0103	0.0054	Bias
		(0.2806)	(0.3449)	(0.0847)	MSE
	α_3	0.0251	0.0081	0.0024	Bias
		(0.3408)	(0.0372)	(0.0165)	MSE
C	α_1	0.0736	0.0178	0.0068	Bias
		(1.0435)	(0.3562)	(0.0357)	MSE
	α_2	0.0582	0.0215	0.0049	Bias
		(0.4127)	(0.0433)	(0.0212)	MSE
	α_2	0.0237	0.0081	0.0031	Bias
		(0.3205)	(0.0547)	(0.0274)	MSE

If you want to estimate all parameters, you can use GMM method based on probability generating functions introduced by BräKnnäK and Hall [28]. But they found covariance matrix of estimators depends on z and the orders besides the model parameters in a highly complex way. Thus we do not use this method here. In next section, simulations are provided to give insight into the finite sample behaviour of these estimators.

4. Simulation Study

Consider the following INMASC(1) model:

$$X_t = \begin{cases} \alpha_1 \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \varepsilon_{t-1} \leq \tau_1 \\ \alpha_2 \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \tau_1 < \varepsilon_{t-1} \leq \tau_2 \\ \alpha_3 \circ \varepsilon_{t-1} + \varepsilon_t, & \text{for } \tau_2 < \varepsilon_{t-1}, \end{cases} \tag{13}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. For fixed t, ε_t follows a Poisson distribution with mean λ.

The parameters values considered in this model are listed as follows:

(model A) $(\alpha_1, \alpha_2, \alpha_3) = (0.1, 0.1, 0.1)$, with $\tau_1 = 3$, $\tau_2 = 10$, $\lambda = 1$;

(model B) $(\alpha_1, \alpha_2, \alpha_3) = (0.2, 0.3, 0.1)$, with $\tau_1 = 8$, $\tau_2 = 17$, $\lambda = 10$;

(model C) $(\alpha_1, \alpha_2, \alpha_3) = (0.4, 0.3, 0.1)$, with $\tau_1 = 21$, $\tau_2 = 43$, $\lambda = 50$.

We use the above models to generate data and then use moment methods to estimate the parameters. We computed the empirical bias and the mean square error (MSE) based on 300 replications for each parameter combination. These values are reported within parenthesis in Table 1.

From the results in Table 1, we can see moment estimation is good estimation methods producing estimators whose bias and MSEs are small when the sample sizes are larger. In addition, this method is fast and easy to implement. It is perhaps not surprising that the MSEs are larger when these sample sizes are smaller. As to be expected, both the bias and the MSEs converge to zero with increasing sample size T.

5. Conclusion

Based on some limitations of the present count data models, a new INMA model is introduced to model structural changes. Expressions for mean, variance, and autocorrelation functions are given. Stationary and other basic statistical properties are also obtained. We derived moment estimators of the unknown parameters. Furthermore, we constructed several simulations to evaluate the performance of the estimators of model parameters.

Conflict of Interests

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 71201126), the Science Foundation of Ministry of Education of China (no. 12XJC910001), and the Fundamental Research Funds for the Central Universities (nos. JBK140211 and JBK120405).

References

[1] E. McKenzie, "Some simple models for discrete variate time series," *Water Resources Bulletin*, vol. 21, no. 4, pp. 635–644, 1985.

[2] E. McKenzie, "Autoregressive moving-average processes with negative-binomial and geometric marginal distributions," *Advances in Applied Probability*, vol. 18, no. 3, pp. 679–705, 1986.

[3] M. A. Al-Osh and A. A. Alzaid, "First order integer-valued autoregressive INAR(1) process," *Journal of Time Series Analysis*, vol. 8, no. 3, pp. 261–275, 1987.

[4] A. A. Alzaid and M. Al-Osh, "An integer-valued pth-order autoregressive structure INAR(p) process," *Journal of Applied Probability*, vol. 27, no. 2, pp. 314–324, 1990.

[5] J. G. Du and Y. Li, "The integer-valued autoregressive (INAR(p)) model," *Journal of Time Series Analysis*, vol. 12, no. 2, pp. 129–142, 1991.

[6] M. Al-Osh and A. A. Alzaid, "Integer-valued moving average (INMA) process," *Statistical Papers*, vol. 29, no. 4, pp. 281–300, 1988.

[7] E. McKenzie, "Some ARMA models for dependent sequences of Poisson counts," *Advances in Applied Probability*, vol. 20, no. 4, pp. 822–835, 1988.

[8] R. Ferland, A. Latour, and D. Oraichi, "Integer-valued GARCH process," *Journal of Time Series Analysis*, vol. 27, no. 6, pp. 923–942, 2006.

[9] K. Fokianos and R. Fried, "Interventions in INGARCH processes," *Journal of Time Series Analysis*, vol. 31, no. 3, pp. 210–225, 2010.

[10] C. H. Weiß, "The INARCH(1) model for overdispersed time series of counts," *Communications in Statistics: Simulation and Computation*, vol. 39, no. 6, pp. 1269–1291, 2010.

[11] F. Zhu, "A negative binomial integer-valued GARCH model," *Journal of Time Series Analysis*, vol. 32, no. 1, pp. 54–67, 2011.

[12] F. Zhu, "Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models," *Journal of Mathematical Analysis and Applications*, vol. 389, no. 1, pp. 58–71, 2012.

[13] F. Zhu and D. Wang, "Diagnostic checking integer-valued ARCH(p) models using conditional residual autocorrelations," *Computational Statistics and Data Analysis*, vol. 54, no. 2, pp. 496–508, 2010.

[14] L. Györfi, M. Ispány, G. Pap, and K. Varga, "Poisson limit of an inhomogeneous nearly critical INAR(1) model," *Acta Universitatis Szegediensis*, vol. 73, no. 3-4, pp. 789–815, 2007.

[15] H. S. Bakouch and M. M. Ristić, "Zero truncated Poisson integer-valued AR(1) model," *Metrika*, vol. 72, no. 2, pp. 265–280, 2010.

[16] M. Kachour and J. F. Yao, "First-order rounded integer-valued autoregressive ((RINAR(1)) process," *Journal of Time Series Analysis*, vol. 30, no. 4, pp. 417–448, 2009.

[17] H.-Y. Kim and Y. Park, "A non-stationary integer-valued autoregressive model," *Statistical Papers*, vol. 49, no. 3, pp. 485–502, 2008.

[18] H. Zheng, I. V. Basawa, and S. Datta, "First-order random coefficient integer-valued autoregressive processes," *Journal of Statistical Planning and Inference*, vol. 137, no. 1, pp. 212–229, 2007.

[19] D. Gomes and L. Canto e Castro, "Generalized integer-valued random coefficient for a first order structure autoregressive (RCINAR) process," *Journal of Statistical Planning and Inference*, vol. 139, no. 12, pp. 4088–4097, 2009.

[20] M. M. Ristić, A. S. Nastić, K. Jayakumar, and H. S. Bakouch, "A bivariate INAR(1) time series model with geometric marginals," *Applied Mathematics Letters*, vol. 25, no. 3, pp. 481–485, 2012.

[21] X. Pedeli and D. Karlis, "A bivariate INAR(1) process with application," *Statistical Modelling*, vol. 11, no. 4, pp. 325–349, 2011.

[22] P. Thyregod, J. Carstensen, H. Madsen, and K. Arnbjerg-Nielsen, "Integer valued autoregressive models for tipping bucket rainfall measurements," *Environmetrics*, vol. 10, no. 4, pp. 395–411, 1999.

[23] M. Monteiro, M. G. Scotto, and I. Pereira, "Integer-valued self-exciting threshold autoregressive processes," *Communications in Statistics: Theory and Methods*, vol. 41, no. 15, pp. 2717–2737, 2012.

[24] Š. Hudecová, "Structural changes in autoregressive models for binary time series," *Journal of Statistical Planning and Inference*, vol. 143, no. 10, pp. 1744–1752, 2013.

[25] A. DasGupta, *Asymptotic Theory of Statistics and Probability*, Springer, New York, NY, USA, 2008.

[26] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer, New York, NY, USA, 1991.

[27] K. Yu, D. Shi, and H. Zou, "The random coefficient discretevalued time series model," *Statistical Research*, vol. 28, no. 4, pp. 106–112, 2011.

[28] K. BräKnnäK and A. Hall, "Estimation in integer-valued moving average models," *Applied Stochastic Models in Business and Industry*, vol. 17, no. 3, pp. 277–291, 2001.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 358601, 10 pages
http://dx.doi.org/10.1155/2014/358601

Research Article

On H_∞ Fault Estimator Design for Linear Discrete Time-Varying Systems under Unreliable Communication Link

Yueyang Li,[1] Shuai Liu,[2,3] Zhonghua Wang,[1] and Xuehua Yan[1]

[1] *School of Electrical Engineering, University of Jinan, Jinan 250022, China*
[2] *School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798*
[3] *School of Control Science and Engineering, Shandong University, Jinan 250061, China*

Correspondence should be addressed to Yueyang Li; cse_liyy@ujn.edu.cn

Received 24 April 2014; Revised 3 July 2014; Accepted 6 July 2014; Published 17 July 2014

Academic Editor: Weihai Zhang

This paper investigates the H_∞ fixed-lag fault estimator design for linear discrete time-varying (LDTV) systems with intermittent measurements, which is described by a Bernoulli distributed random variable. Through constructing a novel partially equivalent dynamic system, the fault estimator design is converted into a deterministic quadratic minimization problem. By applying the innovation reorganization technique and the projection formula in Krein space, a necessary and sufficient condition is obtained for the existence of the estimator. The parameter matrices of the estimator are derived by recursively solving two standard Riccati equations. An illustrative example is provided to show the effectiveness and applicability of the proposed algorithm.

1. Introduction

To satisfy the growing demands for reliability and safety in control systems, more and more research efforts are made for model-based fault detection (FD) during the past decades; see [1–6] and references therein. Basically, the FD issue concerns designing a fault detection filter (FDF) for generating a residual signal such that the sensitivity of residual to fault is intensified by enhancing the robustness to the disturbance. In reviewing of the development of FD, with the aid of linear matrix inequality (LMI) techniques, much attention has been paid to linear time-invariant (LTI) systems with various characteristics such as time-delay, model inaccuracy, time-dependent switching mode, and uncertain observations; see [7–10] and related works. Recently, some contributions are devoted to linear time-varying (LTV) systems since most practical industrial processes can be represented or well approximated by time-varying dynamics [11]. For example, in [11, 12], unified optimal solutions are derived in the framework of maximizing H_-/H_∞ and H_∞/H_∞ FD performance indices for linear continuous time-varying (LCTV) and linear discrete time-varying (LDTV) systems, respectively. In [13–15], the H_∞ filtering based fault estimation methods are proposed for LDTV systems in virtue of the Krein space based reorganized innovation analysis and projection theory in the background of [16–21].

On another front line, with the rapid progress of networked control systems and distributed sensor/actuator systems, the packet dropout caused by sensor gain reductions may happen when transmitting information under unreliable links. The random uncertainty introduced by packet dropouts evidently deteriorates the performance of the FDF. Many contributions are dedicated to FD issue for systems with incomplete measurements by employing the LMI formulated H_∞ fault estimation approach over infinite horizon; we refer to [22–26] and references therein. For finite-horizon case, an H_∞ fault estimator for LDTV systems with multiple packet dropouts is designed in [27] based on the stochastic bounded real lemma (BRL), while a two-objective optimization FD method for LDTV systems with intermittent observations is addressed in [28]. Unfortunately, if there is no sensor fault in the measurement channel or the data packet is not time-stamped, the algorithms proposed in [27, 28] will fail. This indicates that research on FD problem for LDTV systems subject to intermittent measurements has not been fully investigated yet, which is the main motivation of the present study.

To overcome the drawbacks in the existing results, a novel fault estimator design method for LDTV systems with intermittent observations is proposed. The contribution of this paper consists in three aspects as follows:

(1) an H_∞ fixed-lag fault estimator design problem is formulated by establishing an equivalent system and its corresponding deterministic performance index;

(2) by employing the reorganized innovation analysis approach and the projection theory in Krein space, a necessary and sufficient condition of the existence of the estimator is derived;

(3) a recursive fault estimation algorithm is proposed, which is apt to be online applied for finite-horizon.

The rest of the content is organized as follows. Section 2 provides the formulation of the concerned problem. Section 3 presents our main results of designing the fault estimator. The proposed approach is applied to a time-varying model to illustrate its applicability in Section 4. Finally, the paper is ending with some conclusions.

Notations. Throughout this paper, vectors in the Krein space are represented by **boldface** letters, and vectors in the Euclidean space are denoted by normal letters. For a matrix X, X^{T} and X^{-1} stand for the transpose and inverse of X, respectively. $X > 0$ ($X < 0$) denotes X is positive (negative) definite. R^n means the set of n-dimensional real vectors. I and 0 denote identity matrix and zero matrix with appropriate dimensions, respectively. $\mathrm{E}\{\vartheta(k)\}$ means the mathematical expectation of $\vartheta(k)$. $\vartheta(k) \in l_2[0,N]$ means $\sum_{k=0}^{N}\vartheta^{\mathrm{T}}(k)\vartheta(k) < \infty$, where N is a positive integer. The symbol $\mathscr{L}\{\{\vartheta(i)\}_{i=j}^{k}\}$ represents the linear space spanned by the sequence $\vartheta(k)$ taking values in the time interval $[j,k]$. $\mathrm{Prob}\{\Upsilon\}$ denotes the occurrence probability of the event "Υ". δ_{ij} represents the Kronecker delta function, which is equal to unity for $i = j$ and zero for $i \neq j$. $\mathrm{diag}\{S_1, S_2, \ldots, S_n\}$ means a block diagonal matrix with diagonal blocks $S_1, S_2, \ldots, S_n$.

2. Problem Formulation and Preliminaries

Consider the following LDTV system:

$$\begin{gathered} x(k+1) = A(k)x(k) + B_f(k)f(k) + D(k)d(k), \\ y(k) = \theta(k)C(k)x(k) + v(k), \\ x(0) = x_0, \end{gathered} \tag{1}$$

where $x(k) \in R^n$, $y(k) \in R^q$, $d(k) \in R^{n_d}$, $v(k) \in R^{n_v}$, and $f(k) \in R^{n_f}$ denote the state, sensor measurement, process noise, observation noise, and fault, respectively. $f(k)$, $d(k)$, and $v(k)$ belong to $l_2[0,N]$. $A(k)$, $B_f(k)$, $C(k)$, and $D(k)$ are known time-varying matrices with appropriate dimensions. $\theta(k)$ is a Bernoulli distributed binary stochastic variable to describe the measurement packet dropouts, which satisfies

$$\begin{gathered} \mathrm{Prob}\{\theta(k) = 1\} = \mathrm{E}\{\theta(k)\} = \rho, \\ \mathrm{Prob}\{\theta(k) = 0\} = 1 - \mathrm{E}\{\theta(k)\} = 1 - \rho, \end{gathered} \tag{2}$$

with ρ being a known constant. The value of ρ can be obtained by empirical observations, experimentations, and statistical analysis [29].

The main purpose of this paper is as follows: given a prescribed disturbance attenuation level γ, by collecting the observations $y(0), \ldots, y(k)$, find $\check{f}(k-l \mid k)$ as a suitable estimation of the fault signal $f(k)$ such that the following l-step delayed H_∞ performance index is fulfilled with l being a positive integer:

$$\begin{aligned} \sup_{(x_0,f_k,d_k,v_k)\neq 0} \; & \mathrm{E}\left\{\sum_{k=l}^{N}\left(\check{f}(k-l \mid k) - f(k-l)\right)^{\mathrm{T}} \right. \\ & \left. \times \left(\check{f}(k-l \mid k) - f(k-l)\right)\right\} \\ & \times \left(x_0^{\mathrm{T}}P_0^{-1}x_0 + \sum_{k=0}^{N}f^{\mathrm{T}}(k)f(k)\right. \\ & \left. + \sum_{k=0}^{N-1}d^{\mathrm{T}}(k)d(k) + \sum_{k=0}^{N}v^{\mathrm{T}}(k)v(k)\right)^{-1} < \gamma^2, \end{aligned} \tag{3}$$

where $f_k = [f^{\mathrm{T}}(0)\cdots f^{\mathrm{T}}(k)]^{\mathrm{T}}$, $d_k = [d^{\mathrm{T}}(0)\cdots d^{\mathrm{T}}(k)]^{\mathrm{T}}$, $v_k = [v^{\mathrm{T}}(0)\cdots v^{\mathrm{T}}(k)]^{\mathrm{T}}$.

Due to the fact that the denominator of the left side of (3) is positive, (3) can be rewritten as

$$\begin{aligned} J_0 = {} & x_0^{\mathrm{T}}P_0^{-1}x_0 + \sum_{k=0}^{N}f^{\mathrm{T}}(k)f(k) \\ & + \sum_{k=0}^{N-1}d^{\mathrm{T}}(k)d(k) + \sum_{k=0}^{N}v^{\mathrm{T}}(k)v(k) \\ & - \mathrm{E}\left\{\gamma^{-2}\sum_{k=l}^{N}v_s^{\mathrm{T}}(k)v_s(k)\right\} > 0, \end{aligned} \tag{4}$$

where $v_s(k) = \check{f}(k-l \mid k) - f(k-l)$. Consequently, according to [30], the H_∞ fixed-lag fault estimation problem can be restated as follows: given a constant $\gamma > 0$, design an estimator in the following way:

$$\check{f} = \Psi(y) = \overline{\Psi}(f,d,v), \tag{5}$$

where Ψ denotes a stable operator which generates a bounded operator $\overline{\Psi}$ mapping from f, d, v to $\check{f}$, such that the indefinite cost function (4) has a positive minimum with respect to f, d, and v.

Remark 1. In the existing results, for example, [22–28], the Bernoulli distributed random variables are introduced to describe the packet dropping or finite step measurement time-delay phenomenon. It is noteworthy that the designed estimators only depend on the probability, that is, ρ, rather than $\theta(k)$. This indicates that the desired fault estimator does not require the time stamp of the data packet.

Remark 2. Notice that when $y(k)$ is affected by the so-called sensor fault with the following form:

$$y(k) = \theta(k) C(k) x(k) + D_f(k) f(k) + v(k), \quad (6)$$

the existing BRL based H_∞ fault estimation algorithm in [27] is applicable in a "*filter*" manner. In the case that $D_f(k) = 0$, the estimator is supposed to be designed as a "*smoother*" with the proposed performance index (4). In this scenario, the methodology in [27] may induce computational burden via state augmentation approach and the gain matrices of the estimator are arduous to be derived due to some coupled product terms. In what follows, a Krein space based fault estimator design scheme will be addressed to overcome the aforementioned defects.

3. Main Results

In this section, inspired by [31, 32], an equivalent Krein space stochastic system and a corresponding H_∞ performance index are first introduced. Then, by exploiting the reorganized innovation analysis and the projection theory in Krein space, the H_∞ fault estimator is derived.

3.1. Krein Space Model Design. Before we proceed, we would like to propose the following lemma to construct an auxiliary stochastic system in Krein space.

Lemma 3. *Given a scalar $\gamma > 0$ and an integer $l > 0$, then the H_∞ performance (4) is fulfilled if and only if there exists a fault estimator $\check{f}(k-l \mid k)$ such that the following inequality holds:*

$$\begin{aligned} J = {} & x_0^T P_0^{-1} x_0 + \sum_{k=0}^{N} f^T(k) f(k) + \sum_{k=0}^{N} v_0^T(k) v_0(k) \\ & + \sum_{k=0}^{N-1} d^T(k) d(k) + \sum_{k=0}^{N} v_z^T(k) v_z(k) \\ & - \gamma^{-2} \sum_{k=l}^{N} v_s^T(k) v_s(k) > 0, \end{aligned} \quad (7)$$

subject to the following dynamic constraints:

$$\begin{aligned} x(k+1) &= A(k) x(k) + B_f(k) f(k) + D(k) d(k), \\ y_0(k) &= \rho C(k) x(k) + v_0(k), \\ y_z(k) &= \sqrt{\rho(1-\rho)} C(k) x(k) + v_z(k), \\ \check{f}(k-l \mid k) &= f(k-l) + v_s(k), \\ x(0) &= x_0, \end{aligned} \quad (8)$$

where $y_0(k)$ and $y_z(k)$ are the fictitious observations with their corresponding observation noises $v_0(k)$ and $v_z(k)$, respectively. The instantaneous value of y_0 at each time instant k is equal to $y(k)$ along with $y_z(k) \equiv 0$.

Proof. Consider the following.

Necessity. From (1), the state transition matrix Φ is defined as

$$\Phi(k, j) = \begin{cases} A(k-1) \cdots A(j), & 0 < k < j, \\ I, & k = j; \end{cases} \quad (9)$$

hence, we have

$$\begin{aligned} x(k) = {} & \Phi(k, 0) x_0 + \sum_{i=0}^{k-1} \Phi(k, i+1) B_f(i) f(i) \\ & + \sum_{i=0}^{k-1} \Phi(k, i+1) D(i) d(i). \end{aligned} \quad (10)$$

Define

$$\begin{aligned} y_k &= \left[y^T(0) \cdots y^T(k) \right]^T, \\ v_{s,k} &= \left[v_s^T(0) \cdots v_s^T(k) \right]^T, \\ \check{f}_k &= \left[\check{f}^T(0 \mid l) \cdots \check{f}^T(k-l \mid k) \right]^T. \end{aligned} \quad (11)$$

Then, in view of (10), we have

$$\begin{aligned} y_N &= \Xi(k) G_x x_0 + \Xi(k) G_f f_N + \Xi(k) G_d d_N + v_N, \\ \check{f}_N &= f_{N-l} + v_{s,N}, \end{aligned} \quad (12)$$

where

$$\begin{aligned} \Xi(k) &= \operatorname{diag}\{\theta(1), \ldots, \theta(k)\}, \\ G_f(k, i) &= C(k) \Phi(k, i+1) B_f(i), \\ G_d(k, i) &= C(k) \Phi(k, i+1) D(i), \\ G_x &= \begin{bmatrix} C(0)\Phi(0,0) \\ C(1)\Phi(1,0) \\ \vdots \\ C(N)\Phi(N,0) \end{bmatrix}, \\ G_f &= \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ G_f(1,0) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \\ G_f(N,0) & G_f(N,1) & \cdots & 0 \end{bmatrix}, \\ G_d &= \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ G_d(1,0) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \\ G_d(N,0) & G_d(N,1) & \cdots & 0 \end{bmatrix}. \end{aligned} \quad (13)$$

Thus, by substituting (12) into (4) and taking (2) into consideration, we have

$$\begin{aligned}
J_0 = \mathrm{E}\Bigg\{ & x_0^{\mathrm{T}} P_0^{-1} x_0 + \sum_{k=0}^{N} f^{\mathrm{T}}(k) f(k) + \sum_{k=0}^{N-1} d^{\mathrm{T}}(k) d(k) \\
& - \left(y_N - \Xi(k) G_x x_0 - \Xi(k) G_f f_N - \Xi(k) G_d d_N\right)^{\mathrm{T}} \\
& \times \left(y_N - \Xi(k) G_x x_0 - \Xi(k) G_f f_N - \Xi(k) G_d d_N\right) \\
& - \gamma^{-2} \sum_{k=l}^{N} \left(\check{f}(k-l \mid k) - f(k-l)\right)^{\mathrm{T}} \\
& \times \left(\check{f}(k-l \mid k) - f(k-l)\right) \Bigg\} \\
= {} & x_0^{\mathrm{T}} P_0^{-1} x_0 + \sum_{k=0}^{N} f^{\mathrm{T}}(k) f(k) + \sum_{k=0}^{N-1} d^{\mathrm{T}}(k) d(k) \\
& + \left(y_{0,N} - \overline{\Xi} G_x x_0 - \overline{\Xi} G_f f_N - \overline{\Xi} G_d d_N\right)^{\mathrm{T}} \\
& \times \left(y_{0,N} - \overline{\Xi} G_x x_0 - \overline{\Xi} G_f f_N - \overline{\Xi} G_d d_N\right) \\
& + \left(y_{z,N} - \widetilde{\Xi} G_x x_0 - \widetilde{\Xi} G_f f_N - \widetilde{\Xi} G_d d_N\right)^{\mathrm{T}} \\
& \times \left(y_{z,N} - \widetilde{\Xi} G_x x_0 - \widetilde{\Xi} G_f f_N - \widetilde{\Xi} G_d d_N\right) \\
& - \gamma^{-2} \sum_{k=l}^{N} \left(\check{f}(k-l \mid k) - f(k-l)\right)^{\mathrm{T}} \\
& \times \left(\check{f}(k-l \mid k) - f(k-l)\right),
\end{aligned} \tag{14}$$

where

$$\begin{gathered}
y_{0,k} = \left[y_0^{\mathrm{T}}(0) \cdots y_0^{\mathrm{T}}(k)\right]^{\mathrm{T}}, \qquad y_{z,k} = \left[y_z^{\mathrm{T}}(0) \cdots y_z^{\mathrm{T}}(k)\right]^{\mathrm{T}}, \\
y_0(i) = y(i), \quad y_z(i) = 0, \quad (i = 0, \dots, k), \\
\overline{\Xi} = \rho I, \qquad \widetilde{\Xi} = \sqrt{\rho(1-\rho)}\, I.
\end{gathered} \tag{15}$$

Therefore, if the H_∞ performance index (4) is satisfied, then, following the same line with the correlation between (1) and (4), we have $J > 0$ subject to the dynamics (8) over x_0, f_k, and d_k.

Sufficiency. For (8), since the value of $y_0(k)$ is equivalent to $y(k)$ and $y_z(k) \equiv 0$, in light of (14), it is easy to find out that for a given constant $\gamma > 0$ and an integer $l > 0$, $J_0 = J$, which indicates that if $J > 0$ holds, then the H_∞ performance (4) is satisfied. Combing the sufficiency and necessity part, the proof is complete. □

In virtue of Lemma 3, the auxiliary performance index J in (7) can be converted into the following compact form:

$$J = \begin{bmatrix} x_0 \\ d_N \\ f_N \\ v_{a,N} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & Q_{a,N} \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ d_N \\ f_N \\ v_{a,N} \end{bmatrix}, \tag{16}$$

where

$$v_a(k) = \begin{cases} v_1(k) = \begin{bmatrix} v_0(k) \\ v_z(k) \end{bmatrix}, & 0 \le k < l, \\ v_2(k) = \begin{bmatrix} v_0(k) \\ v_z(k) \\ v_s(k) \end{bmatrix}, & k \ge l, \end{cases} \tag{17}$$

$$\begin{gathered}
v_{a,N} = \left[v_a^{\mathrm{T}}(0) \cdots v_a^{\mathrm{T}}(N)\right]^{\mathrm{T}}, \\
Q_a(k) = \begin{cases} Q_{v_1}(k) = \operatorname{diag}\{I, I\}, & 0 \le k < l, \\ Q_{v_2}(k) = \operatorname{diag}\left\{I, I, -\gamma^2 I\right\}, & k \ge l, \end{cases} \\
Q_{a,N} = \operatorname{diag}\{Q_a(0), \dots, Q_a(N)\}.
\end{gathered} \tag{18}$$

From (8) and (17), we have

$$\begin{gathered}
y_f(k) = \begin{bmatrix} y(k) \\ y_z(k) \end{bmatrix} = C_1(k) x(k) + v_1(k), \\
y_a(k) = \begin{cases} y_f(k), & 0 \le k < l, \\ \begin{bmatrix} y_f(k) \\ \check{f}(k-l \mid k) \end{bmatrix} \\ \quad = C_2(k) x(k) \\ \qquad + H f(k-l) + v_2(k), & k \ge l, \end{cases}
\end{gathered} \tag{19}$$

where

$$\begin{gathered}
C_1(k) = \begin{bmatrix} \rho C(k) \\ \sqrt{\rho(1-\rho)}\, C(k) \end{bmatrix}, \qquad C_2(k) = \begin{bmatrix} C_1(k) \\ 0 \end{bmatrix}, \\
H = \begin{bmatrix} 0 & 0 & I \end{bmatrix}^{\mathrm{T}}.
\end{gathered} \tag{20}$$

Thus, according to [20, 21], we introduce the following Krein space system associated with (8), (16), (18), and (19):

$$\begin{gathered}
\mathbf{x}(k+1) = A(k)\mathbf{x}(k) + B_f(k)\mathbf{f}(k) + D(k)\mathbf{d}(k), \\
\mathbf{y}_a(k) = \begin{cases} \mathbf{y}_f(k) = C_1(k)\mathbf{x}(k) + \mathbf{v}_1(k), & 0 \le k < l, \\ \begin{bmatrix} \mathbf{y}_f(k) \\ \check{\mathbf{f}}(k-l \mid k) \end{bmatrix} \\ \quad = C_2(k)\mathbf{x}(k) + H\mathbf{f}(k-l) + \mathbf{v}_2(k), & k \ge l, \end{cases} \\
\mathbf{x}(0) = \mathbf{x}_0,
\end{gathered} \tag{21}$$

where $\mathbf{x}_0(i)$, $\mathbf{d}(i)$, $\mathbf{f}(i)$, $\mathbf{v}_1(i)$, and $\mathbf{v}_2(i)$ are uncorrelated white noises in Krein space satisfying

$$\left\langle \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{d}(i) \\ \mathbf{f}(i) \\ \mathbf{v}_a(i) \end{bmatrix}, \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{d}(i) \\ \mathbf{f}(i) \\ \mathbf{v}_a(i) \end{bmatrix} \right\rangle = \begin{bmatrix} I\delta_{ij} & 0 & 0 & 0 \\ 0 & I\delta_{ij} & 0 & 0 \\ 0 & 0 & I\delta_{ij} & 0 \\ 0 & 0 & 0 & Q_a(i)\delta_{ij} \end{bmatrix}, \tag{22}$$

$$\mathbf{v}_a(k) = \begin{cases} \mathbf{v}_1(k) = \begin{bmatrix} \mathbf{v}_0(k) \\ \mathbf{v}_z(k) \end{bmatrix}, & 0 \le k < l, \\ \mathbf{v}_2(k) = \begin{bmatrix} \mathbf{v}_0(k) \\ \mathbf{v}_z(k) \\ \mathbf{v}_s(k) \end{bmatrix}, & k \ge l, \end{cases} \tag{23}$$

with $\mathbf{v}_0(k)$, $\mathbf{v}_z(k)$, and $\mathbf{v}_s(k)$ being fictitious noise in Krein space corresponding to (17).

Consequently, on the basis of Lemma 4.2.1 in [20], we have the following lemma.

Lemma 4. *For* (8), *given a scalar $\gamma > 0$ and an integer $l > 0$, then the H_∞ performance* (7) *has a minimum over x_0, f, d if and only if $Q_a(k)$ and $Q_w(k)$ have the same inertia, where $Q_w(k) = \langle \mathbf{w}(k), \mathbf{w}(k) \rangle$ is the covariance matrix of innovation sequence $\mathbf{w}(k)$ given by*

$$\mathbf{w}(k) = \mathbf{y}_a(k) - \hat{\mathbf{y}}_a(k), \tag{24}$$

where $\hat{\mathbf{y}}_a(k)$ is the projection of $\mathbf{y}_a(k)$ onto $\mathscr{L}\{\{\mathbf{y}_a(j)\}_{j=0}^{k-1}\}$. Furthermore, the minimum value of J is

$$\begin{aligned} J_{\min} = {} & \sum_{k=0}^{l-1} \left[y_f(k) - C_1(k)\hat{x}(k) \right]^T \\ & \times Q_w^{-1}(k) \left[y_f(k) - C_1(k)\hat{x}(k) \right] \\ & + \sum_{k=l}^{N} \begin{bmatrix} y_f(k) - C_1(k)\hat{x}(k) \\ \check{f}(k-l \mid k) - \check{f}(k-l \mid k-1) \end{bmatrix}^T \\ & \times Q_w^{-1}(k) \begin{bmatrix} y_f(k) - C_1(k)\hat{x}(k) \\ \check{f}(k-l \mid k) - \hat{f}(k-l \mid k-1) \end{bmatrix}, \end{aligned} \tag{25}$$

where $\hat{x}(k)$ and $\hat{f}(k-l \mid k-1)$ are, respectively, calculated from the Krein space projections of $\mathbf{x}(k)$ and $\mathbf{f}(k-l)$ onto $\mathscr{L}\{\{\mathbf{y}_a(j)\}_{j=0}^{k-1}\}$.

Remark 5. According to Lemmas 3 and 4, the purpose of establishing the dynamic model (8) associated with (7) is to derive a positive minimum of the cost function (4) by applying the projection theory in Krein space. Notice that although the measurement $\{y(k)\}_{k=0}^{N}$ is a substantially stochastic sequence, the instantaneous values of $y(k)$ and $\check{f}(k-l \mid k)$ at each instant are available for the estimator. Thus, the equivalent cost function (7) and its corresponding dynamic constraint are constructed in a "*conditional expectation*" sense by gathering up $\{y(k)\}_{k=0}^{N}$ (cf. (14) in the proof of Lemma 3).

3.2. Kalman Filtering in Krein Space. From the analysis above, the key step to achieve our goal is to find a suitable $\hat{x}(k)$ and $\hat{f}(k-l \mid k-1)$. To this end, let

$$\mathbf{y}_1(k) = \mathbf{y}_f(k), \qquad \mathbf{y}_2(k) = \begin{bmatrix} \mathbf{y}_f(k) \\ \check{\mathbf{f}}(k \mid k+l) \end{bmatrix}; \tag{26}$$

then

$$\begin{aligned} \mathbf{y}_1(k-l+i) &= C_1(k-l+i)\mathbf{x}(k-l+i) + \tilde{\mathbf{v}}_1(k-l+i), \\ & \qquad i = 1, \ldots, l, \\ \mathbf{y}_2(i) &= C_2(i)\mathbf{x}(i) + H\mathbf{f}(i) + \tilde{\mathbf{v}}_2(i), \quad i = 0, \ldots, k-l, \end{aligned} \tag{27}$$

where $\tilde{\mathbf{v}}_1(k) = \mathbf{v}_1(k)$ and $\tilde{\mathbf{v}}_2 = \begin{bmatrix} \mathbf{v}_1^{\mathrm{T}}(k) & \mathbf{v}_s^{\mathrm{T}}(k+l) \end{bmatrix}^{\mathrm{T}}$ are zero-mean white noises with the following covariance matrices, respectively:

$$Q_{\tilde{\mathbf{v}}_1}(k) = \operatorname{diag}\{I, I\}, \qquad Q_{\tilde{\mathbf{v}}_2}(k) = \operatorname{diag}\left\{I, I, -\gamma^2 I\right\}. \tag{28}$$

It is easy to check out that $\{\mathbf{y}_2(0), \ldots, \mathbf{y}_2(k-l); \mathbf{y}_1(k-l+1), \ldots, \mathbf{y}_1(k)\}$ span the same linear space as $\mathscr{L}\{\{\mathbf{y}_a(j)\}_{j=0}^{k}\}$.

To proceed, the following definition is introduced.

Definition 6 (see [32]). For $t > k-l$, the estimator $\hat{\eta}(t,1)$ is the optimal estimation of $\eta(t)$ on the observation $\mathscr{L}\{\{\mathbf{y}_2(t)\}_{t=0}^{k-l-1}; \{\mathbf{y}_1(t)\}_{t=k-l}^{t=k-1}\}$. For $0 < t \le k-l$, the estimator $\hat{\eta}(t,2)$ is the optimal estimation of $\eta(t)$ on the observation $\mathscr{L}\{\{\mathbf{y}_2(t)\}_{t=0}^{t=k-1}\}$.

In accordance with (24), the innovation sequence is defined as follows:

$$\begin{aligned} \mathbf{w}_1(k-l+i) &= C_1(k-l+i)\mathbf{e}_1(k-l+i) \\ & \quad + \tilde{\mathbf{v}}_1(k-l+i), \quad i = 0, \ldots, l, \\ \mathbf{w}_2(i) &= C_2(i)\mathbf{e}_2(i) + Hf(i) + \tilde{\mathbf{v}}_2(i), \quad i = 0, \ldots, k-l, \end{aligned} \tag{29}$$

where

$$\begin{aligned} \mathbf{e}_1(k-l+i) &= \mathbf{x}(k-l+i) - \hat{\mathbf{x}}(k-l+i, 1), \quad i = 0, \ldots, l, \\ \mathbf{e}_2(i) &= \mathbf{x}(i) - \hat{\mathbf{x}}(i, 2), \quad i = 0, \ldots, k-l, \end{aligned} \tag{30}$$

with the corresponding covariance matrices given as

$$\begin{aligned} P_1(k-l+i) &= \left\langle \mathbf{e}_1(k-l+i), \mathbf{e}_1(k-l+i) \right\rangle, \quad i = 0, \ldots, l, \\ P_2(i) &= \left\langle \mathbf{e}_2(i), \mathbf{e}_2(i) \right\rangle, \quad i = 0, \ldots, k-l. \end{aligned} \tag{31}$$

In light of Lemma 2.2.1 in [20], the innovation sequences $\mathscr{L}\{\{\mathbf{w}_2(t)\}_{t=0}^{k-l-1};\{\mathbf{w}_1(t)\}_{t=k-l}^{t=k-1}\}$ are uncorrelated white noises and span the same linear space as $\mathscr{L}\{\{\mathbf{y}_a(j)\}_{j=0}^{k}\}$.

For deriving $\hat{\mathbf{x}}(k-l,2)$ $(k=l+1,l+2,\ldots)$, applying the Krein space based projection formula in [21] by taking (21) and (22) into account, we have that

$$\begin{aligned}\hat{\mathbf{x}}(k-l,2) &= A(k-l-1)\hat{\mathbf{x}}(k-l-1,2)\\&\quad+\langle\mathbf{x}(k-l),\mathbf{w}_2(k-l-1)\rangle\\&\quad\times\langle\mathbf{w}_2(k-l-1),\mathbf{w}_2(k-l-1)\rangle^{-1}\\&\quad\times\mathbf{w}_2(k-l-1)\\&= A(k-l-1)\hat{\mathbf{x}}(k-l-1,2)\\&\quad+K_2(k-l-1)\mathbf{w}_2(k-l-1),\\ \hat{\mathbf{x}}(0)&=0,\end{aligned}\tag{32}$$

where

$$\begin{aligned}K_2(k-l-1) = \big(&A(k-l-1)P_2(k-l-1)C_2^{\mathrm{T}}(k-l-1)\\&+B_f(k-l-1)H\big)Q_2^{-1}(k-l-1),\end{aligned}\tag{33}$$

with $Q_2(k-l-1)=C_2(k-l-1)P_2(k-l-1)C_2^{\mathrm{T}}(k-l-1)+HH^{\mathrm{T}}+Q_{\tilde{\mathbf{v}}_2}(k-l-1)$.

In addition, following the definition of $P_2(i)$ and (32), $P_2(i)$ $(i=0,1,\ldots,k-l-1)$ is the solution to the following standard Riccati equation:

$$\begin{aligned}P_2(i+1) &= A(i)P_2(i)A^{\mathrm{T}}(i)+B_f(i)B_f^{\mathrm{T}}(i)\\&\quad+D(i)D^{\mathrm{T}}(i)-K_2(i)Q_2^{-1}(i)K_2^{\mathrm{T}}(i),\\P_2(0)&=P_0.\end{aligned}\tag{34}$$

For calculating $\hat{\mathbf{x}}(k-l+i,1)$ $(i=1,\ldots,l)$ with the initial condition $\hat{\mathbf{x}}(k-l,1)=\hat{\mathbf{x}}(k-l,2)$, we apply the projection formula once again such that

$$\begin{aligned}&\hat{\mathbf{x}}(k-l+i+1,1)\\&\quad= A(k-l+i)\hat{\mathbf{x}}(k-l+i,1)\\&\qquad+A(k-l+i)\langle\mathbf{x}(k-l+i),\mathbf{w}_1(k-l+i)\rangle\\&\qquad\times\langle\mathbf{w}_1(k-l+i),\mathbf{w}_1(k-l+i)\rangle^{-1}\mathbf{w}_1(k-l+i)\\&\quad= A(k-l+i)\hat{\mathbf{x}}(k-l+i,1)+K_1(k-l+i)\mathbf{w}_1(k-l+i),\end{aligned}\tag{35}$$

where

$$\begin{aligned}K_1(k-l-1) = &A(k-l+i)P_1(k-l+i)\\&\times C_1^{\mathrm{T}}(k-l+i)Q_1^{-1}(k-l+i),\end{aligned}\tag{36}$$

with $Q_1(k-l+i)=C_1(k-l+i)P_1(k-l+i)C_1^{\mathrm{T}}(k-l+i)+Q_{\tilde{\mathbf{v}}_1}(k-l+i)$, and $P_1(k-l+i)$ is computed recursively in the following form:

$$\begin{aligned}&P_1(k-l+i+1)\\&\quad= A(k-l+i)P_1(k-l+i)A^{\mathrm{T}}(k-l+i)\\&\qquad+B_f(k-l+i)B_f^{\mathrm{T}}(k-l+i)\\&\qquad+D(k-l+i)D^{\mathrm{T}}(k-l+i)\\&\qquad-K_1(k-l+i)Q_2^{-1}(k-l+i)K_1^{\mathrm{T}}(k-l+i),\\&\qquad\qquad P_1(k-l)=P_2(k-l).\end{aligned}\tag{37}$$

Similarly, the projection formula is reutilized to compute $\hat{\mathbf{f}}(k-l\mid k-1)$; that is,

$$\begin{aligned}&\hat{\mathbf{f}}(k-l\mid k-1)\\&\quad=\sum_{i=0}^{l-1}\langle\mathbf{f}(k-l),\mathbf{w}_1(k-l+i)\rangle Q_1^{-1}(k-l+i)\mathbf{w}_1(k-l+i)\\&\quad=\sum_{i=0}^{l-1}\Omega_{k-l+i}^{k-l}C_1^{\mathrm{T}}(k-l+i)Q_1^{-1}(k-l+i)\mathbf{w}_1(k-l+i),\\&\qquad\qquad i=1,\ldots,l-1,\end{aligned}\tag{38}$$

where Ω_{k-l+i}^{k-l}, $i=1,\ldots,l-1$ is obtained recursively in terms of

$$\begin{aligned}\Omega_{k-l+i}^{k-l}&=\Omega_{k-l+i-1}^{k-l}\big[A(k-l+i-1)-K_1(k-l+i-1)\\&\qquad\times C_1(k-l+i-1)\big]^{\mathrm{T}},\\\Omega_{k-l+1}^{k-l}&=B_f^{\mathrm{T}}(k-l).\end{aligned}\tag{39}$$

Finally, in order to calculate $Q_w(k)$ which is associated with $J_{\min}$ and $\check{f}(k-l)\mid k$, define $\tilde{\mathbf{f}}(k-l)=\mathbf{f}(k-l)-\hat{\mathbf{f}}(k-l\mid k-1)$, and then, from (38), we know that

$$\begin{aligned}&\langle\tilde{\mathbf{f}}(k-l),\tilde{\mathbf{f}}(k-l)\rangle\\&\quad=I-\sum_{i=0}^{l-1}\Omega_{k-l+i}^{k-l}C_1^{\mathrm{T}}(k-l+i)Q_1^{-1}(k-l+i)\\&\qquad\times\left(\Omega_{k-l+i}^{k-l}C_1^{\mathrm{T}}(k-l+i)\right)^{\mathrm{T}}.\end{aligned}\tag{40}$$

Combining (29) and (40), we have

$$Q_w(k) = \begin{cases} C_1(k) P_1(k) C_1^{\mathrm{T}}(k) + I, & 0 < k < l, \\ \begin{bmatrix} C_1(k) P_1(k) C_1^{\mathrm{T}}(k) + I & C_1(k)\left(\Omega_k^{k-l}\right)^{\mathrm{T}} \\ \Omega_k^{k-l} C_1^{\mathrm{T}}(k) & -\gamma^2 I + I - \left\langle \tilde{f}(k-l), \tilde{f}(k-l) \right\rangle \end{bmatrix}, & k \geq l, \end{cases} \tag{41}$$

where $P_1(k)$ and Ω_{l-l+i}^{k-l} are the same as in (37) and (39).

3.3. H_∞ Fault Estimator Design. From analysis and lemmas above, we are now in the position to give our main results for designing the fault estimator, which is summarized in the following theorem.

Theorem 7. *For* (8), *given a scalar $\gamma > 0$ and an integer $l > 0$, then the H_∞ fixed-lag fault estimator that satisfies* (7) *exists if and only if*

$$\begin{aligned} \Lambda_1(k) &= C_1(k) P_1(k) C_1^T(k) + I > 0, \\ \Lambda_3(k) &= -\gamma^2 I + I - \sum_{i=0}^{l-1} \Omega_{k-l+i}^{k-l} C_1^T(k-l+i) Q_1^{-1}(k-l+i) \\ &\quad \times \left(\Omega_{k-l+i}^{k-l} C_1^T(k-l+i)\right)^T \\ &\quad - \Omega_k^{k-l} C_1^T(k) \Lambda_1^{-1}(k) \left(\Omega_k^{k-l} C_1^T(k)\right)^T < 0. \end{aligned} \tag{42}$$

In this case, a feasible fault estimator is given by

$$\begin{aligned} \check{f}(k-l \mid k) &= \sum_{i=0}^{l} \Omega_{k-l+i}^{k-l} C_1^T(k-l+i) Q_1^{-1}(k-l+i) \\ &\quad \times \left[y_f(k-l+i) \right. \\ &\quad \left. - C_1(k-l+i)\hat{x}(k-l+i,1) \right], \end{aligned} \tag{43}$$

where $\hat{x}(k-l+i,1)$, $Q_1(k-l+i)$, and Ω_{k-l+i}^{k-l} are calculated by (32), (34), (35), (37), *and* (39).

Proof. For $k \geq l$, applying the block triangular factorization technique to $Q_w(k)$ in (41), we have

$$\begin{aligned} Q_w(k) &= \begin{bmatrix} I & 0 \\ \Lambda_2(k)^{\mathrm{T}} \Lambda_1(k)^{-1} & I \end{bmatrix} \begin{bmatrix} \Lambda_1(k) & 0 \\ 0 & \Lambda_3(k) \end{bmatrix} \\ &\quad \times \begin{bmatrix} I & 0 \\ \Lambda_2(k)^{\mathrm{T}} \Lambda_1(k)^{-1} & I \end{bmatrix}^{\mathrm{T}}, \end{aligned} \tag{44}$$

where $\Lambda_2(k) = \Omega_k^{k-l} C_1^{\mathrm{T}}(k)$. Thus, from Lemma 4, we know that $Q_a(k)$ and $Q_w(k)$ have the same inertia if and only if $\Lambda_1(k) > 0$ as well as $\Lambda_3(k) < 0$. Furthermore, based on (25) and (44), J has a minimum $J_{\min}$ if (42) are satisfied, where

$$\begin{aligned} J_{\min} &= \sum_{k=0}^{l-1} \left[y_f(k) - C_1(k)\hat{x}(k) \right]^{\mathrm{T}} \\ &\quad \times \Lambda_1^{-1}(k) \left[y_f(k) - C_1(k)\hat{x}(k) \right] \\ &\quad + \sum_{k=l}^{N} \left(\begin{bmatrix} I & 0 \\ -\Lambda_2(k)\Lambda_1^{-1}(k) & I \end{bmatrix} \right. \\ &\quad \left. \times \begin{bmatrix} y_f(k) - C_1(k)\hat{x}(k) \\ \check{f}(k-l \mid k) - \hat{f}(k-l \mid k-1) \end{bmatrix} \right)^{\mathrm{T}} \\ &\quad \times \begin{bmatrix} \Lambda_1^{-1}(k) & 0 \\ 0 & \Lambda_3^{-1}(k) \end{bmatrix} \begin{bmatrix} I & 0 \\ -\Lambda_2(k)\Lambda_1^{-1}(k) & I \end{bmatrix} \\ &\quad \times \begin{bmatrix} y_f(k) - C_1(k)\hat{x}(k) \\ \check{f}(k-l \mid k) - \hat{f}(k-l \mid k-1) \end{bmatrix}. \end{aligned} \tag{45}$$

Since $\Lambda_3(k) < 0$, to guarantee $J_{\min} > 0$, combining (38) with (45), we know that a possible choice of $\check{f}(k-l \mid k)$ is

$$\begin{aligned} \check{f}(k-l \mid k) &= \hat{f}(k-l \mid k-1) \\ &\quad + \Lambda_2(k)\Lambda^{-1}\left(y_f(k) - C_1\hat{x}(k) \right) \\ &= \sum_{i=0}^{l} \Omega_{k-l+i}^{k-l} C_1^{\mathrm{T}}(k-l+i) Q_1^{-1}(k-l+i) \\ &\quad \times \left[y_f(k-l+i) - C_1(k-l+i) \right. \\ &\quad \left. \times \hat{x}(k-l+i,1) \right], \end{aligned} \tag{46}$$

which indicates (43). This completes the proof. □

Remark 8. It can be seen from Theorem 7 that the superiority of the proposed algorithm lies in three aspects:

(i) in contrast with the results in [22–26], the proposed algorithm can be applied to systems with time-varying $\rho(k)$;

(ii) comparing to the result in [27], the parameter matrices of the addressed estimator are given in terms of standard Riccati equations with the same dimension "n" of system (8), where no coupled Lyapunov equation with higher dimension is needed;

(iii) the fault can be estimated in an arbitrary fixed-lag "l."

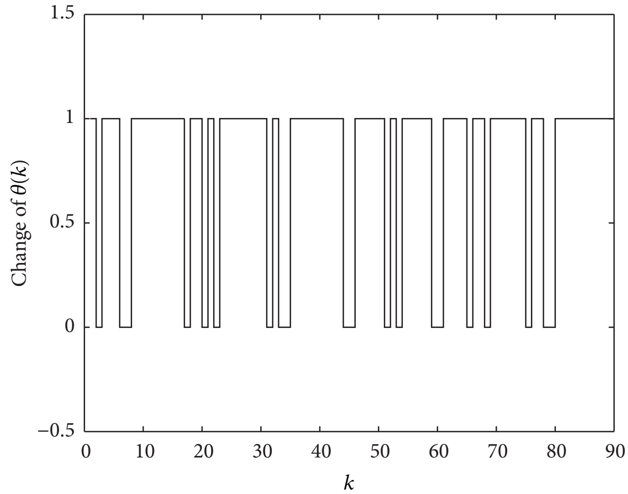


Figure 1: The change mode of $\theta(k)$.

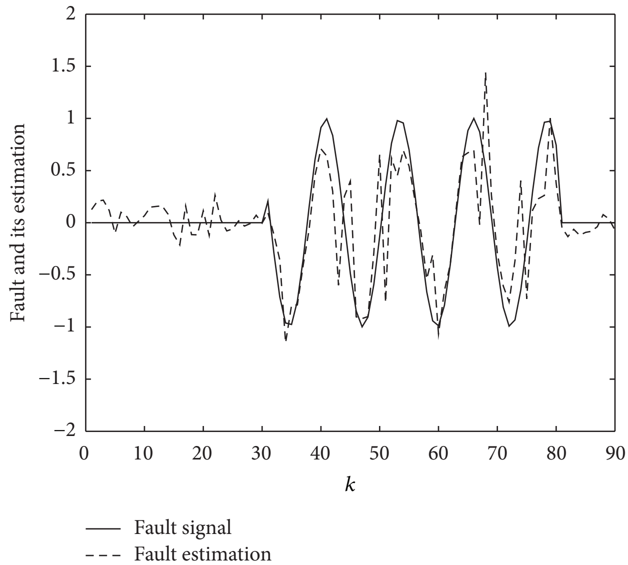


Figure 2: Fault and its estimation.

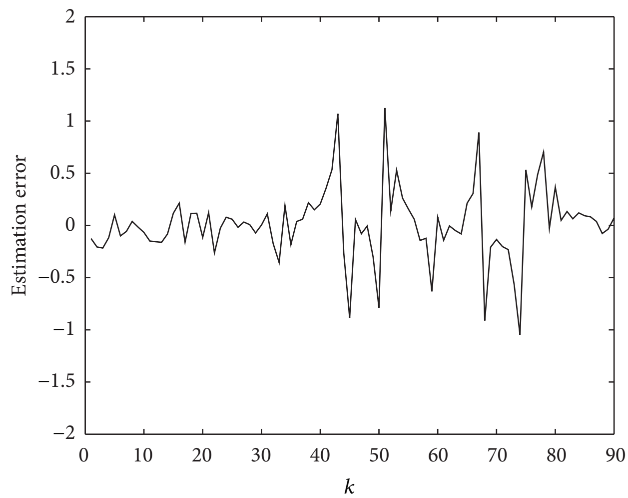


Figure 3: Fault estimation error.

4. An Illustrative Example

To illustrate the effectiveness and the applicability of the proposed method, we will implement our algorithm on a time-varying model. The following system matrices are adopted which are borrowed from [33, 34]:

$$A(k) = (1 + 0.2\sin(0.02k\pi)) \times \begin{bmatrix} 0.8 & 0 \\ 0.9 & 0.2 \end{bmatrix},$$
$$B_f(k) = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}^{\mathrm{T}}, \qquad C(k) = \begin{bmatrix} 1 & 1 \end{bmatrix}, \tag{47}$$
$$D(k) = \begin{bmatrix} 0.3 & 0.25 \end{bmatrix}^{\mathrm{T}}.$$

The process noise $d(k)$ is uniformly randomly chosen from the interval $[-0.5, 0.5]$ and the measurement noise $v(k)$ is assumed as $v(k) = 0.5\sin(0.2k)$. The fault signal $f(k)$ is assumed to be time-varying in the following sinusoidal form:

$$f(k) = \begin{cases} \sin(0.5k), & k \in [30, 80], \\ 0, & \text{otherwise}, \end{cases} \tag{48}$$

and the expectation of $\theta(k)$ is assumed as $\rho = 0.8$, where Figure 1 displays the switching mode of $\theta(k)$.

Set $l = 10$, $\gamma = 1.52$, $x_0 = [0.2 \ \ 0]^{\mathrm{T}}$, and $P_0 = 0.1I$; we design the fault estimator by applying Theorem 7. Figure 2 displays the fault signal and its estimation simultaneously. Figure 3 shows the value of $f(k-l) - \check{f}(k-l \mid k)$ which is the error between the fault and its estimation. It can be seen from the results that our algorithm can track the fault signal no matter whether the random packet dropouts occur.

5. Conclusions

The problem of H_∞ fixed-lag fault estimator design for LDTV systems subject to intermittent observations has been dealt with. Special efforts have been made to handle the multiplicative uncertainty introduced by the random measurement packet dropouts. Through defining a couple of equivalent dynamic system and H_∞ performance index, the fault estimator has been derived by using the projection formula in Krein space based on the reorganized innovation approach. The parameter matrices of the estimator have been calculated by solving two standard Riccati equations. The proposed algorithm has been applied to an LDTV model to illustrate its effectiveness and applicability.

Conflict of Interests

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

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (no. 61203083, no. 61304045, no. 61304013, and no. 61074021), the Doctoral Foundation of University of Jinan (no. XBS1242), and the Science and Technology Development Plan Program of Shandong Province (no. 2013GGX10117).

References

[1] M. Blanke, M. Kinnart, and J. Lunze, *Diagnosis and Fault-Tolerant Control*, Springer, Berlin, Germany, 2006.

[2] S. X. Ding, *Model-based Fault Diagnosis Techniques*, Springer, Berlin, Germany, 2008.

[3] M. Chadli, A. Abdo, and S. X. Ding, "H_-/H_∞ fault detection filter design for discrete-time Takagi-Sugeno fuzzy system," *Automatica*, vol. 49, no. 7, pp. 1996–2005, 2013.

[4] R. M. G. Ferrari, M. M. Polycarpou, and T. Parisini, "Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: an adaptive approximation approach," *IEEE Transaction on Automatic Control*, vol. 57, no. 2, pp. 275–290, 2012.

[5] J. Cai, C. Wen, H. Su, and Z. Liu, "Robust adaptive failure compensation of hysteretic actuators for a class of uncertain nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2388–2394, 2013.

[6] J. Liu and D. Yue, "Event-based fault detection for networked systems with communication delay and nonlinear perturbation," *Journal of the Franklin Institute*, vol. 350, no. 9, pp. 2791–2807, 2013.

[7] N. Meskin and K. Khorasani, "Robust fault detection and isolation of time-delay systems using a geometric approach," *Automatica*, vol. 45, no. 6, pp. 1567–1573, 2009.

[8] A. Casavola, D. Famularo, and G. Franz, "A robust deconvolution scheme for fault detection and isolation of uncertain linear systems: an LMI approach," *Automatica*, vol. 41, no. 8, pp. 1463–1472, 2005.

[9] Y. Zhang, H. Fang, and Z. Liu, "Fault detection for nonlinear networked control systems with Markov data transmission pattern," *Circuits, Systems, and Signal Processing*, vol. 31, no. 4, pp. 1343–1358, 2012.

[10] Y. Zhang, Z. Liu, H. Fang, and H. Chen, "H_∞ fault detection for nonlinear networked systems with multiple channels data transmission pattern," *Information Sciences*, vol. 221, no. 1, pp. 534–543, 2013.

[11] X. Li and K. Zhou, "A time domain approach to robust fault detection of linear time-varying systems," *Automatica*, vol. 45, no. 1, pp. 94–102, 2009.

[12] M. Zhong, S. X. Ding, and E. L. Ding, "Optimal fault detection for linear discrete time-varying systems," *Automatica*, vol. 46, no. 8, pp. 1395–1400, 2010.

[13] M. Zhong, S. Liu, and H. Zhao, "Krein space-based H_∞ fault estimation for linear discrete time-varying systems," *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1529–1533, 2008.

[14] M. Zhong, D. Zhou, and S. X. Ding, "On designing H_∞ fault detection filter for linear discrete time-varying systems," *IEEE Transaction on Automatic Control*, vol. 55, no. 7, pp. 1689–1695, 2010.

[15] B. Shen, S. X. Ding, and Z. Wang, "Finite-horizon H_∞ fault estimation for linear discrete time-varying systems with delayed measurements," *Automatica*, vol. 49, no. 1, pp. 293–296, 2013.

[16] H. Zhang, L. Xie, D. Zhang, and Y. C. Soh, "A reorganized innovation approach to linear estimation," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1810–1814, 2004.

[17] H. Zhang, L. Xie, Y. C. Soh, and D. Zhang, "H_∞ fixed-lag smoothing for discrete linear time-varying system," *Automatica*, vol. 41, no. 5, pp. 839–846, 2005.

[18] H. Zhang, G. Feng, and C. Han, "Linear estimation for random delay systems," *Systems and Control Letters*, vol. 60, no. 7, pp. 450–459, 2011.

[19] H. Zhang, X. Lu, W. Zhang, and W. Wang, "Kalman filtering for linear time-delayed continuous-time systems with stochastic multiplicative noises," *International Journal of Control, Automation and Systems*, vol. 5, no. 4, pp. 355–363, 2007.

[20] H. Zhang and L. Xie, *Control and Estimation of Systems with Input/Output Delays*, Springer, Berlin , Germany, 2007.

[21] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in Krein spaces. I. Theory," *IEEE Transactions on Automatic Control*, vol. 41, no. 1, pp. 18–33, 1996.

[22] H. Dong, Z. Wang, and H. Gao, "On design of quantized fault detection filters with randomly occurring nonlinearities and mixed time-delays," *Signal Processing*, vol. 92, no. 4, pp. 1117–1125, 2012.

[23] J. Yu, M. Liu, W. Yang, and P. Shi, "Robust fault detection for Markovian jump systems with unreliable communication links," *International Journal of Systems Science*, vol. 44, no. 11, pp. 2015–2026, 2013.

[24] X. Wan, H. Fang, and F. Yang, "Fault detection for a class of networked nonlinear systems subject to imperfect measurements," *International Journal of Control, Automation, and Systems*, vol. 10, no. 2, pp. 265–274, 2013.

[25] D. Zhang, Q. Wang, L. Yu, and H. Song, "Fuzzy-model-based fault detection for a class of nonlinear systems with networked measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 12, pp. 3148–3159, 2013.

[26] S. M. M. Alavi and M. Saif, "Fault detection in nonlinear stable systems over lossy networks," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 6, pp. 2129–2142, 2013.

[27] Y. Y. Li and M. Y. Zhong, "Designing robust H_∞ fault detection filter for linear discrete time-varying systems with multiple packet dropouts," *Acta Automatica Sinica*, vol. 36, no. 12, pp. 1788–1796, 2010.

[28] Y. Li, S. Liu, and Z. Wang, "Fault detection for linear discrete time-varying systems with measurement packet dropping," *Mathematical Problems in Engineering*, vol. 2013, Article ID 697345, 9 pages, 2013.

[29] J. Nilsson, *Real-time control systems with delays [Ph.D. dissertation]*, Lund institute of Technology, 1998.

[30] A. A. Stoorvogel, H. H. Niemann, A. Saberi, and B. Sannuti, "Optimal fault signal estimation," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 8, pp. 697–727, 2002.

[31] H. Zhao and C. Zhang, "H_∞ fixed-lag smoothing for linear discrete time-varying systems with uncertain observations," *Applied Mathematics and Computation*, vol. 224, pp. 387–397, 2013.

[32] X. Lu, H. Zhang, and J. Yan, "On the H_∞ deconvolution fixed-lag smoothing," *International Journal of Control, Automation and Systems*, vol. 8, no. 4, pp. 896–902, 2010.

[33] J. Ma and S. Sun, "Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5181–5192, 2011.

[34] M. Moayedi, Y. K. Foo, and Y. C. Soh, "Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1577–1588, 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 653259, 9 pages
http://dx.doi.org/10.1155/2014/653259

Research Article

Iterative Mixture Component Pruning Algorithm for Gaussian Mixture PHD Filter

Xiaoxi Yan

School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China

Correspondence should be addressed to Xiaoxi Yan; yanxiaoxi1981@gmail.com

Received 23 April 2014; Accepted 30 June 2014; Published 15 July 2014

Academic Editor: Suiyang Khoo

As far as the increasing number of mixture components in the Gaussian mixture PHD filter is concerned, an iterative mixture component pruning algorithm is proposed. The pruning algorithm is based on maximizing the posterior probability density of the mixture weights. The entropy distribution of the mixture weights is adopted as the prior distribution of mixture component parameters. The iterative update formulations of the mixture weights are derived by Lagrange multiplier and Lambert W function. Mixture components, whose weights become negative during iterative procedure, are pruned by setting corresponding mixture weights to zeros. In addition, multiple mixture components with similar parameters describing the same PHD peak can be merged into one mixture component in the algorithm. Simulation results show that the proposed iterative mixture component pruning algorithm is superior to the typical pruning algorithm based on thresholds.

1. Introduction

The objective of multitarget tracking is to estimate target number and target states from a sequence of noisy and cluttered measurement sets. The tracked target is generally simplified as a point [1–3]. Most of the existing point target tracking algorithms are based on data association where the correspondence of measurements to targets has to be set up. The simplest data association algorithm is the nearest-neighbour algorithm in which the measurement closest in statistical distance to predicted state is used to update target state estimate. Probabilistic data association is another typical algorithm in which all the measurements close to the predicted state are used to update target state estimate [4]. Joint probabilistic data association is a generalization of probabilistic data association for multiple target tracking in which association probabilities of all the targets and measurements are described by confirmed matrices [5, 6]. Multitarget tracking algorithms based on data association are in individual view, where the problem of multitarget tracking is converted into the multiple problems of single target tracking. In the multitarget tracking, both the measurements and the estimations are gained in the set form. Thus, multitarget tracking is naturally a class of set-valued estimation problems. The probability hypothesis density (PHD) filter derived by Mahler based on random finite sets statistics theory is an elegant and tractable approximate solution to the multitarget tracking problem [7, 8]. Another interpretation of the PHD in bin-occupancy view is presented in [9]. By now, there have been two implementations of PHD filter, Gaussian mixture implementation [10, 11] and sequential Monte Carlo implementation [12–16], which are suitable for linear Gaussian dynamics and nonlinear non-Gaussian dynamics. The convergence of Gaussian mixture implementation is discussed in [17] and the convergence of sequential Monte Carlo implementation in [15, 18, 19]. The cardinalized PHD (CPHD) filter propagating both the PHD and the distribution of target number is developed to improve the performance of the PHD filter [20]. Generally, the CPHD filter is computationally less tractable compared to the PHD filter. There have been the Gaussian mixture implementation of CPHD filter under multitarget linear Gaussian assumptions [21] and the sequential Monte Carlo implementation [22]. As promising and unified methodologies, the PHD and CPHD filters have been widely applied in many fields, such as maneuvering target tracking [23, 24], sonar tracking [25, 26], and visual tracking [27–29]. As the sensor resolution is greatly improved, target tracking should be formulated as extended object tracking [30]. Extended object PHD filter is also derived by Mahler in [31]. There have been some implementations of extended object probability hypothesis density filter by now [32–36]. The convergence of the Gaussian mixture implementation of extended object probability hypothesis density filter is discussed in [37]. When the Gaussian mixture model is applied in set-valued multitarget tracking, the Gaussian mixture reduction is an important topic [10, 38]. The earlier work in Gaussian mixture reduction for target tracking has been done in [39, 40]. As the Gaussian mixture reduction is implemented, there are several criterions such as maximum similarity [41], Euclidean distance [42–44], and Kullback-Leibler divergence measure [45]. The concentrations of this paper are on the Gaussian mixture reduction of the Gaussian mixture implementation of PHD filter.

As far as the Gaussian mixture implementation of the PHD filter is concerned, it approximates the PHD by the summation of weighted Gaussian components under the multitarget linear Gaussian assumptions [10]. In the Gaussian mixture PHD filter, the PHD is presented by a large number of weighted Gaussian components that are propagated over time. The sum of the weights of Gaussian components is the expected target number since the integral of the PHD over the state space is the expected target number. The output of Gaussian mixture PHD filter is weighted Gaussian components. However, the Gaussian mixture PHD filter suffers from computation problems associated with the increasing number of Gaussian components as time progresses, since mixture component number increases both at prediction step and at update step. In fact, component number increases without bound. Thus, the Gaussian mixture PHD filter is infeasible without component pruning operation. The goal of this paper is to prune the Gaussian components to make the Gaussian mixture PHD filter feasible. An iterative mixture component pruning algorithm is proposed for the Gaussian mixture PHD filter. The pruning operation of mixture components is done by setting mixture weights to zeros during the iteration procedure.

The remaining parts of this paper are organized as follows. Section 2 describes the component increasing problem in Gaussian mixture PHD filter. The iterative mixture component pruning algorithm is derived in Section 3. Section 4 is devoted to the simulation study. Conclusion is provided in Section 5.

2. Problem Description

The predictor and connector of PHD filter [7, 8] are

$$\begin{aligned} v_{k|k-1}(x) = &\int p_{S,k}(\zeta) f_{k|k-1}(x \mid \zeta) v_{k-1}(\zeta) d\zeta \\ &+ \int \beta_{k|k-1}(x \mid \zeta) v_{k-1}(\zeta) d\zeta + \gamma_k(x), \end{aligned} \tag{1}$$

$$\begin{aligned} v_k(x) = &\left[1 - p_{D,k}(x)\right] v_{k|k-1}(x) \\ &+ \sum_{z \in Z_k} \frac{\varphi_{z,k}(x) v_{k|k-1}(x)}{\kappa_k(z) + \int \varphi_{z,k}(\xi) v_{k|k-1}(\xi) d\xi}, \end{aligned} \tag{2}$$

respectively, where $v(\cdot)$ is the PHD, $\gamma_k(x)$ is the birth PHD at time step k, $\beta_{k|k-1}(\cdot \mid \zeta)$ is the spawned PHD from ζ at time step $k-1$, $\kappa_k(z)$ is the clutter PHD, $p_{S,k}(\zeta)$ is the survival probability, $p_{D,k}(x)$ is the detection probability, $\varphi_{z,k}(x) = p_{D,k}(x) g_k(z \mid x)$, $g_k(z \mid x)$ is the single target likelihood, and Z_k is the measurements at time step k.

Under the linear Gaussian assumptions, the Gaussian mixture PHD filter is derived in [10]. The main steps of the Gaussian mixture PHD filter are summarized as follows. If the PHD at time step $k-1$ is in the form of Gaussian mixture

$$v_{k-1}(x) = \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}\left(x; m_{k-1}^{(i)}, P_{k-1}^{(i)}\right), \tag{3}$$

where w is the mixture weight, $\mathcal{N}(\cdot)$ is the Gaussian distribution, m is the mean, P is the covariance, and J is the component number, then the predicted PHD for time step k is given by

$$v_{k|k-1}(x) = v_{S,k|k-1}(x) + v_{\beta,k|k-1}(x) + \gamma_k(x), \tag{4}$$

where γ_k is the birth PHD

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}\left(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}\right), \tag{5}$$

$v_{S,k|k-1}$ is the survival PHD

$$v_{S,k|k-1}(x) = p_{S,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} \mathcal{N}\left(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}\right), \tag{6}$$

$m_{S,k|k-1}^{(j)}$ is the predicted mean of the Gaussian component

$$m_{S,k|k-1}^{(j)} = F_{k-1} m_{k-1}^{(j)}, \tag{7}$$

$P_{S,k|k-1}^{(j)}$ is the predicted covariance of the Gaussian component

$$P_{S,k|k-1}^{(j)} = Q_{k-1} + F_{k-1} P_{k-1}^{(j)} F_{k-1}^T, \tag{8}$$

$v_{\beta,k|k-1}$ is the spawned PHD

$$v_{\beta,k|k-1}(x) = \sum_{j=1}^{J_{k-1}} \sum_{l=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(l)} \mathcal{N}\left(x; m_\beta^{(j,l)}, P_\beta^{(j,l)}\right), \tag{9}$$

$m_\beta^{(j,l)}$ is the spawned mean of the Gaussian component

$$m_\beta^{(j,l)} = F_{\beta,k-1}^{(l)} m_{k-1}^{(j)} + d_{\beta,k-1}^{(l)}, \tag{10}$$

and $P_\beta^{(j,l)}$ is the spawned covariance of the Gaussian component

$$P_\beta^{(j,l)} = Q_{\beta,k-1}^{(l)} + F_{\beta,k-1}^{(l)} P_{\beta,k-1}^{(j)} \left(F_{\beta,k-1}^{(l)}\right)^T. \tag{11}$$

If formula (4) is rewritten in the simple form of the Gaussian mixture

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}\left(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}\right), \quad (12)$$

then the posterior PHD at time step k is

$$v_k(x) = \left(1 - p_{D,k}\right) v_{k|k-1}(x) + \sum_{z \in Z_k} v_{D,k}(x; z), \quad (13)$$

where $v_{D,k}$ is the detected PHD

$$v_{D,k}(x; z) = \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) \mathcal{N}\left(x; m_{k|k}^{(j)}(z), P_{k|k}^{(j)}\right), \quad (14)$$

$w_k^{(j)}$ is the updated weight

$$w_k^{(j)}(z) = \frac{p_{D,k} w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)}, \quad (15)$$

$m_{k|k}^{(j)}(z)$ is the updated mean

$$m_{k|k}^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)}\left(z - H_k m_{k|k-1}^{(j)}\right), \quad (16)$$

$P_{k|k}^{(j)}$ is the updated covariance

$$P_{k|k}^{(j)} = \left[I - K_k^{(j)} H_k\right] P_{k|k-1}^{(j)}, \quad (17)$$

and $K_k^{(j)}$ is the gain

$$K_k^{(j)} = P_{k|k-1}^{(j)} H_k^T \left(H_k P_{k|k-1}^{(j)} H_k^T + R_k\right)^{-1}. \quad (18)$$

It can be seen from formula (4) that component number increases from J_{k-1} to $J_{k|k-1}$ by $J_{k-1} \cdot J_{\beta,k} + J_{\gamma,k}$ at the prediction step. It is obvious in formula (13) that component number increases from $J_{k|k-1}$ to J_k by $J_{k|k-1} \cdot |Z_k|$ at the update step. Hence, the number of Gaussian components J_k representing PHD v_k at time step k in Gaussian mixture PHD filter is

$$J_k = \left(J_{k-1}\left(1 + J_{\beta,k-1}\right) + J_{\gamma,k}\right)\left(1 + |Z_k|\right), \quad (19)$$

where J_{k-1} is the number of components of the PHD v_{k-1} at time step $k-1$. In formula (19), the component number increases in $\mathcal{O}(J_{k-1}|Z_k|)$. In particular, the component number mostly increases in $(J_{k-1}(1 + J_{\beta,k}) + J_{\gamma,k})|Z_k|$ at the update step. Indeed, the number of Gaussian components increases without bound so that the computation of the Gaussian mixture PHD filter is intractable after several time steps. Therefore, it is necessary to reduce the number of components to make the Gaussian mixture PHD filter feasible. The goal of this paper is to prune the Gaussian mixture components to reduce component number in Gaussian mixture PHD filter.

3. Iterative Pruning Algorithm

For simplicity, the time index k is neglected and let $M = J_{k|k}$ represent component number. w_S is the sum of the weights of the Gaussian components:

$$w_S = \sum_{j=1}^{M} w_j. \quad (20)$$

In the iterative pruning algorithm, the weights of Gaussian components are normalized by $\{w_1/w_S, \ldots, w_M/w_S\}$ at first so that

$$\sum_{j=1}^{M} w_j = 1. \quad (21)$$

Let $\theta_j = \{m^{(j)}, P^{(j)}\}$ represent the parameters of the jth Gaussian component, where $m^{(j)}$ and $P^{(j)}$ are the mean and covariance, respectively. Then, the whole parameter set of M Gaussian components is $\theta = \{w_1, \ldots, w_M, \theta_1, \ldots, \theta_M\}$.

The entropy distribution of the mixture weights is adopted as the prior of θ:

$$p(\theta) \propto \exp\left(-H\left(w_1, \ldots, w_M\right)\right), \quad (22)$$

where $H(w_1, \ldots, w_M) = -\sum_{j=1}^{M} w_j \log w_j$ is the entropy measure [46, 47]. The goal of this choice of prior distribution, which depends only on the mixture weights, is to reduce mixture components by the adjustment of mixture weights during the iteration procedure. If we define the log-likelihood of the measurements $Z = \{z_1, \ldots, z_n\}$ given the mixture parameters as

$$\log p(Z \mid \theta) = \sum_{i=1}^{n} \log \sum_{j=1}^{M} w_j g\left(z_i \mid \theta_j\right), \quad (23)$$

where $g(z \mid \theta_j)$ is the single target likelihood in jth component, then the MAP estimate of θ is

$$\widehat{\theta} = \arg\max_{\theta} \left\{\log p(Z \mid \theta) + \log p(\theta)\right\}. \quad (24)$$

For the mixture weight w_j, the MAP estimate can be computed by setting the derivative of the log-posterior to zero:

$$\frac{\partial}{\partial w_j}\left(\log p(Z \mid \theta) + \log p(\theta)\right) = 0. \quad (25)$$

The MAP estimate of w_j is computed by maximizing $\log p(Z \mid \theta) + \log p(\theta)$ under the constraint (21):

$$\frac{\partial}{\partial w_j}\left(\log p(Z \mid \theta) + \log p(\theta) + \lambda\left(\sum_{j=1}^{M} w_j - 1\right)\right) = 0, \quad (26)$$

where λ is Lagrange multiplier. Substituting formulas (22) and (23) into formula (26) gives

$$\frac{\sum_{i=1}^{n} \omega_j(z_i)}{w_j} + \log w_j + \lambda + 1 = 0, \quad (27)$$

where $\omega_j(z)$ represents the membership that z is from the jth mixture component:

$$\omega_j(z) = \frac{w_j g(z \mid \theta_j)}{\sum_{l=1}^{M} w_l g(z \mid \theta_l)}. \quad (28)$$

Formula (27) is a simultaneous transcendental equation. We solve it for the w_j using the Lambert W function [48], an inverse mapping satisfying $W(y)e^{W(y)} = y$, and therefore $\log W(y) + W(y) = \log y$. The Lambert W function of complex y is defined as $W(y)$, which is a set of functions. The complex y can be computed by the equation $W(y)e^{W(y)} = y$, where $e^{W(y)}$ is the exponential function. Lambert W function $W(y)$ is a multivalued function defined in general for y complex and assumed $W(y)$ complex. If y is real and $y < -1/e$, then $W(y)$ is multivalued complex. If y is real and $-1/e \le y < 0$, $W(y)$ has two possible real values. If y is real and $y > 0$, $W(y)$ has one real value. Then, for the Lambert W function $W(y)$,

$$-W(y) - \log W(y) + \log y = 0. \quad (29)$$

Setting $y = e^x$, formula (29) can be rewritten as

$$-W(e^x) - \log W(e^x) + x = 0. \quad (30)$$

In formula (27), it is assumed that

$$\omega_j = \sum_{i=1}^{n} \omega_j(z_i). \quad (31)$$

Consequently, formula (30) is

$$\frac{\omega_j}{-\omega_j / W(e^x)} + \log\left(\frac{-\omega_j}{W(e^x)}\right) + x - \log(-\omega_j) = 0. \quad (32)$$

Comparing the Lambert W function (32) to formula (27), (32) can be reduced to (27) by setting $x = 1 + \lambda + \log(-\omega_j)$:

$$\frac{\omega_j}{-\omega_j / W(e^x)} + \log\left(\frac{-\omega_j}{W(e^x)}\right) + 1 + \lambda = 0. \quad (33)$$

Consequently,

$$w_j = \frac{-\omega_j}{W(e^{1+\lambda+\log(-\omega_j)})} = \frac{-\omega_j}{W(-\omega_j e^{1+\lambda})}. \quad (34)$$

Formula (27) and formula (34) constitute an iterative procedure for the MAP estimates of $\{w_1, \ldots, w_M\}$: (1) given λ, $\{w_1, \ldots, w_M\}$ are calculated by formula (34); (2) $\{w_1, \ldots, w_M\}$ are normalized; (3) given normalized $\{w_1, \ldots, w_M\}$, λ is computed by formula (27). The iteration procedure stops when the difference rate of log-posterior is smaller than the given threshold.

At the normalization step of the iteration procedure, if a mixture weight becomes negative, the corresponding component is removed from the mixture components by setting its weight to zero. The removed mixture component will not be considered when the log-posterior is computed in the following iterations. The mixing weights of survival mixture components are normalized at the end of this step.

The effect of entropy distribution of mixing weights is taken during the iterative procedure. The mixture weights of components negligible to the PHD become smaller and smaller iteration by iteration, since the parameter estimates are driven into low-entropy direction by entropy distribution. The low-entropy tendency can also promote competition among the mixture components with similar parameters which can then be merged into one mixture component with larger weight.

For the mean $m^{(j)}$ and covariance $P^{(j)}$ of mixture component with nonzero weight w_j, they are updated by

$$m^{(j)} = (\omega_j)^{-1} \sum_{i=1}^{n} z_i \omega_j(z_i), \quad (35)$$

$$P^{(j)} = (\omega_j)^{-1} \sum_{i=1}^{n} (z_i - m^{(j)})(z_i - m^{(j)})^T \omega_j(z_i). \quad (36)$$

The main steps of iterative mixture component pruning algorithm are summarized in Algorithm 1.

4. Simulation Study

A two-dimensional scenario with unknown and time-varying target number is considered to test the proposed iterative mixture component pruning algorithm. The surveillance region is $[-1000, 1000] \times [-1000, 1000]$ (in meter). The target state consists of position and velocity, while target measurement is the position. Each target moves according to the following dynamics:

$$x_k = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} v_{1,k} \\ v_{2,k} \end{bmatrix}, \quad (37)$$

where $x_k = [x_{1,k}, x_{2,k}, x_{3,k}, x_{4,k}]^{\mathrm{T}}$ is the target state, $[x_{1,k}, x_{2,k}]^{\mathrm{T}}$ is the target position, and $[x_{3,k}, x_{4,k}]^{\mathrm{T}}$ is the target velocity at time step k. The process noises are a zero-mean Gaussian white noise with standard deviations $\sigma_{v_1} = \sigma_{v_2} = 5\,(\mathrm{m/s}^2)$. The survival probability is $p_{S,k} = 0.99$. The number of targets is unknown and variable over all scans. New targets appear spontaneously according to a Poisson point process with PHD function $\gamma_k = 0.2\mathcal{N}(\cdot; \overline{x}, Q)$, where

$$\overline{x} = \begin{bmatrix} -400 \\ -400 \\ 0 \\ 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 25 \end{bmatrix}. \quad (38)$$

$\mathcal{N}(\cdot; \overline{x}, Q)$ is the Gaussian component with mean $\overline{x}$ and covariance Q. The spawned PHD is $\beta_{k|k-1}(x \mid \zeta) = 0.05\mathcal{N}(x; \zeta, Q_\beta)$, where $Q_\beta = \mathrm{diag}([100, 100, 400, 400]^{\mathrm{T}})$.

(1) normalize $w_1, \ldots, w_M$ by formula (20).
(2) $t = 0$.
(3) $t = t + 1$.
(4) **for** $i = 1, \ldots, n$ **do**
(5) **for** $j = 1, \ldots, M$ **do**
(6) compute $\omega_j(z_i)$ by formula (28).
(7) **end for**
(8) **end for**
(9) **for** $j = 1, \ldots, M$ **do**
(10) compute ω_j by formula (31).
(11) **end for**
(12) **for** $j = 1, \ldots, M$ **do**
(13) compute w_j by formula (34).
(14) **end for**
(15) **for** $j = 1, \ldots, M$ **do**
(16) **if** $w_j < 0$ **do**
(17) **for** $l = j, \ldots, M - 1$ **do**
(18) $w_l = w_{l+1}$.
(19) $m^{(l)} = m^{(l+1)}$.
(20) $P^{(l)} = P^{(l+1)}$.
(21) **end for**
(22) $j = j - 1$.
(23) $M = M - 1$.
(24) **else do**
(25) compute $m^{(j)}$ by formula (35).
(26) compute $P^{(j)}$ by formula (36).
(27) **end if**
(28) **end for**
(29) normalize $w_1, \ldots, w_M$;
(30) compute λ by formula (27);
(31) **if** $\log p(\theta(t) \mid Z) - \log p(\theta(t-1) \mid Z) > \varepsilon \cdot \log p(\theta(t-1) \mid Z)$ **do**
(32) goto step 3;
(33) **end if**.

Algorithm 1: Iterative pruning algorithm.

Each target is detected with probability $p_{D,k} = 0.98$. The target-originated measurement model is

$$y_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} w_{1,k} \\ w_{2,k} \end{bmatrix}, \tag{39}$$

where the measurement noise is a zero-mean Gaussian white noise with standard deviation $\sigma_{w_1} = \sigma_{w_2} = 10\,(\mathrm{m})$. Clutter is modelled as a Poisson random finite set with intensity

$$\kappa_k(z_k) = \lambda_c \cdot c_k(z_k), \tag{40}$$

where λ_c is the average number of clutter measurements per scan and $c(z)$ is the probability distribution over surveillance region. Here $c(z)$ is a uniform distribution and λ_c is assumed to be 50.

The means of the Gaussian mixture components with mixing weights greater than 0.5 are chosen as the estimates of multitarget states after the mixture reduction.

The tracking results in one Monte Carlo trial are presented in Figures 1 and 2. It can be seen from Figures 1 and 2 that the Gaussian mixture PHD filter with the proposed iterative mixture component pruning algorithm is able to detect the spontaneous and spawned targets and estimate the multiple target states.

The mixture components with weights larger than 0.0005 at the 86th time step before pruning operation in the above Monte Carlo simulation trial are presented in Figure 3. The mixture components with weights larger than 0.01 after pruning operation are presented in Figure 4. It is obvious that the mixture components with similar parameters describing the same PHD peak can be merged into one mixture component.

The typical mixture component pruning algorithm based on thresholds in [10] is adopted as the comparison algorithm. The thresholds in typical mixture component pruning algorithm are weight pruning threshold 10^{-5}, mixture component merging threshold 4, and maximum allowable mixture component number 100. We evaluate the tracking performance of proposed algorithm against the typical algorithm by Wasserstein distance [49]. The Wasserstein distance is defined as

$$d_p(\widehat{X}, X) = \min_C \sqrt[p]{\sum_{i=1}^{|\widehat{X}|} \sum_{j=1}^{|X|} C^{ij} \left\| \widehat{x}^i - x^j \right\|^p}, \tag{41}$$

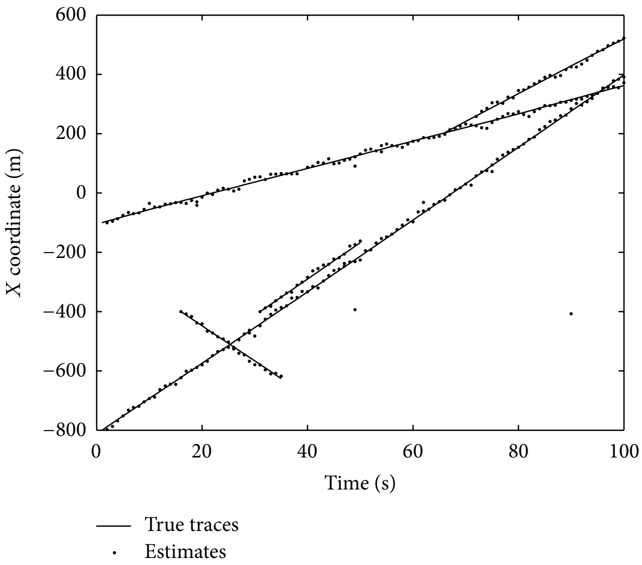


Figure 1: True traces and estimates of X coordinates.

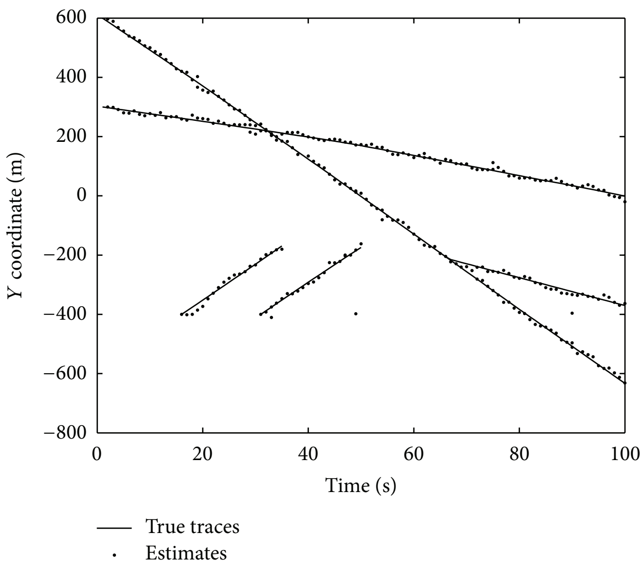


Figure 2: True traces and estimates of Y coordinate.

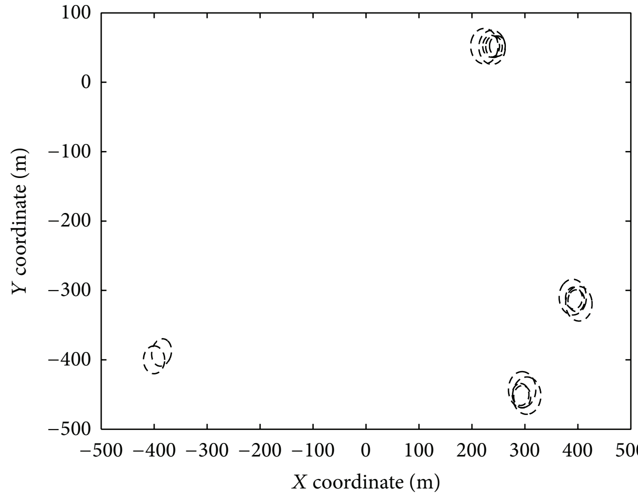


Figure 3: Components before pruning operation.

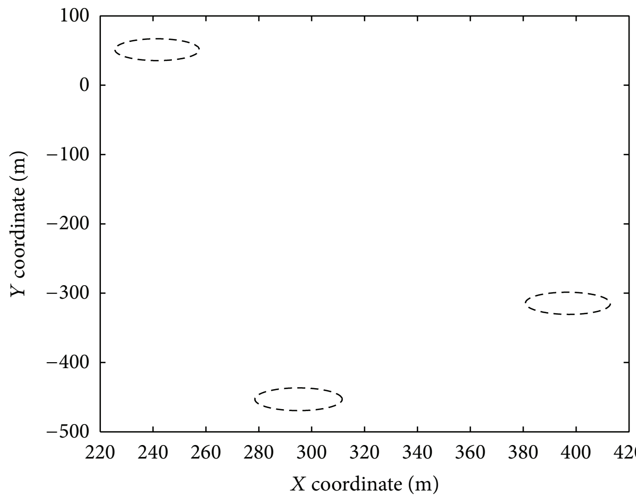


Figure 4: Components after pruning operation.

where $\widehat{X}$ is the estimate of multitarget state set and X is the true multitarget state set. The minimum is taken over the set of all transportation matrices C (a transportation matrix C is one whose entries C^{ij} satisfy $C^{ij} \geq 0$, $\sum_{j=1}^{|X|} C^{ij} = 1/|\widehat{X}|$, and $\sum_{i=1}^{|\widehat{X}|} C^{ij} = 1/|X|$). This distance is not defined if either X or $\widehat{X}$ is not defined. Figure 5 shows the mean Wasserstein distances of two algorithms over 100 simulation trials. Process noise, measurement noise, and clutter are independently generated at each trial. It can be seen from Figure 5 that the proposed iterative mixture component pruning algorithm is superior to the typical algorithm at most time steps. The proposed iterative mixture component pruning algorithm is worse than typical algorithm when spawned target is generated and two or more targets are close to each other. Two PHD peaks of two close targets may be regarded as one PHD peak in the proposed algorithm as a result of low-entropy tendency of entropy distribution. Then, some targets are not detected.

Figure 6 shows the estimates of target numbers of two algorithms. It is obvious that the estimates of target number of proposed algorithm are closer to the ground truth than typical algorithm at most time steps.

Figure 7 shows the mean component numbers of two algorithms after component pruning operations over 100 simulation trials. The component numbers of proposed algorithm are smaller than typical algorithm.

The case of low signal-to-noise rate (SNR) is yet considered for the further comparison of two algorithms. λ_c is assumed 80 in this low SNR case. The corresponding

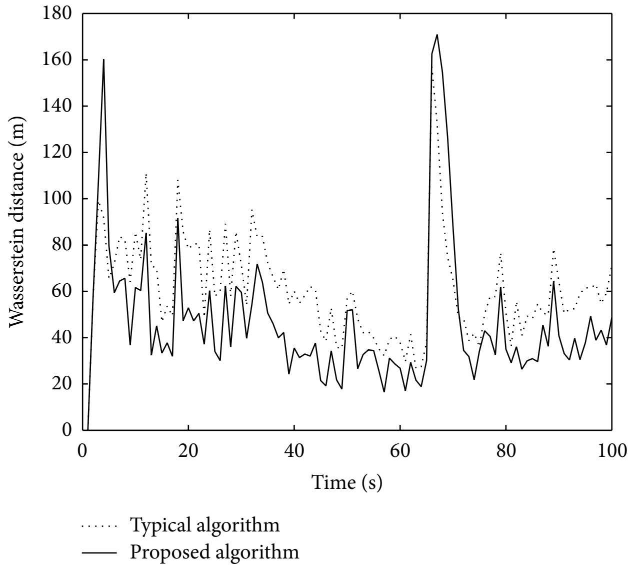


Figure 5: The averaged Wasserstein distances.

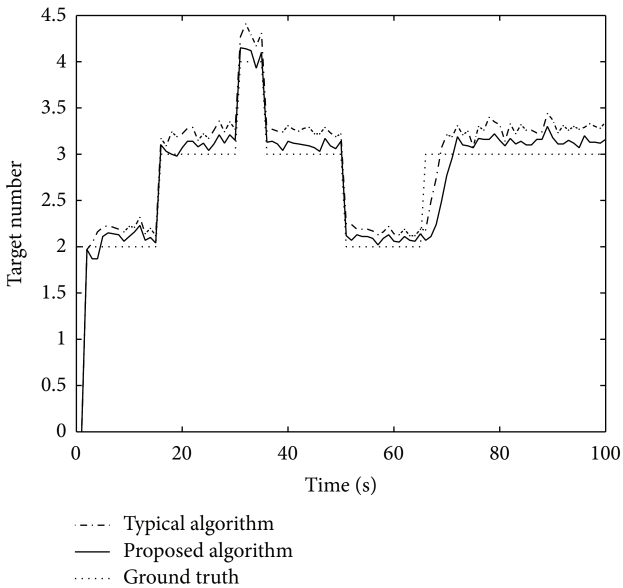


Figure 6: Estimates of target numbers.

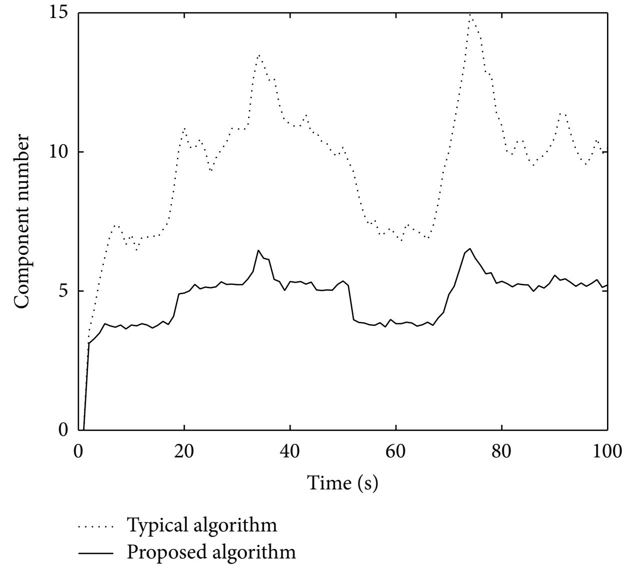


Figure 7: The averaged component numbers.

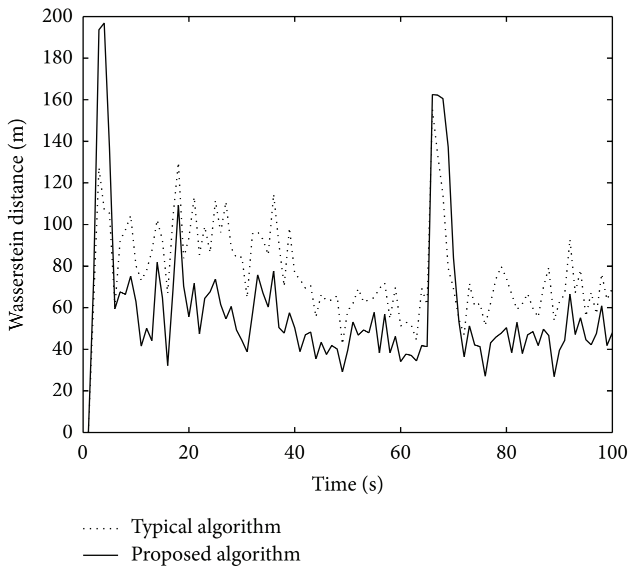


Figure 8: The averaged Wasserstein distances under low SNR.

Wasserstein distances, target number estimates, and component numbers are presented in Figures 8, 9, and 10. It can be seen that the proposed iterative mixture component pruning algorithm is also superior to the typical mixture component pruning algorithm based on thresholds in low SNR case.

5. Conclusion

An iterative mixture component pruning algorithm is proposed for the Gaussian mixture PHD filter. The entropy distribution of the mixture weights is used as the prior distribution of mixture parameters. The update formula of the mixture weight is derived by Lagrange multiplier and Lambert W function. When the mixture weight becomes negative during the iteration procedure, the corresponding mixture component is pruned by setting the weight to zero. Simulation results show that the proposed iterative mixture component pruning algorithm is superior to the typical mixture component pruning algorithm based on thresholds at most time steps.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

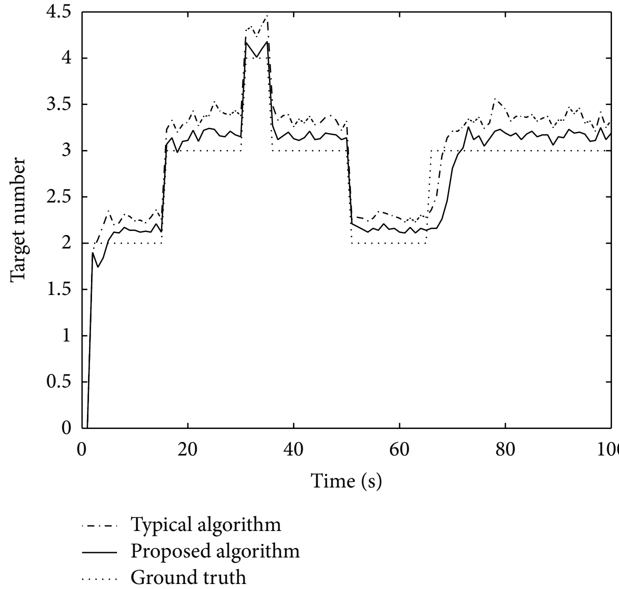


FIGURE 9: Estimates of target numbers under low SNR.

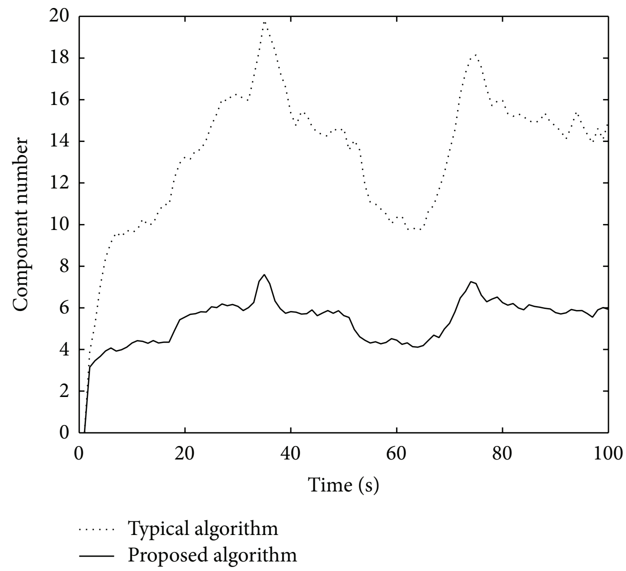


FIGURE 10: The averaged component numbers under low SNR.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant no. 61304261), the Senior Professionals Scientific Research Foundation of Jiangsu University (Grant no. 12JDG076), and a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

[1] G. W. Pulford, "Taxonomy of multiple target tracking methods," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 291–304, 2005.

[2] S. Blackman and R. F. Popoli, *Design and Analysis of Modern Tracking System*, Artech House, Boston, Mass, USA, 1999.

[3] Y. Barshalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, Yaakov Bar-Shalom, Storrs, Conn, USA, 1995.

[4] Y. Barshalom, "Tracking in a cluttered enviroment probabilistic data association," *Automatica*, vol. 11, no. 5, pp. 451–460, 1975.

[5] Y. Barshalom, "Tracking methods in a multitarget enviroment," *IEEE Transactions on Automatic Control*, vol. 23, no. 4, pp. 618–626, 1978.

[6] K. C. Chang and Y. Barshalom, "Joint probabilistic data association for multitarget tracking with possibly unresolved measurements and maneuvers," *IEEE Transactions on Automatic Control*, vol. 29, no. 7, pp. 585–594, 1984.

[7] R. P. S. Mahler, "Multi-target Bayes filtering via first-order multi-target moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[8] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion. 1em Plus 0.5em Minus 0.4em*, Artech House, Norwood, Mass, USA, 2007.

[9] O. Erdinc, P. Willett, and Y. Bar-Shalom, "The bin-occupancy filter and its connection to the PHD filters," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4232–4246, 2009.

[10] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[11] S.-A. Pasha, B.-N. Vo, and H.-D. Tuan, "A Gaussian mixture PHD filter for jump Markov system models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 919–936, 2009.

[12] H. Sidenbladh, "Multi-target particle filtering for the probability hypothesis density," in *Proceedings of the 6th International Conference of Information Fusion*, vol. 2, pp. 800–806, Cairns, Australia, July 2003.

[13] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo implementation of the PHD filter for multi-target tracking," in *Proceedings of the International Conference on Information Fusion*, pp. 792–799, Cairns, Australia, 2003.

[14] T. Zajic and R. P. S. Mahler, "A particle-systems implementation of the PHD multi-target tracking filter," in *Signal Processing, Sensor Fusion and Target Recognition XI*, I. Kadar, Ed., vol. 5096 of *Proceedings of SPIE*, pp. 291–299, 2003.

[15] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multi-target filtering with random finite sets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, 2005.

[16] N. Whiteley, S. Singh, and S. Godsill, "Auxiliary particle implementation of probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 7, pp. 1437–1454, 2010.

[17] D.-E. Clark and B.-N. Vo, "Convergence analysis of the Gaussian mixture PHD filter," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1204–1212, 2007.

[18] D.-E. Clark and J. Bell, "Convergence results for the particle PHD filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2252—2261, 2006.

[19] A. M. Johansen, S. S. Singh, A. Doucet, and B.-N. Vo, "Convergence of the SMC implementation of the PHD filte," *Methodology and Computing in Applied Probability*, vol. 8, no. 2, pp. 265–291, 2006.

[20] R. P. S. Mahler, "PHD filters of higher order in target number," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, 2007.

[21] B. Vo, B. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, part 2, pp. 3553–3567, 2007.

[22] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 409–423, 2009.

[23] K. Punithakumar, T. Kirubarajan, and A. Sinha, "Multiple-model probability hypothesis density filter for tracking maneuvering targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 1, pp. 87–98, 2008.

[24] D. Dunne and T. Kirubarajan, "Multiple model multi-Bernoulli filters for manoeuvering targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 4, pp. 2627–2692, 2013.

[25] D.-E. Clark and J. Bell, "Bayesian multiple target tracking in forward scan sonar images using the PHD filter," *IEE Proceedings-Radar Sonar and Navigation*, vol. 152, no. 5, pp. 327–334, 2005.

[26] D.-E. Clark, I.-T. Ruiz, Y. Petillot et al., "Particle PHD filter multiple target tracking in sonar image," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 409–416, 2007.

[27] E. Maggio, M. Taj, and A. Cavallaro, "Efficient multitarget visual tracking using random finite sets," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1016–1027, 2008.

[28] Y. D. Wang, J. K. Wu, A. A. Kassim et al., "Data-driven probability hypothesis density filter for visual tracking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1085–1095, 2008.

[29] E. Maggio and A. Cavallaro, "Learning scene context for multiple object tracking," *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1873–1884, 2009.

[30] K. Gilholm and D. Salmond, "Spatial distribution model for tracking extended objects," *IEE Proceedings Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 364–371, 2005.

[31] R. P. S. Mahler, "PHD filters for nonstandard targets, I: extended targets," in *Proceedings of the 12th International Conference on Information Fusion, IEEE*, vol. 1–4, pp. 915–921, 2009.

[32] K. Granstrom, C. Lundquist, and U. Orguner, "A Gaussian mixture PHD filter for extended target," in *Proceedings of the International Conference on Information Fusion*, Edinburgh, UK, 2010.

[33] K. Granstrom, C. Lundquist, and U. Orguner, "Extended target tracking using a Gaussian-mixture PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3268–3286, 2012.

[34] K. Granstrom and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5657–5671, 2012.

[35] C. Lundquist, K. Granstrom, and U. Orguner, "An extended target CPHD filter and a gamma gaussian inverse wishart implementation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 472–483, 2013.

[36] K. Granstrom and U. Orguner, "On spawning and combination of extended/group targets modeled with random matrices," *IEEE Transactions on Signal Processing*, vol. 61, no. 3, pp. 678–692, 2013.

[37] F. Lian, C. Han, J. Liu, and H. Chen, "Convergence results for the Gaussian mixture implementation of the extended-target PHD filter and its extended Kalman filtering approximation," *Journal of Applied Mathematics*, vol. 2012, Article ID 141727, 20 pages, 2012.

[38] B.-N. Vo, B.-T. Vo, and R. P. S. Mahler, "Closed-form solutions to forward-backward smoothing," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 2–17, 2012.

[39] D. J. Salmond, "Mixture reduction algorithms for target tracking in clutter," in *Signal and Data Processing of Small Targets Conference*, vol. 1305 of *Proceedings of SPIE*, pp. 434–445, 1990.

[40] D. J. Salmond, "Mixture reduction algorithms for point and extendedobject tracking in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 2, pp. 667–686, 2009.

[41] J. E. Harmse, "Reduction of Gaussian mixture models by maximum similarity," *Journal of Nonparametric Statistics*, vol. 22, no. 5-6, pp. 703–709, 2010.

[42] M. F. Huber and U. D. Hanebeck, "Progressive Gaussian mixture reduction," in *Proceedings of the 11th International Conference on information Fusion*, pp. 1–8, Cologne, Germany, 2008.

[43] O. C. Schrempf, O. Feiermann, and U. D. Hanebeck, "Optimal mixture approximation of the product of mixtures," in *Proceedings of the 8th International Conference on Information Fusion*, pp. 85–92, Philadelphia, Pa, USA, 2005.

[44] D. Schieferdecker and M. F. Huber, "Gaussian mixture reduction via clustering," in *Proceedings of the 12th International Conference on Information Fusion (FUSION '09)*, pp. 1536–1543, Seattle, Wash, USA, July 2009.

[45] A. R. Runnalls, "Kullback-Leibler approach to Gaussian mixture reduction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 989–999, 2007.

[46] A. Caticha and R. Preuss, "Maximum entropy and Bayesian data analysis: entropic prior distributions," *Physical Review E*, vol. 70, no. 4, Article ID 046127, 12 pages, 2004.

[47] M. Brand, "Structure learning in conditional probability models via an entropic prior and parameter extinction," *Neural Computation*, vol. 11, no. 5, pp. 1155–1182, 1999.

[48] R. M. Corless, G. H. Gonnet, D. E. G. Hare, and D. E. Knuth, "On the Lambert *W* function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[49] J. Hoffman and R. P. S. Mahler, "Multitarget miss distance via optimal assignment," *IEEE Transactions on Systems, Man, and Cybernetics-Part A*, vol. 34, no. 3, pp. 327–336, 2004.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 450149, 7 pages
http://dx.doi.org/10.1155/2014/450149

Research Article

The Gerber-Shiu Discounted Penalty Function of Sparre Andersen Risk Model with a Constant Dividend Barrier

Yujuan Huang[1] and Wenguang Yu[2]

[1] *School of Science, Shandong Jiaotong University, Jinan 250357, China*
[2] *School of Insurance, Shandong University of Finance and Economics, Jinan 250014, China*

Correspondence should be addressed to Wenguang Yu; yuwg@mail.sdu.edu.cn

Received 19 January 2014; Revised 1 May 2014; Accepted 18 June 2014; Published 15 July 2014

Academic Editor: Guangchen Wang

This paper constructs a Sparre Andersen risk model with a constant dividend barrier in which the claim interarrival distribution is a mixture of an exponential distribution and an Erlang(n) distribution. We derive the integro-differential equation satisfied by the Gerber-Shiu discounted penalty function of this risk model. Finally, we provide a numerical example.

1. The Risk Model

Consider a Sparre Andersen risk model,

$$U(t) = u + ct - \sum_{i=1}^{N(t)} X_i \quad \text{for } t \ge 0, \tag{1}$$

where $u \ge 0$ represents the initial capital, c is the insurer's rate of premium income per unit time, and $\{N(t), t \ge 0\}$ is the claim number process representing the number of claims up to time t. $\{X_i, i \ge 1\}$ is a sequence of i.i.d. random variables representing the individual claim amounts with distribution function $F(x)$ and density function $f(x)$ with mean μ. We assume that $\{N(t), t \ge 0\}$ and $\{X_i, i \ge 1\}$ are independent. Let $\{T_i, i \ge 1\}$ be sequence i.i.d. random variables, which represent the claim interarrival times, and T_i has a density function $K(t)$,

$$K(t) = \beta_1 \lambda e^{-\lambda t} + \beta_2 e^{-\lambda t} \frac{\lambda^n t^{n-1}}{(n-1)!}, \quad t \ge 0, \tag{2}$$

where $n \ge 1$ is a positive integer, $\lambda \ge 0$, $\beta_1, \beta_2 \ge 0$, and $\beta_1 + \beta_2 = 1$. We further assume that $cE[T_i] > E[X_i]$ for all i, which ensure that $\lim_{t\to\infty} U(t) = \infty$ almost surely. Throughout the paper we use the convention that $\sum_{i=1}^{0} X_i = 0$.

In recent years the Sparre Andersen model has been studied extensively. Ruin probabilities and many ruin related quantities such as the marginal and joint defective distributions of the time to ruin, the deficit at ruin, the surplus prior to ruin, and the claim size causing ruin have received considerable attention. Some related results can be found in Cai and Dickson [1], Sun and Yang [2], Gerber and Shiu [3], and Ko [4]. Li and Garrido [5] consider a compound renewal (Sparre Andersen) risk process in the presence of a constant dividend barrier in which the claim waiting times are generalized Erlang(n) distributed. The Sparre Andersen model with phase-type interclaim times has been studied by Ren [6]. Ng and Yang [7] study the ruin probability and the distribution of the severity of ruin in risk models with phase-type claims. Landriault and Willmot [8] study the Gerber-Shiu function in a Sparre Andersen model with general interclaim times. Yang and Zhang [9] study a Sparre Andersen model in which the interclaim times are generalized Erlang(n) distributed. They assume that the premium rate is a step function depending on the current surplus level. Landriault and Sendova [10] generalize the Sparre Andersen dual risk model with Erlang(n) interinnovation times by adding a budget-restriction strategy. Shi and Landriault [11] utilize the multivariate version of Lagrange expansion theorem to obtain a series expansion for the density of the time to ruin under a more general distribution assumption, namely, the combination of n exponentials. Yang and Sendova [12] study the Sparre Andersen dual risk model in which the times between positive gains are independently and identically distributed and have a generalized Erlang(n) distribution.

The barrier strategy was initially proposed by De Finetti [13] for a binomial model. From then on, barrier strategies have been studied in a number of papers and books, including Lin et al. [14], Dickson and Waters [15], Li and Lu [16], Yu [17–19], Yao et al. [20], Zhu [21], Tan et al. [22], and references therein for details. The purpose of this paper is to extend some results in Li and Garrido [5] and Yang and Zhang [9]. We study the Sparre Andersen risk model with a constant dividend barrier and the claim interarrival distribution is a mixture of an exponential distribution and an Erlang(n) distribution.

The contents of this paper are organized as follows. Section 2 introduces the risk model. In Section 3, we derive the higher-order integro-differential equation for the Gerber-Shiu discounted penalty function. Finally, in the special case we provide the numerical example in Section 4.

2. The Risk Model

Let $U_b(t)$ be the surplus process with initial surplus $U_b(0) = u$ under the barrier strategy. Thus, it can be expressed as

$$dU_b(t) = \begin{cases} cdt - dS(t) & U_b(t) < b \\ -dS(t) & U_b(t) \geq b, \end{cases} \tag{3}$$

where $S(t) = \sum_{i=1}^{N(t)} X_i$. Define $T_b = \inf\{t : U_b(t) < 0\}$ to be the first time that the surplus becomes negative. The stopping time T_b is referred to as the time of ruin. Let $\psi_b(u) = \Pr(T_b < \infty)$ be the ruin probability.

In this paper, we will study the time of ruin T_b and its related functions such as the surplus before ruin $U_b(T_b-)$ and the deficit at ruin $|U_b(T_b)|$. By using a renewal equation approach, we will be able to get a number of analytic and probabilistic properties of these quantities. Our analysis will involve the Gerber-Shiu discounted penalty function that is defined below.

Let $\omega(x, y)$, $0 \leq x, y < \infty$, be a nonnegative function. For $\delta \geq 0$, define

$$\begin{aligned} m_b(u) = E\Big[&e^{-\delta T_b}\omega\left(U(T_b-), |U(T_b)|\right) \\ &I(T_b < \infty) \mid U(0) = u\Big], \end{aligned} \tag{4}$$

where $I(\cdot)$ is the indicator function, $I(T_b < \infty) = 1$ if $T_b < \infty$, and $I(T_b < \infty) = 0$ otherwise. The function $m_b(u)$ in (4) is useful for deriving results in connection with joint and marginal distributions of T_b, $U_b(T_b-)$ and $|U_b(T_b)|$. While δ may be interpreted as a force of interest, function (4) may also be viewed in terms of a Laplace transform with δ serving as the argument. In particular, if we let $\omega(x, y) = 1$, (4) is the Laplace transform of the time of ruin T_b. If we let $\delta = 0$ and $\omega(x, y) = 1$, then $m_b(u)$ becomes the ruin probability $\psi(u)$. If we let $\delta = 0$ and $\omega(x, y) = I(x \leq z_1)I(y \leq z_2)$, (4) becomes the joint df of the surplus before ruin and the deficit at ruin. Furthermore, if $\delta = 0$ and $\omega(x, y) = x_1^n$, we obtain the nth moment of the surplus before ruin. Likewise, if $\delta = 0$ and $\omega(x, y) = x_2^n$, we obtain the nth moment of the deficit at ruin. For other functions of interest, see Gerber and Shiu [23] and Lin and Willmot [24]. Let f^* denote the Laplace transform of the function f, that is, $f^*(s) = \int_0^\infty e^{-sx} f(x)dx$.

3. An Integro-Differential Equation

In this section, we show $m_b(u)$ satisfies a higher-order integro-differential equation.

Lemma 1. *Assume $s > u$; then $H(u,s) = K((s-u)/c)e^{-\delta((s-u)/c)}$ satisfies the following differential equation:*

$$\begin{aligned} &\sum_{j=0}^{n-1} C_{n-1}^j c^j (-\lambda-\delta)^{n-1-j} \frac{\partial^j H(u,s)}{\partial u^j} \\ &\quad = (-1)^{n-1}\beta_2\lambda^n e^{-(\lambda+\delta)((s-u)/c)}, \end{aligned} \tag{5}$$

with the boundary conditions when $s = u$,

$$\begin{aligned} \frac{\partial^k H(u,s)}{\partial u^k} &= \beta_1\lambda\left(\frac{\lambda+\delta}{c}\right)^k, \quad k = 0, 1, 2, \ldots, n-2, \\ \frac{\partial^{n-1} H(u,s)}{\partial u^{n-1}} &= \beta_1\lambda\left(\frac{\lambda+\delta}{c}\right)^{n-1} + \beta_2\lambda^n\left(-\frac{1}{c}\right)^{n-1}. \end{aligned} \tag{6}$$

Proof. Note that $H(u,s) = [\beta_1\lambda + (\lambda^n\beta_2/(n-1)!)((\lambda+\delta)/c)^{n-1}]e^{-(\lambda+\delta)((s-u)/c)}$. Taking derivative with respect to variable u for k times and by induction, we can obtain

$$\begin{aligned} &\frac{\partial^k H(u,s)}{\partial u^k} \\ &\quad = -\sum_{j=0}^{k-1} C_k^j \left(-\frac{\lambda+\delta}{c}\right)^{k-j} \frac{\partial^j H(u,s)}{\partial u^j} \\ &\qquad + \frac{\beta_2\lambda^n}{(n-1-k)!}\left(\frac{s-u}{c}\right)^{n-1-k}\left(-\frac{1}{c}\right)^k e^{-(\lambda+\delta)((s-u)/c)}, \\ &\qquad\qquad 0 \leq k \leq n-1. \end{aligned} \tag{7}$$

When $k = n-1$, one gets (5). On substituting $s = u$ in (7), we get the boundary conditions. □

Theorem 2. *The Gerber-Shiu discounted penalty function $m_b(u)$ satisfies the higher-order integro-differential equation*

$$\begin{aligned} &\sum_{k=0}^{n} C_n^k c^k (-\lambda-\delta)^{n-k} \frac{d^k m_b(u)}{du^k} \\ &\quad = \left[\beta_2(-\lambda)^n + \beta_1(-\lambda)(-\lambda-\delta)^{n-1}\right] \\ &\qquad \times \int_0^\infty m_b(u-x)\,dF(x) \\ &\qquad - \beta_1\lambda\sum_{k=1}^{n-1} C_{n-1}^k c^k (-\lambda-\delta)^{n-k-1}\frac{d^k}{du^k} \\ &\qquad \times \left(\int_0^\infty m_b(u-x)\,dF(x)\right). \end{aligned} \tag{8}$$

Proof. Let t be the time of the first claim and let x be the amount of the claim. There are two possibilities. First, $t < (b-u)/c$ and the surplus has not yet reached the barrier. In this case, the surplus immediately before time t is $u+ct$. Second, $t \geq (b-u)/c$ and the surplus immediately before time t is b. And since the "probability" that the claim occurs at time t is $K(t)dt$ and the "probability" of the claim amount being x is $dF(x)$, we have, for $0 \leq u \leq b$,

$$\begin{aligned} m_b(u) = &\int_0^{(b-u)/c} K(t)\, e^{-\delta t} \\ &\times \left[\int_0^{u+ct} m_b(u+ct-x)\, dF(x) \right. \\ &\left. + \int_{u+ct}^{\infty} w(u+ct, x-u-ct)\, dF(x) \right] dt \\ &+ \int_{((b-u)/c)}^{\infty} K(t)\, e^{-\delta t} \\ &\times \left[\int_0^{b} m_b(b-x)\, dF(x) \right. \\ &\left. + \int_b^{\infty} w(b, x-b)\, dF(x) \right] dt. \end{aligned} \tag{9}$$

Using the substitution $s = u + ct$, we have

$$\begin{aligned} m_b(u) = &\int_u^{b} K\left(\frac{s-u}{c}\right) e^{-\delta((s-u)/c)} \\ &\times \left[\int_0^{s} m_b(s-x)\, dF(x) \right. \\ &\left. + \int_{s}^{\infty} w(s, x-s)\, dF(x) \right] \frac{1}{c} ds \\ &+ \int_b^{\infty} K\left(\frac{s-u}{c}\right) e^{-\delta((s-u)/c)} \\ &\times \left[\int_0^{b} m_b(b-x)\, dF(x) \right. \\ &\left. + \int_b^{\infty} w(b, x-b)\, dF(x) \right] \frac{1}{c} ds \end{aligned} \tag{10}$$

which implies that

$$\begin{aligned} c m_b(u) = &\int_u^{b} H(u,s) \int_0^{\infty} m_b(s-x)\, dF(x)\, ds \\ &+ \int_b^{\infty} H(u,s) \int_0^{\infty} m_b(b-x)\, dF(x)\, ds, \end{aligned} \tag{11}$$

where $H(u,s)$ is defined in Lemma 1. Differentiating the above equation k times and using condition (6) yield

$$\begin{aligned} c\frac{d^k m_b(u)}{du^k} &= -\beta_1 \lambda \sum_{i=0}^{k-1} \left(\frac{\lambda+\delta}{c}\right)^{k-1-i} \frac{d^i}{du^i} \int_0^{\infty} m_b(u-x)\, dF(x) \\ &+ \int_u^{b} \frac{\partial^k H(u,s)}{\partial u^k} \int_0^{\infty} m_b(s-x)\, dF(x)\, ds \\ &+ \int_b^{\infty} \frac{\partial^k H(u,s)}{\partial u^k} \int_0^{\infty} m_b(b-x)\, dF(x)\, ds. \end{aligned} \tag{12}$$

Multiplying (12) by $c^k(-\lambda-\delta)^{n-1-k} C_{n-1}^k$ for $k = 0, 1, 2, \ldots, n-1$, then adding up these equations, and using (5), we obtain

$$\begin{aligned} &\sum_{k=0}^{n-1} C_{n-1}^k (-\lambda-\delta)^{n-1-k} c^{k+1} \frac{d^k m_b(u)}{du^k} \\ &= \beta_2 \lambda^n (-1)^{n-1} \int_u^{b} e^{-(\lambda+\delta)((s-u)/c)} \\ &\times \int_0^{\infty} m_b(s-x)\, dF(x)\, ds \\ &+ \beta_2 \lambda^n (-1)^{n-1} \int_b^{\infty} e^{-(\lambda+\delta)((s-u)/c)} \\ &\times \int_0^{\infty} m_b(b-x)\, dF(x)\, ds \\ &- \beta_1 \lambda \sum_{k=1}^{n-1} C_{n-1}^k (-\lambda-\delta)^{n-1-k} c^k \\ &\times \left[\sum_{i=0}^{k-1} \left(\frac{\lambda+\delta}{c}\right)^{k-1-i} \frac{d^i}{du^i} \left(\int_0^{\infty} m_b(u-x)\, dF(x) \right) \right]. \end{aligned} \tag{13}$$

Differentiating (13) again, we have

$$\begin{aligned} &\sum_{k=0}^{n-1} C_{n-1}^k (-\lambda-\delta)^{n-1-k} c^{k+1} \frac{d^{k+1} m_b(u)}{du^{k+1}} \\ &= (-1)^n \beta_2 \lambda^n \int_0^{\infty} m_b(u-x)\, dF(x) \\ &- \beta_1 \lambda \sum_{k=1}^{n-1} C_{n-1}^k (-\lambda-\delta)^{n-1-k} c^k \end{aligned}$$

$$\times\left[\sum_{i=0}^{k-1}\left(\frac{\lambda+\delta}{c}\right)^{k-1-i}\frac{d^{i+1}}{du^{i+1}}\left(\int_0^\infty m_b(u-x)\,dF(x)\right)\right]$$
$$+\frac{\lambda+\delta}{c}\beta_2\lambda^n(-1)^{n-1}$$
$$\times\left[\int_u^b e^{-(\lambda+\delta)((s-u)/c)}\int_0^\infty m_b(s-x)\,dF(x)\,ds\right.$$
$$\left.+\int_b^\infty e^{-(\lambda+\delta)((s-u)/c)}\int_0^\infty m_b(b-x)\,dF(x)\,ds\right] \quad (14)$$

which, together with (13), implies

$$\sum_{k=0}^{n-1}C_{n-1}^k c^{k+1}(-\lambda-\delta)^{n-1-k}\frac{d^{k+1}m_b(u)}{du^{k+1}}$$
$$+\sum_{k=0}^{n-1}C_{n-1}^k c^k(-\lambda-\delta)^{n-k}\frac{d^k m_b(u)}{du^k}$$
$$=-\beta_1\lambda\sum_{k=1}^{n-1}C_{n-1}^k(-\lambda-\delta)^{n-1-k}c^k$$
$$\times\left[\sum_{i=0}^{k-1}\left(\frac{\lambda+\delta}{c}\right)^{k-1-i}\frac{d^{i+1}}{du^{i+1}}\left(\int_0^\infty m_b(u-x)\,dF(x)\right)\right]$$
$$-\beta_1\lambda\sum_{k=1}^{n-1}C_{n-1}^k(-\lambda-\delta)^{n-k}c^{k-1}$$
$$\times\left[\sum_{i=0}^{k-1}\left(\frac{\lambda+\delta}{c}\right)^{k-1-i}\frac{d^{i}}{du^{i}}\left(\int_0^\infty m_b(u-x)\,dF(x)\right)\right]$$
$$+(-1)^n\beta_2\lambda^n\int_0^\infty m_b(u-x)\,dF(x). \quad (15)$$

Moreover, note that

$$\sum_{k=0}^{n-1}C_{n-1}^k c^{k+1}(-\lambda-\delta)^{n-1-k}\frac{d^{k+1}m_b(u)}{du^{k+1}}$$
$$=c^n\frac{d^n m_b(u)}{du^n}+\sum_{k=1}^{n-1}C_{n-1}^{k-1}c^k(-\lambda-\delta)^{n-k}\frac{d^k m_b(u)}{du^k},$$
$$\sum_{k=0}^{n-1}C_{n-1}^k c^k(-\lambda-\delta)^{n-k}\frac{d^k m_b(u)}{du^k}$$
$$=(-\lambda-\delta)^n m_b(u)+\sum_{k=1}^{n-1}C_{n-1}^k c^k(-\lambda-\delta)^{n-k}\frac{d^k m_b(u)}{du^k}. \quad (16)$$

So, it follows from (16) that

$$\sum_{k=0}^{n-1}C_{n-1}^k c^{k+1}(-\lambda-\delta)^{n-1-k}\frac{d^{k+1}m_b(u)}{du^{k+1}}$$
$$+\sum_{k=0}^{n-1}C_{n-1}^k c^k(-\lambda-\delta)^{n-k}\frac{d^k m_b(u)}{du^k} \quad (17)$$
$$=\sum_{k=0}^{n}C_n^k c^k(-\lambda-\delta)^{n-k}\frac{d^k m_b(u)}{du^k}$$

and thus the result follows from (15) and (17). □

Remark 3. Letting $\beta_1 = 0$, $\beta_2 = 1$, $n = 2$ in (8), we get the integro-differential equation for Erlang (2) risk model with a constant dividend barrier of Li and Garrido [5].

Remark 4. Letting $\beta_1 = 0$, $\beta_2 = 1$, $n = 2$, $b = \infty$ in (8), we obtain the integro-differential equation for Erlang (2) risk model with no dividend barrier, which has been considered in Dickson and Hipp [25].

Remark 5. Letting $n = 1$, $b = \infty$ in (8), we derive the integro-differential equation for classical risk model. For details, see Gerber and Shiu [23].

Remark 6. Letting $n = 1$, the case has been studied in Lin et al. [14].

Remark 7. Letting $b = \infty$, the case has been studied in Zhao and Yin [26].

Theorem 8. *The Laplace transform of $m_b(u)$ is*

$$m_b^*(s)$$
$$=\frac{A\int_0^\infty e^{-su}\int_u^\infty \omega(u,x-u)\,dF(x)\,du+G(s)+D(s)}{(sc-\lambda-\delta)^n-\left[\beta_2(-\lambda)^n-\lambda\beta_1(sc-\lambda-\delta)^{n-1}\right]f^*(s)}, \quad (18)$$

where

$$A=\beta_2(-\lambda)^n+\beta_1(-\lambda)(-\lambda-\delta)^{n-1},$$
$$G(s)=\sum_{k=1}^{n-1}\sum_{j=0}^{k-1}C_n^k(-\lambda-\delta)^{n-k}c^k s^{k-1-j}m_b^{(j)}(0),$$
$$D(s)=\beta_1\lambda\sum_{k=2}^{n-1}C_{n-1}^k(-\lambda-\delta)^{n-1-k}c^k$$
$$\times\sum_{j=1}^{k-1}s^{k-1-j}\sum_{l=0}^{j-1}m_b^{(l)}(0)\,f^{(j-1-l)}(0) \quad (19)$$
$$-\beta_1\lambda\sum_{k=1}^{n-1}C_{n-1}^k(-\lambda-\delta)^{n-1-k}c^k$$
$$\times\int_0^\infty e^{-su}\left[\frac{d^k}{du^k}\int_u^\infty \omega(u,x-u)\,dF(x)\right]du.$$

Proof. It is easy to see that

$$\int_0^\infty e^{-su}\frac{d^k m_b(u)}{du^k}du = s^k m_b^*(s) - \sum_{j=0}^{k-1} s^{k-1-j} m_b^{(j)}(0), \quad (20)$$

$$\begin{aligned}\int_0^\infty e^{-su}\int_0^\infty m_b(u-x)\,dF(x)\,du &= \int_0^\infty e^{-su}\int_0^u m_b(u-x)\,dF(x)\,du \\ &\quad + \int_0^\infty e^{-su}\int_u^\infty \omega(u,x-u)\,dF(x)\,du \\ &= sm_b^*(s)\,f^*(s) + \int_0^\infty e^{-su}\int_u^\infty \omega(u,x-u)\,dF(x)\,du,\end{aligned} \quad (21)$$

$$\begin{aligned}\int_0^\infty e^{-su}\frac{d^k}{du^k}\left(\int_0^\infty m_b(u-x)\,dF(x)\right)du &= \int_0^\infty e^{-su}\frac{d^k}{du^k}\left(\int_0^u m_b(u-x)\,dF(x)\right)du \\ &\quad + \int_0^\infty e^{-su}\frac{d^k}{du^k}\left(\int_u^\infty \omega(u,x-u)\,dF(x)\right)du \\ &= s^k m_b^*(s)\,f^*(s) - \sum_{j=1}^{k-1} s^{k-1-j}\sum_{l=0}^{j-1} m_b^{(l)}(0)\,f^{(j-1-l)}(0) \\ &\quad + \int_0^\infty e^{-su}\frac{d^k}{du^k}\left(\int_u^\infty \omega(u,x-u)\,dF(x)\right)du.\end{aligned} \quad (22)$$

Taking the Laplace transform on both sides of (8), and together with (20), (21), and (22), we have

$$\begin{aligned}&\sum_{k=0}^{n} C_n^k c^k(-\lambda-\delta)^{n-k}\left(s^k m_b^*(s) - \sum_{j=0}^{k-1} s^{k-1-j} m_b^{(j)}(0)\right) \\ &= \left[\beta_2(-\lambda)^n + \beta_1(-\lambda)(-\lambda-\delta)^{n-1}\right] \\ &\quad \times \left[m_b^*(s)\,f^*(s) + \int_0^\infty e^{-su}\int_u^\infty \omega(u,x-u)\,dF(x)\,du\right] \\ &\quad - \beta_1\lambda\left[(sc-\lambda-\delta)^{n-1} - (-\lambda-\delta)^{n-1}\right] m_b^*(s)\,f^*(s) \\ &\quad + \beta_1\lambda\sum_{k=2}^{n-1} C_{n-1}^k c^k(-\lambda-\delta)^{n-1-k} \\ &\quad \times \sum_{j=1}^{k-1} s^{k-1-j}\sum_{l=0}^{j-1} m_b^{(l)}(0)\,f^{(j-1-l)}(0) \\ &\quad - \beta_1\lambda\sum_{k=1}^{n-1} C_{n-1}^k c^k(-\lambda-\delta)^{n-1-k} \\ &\quad \times \int_0^\infty e^{-su}\left[\frac{d^k}{du^k}\int_u^\infty \omega(u,x-u)\,dF(x)\right]du\end{aligned} \quad (23)$$

which implies (8). □

Lemma 9. *Let δ be strictly positive and n is a positive integer; then the equation*

$$\begin{aligned}(sc-\lambda-\delta)^n &= f^*(s)\left[\beta_2(-\lambda)^n + \beta_1(-\lambda)(-\lambda-\delta)^{n-1}\right] \\ &\quad - f^*(s)\,\lambda\beta_1\left[(sc-\lambda-\delta)^{n-1} - (-\lambda-\delta)^{n-1}\right]\end{aligned} \quad (24)$$

has exact n roots $s_l(\delta)$ with $\mathrm{Re}(s_l(\delta)) > 0$ $(l = 1,2,3,\ldots,n)$.

Proof. When $s = 0$,we have

$$\left|\left[\beta_2(-\lambda)^n + \beta_1(-\lambda)(-\lambda-\delta)^{n-1}\right] f^*(0)\right| < \left|(-\lambda-\delta)^n\right|. \quad (25)$$

So for $\rho > 0$ sufficiently big, the inequality

$$\begin{aligned}&\left|(sc-\lambda-\delta)^n\right| \\ &> \left|\left[\beta_2(-\lambda)^n + \beta_1(-\lambda)(-\lambda-\delta)^{n-1}\right.\right. \\ &\quad \left.\left. -\lambda\beta_1(sc-\lambda-\delta)^{n-1} + \lambda\beta_1(-\lambda-\delta)^{n-1}\right] f^*(s)\right|\end{aligned} \quad (26)$$

holds on the imaginary axis and on the semicircle $\{s \in £,\ \mathrm{Re}(s) > 0,\ |s| = \rho\}$. By Rouches theorem (20) has exact n roots on the right-half plane. □

4. Numerical Illustration for Ruin Probability

In this section, we give the numerical illustration for $m_b(u)$ when the claim number process has Erlang (2) process ($\beta_1 = 0$, $\beta_2 = 1$, $n = 2$), $\delta = 0$ and $w(x,y) = 1$. At this time, $m_b(u)$ turns to ruin probability $\psi_b(u)$. By conditioning on the time of the first claim we have, for $0 \le u \le b$,

$$\begin{aligned}m_b(u) &= \int_0^{((b-u)/c)} K_1(t)\,\gamma_b(u+ct)\,dt \\ &\quad + \int_{((b-u)/c)}^\infty K_1(t)\,\gamma_b(b)\,dt,\end{aligned} \quad (27)$$

where

$$\gamma_b(t) = \int_0^t m_b(t-x)\,dF(x) + 1 - F(t). \quad (28)$$

Substituting $K_1(t) = \lambda^2 te^{-\lambda t}$ into (27), we obtain

$$\begin{aligned}m_b(u) &= \left(\frac{\lambda}{c}\right)^2\int_u^b (t-u)\,e^{-(\lambda/c)(t-u)}\gamma_b(t)\,dt \\ &\quad + \gamma_b(b)\,e^{-(\lambda/c)(t-u)}\left[1 + \frac{\lambda}{c}(b-u)\right].\end{aligned} \quad (29)$$

Differentiating (29) with respect to u, we have, for $0 \le u \le b$,

$$\begin{aligned}m_b'(u) &= \frac{\lambda}{c}m_b(u) - \left(\frac{\lambda}{c}\right)^2\int_u^b e^{-(\lambda/c)(t-u)}\gamma_b(t)\,dt \\ &\quad - \frac{\lambda\gamma_b(b)}{c}e^{-(\lambda/c)(b-u)}.\end{aligned} \quad (30)$$

Differentiating (30) again with respect to u, we have

$$\begin{aligned} m_b''(u) &= \frac{\lambda}{c} m_b'(u) \\ &\quad - \frac{\lambda}{c}\left[\left(\frac{\lambda}{c}\right)^2 \int_u^b e^{-(\lambda/c)(t-u)} \gamma_b(t)\,dt \right. \\ &\qquad \left. + \frac{\lambda}{c}\gamma_b(b) e^{-(\lambda/c)(b-u)}\right] \\ &\quad + \left(\frac{\lambda}{c}\right)^2 \gamma_b(u). \end{aligned} \tag{31}$$

Suppose the claim size distribution is exponential. Let $F(x) = 1 - e^{-\alpha x}$, $\alpha > 0$; then substituting (30) into (31), we have

$$\begin{aligned} m_b''(u) &= \frac{2\lambda}{c} m_b'(u) - \left(\frac{\lambda}{c}\right)^2 m_b(u) \\ &\quad + \left(\frac{\lambda}{c}\right)^2 \alpha e^{-\alpha u} \int_0^u m_b(t) e^{-\alpha t} dt + \left(\frac{\lambda}{c}\right)^2 e^{-\alpha u}. \end{aligned} \tag{32}$$

Differentiating (32) with respect to u, we have

$$\begin{aligned} m_b'''(u) &= \frac{2\lambda}{c} m_b''(u) - \left(\frac{\lambda}{c}\right)^2 m_b'(u) \\ &\quad - \left(\frac{\lambda}{c}\right)^2 \alpha^2 e^{-\alpha u} \int_0^u m_b(t) e^{-\alpha t} dt \\ &\quad + \left(\frac{\lambda}{c}\right)^2 \alpha m_b(u) - \left(\frac{\lambda}{c}\right)^2 \alpha e^{-\alpha u}. \end{aligned} \tag{33}$$

(32) $\times\, \alpha +$ (33) implies

$$m_b'''(u) + \left(\alpha - \frac{2\lambda}{c}\right) m_b''(u) + \frac{\lambda^2 - 2\alpha c\lambda}{c^2} m_b'(u) = 0. \tag{34}$$

This is a three-order differential equation with constant coefficients, so we can carry on the numerical solution. Suppose $\alpha = 10000$, $c = 200$, $\lambda = 0.0001$, $b = 20$; then by the Matlab, we obtain the curve of ruin probability (see Figure 1). As is known to all ruin must occur under the constant dividend barrier. From Figure 1, we know that ruin probability $\psi_b(u)$ is an increasing function of the initial surplus u (convex function) and the function value of 1 is its asymptote.

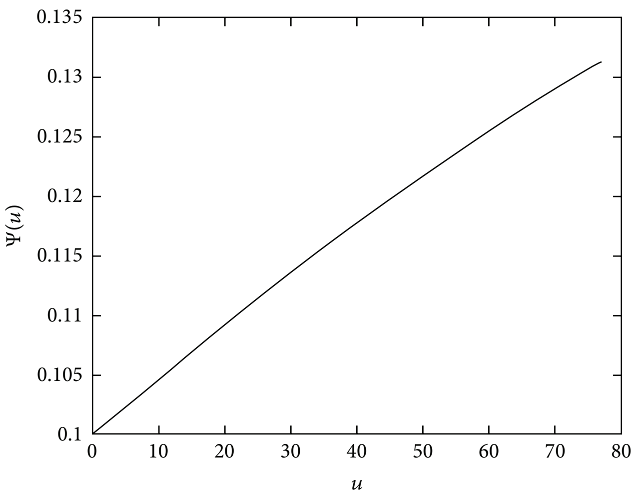


Figure 1: The curve of ruin probability.

Conflict of Interests

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

Acknowledgments

The authors would like to thank an anonymous referee for the valuable comments and suggestions that improved the paper presentation. This work was supported by National Natural Science Foundation of China (no. 11301303), Natural Science Foundation of Shandong Province (nos. ZR2012AQ013 and ZR2010GL013), Humanities and Social Sciences Project of the Ministry of Education of China (no. 14YJA630088, no. 13YJC630150, no. 10YJC630092, and no. 09YJC910004), and the Doctorate Scientific Research Foundation of Shandong Jiaotong University.

References

[1] J. Cai and D. C. M. Dickson, "Upper bounds for ultimate ruin probabilities in the Sparre Andersen model with interest," *Insurance: Mathematics & Economics*, vol. 32, no. 1, pp. 61–71, 2003.

[2] L. J. Sun and H. L. Yang, "On the joint distributions of surplus immediately before ruin and the deficit at ruin for Erlang(2) risk processes," *Insurance: Mathematics & Economics*, vol. 34, no. 1, pp. 121–125, 2004.

[3] H. U. Gerber and E. S. W. Shiu, "The time value of ruin in a Sparre Andersen model," *North American Actuarial Journal*, vol. 9, no. 2, pp. 49–69, 2005.

[4] B. Ko, "Discussion of the discounted joint distribution of the surplus prior to ruin and the deficit at ruin in a Sparre Andersen model," *North American Actuarial Journal*, vol. 11, no. 3, pp. 136–137, 2007.

[5] S. Li and J. Garrido, "On a class of renewal risk models with a constant dividend barrier," *Insurance: Mathematics & Economics*, vol. 35, no. 3, pp. 691–701, 2004.

[6] J. Ren, "The discounted joint distribution of the surplus prior to ruin and the deficit at ruin in a Sparre Andersen model," *North American Actuarial Journal*, vol. 11, no. 3, pp. 128–136, 2007.

[7] A. C. Y. Ng and H. Yang, "Lundberg-type bounds for the joint distribution of surplus immediately before and at ruin under the Sparre Andersen model," *North American Actuarial Journal*, vol. 9, no. 2, pp. 85–107, 2005.

[8] D. Landriault and G. E. Willmot, "On the Gerber- Shiu discounted penalty function in the Sparre Andersen model with an arbitrary inter-claim time distribution," *Insurance: Mathematics & Economics*, vol. 42, no. 2, pp. 600–608, 2008.

[9] H. Yang and Z. Zhang, "Gerber-Shiu discounted penalty function in a Sparre Andersen model with multi-layer dividend strategy," *Insurance: Mathematics & Economics*, vol. 42, no. 3, pp. 984–991, 2008.

[10] D. Landriault and K. P. Sendova, "A direct approach to a first-passage problem with applications in risk theory," *Stochastic Models*, vol. 27, no. 3, pp. 388–406, 2011.

[11] T. X. Shi and D. Landriault, "Distribution of the time to ruin in some Sparre Andersen risk models," *ASTIN Bulletin*, vol. 43, no. 1, pp. 39–59, 2013.

[12] C. Yang and K. P. Sendova, "The ruin time under the Sparre-Andersen dual model," *Insurance: Mathematics & Economics*, vol. 54, no. 1, pp. 28–40, 2014.

[13] B. De Finetti, "Su un'impostazione alternativa dell teoria colletiva del rischio," in *Proceedings of the Transactions of the 15th International Congress of Actuaries*, vol. 2, pp. 433–443, 1957.

[14] X. S. Lin, G. E. Willmot, and S. Drekic, "The classical risk model with a constant dividend barrier: analysis of the Gerber-Shiu discounted penalty function," *Insurance: Mathematics & Economics*, vol. 33, no. 3, pp. 551–566, 2003.

[15] D. C. M. Dickson and H. R. Waters, "Some optimal dividends problems," *Astin Bulletin*, vol. 34, no. 1, pp. 49–74, 2004.

[16] S. Li and Y. Lu, "Moments of the dividend payments and related problems in a Markov-modulated risk model," *North American Actuarial Journal*, vol. 11, no. 2, pp. 65–76, 2007.

[17] W. G. Yu, "Some results on absolute ruin in the perturbed insurance risk model with investment and debit interests," *Economic Modelling*, vol. 31, no. 1, pp. 625–634, 2013.

[18] W. G. Yu, "Randomized dividends in a discrete insurance risk model with stochastic premium income," *Mathematical Problems in Engineering*, vol. 2013, Article ID 579534, 9 pages, 2013.

[19] W. G. Yu, "On the expected discounted penalty function for a Markov regime-switching insurance risk model with stochastic premium income," *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 320146, 9 pages, 2013.

[20] D. J. Yao, H. L. Yang, and R. M. Wang, "Optimal risk and dividend control problem with fixed costs and salvage value: variance premium principle," *Economic Modelling*, vol. 37, pp. 53–64, 2014.

[21] J. Zhu, "Optimal dividend control for a generalized risk model with investment incomes and debit interest," *Scandinavian Actuarial Journal*, vol. 2013, no. 2, pp. 140–162, 2013.

[22] J. Y. Tan, P. T. Yuan, Y. J. Cheng, and Z. Q. Li, "An optimal dividend strategy in the discrete Sparre Andersen model with bounded dividend rates," *Journal of Computational and Applied Mathematics*, vol. 258, pp. 1–16, 2014.

[23] H. U. Gerber and E. S. W. Shiu, "On the time value of ruin," *North American Actuarial Journal*, vol. 2, no. 1, pp. 48–78, 1998.

[24] X. S. Lin and G. E. Willmot, "The moments of the time of ruin, the surplus before ruin, and the deficit at ruin," *Insurance: Mathematics & Economics*, vol. 27, no. 1, pp. 19–44, 2000.

[25] D. C. M. Dickson and C. Hipp, "On the time to ruin for Erlang(2) risk processes," *Insurance: Mathematics & Economics*, vol. 29, no. 3, pp. 333–344, 2001.

[26] X. H. Zhao and C. C. Yin, "Ruin problems for a Sparre Andersen risk model," *Journal of Qufu Normal University*, vol. 31, no. 2, pp. 9–12, 2005.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 718948, 18 pages
http://dx.doi.org/10.1155/2014/718948

Research Article

Mean-Field Backward Stochastic Evolution Equations in Hilbert Spaces and Optimal Control for BSPDEs

Ruimin Xu[1,2] and Tingting Wu[3]

[1] *School of Mathematics, Shandong University, Jinan 250100, China*
[2] *School of Mathematics, Shandong Polytechnic University, Jinan 250353, China*
[3] *School of Mathematical Sciences, Shandong Normal University, Jinan 250014, China*

Correspondence should be addressed to Ruimin Xu; ruiminx@126.com

Received 31 March 2014; Revised 23 May 2014; Accepted 18 June 2014; Published 13 July 2014

Academic Editor: Guangchen Wang

We obtain the existence and uniqueness result of the mild solutions to mean-field backward stochastic evolution equations (BSEEs) in Hilbert spaces under a weaker condition than the Lipschitz one. As an intermediate step, the existence and uniqueness result for the mild solutions of mean-field BSEEs under Lipschitz condition is also established. And then a maximum principle for optimal control problems governed by backward stochastic partial differential equations (BSPDEs) of mean-field type is presented. In this control system, the control domain need not to be convex and the coefficients, both in the state equation and in the cost functional, depend on the law of the BSPDE as well as the state and the control. Finally, a linear-quadratic optimal control problem is given to explain our theoretical results.

1. Introduction

Backward stochastic evolution equations (BSEEs) in their general nonlinear form were introduced by Hu and Peng [1] in 1991. By the stochastic Fubini theorem and an extended martingale representation theorem, Hu and Peng [1] obtained the existence and uniqueness result of a so-called "mild solution" under Lipschitz coefficients for semilinear BSEEs. Since then, BSEEs have been studied by a lot of authors and have found various applications, namely, in the theory of infinite dimensional optimal control and the controllability for stochastic partial differential equations (see e.g., [1–4] and the papers cited therein). To relax the Lipschitz condition of the coefficients, Mahmudov and Mckibben [2] studied BSEEs under a weaker condition than the Lipschitz one in Hilbert spaces. Their approach extended the method proposed by Mao [5], in which the author investigated BSDEs under a weaker condition which contains Lipschitz condition as a special case. Our present work also investigates backward stochastic evolution equations, but with one main difference to the setting chosen by the papers mentioned above: the coefficients of the BSEEs are allowed to depend on the law of the BSEEs.

Recently, mean-field approaches, which can be used to describe particle systems at the mesoscopic level, have attracted more and more researchers' attention because of their great importance in applications. For example, mean-field approach can be used in statistical mechanics and physics, quantum mechanics and quantum chemistry, economics, finance, game theory, and optimal control theory (refer to [6–8] and the references therein). Mean-field BSDEs were deduced by Buckdahn et al. [9] when they investigated a special mean-field problem in a purely stochastic approach. Buckdahn et al. [7] studied the well posedness of mean-field BSDEs and gave a probabilistic interpretation to semilinear McKean-Vlasov partial differential equations. To give a probabilistic representation of the solutions for a class of Mckean-Vlasov stochastic partial differential equations, Xu [10] investigated the well-posedness of mean-field backward doubly stochastic differential equations with locally monotone coefficients.

In this paper, we investigate a new type of backward stochastic evolution equations in Hilbert spaces which we call mean-field BSEEs. Mean-field implies that the coefficient of the BSEE depends on the law of the BSEE. Specifically, the BSEE we consider is defined as

$$\begin{aligned} dY(s) = &- AY(s)\,ds \\ &- \mathbb{E}'\left[f\left(s, Y'(s), Z'(s), Y(s), Z(s)\right)\right] ds \\ &+ Z(s)\,dW(s), \\ &\quad Y(T) = \xi, \quad s \in [0,T], \end{aligned} \tag{1}$$

in a Hilbert space H, where f denotes a given measurable mapping, T is a fixed positive real number, $W(s)$ is a cylindrical Wiener process, and A represents the generator of a strongly continuous semigroup e^{tA} in H with $t \geq 0$. Precise interpretation of $\mathbb{E}'[f(s, Y'(s), Z'(s), Y(s), Z(s))]$ is given in the following sections. Based on the contraction mapping, we firstly prove that (1) admits a unique mild solution if the function f is Lipschitz continuous. Secondly, under non-Lipschitz assumptions, we obtain the existence and uniqueness of the mild solution for mean-field BSEE by constructing a special Cauchy sequence. The Lipschitz condition is a special case of this non-Lipschitz condition (see Mao [5]). In addition, we investigate the well-posedness of mean-field stochastic evolution equations.

We also study optimal control problems of stochastic systems governed by mean-field BSPDEs in Hilbert spaces. Our objective is to formulate a stochastic maximum principle (SMP) for the optimal control problem with an initial state constraint. There is a vast literature on the theory of SMP. Among these papers, Andersson and Djehiche [8] studied the optimal control problem for mean-field stochastic system when the control domain is convex. They obtained the maximum principle by a convex variational method. By a spike variational technique, Buckdahn et al. [11] obtained a general maximum principle for a special mean-field stochastic differential equation (SDE) where the action space is not convex. Later, Li [12] investigated the maximum principle for more general SDEs of mean-field type with a convex control domain. Wang et al. [13] were concerned with a partially observed optimal control problem of mean-field type. By using Girsanov's theorem and convex variation, they derived the corresponding maximum principle and gave an illustrative example to demonstrate the application of the obtained SMP. Hafayed studied the mean-field SMP for singular stochastic control in [14] and mean-field SMP for FBSDEs with Poisson jump processes in [15].

For the case of stochastic control systems in infinite dimensions, on the assumption that the control domain is not necessarily convex while the diffusion coefficient does not contain the control variable, Hu and Peng [16] used spike variation approach and Ekeland's variational principle to establish the maximum principle for semilinear stochastic evolution control systems with a final state constraint. Mahmudov and Mckibben [2] obtained an SMP for stochastic control systems governed by BSEEs in Hilbert spaces. Recently, Fuhrman et al. [17] deduced the maximum principle for optimal control of stochastic PDEs when the control domain is not necessarily convex.

We establish necessary optimality conditions for the control problem in the form of a maximum principle on the assumption that the control domain is not necessarily convex. Due to the initial state constraint, we first need to apply Ekeland's variational principle to convert the given control problem into a free initial state optimal control problem. Then spike variation approach is used to deduce the SMP in the mean-field framework. In our control system, not only the state processes which are the unique mild solution of the given BSPDE, but also the cost functional are of mean-field type. In other words, they depend on the law of the BSPDE as well as the state and the control. For this new controlled system, the adjoint equation will turn out to be a mean-field stochastic evolution equation.

The plan of this paper is organized as follows. In Section 2, we introduce some notations which are needed in what follows. In Section 3, the well-posedness of mean-field BSEE (1) is studied; we first prove the existence and uniqueness of a mild solution under the Lipschitz condition and investigate the regular dependence of the solution (Y, Z) on (ξ, f). And then, under the assumption that the coefficient is non-Lipschitz continuous, a new result on the existence and uniqueness of the mild solution to (1) in Hilbert space is established, which generalizes the result for the Lipschitz case. Section 4 is devoted to the regularity of mean-field stochastic evolution equations. In Section 5, we derive the stochastic maximum principle for the BSPDE systems of mean-field type with an initial state constraint, and at the last part of Section 5, an LQ example is given to show the application of our maximum principle. An explicit optimal control is obtained in this example.

2. Preliminaries

The norm of an element x in a Banach space F is denoted by $|x|_F$ or simply $|x|$, if no confusion is possible. Γ, H, and K are three real and separable Hilbert spaces. Scalar product is denoted by $\langle \cdot, \cdot \rangle$, with a subscript to specify the space, if necessary. $\mathscr{L}(\Gamma, K)$ is the space of Hilbert-Schmidt operators from Γ to K, endowed with the Hilbert-Schmidt norm.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space. A cylindrical Wiener process $\{W(t), t \geq 0\}$ in a Hilbert space Γ is a family of linear mappings $\Gamma \to L^2(\Omega, \mathscr{F}, \mathbb{P})$ such that

(i) for every $u \in \Gamma$, $\{W(t)u, t \geq 0\}$ is a real (continuous) Wiener process;

(ii) for every u, $v \in \Gamma$ and t, $s \geq 0$, $\mathbb{E}(W(t)u \cdot W(s)v) = (t \wedge s)\langle u, v\rangle_\Gamma$.

By $\mathscr{F}_t$, $t \in [0, T]$, we denote the natural filtration of W, augmented with the family $\mathscr{N}$ of $\mathbb{P}$-null sets of $\mathscr{F}_T$:

$$\mathscr{F}_t = \sigma(W(s) : s \in [0, t]) \vee \mathscr{N}. \tag{2}$$

The filtration $(\mathscr{F}_t)_{t\geq 0}$ satisfies the usual conditions. All the concepts of measurability for stochastic processes (e.g., adapted, etc.) refer to this filtration.

Next we define several classes of stochastic processes with values in a Hilbert space H.

(I) $\mathscr{H}^2_{\mathbb{F}}([0,T];H)$ denotes the set of (classes of $d\mathbb{P}\times dt$ a.e. equal) measurable random processes $\{\psi_t; t\in[0,T]\}$ which satisfy

(i) $\mathbb{E}\int_0^T |\psi_t|^2 dt < +\infty$,

(ii) ψ_t is $\mathscr{F}_t$ measurable, for a.e. $0\le t\le T$.

Evidently, $\mathscr{H}^2_{\mathbb{F}}(0,T;H)$ is a Banach space endowed with the canonical norm

$$\|\psi\| = \left\{\mathbb{E}\int_0^T |\psi_s|^2 ds\right\}^{1/2}. \tag{3}$$

(II) $\mathcal{S}^2_{\mathbb{F}}([0,T];H)$ denotes the set of continuous random processes $\{\psi_t; t\in[0,T]\}$ which satisfy

(i) $\mathbb{E}(\sup_{0\le t\le T}|\psi_t|^2) < +\infty$,

(ii) ψ_t is $\mathscr{F}_t$ measurable, for a.e. $0\le t\le T$.

(III) $L^0(\Omega,\mathscr{F},\mathbb{P};H)$ denotes the space of all H valued $\mathscr{F}$-measurable random variables.

(IV) For $1\le p<\infty$, $L^p(\Omega,\mathscr{F},\mathbb{P};H)$ is the space of all $\mathscr{F}$-measurable random variables such that $\mathbb{E}[|\xi|^p]<\infty$.

(V) For any $\beta\in\mathbb{R}$, introduce the norm

$$\|(y,z)\|^2_{\beta,t} = \mathbb{E}\int_t^T e^{2\beta s}\left(|y(s)|^2 + |z(s)|^2\right) ds \tag{4}$$

on the Banach space

$$\mathscr{K}_\beta[t,T] = \mathcal{S}^2_{\mathbb{F}}([t,T];H)\times\mathscr{H}^2_{\mathbb{F}}([t,T];\mathscr{L}(\Gamma,H)). \tag{5}$$

For $0<T<\infty$, all the norms $\|\cdot\|_{\beta,t}$ with different $\beta\in\mathbb{R}$ are equivalent. $\mathscr{K}[0,T]=\mathscr{K}_0[0,T]$ is the Banach space endowed with the norm

$$\|(y,z)\|^2 = \mathbb{E}\int_0^T \left(|y(s)|^2 + |z(s)|^2\right) ds. \tag{6}$$

The following result on BSEEs (see Lemma 2 in Mahmudov and McKibben [2]) will play a key role in proving the well-posedness of mean-field BSEEs.

Lemma 1. *Let H be a Hilbert space, and let $A: D(A)\subset H\to H$ be a linear operator which generates a C_0-semigroup $\{S(t), 0\le t\le T\}$ on H. For any $(\xi,f)\in L^2(\Omega,\mathscr{F}_T,\mathbb{P};H)\times\mathscr{H}^2_{\mathbb{F}}([0,T],H)$ the following equation*

$$\begin{aligned} Y(t) &= S(T-t)\xi + \int_t^T S(s-t) f(s)\,ds \\ &\quad + \int_t^T S(s-t) Z(s)\,dW(s), \quad P\text{-a.s.} \end{aligned} \tag{7}$$

has a unique solution in $\mathscr{K}_\beta[0,T]$; moreover,

$$\begin{aligned} &\mathbb{E}\sup_{t\le s\le T} e^{2\beta s}|Y(s)|^2 + \mathbb{E}\int_t^T e^{2\beta s}|Z(s)|^2 ds \\ &\quad\le 24M_S^2\left(e^{2\beta T}\mathbb{E}|\xi|^2 + \frac{1}{2\beta}\int_t^T e^{2\beta r}\mathbb{E}|f(r)|^2 dr\right), \end{aligned} \tag{8}$$

where $M_S=\sup\{\|S(t)\|_{\mathfrak{B}(H)}, 0\le t\le T\}$ and $\mathfrak{B}(H)$ is the space of bounded, linear operators on H.

3. Mean-Field Backward Stochastic Evolution Equations

In this section, we study the existence and uniqueness result of mild solutions to mean-field BSEEs in a Hilbert space H. To this end, we firstly recall some notations introduced by Buckdahn et al. [7].

Let $(\overline{\Omega},\overline{\mathscr{F}},\overline{\mathbb{P}}) = (\Omega\times\Omega, \mathscr{F}\otimes\mathscr{F}, \mathbb{P}\otimes\mathbb{P})$ be the (noncompleted) product of $(\Omega,\mathscr{F},\mathbb{P})$ with itself and we define $\overline{\mathbb{F}}=\{\overline{\mathscr{F}}_t=\mathscr{F}\otimes\mathscr{F}_t, 0\le t\le T\}$ on this product space. A random variable $\xi\in L^0(\Omega,\mathscr{F},\mathbb{P};H)$ originally defined on Ω is extended canonically to $\overline{\Omega}$: $\xi'(\omega',\omega)=\xi(\omega')$, $(\omega',\omega)\in\overline{\Omega}=\Omega\times\Omega$. For any $\eta\in L^1(\overline{\Omega},\overline{\mathscr{F}},\overline{\mathbb{P}})$, the variable $\eta(\cdot,\omega):\Omega\to K$ belongs to $L^1(\Omega,\mathscr{F},\mathbb{P})$, $\mathbb{P}(d\omega)$ a.s., whose expectation is denoted by

$$\mathbb{E}'[\eta(\cdot,\omega)] = \int_\Omega \eta(\omega',\omega)\,\mathbb{P}(d\omega'). \tag{9}$$

Note that $\mathbb{E}'[\eta]=\mathbb{E}'[\eta(\cdot,\omega)]\in L^1(\Omega,\mathscr{F},\mathbb{P})$ and

$$\overline{\mathbb{E}}[\eta]\left(=\int_{\overline{\Omega}}\eta\,d\overline{\mathbb{P}} = \int_\Omega \mathbb{E}'[\eta(\cdot,\omega)]\,\mathbb{P}(d\omega)\right) = \mathbb{E}\left[\mathbb{E}'[\eta]\right]. \tag{10}$$

The mean-field BSEE we consider has the following form: for any given measurable mapping $f:[0,T]\times H\times\mathscr{L}(\Gamma,H)\times H\times\mathscr{L}(\Gamma,H)\to H$ and $\xi\in L^2(\Omega,\mathscr{F}_T,\mathbb{P};H)$,

$$\begin{aligned} dY(s) &= -AY(s)\,ds \\ &\quad - \mathbb{E}'\left[f\left(s, Y'(s), Z'(s), Y(s), Z(s)\right)\right] ds \\ &\quad + Z(s)\,dW(s), \\ Y(T) &= \xi, \quad s\in[0,T], \end{aligned} \tag{11}$$

where $A: D(A)\subset H\to H$ is the generator of a strongly continuous semigroup e^{tA}, $t\ge 0$, in the Hilbert space H, with the notation $M_A \triangleq \sup_{t\in[0,T]}|e^{tA}|$.

Definition 2. We say that a pair of adapted processes (Y,Z) is a mild solution of mean-field BSEE (11) if $(Y,Z) \in \mathcal{S}^2_{\mathbb{F}}([0,T];H)\times\mathcal{H}^2_{\mathbb{F}}([0,T];\mathcal{L}(\Gamma,H))$ and for all $t\in[0,T]$

$$Y(t) = e^{A(T-t)}\xi + \int_t^T e^{A(s-t)}\mathbb{E}'\left[f\left(s,Y'(s),Z'(s),Y(s),Z(s)\right)\right]ds - \int_t^T e^{A(s-t)}Z(s)\,dW(s). \tag{12}$$

Remark 3. We emphasize that the coefficient of (11) can be interpreted as

$$\begin{aligned}\mathbb{E}'&\left[f\left(s,Y'(s),Z'(s),Y(s),Z(s)\right)\right](\omega)\\ &= \mathbb{E}'\left[f\left(s,Y'(s),Z'(s),Y(\omega,s),Z(\omega,s)\right)\right]\\ &= \int_\Omega f\left(\omega',\omega,s,Y\left(\omega',s\right),Z\left(\omega',s\right),Y(\omega,s),Z(\omega,s)\right)\\ &\quad\times\mathbb{P}\left(d\omega'\right).\end{aligned} \tag{13}$$

3.1. Lipschitz Case. Now we study the existence and uniqueness of mild solutions to mean-field BSEE (11) under Lipschitz conditions. For $f:[0,T]\times H\times\mathcal{L}(\Gamma,H)\times H\times\mathcal{L}(\Gamma,H)\to H$, assume the following.

(A1) There exists an $L>0$ such that

$$\begin{aligned}&\left|f\left(t,y_1',z_1',y_1,z_1\right)-f\left(t,y_2',z_2',y_2,z_2\right)\right|^2\\ &\quad\le L\left(\left|y_1'-y_2'\right|^2+\left|z_1'-z_2'\right|^2\right.\\ &\qquad\left.+\left|y_1-y_2\right|^2+\left|z_1-z_2\right|^2\right),\end{aligned} \tag{14}$$

for all $t\in[0,T]$, $y_i', y_i\in H$, $z_i', z_i\in\mathcal{L}(\Gamma,H)$, $(i=1,2)$.

(A2) $f(\cdot,0,0,0,0)\in\mathcal{H}^2_{\mathbb{F}}([0,T];H)$.

We have the following theorem.

Theorem 4. *For any random variable $\xi\in L^2(\Omega,\mathcal{F}_T,\mathbb{P};H)$, under (A1) and (A2), mean-field BSEE (11) admits a unique mild solution $(Y,Z)\in\mathcal{S}^2_{\mathbb{F}}([0,T];H)\times\mathcal{H}^2_{\mathbb{F}}([0,T];\mathcal{L}(\Gamma,H))$.*

Proof. Consider the following.

Step 1. For any $(y,z)\in\mathcal{S}^2_{\mathbb{F}}([0,T];H)\times\mathcal{H}^2_{\mathbb{F}}([0,T];\mathcal{L}(\Gamma,H))$, BSEE

$$Y(t) = e^{A(T-t)}\xi + \int_t^T \mathbb{E}'\left[e^{A(s-t)}f\left(s,y'(s),z'(s),Y(s),Z(s)\right)\right]ds - \int_t^T e^{A(s-t)}Z(s)\,dW(s),\quad 0\le t\le T \tag{15}$$

has a unique solution. In order to get this conclusion, we define

$$f^{(y,z)}(s,\mu,\nu) := \mathbb{E}'\left[f\left(s,y'(s),z'(s),\mu,\nu\right)\right]. \tag{16}$$

Then (15) can be rewritten as

$$Y(t) = e^{A(T-t)}\xi + \int_t^T e^{A(s-t)}f^{(y,z)}(Y(s),Z(s))\,ds - \int_t^T e^{A(s-t)}Z(s)\,dW(s). \tag{17}$$

Due to (A1), for all $(\mu_1,\nu_1),(\mu_2,\nu_2)\in H\times\mathcal{L}(\Gamma,H)$, f satisfies

$$\begin{aligned}&\left|f^{(y,z)}(\mu_1,\nu_1)-f^{(y,z)}(\mu_2,\nu_2)\right|^2\\ &\quad\le L\left(\left|\mu_1-\mu_2\right|^2+\left|\nu_1-\nu_2\right|^2\right).\end{aligned} \tag{18}$$

According to Theorem 3.1 in [1], BSEE (15) has a unique solution.

Step 2. From Step 1, we can define a mapping $\Phi:(Y(\cdot),Z(\cdot))=\Phi[(y'(\cdot),z'(\cdot))]:\mathcal{K}[0,T]\to\mathcal{K}[0,T]$ through

$$Y(t) = e^{A(T-t)}\xi + \int_t^T \mathbb{E}'\left[e^{A(s-t)}f\left(s,y'(s),z'(s),Y(s),Z(s)\right)\right]ds - \int_t^T e^{A(s-t)}Z(s)\,dW(s),\quad 0\le t\le T. \tag{19}$$

For any $(y^i,z^i)\in\mathcal{K}[0,T]$, we set $(Y^i,Z^i)=\Phi[(y^i,z^i)]$, $i=1,2$, $(\overline{y},\overline{z})=(y^1-y^2,z^1-z^2)$, and $(\overline{Y},\overline{Z})=(Y^1-Y^2,Z^1-Z^2)$. Then, from Lemma 1, we have

$$\begin{aligned}&\mathbb{E}\sup_{0\le s\le T}e^{2\beta s}\left|\overline{Y}(s)\right|^2+\mathbb{E}\int_0^T e^{2\beta s}\left|\overline{Z}(s)\right|^2ds\\ &\le\frac{12M_A^2}{\beta}\mathbb{E}\end{aligned}$$

$$\times \left[\int_0^T e^{2\beta s} \left|\mathbb{E}'\left[f\left(s, \left(y^1(s)\right)', \left(z^1(s)\right)', Y^1(s), Z^1(s)\right)\right.\right.\right.$$
$$- f\left(s, \left(y^2(s)\right)',\right.$$
$$\left.\left.\left.\left(z^2(s)\right)', Y^2(s), Z^2(s)\right)\right]\right|^2 ds\right]$$
$$\leq \frac{12M_A^2 L^2}{\beta}\mathbb{E}$$
$$\times \left[\int_0^T e^{2\beta s}\left(\mathbb{E}\left[|\overline{y}(s)|^2 + |\overline{z}(s)|^2\right]\right.\right.$$
$$\left.\left. + \left|\overline{Y}(s)\right|^2 + \left|\overline{Z}(s)\right|^2\right) ds\right]$$
$$= \frac{12M_A^2 L^2}{\beta}\mathbb{E}$$
$$\times \left[\int_0^T e^{2\beta s}\left(|\overline{y}(s)|^2 + |\overline{z}(s)|^2\right.\right.$$
$$\left.\left. + \left|\overline{Y}(s)\right|^2 + \left|\overline{Z}(s)\right|^2\right) ds\right]. \tag{20}$$

If we set $\beta = 36M_A^2 L^2 \max\{T, 1\}$, then

$$\mathbb{E}\int_0^T e^{2\beta s}\left(\left|\overline{Y}(s)\right|^2 + \left|\overline{Z}(s)\right|^2\right) ds$$
$$\leq T \cdot \mathbb{E}\sup_{0\leq s\leq T} e^{2\beta s}\left|\overline{Y}(s)\right|^2 + \mathbb{E}\int_0^T e^{2\beta s}\left|\overline{Z}(s)\right|^2 ds$$
$$\leq \frac{12M_A^2 L^2 \max\{T,1\}}{\beta}\mathbb{E}$$
$$\times \left[\int_0^T e^{2\beta s}\left(|\overline{y}(s)|^2 + |\overline{z}(s)|^2\right.\right. \tag{21}$$
$$\left.\left. + \left|\overline{Y}(s)\right|^2 + \left|\overline{Z}(s)\right|^2\right) ds\right]$$
$$= \frac{1}{3}\mathbb{E}\left[\int_0^T e^{2\beta s}\left(|\overline{y}(s)|^2 + |\overline{z}(s)|^2\right.\right.$$
$$\left.\left. + \left|\overline{Y}(s)\right|^2 + \left|\overline{Z}(s)\right|^2\right) ds\right].$$

That is,

$$\mathbb{E}\int_0^T e^{2\beta s}\left(\left|\overline{Y}(s)\right|^2 + \left|\overline{Z}(s)\right|^2\right) ds$$
$$\leq \frac{1}{2}\mathbb{E}\int_0^T e^{2\beta s}\left(|\overline{y}(s)|^2 + |\overline{z}(s)|^2\right) ds. \tag{22}$$

The estimate (22) shows that Φ is a contraction on the space $\mathscr{K}_\beta[0,T]$ with the norm

$$\|(Y,Z)\|_\beta^2 = \mathbb{E}\int_0^T e^{2\beta s}\left(|Y(s)|^2 + |Z(s)|^2\right) ds. \tag{23}$$

With the contraction mapping theorem, there admits a unique fixed point $(Y,Z) \in \mathscr{K}_\beta[0,T]$ such that $\Phi(Y,Z) = (Y,Z)$. On the other hand, from Step 1, we know that if $\Phi(Y,Z) = (Y,Z)$, then $(Y,Z) \in \mathcal{S}_{\mathbb{F}}^2([0,T];H) \times \mathscr{H}_{\mathbb{F}}^2([0,T];\mathscr{L}(\Gamma,H))$, which is the unique mild solution of (11). □

Arguing as the previous proof, we arrive at the following assertion in a straightforward way.

Corollary 5. *Suppose that, for all α in a metric space F, f_α is a given function satisfying (A1) and (A2) with L independent on α. Also suppose that*

$$\mathbb{E}'\left[f_\alpha\left(s, Y'(s), Z'(s), Y(s), Z(s)\right)\right]$$
$$\longrightarrow \mathbb{E}'\left[f_{\alpha_0}\left(s, Y'(s), Z'(s), Y(s), Z(s)\right)\right] \tag{24}$$

in $L^2([0,T];H)$ as $\alpha \to \alpha_0$ for all $(Y,Z) \in \mathcal{S}_{\mathbb{F}}^2([0,T];H) \times \mathscr{H}_{\mathbb{F}}^2([0,T];\mathscr{L}(\Gamma,H))$.

If we denote by $(Y(\xi,\alpha), Z(\xi,\alpha))$ the mild solution of (11) *corresponding to the functions f_α and to the final data $\xi \in L^2(\Omega,\mathscr{F}_T,\mathbb{P};H)$, then the map $(\alpha,\xi) \to (Y(\xi,\alpha), Z(\xi,\alpha))$ is continuous from $F \times L^2(\Omega,\mathscr{F}_T,\mathbb{P};H)$ to $\mathcal{S}_{\mathbb{F}}^2([0,T];H) \times \mathscr{H}_{\mathbb{F}}^2([0,T];\mathscr{L}(\Gamma,H))$.*

3.2. Non-Lipschitz Case. This subsection is devoted to finding some weaker conditions than the Lipschitz one under which the mean-field BSEE has a unique solution. To state our main result in this section, we suppose the following.

(A3) For all $t \in [0,T]$, $y_i', y_i \in H$, $z_i', z_i \in \mathscr{L}(\Gamma,H)$, $(i = 1,2)$, there exists an $l > 0$, such that

$$\left|f\left(t, y_1', z_1', y_1, z_1\right) - f\left(t, y_2', z_2', y_2, z_2\right)\right|^2$$
$$\leq \theta\left(\left|y_1' - y_2'\right|^2\right) + \theta\left(\left|y_1 - y_2\right|^2\right) \tag{25}$$
$$+ l\left(\left|z_1' - z_2'\right|^2 + \left|z_1 - z_2\right|^2\right),$$

where $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ is a concave increasing function such that $\theta(0) = 0$, $\theta(u) > 0$ for $u > 0$ and $\int_{0^+}(du/\theta(u)) = \infty$.

In Mao [5], the author gave three examples of the function $\theta(\cdot)$ to show the generality of condition (A3). From these examples, we can see that Lipschitz condition (A1) is a special case of the given condition (A3).

Since θ is concave and $\theta(0) = 0$, there exists a pair of positive constants a and b such that

$$\theta(u) \leq a + bu \tag{26}$$

for all $u \geq 0$. Therefore, under assumptions (A2) and (A3), $f(\cdot, y'(\cdot), z'(\cdot), y(\cdot), z(\cdot)) \in \mathscr{H}_{\mathbb{F}}^2([0,T];H)$ whenever

$y'(\cdot), y(\cdot) \in \mathcal{S}^2_{\mathbb{F}}([0,T]; H)$ and $z'(\cdot), z(\cdot) \in \mathcal{H}^2_{\mathbb{F}}([0,T]; \mathcal{L}(\Gamma, H))$.

By Picard-type iteration, we now construct an approximate sequence, using which we obtain the desired result. Let $Y_0(t) \equiv 0$, and, for $n \in \mathbb{N}$, let $\{Y_n, Z_n\}$ be a sequence in $\mathcal{S}^2_{\mathbb{F}}([0,T]; H) \times \mathcal{H}^2_{\mathbb{F}}([0,T]; \mathcal{L}(\Gamma, H))$ defined recursively by

$$\begin{aligned} Y_n(t) &= e^{A(T-t)}\xi \\ &\quad + \int_t^T \mathbb{E}'\left[e^{A(s-t)} \right.\\ &\qquad \left.\times f\left(s, Y'_{n-1}(s), Z'_n(s), Y_{n-1}(s), Z_n(s)\right)\right] ds \\ &\quad - \int_t^T e^{A(s-t)} Z_n(s)\, dW(s), \end{aligned} \tag{27}$$

on $0 \le t \le T$. From Theorem 4, (27) has a unique mild solution $(Y_n(t), Z_n(t))$.

In order to give the main result, we need to prepare the following lemmas about the properties of $(Y_n(t), Z_n(t))$, $t \in [0,T]$.

Lemma 6. *Under hypotheses (A2) and (A3), there exist positive constants C_1 and C_2 such that*

$$\begin{aligned} &\text{(i) } \mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_n(s)\right|^2\right) \le 2C_1 \exp(T-t), \\ &\text{(ii) } \mathbb{E}\int_t^T e^{2\beta s}\left|Z_n(s)\right|^2 ds \le 2C_1 \exp(T-t), \\ &\text{(iii) } \mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_{n+1}(s) - Y_n(s)\right|^2\right) \\ &\qquad \le C_2 \int_t^T \theta\left(\mathbb{E}\sup_{s\le r\le T} e^{2\beta r}\left|Y_n(r) - Y_{n-1}(r)\right|^2\right) ds, \end{aligned} \tag{28}$$

for all $t \in [0,T]$ and $n \ge 1$.

Proof. Using the hypotheses (A2) and (A3) with $\theta(u) \le a + bu$ yields

$$\begin{aligned} &\left|f\left(s, Y'_{n-1}(s), Z'_n(s), Y_{n-1}(s), Z_n(s)\right)\right|^2 \\ &\quad \le 2\theta\left(\left|Y'_{n-1}(s)\right|^2\right) + 2\theta\left(\left|Y_{n-1}(s)\right|^2\right) \\ &\qquad + 2l\left(\left|Z'_n(s)\right|^2 + \left|Z_n(s)\right|^2\right) + 2\left|f(s,0,0,0,0)\right|^2 \\ &\quad \le 4a + 2b\left|Y'_{n-1}(s)\right|^2 + 2b\left|Y_{n-1}(s)\right|^2 \\ &\qquad + 2l\left(\left|Z'_n(s)\right|^2 + \left|Z_n(s)\right|^2\right) + 2\left|f(s,0,0,0,0)\right|^2. \end{aligned} \tag{29}$$

Then, it follows from Lemma 1 that

$$\begin{aligned} &\mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_n(s)\right|^2\right) + \mathbb{E}\int_t^T e^{2\beta s}\left|Z_n(s)\right|^2 ds \\ &\quad \le 24M_A^2 e^{2\beta T}\mathbb{E}|\xi|^2 + \frac{12M_A^2}{\beta}\mathbb{E} \\ &\qquad \times \int_t^T e^{2\beta s}\left|\mathbb{E}'\left[f\left(s, Y'_{n-1}(s),\right.\right.\right. \\ &\qquad\qquad \left.\left.\left. Z'_n(s), Y_{n-1}(s), Z_n(s)\right)\right]\right|^2 ds \\ &\quad \le 24M_A^2 e^{2\beta T}\mathbb{E}|\xi|^2 + \frac{12M_A^2}{\beta}\mathbb{E} \\ &\qquad \times \int_t^T \mathbb{E}'\left[e^{2\beta s}\left|f\left(s, Y'_{n-1}(s),\right.\right.\right. \\ &\qquad\qquad \left.\left.\left. Z'_n(s), Y_{n-1}(s), Z_n(s)\right)\right|^2\right] ds \\ &\quad \le C_1 + \frac{24M_A^2}{\beta}\mathbb{E} \\ &\qquad \times \int_t^T \mathbb{E}'\left[e^{2\beta s}\left[b\left|Y'_{n-1}(s)\right|^2 + b\left|Y_{n-1}(s)\right|^2\right.\right. \\ &\qquad\qquad \left.\left. + l\left|Z'_n(s)\right|^2 + l\left|Z_n(s)\right|^2\right]\right] ds \\ &\quad = C_1 + \frac{48M_A^2}{\beta}\mathbb{E} \\ &\qquad \times \int_t^T e^{2\beta s}\left[b\left|Y_{n-1}(s)\right|^2 + l\left|Z_n(s)\right|^2\right] ds, \end{aligned} \tag{30}$$

where

$$\begin{aligned} C_1 &= 24M_A^2 e^{2\beta T}\mathbb{E}|\xi|^2 \\ &\quad + \frac{24M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\left[2a + \left|f(s,0,0,0,0)\right|^2\right] ds + 1. \end{aligned} \tag{31}$$

If we set $\beta = 96M_A^2 \max\{b, l\}$, we can obtain

$$\begin{aligned} &\sup_{n\in\mathbb{N}}\mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_n(s)\right|^2\right) + \frac{1}{2}\int_t^T \sup_{n\in\mathbb{N}}\mathbb{E}\left[e^{2\beta s}\left|Z_n(s)\right|^2\right] ds \\ &\quad \le C_1 + \frac{1}{2}\int_t^T \sup_{n\in\mathbb{N}}\mathbb{E}\left[e^{2\beta s}\left|Y_{n-1}(s)\right|^2\right] ds \\ &\quad \le C_1 + \frac{1}{2}\int_t^T \sup_{n\in\mathbb{N}}\mathbb{E}\left[\sup_{s\le r\le T} e^{2\beta r}\left|Y_{n-1}(r)\right|^2\right] ds. \end{aligned} \tag{32}$$

An application of the Gronwall inequality now implies

$$\begin{aligned} \sup_{n\in\mathbb{N}}\mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_n(s)\right|^2\right) &\le 2C_1 \exp\left(\frac{T-t}{2}\right) \\ &\le 2C_1 \exp(T-t). \end{aligned} \tag{33}$$

Point (i) of Lemma 6 is now proved.

From formula (32), we know that

$$\begin{aligned}
&\int_t^T \sup_{n\in\mathbb{N}} \mathbb{E}\left[e^{2\beta s}\left|Z_n(s)\right|^2\right] ds \\
&\quad \le 2C_1 + \int_t^T \sup_{n\in\mathbb{N}} \mathbb{E}\left[\sup_{s\le r\le T} e^{2\beta r}\left|Y_{n-1}(r)\right|^2\right] ds \\
&\quad \le 2C_1 + 2C_1 \int_t^T \exp(T-s)\, ds \\
&\quad = 2C_1 \exp(T-t).
\end{aligned} \tag{34}$$

This proves point (ii) of the Lemma.

To prove point (iii), we note that

$$\begin{aligned}
&\mathbb{E}'\left[\left|f\left(s, Y_n'(s), Z_{n+1}'(s), Y_n(s), Z_{n+1}(s)\right)\right.\right. \\
&\qquad \left.\left. - f\left(s, Y_{n-1}'(s), Z_n'(s), Y_{n-1}(s), Z_n(s)\right)\right|^2\right] \\
&\quad \le \mathbb{E}'\left[\theta\left(\left|Y_n'(s) - Y_{n-1}'(s)\right|^2\right) + l\left|Z_{n+1}'(s) - Z_n'(s)\right|^2\right] \\
&\qquad + \theta\left(\left|Y_n(s) - Y_{n-1}(s)\right|^2\right) + l\left|Z_{n+1}(s) - Z_n(s)\right|^2 \\
&\quad = \mathbb{E}\left[\theta\left(\left|Y_n(s) - Y_{n-1}(s)\right|^2\right) + l\left|Z_{n+1}(s) - Z_n(s)\right|^2\right] \\
&\qquad + \theta\left(\left|Y_n(s) - Y_{n-1}(s)\right|^2\right) + l\left|Z_{n+1}(s) - Z_n(s)\right|^2.
\end{aligned} \tag{35}$$

By Lemma 1 we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_{n+1}(s) - Y_n(s)\right|^2\right) \\
&\qquad + \mathbb{E}\int_t^T e^{2\beta s}\left|Z_{n+1}(s) - Z_n(s)\right|^2 ds \\
&\quad \le \frac{12M_A^2}{\beta}\mathbb{E} \\
&\qquad \times \int_t^T e^{2\beta s}\left|\mathbb{E}'\left[f\left(s, Y_n'(s), Z_{n+1}'(s), Y_n(s), Z_{n+1}(s)\right)\right.\right. \\
&\qquad\qquad - f\left(s, Y_{n-1}'(s), Z_n'(s),\right. \\
&\qquad\qquad\qquad \left.\left.\left. Y_{n-1}(s), Z_n(s)\right)\right]\right|^2 ds \\
&\quad \le \frac{12M_A^2}{\beta}\mathbb{E} \\
&\qquad \times \int_t^T e^{2\beta s}\mathbb{E}'\left[\left|f\left(s, Y_n'(s), Z_{n+1}'(s), Y_n(s), Z_{n+1}(s)\right)\right.\right. \\
&\qquad\qquad - f\left(s, Y_{n-1}'(s), Z_n'(s),\right. \\
&\qquad\qquad\qquad \left.\left.\left. Y_{n-1}(s), Z_n(s)\right)\right|^2\right] ds \\
&\quad \le \frac{24M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\left[\theta\left(\left|Y_n(s) - Y_{n-1}(s)\right|^2\right)\right. \\
&\qquad\qquad \left. + l\left|Z_{n+1}(s) - Z_n(s)\right|^2\right] ds.
\end{aligned} \tag{36}$$

We can choose $\beta > 0$ sufficiently large such that

$$\left(1 - \frac{24M_A^2 l}{\beta}\right)\mathbb{E}\int_t^T e^{2\beta s}\left\|Z_{n+1}(s) - Z_n(s)\right\|^2 ds \ge 0. \tag{37}$$

Then

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_{n+1}(s) - Y_n(s)\right|^2\right) \\
&\quad \le \frac{24M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\theta\left(\left|Y_n(s) - Y_{n-1}(s)\right|^2\right) ds \\
&\quad \le C_2\mathbb{E}\int_t^T \theta\left(\mathbb{E}\sup_{s\le r\le T} e^{2\beta r}\left|Y_n(r) - Y_{n-1}(r)\right|^2\right) ds,
\end{aligned} \tag{38}$$

where we set $C_2 = (24M_A^2/\beta)e^{2\beta T}$. □

We divide the interval $[0, T]$ into subintervals $0 = \tau_0 < \tau_1 < \cdots < \tau_m = T$ by setting $\tau_k = k\delta$, $k = 1, 2, 3, \ldots, m$ with $\delta = T/m$.

Lemma 7. *For all $t \in [\tau_{k-1}, \tau_k]$, define*

$$\begin{aligned}
C_3 &= C_2\theta\left(2C_1\exp(T)\right), \\
\varphi_{k,1}(t) &= C_3\left(\tau_k - t\right), \\
\varphi_{k,n+1}(t) &= C_2\int_t^{\tau_k}\theta\left(\varphi_{k,n}(s)\right) ds, \quad n \ge 1.
\end{aligned} \tag{39}$$

Then, for all $n \ge 1$, the following inequality holds for a suitable $\delta > 0$:

$$0 \le \varphi_{k,n}(t) \le \varphi_{k,n-1}(t) \le \cdots \le \varphi_{k,1}(t). \tag{40}$$

Proof. Firstly, it needs to be verified that for all $t \in [\tau_{k-1}, \tau_k]$ the following inequality

$$\begin{aligned}
\varphi_{k,2}(t) &= C_2\int_t^{\tau_k}\theta\left(\varphi_{k,1}(s)\right) ds = C_2\int_t^{\tau_k}\theta\left(C_3\left(\tau_k - s\right)\right) ds \\
&\le C_3\left(\tau_k - t\right) = \varphi_{k,1}(t)
\end{aligned} \tag{41}$$

holds provided $\delta > 0$ is chosen sufficiently small.

Actually, this inequality holds provided that

$$C_2\theta\left(C_3\left(\tau_k - t\right)\right) \le C_3 = C_2\theta\left(2C_1\exp(T)\right) \tag{42}$$

or

$$C_3\left(\tau_k - t\right) = C_2\theta\left(2C_1 \exp\left(T\right)\right)\left(\tau_k - t\right) \leq 2C_1 \exp\left(T\right). \tag{43}$$

Since $C_1 > 1$, from $\theta(u) \leq a + bu$, the above inequality holds if

$$C_2\left(a+b\right)\left(\tau_k - t\right) \leq 1. \tag{44}$$

Thus, (41) holds for any $t \in [\tau_{k-1}, \tau_k]$, $k = 1, 2, \ldots, m$ if $\tau_k - \tau_{k-1} \leq 1/C_2(a+b)$. Therefore, we can choose a sufficiently large $m \in \mathbb{N}$ such that $\delta = T/m \leq 1/C_2(a+b)$. Clearly, such a δ only depends on a, b, l, T, and M_A.

Now, assume that (40) holds for some $n \geq 2$. Then, we have

$$\begin{aligned}\varphi_{k,n+1}(t) &= C_2\int_t^{\tau_k}\theta\left(\varphi_{k,n}(s)\right)ds \leq C_2\int_t^{\tau_k}\theta\left(\varphi_{k,n-1}(s)\right)ds\\ &= \varphi_{k,n}(t), \quad \forall t \in \left[\tau_{k-1}, \tau_k\right].\end{aligned} \tag{45}$$

This completes the proof. □

Now, we can give the main result of this section.

Theorem 8. *Assume that (A2) and (A3) hold. Then, there exists a unique mild solution* (Y, Z) *to* (11).

Proof. Consider the following.

Uniqueness. To show the uniqueness, let both (Y, Z) and $(\widetilde{Y}, \widetilde{Z})$ be solutions of (11). For any $\beta > 0$, similar to the proof of (36), one can obtain

$$\begin{aligned}&\mathbb{E}\left(\sup_{t\leq s\leq T} e^{2\beta s}\left|Y(s) - \widetilde{Y}(s)\right|^2\right) + \mathbb{E}\int_t^T e^{2\beta s}\left|Z(s) - \widetilde{Z}(s)\right|^2 ds\\ &\quad\leq \frac{24M_A^2}{\beta}\mathbb{E}\\ &\qquad\times\int_t^T e^{2\beta s}\left[\theta\left(\left|Y(s) - \widetilde{Y}(s)\right|^2\right) + l\left|Z(s) - \widetilde{Z}(s)\right|^2\right]ds.\end{aligned} \tag{46}$$

That is, if β is sufficiently large,

$$\begin{aligned}&\mathbb{E}\left(\sup_{t\leq s\leq T} e^{2\beta s}\left|Y(s) - \widetilde{Y}(s)\right|^2\right)\\ &\qquad\leq \frac{24M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\left[\theta\left(\left|Y(s) - \widetilde{Y}(s)\right|^2\right)\right]ds\\ &\qquad\leq C_2\mathbb{E}\int_t^T \theta\left(\mathbb{E}\sup_{s\leq r\leq T} e^{2\beta r}\left|Y(s) - \widetilde{Y}(s)\right|^2\right)ds.\end{aligned} \tag{47}$$

An application of Bihari inequality yields

$$\mathbb{E}\left(\sup_{t\leq s\leq T} e^{2\beta s}\left|Y(s) - \widetilde{Y}(s)\right|^2\right) = 0. \tag{48}$$

So $Y(t) = \widetilde{Y}(t)$ for all $t \in [0, T]$ a.s. It then follows from (46) that $Z(t) = \widetilde{Z}(t)$ for all $t \in [0, T]$ a.s. as well. This establishes the uniqueness.

Existence. We claim that the sequence (Y_n, Z_n) defined by (27) satisfies

$$\mathbb{E}\sup_{t\leq s\leq T} e^{2\beta s}\left|Y_{n+1}(s) - Y_n(s)\right|^2 \longrightarrow 0, \quad \forall 0 \leq t \leq T, \tag{49}$$

as $n \to \infty$.

Indeed, for all $t \in [\tau_{k-1}, \tau_k]$, we set $\widetilde{\varphi}_{k,n}(t) = \mathbb{E}\sup_{s\in[t,\tau_k]} e^{2\beta s}|Y_{n+1}(s) - Y_n(s)|^2$. By Lemmas 6 and 7,

$$\begin{aligned}\widetilde{\varphi}_{k,1}(t) &= \mathbb{E}\sup_{s\in[t,\tau_k]} e^{2\beta s}\left|Y_2(s) - Y_1(s)\right|^2\\ &\leq C_2\int_t^{\tau_k}\theta\left(\mathbb{E}\sup_{s\leq r\leq\tau_k} e^{2\beta r}\left|Y_1(r) - Y_0(r)\right|^2\right)ds\\ &\leq C_2\int_t^{\tau_k}\theta\left(2C_1\exp\left(\tau_k - t\right)\right)ds\\ &\leq C_2\theta\left(2C_1\exp\left(T\right)\right)\left(\tau_k - t\right) = C_3\left(\tau_k - t\right) = \varphi_{k,1}(t).\end{aligned} \tag{50}$$

Suppose that $\widetilde{\varphi}_{k,n}(t) \leq \varphi_{k,n}(t)$ holds for some $n \geq 1$. According to Lemma 6(iii) and Lemma 7, for all $t \in [\tau_{k-1}, \tau_k]$, we obtain

$$\begin{aligned}\widetilde{\varphi}_{k,n+1}(t) &= \mathbb{E}\sup_{s\in[t,\tau_k]} e^{2\beta s}\left|Y_{n+2}(s) - Y_{n+1}(s)\right|^2\\ &\leq C_2\int_t^{\tau_k}\theta\left(\mathbb{E}\sup_{r\in[s,\tau_k]} e^{2\beta r}\left|Y_{n+1}(r) - Y_n(r)\right|^2\right)ds\\ &= C_2\int_t^{\tau_k}\theta\left(\widetilde{\varphi}_{k,n}(s)\right)ds\\ &\leq C_2\int_t^{\tau_k}\theta\left(\varphi_{k,n}(s)\right)ds = \varphi_{k,n+1}(t).\end{aligned} \tag{51}$$

This implies that, for all $n \in \mathbb{N}$,

$$\widetilde{\varphi}_{k,n}(t) \leq \varphi_{k,n}(t). \tag{52}$$

By definition, $\varphi_{k,n}(\cdot)$ is continuous on $[\tau_{k-1}, \tau_k]$. Note that for each $n \geq 1$, $\varphi_{k,n}(\cdot)$ is decreasing on $[\tau_{k-1}, \tau_k]$, and for each t, k, $\varphi_{k,n}(t)$ is a nonincreasing sequence. Therefore, we define the function $\varphi_k(t)$ by $\varphi_{k,n}(t) \downarrow \varphi_k(t)$. It is easy to verify that $\varphi_k(t)$ is continuous and nonincreasing on $[\tau_{k-1}, \tau_k]$. By the definitions of $\varphi_{k,n}(t)$ and $\varphi_k(t)$ we get

$$\varphi_k(t) = \lim_{n\to\infty} C_2\int_t^{\tau_k}\theta\left(\varphi_{k,n}(s)\right)ds = C_2\int_t^{\tau_k}\theta\left(\varphi_k(s)\right)ds \tag{53}$$

for all $\tau_{k-1} \leq t \leq \tau_k$. Since $\int_{0^+}(du/\theta(u)) = \infty$, the Bihari inequality implies

$$\varphi_k(t) = 0, \quad \text{for each } t \in \left[\tau_{k-1}, \tau_k\right]. \tag{54}$$

For each $k \in \{1, 2, \ldots, m\}$, (52) and (54) yield

$$\lim_{n\to\infty}\widetilde{\varphi}_{k,n}(t) = 0. \tag{55}$$

Then,

$$\begin{aligned}
&\mathbb{E}\sup_{t\le s\le T} e^{2\beta s}\left|Y_{n+1}(s)-Y_n(s)\right|^2 \\
&\qquad \le \max_{1\le k\le m}\mathbb{E}\sup_{s\in[\tau_{k-1},\tau_k]} e^{2\beta s}\left|Y_{n+1}(s)-Y_n(s)\right|^2 \\
&\qquad = \max_{1\le k\le m}\widetilde{\varphi}_{k,n}(t)\longrightarrow 0,
\end{aligned}\tag{56}$$

as $n\to\infty$, and this proves the assertion (49).

By (36), we obtain

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\le s\le T} e^{2\beta s}\left|Y_{n+1}(s)-Y_n(s)\right|^2\right)+\left(1-\frac{24M_A^2 l}{\beta}\right)\mathbb{E} \\
&\quad\times\int_t^T e^{2\beta s}\left|Z_{n+1}(s)-Z_n(s)\right|^2 ds \\
&\le \frac{24M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\left[\theta\left(\left|Y_n(s)-Y_{n-1}(s)\right|^2\right)\right]ds.
\end{aligned}\tag{57}$$

Applying (49) to the above formula, we see that (Y_n, Z_n) is a Cauchy (hence convergent) sequence in $\mathcal{S}^2_{\mathbb{F}}([0,T];H)\times\mathcal{H}^2_{\mathbb{F}}([0,T];\mathcal{L}(\Gamma,H))$; denote the limit by (Y,Z). Now letting $n\to\infty$ in (27), we obtain that

$$\begin{aligned}
Y(t) &= e^{A(T-t)}\xi \\
&\quad+\int_t^T \mathbb{E}'\left[e^{A(s-t)}f\left(s,Y'(s),Z'(s),Y(s),Z(s)\right)\right]ds \\
&\quad-\int_t^T e^{A(s-t)}Z(s)\,dW(s)
\end{aligned}\tag{58}$$

holds on the entire interval $[0,T]$. The theorem is now proved. □

To illustrate the application of the obtained existence and uniqueness result, we consider the example of backward stochastic partial differential equations (BSPDEs) of mean-field type.

Example 9. Let $\mathcal{O}$ be an open bounded domain in $\mathbb{R}^n$ with uniformly C^2 boundary $\partial\mathcal{O}$, let $B(t)$ be a standard n-dimensional Brownian motion (equipped with the normal filtration), and let $\xi:\mathcal{O}\to\mathbb{R}$ be an $\mathcal{F}_T$-measurable random variable. We also let L denote the semielliptic partial differential operator on $C^2(\mathbb{R})$ of the form

$$L=\sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}+\sum_{i=1}^n b_i(x)\frac{\partial}{\partial x_i}.\tag{59}$$

The aim is to study the solvability of the following initial boundary value problem:

$$\begin{aligned}
&dY(t,x) \\
&\quad=\Big(LY(t,x)+\mathbb{E}' \\
&\qquad\times\left[g\left(t,x,Y'(t,x),Z'(t,x),Y(t,x),Z(t,x)\right)\right]\Big)dt \\
&\qquad+Z(t,x)\,dB(t),\quad \text{a.e. on } (0,T)\times\mathcal{O} \\
&\qquad Y(t,x)=0,\quad \text{a.e. on } (0,T)\times\partial\mathcal{O}, \\
&\qquad Y(T,x)=\xi(T,x),\quad \text{a.e. on } \mathcal{O},
\end{aligned}\tag{60}$$

where

$$\begin{aligned}
&Y:[0,T]\times\mathcal{O}\longrightarrow\mathbb{R}, \\
&Z:[0,T]\times\mathcal{O}\longrightarrow\mathcal{L}\left(\mathbb{R}^n;L^2(\mathcal{O})\right), \\
&g:[0,T]\times\mathcal{O}\times\mathbb{R}\times\mathcal{L}\left(\mathbb{R}^n;L^2(\mathcal{O})\right) \\
&\qquad\times\mathbb{R}\times\mathcal{L}\left(\mathbb{R}^n;L^2(\mathcal{O})\right)\longrightarrow\mathbb{R}.
\end{aligned}\tag{61}$$

The following assumptions will have to be in force.

(H1) a_{ij}, $b_i:\mathbb{R}^n\to\mathbb{R}$ are uniformly continuous and bounded and satisfy the usual uniform ellipticity condition: $\sum_{i,j=1}^n a_{ij}(x)w_iw_j\ge\lambda|w|^2$, for some $\lambda>0$ and all $x\in\overline{\mathcal{O}}$, $w\in\mathbb{R}^n$.

(H2) g is measurable in $(t,x,\widetilde{y},\widetilde{z},y,z)$ and continuous in $(\widetilde{z},z)$, and there exists $C>0$ such that

$$\begin{aligned}
&\left|g\left(t,x,\widetilde{y}_1,\widetilde{z}_1,y_1,z_1\right)-g\left(t,x,\widetilde{y}_2,\widetilde{z}_2,y_2,z_2\right)\right| \\
&\quad\le C\left[\left|\widetilde{y}_1-\widetilde{y}_2\right|+\left|\widetilde{z}_1-\widetilde{z}_2\right|+\left|y_1-y_2\right|+\left|z_1-z_2\right|\right],
\end{aligned}\tag{62}$$

for all $0\le t\le T$, $x\in\mathcal{O}$, $\widetilde{y}_1,\widetilde{y}_2,y_1,y_2\in\mathbb{R}$, $\widetilde{z}_1,\widetilde{z}_2,z_1,z_2\in\mathcal{L}(\mathbb{R}^n;L^2(\mathcal{O}))$.

Then, we are now in a position of showing existence and uniqueness of the solution of BSPDEs (60).

Theorem 10. *If (H1) and (H2) are satisfied, then the mean-field BSPDE* (60) *has a unique mild solution* $(Y,Z)\in L^2(0,T;L^2(\Omega,L^2(\mathcal{O})))\times L^2_{\mathbb{F}}(0,T;L^2(\mathbb{R}^n,L^2(\Omega,L^2(\mathcal{O}))))$.

Proof. Let $H=L^2(\mathcal{O})$ and $K=\mathbb{R}^n$. Define the operator A by

$$AY(t,\cdot)=LY(t,\cdot).\tag{63}$$

It is shown in [17] (see Example 2.1 in [17]) that A generates a strongly continuous semigroup on H. Define the maps $f:[0,T]\times H\times\mathcal{L}(K,H)\times H\times\mathcal{L}(K,H)\to H$ by

$$\begin{aligned}
&f\left(t,Y'(t),Z'(t),Y(t),Z(t)\right)(x) \\
&\quad=g\left(t,x,Y'(t,x),Z'(t,x),Y(t,x),Z(t,x)\right)
\end{aligned}\tag{64}$$

for all $0 \leq t \leq T$, $x \in \mathcal{O}$. With these identifications, (60) can be written in the form of (11). By (H2), we know f satisfy condition (A1). Hence, an application of Theorem 4 concludes that (60) has a unique mild solution $(Y, Z) \in L^2(0,T;L^2(\Omega,L^2(\mathcal{O}))) \times L^2_{\mathbb{F}}(0,T;L^2(\mathbb{R}^n,L^2(\Omega,L^2(\mathcal{O}))))$. □

4. Mean-Field Stochastic Evolution Equations

Let $W(t)$, $t \in [0,T]$, be a cylindrical Wiener process with values in a Hilbert space Γ, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We fix an interval $[t,T] \subset [0,T]$ and consider the stochastic evolution equations of mean-field type for an unknown process $X(s)$, $s \in [t,T]$ with values in a Hilbert space K:

$$\begin{aligned} dX(s) &= BX(s)\,ds + \mathbb{E}'\left[b\left(s, X'(s), X(s)\right)\right] ds \\ &\quad + \mathbb{E}'\left[\sigma\left(s, X'(s), X(s)\right)\right] dW(s), \\ X(t) &= x \in K, \end{aligned} \tag{65}$$

where operator B is the generator of a strongly continuous semigroup e^{tB}, $t \geq 0$, in the Hilbert space K, with $M_B \triangleq \sup_{t\in[0,T]}|e^{tB}|$.

By a mild solution of (65) we mean an $\mathcal{F}_s$-measurable process $X(s)$, $s \in [t,T]$, with continuous paths in K, such that, $\mathbb{P}$-a.s.,

$$\begin{aligned} X(s) &= e^{B(s-t)}x \\ &\quad + \int_t^s e^{B(s-\tau)}\mathbb{E}'\left[b\left(\tau, X'(\tau), X(\tau)\right)\right] d\tau \\ &\quad + \int_t^s e^{B(s-\tau)}\mathbb{E}'\left[\sigma\left(\tau, X'(\tau), X(\tau)\right)\right] dW(\tau), \\ &\qquad s \in [t,T]. \end{aligned} \tag{66}$$

We suppose the following.

(A4) $b : [0,T] \times K \times K \to K$ is a measurable mapping which satisfies

$$\begin{aligned} &\left|b\left(t,x',x\right) - b\left(t,y',y\right)\right|^2 \\ &\quad \leq L_1\left(\left|x'-y'\right|^2 + |x-y|^2\right), \\ &\quad t \in [0,T],\ x', x, y', y \in K, \end{aligned} \tag{67}$$

for some constant $L_1 > 0$.

(A5) The mapping $\sigma : [0,T] \times K \times K \to \mathcal{L}(\Gamma, K)$ fulfills that for every $v \in \Gamma$ the map $\sigma v : [0,T] \times K \times K \to K$ is measurable, for every $s > 0$, $t \in [0,T]$, x', y', x, $y \in K$, $e^{sB}\sigma(t,x',x) \in \mathcal{L}(\Gamma,K)$, and

$$\begin{aligned} &\left|e^{sB}\sigma\left(t,x',x\right)\right|_{\mathcal{L}(\Gamma,K)} \leq L_2 s^{-\gamma}\left(1+\left|x'\right|+|x|\right), \\ &\left|e^{sB}\sigma\left(t,x',x\right) - e^{sB}\sigma\left(t,y',y\right)\right|_{\mathcal{L}(\Gamma,K)} \\ &\quad \leq L_2 s^{-\gamma}\left(\left|x'-y'\right|+|x-y|\right), \\ &\left|\sigma\left(t,x',x\right)\right|_{\mathcal{L}(\Gamma,K)} \leq L_2\left(1+\left|x'\right|+|x|\right), \end{aligned} \tag{68}$$

for some constants $L_2 > 0$ and $\gamma \in [0, 1/2)$.

Theorem 11. *Under assumptions (A3) and (A4),* (65) *has a unique mild solution* $X(\cdot) \in \mathcal{S}^2_{\mathbb{F}}([t,T];K)$.

The proof is constructed in two steps like that of Theorem 4 and it uses standard arguments for stochastic evolution equations introduced in the proof of Proposition 3.2 in [3]. Since the proof is straightforward, we prefer to omit it.

Remark 12. In our paper, Lipchitz condtion (A4) is given to get the well-posedness of mean-field stochastic evolution equations. In fact, (A4) can be replaced by a weaker condition such as (A3). We just give the condition (A4) for simplicity.

From standard arguments, we can also get the following continuous dependence theorem.

Corollary 13. *Assume that for all α in a metric space F, $(b_\alpha, \sigma_\alpha)$ satisfy (A4) and (A5) with L_1 and L_2 independent of α. Also assume that*

$$\begin{aligned} &\mathbb{E}\int_0^T \Big|\mathbb{E}'\left[b_\alpha\left(s, X'(s), X(s)\right)\right] \\ &\qquad - \mathbb{E}'\left[b_{\alpha_0}\left(s, X'(s), X(s)\right)\right]\Big|^2 ds \longrightarrow 0, \\ &\mathbb{E}\int_0^T \Big|\mathbb{E}'\left[\sigma_\alpha\left(s, X'(s), X(s)\right)\right] \\ &\qquad - \mathbb{E}'\left[\sigma_{\alpha_0}\left(s, X'(s), X(s)\right)\right]\Big|^2 ds \longrightarrow 0, \end{aligned} \tag{69}$$

as $\alpha \to \alpha_0$ for all $X \in \mathcal{S}^2_{\mathbb{F}}([0,T];K)$.

If we denote by $X^\alpha(\cdot)$ the mild solution of mean-field SEE (65) *corresponding to the functions $(b_\alpha, \sigma_\alpha)$ and to the initial data x, then we have*

$$\sup_{s\in[0,T]} \mathbb{E}\left|X^\alpha(s) - X^{\alpha_0}(s)\right|^2 \longrightarrow 0, \quad \text{as } \alpha \longrightarrow \alpha_0. \tag{70}$$

5. Maximum Principle for BSPDEs of Mean-Field Type

5.1. Formulation of the Problem. Let $\mathcal{O} \in \mathbb{R}^n$ be a bounded open set with smooth boundary $\partial\mathcal{O}$ and let U, the space of controls, be a separable real Hilbert space. We denote

$$\begin{aligned} \mathcal{U} = \Big\{ &v(\cdot) \in L^2_{\mathcal{F}}(0,T;U) \\ &\mid v_t\left(\omega', \omega\right) : [0,T] \times \Omega \times \Omega \\ &\longrightarrow U \text{ is } \mathcal{F} \otimes \mathcal{F}_t\text{-progressively measurable}\Big\}. \end{aligned} \tag{71}$$

An element of $\mathscr{U}$ is called an admissible control.

For any $v \in \mathscr{U}$, we consider the following controlled BSPDE system in the state space $H = L^2(\mathscr{O})$ (norm $|\cdot|$, scalar product $\langle\cdot,\cdot\rangle$):

$$\begin{aligned} dY_t(x) &= -AY_t(x)\,dt \\ &\quad - \mathbb{E}'\left[f\left(t,x,(Y_t(x))',(Z_t(x))',\right.\right. \\ &\qquad\left.\left. Y_t(x),Z_t(x),v_t\right)\right]dt \\ &\quad + Z_t(x)\,dW(t), \quad t\in[0,T], \\ Y_T(x) &= \xi(x), \quad x\in\mathscr{O}, \end{aligned} \tag{72}$$

where A is a partial differential operator, $f:[0,T]\times\mathscr{O}\times H\times\mathscr{L}(\Gamma,H)\times H\times\mathscr{L}(\Gamma,H)\times U \to H$, and $\xi\in L^2(\Omega,\mathscr{F}_T,\mathbb{P};H)$.

The cost functional is given by

$$\begin{aligned} J(v) &= \mathbb{E}\left\{\int_0^T\int_{\mathscr{O}}\mathbb{E}'\left[h\left(s,x,(Y_s(x))',(Z_s(x))',\right.\right.\right. \\ &\qquad\left.\left. Y_s(x),Z_s(x),v_s\right)\right]dx\,ds \\ &\quad\left. + \mathbb{E}'\int_{\mathscr{O}} g\left(x,(Y_0(x))',Y_0(x)\right)dx\right\}, \end{aligned} \tag{73}$$

where

$$\begin{gathered} h:[0,T]\times\mathscr{O}\times H\times\mathscr{L}(\Gamma,H)\times H\times\mathscr{L}(\Gamma,H)\times U\longrightarrow\mathbb{R}, \\ g:\mathscr{O}\times H\times H\longrightarrow\mathbb{R}. \end{gathered} \tag{74}$$

Our purpose is to minimize the functional $J(\cdot)$ over $\mathscr{U}_{\mathrm{ad}}$, subject to the following state constraint:

$$\overline{\mathbb{E}}\int_{\mathscr{O}}\Phi\left(x,(Y_0(x))',Y_0(x)\right)dx = 0, \tag{75}$$

where

$$\Phi:\mathscr{O}\times H\times H\longrightarrow\mathbb{R}. \tag{76}$$

An admissible control $\overline{u}\in\mathscr{U}_{\mathrm{ad}}$ that satisfies

$$J(\overline{u}) = \min_{v\in\mathscr{U}_{\mathrm{ad}}} J(v) \tag{77}$$

is called optimal.

Through what follows, the following assumptions will be in force.

(L1) A is a partial differential operator with appropriate boundary conditions. We assume that A is the infinitesimal generator of a strongly continuous semigroup e^{tA}, $t\ge 0$ in H. Moreover, for every $t\in[0,T]$, $\|e^{tA}f\|_{L^2(\mathscr{O})} \le M_A\|f\|_{L^2(\mathscr{O})}$ for some constant M_A independent of t and f.

(L2) f, h, g, and Φ are continuously Gâteaux differentiable with respect to (y',z',y,z). f is continuously Gâteaux differentiable with respect to v and h is continuous with respect to v.

(L3) The derivatives of f, h, g, and Φ are Lipschitz continuous and bounded by

$$\begin{gathered} \left|f_{y'}\right|+\left|f_{z'}\right|+\left|f_y\right|+\left|f_z\right|+\left|f_v\right|+\left|\Phi_{y'}\right|+\left|\Phi_y\right|\le C, \\ \left|h_{y'}\right|+\left|h_{z'}\right|+\left|h_y\right|+\left|h_z\right|+\left|g_{y'}\right|+\left|g_y\right| \\ \le C\left(1+\left|y'\right|+\left|z'\right|+|y|+|z|+|v|\right), \end{gathered} \tag{78}$$

where C is a positive constant.

Obviously, according to Theorem 4, state equation (72) has a unique mild solution under the above assumptions.

Remark 14. We can define the second order differential operator:

$$(Af)(x) = \sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x) + \sum_{i=1}^n b_i(x)\frac{\partial f}{\partial x_i}(x). \tag{79}$$

By Example 9, A fulfills assumption (L1) if a_{ij}, b_i satisfy condition (H1).

5.2. Variation of the Trajectory. Let $\overline{u}$ be an optimal control with $(\overline{Y}(\cdot),\overline{Z}(\cdot))$ being the corresponding optimal state. Let $\varepsilon>0$ and $[r,r+\varepsilon]\subseteq[0,T]$. For any given $v\in\mathscr{U}_{\mathrm{ad}}$, we introduce the spike variation of the control $\overline{u}(\cdot)$:

$$u_t^\varepsilon = \begin{cases} v_t, & t\in[r,r+\varepsilon], \\ \overline{u}_t, & t\in[0,T]\setminus[s,s+\varepsilon]. \end{cases} \tag{80}$$

It is clear that $u^\varepsilon(\cdot)\in\mathscr{U}_{\mathrm{ad}}$.

Let $(Y^\varepsilon(\cdot),Z^\varepsilon(\cdot))$ be the trajectory corresponding to $u^\varepsilon(\cdot)$. We use the following short notation for brevity:

$$\begin{aligned} f(t) &= f\left(t,x,\left(\overline{Y}_t(x)\right)',\left(\overline{Z}_t(x)\right)',\overline{Y}_t(x),\overline{Z}_t(x),\overline{u}_t\right), \\ f(u_t^\varepsilon) &= f\left(t,x,\left(\overline{Y}_t(x)\right)',\left(\overline{Z}_t(x)\right)',\overline{Y}_t(x),\overline{Z}_t(x),u_t^\varepsilon\right). \end{aligned} \tag{81}$$

Consider the following equation:

$$\begin{aligned} dK_t^\varepsilon(x) &= -AK_t^\varepsilon(x)\,dt \\ &\quad - \mathbb{E}'\left[f_{y'}(t)\left(K_t^\varepsilon(x)\right)' + f_y(t)K_t^\varepsilon(x)\right. \\ &\qquad + f_{z'}(t)\left(Q_t^\varepsilon(x)\right)' + f_z(t)Q_t^\varepsilon(x) \\ &\qquad\left. + \frac{1}{\varepsilon}\left(f(u_t^\varepsilon)-f(t)\right)\right]dt + Q_t^\varepsilon(x)\,dW(t), \\ K_T^\varepsilon(x) &= 0. \end{aligned} \tag{82}$$

Since the coefficients in (82) are bounded, it is easy to check that there exists a unique mild solution such that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|K_t^\varepsilon\right|^2 + \int_0^T\left|Q_t^\varepsilon\right|^2dt\right] < \infty. \tag{83}$$

We have the following estimate.

Theorem 15. *There holds*

$$\lim_{\varepsilon\to 0}\mathbb{E}\left[\sup_{s\in[t,T]}\left|\frac{Y_s^\varepsilon-\overline{Y}_s}{\varepsilon}-K_s^\varepsilon\right|^2\right]=0,\quad \forall t\in[0,T],$$
$$\lim_{\varepsilon\to 0}\mathbb{E}\int_0^T\left|\frac{Z_s^\varepsilon-\overline{Z}_s}{\varepsilon}-Q_s^\varepsilon\right|^2 ds=0. \tag{84}$$

Proof. We define

$$\eta_s^\varepsilon=\frac{Y_s^\varepsilon-\overline{Y}_s}{\varepsilon}-K_s^\varepsilon,\quad \zeta_s^\varepsilon=\frac{Z_s^\varepsilon-\overline{Z}_s}{\varepsilon}-Q_s^\varepsilon,\quad s\in[0,T]. \tag{85}$$

For simplicity, let us define

$$\begin{aligned}\Lambda_s^\varepsilon&=\left((Y_s^\varepsilon)',(Z_s^\varepsilon)',Y_s^\varepsilon,Z_s^\varepsilon\right),\\ f(s,\lambda)&=f\Big(s,\overline{Y}_s'+\lambda\big(Y_s^\varepsilon-\overline{Y}_s\big)',\\ &\qquad\overline{Z}_s'+\lambda\big(Z_s^\varepsilon-\overline{Z}_s\big)',\overline{Y}_s+\lambda\left(Y_s^\varepsilon-\overline{Y}_s\right),\\ &\qquad\overline{Z}_s+\lambda\left(Z_s^\varepsilon-\overline{Z}_s\right),u_s^\varepsilon\Big).\end{aligned} \tag{86}$$

By the definition of $(Y_s^\varepsilon,Z_s^\varepsilon)$, $(\overline{Y}_s,\overline{Z}_s)$, and $(K_s^\varepsilon,Q_s^\varepsilon)$, $(\eta_s^\varepsilon,\zeta_s^\varepsilon)$ is the mild solution of

$$d\eta_s^\varepsilon=-A\eta_s^\varepsilon ds-\mathbb{E}'\left[L(s,\varepsilon)\right]ds+\zeta_s^\varepsilon dW(s),\qquad \eta_T^\varepsilon=0 \tag{87}$$

with

$$\begin{aligned}L(s,\varepsilon)&=\frac{1}{\varepsilon}\left(f\left(s,\Lambda_s^\varepsilon,u_s^\varepsilon\right)-f\left(u_s^\varepsilon\right)\right)\\ &\quad-f_{y'}(s)\left(K_s^\varepsilon\right)'-f_y(s)K_s^\varepsilon-f_{z'}(s)\left(Q_s^\varepsilon\right)'-f_z(s)Q_s^\varepsilon\\ &=\left((\eta_s^\varepsilon)'+(K_s^\varepsilon)'\right)\int_0^1 f_{y'}(s,\lambda)\,d\lambda+\left(\eta_s^\varepsilon+K_s^\varepsilon\right)\\ &\quad\times\int_0^1 f_y(s,\lambda)\,d\lambda-f_{y'}(s)\left(K_s^\varepsilon\right)'-f_y(s)K_s^\varepsilon\\ &\quad+\left((\zeta_s^\varepsilon)'+(Q_s^\varepsilon)'\right)\int_0^1 f_{z'}(s,\lambda)\,d\lambda+\left(\zeta_s^\varepsilon+Q_s^\varepsilon\right)\\ &\quad\times\int_0^1 f_z(s,\lambda)\,d\lambda-f_{z'}(s)\left(Q_s^\varepsilon\right)'-f_z(s)Q_s^\varepsilon\\ &=\left(\eta_s^\varepsilon\right)'\int_0^1 f_{y'}(s,\lambda)\,d\lambda+\eta_s^\varepsilon\int_0^1 f_y(s,\lambda)\,d\lambda\\ &\quad+\left(\zeta_s^\varepsilon\right)'\int_0^1 f_{z'}(s,\lambda)\,d\lambda+\zeta_s^\varepsilon\int_0^1 f_z(s,\lambda)\,d\lambda+\gamma_s^\varepsilon,\end{aligned} \tag{88}$$

where we denote

$$\begin{aligned}\gamma_s^\varepsilon&=\left(K_s^\varepsilon\right)'\int_0^1\left(f_{y'}(s,\lambda)-f_{y'}(s)\right)d\lambda\\ &\quad+K_s^\varepsilon\int_0^1\left(f_y(s,\lambda)-f_y(s)\right)d\lambda\\ &\quad+\left(Q_s^\varepsilon\right)'\int_0^1\left(f_{z'}(s,\lambda)-f_{z'}(s)\right)d\lambda\\ &\quad+Q_s^\varepsilon\int_0^1\left(f_z(s,\lambda)-f_z(s)\right)d\lambda.\end{aligned} \tag{89}$$

For any $\beta>0$, according to Lemma 1, we obtain

$$\begin{aligned}&\mathbb{E}\sup_{t\le s\le T}e^{2\beta s}\left|\eta_s^\varepsilon\right|^2+\mathbb{E}\int_t^T e^{2\beta s}\left|\zeta_s^\varepsilon\right|^2 ds\\ &\quad\le\frac{12M_A^2}{\beta}\int_t^T e^{2\beta s}\mathbb{E}\left[\left|\mathbb{E}'\left[L(s,\varepsilon)\right]\right|^2\right]ds\\ &\quad\le\frac{12M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\mathbb{E}'\left[\left|L(s,\varepsilon)\right|^2\right]ds.\end{aligned} \tag{90}$$

By condition (L3), we have

$$\begin{aligned}&\mathbb{E}\left[\mathbb{E}'\left[\left|L(s,\varepsilon)\right|^2\right]\right]\\ &\quad=\mathbb{E}\left[\mathbb{E}'\left[\left|L(s,\varepsilon)-\gamma_s^\varepsilon+\gamma_s^\varepsilon\right|^2\right]\right]\\ &\quad\le 8C^2\mathbb{E}\left[\mathbb{E}'\left[\left|(\eta_s^\varepsilon)'\right|^2+\left|\eta_s^\varepsilon\right|^2+\left|(\zeta_s^\varepsilon)'\right|^2+\left|\zeta_s^\varepsilon\right|^2\right]\right]\\ &\qquad+2\mathbb{E}\left[\mathbb{E}'\left[\left|\gamma_s^\varepsilon\right|^2\right]\right]\\ &\quad\le 16C^2\mathbb{E}\left[\left|\eta_s^\varepsilon\right|^2+\left|\zeta_s^\varepsilon\right|^2\right]+2\mathbb{E}\left[\mathbb{E}'\left[\left|\gamma_s^\varepsilon\right|^2\right]\right].\end{aligned} \tag{91}$$

Combined with (91), (90) yields

$$\begin{aligned}&\mathbb{E}\sup_{t\le s\le T}e^{2\beta s}\left|\eta_s^\varepsilon\right|^2+\left(1-\frac{192M_A^2C^2}{\beta}\right)\mathbb{E}\int_t^T e^{2\beta s}\left|\zeta_s^\varepsilon\right|^2 ds\\ &\quad\le\frac{192M_A^2C^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\left|\eta_s^\varepsilon\right|^2 ds\\ &\qquad+\frac{24M_A^2}{\beta}\mathbb{E}\int_t^T e^{2\beta s}\left[\mathbb{E}'\left[\left|\gamma_s^\varepsilon\right|^2\right]\right]ds.\end{aligned} \tag{92}$$

We claim that

$$\mathbb{E}\int_t^T\mathbb{E}'\left[\left|\gamma_s^\varepsilon\right|^2\right]ds\longrightarrow 0,\quad\text{as }\varepsilon\longrightarrow 0. \tag{93}$$

From (89)

$$\gamma_s^\varepsilon=I_s^1+I_s^2+I_s^3+I_s^4, \tag{94}$$

where

$$I_s^1 = (K_s^\varepsilon)' \int_0^1 \left(f_{y'}(s,\lambda) - f_{y'}(s)\right) d\lambda,$$
$$I_s^2 = K_s^\varepsilon \int_0^1 \left(f_y(s,\lambda) - f_y(s)\right) d\lambda,$$
$$I_s^3 = (Q_s^\varepsilon)' \int_0^1 \left(f_{z'}(s,\lambda) - f_{z'}(s)\right) d\lambda,$$
$$I_s^4 = Q_s^\varepsilon \int_0^1 \left(f_z(s,\lambda) - f_z(s)\right) d\lambda. \tag{95}$$

Then,

$$\mathbb{E}\int_t^T \mathbb{E}'\left[|\gamma_s^\varepsilon|^2\right] ds \le 4\mathbb{E}\int_t^T \mathbb{E}'\left[\left|I_s^1\right|^2 + \left|I_s^2\right|^2 + \left|I_s^3\right|^2 + \left|I_s^4\right|^2\right] ds. \tag{96}$$

Take $\mathbb{E}\int_t^T \mathbb{E}'[|I_s^2|^2]ds$; for example,

$$\begin{aligned}
\mathbb{E}\int_t^T \mathbb{E}'\left[\left|I_s^2\right|^2\right] ds
&= \mathbb{E}\int_t^T \mathbb{E}'\left[\left(K_s^\varepsilon \int_0^1 \left(f_y(s,\lambda) - f_y(s)\right) d\lambda\right)^2\right] ds \\
&\le \mathbb{E}\int_t^T \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(s,\lambda) - f_y(s)\right|^2 d\lambda\right] ds \\
&\le 2\mathbb{E}\int_t^T \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(s,\lambda) - f_y(u_t^\varepsilon)\right|^2 d\lambda\right] ds \\
&\quad + 2\mathbb{E}\int_t^T \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(u_t^\varepsilon) - f_y(s)\right|^2 d\lambda\right] ds.
\end{aligned} \tag{97}$$

Note that

$$\begin{aligned}
&\mathbb{E}\int_t^T \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(u_t^\varepsilon) - f_y(s)\right|^2 d\lambda\right] ds \\
&= \mathbb{E}\int_r^{r+\varepsilon} \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(u_t^\varepsilon) - f_y(s)\right|^2 d\lambda\right] ds \\
&\le \sup_{s\in[t,T]} \mathbb{E}\left[\mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(u_t^\varepsilon) - f_y(s)\right|^2 d\lambda\right]\right] \varepsilon \\
&\longrightarrow 0, \quad \text{as } \varepsilon \longrightarrow 0.
\end{aligned} \tag{98}$$

The inequality above holds due to the boundedness of $|K_s^\varepsilon|^2 \int_0^1 |f_y(u_t^\varepsilon) - f_y(s)|^2 d\lambda$. Indeed, Assumption (L3) implies the boundedness of $f_y(u_t^\varepsilon) - f_y(s)$. Meanwhile K_s^ε is the solution of mean-field BSEE (82). It can be easy to check K_s^ε is bounded since the coefficients in (82) are bounded.

On the other hand,

$$\begin{aligned}
&\mathbb{E}\int_t^T \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(s,\lambda) - f_y(u_t^\varepsilon)\right|^2 d\lambda\right] ds \\
&\le C^2\lambda^2 \mathbb{E}\int_t^T |K_s^\varepsilon|^2 \\
&\qquad \times \int_0^1 E'\left[\left|\left(Y_s^\varepsilon - \overline{Y}_s\right)'\right|^2 + \left|\left(Z_s^\varepsilon - \overline{Z}_s\right)'\right|^2\right. \\
&\qquad \left. \times \left|Y_s^\varepsilon - \overline{Y}_s\right|^2 + \left|Z_s^\varepsilon - \overline{Z}_s\right|^2\right] d\lambda\, ds,
\end{aligned} \tag{99}$$

where $(Y_s^\varepsilon, Z_s^\varepsilon)$ is the mild solution of the following equation:

$$\begin{aligned}
dY_t^\varepsilon &= -AY_t^\varepsilon dt \\
&\quad - \mathbb{E}'\left[f\left(t, (Y_t^\varepsilon)', (Z_t^\varepsilon)', Y_t^\varepsilon, Z_t^\varepsilon, u_t^\varepsilon\right)\right] dt \\
&\quad + Z_t^\varepsilon dW(t), \quad t \in [0,T], \\
Y_T^\varepsilon &= \xi
\end{aligned} \tag{100}$$

and $(\overline{Y}_s, \overline{Z}_s)$ is the mild solution of

$$\begin{aligned}
d\overline{Y}_t &= -A\overline{Y}_t dt \\
&\quad - \mathbb{E}'\left[f\left(t, \left(\overline{Y}_t\right)', \left(\overline{Z}_t\right)', \overline{Y}_t, \overline{Z}_t, \overline{u}_t\right)\right] dt \\
&\quad + \overline{Z}_t dW(t), \quad t \in [0,T], \\
\overline{Y}_T &= \xi.
\end{aligned} \tag{101}$$

By the definition of u_t^ε, according to (L2), we have

$$\begin{aligned}
&\mathbb{E}'\left[f\left(s, Y'(s), Z'(s), Y(s), Z(s), u_t^\varepsilon\right)\right] \\
&\quad \longrightarrow \mathbb{E}'\left[f\left(s, Y'(s), Z'(s), Y(s), Z(s), \overline{u}_t\right)\right]
\end{aligned} \tag{102}$$

in $L^2([0,T]; H)$ as $\varepsilon \to 0$. Using the continuous dependence theorem Corollary 5, we obtain

$$(Y_s^\varepsilon, Z_s^\varepsilon) \longrightarrow \left(\overline{Y}_s, \overline{Z}_s\right) \quad \text{as } \varepsilon \longrightarrow 0. \tag{103}$$

Then,

$$\begin{aligned}
&\mathbb{E}\int_t^T \mathbb{E}'\left[|K_s^\varepsilon|^2 \int_0^1 \left|f_y(s,\lambda) - f_y(u_t^\varepsilon)\right|^2 d\lambda\right] ds \\
&\le C^2\lambda^2 \mathbb{E}\int_t^T |K_s^\varepsilon|^2 \\
&\qquad \times \int_0^1 \mathbb{E}'\left[\left|\left(Y_s^\varepsilon - \overline{Y}_s\right)'\right|^2 + \left|\left(Z_s^\varepsilon - \overline{Z}_s\right)'\right|^2\right. \\
&\qquad \left. \times \left|Y_s^\varepsilon - \overline{Y}_s\right|^2 + \left|Z_s^\varepsilon - \overline{Z}_s\right|^2\right] d\lambda\, ds \\
&\longrightarrow 0, \quad \text{as } \varepsilon \longrightarrow 0.
\end{aligned} \tag{104}$$

Combining (98) with (104), we finally have $\mathbb{E}\int_t^T \mathbb{E}'[|I_s^2|^2]ds \to 0$, as $\varepsilon \to 0$.

The required result (93) follows by using the similar estimations for I_s^1, I_s^3, and I_s^4.

Note that $1 - 192M_A^2C^2/\beta > 0$ if β is sufficiently large. Now to prove the desired result (84) it suffices to apply Gronwall's lemma and estimate (93) to inequality (92). □

To deal with the state constraint (75), we need to recall the Ekeland variational principle.

Lemma 16 (Ekeland's variational principle, see [16, Lemma 4.1]). *Let (S,d) be a complete metric space and let $F(\cdot): S \to \mathbb{R}$ be lower semicontinuous and bounded from below. If, for $\rho > 0$, there exists $u \in S$, such that*

$$F(u) \leq \inf_{v\in S} F(v) + \rho, \tag{105}$$

then there exists $u^\rho \in S$, satisfying

$$\begin{aligned}&\text{(i) } F(u^\rho) \leq F(u),\\ &\text{(ii) } d(u^\rho, u) \leq \rho,\\ &\text{(iii) } F(u^\rho) \leq F(v) + \rho \cdot d(u^\rho, u), \quad \forall v \neq u^\rho.\end{aligned} \tag{106}$$

Now fix $v \in \mathscr{U}_{\text{ad}}$, and set

$$\begin{aligned}S &= \left\{ v(\cdot) \in \mathscr{U}_{\text{ad}} \mid \sup_{0\leq t\leq T} \mathbb{E}|v_t|^2 \leq \mathbb{E}|\overline{u}_t|^2 + |v|^2 \right\},\\ d(\overline{v}(\cdot), v(\cdot)) &= m\left\{ t \in [0,T] \mid \mathbb{E}|\overline{v}_t - v_t|^2 > 0 \right\},\\ &\qquad \forall \overline{v}(\cdot), v(\cdot) \in S,\end{aligned} \tag{107}$$

where m denotes the Lebesgue measure on $\mathbb{R}$.

The following result is proved as Proposition 4.1 in [16].

Lemma 17. *$(S, d(\cdot,\cdot))$ is a complete metric space and J^ρ is continuous and bounded on S, where*

$$\begin{aligned}J^\rho(v(\cdot)) = &\left\{ (J(v(\cdot)) - J(\overline{u}(\cdot)) + \rho)^2 \right.\\ &\left. + \left| \overline{\mathbb{E}} \int_{\mathscr{O}} \Phi\left(x, (Y_0(x))', Y_0(x)\right) dx \right|^2 \right\}^{1/2},\\ &\qquad \forall v(\cdot) \in S\end{aligned} \tag{108}$$

and (Y, Z) is the mild solution of (72) corresponding to the control v.

Now we consider the following free initial state optimal control problem:

$$\inf_{v(\cdot)\in S} J^\rho(v(\cdot)). \tag{109}$$

It is easy to check that

$$0 \leq \inf_{v(\cdot)\in S} J^\rho(v(\cdot)) \leq J^\rho(\overline{u}(\cdot)) = \rho. \tag{110}$$

According to Ekeland's variational principle, there exists a $u^\rho(\cdot) \in V$ such that

$$\begin{aligned}&\text{(i) } J^\rho(u^\rho(\cdot)) \leq \rho,\\ &\text{(ii) } d(u^\rho(\cdot), \overline{u}(\cdot)) \leq \rho,\\ &\text{(iii) } J^\rho(u^\rho(\cdot)) \leq J^\rho(v(\cdot)) + \rho d(u^\rho(\cdot), \overline{u}(\cdot)), \quad \forall v(\cdot) \in S.\end{aligned} \tag{111}$$

Using the spike variation method, we can construct $u^{\varepsilon\rho}(\cdot) \in S$ as follows:

$$u_t^{\varepsilon\rho} = \begin{cases} v_t, & t \in [s, s+\varepsilon],\\ u_t^\rho, & t \in [0,T] \setminus [s, s+\varepsilon]. \end{cases} \tag{112}$$

It is clear that $d(u^{\varepsilon\rho}(\cdot), u^\rho(\cdot)) \leq \varepsilon$. Let $(Y^{\varepsilon\rho}(\cdot), Z^{\varepsilon\rho}(\cdot))$ (resp., $(Y^\rho(\cdot), Z^\rho(\cdot))$) be the solution of (72) with respect to the control $u^{\varepsilon\rho}(\cdot)$ (resp., $u^\rho(\cdot)$). Following (82), $(K_t^{\varepsilon\rho}, Q_t^{\varepsilon\rho})$ is the mild solution of

$$\begin{aligned}K_t^{\varepsilon\rho} = &\int_t^T e^{A(s-t)}\mathbb{E}'\\ &\times \Big[f_{y'}(s, \Lambda_s^\rho, u_s^\rho)(K_s^{\varepsilon\rho})' + f_y(s, \Lambda_s^\rho, u_s^\rho) K_s^{\varepsilon\rho}\\ &\quad + f_{z'}(s, \Lambda_s^\rho, u_s^\rho)(Q_s^{\varepsilon\rho})' + f_z(s, \Lambda_s^\rho, u_s^\rho) Q_s^{\varepsilon\rho}\\ &\quad + \frac{1}{\varepsilon}\left(f(s, \Lambda_s^\rho, u_s^{\varepsilon\rho}) - f(s, \Lambda_s^\rho, u_s^\rho) \right) \Big] ds\\ &- \int_t^T e^{A(s-t)} Q_s^{\varepsilon\rho} dW(s).\end{aligned} \tag{113}$$

By Theorem 15, we know that

$$\begin{aligned}&\lim_{\varepsilon\to 0}\mathbb{E}\left[\sup_{s\in[t,T]} \left| \frac{Y_s^{\varepsilon\rho} - Y_s^\rho}{\varepsilon} - K_s^{\varepsilon\rho} \right|^2 \right] = 0, \quad \forall t \in [0,T],\\ &\lim_{\varepsilon\to 0}\mathbb{E}\int_0^T \left| \frac{Z_s^{\varepsilon\rho} - Z_s^\rho}{\varepsilon} - Q_s^{\varepsilon\rho} \right|^2 ds = 0.\end{aligned} \tag{114}$$

The proof of the following proposition is technical but based on the arguments above and we omit it.

Proposition 18. *One has*

$$\begin{aligned}&\frac{1}{\varepsilon}\overline{\mathbb{E}}\left[\Phi\left((Y_0^{\varepsilon\rho})', Y_0^{\varepsilon\rho}\right) - \Phi\left((Y_0^\rho)', Y_0^\rho\right) \right]\\ &\quad = \overline{\mathbb{E}}\left[\Phi_{y'}\left((Y_0^\rho)', Y_0^\rho\right)(K_0^{\varepsilon\rho})' + \Phi_y\left((Y_0^\rho)', Y_0^\rho\right) K_0^{\varepsilon\rho} \right]\\ &\qquad + o(\varepsilon),\end{aligned}$$

$$\begin{aligned}\frac{1}{\varepsilon}&\left(J\left(u^{\varepsilon\rho}\right)-J\left(u^{\rho}\right)\right)\\&=\overline{\mathbb{E}}\left[g_{y'}\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\left(K_0^{\varepsilon\rho}\right)'+g_y\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)K_0^{\varepsilon\rho}\right]\\&\quad+\Delta^{\varepsilon}+\frac{1}{\varepsilon}\mathbb{E}\int_0^T\mathbb{E}'\left[h\left(s,\Lambda_s^{\rho},u_s^{\varepsilon\rho}\right)\right.\\&\qquad\left.-h\left(s,\Lambda_s^{\rho},u_s^{\rho}\right)\right]ds+o\left(\varepsilon\right),\end{aligned}\tag{115}$$

where

$$\begin{aligned}\Delta^{\varepsilon}=\mathbb{E}\int_0^T\mathbb{E}'&\left[h_{y'}\left(s,\Lambda_s^{\rho},u_s^{\varepsilon\rho}\right)\left(K_s^{\varepsilon\rho}\right)'\right.\\&+h_{z'}\left(s,\Lambda_s^{\rho},u_s^{\varepsilon\rho}\right)\left(Q_s^{\varepsilon\rho}\right)'+h_y\left(s,\Lambda_s^{\rho},u_s^{\varepsilon\rho}\right)K_s^{\varepsilon\rho}\\&\left.+h_z\left(s,\Lambda_s^{\rho},u_s^{\varepsilon\rho}\right)Q_s^{\varepsilon\rho}\right]ds.\end{aligned}\tag{116}$$

5.3. Variational Inequality and Adjoint Equation. In this subsection, the adjoint process is introduced to deduce the variational inequality.

If we set $v(\cdot) = u^{\varepsilon\rho}(\cdot)$ in (111) and notice that $d(u^{\varepsilon\rho}(\cdot),u^{\rho}(\cdot))\le\varepsilon$, we get

$$-\rho\le\frac{1}{\varepsilon}\left(J^{\rho}\left(u^{\varepsilon\rho}\left(\cdot\right)\right)-J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)\right).\tag{117}$$

By Lemma 17,

$$\begin{aligned}\frac{1}{\varepsilon}&\left(J^{\rho}\left(u^{\varepsilon\rho}\left(\cdot\right)\right)-J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)\right)\\&=\frac{\left(J^{\rho}\left(u^{\varepsilon\rho}\left(\cdot\right)\right)\right)^2-\left(J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)\right)^2}{\varepsilon\left(J^{\rho}\left(u^{\varepsilon\rho}\left(\cdot\right)\right)+J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)\right)}\\&=\frac{J\left(u^{\varepsilon\rho}\left(\cdot\right)\right)+J\left(u^{\rho}\left(\cdot\right)\right)-2J\left(\overline{u}\left(\cdot\right)\right)+2\rho}{J^{\rho}\left(u^{\varepsilon\rho}\left(\cdot\right)\right)+J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)}\\&\quad\times\frac{J\left(u^{\varepsilon\rho}\left(\cdot\right)\right)-J\left(u^{\rho}\left(\cdot\right)\right)}{\varepsilon}\\&\quad+\frac{\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\varepsilon\rho}\right)',Y_0^{\varepsilon\rho}\right)\right]+\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\right]}{J^{\rho}\left(u^{\varepsilon\rho}\left(\cdot\right)\right)+J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)}\\&\quad\times\frac{\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\varepsilon\rho}\right)',Y_0^{\varepsilon\rho}\right)\right]-\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\right]}{\varepsilon}\\&\longrightarrow l_1^{\rho}\frac{J\left(u^{\varepsilon\rho}\left(\cdot\right)\right)-J\left(u^{\rho}\left(\cdot\right)\right)}{\varepsilon}\\&\quad+l_2^{\rho}\frac{\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\varepsilon\rho}\right)',Y_0^{\varepsilon\rho}\right)\right]-\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\right]}{\varepsilon},\end{aligned}\tag{118}$$

where we set

$$\begin{aligned}l_1^{\rho}&=\frac{J\left(u^{\rho}\left(\cdot\right)\right)-J\left(\overline{u}\left(\cdot\right)\right)+\rho}{J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)},\\l_2^{\rho}&=\frac{\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\right]}{J^{\rho}\left(u^{\rho}\left(\cdot\right)\right)}\end{aligned}\tag{119}$$

and use the limit

$$\begin{aligned}J\left(u^{\varepsilon\rho}\left(\cdot\right)\right)&\longrightarrow J\left(u^{\rho}\left(\cdot\right)\right),\\\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\varepsilon\rho}\right)',Y_0^{\varepsilon\rho}\right)\right]&\longrightarrow\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\right],\end{aligned}\tag{120}$$

as $\varepsilon\to 0$ according to (115).

As $|l_1^{\rho}|^2+|l_2^{\rho}|^2=1$ for all $\rho>0$, we know that there exists a subsequence of $\{l_1^{\rho},l_2^{\rho}\}$ (still denoted by $\{l_1^{\rho},l_2^{\rho}\}$) such that

$$\begin{aligned}\lim_{\rho\to 0}\left\{l_1^{\rho},l_2^{\rho}\right\}&=\left\{l_1,l_2\right\},\\\left|l_1\right|^2+\left|l_2\right|^2&=1.\end{aligned}\tag{121}$$

Combining (115), (117) with (118), we get

$$\begin{aligned}-\rho&\le l_1^{\rho}\frac{J\left(u^{\varepsilon\rho}\left(\cdot\right)\right)-J\left(u^{\rho}\left(\cdot\right)\right)}{\varepsilon}\\&\quad+l_2^{\rho}\frac{\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\varepsilon\rho}\right)',Y_0^{\varepsilon\rho}\right)\right]-\overline{\mathbb{E}}\left[\Phi\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\right]}{\varepsilon}\\&=l_1^{\rho}\Delta^{\varepsilon}+l_1^{\rho}\overline{\mathbb{E}}\left[g_{y'}\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\left(K_0^{\varepsilon\rho}\right)'\right.\\&\qquad\left.+g_y\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)K_0^{\varepsilon\rho}\right]\\&\quad+\frac{l_1^{\rho}}{\varepsilon}\mathbb{E}\int_0^T\mathbb{E}'\left[h\left(s,\Lambda_s^{\rho},u_s^{\varepsilon\rho}\right)-h\left(s,\Lambda_s^{\rho},u_s^{\rho}\right)\right]ds\\&\quad+l_2^{\rho}\overline{\mathbb{E}}\left[\Phi_{y'}\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)\left(K_0^{\varepsilon\rho}\right)'\right.\\&\qquad\left.+\Phi_y\left(\left(Y_0^{\rho}\right)',Y_0^{\rho}\right)K_0^{\varepsilon\rho}\right]+\left(l_1^{\rho}+l_2^{\rho}\right)o\left(\varepsilon\right).\end{aligned}\tag{122}$$

Next, we introduce the adjoint equation corresponding to variational equation (113), whose solution is denoted by $P^{\rho}(t)$:

$$\begin{aligned}dP^{\rho}&\left(t\right)\\&=A^{*}P^{\rho}\left(t\right)dt\\&\quad+\mathbb{E}'\left[f_{y'}\left(t,\Lambda_t^{\rho},u_t^{\rho}\right)\left(P^{\rho}\left(t\right)\right)'\right.\\&\qquad+f_y\left(t,\Lambda_t^{\rho},u_t^{\rho}\right)P^{\rho}\left(t\right)+l_1^{\rho}h_{y'}\left(t,\Lambda_t^{\rho},u_t^{\rho}\right)\\&\qquad\left.+l_1^{\rho}h_y\left(t,\Lambda_t^{\rho},u_t^{\rho}\right)\right]dt\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}'\left[f_{z'}\left(t,\Lambda_t^\rho,u_t^\rho\right)\left(P^\rho(t)\right)'+f_z\left(t,\Lambda_t^\rho,u_t^\rho\right)P^\rho(t)\right.\\
&\left.+l_1^\rho h_{z'}\left(t,\Lambda_t^\rho,u_t^\rho\right)+l_1^\rho h_z\left(t,\Lambda_t^\rho,u_t^\rho\right)\right]dW(t),\\
P^\rho(0)&=l_1^\rho\overline{\mathbb{E}}\left[g_{y'}\left(\left(Y_0^\rho(x)\right)',Y_0^\rho(x)\right)\right.\\
&\left.+g_y\left(\left(Y_0^\rho(x)\right)',Y_0^\rho(x)\right)\right]\\
&+l_2^\rho\overline{\mathbb{E}}\left[\Phi_{y'}\left(\left(Y_0^\rho(x)\right)',Y_0^\rho(x)\right)\right.\\
&\left.+\Phi_y\left(\left(Y_0^\rho(x)\right)',Y_0^\rho(x)\right)\right],
\end{aligned}
\tag{123}
$$

where A^* is the $L^2(\mathcal{O})$-adjoint operator of A. Under assumptions (L1)–(L3), this is a linear mean-field SEE with bounded coefficients. An application of Theorem 11 implies that it has a unique adapted mild solution such that $P^\rho(t)\in\mathcal{S}_{\mathbb{F}}^2([0,T];K)$.

When $\rho\to 0$, according to Corollaries 5 and 13, $P^\rho(t)$ converges to $\overline{P}(t)$, where

$$\overline{P}(t)\in\mathcal{S}_{\mathbb{F}}^2([0,T];K) \tag{124}$$

is the solution of the following equation:

$$
\begin{aligned}
\overline{P}(t)&=l_1\overline{\mathbb{E}}\left[g_{y'}\left(\overline{Y}_0',\overline{Y}_0\right)+g_y\left(\overline{Y}_0',\overline{Y}_0\right)\right]\\
&+l_2\overline{\mathbb{E}}\left[\Phi_{y'}\left(\overline{Y}_0',\overline{Y}_0\right)+\Phi_y\left(\overline{Y}_0',\overline{Y}_0\right)\right]\\
&+\int_0^t e^{A^*(t-s)}\mathbb{E}'\left[f_{y'}(s)\overline{P}'(s)+f_y(s)\overline{P}(s)\right.\\
&\left.+l_1h_{y'}(s)+l_1h_y(s)\right]ds\\
&+\int_0^t e^{A^*(t-s)}\mathbb{E}'\left[f_{z'}(s)\overline{P}'(s)+f_z(s)\overline{P}(s)\right.\\
&\left.+l_1h_{z'}(s)+l_1h_z(s)\right]dW(s).
\end{aligned}
\tag{125}
$$

The following proposition, which formally follows from Proposition 18, gives the relation between $P^\rho(t)$ and $K_t^{\varepsilon\rho}$.

Proposition 19. *Consider the following:*

$$
\begin{aligned}
&\mathbb{E}\left[P^\rho(0)K_0^{\varepsilon\rho}\right]\\
&=\mathbb{E}\int_0^T\left\{\frac{1}{\varepsilon}P^\rho(s)\mathbb{E}'\left[f\left(s,\Lambda_s^\rho,u_s^{\varepsilon\rho}\right)-f\left(s,\Lambda_s^\rho,u_s^\rho\right)\right]\right.\\
&-l_1^\rho K_s^{\varepsilon\rho}\mathbb{E}'\left[h_{y'}\left(s,\Lambda_s^\rho,u_s^\rho\right)+h_y\left(s,\Lambda_s^\rho,u_s^\rho\right)\right]\\
&\left.-l_1^\rho Q_s^{\varepsilon\rho}\mathbb{E}'\left[h_{z'}\left(s,\Lambda_s^\rho,u_s^\rho\right)+h_z\left(s,\Lambda_s^\rho,u_s^\rho\right)\right]\right\}ds.
\end{aligned}
\tag{126}
$$

The following theorem constitutes the main contribution of this section, the maximum principle for the BSPDE control system.

Theorem 20. *Let assumptions (L1)–(L3) hold. Suppose $\overline{u}(\cdot)$ is an optimal control and $(\overline{Y}(\cdot),\overline{Z}(\cdot))$ is the corresponding optimal state trajectory for the BSPDE control systems (72) and (73) with the initial state constraint (75). Then there exists $\overline{P}(t)\in\mathcal{S}_{\mathbb{F}}^2([0,T];K)$ which satisfies (125), such that*

$$
\begin{aligned}
&\mathcal{H}\left(t,\overline{Y}_t',\overline{Z}_t',\overline{Y}_t,\overline{Z}_t,v_t,\overline{P}(t)\right)\\
&\geq\mathcal{H}\left(t,\overline{Y}_t',\overline{Z}_t',\overline{Y}_t,\overline{Z}_t,\overline{u}_t,\overline{P}(t)\right),\\
&\text{a.e., a.s. }\forall v\in\mathcal{U}_{ad},
\end{aligned}
\tag{127}
$$

where $\mathcal{H}:[0,T]\times H\times\mathcal{L}(\Gamma,H)\times H\times\mathcal{L}(\Gamma,H)\times U\times K\to\mathbb{R}$ is the Hamiltonian function defined by

$$
\begin{aligned}
\mathcal{H}(t,\tilde{y},\tilde{z},y,z,v,p)&=l_1h(t,\tilde{y},\tilde{z},y,z,v)\\
&+pf(t,\tilde{y},\tilde{z},y,z,v).
\end{aligned}
\tag{128}
$$

Proof. By (122) and Proposition 19, we obtain

$$
\begin{aligned}
-\rho&\leq\frac{1}{\varepsilon}\mathbb{E}\int_0^T P^\rho(s)\mathbb{E}'\left[f\left(s,\Lambda_s^\rho,u_s^{\varepsilon\rho}\right)\right.\\
&\left.-f\left(s,\Lambda_s^\rho,u_s^\rho\right)\right]ds\\
&+\frac{l_1^\rho}{\varepsilon}\mathbb{E}\int_0^T\mathbb{E}'\left[h\left(s,\Lambda_s^\rho,u_s^{\varepsilon\rho}\right)\right.\\
&\left.-h\left(s,\Lambda_s^\rho,u_s^\rho\right)\right]ds+\left(l_1^\rho+l_2^\rho\right)o(\varepsilon).
\end{aligned}
\tag{129}
$$

Letting $\varepsilon\to 0^+$ in (129), we derive, for a.e. $\tau\in[0,T]$,

$$
\begin{aligned}
-\rho&\leq l_1^\rho\mathbb{E}'\left[h\left(\tau,\Lambda_\tau^\rho,v_\tau\right)-h\left(\tau,\Lambda_\tau^\rho,u_\tau^\rho\right)\right]\\
&+P^\rho(s)\mathbb{E}'\left[f\left(\tau,\Lambda_\tau^\rho,v_\tau\right)-f\left(\tau,\Lambda_\tau^\rho,u_\tau^\rho\right)\right],
\end{aligned}
\tag{130}
$$

for all $v\in\mathcal{U}_{\mathrm{ad}}$.

Finally, taking $\rho\to 0$ in (130), we derive the desired result. ☐

Remark 21. We note that if the coefficients do not depend explicitly on the marginal law of the underlying diffusion, the result reduces to the classical case, that is, the SMP for BSPDEs without mean-field term.

Remark 22. When we remove the initial state constraint (75), we obtain the general maximum principle for the mean-field BSPDEs system (i.e., without the constraint) with $l_1=1$.

5.4. Application: A Backward Linear Quadratic Control Problem. Now, we apply our maximum principle to solve an LQ problem. For notational simplicity, we restrict ourselves to the free case (i.e., without the initial state constraint (75)), the general case being handled in a similar way.

Consider the following problem:

$$J(u) = \frac{1}{2}\mathbb{E}\left[(Y(0))^2\right] + \frac{1}{2}\mathbb{E}\int_0^T \left\{N v^2(t)\right\} dt \longrightarrow \min \tag{131}$$

subject to

$$\begin{aligned} dY(t) &= -AY(t)\,dt \\ &\quad - \left\{BY(t) + \widetilde{B}\mathbb{E}[Y(t)] \right. \\ &\quad \left. + Cv(t) + DZ(t) + \widetilde{D}\mathbb{E}[Z(t)]\right\} dt \\ &\quad + Z(t)\,dW(t), \\ Y(T) &= \xi, \quad t \in [0,T], \end{aligned} \tag{132}$$

where A is a partial differential operator satisfying condition (L1) and $B, \widetilde{B}, C, D, \widetilde{D}$, and N are bounded and deterministic constants. We also assume that $N > 0$ and $v \in L^2_{\mathbb{F}}(0,T;U)$. It is easy to verify that BSPDE (132) admits a unique mild solution $(Y(t), Z(t))$.

$P(t)$, the adjoint process of state equation (132), is the solution of

$$\begin{aligned} dP(t) &= A^* P(t)\,dt + \left\{BP(t) + \widetilde{B}\mathbb{E}[P(t)]\right\} dt \\ &\quad + \left\{DP(t) + \widetilde{D}\mathbb{E}[P(t)]\right\} dW(t), \\ P(0) &= Y(0). \end{aligned} \tag{133}$$

Let $\overline{u}(\cdot)$ be an optimal control, and let $(\overline{Y}(\cdot), \overline{Z}(\cdot))$ be the corresponding state process. By maximum principle of Theorem 20 (note that $l_1 = 1$ in this problem),

$$\frac{1}{2}Nv^2(t) + P(t)Cv(t) \geq \frac{1}{2}N\overline{u}^2(t) + P(t)C\overline{u}(t) \tag{134}$$

for all $v \in U_{\mathrm{ad}}$ since the state equation has the form (132). This in turn implies

$$\overline{u}(t) = -\frac{C}{N}P(t). \tag{135}$$

It is clear that (131) is a positive quadratic functional of control because of $N > 0$. Hence, an optimal control exists. The candidate optimal control (135) is indeed an optimal control of this LQ problem for it is the only control which satisfies the maximum principle.

Next, we want to obtain a more explicit representation of the optimal control (135) from the state equation (132). Substituting (135) into state equation (132) yields

$$\begin{aligned} dY(t) &= -AY(t)\,dt \\ &\quad - \left\{BY(t) + \widetilde{B}\mathbb{E}[Y(t)] - \frac{C^2}{N}P(t) \right. \\ &\quad \left. + DZ(t) + \widetilde{D}\mathbb{E}[Z(t)]\right\} dt + Z(t)\,dW(t), \\ Y(T) &= \xi, \quad t \in [0,T]. \end{aligned} \tag{136}$$

Combining the above equation with (133), we obtain the following related feedback control system:

$$\begin{aligned} dY(t) &= -AY(t)\,dt \\ &\quad - \left\{BY(t) + \widetilde{B}\mathbb{E}[Y(t)] - \frac{C^2}{N}P(t) \right. \\ &\quad \left. + DZ(t) + \widetilde{D}\mathbb{E}[Z(t)]\right\} dt + Z(t)\,dW(t), \\ Y(T) &= \xi, \\ dP(t) &= A^* P(t)\,dt + \left\{BP(t) + \widetilde{B}\mathbb{E}[P(t)]\right\} dt \\ &\quad + \left\{DP(t) + \widetilde{D}\mathbb{E}[P(t)]\right\} dW(t), \\ P(0) &= Y(0). \end{aligned} \tag{137}$$

Looking at the terminal condition of $P(t)$ in (133) and considering the mean-field type of (132), it is reasonable to conjecture that $P(t)$ has the following form:

$$P(t) = \varphi(t)Y(t) + \phi(t)\mathbb{E}[Y(t)], \tag{138}$$

where $\varphi(t), \phi(t)$ are deterministic differential functions which will be specified below. Moreover, $\varphi(0) = 1$, $\phi(0) = 0$.

Inserting this form into adjoint equation (133) and noticing that $Y(t)$ satisfies (136), we can compare the coefficients of dt and $dW(t)$ to obtain the following equation:

$$\begin{aligned} &2\varphi(t)AY(t) + 2\phi(t)A\mathbb{E}[Y(t)] + B\varphi(t)Y(t) \\ &\quad + 2\left(B\phi(t) + \widetilde{B}\varphi(t) + \widetilde{B}\phi(t)\right)\mathbb{E}[Y(t)] \\ &= -\varphi(t)\left\{BY(t) - \frac{C^2}{N}\varphi(t)Y(t) + DZ(t)\right\} \\ &\quad + Y(t)\frac{d\varphi(t)}{dt} + \mathbb{E}[Y(t)]\frac{d\phi(t)}{dt} \\ &\quad + \left(\phi^2(t)\frac{C^2}{N} + 2\frac{C^2}{N}\varphi(t)\phi(t)\right)\mathbb{E}[Y(t)] \\ &\quad - \left(D\phi(t) + \widetilde{D}\phi(t) + \widetilde{D}\varphi(t)\right)\mathbb{E}[Z(t)], \\ &\varphi(t)Z(t) = D\varphi(t)Y(t) \\ &\quad + \left(D\phi(t) + \widetilde{D}\varphi(t) + \widetilde{D}\phi(t)\right)\mathbb{E}[Y(t)]. \end{aligned} \tag{139}$$

Then, subtracting $Z(t)$ we have

$$\begin{aligned} &2\varphi(t)AY(t) + 2\phi(t)A\mathbb{E}[Y(t)] + B\varphi(t)Y(t) \\ &\quad + 2\left(B\phi(t) + \widetilde{B}\varphi(t) + \widetilde{B}\phi(t)\right)\mathbb{E}[Y(t)] \end{aligned}$$

$$
\begin{aligned}
&= \varphi(t)\left\{-B + \frac{C^2}{N}\varphi(t) - D^2\right\} Y(t) \\
&\quad + Y(t)\frac{d\varphi(t)}{dt} + \mathbb{E}[Y(t)]\frac{d\phi(t)}{dt} \\
&\quad + \left(\phi^2(t)\frac{C^2}{N} + 2\frac{C^2}{N}\varphi(t)\phi(t)\right)\mathbb{E}[Y(t)] \\
&\quad - \frac{1}{\varphi(t)}\left(D\phi(t) + \widetilde{D}\varphi(t) + \widetilde{D}\phi(t)\right)^2\mathbb{E}[Y(t)] \\
&\quad - 2D\left(D\phi(t) + \widetilde{D}\varphi(t) + \widetilde{D}\phi(t)\right)\mathbb{E}[Y(t)].
\end{aligned}
\tag{140}
$$

Comparing the coefficients of $Y(t)$ and $\mathbb{E}[Y(t)]$, respectively, we get

$$
\begin{gathered}
\frac{d}{dt}\varphi(t) = 2A^*\varphi(t) + 2B\varphi(t) + D^2\varphi(t) - \frac{C^2}{N}\varphi^2(t), \\
\varphi(0) = 1,
\end{gathered}
\tag{141}
$$

$$
\begin{aligned}
\frac{d}{dt}\phi(t) &= 2A^*\phi(t) + \left(\frac{\left(D+\widetilde{D}\right)^2}{\varphi(t)} - \frac{C^2}{N}\right)\phi^2(t) \\
&\quad + 2R(t)\phi(t) + \left(\widetilde{D}^2 + 2\widetilde{B} + 2D\widetilde{D}\right)\varphi(t), \\
&\phi(0) = 0,
\end{aligned}
\tag{142}
$$

where $R(t) = B + \widetilde{B} + (D+\widetilde{D})^2 - (C^2/N)\varphi(t)$.

We solve (141) to get

$$
\begin{aligned}
\varphi(t) = \Bigg(& e^{-(2A^*+2B+D^2)t}\left(1 - \frac{C^2}{N(2A^*+2B+D^2)}\right) \\
&+ \frac{C^2}{N(2A^*+2B+D^2)}\Bigg)^{-1}.
\end{aligned}
\tag{143}
$$

Then (142) exists, a unique solution from the classical Riccati equation theory.

We now conclude the above discussions in the following result.

Theorem 23. *For one's linear quadratic stochastic partial differential control problem* (131)-(132)*, the unique optimal control* $\overline{u}(\cdot) \in U_{ad}$ *is given by*

$$
\overline{u}(t) = -\frac{C}{N}\left(\varphi(t)Y(t) + \phi(t)\mathbb{E}[Y(t)]\right),
\tag{144}
$$

with $\varphi(t)$ *satisfying* (143) *and* $\phi(t)$ *solving* (142).

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

Acknowledgment

This research is partially supported by the Natural Science Foundation of China under Grant no. 11301310.

References

[1] Y. Hu and S. Peng, "Adapted solution of a backward semilinear stochastic evolution equation," *Stochastic Analysis and Applications*, vol. 9, no. 4, pp. 445–459, 1991.

[2] N. I. Mahmudov and M. A. McKibben, "On backward stochastic evolution equations in Hilbert spaces and optimal control," *Nonlinear Analysis*, vol. 67, no. 4, pp. 1260–1274, 2007.

[3] M. Fuhrman and G. Tessitore, "Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control," *The Annals of Probability*, vol. 30, no. 3, pp. 1397–1465, 2002.

[4] M. Fuhrman and G. Tessitore, "Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces," *Annals of Probability*, vol. 32, no. 1B, pp. 607–660, 2004.

[5] X. Mao, "Adapted solutions of backward stochastic differential equations with non-Lipschitz coefficients," *Stochastic Processes and their Applications*, vol. 58, no. 2, pp. 281–292, 1995.

[6] J. Lasry and P. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[7] R. Buckdahn, J. Li, and S. Peng, "Mean-field backward stochastic differential equations and related partial differential equations," *Stochastic Processes and their Applications*, vol. 119, no. 10, pp. 3133–3154, 2009.

[8] D. Andersson and B. Djehiche, "A maximum principle for SDEs of mean-field type," *Applied Mathematics and Optimization*, vol. 63, no. 3, pp. 341–356, 2011.

[9] R. Buckdahn, B. Djehiche, J. Li, and S. Peng, "Mean-field backward stochastic differential equations: a limit approach," *The Annals of Probability*, vol. 37, no. 4, pp. 1524–1565, 2009.

[10] R. Xu, "Mean-field backward doubly stochastic differential equations and related SPDEs," *Boundary Value Problems*, vol. 2012, article 114, 2012.

[11] R. Buckdahn, B. Djehiche, and J. Li, "A general stochastic maximum principle for SDEs of mean-field type," *Applied Mathematics and Optimization*, vol. 64, no. 2, pp. 197–216, 2011.

[12] J. Li, "Stochastic maximum principle in the mean-field controls," *Automatica*, vol. 48, no. 2, pp. 366–373, 2012.

[13] G. Wang, C. Zhang, and W. Zhang, "Stochastic maximum principle for mean-field type optimal control under partial information," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 522–528, 2014.

[14] M. Hafayed, "A mean-field necessary and sufficient conditions for optimal singular stochastic control," *Communications in Mathematics and Statistics*, vol. 1, no. 4, pp. 417–435, 2013.

[15] M. Hafayed, "A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes," *International Journal of Dynamics and Control*, vol. 1, no. 4, pp. 300–315, 2013.

[16] Y. Hu and S. Peng, "Maximum principle for semilinear stochastic evolution control systems," *Stochastics and Stochastics Reports*, vol. 33, no. 3-4, pp. 159–180, 1990.

[17] M. Fuhrman, Y. Hu, and G. Tessitore, "Stochastic maximum principle for optimal control of SPDEs," *Comptes Rendus Mathematique*, vol. 350, no. 13-14, pp. 683–688, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 902395, 7 pages
http://dx.doi.org/10.1155/2014/902395

Research Article

Stochastic Resonance in Neuronal Network Motifs with Ornstein-Uhlenbeck Colored Noise

Xuyang Lou

Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University, Wuxi 214122, China

Correspondence should be addressed to Xuyang Lou; Louxy@126.com

Received 24 April 2014; Accepted 28 June 2014; Published 10 July 2014

Academic Editor: Ming Gao

We consider here the effect of the Ornstein-Uhlenbeck colored noise on the stochastic resonance of the feed-forward-loop (FFL) network motif. The FFL motif is modeled through the FitzHugh-Nagumo neuron model as well as the chemical coupling. Our results show that the noise intensity and the correlation time of the noise process serve as the control parameters, which have great impacts on the stochastic dynamics of the FFL motif. We find that, with a proper choice of noise intensities and the correlation time of the noise process, the signal-to-noise ratio (SNR) can display more than one peak.

1. Introduction

Recently, the dynamics of networks of bioinspired neurons has received wide attentions in all branches of science. Stochastic resonance, as a nonlinear effect in which noise can enhance the detection of weak signals [1], is one of the central topics in theoretical and computational neuroscience. Kwon and Moon [2] investigated the role of different connectivity regimes on the coherence resonance of Hodgkin-Huxley neurons and found that spatial synchronization increases as characteristic path length shortens and firing frequency increases as clustering coefficient decreases. Ozer et al. [3] introduced a subthreshold periodic driving to a single neuron of the Newman-Watts small-world network consisting of biophysically realistic Hodgkin-Huxley neurons and found that the stochastic resonance phenomenon depends significantly also on the coupling strength among neurons and the driving frequency.

Noise can provide improvements in the representation of weak signals through stochastic resonance [4]. Gong et al. [5] analyzed the temporal coherence and the spatial synchronization of the stochastic Hodgkin-Huxley networks with channel noise and found that the random shortcuts can enhance the coherence and synchronization, which are absent in the regular network. It has been found that many studies have been carried out in an attempt to the role of noise in stochastic resonance of neural systems ranging from the microscale to the macroscale [6–9].

However, few reports are available on the stochastic resonance of the small-scale neural motifs [10] which are subnetworks occurring frequently in complex networks and believed to be basic building blocks of many networks. It is demonstrated that network motifs can perform specific functional roles and do exist in real biological networks, such as protein-protein interaction networks [11], brain functional networks [12], neuronal networks [13], and transcription regulatory networks [14]. Among the neural motifs, the feed-forward-loop (FFL) motif commonly occurs in dendrite and feed-forward structure, in which two neurons are unidirectionally coupled to the third one. Therefore, it is of practical importance to understand how the noise, especially the Ornstein-Uhlenbeck colored noise, contributes to the neuronal information processing in FFL motif.

The structure of the paper is as follows. First, we introduce the basic FitzHugh-Nagumo equation formulation of neuronal population dynamics in Section 2. The dynamical model is equipped with a noise current modeled by an Ornstein-Uhlenbeck (OU) process. In Section 3, the effects of noise on the signal-to-noise ratio (SNR) are analyzed in the single neuronal population. We extend the model and explore

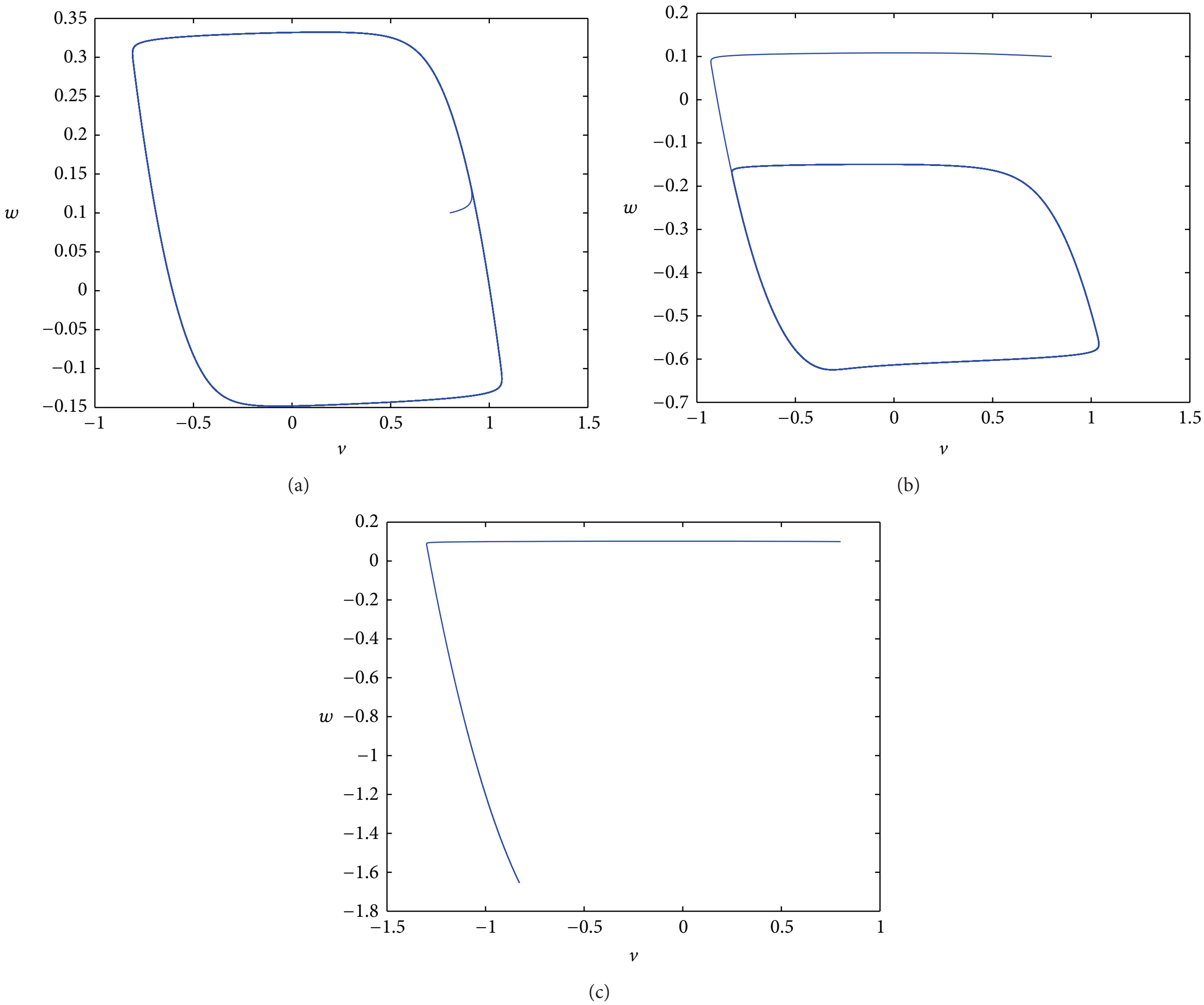


FIGURE 1: Phase plots of the FH model with initial values $v(0) = 0.8$ and $w(0) = 0.1$ and different current values: (a) $I = 0.0$, (b) $I = -0.5$, and (c) $I = -2.0$.

the effects of colored noise on the SNR in the FFL neuronal motif in Section 4. Finally, a brief conclusion and discussion of our work are given.

2. The Model

The Hodgkin-Huxley model is of importance in describing the transmission of an action potential through a cell membrane [15]. However, due to the large number of variables, the phase space dynamics of the equation is hard to visualize. The FitzHugh-Nagumo (FH) neuron model has been proposed as a simplified model of the cell membrane [16], experimentally demonstrated by Nagumo et al. [17] using electrical circuits.

The FH model considered here is defined in a dimensionless form taken from [18]:

$$\begin{aligned} \epsilon \dot{v} &= v(v+a)(1-v) - w + I, \\ \dot{w} &= v - 0.5w, \end{aligned} \tag{1}$$

where v is the voltage potential of the neuron membrane, w is the inactivation of the sodium channels, and I represents the input current. ϵ acts as the singular perturbation parameter and w evolves on a much slower time scale than the voltage potential v. Here, we take the time scale separation variable $\epsilon = 0.01$ and set the parameter $a = 0.6$. For different current values, the neural system may exhibit complex dynamics (see Figure 1).

The total current input consists of the external applied current I^{ext}, the total synaptic current I^{syn}, and the noise current η; that is, $I = I^{\text{ext}} + I^{\text{syn}} + \eta$. In the single neuron level, the total synaptic current I^{syn} will be taken as zero. For the noise current, the independent noise processes are governed by [19]

$$\dot{\eta} = -\frac{1}{\tau}\frac{d}{d\eta}U_p(\eta_i) + \frac{\sqrt{D}}{\tau}\xi(t), \tag{2}$$

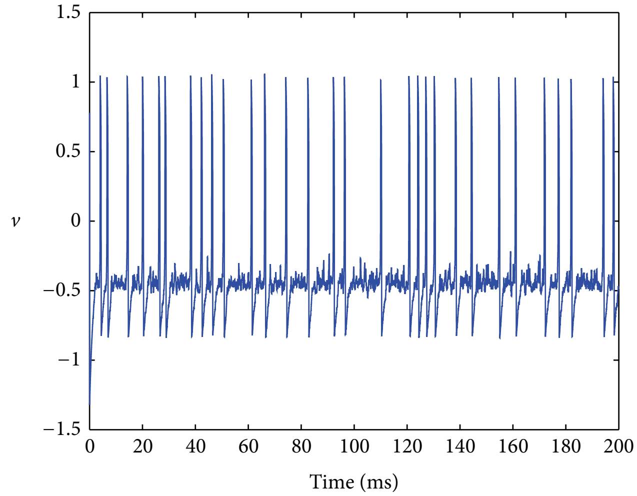


Figure 2: The time series of v of the FH model with $D = 6$ and $\tau = 0.5$.

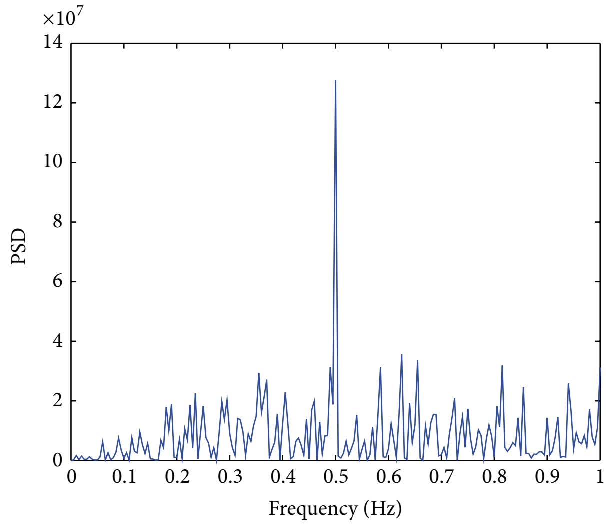


Figure 3: The power spectrum density graph of the time series of v in Figure 2.

where the potential function is

$$U_p(\eta) = \left(\frac{D}{\tau}(p-1)\right)\ln\left[1 + \frac{\alpha(p-1)\eta^2}{2}\right] \quad (3)$$

with $\alpha = \tau/D$. $\xi(t)$ is the Gaussian white noise process defined via $\langle\xi(t)\xi(t')\rangle = 2\delta(t-t')$ and $\langle\xi(t)\rangle = 0$. D and τ define the intensity and the correlation time of the noise process. The form of the noise η allows us to control the deviation from the Gaussian behavior by changing a single parameter p. For $p = 1$, (2) becomes

$$\dot{\eta} = -\frac{\eta}{\tau} + \frac{\sqrt{D}}{\tau}\xi(t), \quad (4)$$

which is a well-known time evolution equation for the OU noise process [20].

3. Single Neuron Level

Before analyzing the effects of OU noise on the FFL neuron motif, let us now discuss the stochastic resonance in single neuron level. To do this, a localized weak rhythmic activity is introduced in the form of $I^{\text{ext}} = I_0 + A\sin(\omega t)$, which is added additively to the neuron. Here, $I_0 = -2$ is the bias current and $A = 0.01$ denotes the amplitude of the sinusoidal forcing current, whereas $\omega = 0.5$ is the corresponding angular frequency. In this scenario, the external current is not large enough to excite the neuron without the induction of noise. We investigated the effects of noise as perturbation on the FH neuron in order to examine the stochastic oscillation phenomena. We set the noise intensity $D = 6$ and the correlation time of the noise process $\tau = 0.5$. As we see in Figure 2, regular spiking oscillations are induced due to the presence of the noise.

Next, we solved (1) and (4) numerically using a stochastic version of the Euler discretization scheme with a time step of 1 ms much smaller than the time constants of the problem. To quantitatively evaluate the performances of stochastic resonance, all data shown in our work refer to averages over 50 different realizations. Here, similar to [10], we use the SNR to measure the relative performance of stochastic resonance quantitatively. The SNR is one of the important measures for reducing the noise. There are different definitions for SNR. To calculate the SNR, the power spectral density (PSD) is first obtained from the time series of the membrane potential. An example of the PSD for the time series of v in Figure 2 is depicted in Figure 3. It is seen that the PSD contains a main peak located at the forcing frequency $f = 0.5$ Hz and a background noise. This indicates that the frequency characteristic of the output spike train is induced by the local weak periodic forcing [10]. Define the SNR as $R = 10\log_{10}(S(f)/N(f))$, where f is the input signal frequency, $S(f)$ denotes the value of the peak power at the frequency f in the power spectrum of the time series of the output spike train, and $N(f)$ is the averaged power at nearby frequencies. A large SNR means that there is a larger variation of the signal amplitude than of the noise.

Now, we are ready to study the effects of different D values on the relative performance of stochastic resonance. Figure 4 shows the SNR diagram with respect to the noise intensity as the control parameter. When D is too small, there is almost no effect of noise on the spiking activities of the neuron. For an appropriate noise intensity, due to the excitatory effect of noise, the neuron starts to fire spikes. As a result of the increasing noise intensity, the SNR curves all first rise, then drop, and finally are maintained at a certain level, indicating that there exists an optimal noise intensity for the best performance.

4. Motif Neuron Level

Consider the FFL neuronal network motif based on coupled FH neurons given by

$$\begin{aligned} \epsilon\dot{v}_i &= v_i(v_i + a)(1 - v_i) - w_i + I_i, \\ \dot{w}_i &= v_i - 0.5w_i, \end{aligned} \quad (5)$$

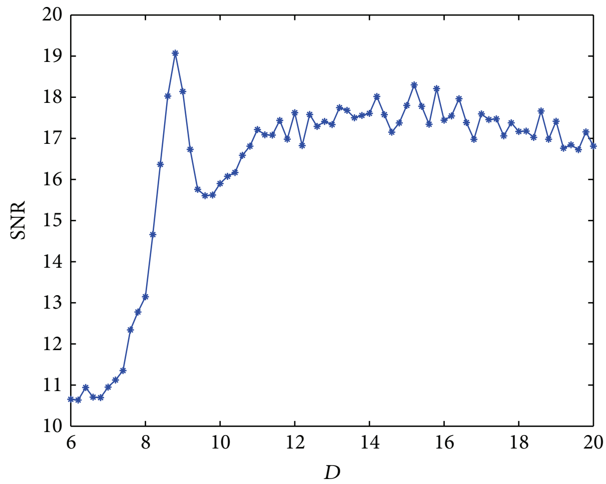


Figure 4: The SNR versus the noise intensity when $\tau = 0.5$.

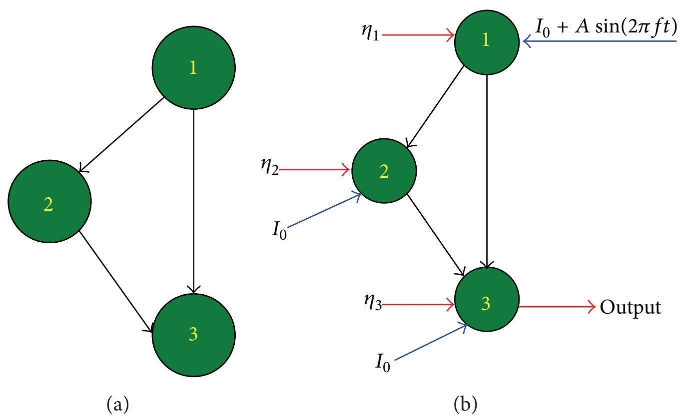


Figure 5: The FFL neuronal network motif. (a) The FFL motif. (b) Input-output structure of the FFL motif.

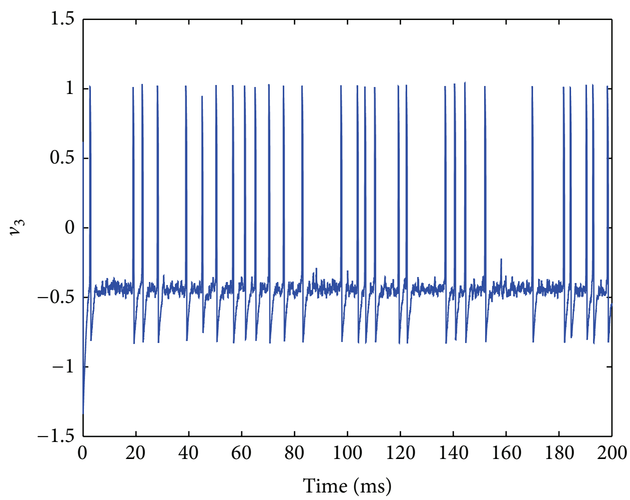


Figure 6: The stochastic oscillation of v_3 in the FFL motif, with $D = 6$, $\tau = 0.5$, and $g = 0.113$.

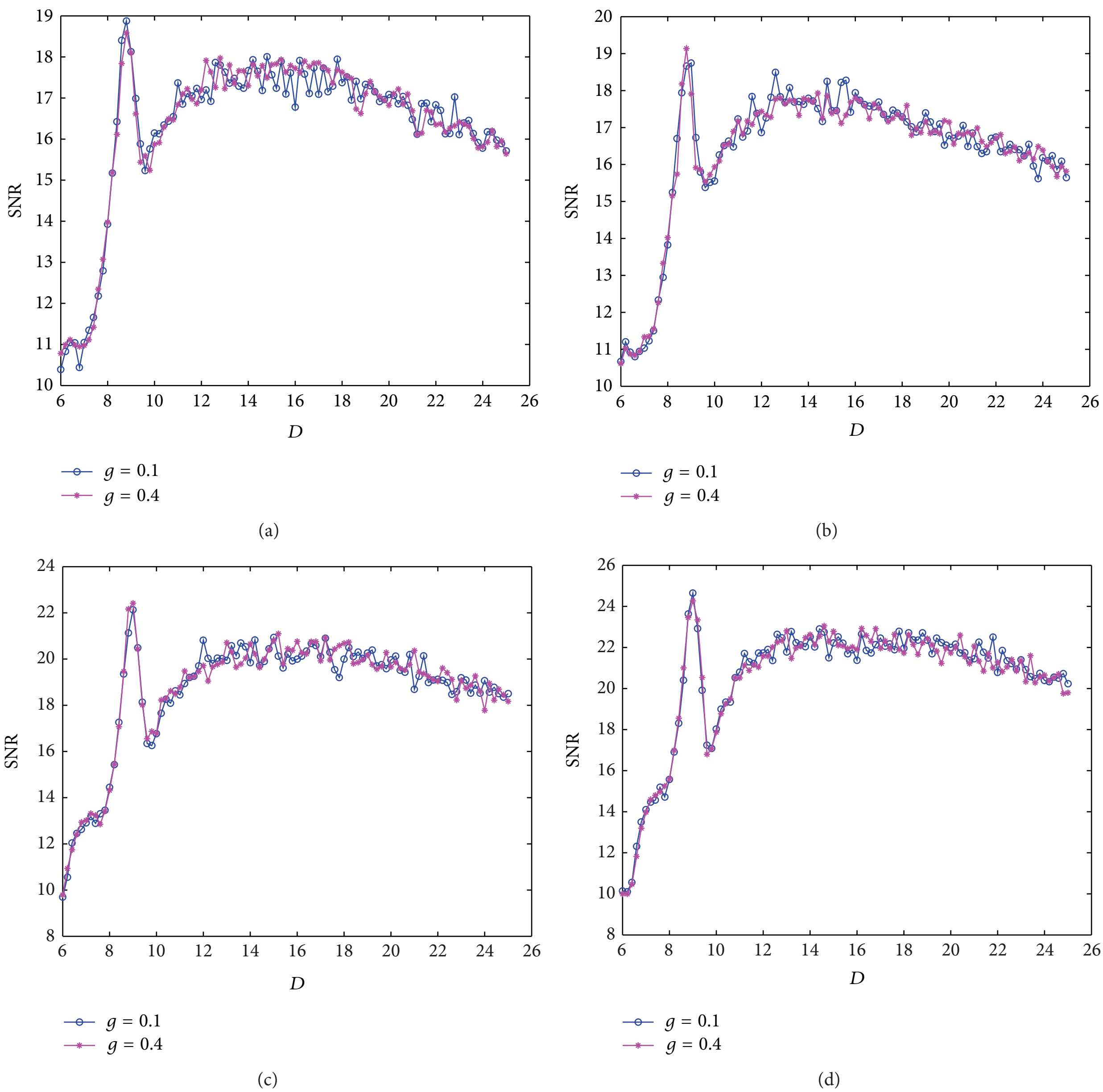


FIGURE 7: The SNR versus the noise intensity for different coupling strengths. (a) $\tau = 0.2$, (b) $\tau = 0.5$, (c) $\tau = 0.8$, and (d) $\tau = 1.1$.

where $i = 1, 2, 3$ index the neurons, v_i is the membrane potential of the ith neuron, and w_i is the inactivation of the sodium channels. The total current input I_i consists of the external applied current I_i^{ext}, the total synaptic current I_i^{syn}, and the noise current η_i; that is, $I_i = I_i^{\text{ext}} + I_i^{\text{syn}} + \eta_i$. The external applied current I_i^{ext} and the noise current η_i are defined the same as in single neuron level. In this work, the synaptic current onto neuron i will be the linear sum of the currents of all incoming synapses, $I_i^{\text{syn}} = -g\sum_{j=1}^{3} w_{ij}x_j(t)$, where g describes the coupling strength of the synapse between neurons. $W = (w_{ij})_{3\times 3}$ is the Laplacian matrix of the network motif, where $w_{ij} \neq 0$ $(i \neq j)$ implies that there is a connection from neuron j to neuron i. x_i represents membrane potential of neuron i. For simplicity, assume that the coupling strength is identical for all connections; that is, $g_{ij} = g$. For the noise current, the OU noise η_i is given by

$$\dot{\eta}_i = -\frac{\eta_i}{\tau} + \frac{\sqrt{D}}{\tau}\xi_i(t). \tag{6}$$

In this section, we focus on the stochastic resonance of the coupled neurons in FFL motif structure. There are different configurations for FFL depending on the excitatory or inhibitory of the neurons in the motif. Here, we consider the FFL motif shown in Figure 5(a), that is, the T1-FFL type in [10], in which the first neuron and the second neuron are unidirectionally coupled to the third neuron.

In the following simulations, we use the same discretization scheme to solve (5) and (6) numerically. And the data

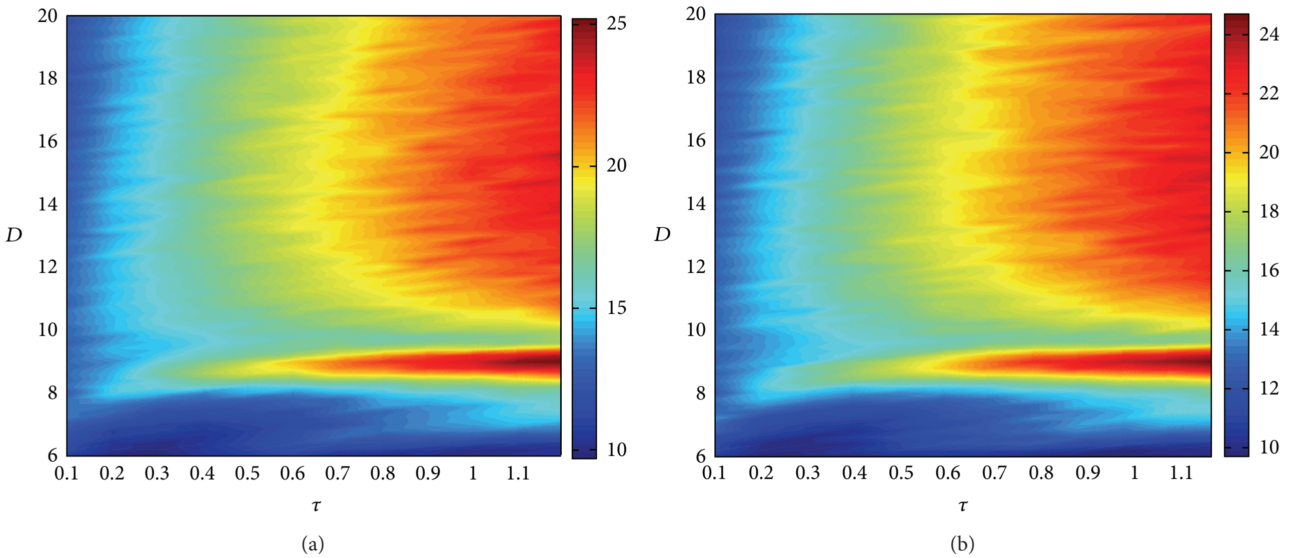


Figure 8: Color-featured SNR in dependence on D and τ for different g. (a) $g = 0.1$ and (b) $g = 3.0$.

are averaged results of 50 independent runs in order to quantitatively evaluate the performances. As stated in [10], we only examine the response of neuron 3 to the external applied current of neuron 1 as neurons 1 and 3 are repetitively regarded as the input and output neurons of the FFL neuronal network motif. The input-output structure of the FFL motif is shown in Figure 5(b).

To examine whether the FFL neuronal network motif exhibits the stochastic resonance, we set $I_2^{\text{ext}} = I_3^{\text{ext}} = -2$, and a localized weak rhythmic activity is introduced in the form of $I_1^{\text{ext}} = I_0 + A\sin(\omega t)$, which is added additively to the first neuron in (5). Here, $I_0 = -2$ is the bias current and $A = 0.01$ denotes the amplitude of the sinusoidal forcing current, whereas $\omega = 0.5$ is the corresponding angular frequency. Different from the simulations in single neuron level, we will investigate the effects of the SNR not only on the strength of noise D but also on the coupling strength g and correlation time τ has been thoroughly studied. To examine the stochastic oscillation phenomena in the FFL motif level, we set the coupling strength $g = 0.113$, the noise intensity $D = 6$, and the correlation time of the noise process $\tau = 0.5$. As we see in Figure 6, the output v_3 of the FFL motif exhibits spiking oscillations induced by the noise in the first neuron.

In what follows, we will systemically analyze effects of different D, τ, and g on the relative performance of stochastic resonance via SNR. First, we examine the dependence of SNR on D and g with fixed values τ. Figure 7 shows the SNR diagram with respect to the noise intensity as the control parameter. Similar phenomenon in Figure 4 can be observed from the curves. Due to the low input stimulus, there is almost no effect of noise on the spiking activities of the coupled neurons. However, along with the increasing noise intensity, the SNR starts to enhance and reaches a peak corresponding to an "optimal" noise intensity. In this case, the first neuron repeatedly fires spikes and then excitedly stimulates the other two neurons. Thus, the firing behaviors of these neurons almost display tonic firing activities and no quiescent state emerges. While further increasing the noise intensity larger than the "optimal" noise intensity, the SNR decreases quickly, indicating that the performance becomes worse. When the noise intensity D is around 10, the SNR reaches a valley point and then raises gradually again by further increasing the noise intensity D. Intuitively, the maximal SNR should turn larger while increasing values of the coupling strength g. However, this is not the case for the FFL motif with Ornstein-Uhlenbeck colored noise here. As we see, as g increases from 0.1 up to 0.4, no explicit differences happen for the maximal SNR and the SNR curves. We further observe that how the scope of noise intensity for the maximal SNR changes with τ. Evidently, the scope of noise intensity for the maximal SNR enlarges with increasing the correlation time of the noise process τ (see Figures 7(a)–7(d)).

Second, to gain more insights into the dependence of the SNR on D and τ, we calculate the R on D and τ by two different coupling strengths g. Figure 8 features the resulting color-contour plots for increasing values of g from top to bottom. It is evident that there exists an optimal area for $9.0 < D < 9.6$ and $1.0 < \tau = 1.2$ where the SNR is maximal, indicating the existence of noise-induced stochastic resonance and the optimal outreach of the localized activity of the output of neuron 3.

5. Conclusions

In summary, we have considered the stochastic resonance of the FFL neuronal network motif subject to non-Gaussian noise. The FFL neuron motif has been built through the FitzHugh-Nagumo neuron model and the directed synapse couplings. We have mainly focused on the influence of noise on the stochastic resonance. The dependence of the SNR on the strength of noise D, coupling strength g, and correlation time τ has been thoroughly studied. It has been observed that the FFL motif can obtain high values of SNR at optimal noise intensities and proper coupling strengths.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61104155), the Fundamental Research Funds for the Central Universities (JUSRP51322B), and the 111 Project (B12018).

References

[1] G. Mato, "Stochastic resonance in neural systems: effect of temporal correlation in the spike trains," *Physical Review E*, vol. 58, no. 1, pp. 876–880, 1998.

[2] O. Kwon and H. T. Moon, "Coherence resonance in small-world networks of excitable cells," *Physics Letters A: General, Atomic and Solid State Physics*, vol. 298, no. 5-6, pp. 319–324, 2002.

[3] M. Ozer, M. Perc, and M. Uzuntarla, "Stochastic resonance on Newman-Watts networks of Hodgkin-Huxley neurons with local periodic driving," *Physics Letters A*, vol. 373, no. 10, pp. 964–968, 2009.

[4] J. E. Levin and J. P. Miller, "Broadband neural encoding in the cricket cercal sensory system enhanced by stochastic resonance," *Nature*, vol. 380, no. 6570, pp. 165–168, 1996.

[5] Y. Gong, M. Wang, Z. Hou, and H. Xin, "Optimal spike coherence and synchronization on complex hodgkin-huxley neuron networks," *ChemPhysChem*, vol. 6, no. 6, pp. 1042–1047, 2005.

[6] T. P. Vogels and L. F. Abbott, "Signal propagation and logic gating in networks of integrate-and-fire neurons," *The Journal of Neuroscience*, vol. 25, no. 46, pp. 10786–10795, 2005.

[7] S. Wang, W. Wang, and F. Liu, "Propagation of firing rate in a feed-forward neuronal Network," *Physical Review Letters*, vol. 96, no. 1, Article ID 018103, 2006.

[8] E. M. Izhikevich, *Dynamical Systems in Neuroscience: the Geometry of Excitability And Bursting*, MIT Press, Cambridge, Mass, USA, 2007.

[9] O. V. Popovych, S. Yanchuk, and P. A. Tass, "Delay- and coupling-induced firing patterns in oscillatory neural loops," *Physical Review Letters*, vol. 107, no. 22, Article ID 228102, 2011.

[10] D. Q. Guo and C. G. Li, "Stochastic and coherence resonance in feed-forward-loop neuronal network motifs," *Physical Review E*, vol. 79, no. 5, 2009.

[11] S. Wuchty, Z. Oltvai, and A. Barabási, "Evolutionary conservation of motif constituents in the yeast protein interaction network," *Nature Genetics*, vol. 35, no. 2, pp. 176–179, 2003.

[12] O. Sporns and R. Kötter, "Motifs in brain networks," *PLoS Biology*, vol. 2, no. 11, pp. 1910–1918, 2004.

[13] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," *PLoS Biology*, vol. 3, no. 3, pp. 507–519, 2005.

[14] U. Alon, *An Introduction to System Biology: Design Principles of Biological Circuits*, Chapman & Hall, London, UK, 2006.

[15] G. Reddy and C. Murthy, *Coherence Resonance in the FitzHugh-Nagumo System*, University of California, Physics 210B: Nonequilibrium Statistical Physics, Student Project, San Diego, Calif, USA, 2013.

[16] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical Journal*, vol. 1, no. 6, pp. 445–466, 1961.

[17] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proceedings of the IRE*, vol. 50, no. 10, pp. 2061–2070, 1962.

[18] J. Keener and J. Sneyd, *Mathematical Physiology*, Springer, New York, NY, USA, 1998.

[19] B. C. Bag, K. G. Petrosyan, and C. Hu, "Influence of noise on the synchronization of the stochastic Kuramoto model," *Physical Review E*, vol. 76, no. 5, Article ID 056210, 2007.

[20] B. C. Bag and C. K. Hu, "Escape through an unstable limit cycle: resonant activation," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 6, Article ID 061107, 2006.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 561382, 12 pages
http://dx.doi.org/10.1155/2014/561382

Research Article

Nonzero Sum Differential Game of Mean-Field BSDEs with Jumps under Partial Information

Xiaolan Chen[1] and Qingfeng Zhu[1,2]

[1] *School of Mathematics and Quantitative Economics, Shandong University of Finance and Economics, Jinan 250014, China*
[2] *Institute for Financial Studies and School of Mathematics, Shandong University, Jinan 250100, China*

Correspondence should be addressed to Qingfeng Zhu; zhuqf508@sohu.com

Received 21 March 2014; Revised 11 May 2014; Accepted 5 June 2014; Published 3 July 2014

Academic Editor: Guangchen Wang

This paper is concerned with a kind of nonzero sum differential game of mean-field backward stochastic differential equations with jump (MF-BSDEJ), in which the coefficient contains not only the state process but also its marginal distribution. Moreover, the cost functional is also of mean-field type. It is required that the control is adapted to a subfiltration of the filtration generated by the underlying Brownian motion and Poisson random measure. We establish a necessary condition in the form of maximum principle with Pontryagin's type for open-loop Nash equilibrium point of this type of partial information game and then give a verification theorem which is a sufficient condition for Nash equilibrium point. The theoretical results are applied to study a partial information linear-quadratic (LQ) game.

1. Introduction

Game theory had been an active area of research and a useful tool in many applications, particularly in biology and economics. The study of differential games was originally stated by Isaacs [1] and then summed up and developed by Basar and Olsder [2], Yeung and Petrosyan [3], and so forth. Berkovitz [4], Fleming [5], Elliott and Kalton [6], and Friedman [7] established the foundations for zero sum differential games and Varaiya [8] and Elliott and Davis [9] for stochastic differential games. Next, the advances in stochastic differential games continue to appear over a large number of fields. Please refer to Hamadène [10], Hamadène et al. [11], Altman [12], Wu and Yu [13], Yu and Ji [14], and Wang and Yu [15] for more information.

For the partial information two-person zero sum (or nonzero sum) stochastic differential games, the objective is to find a saddle point (or equilibrium point) for which the controller has less information than the complete information filtration $\{\mathscr{F}_t\}_{t\geq 0}$. Recently, An and Øksendal [16, 17] and An et al. [18] established a maximum principle for partial information differential games of stochastic differential equations with jump (SDEJ). Wang and Yu [19] developed some results for optimal control of BSDEs and established a maximum principle for partial information differential games of backward stochastic differential equations (BSDEs). They established a necessary condition in the form of maximum principle with Pontryagin's type for open-loop Nash equilibrium point of this type of partial information game and gave a verification theorem which is a sufficient condition for Nash equilibrium point. Meng and Tang [20] and Hui and Xiao [21] established a maximum principle for differential games of forward-backward SDE under partial information. Øksendal and Sulem [22] established a general maximum principle for forward-backward stochastic differential games for Itô-Lévy processes with partial information and applied the theory to optimal portfolio and consumption problems under model uncertainty, in markets modeled by Itô-Lévy processes.

To the best of our knowledge, there are few results about the partial information differential games of the discontinuous mean-field backward stochastic system. In the present paper we will research this topic. This paper is concerned with a new kind of nonzero sum differential game of mean-field backward stochastic differential equations with jump (MF-BSDEJ) under partial information. It is required that the control is adapted to a subfiltration of the filtration generated by the underlying Brownian motion and Poisson random measure. We establish a necessary condition in the form of maximum principle with Pontryagin's type for open-loop Nash equilibrium point of this type of partial information game and then give a verification theorem which is a sufficient condition for Nash equilibrium point. We note that the state system and the cost function in [22] are not mean-field, and the game systems in [15, 19] are BSDEs. The theoretical results are applied to study a partial information linear-quadratic (LQ) game.

The rest of this paper is organized as follows. In Section 2, we state our partial information differential game of MF-BSDEJ and the main assumptions. Section 3 is devoted to the necessary optimality conditions. In Section 4, we obtain the sufficient maximum principle of differential game of MF-BSDEJ under partial information. In Section 5, we give a partial information linear-quadratic (LQ) game as example to show the applications of our theoretical results.

2. Statement of the Problems

Let $(\Omega, \mathcal{F}, P)$ be a completed probability space. We suppose that the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is generated by the following two mutually independent processes: a d-dimensional standard Brownian motion $\{B(t)\}_{t\geq 0}$ and a Poisson random measure N on $\mathbb{R}_+ \times E$, where $E \subset \mathbb{R}^l$ is a nonempty open set equipped with its Borel field $\mathscr{B}(E)$, with compensator $\widehat{N}(de\,dt) = \pi(de)dt$, such that $\widetilde{N}(A \times [0,t]) = (N - \widehat{N})(A \times [0,t])_{t\geq 0}$ is a martingale for $\forall A \in \mathscr{B}(E)$ satisfying $\pi(A) < \infty$. π is assumed to be a σ-finite measure on $(E, \mathscr{B}(E))$ and called the characteristic measure. Let $\mathcal{N}$ denote the class of P-null elements of $\mathcal{F}$. For each $t \in [0,T]$, we define $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^N$, where for any process $\{\eta(t)\}$, $\mathcal{F}_{s,t}^\eta = \sigma\{\eta(r) - \eta(s); s \leq r \leq t\} \vee \mathcal{N}$, $\mathcal{F}_t^\eta = \mathcal{F}_{0,t}^\eta$.

Let $(\Omega^2, \mathcal{F}^2, P^2) = (\Omega\times\Omega, \mathcal{F}\otimes\mathcal{F}, P\otimes P)$ be the completion of the product probability space of the above $(\Omega, \mathcal{F}, P)$ with itself, where we define $\mathcal{F}_t^2 = \mathcal{F}_t \otimes \mathcal{F}_t$ with $t \in [0,T]$ and $\mathcal{F}_t \otimes \mathcal{F}_t$ being the completion of $\mathcal{F}_t \times \mathcal{F}_t$. It is worthy of noting that any random variable $\xi = \xi(\omega)$ defined on Ω can be extended naturally to Ω^2 as $\xi'(\omega, \omega') = \xi(\omega)$ with $(\omega, \omega') \in \Omega^2$. For $H = \mathbb{R}^n$ and so on, let $L^1(\Omega^2, \mathcal{F}^2, P^2; H)$ be the set of random variable $\xi : \Omega^2 \to H$ which is $\mathcal{F}^2$-measurable such that $\mathbb{E}^2|\xi| \equiv \int_{\Omega^2} |\xi(\omega', \omega)| P(d\omega')P(d\omega) < \infty$. For any $\eta \in L^1(\Omega^2, \mathcal{F}^2, P^2; H)$, we denote

$$\mathbb{E}'\eta(\omega, \cdot) \doteq \int_\Omega \eta(\omega, \omega') P(d\omega'). \tag{1}$$

Particularly, for example, if $\eta_1(\omega, \omega') = \eta_1(\omega')$, then

$$\mathbb{E}'\eta_1 = \int_\Omega \eta_1(\omega') P(d\omega') = \mathbb{E}\eta_1. \tag{2}$$

We introduce the following notations:

$$\begin{aligned}
&M^2(0,T;\mathbb{R}^n) \\
&\quad = \Big\{ v(t,\omega) : v(t,\omega) \text{ is an } \mathbb{R}^n\text{-valued}, \\
&\qquad \mathcal{F}_t\text{-measurable process} \\
&\qquad \text{such that } \mathbb{E}\int_0^T |v(t,\omega)|^2 dt < \infty \Big\}, \\
&F_N^2(0,T;\mathbb{R}^n) \\
&\quad = \Big\{ r(t,e,\omega) : r(t,e,\omega) \text{ is an } \mathbb{R}^n\text{-valued}, \\
&\qquad \mathcal{F}_t\text{-measurable process} \\
&\qquad \text{such that } \mathbb{E}\int_0^T \int_{\mathbf{E}} |r(t,e,\omega)|^2 \pi(de)\, dt < \infty \Big\}, \\
&L^2_{\pi(\cdot)}(\mathbb{R}^n) \\
&\quad = \Big\{ r(e) : r(e) \text{ is an } \mathbb{R}^n\text{-valued}, \\
&\qquad \mathscr{B}(\mathbf{E})\text{-measurable function} \\
&\qquad \text{such that } \|r\| = \Big(\int_{\mathbf{E}} |r(e)|^2 \pi(de) \Big)^{1/2} < \infty \Big\}, \\
&L^2(\Omega, \mathcal{F}_T, P; \mathbb{R}^n) \\
&\quad = \Big\{ \xi : \xi \text{ is an } \mathbb{R}^n\text{-valued}, \\
&\qquad \mathcal{F}_T\text{-measurable random variable} \\
&\qquad \text{such that } \mathbb{E}|\xi|^2 < \infty \Big\}.
\end{aligned} \tag{3}$$

We use the usual inner product $\langle \cdot, \cdot \rangle$ and Euclidean norm $|\cdot|$ in $\mathbb{R}^n$, $\mathbb{R}^{n\times d}$, and $\mathbb{R}^{n\times l}$. The notation "T" appearing in the superscripts denotes the transpose of a matrix. All the equalities and inequalities mentioned in this paper are in the sense of $dt \times dP$ almost surely on $[0,T] \times \Omega$.

This work is interested in a class of partial information nonzero sum differential games of MF-BSDEJ, which is inspirited by some interesting financial phenomena. For simplicity, we only consider the case of two players, which is similar for n players. Let us now give a detailed formulation of the problem. Consider the following MF-BSDEJ:

$$\begin{aligned}
-dy^v(t) &= \mathbb{E}' f\big(t, y^v(t), z^v(t), r^v(t,\cdot), (y^v(t))', \\
&\qquad (z^v(t))', (r^v(t,\cdot))', v(t)\big) dt \\
&\quad - z^v(t)\, dW(t) - \int_{\mathbf{E}} r^v(t,e) \widetilde{N}(de\,dt), \\
y^v(T) &= \xi,
\end{aligned} \tag{4}$$

where $\xi \in L^2(\Omega, \mathcal{F}_T, P; \mathbb{R}^n)$, $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \times \mathbb{R}^{k_1} \times \mathbb{R}^{k_2} \to \mathbb{R}^n$, $v_1(\cdot)$ and $v_2(\cdot)$ are the control processes of Player 1 and Player 2, and $v(\cdot) = (v_1(\cdot), v_2(\cdot))$. We always use the subscript 1 (resp., the subscript 2) to characterize the variables corresponding to Player 1 (resp., Player 2). The mean-field backward game system (4) has the meaning that the two players work together to achieve a goal ξ at the terminal time T.

To study our problem, we give some assumptions on $v_1(\cdot)$, $v_2(\cdot)$, and f. Let U_i be a nonempty convex subset of $\mathbb{R}^{k_i}$ $(i = 1,2)$ and $\mathscr{E}_t^i \subseteq \mathscr{F}_t$ $(i = 1,2)$ a given subfiltration which represents the information available to Player i at time $t \in [0,T]$, respectively. Now we introduce the admissible control set

$$\mathscr{U}_i = \Big\{ v_i : [0,T]\times\Omega \longrightarrow U_i \mid v_i \text{ is } \mathscr{E}_t^i\text{-adapted}, \\ \mathbb{E}\int_0^T |v_i(t)|^2 dt < \infty \Big\}, \quad i = 1,2. \tag{5}$$

Each element of $\mathscr{U}_i$ is called an open-loop admissible control for Player i $(i = 1,2)$. And $\mathscr{U}_1 \times \mathscr{U}_2$ is called the set of open-loop admissible controls for the players.

We assume that

(H1) f is continuously differentiable with respect to (y,z,r,y',z',r',v_1,v_2). Moreover, the norm of f_y, f_z, f_r, $f_{y'}$, $f_{z'}$, $f_{r'}$, f_{v_1}, f_{v_2} is bounded by $c > 0$.

Now, if both $v_1(\cdot)$ and $v_2(\cdot)$ are admissible controls and assumption (H1) holds, then MF-BSDEJ (4) admits a unique solution $(y^{v_1,v_2}(\cdot), z^{v_1,v_2}(\cdot), r^{v_1,v_2}(\cdot,\cdot)) \in M^2(0,T;\mathbb{R}^n)\times M^2(0,T;\mathbb{R}^{n\times d})\times F_N^2(0,T;\mathbb{R}^n)$ (see Shen and Siu [23]). Ensuring to achieve the goal ξ, the players have their own benefits, which are described by the following cost functionals:

$$J_i(v(\cdot)) \\ = \mathbb{E}\Big[\int_0^T \mathbb{E}' l_i\big(t, y^v(t), z^v(t), r^v(t,\cdot), (y^v(t))', (z^v(t))', \\ (r^v(t,\cdot))', v(t)\big) dt + \Phi_i(y^v(0))\Big], \tag{6}$$

where $v(\cdot) = (v_1(\cdot), v_2(\cdot))$, $l_i : [0,T]\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n)\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n)\times\mathbb{R}^{k_1}\times\mathbb{R}^{k_2} \to \mathbb{R}$, $\Phi_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1,2$, satisfying the condition

$$\mathbb{E}\Big[\int_0^T \big|\mathbb{E}' l_i\big(t, y^v(t), z^v(t), r^v(t,\cdot), (y^v(t))', \\ (z^v(t))', (r^v(t,\cdot))', v(t)\big)\big| dt \\ + |\Phi_i(y^v(0))|\Big] < \infty, \quad i = 1,2. \tag{7}$$

We also assume that

(H2) l_i is continuously differentiable in (y,z,r,y',z',r',v_1,v_2) and its partial derivatives are continuous in (y,z,r,y',z',r',v_1,v_2) and bounded by $c(1 + |y| + |z| + \|r\| + |y'| + |z'| + \|r'\| + |v_1| + |v_2|)$. Moreover, Φ_i is continuously differentiable and Φ_{iy} is bounded by $c(1 + |y|)$.

Suppose each player hopes to minimize her/his cost functional $J_i(v_1(\cdot), v_2(\cdot))$ by selecting an appropriate admissible control $v_i(\cdot)$ $(i = 1,2)$. Then the problem is to find a pair of admissible controls $(u_1(\cdot), u_2(\cdot)) \in \mathscr{U}_1\times\mathscr{U}_2$ such that

$$J_1(u_1(\cdot), u_2(\cdot)) = \min_{v_1(\cdot)\in\mathscr{U}_1} J_1(v_1(\cdot), u_2(\cdot)), \\ J_2(u_1(\cdot), u_2(\cdot)) = \min_{v_2(\cdot)\in\mathscr{U}_2} J_2(u_1(\cdot), v_2(\cdot)). \tag{8}$$

We call the problem above a backward nonzero sum stochastic differential game, where the word backward means that the game system is described by a MF-BSDEJ. For simplicity, we denote it by Problem BNZ. If we can find an admissible control $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ satisfying (8), then we call it an equilibrium point of Problem BNZ and denote the corresponding state trajectory by $(y(\cdot), z(\cdot), r(\cdot,\cdot)) = (y^u(\cdot), z^u(\cdot), r^u(\cdot,\cdot))$.

3. A Partial Information Necessary Maximum Principle

For the convex admissible control set, the classical way to derive necessary optimality conditions is to use the convex perturbation method. Let $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ be an equilibrium point of Problem BNZ and let $(y(\cdot), z(\cdot), r(\cdot,\cdot))$ be the corresponding optimal trajectory. Let $(v_1(\cdot), v_2(\cdot))$ be such that $(u_1(\cdot) + v_1(\cdot), u_2(\cdot) + v_2(\cdot)) \in \mathscr{U}_1\times\mathscr{U}_2$. Since $\mathscr{U}_1$ and $\mathscr{U}_2$ are convex, for any $0 \le \rho \le 1$, $(u_1^\rho(\cdot), u_2^\rho(\cdot)) = (u_1(\cdot) + \rho v_1(\cdot), u_1(\cdot) + \rho v_1(\cdot))$ is also in $\mathscr{U}_1\times\mathscr{U}_2$. As illustrated before, we denote by $(y^{u_1^\rho}(\cdot), z^{u_1^\rho}(\cdot), r^{u_1^\rho}(\cdot,\cdot))$ and $(y^{u_2^\rho}(\cdot), z^{u_2^\rho}(\cdot), r^{u_2^\rho}(\cdot,\cdot))$ the corresponding state trajectories of game system (4) along with the controls $(u_1^\rho(\cdot), u_2(\cdot))$ and $(u_1(\cdot), u_2^\rho(\cdot))$.

For convenience, we introduce the notations

$$\begin{aligned}
\varphi(t,\cdot) &= \varphi\big(t, y(t), z(t), r(t,\cdot), y'(t), z'(t), \\ &\qquad r'(t,\cdot), u_1(t), u_2(t)\big), \\
\varphi^v(t,\cdot) &= \varphi\big(t, y(t), z(t), r(t,\cdot), y'(t), z'(t), \\ &\qquad r'(t,\cdot), v_1(t), v_2(t)\big), \\
\varphi^{u_1^\rho}(t,\cdot) &= \varphi\big(t, y(t), z(t), r(t,\cdot), y'(t), z'(t), \\ &\qquad r'(t,\cdot), u_1^\rho(t), u_2(t)\big), \\
\varphi^{u_2^\rho}(t,\cdot) &= \varphi\big(t, y(t), z(t), r(t,\cdot), y'(t), z'(t), \\ &\qquad r'(t,\cdot), u_1(t), u_2^\rho(t)\big),
\end{aligned} \tag{9}$$

where φ denotes one of f and l.

We introduce the variational equations as follows:

$$\begin{aligned}
&-dy_i^1(t)\\
&=\mathbb{E}'\Big[f_y(t,\cdot)\,y_i^1(t)+f_z(t,\cdot)\,z_i^1(t)\\
&\quad+\int_{\mathbf{E}}f_r(t,e)\,r_i^1(t,e)\,\pi(de)\\
&\quad+f_{y'}(t,\cdot)\left(y_i^1(t)\right)'+f_{z'}(t,\cdot)\left(z_i^1(t)\right)'\\
&\quad+\int_{\mathbf{E}}f_{r'}(t,e)\left(r_i^1(t,e)\right)'\pi(de)+f_{v_i}(t,\cdot)\,v_i(t)\Big]\,dt\\
&\quad-z_i^1(t)\,dW(t)-\int_{\mathbf{E}}r_i^1(t,e)\,\widetilde{N}(de\,dt),\\
&y_i^1(T)=0,\quad(i=1,2).
\end{aligned}\tag{10}$$

By (H1), it is easy to know that (10) admits unique adapted solution $(y_i^1(t),z_i^1(t),r_i^1(t,\cdot))\in M^2(0,T;\mathbb{R}^n)\times M^2(0,T;\mathbb{R}^{n\times d})\times F_N^2(0,T;\mathbb{R}^n)$, $i=1,2$.

For $t\in[0,T]$, $\rho>0$, we set

$$\begin{aligned}
\widetilde{y}_i^\rho(t)&=\frac{y^{u_i^\rho}(t)-y(t)}{\rho}-y_i^1(t),\\
\widetilde{z}_i^\rho(t)&=\frac{z^{u_i^\rho}(t)-z(t)}{\rho}-z_i^1(t),\\
\widetilde{r}_i^\rho(t,\cdot)&=\frac{r^{u_i^\rho}(t,\cdot)-r(t,\cdot)}{\rho}-r_i^1(t,\cdot),\quad(i=1,2).
\end{aligned}\tag{11}$$

We have the following.

Lemma 1. *Let assumptions (H1) and (H2) hold. Then, for $i=1,2$,*

$$\lim_{\rho\to0}\sup_{0\le t\le T}\mathbb{E}\left|\widetilde{y}_i^\rho(t)\right|^2=0,\tag{12}$$

$$\begin{aligned}
&\lim_{\rho\to0}\mathbb{E}\int_0^T\left|\widetilde{z}_i^\rho(t)\right|^2dt=0,\\
&\lim_{\rho\to0}\mathbb{E}\int_0^T\left\|\widetilde{r}_i^\rho(t)\right\|^2dt=0.
\end{aligned}\tag{13}$$

Proof. For $i=1$, we have

$$\begin{aligned}
&-d\widetilde{y}_1^\rho(t)\\
&=\Big[\frac{1}{\rho}\mathbb{E}'\left(f^{u_1^\rho}(t,\cdot)-f(t,\cdot)\right)\\
&\quad-\mathbb{E}'\Big(f_y(t,\cdot)\,y_1^1(t)+f_z(t,\cdot)\,z_1^1(t)\\
&\quad+\int_{\mathbf{E}}f_r(t,e)\,r_1^1(t,e)\,\pi(de)+f_{y'}(t,\cdot)\left(y_1^1(t)\right)'\\
&\quad+f_{z'}(t,\cdot)\left(z_1^1(t)\right)'\\
&\quad+\int_{\mathbf{E}}f_{r'}(t,e)\left(r_1^1(t,e)\right)'\pi(de)\\
&\quad+f_{v_1}(t,\cdot)\,v_1(t)\Big)\Big]\,dt\\
&\quad-\widetilde{z}_1^\rho(t)\,dW(t)-\int_{\mathbf{E}}\widetilde{r}_1^\rho(t,e)\,\widetilde{N}(de\,dt),\\
&\widetilde{y}_1^\rho(T)=0,
\end{aligned}\tag{14}$$

or

$$\begin{aligned}
&-d\widetilde{y}_1^\rho(t)\\
&=\mathbb{E}'\Big[A_1^\rho(t,\cdot)\,\widetilde{y}_1^\rho(t)+B_1^\rho(t,\cdot)\,\widetilde{z}_1^\rho(t)\\
&\quad+\int_{\mathbf{E}}C_1^\rho(t,e)\,\widetilde{r}_1^\rho(t,e)\,\pi(de)\\
&\quad+D_1^\rho(t,\cdot)\left(\widetilde{y}_1^\rho(t)\right)'+E_1^\rho(t,\cdot)\left(\widetilde{z}_1^\rho(t)\right)'\\
&\quad+\int_{\mathbf{E}}F_1^\rho(t,e)\left(\widetilde{r}_1^\rho(t,e)\right)'\pi(de)+G_1^\rho(t,\cdot)\Big]\,dt\\
&\quad-\widetilde{z}_1^\rho(t)\,dW(t)-\int_{\mathbf{E}}\widetilde{r}_1^\rho(t,e)\,\widetilde{N}(de\,dt),\\
&\widetilde{y}_1^\rho(T)=0,
\end{aligned}\tag{15}$$

where we denote

$$\begin{aligned}
&(\Theta)\\
&=\Big(t,y(t)+\lambda\rho\left(y_1^1(t)+\widetilde{y}_1^\rho(t)\right),z(t)\\
&\quad+\lambda\rho\left(z_1^1(t)+\widetilde{z}_1^\rho(t)\right),r(t,\cdot)\\
&\quad+\lambda\rho\left(r_1^1(t,\cdot)+\widetilde{r}_1^\rho(t,\cdot)\right),(y(t))'\\
&\quad+\lambda\rho\left(\left(y_1^1(t)\right)'+\left(\widetilde{y}_1^\rho(t)\right)'\right),(z(t))'\\
&\quad+\lambda\rho\left(\left(z_1^1(t)\right)'+\left(\widetilde{z}_1^\rho(t)\right)'\right),(r(t,\cdot))'\\
&\quad+\lambda\rho\left(\left(r_1^1(t,\cdot)\right)'+\left(\widetilde{r}_1^\rho(t,\cdot)\right)'\right),u_1(t)\\
&\quad+\lambda\rho v_1(t),u_2(t)\Big),
\end{aligned}$$

$$A_1^\rho(t,\cdot)=\int_0^1f_y(\Theta)\,d\lambda,\qquad B_1^\rho(t,\cdot)=\int_0^1f_z(\Theta)\,d\lambda,$$

$$C_1^\rho(t,\cdot)=\int_0^1f_r(\Theta)\,d\lambda,\qquad D_1^\rho(t,\cdot)=\int_0^1f_{y'}(\Theta)\,d\lambda,$$

$$E_1^\rho(t,\cdot)=\int_0^1f_{z'}(\Theta)\,d\lambda,\qquad F_1^\rho(t,\cdot)=\int_0^1f_{r'}(\Theta)\,d\lambda,$$

$$\begin{aligned}
&G_1^\rho(t,\cdot)\\
&=\int_0^1\left(f_{v_1}(\Theta)-f_{v_1}(t,\cdot)\right)v_1(t)\,d\lambda
\end{aligned}$$

$$\begin{aligned}&+\left[A_1^\rho(t,\cdot)-f_y(t,\cdot)\right]y_1^1(t)+\left[B_1^\rho(t,\cdot)-f_z(t,\cdot)\right]z_1^1(t)\\&+\left[C_1^\rho(t,\cdot)-f_r(t,\cdot)\right]r_1^1(t,\cdot)\\&+\left[D_1^\rho(t,\cdot)-f_{y'}(t,\cdot)\right]\left(y_1^1(t)\right)'\\&+\left[E_1^\rho(t,\cdot)-f_{z'}(t,\cdot)\right]\left(z_1^1(t)\right)'\\&+\left[F_1^\rho(t,\cdot)-f_{r'}(t,\cdot)\right]\left(r_1^1(t,\cdot)\right)'.\end{aligned}\tag{16}$$

Applying Itô's formula to $|\tilde{y}_1^\rho(t)|^2$ on $[t,T]$, by virtue of (H1), we get

$$\begin{aligned}&\mathbb{E}\left|\tilde{y}_1^\rho(t)\right|^2+\mathbb{E}\int_t^T\left(\left|\tilde{z}_1^\rho(s)\right|^2+\left\|\tilde{r}_1^\rho(s)\right\|^2\right)ds\\&=2\mathbb{E}\mathbb{E}'\int_t^T\int_{\mathbf{E}}\left|\left\langle\tilde{y}_1^\rho(s),A_1^\rho(s,\cdot)\tilde{y}_1^\rho(s)+B_1^\rho(s,\cdot)\tilde{z}_1^\rho(s)\right.\right.\\&\quad+\int_{\mathbf{E}}C_1^\rho(s,e)\tilde{r}_1^\rho(s,e)\pi(de)\\&\quad+D_1^\rho(s,\cdot)\left(\tilde{y}_1^\rho(s)\right)'+E_1^\rho(s,\cdot)\left(\tilde{z}_1^\rho(s)\right)'\\&\quad+\int_{\mathbf{E}}F_1^\rho(s,e)\left(\tilde{r}_1^\rho(s,e)\right)'\pi(de)\\&\quad\left.\left.+G_1^\rho(s,\cdot)\right\rangle\right|\pi(de)\;ds\\&\le C_0\mathbb{E}\int_t^T\left|\tilde{y}_1^\rho(s)\right|^2ds\\&\quad+\frac{1}{2}\mathbb{E}\int_t^T\left(\left|\tilde{z}_1^\rho(s)\right|^2+\left\|\tilde{r}_1^\rho(s)\right\|^2\right)ds\\&\quad+C_1\alpha\left(\mathbb{E}\int_t^T\left|G_1^\rho(s)\right|^2\right)ds.\end{aligned}\tag{17}$$

By Gronwall's inequality, we easily obtain the desired result. Similarly, we can show that the conclusion holds for $i=2$. □

Since $(u_1(\cdot),u_2(\cdot))$ is an equilibrium point of Problem BNZ, then

$$\rho^{-1}\left[J_1\left(u_1^\rho(\cdot),u_2(\cdot)\right)-J_1\left(u_1(\cdot),u_2(\cdot)\right)\right]\ge 0,\tag{18}$$

$$\rho^{-1}\left[J_2\left(u_1(\cdot),u_2^\rho(\cdot)\right)-J_2\left(u_1(\cdot),u_2(\cdot)\right)\right]\ge 0.\tag{19}$$

From this and Lemma 1, we have the following variational inequality.

Lemma 2. *Let assumption (H1) hold. Then,*

$$\begin{aligned}&\mathbb{E}\int_0^T\mathbb{E}'\left[l_{iy}(t,\cdot)y_i^1(t)+l_{iz}(t,\cdot)z_i^1(t)\right.\\&\quad+\int_{\mathbf{E}}l_{ir}(t)r_i^1(t,e)\pi(de)+l_{iy'}(t,\cdot)\left(y_i^1(t)\right)'\\&\quad+l_{iz'}(t,\cdot)\left(z_i^1(t)\right)'+\int_{\mathbf{E}}l_{ir'}(t)\left(r_i^1(t,e)\right)'\pi(de)\\&\quad\left.+l_{iv_i}(t,\cdot)v_i(t)\right]dt\\&+\mathbb{E}\left[\Phi_{iy}\left(y(0)\right)y_i^1(0)\right]\ge 0,\quad(i=1,2).\end{aligned}\tag{20}$$

Proof. For $i=1$, from (12), we derive

$$\begin{aligned}&\rho^{-1}\left[\Phi_1\left(y^{u_1^\rho}(0)\right)-\Phi_1\left(y(0)\right)\right]\\&=\rho^{-1}\mathbb{E}\int_0^1\Phi_{1y}\left(y(0)+\lambda\left(y^{u_1^\rho}(0)-y(0)\right)\right)\\&\qquad\times\left(y^{u_1^\rho}(0)-y(0)\right)d\lambda\\&\longrightarrow\mathbb{E}\left[\Phi_{1y}\left(y(0)\right)y_1^1(0)\right],\quad\rho\longrightarrow 0.\end{aligned}\tag{21}$$

Similarly, we have

$$\begin{aligned}&\rho^{-1}\left\{\mathbb{E}\int_0^T\mathbb{E}'\left[l_1^{u_1^\rho}(t,\cdot)-l_1(t,\cdot)\right]dt\right\}\\&\longrightarrow\mathbb{E}\int_0^T\mathbb{E}'\left[l_{1y}(t,\cdot)y_1^1(t)+l_{1z}(t,\cdot)z_1^1(t)\right.\\&\quad+\int_{\mathbf{E}}l_{1r}(t)r_1^1(t,e)\pi(de)+l_{1y'}(t,\cdot)\left(y_1^1(t)\right)'\\&\quad+l_{1z'}(t,\cdot)\left(z_1^1(t)\right)'\\&\quad+\int_{\mathbf{E}}l_{1r'}(t)\left(r_1^1(t,e)\right)'\pi(de)\\&\quad\left.+l_{1v_1}(t,\cdot)v_1(t)\right]dt,\\&\qquad\rho\longrightarrow 0.\end{aligned}\tag{22}$$

Let $\rho\to 0$ in (18); then, it follows that, for $i=1$, (20) holds. Similarly, we can show that the conclusion holds for $i=2$. □

We define the Hamiltonian function $H_i:[0,T]\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n)\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n)\times\mathbb{R}^{k_1}\times\mathbb{R}^{k_2}\times\mathbb{R}^n\to\mathbb{R}$, $i=1,2$, as follows:

$$\begin{aligned}&H_i\left(t,y,z,r,y',z',r',v_1,v_2,p_i\right)\\&=-\left\langle f\left(t,y,z,r(\cdot),y',z',r'(\cdot),v_1,v_2\right),p_i\right\rangle\\&\quad+l_i\left(t,y,z,r(\cdot),y',z',r'(\cdot),v_1,v_2\right),\\&\qquad i=1,2.\end{aligned}\tag{23}$$

Let

$$H_i(t,\cdot) = H_i\left(t, y, z, r, y', z', r', u_1, u_2, p_i\right),$$
$$H_i^{v_1,v_2}(t,\cdot) = H_i\left(t, y, z, r, y', z', r', v_1, v_2, p_i\right), \quad i = 1,2. \tag{24}$$

We introduce the following adjoint equation:

$$\begin{aligned} dp_i^{v_1,v_2}(t) &= -\mathbb{E}'\left[H_{iy}^{v_1,v_2}(t,\cdot) + H_{iy'}^{v_1,v_2}(t,\cdot)\right]dt \\ &\quad - \mathbb{E}'\left[H_{iz}^{v_1,v_2}(t,\cdot) + H_{iz'}^{v_1,v_2}(t,\cdot)\right]dW(t) \\ &\quad - \int_{\mathbf{E}} \mathbb{E}'\left[H_{ir}^{v_1,v_2}(t,e) + H_{ir'}^{v_1,v_2}(t,e)\right]\widetilde{N}(de\,dt), \\ p_i^{v_1,v_2}(0) &= -\Phi_{iy}\left(y(0)\right), \quad (i = 1,2). \end{aligned} \tag{25}$$

Starting from the variational inequality (20), we can now state necessary optimality conditions.

Theorem 3 (partial information necessary maximum principle). *Suppose (H1) and (H2) hold. Suppose $(u_1(\cdot), u_2(\cdot))$ is an equilibrium point of Problem BNZ and $(y(\cdot), z(\cdot), r(\cdot,\cdot))$ is the corresponding state trajectory. Then one has that*

$$\begin{aligned} &\mathbb{E}\left[\left\langle H_{1v_1}(t,\cdot), v_1 - u_1(t)\right\rangle \mid \mathscr{E}_t^1\right] \ge 0 \\ &\mathbb{E}\left[\left\langle H_{2v_2}(t,\cdot), v_2 - u_2(t)\right\rangle \mid \mathscr{E}_t^2\right] \ge 0, \end{aligned} \tag{26}$$

hold for any $(v_1, v_2) \in U_1 \times U_2$, a.e., a.s., where $p_i(\cdot)$ $(i = 1,2)$ is the solution of the adjoint equation (25).

Proof. For $i = 1$, applying Itô's formula to $\langle y_1^1(t), p_1(t)\rangle$, we obtain

$$\begin{aligned} &\mathbb{E}\int_0^T \mathbb{E}'\left[l_{1y}(t,\cdot)\, y_1^1(t) + l_{1z}(t,\cdot)\, z_1^1(t)\right. \\ &\qquad + \int_{\mathbf{E}} l_{1r}(t,e)\, r_1^1(t,e)\,\pi(de) + l_{1y'}(t,\cdot)\left(y_1^1(t)\right)' \\ &\qquad + l_{1z'}(t,\cdot)\left(z_1^1(t)\right)' + \int_{\mathbf{E}} l_{1r'}(t,e)\left(r_1^1(t,e)\right)'\pi(de) \\ &\qquad \left. + l_{1v_1}(t)\, v_1(t)\right]dt + \mathbb{E}\left[\Phi_{1y}\left(y(0)\right) y_1^1(0)\right] \\ &= \mathbb{E}\left\langle -f_{v_1}^T(t)\, p_1(t) + l_{1v_1}(t), v_1(t)\right\rangle dt. \end{aligned} \tag{27}$$

From Lemma 2, it follows that we have

$$\mathbb{E}\int_0^T \left\langle H_{1v_1}(t,\cdot), v_1(t)\right\rangle dt \ge 0. \tag{28}$$

Because $v_1(t)$ satisfies $u_1(t) + v_1(t) \in \mathscr{U}_1$, we have

$$\mathbb{E}\int_0^T \left\langle H_{1v_1}(t,\cdot), v_1 - u_1(t)\right\rangle dt \ge 0, \quad \forall v_1 \in U_1. \tag{29}$$

This implies that

$$\mathbb{E}\left\langle H_{1v_1}(t,\cdot), v_1 - u_1(t)\right\rangle \ge 0, \quad \forall v_1 \in U_1. \tag{30}$$

Now, let $v_1(t) \in U_1$ be a deterministic element and let F be an arbitrary element of the σ-algebra $\mathscr{E}_t^1$. And set

$$w_1(t) = v_1(t)\,\mathbf{1}_F + u_1(t)\,\mathbf{1}_{\Omega - F}. \tag{31}$$

It is obvious that w_1 is an admissible control.

Applying the above inequality with w_1, we get

$$\mathbb{E}\left[\mathbf{1}_F\left\langle H_{1v_1}(t,\cdot), v_1 - u_1(t)\right\rangle\right] \ge 0, \quad \forall F \in \mathscr{E}_t^1, \tag{32}$$

which implies that

$$\mathbb{E}\left[\left\langle H_{1v_1}(t,\cdot), v_1 - u_1(t)\right\rangle \mid \mathscr{E}_t^1\right] \ge 0, \quad \forall v_1 \in U_1, \text{ a.e., a.s.} \tag{33}$$

Proceeding in the same way as the above proof, we can show that the other inequality holds for any $v_2 \in U_2$. Then the proof is completed. □

4. A Partial Information Sufficient Maximum Principle

In this section, we investigate a sufficient maximum principle for Problem BNZ. Let $(y(t), z(t), r(t,\cdot), u_1(t), u_2(t))$ be a quintuple satisfying (4) and suppose that there exists a solution $p_i(t)$ of the corresponding adjoint forward SDE (25). We assume that

(H3) for $i = 1,2$, for all $t \in [0,T]$, $H_i(t, y, z, r, y', z', r', v_1, v_2, p_i)$ is convex in $(y, z, r, y', z', r', v_1, v_2)$, and $\Phi_i(y)$ is convex in y.

Let

$$\begin{aligned} H_i(t,\cdot) &= H_i\left(t, y(t), z(t), r(t,\cdot), y'(t), z'(t),\right. \\ &\qquad \left. r'(t,\cdot), u_1(t), u_2(t), p_i(t)\right), \\ H_i^{v_1}(t,\cdot) &= H_i\left(t, y(t), z(t), r(t,\cdot), y'(t), z'(t),\right. \\ &\qquad \left. r'(t,\cdot), v_1(t), u_2(t), p_i(t)\right), \\ H_i^{v_2}(t,\cdot) &= H_i\left(t, y(t), z(t), r(t,\cdot), y'(t), z'(t),\right. \\ &\qquad \left. r'(t,\cdot), u_1(t), v_2(t), p_i(t)\right), \quad i = 1,2, \\ \varphi(t,\cdot) &= \varphi\left(t, y(t), z(t), r(t,\cdot), y'(t), z'(t),\right. \\ &\qquad \left. r'(t,\cdot), u_1(t), u_2(t)\right), \\ \varphi^{v_1}(t,\cdot) &= \varphi\left(t, y(t), z(t), r(t,\cdot), y'(t), z'(t),\right. \\ &\qquad \left. r'(t,\cdot), v_1(t), u_2(t)\right), \\ \varphi^{v_2}(t,\cdot) &= \varphi\left(t, y(t), z(t), r(t,\cdot), y'(t), z'(t),\right. \\ &\qquad \left. r'(t,\cdot), u_1(t), v_2(t)\right), \end{aligned} \tag{34}$$

where φ denotes one of f and l.

Theorem 4 (partial information sufficient maximum principle). *Assume that (H1)–(H3) are satisfied. Moreover, the following partial information maximum conditions hold:*

$$\mathbb{E}\left[H_1(t,\cdot)\mid\mathscr{E}_t^1\right]=\min_{v_1\in\mathscr{U}_1}\mathbb{E}\left[H_1^{v_1}(t,\cdot)\mid\mathscr{E}_t^1\right], \quad (35)$$

$$\mathbb{E}\left[H_2(t,\cdot)\mid\mathscr{E}_t^2\right]=\min_{v_2\in\mathscr{U}_2}\mathbb{E}\left[H_2^{v_2}(t,\cdot)\mid\mathscr{E}_t^2\right]. \quad (36)$$

Then $(u_1(\cdot),u_2(\cdot))$ is an equilibrium point of Problem BNZ.

Proof. For any $v_1(\cdot)\in\mathscr{U}_1$, we consider

$$J_1(v_1(\cdot),u_2(\cdot))-J_1(u_1(\cdot),u_2(\cdot))=\mathbf{I}_1+\mathbf{I}_2, \quad (37)$$

where

$$\begin{aligned}\mathbf{I}_1&=\mathbb{E}\int_0^T\mathbb{E}'\left[l_1^{v_1}(t,\cdot)-l_1(t,\cdot)\right]dt,\\ \mathbf{I}_2&=\mathbb{E}\left[\Phi_1(y^{v_1}(0))-\Phi_1(y(0))\right].\end{aligned} \quad (38)$$

Now applying Itô's formula to $\langle p_1(t),y^{v_1}(t)-y(t)\rangle$ on $[0,T]$, we get

$$\begin{aligned}&\mathbb{E}\left\langle\Phi_{1y}(y(0)),y^{v_1}(0)-y(0)\right\rangle\\ &=\mathbb{E}\int_0^T\mathbb{E}'\Big[\left\langle y^{v_1}(t)-y(t),-H_{1y}(t,\cdot)\right\rangle\\ &\qquad+\left\langle(y^{v_1}(t))'-y'(t),-H_{1y'}(t,\cdot)\right\rangle\Big]dt\\ &\quad+\mathbb{E}\int_0^T\mathbb{E}'\Big[\langle z^{v_1}(t)-z(t),-H_{1z}(t,\cdot)\rangle\\ &\qquad+\left\langle(z^{v_1}(t))'-z'(t),-H_{1z'}(t,\cdot)\right\rangle\Big]dt\\ &\quad+\mathbb{E}\int_0^T\int_{\mathbf{E}}\mathbb{E}'\Big[\langle r^{v_1}(t,e)-r(t,e),-H_{1r}(t,e)\rangle\\ &\qquad+\Big\langle(r^{v_1}(t,e))'-(r(t,e))',\\ &\qquad-H_{1r'}(t,e)\Big\rangle\Big]\pi(de)\,dt\\ &\quad-\mathbb{E}\int_0^T\mathbb{E}'\left[\langle p_1(t),f^{v_1}(t,\cdot)-f(t,\cdot)\rangle\right]dt.\end{aligned} \quad (39)$$

Moreover, by virtue of (39) and convexity of Φ_1, it instantly follows that

$$\mathbf{I}_2\geq\mathbb{E}\left\langle\Phi_{1y}(y(0)),y^{v_1}(0)-y(0)\right\rangle=-\Xi_1+\Xi_2, \quad (40)$$

where

$$\begin{aligned}\Xi_1&=\mathbb{E}\int_0^T\mathbb{E}'\Big[\left\langle y^{v_1}(t)-y(t),H_{1y}(t,\cdot)\right\rangle\\ &\qquad+\left\langle(y^{v_1}(t))'-y'(t),H_{1y'}(t,\cdot)\right\rangle\Big]dt\\ &\quad+\mathbb{E}\int_0^T\mathbb{E}'\Big[\langle z^{v_1}(t)-z(t),H_{1z}(t,\cdot)\rangle\\ &\qquad+\left\langle(z^{v_1}(t))'-z'(t),H_{1z'}(t,\cdot)\right\rangle\Big]dt\\ &\quad+\mathbb{E}\int_0^T\int_{\mathbf{E}}\mathbb{E}'\Big[\langle r^{v_1}(t,e)-r(t,e),H_{1r}(t,e)\rangle\\ &\qquad+\left\langle(r^{v_1}(t,e))'-(r(t,e))',H_{1r'}(t,e)\right\rangle\Big]\\ &\qquad\times\pi(de)\,dt,\\ \Xi_2&=-\mathbb{E}\int_0^T\mathbb{E}'\left[\langle p_1(t),f^{v_1}(t,\cdot)-f(t,\cdot)\rangle\right]dt.\end{aligned} \quad (41)$$

Noting the definition of H_1 and $\mathbf{I}_1$, we have

$$\begin{aligned}\mathbf{I}_1&=\mathbb{E}\int_0^T\mathbb{E}'\left[H_1^{v_1}(t,\cdot)-H_1(t,\cdot)\right]dt\\ &\quad+\mathbb{E}\int_0^T\mathbb{E}'\left[\langle p_1(t),f^{v_1}(t,\cdot)-f(t,\cdot)\rangle\right]dt\\ &=\Xi_3-\Xi_2,\end{aligned} \quad (42)$$

where

$$\Xi_3=\mathbb{E}\int_0^T\mathbb{E}'\left[H_1^{v_1}(t,\cdot)-H_1(t,\cdot)\right]dt. \quad (43)$$

Using convexity of $H_1(t,y,z,r,y',z',r',v_1,v_2,p_1)$ with respect to (y,z,r,y',z',r',v_1,v_2), we obtain

$$\begin{aligned}&H_1^{v_1}(t,\cdot)-H_1(t,\cdot)\\ &\quad\geq H_{1y}(t)(y^{v_1}(t)-y(t))+H_{1z}(t,\cdot)(z^{v_1}(t)-z(t))\\ &\qquad+\int_{\mathbf{E}}H_{1r}(t,e)(r^{v_1}(t,e)-r(t,e))\pi(de)\\ &\qquad+H_{1y'}(t,\cdot)\left((y^{v_1}(t))'-y'(t)\right)\\ &\qquad+H_{1z'}(t,\cdot)\left((z^{v_1}(t))'-z'(t)\right)\\ &\qquad+\int_{\mathbf{E}}H_{1r'}(t,e)\left((r^{v_1}(t,e))'-r(t,e)\right)\pi(de)\\ &\qquad+H_{1u_1}(t,\cdot)(v_1(t)-u_1(t)).\end{aligned} \quad (44)$$

Since $v_1\to\mathbb{E}[H_1^{v_1}(t,\cdot)\mid\mathscr{E}_t^1]$, $v_1\in\mathscr{U}_1$, is minimal for $u_1(t)$ and $v_1(t)$ and $u_1(t)$ are $\mathscr{E}_t^1$-measurable, we get

$$\begin{aligned}&\mathbb{E}\left[H_{1u_1}(t,\cdot)\mid\mathscr{E}_t^1\right](v_1(t)-u_1(t))\\ &\quad=\mathbb{E}\left[H_{1u_1}(t,\cdot)(v_1(t)-u_1(t))\mid\mathscr{E}_t^1\right]\geq 0.\end{aligned} \quad (45)$$

Hence combining (43), (44), and (45), we obtain

$$\begin{aligned}\Xi_3 \\ &\geq \mathbb{E}\int_0^T \mathbb{E}'\left[\left\langle y^{v_1}(t,\cdot)-y(t),-H_{1y}(t,\cdot)\right\rangle\right. \\ &\qquad\left.+\left\langle (y^{v_1}(t))'-y'(t),-H_{1y'}(t,\cdot)\right\rangle\right]dt \\ &\quad+\mathbb{E}\int_0^T \mathbb{E}'\left[\left\langle z^{v_1}(t)-z(t),-H_{1z}(t,\cdot)\right\rangle\right. \\ &\qquad\left.+\left\langle (z^{v_1}(t))'-z'(t),-H_{1z'}(t,\cdot)\right\rangle\right]dt \\ &\quad+\mathbb{E}\int_0^T\int_{\mathbf{E}} \mathbb{E}'\left[\left\langle r^{v_1}(t,e)-r(t,e),H_{1r}(t,e)\right\rangle\right. \\ &\qquad\left.+\left\langle (r^{v_1}(t,e))'-(r(t,e))',H_{1r'}(t,e)\right\rangle\right] \\ &\qquad\times\pi(de)\,dt=\Xi_1.\end{aligned}\tag{46}$$

Therefore, it follows from (35), (40), and (46) that

$$\begin{aligned}J_1(v_1(\cdot),u_2(\cdot))-J_1(u_1(\cdot),u_2(\cdot)) &\geq \Xi_3-\Xi_2-\Xi_1+\Xi_2 \\ &\geq \Xi_1-\Xi_2-\Xi_1+\Xi_2=0.\end{aligned}\tag{47}$$

Then it implies that

$$J_1(u_1(\cdot),u_2(\cdot))=\min_{v_1(\cdot)\in\mathcal{U}_1} J_1(v_1(\cdot),u_2(\cdot)).\tag{48}$$

In the same way

$$J_2(u_1(\cdot),u_2(\cdot))=\min_{v_2(\cdot)\in\mathcal{U}_2} J_2(u_1(\cdot),v_2(\cdot)).\tag{49}$$

Hence, we draw the desired conclusion. The proof is completed. □

5. Application in a Partial Information LQ Case

In this section we work out an example of partial information linear-quadratic differential games of MF-BSDEJ to illustrate the application of the theoretical results. For notational simplification, we assume $n=d=k_1=k_2=1$, $U_1=U_2=\mathbb{R}$, and $\mathscr{E}_t^1=\mathscr{E}_t^2=\mathscr{E}_t\subseteq\mathscr{F}_t$.

Consider the following:

$$\begin{aligned}&-dy^{v_1,v_2}(t) \\ &\quad=\mathbb{E}'\left[A(t)\,y^{v_1,v_2}(t)+C(t)\,z^{v_1,v_2}(t)+D(t)\,r^{v_1,v_2}(t,\cdot)\right. \\ &\qquad+\overline{A}(t)\left(y^{v_1,v_2}(t)\right)'+\overline{C}(t)\left(z^{v_1,v_2}(t)\right)' \\ &\qquad+\overline{D}(t)\left(r^{v_1,v_2}(t,\cdot)\right)'+B_1(t)\,v_1(t) \\ &\qquad\left.+B_2(t)\,v_2(t)\right]dt \\ &\qquad-z^{v_1,v_2}(t)\,dW(t)-\int_{\mathbf{E}} r^{v_1,v_2}(t,e)\,\widetilde{N}(de\,dt), \\ &\qquad y^{v_1,v_2}(T)=\xi.\end{aligned}\tag{50}$$

The cost functional is

$$\begin{aligned}&J_i(v_1(\cdot),v_2(\cdot)) \\ &\quad=\frac{1}{2}\mathbb{E}\left[\int_0^T \mathbb{E}'\left(M_i(t)\,v_i^2(t)+N_i(t)\left(y^{v_1,v_2}(t)\right)^2\right.\right. \\ &\qquad\left.+\overline{N}_i(t)\left(\left(y^{v_1,v_2}(t)\right)'\right)^2\right)dt \\ &\qquad\left.+L_i\left(y^{v_1,v_2}(0)\right)^2\right],\quad i=1,2,\end{aligned}\tag{51}$$

where constants $L_i\geq 0$, $i=1,2$. Functions $A(\cdot)$, $\overline{A}(\cdot)$, $B_1(\cdot)$, $B_2(\cdot)$, $C(\cdot)$, $\overline{C}(\cdot)$, $D(\cdot)$, $\overline{D}(\cdot)$ are bounded and deterministic; $N_i(\cdot)$, $\overline{N}_i(\cdot)$, $i=1,2$, are nonnegative, bounded, and deterministic; $M_i(\cdot)$, $i=1,2$, are positive, bounded, and deterministic; $M_i^{-1}(\cdot)$, $i=1,2$, are also bounded. Our task is to find $(u_1(\cdot),u_2(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2$ such that

$$\begin{aligned}J_1(u_1(\cdot),u_2(\cdot))&=\min_{v_1(\cdot)\in\mathcal{U}_1} J_1(v_1(\cdot),u_2(\cdot)), \\ J_2(u_1(\cdot),u_2(\cdot))&=\min_{v_2(\cdot)\in\mathcal{U}_2} J_2(u_1(\cdot),v_2(\cdot)).\end{aligned}\tag{52}$$

Theorem 5. *The mapping*

$$\begin{aligned}u_1(t)&=M_1^{-1}(t)\,B_1(t)\,\mathbb{E}\left[p_1(t)\mid\mathscr{E}_t\right], \\ u_2(t)&=M_2^{-1}(t)\,B_2(t)\,\mathbb{E}\left[p_2(t)\mid\mathscr{E}_t\right],\end{aligned}\tag{53}$$

is one Nash equilibrium point for the above game problem, where $(p_1(t),p_2(t),y(t),z(t),r(t,\cdot))$ is the solution of the following mean-field forward-backward stochastic differential equations with jumps (MF-FBSDEJ):

$$\begin{aligned}&dp_i(t) \\ &\quad=\mathbb{E}'\left[A(t)\,p_i(t)+\overline{A}(t)\,p_i'(t)-N_i(t)\,y(t)\right. \\ &\qquad\left.-\overline{N}_i(t)\,y'(t)\right]dt \\ &\qquad+\mathbb{E}'\left[C(t)\,p_i(t)+\overline{C}(t)\,p_i'(t)\right]dW(t) \\ &\qquad+\int_{\mathbf{E}}\mathbb{E}'\left[D(t)\,p_i(t)+\overline{D}(t)\,p_i'(t)\right]\widetilde{N}(de\,dt), \\ &-dy(t) \\ &\quad=\mathbb{E}'\left\{A(t)\,y(t)+C(t)\,z(t)+D(t)\,r(t,\cdot)+\overline{A}(t)\,y'(t)\right. \\ &\qquad+\overline{C}(t)\,z'(t)+\overline{D}(t)\,r'(t,\cdot)+B_1^2(t)\,M_1^{-1}(t) \\ &\qquad\left.\times\mathbb{E}\left[p_1(t)\mid\mathscr{E}_t\right]+B_2^2(t)\,M_2^{-1}(t)\,\mathbb{E}\left[p_2(t)\mid\mathscr{E}_t\right]\right\}dt \\ &\qquad-z(t)\,dW(t)-\int_{\mathbf{E}} r(t,e)\,\widetilde{N}(de\,dt), \\ &\qquad p_i(0)=-L_i\,y(0),\qquad y(T)=\xi,\quad i=1,2.\end{aligned}\tag{54}$$

Proof. We first prove the existence of the solution of (54). We set

$$\widehat{\theta}(t)=\mathbb{E}\left[\theta(t)\mid\mathscr{E}_t\right],\quad \theta=y,z,r,y',z',r',p_1,p_2. \tag{55}$$

Similar to Lemma 5.4 of [24], the optimal filter $(\widehat{y}(t),\widehat{z}(t),\widehat{r}(t,\cdot),\widehat{y}'(t),\widehat{z}'(t),\widehat{r}'(t,\cdot),\widehat{p}_1(t),\widehat{p}_2(t))$ of $(y(t),z(t),r(t,\cdot),y'(t),z'(t),r'(t,\cdot),p_1(t),p_2(t))$ satisfies

$$\begin{aligned}
&d\widehat{p}_i(t)\\
&\quad=\mathbb{E}'\left[A(t)\widehat{p}_i(t)+\overline{A}(t)\widehat{p}_i'(t)-N_i(t)\widehat{y}(t)\right.\\
&\qquad\left.-\overline{N}_i(t)\widehat{y}'(t)\right]dt\\
&\qquad+\mathbb{E}'\left[C(t)\widehat{p}_i(t)+\overline{C}(t)\widehat{p}_i'(t)\right]dW(t)\\
&\qquad+\int_{\mathbf{E}}\mathbb{E}'\left[D(t)\widehat{p}_i(t)+\overline{D}(t)\widehat{p}_i'(t)\right]\widetilde{N}(de\,dt),\\
&-d\widehat{y}(t)\\
&\quad=\mathbb{E}'\left[A(t)\widehat{y}(t)+C(t)\widehat{z}(t)+D(t)\widehat{r}(t,\cdot)\right.\\
&\qquad+\overline{A}(t)\widehat{y}'(t)+\overline{C}(t)\widehat{z}'(t)+\overline{D}(t)\widehat{r}'(t,\cdot)\\
&\qquad\left.+B_1^2(t)M_1^{-1}(t)\widehat{p}_1(t)+B_2^2(t)M_2^{-1}(t)\widehat{p}_2(t)\right]dt\\
&\qquad-\widehat{z}(t)\,dW(t)-\int_{\mathbf{E}}\widehat{r}(t,e)\widetilde{N}(de\,dt),\\
&\widehat{p}_i(0)=-L_i\widehat{y}(0),\qquad \widehat{y}(T)=\mathbb{E}\left[\xi\mid\mathscr{E}_T\right],\quad i=1,2.
\end{aligned}\tag{56}$$

Due to the above analysis, the candidate equilibrium point $(u_1(\cdot),u_2(\cdot))$ can be rewritten as

$$\begin{aligned}
u_1(t)&=M_1^{-1}(t)B_1(t)\widehat{p}_1(t),\\
u_2(t)&=M_2^{-1}(t)B_2(t)\widehat{p}_2(t),
\end{aligned}\tag{57}$$

where $\widehat{p}_i(t)$ $(i=1,2)$ admits MF-FBSDEJ (56). We introduce a new MF-FBSDEJ:

$$\begin{aligned}
&dP(t)\\
&\quad=\mathbb{E}'\left[A(t)P(t)+\overline{A}(t)P'(t)\right.\\
&\qquad-\left(B_1^2(t)M_1^{-1}(t)N_1(t)+B_2^2(t)M_2^{-1}(t)N_2(t)\right)\\
&\qquad\times Y(t)-\left(B_1^2(t)M_1^{-1}(t)\overline{N}_1(t)+B_2^2(t)M_2^{-1}(t)\right.\\
&\qquad\left.\left.\times\overline{N}_2(t)\right)Y'(t)\right]dt\\
&\qquad+\mathbb{E}'\left[C(t)P(t)+\overline{C}(t)P'(t)\right]dW(t)\\
&\qquad+\int_{\mathbf{E}}\mathbb{E}'\left[D(t)P(t)+\overline{D}(t)P'(t)\right]\widetilde{N}(de\,dt),\\
&-dY(t)\\
&\quad=\mathbb{E}'\left[A(t)Y(t)+C(t)Z(t)+D(t)R(t,\cdot)\right.\\
&\qquad+\overline{A}(t)Y'(t)+\overline{C}(t)Z'(t)+\overline{D}(t)R'(t,\cdot)\\
&\qquad\left.+P(t)\right]dt-Z(t)\,dW(t)\\
&\qquad-\int_{\mathbf{E}}R(t,e)\widetilde{N}(de\,dt),\\
&P(0)=-\left[B_1^2(0)M_1^{-1}(0)L_1+B_2^2(0)M_2^{-1}(0)L_2\right]Y(0),\\
&Y(T)=\xi.
\end{aligned}\tag{58}$$

Based on the analysis above, we can say the existence and uniqueness of MF-FBSDEJ (56) are equivalent to those of MF-FBSDEJ (58). It is easy to check that MF-FBSDEJ (58) satisfies assumptions (A3)–(A5) with $H=1,\mu_1=1$, and $\mu_2=\beta_2=0$. According to Theorem 7 in Appendix, there exists a unique solution $(P(t),Y(t),Z(t),R(t,\cdot))$ of (58); here,

$$\begin{aligned}
&P(t)=B_1^2(t)M_1^{-1}(t)\widehat{p}_1(t)+B_2^2(t)M_2^{-1}(t)\widehat{p}_2(t),\\
&Y(t)=\widehat{y}(t),\qquad Z(t)=\widehat{z}(t),\qquad R(t,\cdot)=\widehat{r}(t,\cdot).
\end{aligned}\tag{59}$$

Then there exists a unique solution $(\widehat{p}_1(t),\widehat{p}_2(t),\widehat{y}(t),\widehat{z}(t),\widehat{r}(t,\cdot))$ of MF-FBSDEJ (56). Furthermore, there exists at most one equilibrium point for the underlying game.

Now we try to prove that $(u_1(\cdot),u_2(\cdot))$ is a Nash equilibrium point for our backward LQ game problem. We only prove

$$J_1\left(u_1(\cdot),u_2(\cdot)\right)=\min_{v_1(\cdot)\in\mathscr{U}_1}J_1\left(v_1(\cdot),u_2(\cdot)\right). \tag{60}$$

It is similar to getting another inequality of (52). $(y^{v_1}(t),z^{v_1}(t),r^{v_1}(t,\cdot))$ denotes the solution of the system

$$\begin{aligned}
&-dy^{v_1}(t)\\
&\quad=\mathbb{E}'\left[A(t)y^{v_1}(t)+C(t)z^{v_1}(t)+D(t)r^{v_1}(t,\cdot)\right.\\
&\qquad+\overline{A}(t)\left(y^{v_1}(t)\right)'+\overline{C}(t)\left(z^{v_1}(t)\right)'+\overline{D}(t)\left(r^{v_1}(t,\cdot)\right)'\\
&\qquad\left.+B_1(t)v_1(t)+B_2(t)u_2(t)\right]dt\\
&\qquad-z^{v_1}(t)\,dW(t)-\int_{\mathbf{E}}r^{v_1}(t,e)\widetilde{N}(de\,dt),\\
&\qquad y^{v_1}(T)=\xi.
\end{aligned}\tag{61}$$

Then

$$\begin{aligned} &J_1(v_1(\cdot),u_2(\cdot)) - J_1(u_1(\cdot),u_2(\cdot)) \\ &\quad = \frac{1}{2}\mathbb{E}\Big[\int_0^T \mathbb{E}'\big(M_1(t)(v_1(t)-u_1(t))^2 \\ &\qquad + 2M_1(t)u_1(t)(v_1(t)-u_1(t)) \\ &\qquad + N_1(t)(y^{v_1}(t)-y(t))^2 \\ &\qquad + 2N_1(t)y(t)(y^{v_1}(t)-y(t)) \\ &\qquad + \overline{N}_1(t)\left((y^{v_1}(t))'-y'(t)\right)^2 \\ &\qquad + 2\overline{N}_1(t)y'(t)\left((y^{v_1}(t))'-y'(t)\right)dt \\ &\qquad + L_1(y^{v_1}(0)-y^u(0))^2 \\ &\qquad + 2L_1y^u(0)(y^v(0)-y^u(0))\Big]. \end{aligned} \tag{62}$$

Applying Itô's formula to $(y^{v_1}(t)-y(t))\widehat{p}_1(t)$, we have

$$\begin{aligned} &\mathbb{E}\{L_1y^v(0)(y^v(0)-y^u(0))\} \\ &\quad = -\mathbb{E}\int_0^T \mathbb{E}'\big[B_1(t)\widehat{p}_1(t)(v_1(t)-u_1(t)) \\ &\qquad + N_1(t)y(t)(y^{v_1}(t)-y(t)) \\ &\qquad + \overline{N}_1(t)y'(t)\left((y^{v_1}(t))'-y'(t)\right)\big]dt. \end{aligned} \tag{63}$$

As $M_1(t) > 0$, $N_1(t) \geq 0$, $\overline{N}_1(t) \geq 0$, $\forall t \in [0,T]$, $L_1 \geq 0$, noting that $u_1(t) = M_1^{-1}(t)B_1(t)\widehat{p}_1(t)$, we have

$$\begin{aligned} &J_1(v_1(\cdot),u_2(\cdot)) - J_1(u_1(\cdot),u_2(\cdot)) \\ &\quad \geq \mathbb{E}\int_0^T \mathbb{E}'\big[(M_1(t)u_1(t) - B_1(t)\widehat{p}_1(t)) \\ &\qquad \times (v_1(t)-u_1(t))\big]dt = 0. \end{aligned} \tag{64}$$

So $(u_1(t),u_2(t)) = (M_1^{-1}(t)B_1(t)\widehat{p}_1(t), M_2^{-1}(t)B_2(t)\widehat{p}_2(t))$ is a Nash equilibrium point for our backward LQ nonzero sum differential game problem. □

6. Appendix

For the sake of convenience and completeness, we cite the existence and uniqueness theorem of MF-BSDEJ obtained by Shen and Siu [23]. They studied the following MF-BSDEJ:

$$\begin{aligned} &-dy(t) \\ &\quad = \mathbb{E}'f\left(t,\xi(t),\xi'(t),v(t)\right)dt \\ &\qquad - z(t)\,dW(t) - \int_{\mathbf{E}} r(t,e)\widetilde{N}(de\,dt), \\ &\qquad y(T) = \xi, \end{aligned} \tag{65}$$

where $(\xi(t),\xi'(t)) = (y(\cdot),z(\cdot),r(\cdot,\cdot),y'(\cdot),z'(\cdot),r'(\cdot,\cdot)) \in \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n)$, $\xi \in L^2(\Omega,\mathcal{F}_T,P;\mathbb{R}^n)$, is a random variable, and $T > 0$;

$$\begin{aligned} f : [0,T]\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n) \\ \times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n) \longrightarrow \mathbb{R}^n. \end{aligned} \tag{66}$$

They assumed that

(A1) for each ξ, $\xi' \in \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n)$, $f(\cdot,\xi,\xi')$ is an $\mathcal{F}_t$-measurable process defined on $[0,T]$ with $f(\cdot,0,0) \in M^2(0,T;\mathbb{R}^n\times\mathbb{R}^{n\times d}) \times F_N^2(0,T;\mathbb{R}^n)$;

(A2) $f(t,\zeta,\zeta')$ satisfies Lipschitz condition: there exists a constant $l > 0$, such that

$$\begin{aligned} &\left|f(t,\xi,\xi') - f(t,\overline{\xi},\overline{\xi}')\right| \leq l\left(\left|\xi-\overline{\xi}\right| + \left|\xi'-\overline{\xi}'\right|\right), \\ &\forall\xi = (y,z,r)^T, \quad \xi' = (y',z',r')^T, \quad \overline{\xi} = (\overline{y},\overline{z},\overline{r})^T, \\ &\overline{\xi}' = (\overline{y}',\overline{z}',\overline{r}')^T \in \mathbb{R}^n\times\mathbb{R}^{n\times d}\times L^2_{\pi(\cdot)}(\mathbb{R}^n), \quad \forall t \in [0,T]. \end{aligned} \tag{67}$$

Based on the fixed-point theorem, Shen and Siu [23] obtained the following existence and uniqueness result.

Proposition 6. *One assumes that (A1) and (A2) hold. Then MF-BSDEJ (65) has a unique solution* $(y(t),z(t),r(t,\cdot)) \in M^2(0,T;\mathbb{R}^n\times\mathbb{R}^{n\times d}) \times F_N^2(0,T;\mathbb{R}^n)$.

In the present paper we research the game problem of a MF-BSDEJ, so the game system and the adjoint equation constitute exactly a kind of initial coupled MF-FBSDEJ. Due to this, we give an existence and uniqueness theorem of MF-FBSDEJ. Consider the following MF-FBSDEJ:

$$\begin{aligned} &dx(t) \\ &\quad = \mathbb{E}'\big[b\big(t,x(t),y(t),z(t),r(t,\cdot),x'(t), \\ &\qquad y'(t),z'(t),r'(t,\cdot)\big)\big]dt \\ &\qquad + \mathbb{E}'\big[\sigma\big(t,x(t),y(t),z(t),r(t,\cdot),x'(t), \\ &\qquad y'(t),z'(t),r'(t,\cdot)\big)\big]dW(t) \\ &\qquad + \int_{\mathbf{E}}\mathbb{E}'\big[\gamma\big(t,x(t),y(t),z(t),r(t,e),x'(t), \\ &\qquad y'(t),z'(t),r'(t,e),e\big)\big]\widetilde{N}(de\,dt), \\ &-dy(t) \\ &\quad = \mathbb{E}'\big[f\big(t,x(t),y(t),r(t,\cdot),x'(t), \\ &\qquad y'(t),r'(t,\cdot)\big)\big]dt - z(t)\,dW(t) \\ &\qquad - \int_{\mathbf{E}} r(t,e)\widetilde{N}(de\,dt), \\ &\qquad x(0) = \psi(y(0)), \quad y(T) = \xi, \end{aligned} \tag{68}$$

where $(x(\cdot), y(\cdot), z(\cdot), r(\cdot,\cdot), x'(\cdot), y'(\cdot), z'(\cdot), r'(\cdot,\cdot)) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n)$, $\xi \in L^2(\Omega, \mathcal{F}_T, P; \mathbb{R}^n)$, is a random variable, and $T > 0$;

$$\begin{aligned}
b &: \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \\
&\quad \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \longrightarrow \mathbb{R}^m, \\
\sigma &: \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \\
&\quad \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \longrightarrow \mathbb{R}^{m\times d}, \\
\gamma &: \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \\
&\quad \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \times \mathbf{E} \longrightarrow \mathbb{R}^m, \\
f &: \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \\
&\quad \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n) \longrightarrow \mathbb{R}^n, \\
\psi &: \mathbb{R}^n \longrightarrow \mathbb{R}^m.
\end{aligned} \tag{69}$$

Given an $n \times m$ full-rank matrix H, let us introduce some notations

$$\zeta = \begin{pmatrix} x \\ y \\ z \\ r \end{pmatrix}, \qquad \zeta' = \begin{pmatrix} x' \\ y' \\ z' \\ r' \end{pmatrix},$$
$$A(t,\zeta,\zeta') = \begin{pmatrix} -H^T f \\ Hb \\ H\sigma \\ H\gamma \end{pmatrix} (t,\zeta,\zeta'), \tag{70}$$

where $H\sigma = (H\sigma_1 \cdots H\sigma_d)$. Assume that

(A3) for each $\zeta, \zeta' \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n)$, $A(\cdot,\zeta,\zeta')$ is an $\mathcal{F}_t$-measurable process defined on $[0,T]$ with $A(\cdot,0,0) \in M^2(0,T;\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d}) \times F^2_N(0,T;\mathbb{R}^n)$;

(A4) $A(t,\zeta,\zeta')$ and $\psi(y)$ satisfy Lipschitz conditions: there exists a constant $k > 0$, such that

$$\begin{aligned}
&\left|A(t,\zeta,\zeta') - A(t,\overline{\zeta},\overline{\zeta}')\right| \le k\left(\left|\zeta - \overline{\zeta}\right| + \left|\zeta' - \overline{\zeta}'\right|\right), \\
&\quad \forall \zeta = (x,y,z,r)^T, \qquad \zeta' = (x',y',z',r')^T, \\
&\quad \overline{\zeta} = (\overline{x},\overline{y},\overline{z},\overline{r})^T, \\
&\overline{\zeta}' = (\overline{x}',\overline{y}',\overline{z}',\overline{r}')^T \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n), \\
&\qquad\qquad\qquad \forall t \in [0,T], \\
&\quad \left|\psi(y) - \psi(\overline{y})\right| \le k\left|y - \overline{y}\right|, \quad \forall y, \overline{y} \in \mathbb{R}^n;
\end{aligned} \tag{71}$$

(A5) $A(t,\zeta,\zeta')$ and $\psi(y)$ satisfy monotonic conditions:

$$\begin{aligned}
&\left\langle A(t,\zeta,\zeta') - A(t,\overline{\zeta},\overline{\zeta}'), \zeta - \overline{\zeta} \right\rangle \\
&\quad \le -\mu_1 \left|H(x - \overline{x})\right|^2 \\
&\qquad - \mu_2 \left(\left|H^T(y - \overline{y})\right|^2 + \left|H^T(z - \overline{z})\right|^2 \right. \\
&\qquad\qquad \left. + \left\|H^T(r - \overline{r})\right\|^2\right), \\
&\quad \forall \zeta = (x,y,z,r)^T, \qquad \zeta' = (x',y',z',r')^T, \\
&\quad \overline{\zeta} = (\overline{x},\overline{y},\overline{z},\overline{r})^T, \\
&\overline{\zeta}' = (\overline{x}',\overline{y}',\overline{z}',\overline{r}')^T \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times L^2_{\pi(\cdot)}(\mathbb{R}^n), \\
&\qquad\qquad\qquad \forall t \in [0,T]. \\
&\left\langle \psi(y) - \psi(\overline{y}), y - \overline{y} \right\rangle \le -\beta_2 \left|H^T(y - \overline{y})\right|^2, \quad \forall y, \overline{y} \in \mathbb{R}^n,
\end{aligned} \tag{72}$$

where μ_1, μ_2, and β_2 are given nonnegative constants with $\mu_1 + \mu_2 > 0$, $\mu_1 + \beta_2 > 0$. Moreover we have $\mu_1 > 0$ (resp., $\mu_2 > 0$, $\beta_2 > 0$) when $m < n$ (resp., $m > n$).

By similar arguments of Yu and Ji [14], Wang and Yu [19], and Min et al. [25], we have the following existence and uniqueness theorem.

Theorem 7. *One assumes that (A3), (A4), and (A5) hold. Then MF-FBSDEJ* (68) *has a unique solution* $(x(t), y(t), z(t), r(t,\cdot)) \in M^2(0,T;\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d}) \times F^2_N(0,T;\mathbb{R}^n)$.

7. Conclusion

In this paper, we investigate a new differential game problem of mean-field BSDE with jump (MF-BSDEJ). Compared with the previous literature, our game systems are mean-field BSDE with jump and are under the framework of partial information. We established a maximum principle and a verification theorem for an equilibrium point of nonzero sum differential games. We also give a partial information linear-quadratic (LQ) game as example to show the applications of our theoretical results.

The subject issue studied in this paper possesses fine generality. Firstly, the mean-field BSDEJ game system covers many systems as its particular case. For example, if we drop the terms on jump or mean-field or both of them, then the MF-BSDEJ can be reduced to MF-BSDE or BSDEJ or BSDE. Secondly, if we suppose that $\mathcal{E}_t = \mathcal{F}_t$, for all $t \in [0,T]$, all the results are reduced to the case of full information. Thirdly, if the present nonzero sum stochastic differential game has only one player, the game problem is reduced to some related optimal control. Particularly, our results are a partial extension to differential games of full information BSDEs [15] and partial information BSDEs [19]. Finally, since many optimization and game problems in finance and economics can be associated with mean-field BSDE with jump, the outcomes of this paper can be widely applied in these areas.

Conflict of Interests

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

Acknowledgments

The authors would like to thank the editors and the reviewers for their constructive comments and suggestions which helped them to improve this paper. This work was supported by the National Natural Science Foundation of China (Grant nos. 11371226, 11071145, 11301298, and 11231005), the Foundation for Innovative Research Groups of National Natural Science Foundation of China (Grant no. 11221061), the 111 Project (Grant no. B12023), and the Development Project of Shandong Province Science and Technology (Grant no. 2012GGB01241).

References

[1] R. Isaacs, "Differential Games, Parts 1–4," The Rand Corpration, Research Memorandums Nos. RM-1391, RM-1411, RM-1486, 1954-1955.

[2] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, vol. 160 of *Mathematics in Science and Engineering*, Academic Press, London, UK, 1982.

[3] D. W. K. Yeung and L. A. Petrosyan, *Cooperative Stochastic Differential Games*, Springer, New York, NY, USA, 2005.

[4] L. D. Berkovitz, "Necessary conditions for optimal strategies in a class of differential games and control problems," *SIAM Journal on Control*, vol. 5, no. 1, pp. 1–24, 1967.

[5] W. H. Fleming, "The convergence problem for differential games," *Journal of Mathematical Analysis and Applications*, vol. 3, pp. 102–116, 1961.

[6] R. J. Elliott and N. J. Kalton, "The existence of value in differential games," *Bulletin of the American Mathematical Society*, vol. 126, pp. 1–67, 1972.

[7] A. Friedman, *Differential Games*, Wiley-Interscience, New York, NY, USA, 1971.

[8] P. Varaiya, "N-player stochastic differential games," *SIAM Journal on Control and Optimization*, vol. 14, no. 3, pp. 538–545, 1976.

[9] R. J. Elliott and M. H. A. Davis, "Optimal play in a stochastic differential game," *SIAM Journal on Control and Optimization*, vol. 19, no. 4, pp. 543–554, 1981.

[10] S. Hamadène, "Nonzero sum linear-quadratic stochastic differential games and backward-forward equations," *Stochastic Analysis and Applications*, vol. 17, no. 1, pp. 117–130, 1999.

[11] S. Hamadène, J. P. Lepeltier, and S. Peng, "BSDE with continuous coefficients and applications to Markovian nonzero sum stochastic differential games," in *Pitman Research Notes Mathematics Series: Vol. 364. Backward Stochastic Differential Equations*, N. El Karoui and L. Mazliak, Eds., pp. 115–128, Longman, Harlow, UK, 1997.

[12] E. Altman, "Applications of dynamic games in queues," in *Advances in Dynamic Games*, vol. 7, pp. 309–342, Birkhauser, Boston, Mass, USA, 2005.

[13] Z. Wu and Z. Y. Yu, "Linear quadratic nonzero-sum differential games with random jumps," *Applied Mathematics and Mechanics*, vol. 26, no. 8, pp. 1034–1039, 2005.

[14] Z. Yu and S. Ji, "Linear-quadratic nonzero-sum differential game of backward stochastic differential equations," in *Proceedings of the 27th Chinese Control Conference (CCC '08)*, pp. 562–566, Kunming, China, July 2008.

[15] G. Wang and Z. Yu, "A Pontryagin's maximum principle for non-zero sum differential games of BSDEs with applications," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1742–1747, 2010.

[16] T. T. K. An and B. Øksendal, "Maximum principle for stochastic differential games with partial information," *Journal of Optimization Theory and Applications*, vol. 139, no. 3, pp. 463–483, 2008.

[17] T. T. K. An and B. Øksendal, "A maximum principle for stochastic differential games with g-expectations and partial information," *Stochastics*, vol. 84, no. 2-3, pp. 137–155, 2012.

[18] T. T. K. An, B. Øksendal, and Y. Y. Okur, "A Malliavin calculus approach to general stochastic differential games with partial information," in *Malliavin Calculus and Stochastic Analysis*, vol. 34 of *Springer Proceedings in Mathematics & Statistics*, pp. 489–510, Springer, Berlin, Germany.

[19] G. Wang and Z. Yu, "A partial information non-zero sum differential game of backward stochastic differential equations with applications," *Automatica*, vol. 48, no. 2, pp. 342–352, 2012.

[20] Q. Meng and M. Tang, "Stochastic differential games of fully coupled forward-backward stochastic systems under partial information," in *Proceedings of the 29th Chinese Control Conference (CCC '10)*, pp. 1150–1155, Beijing, China, July 2010.

[21] E. C. M. Hui and H. Xiao, "Differential games of partial information forward-backward doubly SDE and applications," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 20, no. 1, pp. 78–94, 2014.

[22] B. Øksendal and A. Sulem, "Forward-backward stochastic differential games and stochastic control under model uncertainty," *Journal of Optimization Theory and Applications*, vol. 161, no. 1, pp. 22–55, 2014.

[23] Y. Shen and T. K. Siu, "The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 86, pp. 58–73, 2013.

[24] J. Xiong, *An Introduction to Stochastic Filtering Theory*, Oxford University Press, London, UK, 2008.

[25] H. Min, Y. Peng, and Y. Qin, "Fully coupled mean-field forward-backward stochastic differential equations and stochastic maximum principle," *Abstract and Applied Analysis*, vol. 2014, Article ID 839467, 15 pages, 2014.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 494575, 9 pages
http://dx.doi.org/10.1155/2014/494575

Research Article

A Closed-Form Solution for Robust Portfolio Selection with Worst-Case CVaR Risk Measure

Le Tang[1] and Aifan Ling[2]

[1] *School of Foreign Languages and International Business, Jiangxi University of Technology, Nanchang 330098, China*
[2] *School of Finance, Jiangxi University of Finance and Economics, Nanchang 330013, China*

Correspondence should be addressed to Aifan Ling; aiffling@163.com

Received 24 March 2014; Accepted 11 June 2014; Published 2 July 2014

Academic Editor: Guangchen Wang

With the uncertainty probability distribution, we establish the worst-case CVaR (WCCVaR) risk measure and discuss a robust portfolio selection problem with WCCVaR constraint. The explicit solution, instead of numerical solution, is found and two-fund separation is proved. The comparison of efficient frontier with mean-variance model is discussed and finally we give numerical comparison with VaR model and equally weighted strategy. The numerical findings indicate that the proposed WCCVaR model has relatively smaller risk and greater return and relatively higher accumulative wealth than VaR model and equally weighted strategy.

1. Introduction

VaR (value at risk) has been a popular risk measure in finance industry and academic research and is written in New Basel Accord. But two difficulties are faced by user: (1) the explicit expression of VaR is unavailable unless the normal distribution assumption is done and (2) VaR, as a risk measure tool, does not satisfy the coherent axiom [1]. Hence, an approximation of VaR is often considered in practice by either assuming normal distribution or simulation method based historical data. Rockafellar and Uryasev [2] proposed an alternative risk measure, namely, conditional VaR (CVaR), which is coherent and provided a linear programming approximate with historical data. But the assumption that the return of risky asset follows the normal distribution is usually done when one computes CVaR by parameterized approaches. As we know, normal distribution can usually underestimate the loss of the rare event and is not clearly a very good approximation of the return of risky asset. This is still a challenge for computing an explicit expression of CVaR without any special distribution information. The current paper will explore this problem and establish the mean-CVaR portfolio model without probability distribution assumption.

Robust portfolio problems with parameters uncertainty are recently paid close attention to. Goldfarb and Iyengar [3], for instance, considered a class of robust portfolio problem with risk factors in which they solve numerically robust mean-variance portfolio problem, robust downside risk portfolio problem with normal distribution, and robust Sharpe ratio portfolio problem; see also, Costa and Paiva [4], Halldórsson and Tütüncü [5], Tütüncü and Koenig [6], Lu [7], and Ling and Xu [8] for the relative researches. The uncertainty of models above is only from the parameters under the deterministic distribution and cannot capture the uncertainty in distribution. El Ghaoui et al. [9] proposed the worst-case VaR (WCVaR) risk measure and considered a portfolio selection problem with minimization of WCVaR. Zhu and Fukushima [10] proposed the worst-case CVaR risk measure and discussed a robust mean-CVaR portfolio model with uncertainty discrete distribution. Some similar researches can be found in Zhu et al. [11] and Huang et al. [12]. A richer literature can be referred to in Fabozzi et al. [13].

We define the worst-case CVaR with uncertainty distribution including the continuous and discrete distribution and consider a portfolio selection problem with WCCVaR as risk measure. Our results extend that of Zhu and Fukushima [10] for which they considered only the discrete case to the case including the continuous and discrete distribution. Most of methods for robust portfolio problems are that one converts first the problems into convex cone (e.g., linear programming, second-order cone programming, or positive semidefinite programming) and then solves them numerically. Differently from these numerical methods, we consider an analytic solution approach for the proposed robust mean-WCCVaR problem. We discuss two cases of the proposed robust problems with and without risky-free asset and prove two-fund separable theorem. Numerical results and comparisons with VaR and equally weighted strategy for real market data are reported.

The outline of this paper is arranged as follows. We introduce the definition of worst-case CVaR, establish mean-WCCVaR portfolio model, and give the closed-form solution in Section 2 and Section 3 proves the two-fund separation theorem. The extension of the model with risky-free asset is considered in Section 4. Numerical results are reported in Section 5.

2. Mean-WCCVaR Portfolio Model

We consider mainly an investing and holding strategy in this paper for which the investor allocates his (her) assets at time 0 and collects his (her) returns of portfolio at time 1. Generally speaking, two things must be done at time 0: one is that the investor needs to estimate the returns of risky assets at time 1 using the available information at time 0 and another is that the investor must choose an optimal decision to allocate his (her) wealth.

Let there be n available risky assets in the market and let their random returns vector be denoted by $\mathbf{r} = (r_1, \ldots, r_n)^T \in \mathbb{R}^n$. The expected returns and covariance matrix are denoted, respectively, by $\boldsymbol{\mu} = \mathbb{E}[\mathbf{r}] = (\mu_1, \ldots, \mu_n)^T \in \mathbb{R}^n$ and $\Sigma = (\sigma_{ij})_{n\times n}$, where $\sigma_{ii} = \sigma_i^2 = \operatorname{var}(r_i)$ is variance of asset i $(i = 1, \ldots, n)$. The rate of return of risky-free asset is denoted by r_f. The portfolio vector is $\mathbf{w} = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$ with w_i the proportion of wealth invested in asset i. The weight of wealth invested in risky-free asset is denoted by w_f. Let $f(\mathbf{w}, \mathbf{r})$ be the loss for portfolio vector $\mathbf{w}$ and satisfy $\mathbb{E}[|f(\mathbf{w}, \mathbf{r})|] < +\infty$, and let $F(\cdot)$ be the joint cumulative probability distribution of random vector $\mathbf{r}$. Then the probability that the loss $f(\mathbf{w}, \mathbf{r})$ is not greater than a given constant α is

$$\Psi(\mathbf{w}, \alpha) = \mathbb{P}\{f(\mathbf{w}, \mathbf{r}) \le \alpha\} = \int_{f(\mathbf{w},\mathbf{r})\le\alpha} dF(\mathbf{r}). \tag{1}$$

Let $\beta \in (0, 1)$; then, with confidence level β, VaR_β can be expressed as

$$\mathrm{VaR}_\beta(\mathbf{w}) = \min\{\alpha \in \mathbb{R} \mid \Psi(\mathbf{w}, \alpha) \ge \beta\}. \tag{2}$$

Rockafellar and Uryasev [2] defined CVaR as the conditional expectation of loss $f(\mathbf{w}, \mathbf{r})$ greater than VaR_β. With the definition of CVaR and the given confidence level β, the mathematic formulation of CVaR can be written as

$$\begin{aligned}\mathrm{CVaR}_\beta(\mathbf{w}) &= \mathbb{E}\left[f(\mathbf{w}, \mathbf{r}) \mid f(\mathbf{w}, \mathbf{r}) \ge \mathrm{VaR}_\beta(\mathbf{w})\right] \\ &= \frac{1}{1-\beta}\int_{f(\mathbf{w},\mathbf{r})\ge\mathrm{VaR}_\beta(\mathbf{w})} f(\mathbf{w}, \mathbf{r})\, dF(\mathbf{r}).\end{aligned} \tag{3}$$

Let

$$H_\beta(\mathbf{w}, \alpha) = \alpha + \frac{1}{1-\beta}\mathbb{E}\left[(f(\mathbf{w}, \mathbf{r}) - \alpha)_+\right]. \tag{4}$$

Then $\mathrm{CVaR}_\beta(\mathbf{w})$ can be expressed further as [2]

$$\mathrm{CVaR}_\beta(\mathbf{w}) = \min_\alpha H_\beta(\mathbf{w}, \alpha), \tag{5}$$

where

$$(a)_+ = \begin{cases} a, & a > 0; \\ 0, & a \le 0. \end{cases} \tag{6}$$

Clearly, it is not possible to get an exact result of CVaR by (4) if we have not any information on the distribution of random vector $\mathbf{r}$. Some sampling or simulation methods are used to computes the approximation of CVaR in the literature. We explore a closed-form solution in this paper with only partial distribution assumptions for random vector $\mathbf{r}$, for which we assume that random vector $\mathbf{r}$ follows a family of distributions $\mathscr{D}$ defined by

$$\mathscr{D} = \{\mathbf{r} \mid \mathbb{E}[\mathbf{r}] = \boldsymbol{\mu}, \operatorname{Cov}(\mathbf{r}) = \Sigma \succ 0\}, \tag{7}$$

where $\Sigma \succ 0$ means that Σ is a positive definite matrix and $\mathscr{D}$ is called the uncertainty set of distribution of random vector $\mathbf{r}$. Clearly, $\mathscr{D}$ is a distribution family with given mean value $\boldsymbol{\mu}$ and covariance matrix Cov, where $\boldsymbol{\mu}$, Cov are assumed to be known. We then can compute CVaR_β when the worst-case probability distribution in $\mathscr{D}$ occurs. To this end, we define worst-case CVaR as follows.

Definition 1. Let $\beta \in (0, 1)$, worst-case CVaR (WCCVaR) of portfolio $\mathbf{w}$ under uncertainty set $\mathscr{D}$ is defined by

$$\mathrm{WCCVaR}_\beta(\mathbf{w}) = \sup_{F\in\mathscr{D}} \mathrm{CVaR}_\beta(\mathbf{w}) = \sup_{F\in\mathscr{D}} \min_\alpha H_\beta(\mathbf{w}, \alpha). \tag{8}$$

Noticing that $H_\beta(\mathbf{w}, \alpha)$ is the convex function of α [14], then we have from max-min theorem [15]

$$\mathrm{WCCVaR}_\beta(\mathbf{w}) = \sup_{F\in\mathscr{D}} \min_\alpha H_\beta(\mathbf{w}, \alpha) = \min_\alpha \sup_{F\in\mathscr{D}} H_\beta(\mathbf{w}, \alpha). \tag{9}$$

The following results are straightforward and a similar proof can be found in [10].

Theorem 2. $\mathrm{WCCVaR}_\beta(\mathbf{w})$ *is a coherent risk measure and satisfies that*

$$\mathrm{WCCVaR}_\beta(\mathbf{w}) \ge \mathrm{CVaR}_\beta(\mathbf{w}) \ge \mathrm{VaR}_\beta(\mathbf{w}). \tag{10}$$

Hence, $\mathrm{WCCVaR}_\beta(\mathbf{w})$ can be used as a risk measure and if the investor measures the risk of portfolio based on WCCVaR, then we can establish the mean-WCCVaR portfolio model by

$$\begin{aligned} \max_{\mathbf{w}} \quad & \boldsymbol{\mu}^T\mathbf{w} \\ \text{s.t.} \quad & \mathrm{WCCVaR}_\beta(\mathbf{w}) \le \tau, \\ & \mathbf{e}^T\mathbf{w} = 1, \end{aligned} \tag{RP1}$$

where $\tau > 0$ is a preset constant. We solve mainly problem (RP1) by exploring an explicit approach. To this end, for convenience, we denote sometime $\mathbf{r} \in \mathscr{D}$ by $\mathbf{r} \sim (\boldsymbol{\mu}, \Sigma)$. The following result is helpful for our analysis later.

Lemma 3 (see [15]). *Let ξ be a random variable with mean value μ and variance σ. ρ is any real number. Then, for the supper bound of $\mathbb{E}[(\rho-\xi)_+]$, we have*

$$\sup_{\xi\sim(\mu,\sigma)} \mathbb{E}\left[(\rho-\xi)_+\right] = \frac{\rho-\mu+\sqrt{\sigma^2+(\rho-\mu)^2}}{2}. \tag{11}$$

Lemma 4 (see [14, 16]). *Let for any vector $\mathbf{a} \in \mathbb{R}^n$,*

$$\begin{aligned} S_1 &= \left\{\mathbf{a}^T\mathbf{r} \mid \mathbf{r} \in \mathscr{D}\right\}, \\ S_2 &= \left\{\eta \mid \mathbb{E}(\eta) = \mathbf{a}^T\boldsymbol{\mu}, \operatorname{Var}(\eta) = \mathbf{a}^T\Sigma\mathbf{a}\right\}. \end{aligned} \tag{12}$$

Then $S_1 = S_2$.

Lemma 3 provides an upper bound of 1-order lower partial moment for one dimensional random variable and Lemma 4 provides a relationship of uncertainty set between one dimensional and several dimensional random variable with given mean value and variance. Lemma 4 indicates also that S_1 (or S_2) is in fact a single variable distribution family with given mean $\mathbf{a}^T\boldsymbol{\mu}$ and variance $\mathbf{a}^T\Sigma\mathbf{a}$, where $\mathbf{a}$, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$ are known. Further, we can obtain an explicit expression for sup in (4) if the loss $f(\mathbf{w},\mathbf{r})$ is linear.

Lemma 5. *If $f(\mathbf{w},\mathbf{r}) = -\mathbf{r}^T\mathbf{w}$, then for any $\mathbf{r} \in \mathscr{D}$ and $\alpha \in \mathbb{R}$, we have*

$$\begin{aligned} &\sup_{\mathbf{r}\in\mathscr{D}} \mathbb{E}\left[(f(\mathbf{w},\mathbf{r})-\alpha)_+\right] \\ &\quad = \sup_{\mathbf{r}\in\mathscr{D}} \mathbb{E}\left[\left(-\alpha-\mathbf{r}^T\mathbf{w}\right)_+\right] \\ &\quad = \frac{1}{2}\left(\sqrt{\mathbf{w}^T\Sigma\mathbf{w}+(\boldsymbol{\mu}^T\mathbf{w}+\alpha)^2}-\left(\boldsymbol{\mu}^T\mathbf{w}+\alpha\right)\right). \end{aligned} \tag{13}$$

Proof. Let $\mathbf{a} = \mathbf{w}$ and $\eta = \mathbf{w}^T\mathbf{r}$. Then, it follows from Lemma 4 that

$$\begin{aligned} &\sup_{\mathbf{r}\in\mathscr{D}} \mathbb{E}\left[(f(\mathbf{w},\mathbf{r})-\alpha)_+\right] \\ &\quad = \sup_{\mathbf{r}\in\mathscr{D}} \mathbb{E}\left[\left(-\alpha-\mathbf{a}^T\mathbf{r}\right)_+\right] \\ &\quad = \sup_{\mathbf{w}^T\mathbf{r}\in S_1} \mathbb{E}\left[\left(-\alpha-\mathbf{w}^T\mathbf{r}\right)_+\right] \\ &\quad = \sup_{\eta\in S_2} \mathbb{E}\left[(-\alpha-\eta)_+\right] \\ &\quad = \sup_{\eta\sim(\mathbf{a}^T\boldsymbol{\mu},\mathbf{a}^T\Sigma\mathbf{a})} \mathbb{E}\left[(-\alpha-\eta)_+\right] \\ &\quad = \sup_{\eta\sim(\mathbf{w}^T\boldsymbol{\mu},\mathbf{w}^T\Sigma\mathbf{w})} \mathbb{E}\left[(-\alpha-\eta)_+\right] \\ &\quad = \frac{1}{2}\left(\sqrt{\mathbf{w}^T\Sigma\mathbf{w}+(\boldsymbol{\mu}^T\mathbf{w}+\alpha)^2}-\left(\boldsymbol{\mu}^T\mathbf{w}+\alpha\right)\right). \end{aligned} \tag{14}$$

The final equality follows from Lemma 3; this is the desired result. □

Hence, from Lemma 5, $\mathrm{WCCVaR}_\beta(\mathbf{w})$ can be expressed by

$$\begin{aligned} &\mathrm{WCCVaR}_\beta(\mathbf{w}) \\ &\quad = \min_{\alpha\in\mathbb{R}} \left\{\alpha + \frac{1}{2(1-\beta)}\left(\sqrt{\mathbf{w}^T\Sigma\mathbf{w}+(\boldsymbol{\mu}^T\mathbf{w}+\alpha)^2} - \left(\boldsymbol{\mu}^T\mathbf{w}+\alpha\right)\right)\right\}. \end{aligned} \tag{15}$$

Clearly, the right side of the equation above is convex function in α and we can prove that $\mathrm{WCCVaR}_\beta(\mathbf{w})$ can be attained at

$$\alpha^* = -\boldsymbol{\mu}^T\mathbf{w} + \frac{2\beta-1}{2\sqrt{\beta(1-\beta)}}\sqrt{\mathbf{w}^T\Sigma\mathbf{w}}; \tag{16}$$

that is,

$$\mathrm{WCCVaR}_\beta(\mathbf{w}) = -\boldsymbol{\mu}^T\mathbf{w} + \sqrt{\frac{\beta}{1-\beta}}\sqrt{\mathbf{w}^T\Sigma\mathbf{w}}. \tag{17}$$

Then, robust mean-WCCVaR portfolio problem (RP1) can be expressed further as

$$\begin{aligned} \max_{\mathbf{w}} \quad & \boldsymbol{\mu}^T\mathbf{w} \\ \text{s.t.} \quad & -\boldsymbol{\mu}^T\mathbf{w} + \sqrt{\frac{\beta}{1-\beta}}\sqrt{\mathbf{w}^T\Sigma\mathbf{w}} \le \tau, \\ & \mathbf{e}^T\mathbf{w} = 1. \end{aligned} \tag{RP2}$$

In the rest of this paper, we discuss mainly the solution of (RP2). To this end, let $a = \mathbf{e}^T\Sigma^{-1}\mathbf{e}$, $b = \boldsymbol{\mu}^T\Sigma^{-1}\mathbf{e}$, $c = \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}$, $a_0 = (ac-b^2)/a$, and $\beta_0 = \sqrt{(1-\beta)/\beta}$. Then, the following result gives feasible conditions of problem (RP2).

Lemma 6. *If $a_0\beta_0^2 < 1$ and*

$$\tau \ge \max\left\{\tau^*, 0\right\}, \tag{18}$$

then problem (RP2) *is feasible, where $\tau^* = \sqrt{1-a_0\beta_0^2}/(\beta_0\sqrt{a}) - b/a$.*

Proof. For any given $\beta \in (0,1)$, consider the problem

$$\tau^* = \min_{\mathbf{w}} \left\{-\boldsymbol{\mu}^T\mathbf{w} + \sqrt{\frac{\beta}{1-\beta}}\sqrt{\mathbf{w}^T\Sigma\mathbf{w}} \mid \text{s.t. } \mathbf{e}^T\mathbf{w} = 1\right\}. \tag{19}$$

Let $v = \sqrt{\mathbf{w}\Sigma\mathbf{w}}$ and let v be viewed as a new variable. Denote the optimal solution of problem (19) by $\widehat{\mathbf{w}}$; then $\widehat{\mathbf{w}}$ satisfies the first order condition:

$$\begin{gathered}
-\boldsymbol{\mu} + \lambda'\mathbf{e} + 2\lambda''\Sigma^{-1}\widehat{\mathbf{w}} = 0, \\
2\lambda'' v = \sqrt{\frac{\beta}{1-\beta}}, \\
\mathbf{e}^T\widehat{\mathbf{w}} = 1, \\
\widehat{\mathbf{w}}^T\Sigma\widehat{\mathbf{w}} = v^2,
\end{gathered} \tag{20}$$

where $\lambda', \lambda'' \in \mathbb{R}$ are the Lagrange multipliers. It is not hard to obtain from the first and third equations of (20)

$$\widehat{\mathbf{w}} = \frac{1}{2\lambda''}\left(\Sigma^{-1}\boldsymbol{\mu} - \frac{b}{a}\Sigma^{-1}\mathbf{e}\right) + \frac{1}{a}\Sigma^{-1}\mathbf{e}. \tag{21}$$

Substituting $\widehat{\mathbf{w}}$ in the fourth equation and combining the second equation of (20), then we can get the optimal solution of (19) when $a_0\beta_0^2 < 1$

$$\widehat{\mathbf{w}} = \frac{\beta_0}{\sqrt{a\left(1 - a_0\beta_0^2\right)}}\left(\Sigma^{-1}\boldsymbol{\mu} - \frac{b}{a}\Sigma^{-1}\mathbf{e}\right) + \frac{1}{a}\Sigma^{-1}\mathbf{e}. \tag{22}$$

The optimal value

$$\tau^* = \frac{\sqrt{1 - a_0\beta_0^2}}{\beta_0\sqrt{a}} - \frac{b}{a}. \tag{23}$$

If $\tau = \tau^* > 0$, then $\widehat{\mathbf{w}}$ is unique solution of problem. This means that the feasible condition is $\tau \geq \max\{\tau^*, 0\}$. The proof is finished. □

Lemma 6 means that the portfolio problem (RP2) is well defined if the investor chooses an appropriate risk tolerance parameter τ. The following theorem gives the main results of the current paper.

Theorem 7. *If $a_0\beta_0^2 < 1$ and $\tau > \max\{\tau^*, 0\}$, the optimal solution of the problem* (RP2) can explicitly be expressed

$$\mathbf{w}^* = f(\tau)\left(\Sigma^{-1}\boldsymbol{\mu} - \frac{b}{a}\Sigma^{-1}\mathbf{e}\right) + \frac{1}{a}\Sigma^{-1}\mathbf{e}, \tag{24}$$

where

$$f(\tau) = \frac{a_0 b\beta_0^2(\tau + b/a) + b\sqrt{a_0\beta_0^2(\tau + b/a)^2 - (a_0/a)\left(1 - a_0\beta_0^2\right)}}{a_0\left(1 - a_0\beta_0^2\right)}. \tag{25}$$

Proof. Let $v = \sqrt{\mathbf{w}\Sigma\mathbf{w}}$; the optimization problem (RP2) can be rewritten as

$$\begin{aligned}
\max_{\mathbf{w}, v>0} \quad & \mathbb{E}\left[\mathbf{r}^T\mathbf{w}\right] \\
\text{s.t.} \quad & v - \beta_0\boldsymbol{\mu}^T\mathbf{w} \leq \beta_0\tau, \\
& \mathbf{e}^T\mathbf{w} = 1, \\
& \mathbf{w}^T\Sigma\mathbf{w} = v^2.
\end{aligned} \tag{26}$$

Let $\mathbf{w}^*$ be the optimal solution of problem (26). Then from KKT condition, $\mathbf{w}^*$ satisfies that

$$\begin{gathered}
\left(1 + \beta_0\lambda_1\right)\boldsymbol{\mu} - \lambda_2\mathbf{e} - 2\lambda_3\Sigma\mathbf{w}^* = 0, \\
-\lambda_1 + 2\lambda_3 v = 0, \\
\mathbf{e}^T\mathbf{w}^* = 1, \\
\left(\mathbf{w}^*\right)^T\Sigma\mathbf{w}^* = v^2, \\
v - \beta_0\boldsymbol{\mu}^T\mathbf{w}^* - \beta_0\tau = 0,
\end{gathered} \tag{27}$$

where $\lambda_1 \geq 0$, $\lambda_2 \in \mathbb{R}$ and $\lambda_3 \in \mathbb{R}$ are Lagrange multipliers. It follows from the first and third equations in (27) that

$$\mathbf{w}^* = \frac{\left(1 + \beta_0\lambda_1\right)}{2\lambda_3}\left(\Sigma^{-1}\boldsymbol{\mu} - \frac{b}{a}\Sigma^{-1}\mathbf{e}\right) + \frac{1}{a}\Sigma^{-1}\mathbf{e}. \tag{28}$$

Substituting $\mathbf{w}^*$ in the fourth and fifth equation in (27), we have

$$\begin{gathered}
a_0\left[\frac{\left(1 + \beta_0\lambda_1\right)}{2\lambda_3}\right]^2 + \frac{1}{a} = v^2, \\
a_0\beta_0\frac{\left(1 + \beta_0\lambda_1\right)}{2\lambda_3} = v - \beta_0\left(\tau + \frac{b}{a}\right).
\end{gathered} \tag{29}$$

Eliminating v from (29), it follows that the quadratic equation with respect to $(1 + \beta_0\lambda_1)/2\lambda_3$ is

$$\begin{aligned}
& a_0\left(1 - a_0\beta_0^2\right)\left(\frac{\left(1 + \beta_0\lambda_1\right)}{2\lambda_3}\right)^2 \\
& \quad - 2a_0\beta_0^2\left(\tau + \frac{b}{a}\right)\left(\frac{\left(1 + \beta_0\lambda_1\right)}{2\lambda_3}\right) + \frac{1}{a} - \beta_0^2\left(\tau + \frac{b}{a}\right)^2 = 0.
\end{aligned} \tag{30}$$

Notice that $v > 0$ and $\lambda_1 \geq 0$; then from the second equation in (27), $\lambda_3 > 0$, this means that $(1 + \beta_0\lambda_1)/2\lambda_3 > 0$. Solving directly the quadratic equation above in $(1 + \beta_0\lambda_1)/2\lambda_3$, we have that

$$\begin{aligned}
& \frac{\left(1 + \beta_0\lambda_1\right)}{2\lambda_3} \\
& \quad = \frac{a_0\beta_0^2(\tau + b/a) + \sqrt{a_0\beta_0^2(\tau + b/a)^2 - (a_0/a)\left(1 - a_0\beta_0^2\right)}}{a_0\left(1 - a_0\beta_0^2\right)} \\
& \quad > 0.
\end{aligned} \tag{31}$$

Then the optimal solution of problem (RP2) can be obtained by substituting it in (27). This finishes the proof. □

We need that condition $a_0\beta_0^2 < 1$ holds in Lemma 6 and Theorem 7. This condition can in fact be easily attained. Notice that $\beta_0 = \sqrt{(1-\beta)/\beta}$. Thus, for any input data a_0, in order to have $a_0\beta_0^2 < 1$, we only require that β satisfies

$$\beta > \frac{a_0}{1+a_0}. \tag{32}$$

If $\beta \in (1/2, 1)$, the condition $a_0\beta_0^2 < 1$ can be satisfied while

$$\beta > \max\left\{\frac{1}{2}, \frac{a_0}{1+a_0}\right\}. \tag{33}$$

3. Two-Fund Separation Theorem

In our analysis of this section, we view $\mathbf{w}^*$ as a function of input parameter τ; that is, we denote it by $\mathbf{w}^* = \mathbf{w}(\tau)$. Let

$$S(\mathbf{w}) = \{\mathbf{w}(\tau) : \max\{\tau^*, 0\} \le \tau < \infty\}. \tag{34}$$

Then, set $S(\mathbf{w})$ is the solution space of problem (RP2). Now we are interested in the question that whether the solution of problem (RP2) satisfies the two-fund separation theorem or not.

Theorem 8. *Let $\mathbf{w}(\tau_1)$, $\mathbf{w}(\tau_2)(\tau_1 \neq \tau_2)$ be two solutions of mean-WCCVaR model; that is, $\mathbf{w}(\tau_1), \mathbf{w}(\tau_2) \in S(\mathbf{w})$. Then for any $\tau \in [\max\{\tau^*, 0\}, \infty)$ and the corresponding solution $\mathbf{w}(\tau)$, there exists a real number θ, such that*

$$\mathbf{w}(\tau) = \theta\mathbf{w}(\tau_1) + (1-\theta)\mathbf{w}(\tau_2); \tag{35}$$

that is, the two-fund separation theorem holds.

Proof. Noting $f(\tau)$ in Theorem 7, the optimal solution of mean-WCCVaR portfolio problem (RP2) can be written in the simple form

$$\mathbf{w}(\tau) = f(\tau)\mathbf{w}_A + (1-f(\tau))\mathbf{w}_\sigma, \tag{36}$$

where $\mathbf{w}_\sigma = (\Sigma^{-1}\mathbf{e})/a$ is the portfolio with the minimum variance and $\mathbf{w}_A = (\Sigma^{-1}\boldsymbol{\mu})/b$. For given $\tau_1, \tau_2 \in [\max\{\tau^*, 0\}, \infty)$, notice that $\tau_1 \neq \tau_2$; it follows that $f(\tau_1) \neq f(\tau_2)$. Then for any $\tau \in [\max\{\tau^*, 0\}, \infty)$, let

$$\theta = \frac{f(\tau) - f(\tau_2)}{f(\tau_1) - f(\tau_2)}; \tag{37}$$

we have

$$f(\tau) = \theta f(\tau_1) + (1-\theta) f(\tau_2). \tag{38}$$

Thus

$$\begin{aligned}\mathbf{w}(\tau) &= f(\tau)\mathbf{w}_A + (1-f(\tau))\mathbf{w}_\sigma \\ &= [\theta f(\tau_1) + (1-\theta) f(\tau_2)]\mathbf{w}_A \\ &\quad + [1 - [\theta f(\tau_1) + (1-\theta) f(\tau_2)]]\mathbf{w}_\sigma \\ &= \theta[f(\tau_1)\mathbf{w}_A + (1-f(\tau_1))\mathbf{w}_\sigma] \\ &\quad + (1-\theta)[f(\tau_2)\mathbf{w}_A + (1-f(\tau_2))\mathbf{w}_\sigma] \\ &= \theta\mathbf{w}(\tau_1) + (1-\theta)\mathbf{w}(\tau_2).\end{aligned} \tag{39}$$

This gets the desired conclusion. □

4. Efficient Frontier

We discuss the efficient frontier of optimal solution of problem (RP2) and analyze the relationship of the efficient frontier between problem (RP2) and mean-variance (MV) model. For any given parameter τ, clearly, the optimal solution $\mathbf{w}^* = \mathbf{w}(\tau)$ and the expected return of portfolio are the function of τ. Hence, we have

$$\begin{aligned}R(\tau) &= \frac{a_0\beta_0^2(\tau + b/a) + \sqrt{a_0\beta_0^2(\tau+b/a)^2 - (a_0/a)(1-a_0\beta_0^2)}}{(1-a_0\beta_0^2)} \\ &\quad + \frac{b}{a}.\end{aligned} \tag{40}$$

Rearranging this equality, we have the quadratic equation in $R(\tau)$ and τ

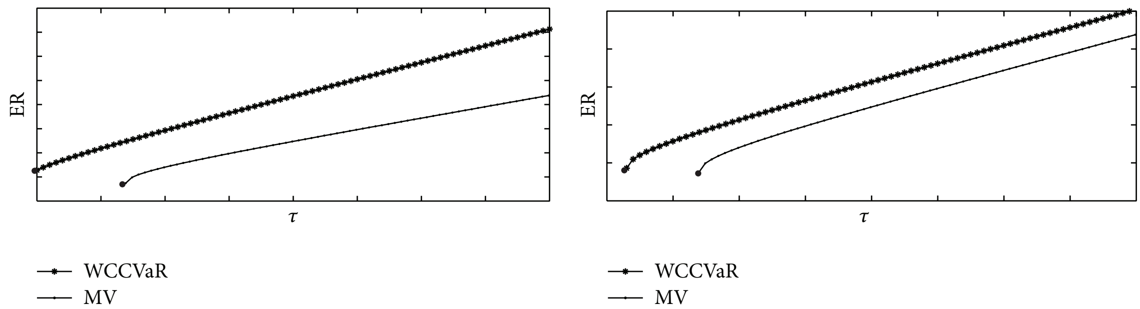
$$\begin{aligned}&\left(1-a_0\beta_0^2\right)^2\left(R(\tau) - \frac{b}{a}\right)^2 \\ &\quad - 2a_0\beta_0^2\left(1-a_0\beta_0^2\right)\left(R(\tau) - \frac{b}{a}\right)\left(\tau + \frac{b}{a}\right) \\ &\quad + a_0\beta_0^2\left(a_0\beta_0^2 - 1\right)\left(\tau + \frac{b}{a}\right)^2 + \frac{a_0}{a}\left(1-a_0\beta_0^2\right) = 0.\end{aligned} \tag{41}$$

The determinant of quadratic term coefficient can be expressed as

$$\begin{aligned}&\begin{vmatrix}\left(1-a_0\beta_0^2\right)^2 & -a_0\beta_0^2\left(1-a_0\beta_0^2\right) \\ -a_0\beta_0^2\left(1-a_0\beta_0^2\right) & a_0\beta_0^2\left(a_0\beta_0^2-1\right)\end{vmatrix} \\ &\quad = -a_0\beta_0^2\left(1-a_0\beta_0^2\right)^2 < 0.\end{aligned} \tag{42}$$

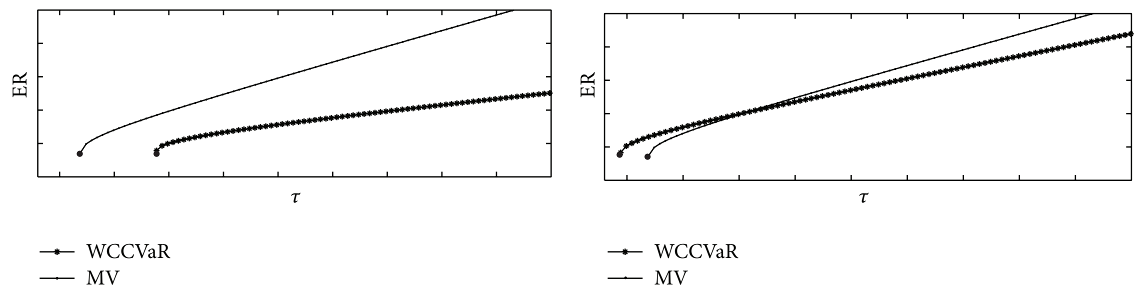
This means from the theory of quadratic curve that the efficient frontier determined by (41) is a branch of the hyperbola and the portfolio at the efficient frontier has the maximum expected return for given τ. The asymptotic line equation of efficient frontier is

$$R(\tau) = \frac{1+\sqrt{a_0\beta_0^2}}{1-a_0\beta_0^2}\left(\sqrt{a_0\beta_0^2}\tau + \frac{b}{a}\right), \tag{43}$$



(a) $\beta_0 > \beta_1$ (b) $\beta_0 = \beta_1$

Figure 1: The relationship of efficient frontiers of two models $\beta_0 \geq \beta_1$.



(a) $\beta_0 < \beta_1, \tau^* > \tau_{\mathrm{MV}}$ (b) $\beta_0 < \beta_1, \tau^* < \tau_{\mathrm{MV}}$

Figure 2: The relationship of efficient frontiers of two models in the case of $\beta_0 < \beta_1$.

the intercept at $R(\tau)$ axis is

$$R_b = \frac{b}{a} \cdot \frac{1 + \sqrt{a_0 \beta_0^2}}{1 - a_0 \beta_0^2}, \tag{44}$$

and the center is $(-b/a, b/a)$. Hence, the location of the hyperbola is determined by sign of b.

Now, we will discuss a relationship of efficient frontiers between the mean-variance (MV) model and the proposed mean-WCCVaR model. To this end, we compare the mean-WCCVaR model with the following MV model:

$$\max \left\{ E\left[\boldsymbol{\mu}^T \mathbf{w}\right] : \mathbf{w}^T \Sigma \mathbf{w} \leq \tau^2, \mathbf{e}^T \mathbf{w} = 1 \right\}. \tag{45}$$

It is not hard to compute that the optimal solution of MV model is

$$\begin{aligned} \mathbf{w}_{\mathrm{MV}}^* &= \mathbf{w}_{\mathrm{MV}}(\tau) \\ &= \left(1 - b\sqrt{\frac{\tau^2 - 1/a}{a_0}}\right) \frac{\Sigma^{-1}\mathbf{e}}{a} + \sqrt{\frac{\tau^2 - 1/a}{a_0}} \Sigma^{-1} \boldsymbol{\mu}. \end{aligned} \tag{46}$$

The corresponding expected value at the optimal solution is

$$R_{\mathrm{MV}}(\tau) = \boldsymbol{\mu}^T \mathbf{w}_{\mathrm{MV}}^* = \frac{ac - b^2}{a} \sqrt{\frac{\tau^2 - 1/a}{a_0}} + \frac{b}{a}. \tag{47}$$

In $(\tau, R_{\mathrm{MV}}(\tau))$ plane, function $R_{\mathrm{MV}}(\tau)$ plots the efficient frontier of MV:

$$a\tau_{\mathrm{MV}}^2 - \frac{a}{a_0}\left(R_{\mathrm{MV}}(\tau) - \frac{b}{a}\right)^2 = 1, \tag{48}$$

whose slope of asymptotic line is $\sqrt{a_0}$. Let

$$\frac{1 + \sqrt{a_0 \beta_0^2}}{1 - a_0 \beta_0^2} \sqrt{a_0 \beta_0^2} = \sqrt{a_0}. \tag{49}$$

Then,

$$\beta_0 = \frac{\sqrt{1 + 4\left(a_0 + \sqrt{a_0}\right)} - 1}{2\left(a_0 + \sqrt{a_0}\right)} =: \beta_1. \tag{50}$$

If $\beta_0 \geq \beta_1$, then for the case of $\tau^* < \tau_{\mathrm{MV}}$, two asymptotic lines are parallel and therefore efficient frontiers are not intersection; see Figure 1. If $\beta_0 < \beta_1$, then the slope of asymptotic line of efficient frontier for mean-WCCVaR model is less than that of MV model and two efficient frontiers can be intersected when $\tau^* < \tau_{\mathrm{MV}}$; see Figure 2(b).

Generally speaking, it is more conservative for mean-WCCVaR model than MV model. But the conservative performance can be improved by adjusted the confidence level β. For example, if β is small, such that $\beta_0 > \beta_1$, then mean-WCCVaR has the higher expectation return at the same τ; see Figure 1. This means that ER/τ is also higher; that is, the expectation return of per unit WCCVaR risk (reflected by parameter τ) is higher. The other case is that if we choose β, such that $\beta_0 < \beta_1$, then mean-WCCVaR model can obtain still the large expected return at the same τ if τ is not large; see Figure 2.

5. An Extension with Risky-Free Asset

We consider the portfolio with risky-free asset in this section and therefore the optimization problem can be expressed as

$$\begin{aligned} \max_{\mathbf{w},w_f} \quad & \mathbb{E}\left[\mathbf{r}^T\mathbf{w}+w_f r_f\right] \\ \text{s.t.} \quad & -\boldsymbol{\mu}^T\mathbf{w}-w_f r_f+\sqrt{\frac{\beta}{1-\beta}}\sqrt{\mathbf{w}^T\Sigma\mathbf{w}}\le\tau, \\ & \mathbf{e}^T\mathbf{w}+w_f=1. \end{aligned} \tag{RPF}$$

Clearly, $(\mathbf{w},w_f) = (\mathbf{0},1)$ is a strictly feasible solution of problem (RPF). Hence, for any $\tau>0$, problem (RPF) is always feasible. The following theorem gives the explicit solution of problem (RPF).

Theorem 9. *If*

$$\beta_0^2\left[a_0+\frac{\left(b-ar_f\right)^2}{a}\right]<1, \tag{51}$$

then problem (RPF) has the optimal solution

$$\mathbf{w}^* = \frac{\beta_0\left(\tau+r_f\right)}{\sqrt{a_0+\left(b-ar_f\right)^2/a}\left(1-\beta_0\sqrt{a_0+\left(b-ar_f\right)^2/a}\right)} \times\left(\Sigma^{-1}\boldsymbol{\mu}-r_f\Sigma^{-1}\mathbf{e}\right), \tag{52}$$

where

$$w_f = 1-\frac{\beta_0\left(\tau+r_f\right)\left(b-ar_f\right)}{\sqrt{a_0+\left(b-ar_f\right)^2/a}\left(1-\beta_0\sqrt{a_0+\left(b-ar_f\right)^2/a}\right)}. \tag{53}$$

Proof. Let $\nu=\sqrt{\mathbf{w}\Sigma\mathbf{w}}$; similar to (RP2), optimization problem (RPF) can be rewritten as

$$\begin{aligned} \max_{\mathbf{w},\nu>0} \quad & \mathbb{E}\left[\mathbf{r}^T\mathbf{w}+r_f w_f\right] \\ \text{s.t.} \quad & \nu-\beta_0\left(\boldsymbol{\mu}^T\mathbf{w}+r_f w_f\right)\le\beta_0\tau, \\ & \mathbf{e}^T\mathbf{w}+r_f w_f=1, \\ & \mathbf{w}^T\Sigma\mathbf{w}=\nu^2. \end{aligned} \tag{54}$$

Then, from KKT condition, the optimal solution $(\mathbf{w}^*,w_f^*)$ must satisfy the first-order condition

$$\begin{aligned} \left(1+\beta_0\lambda'_f\right)\boldsymbol{\mu}-\lambda''_f\mathbf{e}-2\lambda'''_f\Sigma\mathbf{w}^* &= 0, \\ -\lambda'_f+2\lambda'''_f\nu &= 0, \\ \mathbf{e}^T\mathbf{w}^* &= 1-w_f^*, \\ \left(\mathbf{w}^*\right)^T\Sigma\mathbf{w}^* &= \nu^2, \\ \nu-\beta_0\left(\boldsymbol{\mu}^T\mathbf{w}^*+r_f w_f^*\right) &= \beta_0\tau, \\ \left(1+\lambda'_f\beta_0\right)r_f-\lambda''_f &= 0, \end{aligned} \tag{55}$$

where $\lambda'_f\ge 0$, $\lambda''_f\in\mathbb{R}$ and $\lambda'''_f\in\mathbb{R}$ are Lagrangian multipliers. From the first and third equalities, we have

$$\mathbf{w}^* = \frac{\left(1+\beta_0\lambda'_f\right)}{2\lambda'''_f}\left(\Sigma^{-1}\boldsymbol{\mu}-\frac{b}{a}\Sigma^{-1}\mathbf{e}\right)+\frac{1-w_f}{a}\Sigma^{-1}\mathbf{e}. \tag{56}$$

Substituting $\mathbf{w}^*$ into the fourth and fifth equalities in (55), then

$$a_0\left[\frac{(1+\beta_0\lambda'_f)}{2\lambda'''_f}\right]^2+\frac{1-w_f}{a}=\nu^2, \tag{57}$$

$$a_0\beta_0\frac{\left(1+\beta_0\lambda'_f\right)}{2\lambda'''_f}+\beta_0\left(\tau+\frac{b}{a}\left(1-w_f\right)+r_f w_f\right)=\nu. \tag{58}$$

Combining (56), (57), and the third and sixth equations in system of equations (55), we have

$$\frac{\left(1+\beta_0\lambda'_f\right)}{2\lambda'''_f}=\frac{1-w_f}{b-ar_f}. \tag{59}$$

Eliminating ν and $1-w_f$ from (56), (57), and (58), we obtain a quadratic equation in $(1+\beta_0\lambda'_f)/2\lambda'''_f$:

$$\begin{aligned} &\left(a_0+\frac{\left(b-ar_f\right)^2}{a}\right)\left[1-\beta_0^2\left(a_0+\frac{\left(b-ar_f\right)^2}{a}\right)\right] \times\left(\frac{(1+\beta_0\lambda'_f)}{2\lambda'''_f}\right)^2 \\ &-2\beta_0^2\left(\tau+r_f\right)\left(a_0+\frac{\left(b-ar_f\right)^2}{a}\right)\left(\frac{\left(1+\beta_0\lambda'_f\right)}{2\lambda'''_f}\right) \\ &-\beta_0^2\left(\tau+r_f\right)^2=0. \end{aligned} \tag{60}$$

Table 1: The mean and covariance matrix of returns for the chosen 9 indexes.

	II	CI1	PI	UI	CI2	HS	N225	FTSE	DJIA
$\boldsymbol{\mu}$	0.1813	0.1875	0.1924	0.2214	0.1636	0.1018	−0.0096	0.0503	0.0849
	0.2676	0.2785	0.3052	0.2909	0.2416	0.0992	0.0210	0.0330	0.0363
	0.2785	0.3109	0.3246	0.2859	0.2533	0.0983	0.0184	0.0330	0.0301
	0.3052	0.3246	0.3924	0.3177	0.3063	0.1147	0.0309	0.0314	0.0324
	0.2909	0.2859	0.3177	0.3611	0.2287	0.0964	0.0081	0.0315	0.0391
Σ	0.2416	0.2533	0.3063	0.2287	0.3115	0.0881	0.0296	0.0370	0.0382
	0.0992	0.0983	0.1147	0.0964	0.0881	0.0950	0.0539	0.0332	0.0369
	0.0210	0.0184	0.0309	0.0081	0.0296	0.0539	0.0526	0.0248	0.0208
	0.0330	0.0330	0.0314	0.0315	0.0370	0.0332	0.0248	0.0262	0.0246
	0.0363	0.0301	0.0324	0.0391	0.0382	0.0369	0.0208	0.0246	0.0290

Table 2: The out-of-sample returns statistics of portfolios obtained by WCCVaR, VaR, MV, and equally weighted strategy ($1/N$), where $1/N$ is the equally weighted strategy. Std: standard deviation, tv: terminal value of wealth, and cv: coefficient of variation in this table.

Model	Mean	Std	Max.	Min.	tv	cv
WCCVaR	0.0377	0.0131	0.0707	−0.0002	1.0269	1.26%
VaR	0.0037	0.1169	0.2342	−0.1987	0.8141	11.65%
$1/N$	−0.1238	0.0759	0.0127	−0.2485	0.8369	8.66%
MV	0.0046	0.0986	0.0457	−0.1254	0.8433	8.53%

Solving directly this equation, we have that

$$\frac{\left(1+\beta_0\lambda'_f\right)}{2\lambda'''_f} = \frac{\beta_0\left(\tau+r_f\right)}{\sqrt{a_0+\left(b-ar_f\right)^2/a}\left(1-\beta_0\sqrt{a_0+\left(b-ar_f\right)^2/a}\right)}. \tag{61}$$

And further, we can obtain w_f from (58). Then, substituting $(1+\beta_0\lambda'_f)/2\lambda'''_f$ and w_f in (55), we get the results of this theorem. □

It is not hard to compute the relationship between the expectation return E_f of portfolio with risky-free asset and the parameter τ:

$$E_f(\tau) = \frac{\beta_0\left(c-ar_f^2-2br_f\right)}{A}\tau + \left[\frac{\beta_0\left(c-ar_f^2-2br_f\right)}{A}+1\right]r_f, \tag{62}$$

where $A = \sqrt{a_0+(b-ar_f)^2/a}(1-\beta_0\sqrt{a_0+(b-ar_f)^2/a})$ is independent of τ.

6. Numerical Results

We take five domestic risky assets, industrial index (II), commercial index (CI1), properties index (PI), utilities index (UI), and composite index (CI2), and four overseas risky assets, Hengsheng index (HS), Tokyo Nikkei-225 Index (N225), FTSE Index, and Dow-Jones industrial average index (DJIA). The time interval is from January 2, 1995, to December 31, 2012. The returns and covariance matrix of all risky assets can be found in Table 1. For simplicity, we take the risky-free annual interest rate is $r_f = 3\%$.

The rolling procedure is used to test the proposed model as follows.

(i) We estimate first the parameters $\boldsymbol{\mu}$ and Σ using the 15-year data from January 2, 1995, to December 31, 2009, and test the out-of-sample performance at the whole 2010 year.

(ii) And then we further estimate the parameters $\boldsymbol{\mu}$ and Σ using the next 15-year data from January 2, 1996, to December 31, 2010, and test the out-of-sample performance at the whole 2011 year.

(iii) We finally estimate the parameters $\boldsymbol{\mu}$ and Σ using the next 15-year data from January 2, 1997, to December 31, 2011, and test the out-of-sample performance at the whole 2012 year.

In our numerical reports, we compare our WCCVaR model with the classical VaR model under normal distribution assumption, equally weighted strategy [17], and MV model. Table 2 gives the results of three models with $\beta = 95\%$ and $\tau = 0.05$. The following observations can be found from Table 2.

(1) The standard deviation of portfolio obtained by WCCVaR model is clearly less than that of VaR model and equally weighted strategy [17]; moreover, the expected return of portfolio obtained by WCCVaR model is greater than that of VaR model and equally weighted strategy [17]. This means that the proposed

WCCVaR model has the better performance (such as Sharpe ratio).

(2) The real wealth of WCCVaR is at least 0.9998 that is very close to the initial wealth one and the coefficient of variation is only 1.26% which is far less than the coefficient of variation of VaR model and equally weighted strategy. The stable performance of WCCVaR is obvious.

(3) At the beginning period, within the 300 trade dates, the accumulation wealth of VaR and MV model is better than that of WCCVaR, but, after about 300 trade dates, the accumulation wealth of VaR and MV falls rapidly while WCCVaR still holds the stable wealth. We find that the 300th trade date is about corresponding to the first two months of 2011; the real market at that time is a bear market. This is the main reason that the accumulation wealth of VaR, equally weighted strategy and MV falls rapidly.

7. Conclusions

We discuss the worst-case CVaR risk measure without the distribution assumption and consider an application in robust portfolio selection problem. The explicit solution is obtained and two-fund separation theorem is proved for the solutions. The theoretical comparison with classical mean-variance model is first discussed by the efficient frontier and the numerical comparison with VaR, MV model and equally weighted strategy using domestic and overseas assets. The numerical results indicate that the proposed WCCVaR has the better expected return and smaller standard deviation than VaR, MV model and equally weighted strategy and therefore can obtain the better performance, such as Sharpe ratio.

Conflict of Interests

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

Acknowledgments

This work is supported by National Natural Science Foundations of China (71371090), Science Foundation of Ministry of Education of China (13YJCZH160), Natural Science Foundation of Jiangxi Province of China (20114BAB211008), and Jiangxi University of Finance and Economics Support Program Funds for Outstanding Youths.

References

[1] P. Artzner, F. Delbaen, and J.-M. Eber, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[2] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *The Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[3] D. Goldfarb and G. Iyengar, "Robust portfolio selection problems," *Mathematics of Operations Research*, vol. 28, no. 1, pp. 1–38, 2003.

[4] O. L. V. Costa and A. C. Paiva, "Robust portfolio selection using linear-matrix inequalities," *Journal of Economic Dynamics & Control*, vol. 26, no. 6, pp. 889–909, 2002.

[5] B. V. Halldórsson and R. H. Tütüncü, "An interior-point method for a class of saddle-point problems," *Journal of Optimization Theory and Applications*, vol. 116, no. 3, pp. 559–590, 2003.

[6] R. H. Tütüncü and M. Koenig, "Robust asset allocation," *Annals of Operations Research*, vol. 132, no. 1–4, pp. 157–187, 2004.

[7] Z. Lu, "Robust portfolio selection based on a joint ellipsoidal uncertainty set," *Optimization Methods and Software*, vol. 26, no. 1, pp. 89–104, 2011.

[8] A.-F. Ling and C.-X. Xu, "Robust portfolio selection involving options under a'marginal + joint" ellipsoidal uncertainty set," *Journal of Computational and Applied Mathematics*, vol. 236, no. 14, pp. 3373–3393, 2012.

[9] L. El Ghaoui, M. Oks, and F. Oustry, "Worst-case value-at-risk and robust portfolio optimization: a conic programming approach," *Operations Research*, vol. 51, no. 4, pp. 543–556, 2003.

[10] S. Zhu and M. Fukushima, "Worst-case conditional value-at-risk with application to robust portfolio management," *Operations Research*, vol. 57, no. 5, pp. 1155–1168, 2009.

[11] S. Zhu, D. Li, and S. Wang, "Robust portfolio selection under downside risk measures," *Quantitative Finance*, vol. 9, no. 7, pp. 869–885, 2009.

[12] D. Huang, S. Zhu, F. J. Fabozzi, and M. Fukushima, "Portfolio selection with uncertain exit time: a robust CVaR approach," *Journal of Economic Dynamics and Control*, vol. 32, no. 2, pp. 594–623, 2008.

[13] F. J. Fabozzi, D. Huang, and G. Zhou, "Robust portfolios: contributions from operations research and finance," *Annals of Operations Research*, vol. 176, no. 1, pp. 191–220, 2010.

[14] L. Chen, S. He, and S. Zhang, "Tight bounds for some risk measures, with applications to robust portfolio selection," *Operations Research*, vol. 59, no. 4, pp. 847–865, 2011.

[15] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms, Second Edition*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1993.

[16] I. Popescu, "Robust mean-covariance solutions for stochastic optimization," *Operations Research*, vol. 55, no. 1, pp. 98–112, 2007.

[17] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy?" *The Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 2009.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 365240, 11 pages
http://dx.doi.org/10.1155/2014/365240

Research Article

Terminal-Dependent Statistical Inference for the FBSDEs Models

Yunquan Song[1,2]

[1] China University of Petroleum, Qingdao 266580, China
[2] Shandong University Qilu Securities Institute for Financial Studies, Shandong University, Jinan 250100, China

Correspondence should be addressed to Yunquan Song; math1212@163.com

Received 12 March 2014; Accepted 27 May 2014; Published 25 June 2014

Academic Editor: Guangchen Wang

The original stochastic differential equations (OSDEs) and forward-backward stochastic differential equations (FBSDEs) are often used to model complex dynamic process that arise in financial, ecological, and many other areas. The main difference between OSDEs and FBSDEs is that the latter is designed to depend on a terminal condition, which is a key factor in some financial and ecological circumstances. It is interesting but challenging to estimate FBSDE parameters from noisy data and the terminal condition. However, to the best of our knowledge, the terminal-dependent statistical inference for such a model has not been explored in the existing literature. We proposed a nonparametric terminal control variables estimation method to address this problem. The reason why we use the terminal control variables is that the newly proposed inference procedures inherit the terminal-dependent characteristic. Through this new proposed method, the estimators of the functional coefficients of the FBSDEs model are obtained. The asymptotic properties of the estimators are also discussed. Simulation studies show that the proposed method gives satisfying estimates for the FBSDE parameters from noisy data and the terminal condition. A simulation is performed to test the feasibility of our method.

1. Introduction

Since 1973, when the world's first options exchange opened in Chicago, a large number of new financial products have been introduced to meet the customer's demands from the derivative markets. In the same year, Black and Scholes [1] provided their celebrated formula for option pricing and Merton [2] proposed a general equilibrium model for security prices. Since then, modern finance has adopted stochastic differential equations as its basic instruments for portfolio management, asset pricing, risk management, and so on. Among these models, the backward stochastic differential equations (BSDEs for short) are a desirable choice for hedging and pricing an option. Its general form is as follows:

$$dY_s = -g(s, Y_s, Z_s)\,ds + Z_s dB_s,$$
$$Y_T = \xi, \quad s \in [t, T], \tag{1}$$

where g is the generator, B_t is a Brownian motion, and ξ is a $\mathfrak{R}$-valued Borel function as the terminal condition. Usually the terminal condition is designed as a random variable with given distribution. If g meets certain conditions, the BSDE has a unique solution.

In terms of the backward equation, within a complete market, it serves to characterize the dynamic value of replicating portfolio Y_s with a final wealth ξ and a special quantity Z_s that depends on the hedging portfolio. In particular, while the generator consists of diffusion process, the corresponding equation is proved to be a forward-backward stochastic differential equation (FBSDE), which can be expressed as

$$dY_s = -g(s, X_s, Y_s, Z_s)\,ds + Z_s dB_s, \qquad Y_T = \xi, \tag{2}$$

where X_s satisfies the following ordinary stochastic differential equation (OSDE):

$$dX_s = \mu(s, X_s)\,dt + \sigma(s, X_s)\,dB_s, \quad s \in [t, T]. \tag{3}$$

Compared to the OSDE that contains an initial condition, the solution of the FBSDE is affected by the terminal condition $Y_T = \xi(X_T)$. As is well known, there exist a number of parametric and nonparametric methods to deal with estimation and test for the OSDE. However, these methods cannot be directly employed to infer the BSDE and FBSDE because the two models are related to a terminal condition. Forward-backward stochastic differential equations are used in biology systems, mathematical finance, insurance, real estate, multiagent, and network control. See Antonelli [3], Wang et al. [4], Zhang and Li [5], and so on.

For the FBSDE defined above, the statistical inference was investigated initially by Su and Lin [6] and Chen and Lin [7]. Furthermore, by financial and ecological problems, a relevant statistical model was proposed by Lin et al. [8]. However, they did not take the terminal condition into account in the inference procedure. In the framework of the FBSDE mentioned above, the terminal condition is additional, which is not nested into the equation. Thus, there is an essential difficulty to use the terminal condition to refine the inference procedure.

As a result, their methods fail to cover the full problems given in the FBSDE. Zhang and Lin [9] proposed two terminal-dependent estimation methods via terminal control variable for the integral form models of FBSDE. However, they only considered the parametric form of the generator g in their paper.

This paper intends to explore the method to fulfill the terminal-dependent inference: quasi-instrumental variable methods. It is worth mentioning that the key point of our method is the use of the terminal condition information rather than neglecting it. This change leads to a completely new work among the existing researches. The key technique in our method is the use of quasi-instrumental variable which is similar but not the same as instrumental variable (IV). It is known that IV is widely employed in applied econometrics to achieve identification and carry out estimation and inference in the model containing endogenous explanatory variables or panel data; see Hsiao [10] for an overview of the relevant statistical inference and econometric interpretation and see Hall and Horowitz [11] for recent work on nonparametric instrumental variable estimation.

Through the backward equation (2) of FBSDE, we get a regression model. To use the terminal condition information, we put the terminal condition as a quasi-instrumental variable and introduce it into our model. However, when a constraint is appended artificially, the original model may change to be biased in the sense of $E(Z_s dB_s \mid X_s, \xi) \neq 0$, because the constraint condition influences the increase trend of wealth so that $Z_s dB_s$ may deviate from the original center zero; in other words, due to the constraint, the trajectory of Y_s may departure from the original expectation so that $Z_s dB_s$ cannot be regarded as error. Therefore, some problems arise naturally, including how to correct the bias of the model and how to construct the constraint-dependent estimation. To solve these problems, we will use remodeling method to draw terminal condition into differential equation, similar but not the same as IV, called quasi-instrumental variable methods; in other words, the terminal condition ξ enters into the equation as a control variable. This remodeling method takes advantage of the terminal information naturally, and the estimator performs quite well.

We use the nonparametric form of the generator g in model (2) because the correct FBSDEs model for a specific topic can neither be provided automatically by financial market nor be derived from theory of mathematical finance, and in lack of prior information about the structure of a model, nonparametric inference can provide a flexible as well as robust description of a data-generating process. Even in some cases when parametric models are available, nonparametric methods are still employed to avoid the model misspecification that may lead to large errors in option pricing and other problems from financial market. So we adopt the nonparametric form that can endow the model (2) with flexibility and robustness.

Note that Z_s is usually unobservable and g cannot be completely specified in the financial market. The problems of interest are therefore to give both proper estimations of the generator g and the process Z_s based on the observed data (X_s, Y_s) and the terminal expectation ξ.

The remainder of the paper is organized as follows. In Section 2, the FBSDE is rebuilt as a nonparametric model that contains the terminal condition as a quasi-instrumental variable. Consequently, a terminal-dependent estimation procedure is proposed. Next we discuss the asymptotic properties of the newly proposed estimations in Section 3. Simulation study is proposed in Section 4 to illustrate our methods. The proofs of the theorems are presented in Appendix.

2. Model and Method

In this section, we propose a nonparametric estimator with the help of quasi-instrumental variable.

2.1. Model and Its Statistical Version. We begin the following original model by combining (2)-(3):

$$\begin{aligned} dY_s &= -g(s, X_s, Y_s, Z_s)\, ds + Z_s dB_s, \quad Y_T = \xi, \\ dX_s &= \mu(s, X_s)\, dt + \sigma(s, X_s)\, dB_s, \quad s \in [t, T], \end{aligned} \tag{4}$$

where B_t is the standard Brownian motion and ξ is a $\mathfrak{R}$-valued Borel function. Here the generator g is a function of s, X_s, Y_s, and Z_s. For the FBSDEs model (4), only one of the backward components, Y_s, and the forward components, X_s, can be observed. Another backward component Z_s is totally unobservable. Furthermore, the adapted process Z_s and terminal condition could be indicated as a function of X_s.

In this section, we present the statistical structure of FBSDEs by taking advantage of quasi-instrumental variable and obtain the consistent asymptotically normal estimators of g and Z_s based on observed data $\{X_s, Y_s\}$ and the terminal condition ξ.

2.2. Remodeling for Model (4). To construct terminal-dependent estimation for the generator g and process Z_s, the key technique is how to plug the terminal condition into the equation. When ξ is plugged into the model, we call it the quasi-IV, similar but not the same as IV. Evidently, the property of Brownian motion shows that $E(Z_s dB_s \mid X_s) = 0$, but $E(Z_s dB_s \mid X_s, \xi) \neq 0$, which means drawing the terminal control directly into the equation as the condition should not be encouraged at the cost of model bias. Rewriting the first equation of (4) enables us to construct an unbiased model:

$$dY_s = -g(s, X_s, Y_s, Z_s)\, ds + m(X_s, \xi) + U_s, \quad (5)$$

where $m(X_s, \xi) = E(Z_s dB_s \mid X_s, \xi)$, $U_s = Z_s dB_s - m(X_s, \xi)$, and $E(U_s \mid X_s, \xi) = 0$. The newly defined model (5), together with the second equation in (4), can be thought of as a quasi-IV FBSDE. Because the equation in (5) contains the terminal condition ξ, we can construct the terminal-dependent estimation. From the above definitions, we see that, by bias correction, the original model changes to be an additive nonparametric model with nonparametric components $-g(s, X_s, Y_s, Z_s)ds$ and $m(X_s, \xi)$. It shows that when terminal condition is regarded as a quasi-IV and then appended to the model, the result model is unbiased and changes to be nonparametric additive model.

2.3. Estimation for Z_s. Before estimating the model function $m(x_s, \xi)$ and the generator g, we need to estimate Z_s firstly because Z_s is unobservable and it will be seen that the estimators of the model function $m(x_s, \xi)$ and the generator g depend on Z_s. Since the distribution of ξ is supposed to be known, let $\{\xi_i, 1 \le i \le k\}$ for $k \ge 1/\Delta$ be a sample of ξ. Suppose that, for each terminal data ξ_j and equally spaced time points $\{s_i = s_1 + (i-1)\Delta, i = 1, \ldots, n\} \subseteq [0, T]$, we record the observed time series data:

$$\begin{aligned} &\left\{X_{s_i,j}, Y_{s_i,j},\ i = 1, \ldots, n, j = 1, \ldots, k\right\} \\ &\quad = \left\{X_{i,j}, Y_{i,j},\ i = 1, \ldots, n, j = 1, \ldots, k\right\}. \end{aligned} \quad (6)$$

At any time point $s \in [t, T]$, $Z_s^{t,x}$, denoting Z_s and satisfying the initial condition (t, x), is a determined function of $X_s^{t,x}$. As was shown by Su and Lin [6] and Chen and Lin [7], we can adopt a difference-based method to approximate Z^2 as

$$\left(Z_s^{t,X_t}\right)^2 = \frac{1}{\Delta} E\left(Y_{s+\Delta}^{t+\Delta, X_t+\Delta} - Y_s^{t,X_t} \mid X_t, t\right)^2 + O(\Delta). \quad (7)$$

It shows that the numerical approximation error to Z_t^2 converges to zero at rate of order $O_p(\Delta)$.

For each ξ_j, if Z_t depends on t only via variable X_t, by (7) and N-W kernel estimation method, we estimate Z_t^2 at x_0 by

$$\widehat{Z}^2_{x_0,j} = \frac{\sum_{i=1}^{n-1} \Delta^{-1} \left(Y_{i+1,j} - Y_{i,j}\right)^2 K_{h_X}\left(X_{i,j} - x_0\right)}{\sum_{i=1}^{n-1} K_{h_X}\left(X_{i,j} - x_0\right)}. \quad (8)$$

Otherwise, we estimate Z_t^2 at (x_0, t_0) by

$$\begin{aligned} &\widehat{Z}^2_{x_0,t_0,j} \\ &\quad = \frac{\sum_{i=1}^{n-1} \Delta^{-1} \left(Y_{i+1,j} - Y_{i,j}\right)^2 K_{h_X}\left(X_{i,j} - x_0\right) K_{h_t}\left(t_i - t_0\right)}{\sum_{i=1}^{n-1} K_{h_X}\left(X_{i,j} - x_0\right) K_{h_t}\left(t_i - t_0\right)}, \end{aligned} \quad (9)$$

where $K_{h_X} = K(\cdot/h_X)/h_X$ and $K_{h_t} = K(\cdot/h_t)/h_t$, $K(\cdot)$ are regular kernel functions, with h_X and h_t being the corresponding bandwidths.

2.4. Estimation for $m(X_s,\xi)$. After plugging the estimator $\widehat{Z}_s$ into model (5), we still need to consider inference of $m(x_s, \xi)$. As we all know, the nonparametric function $m(X_s, \xi)$ in (5) can be acquired as $m(X_s, \xi) = E(dY_s + g(s, X_s, Y_s, Z_s)ds \mid X_s, \xi)$. We note that $g(s, X_s, Y_s, Z_s)ds$ is a higher order infinitesimal of $Z_s dB_s$ when Δ tends to zero. Under this situation, if $g(s, X_s, Y_s, Z_s)ds$ is ignored, then

$$m(X_s, \xi) \doteq E(dY_s \mid X_s, \xi). \quad (10)$$

It implies that we can use ordinary nonparametric method to estimate function m. For example, we use the N-W ordinary nonparametric method to estimate $m(X_s, \xi)$ valued at (x_0, ξ_0):

$$\begin{aligned} &\widehat{m}(x_0, \xi_0) \\ &\quad = \frac{\sum_{i=1}^{n-1} \sum_{j=1}^{m} \left(Y_{i+1,j} - Y_{i,j}\right) K_{h_X}\left(X_{i,j} - x_0\right) K_{h_\xi}\left(\xi_{i,j} - \xi_0\right)}{\sum_{i=1}^{n-1} \sum_{j=1}^{m} K_{h_X}\left(X_{i,j} - x_0\right) K_{h_\xi}\left(\xi_{i,j} - \xi_0\right)}, \end{aligned} \quad (11)$$

where $K_{h_X} = K(\cdot/h_X)/h_X$ and $K_{h_\xi} = K(\cdot/h_\xi)/h_\xi$ are regular kernel functions, with h_X and h_ξ being the corresponding bandwidths.

2.5. Estimation for Generator g. As was shown in the nonparametric instrumental variables estimator of Hall and Horowitz [11] (hereinafter HH), we can adopt a nonparametric quasi-instrumental variables estimation to estimate the generator g. So in the section we summarize the HH estimator of g in the model:

$$E\left[-dY_t - m(X_t, \xi) \mid X_t, \xi\right] = E\left[g(t, X_t, Y_t, Z_t)\, dt \mid X_t, \xi\right]. \quad (12)$$

Since $\widehat{m}(x_0, \xi_0)$ and $\widehat{Z}^2_{x_0,j}$ are the consistent estimator of $m(x_0, \xi_0)$ and $Z^2_{x_0,j}$, respectively, we use them instead of $m(X_s, \xi)$ and Z_s in the above model and we get

$$\begin{aligned} &E\left[-dY_s - \widehat{m}(X_s, \xi) \mid X_s, \xi\right] \\ &\quad = E\left[g\left(s, X_s, Y_s, \widehat{Z}_s\right) dt \mid X_s, \xi\right]. \end{aligned} \quad (13)$$

Because $\widehat{Z}_s$ is function of X_s and Y_s, for simplicity of presentation, we denote $g(s, X_s, Y_s, \widehat{Z}_s) = g(X_s, Y_s)$. Thus, the model becomes

$$E\left[-dY_s - \widehat{m}(X_s, \xi) \mid X_s, \xi\right] = E\left[g(X_s, Y_s)\, dt \mid X_s, \xi\right]. \quad (14)$$

Let $\mathbb{Y}_i = ((Y_{i+\Delta} - Y_i) - \widehat{m}(X_i, \xi))/\Delta$, $\mathbb{X}_i = X_i$, $\mathbb{Z}_i = Y_i$, $\mathbb{W} = \xi$, and $\mathbb{U}_i = V_i/\sqrt{\Delta}$; the model becomes

$$\mathbb{Y}_i = g(\mathbb{X}_i, \mathbb{Z}_i) + \mathbb{U}_i, \qquad E(\mathbb{U}_i \mid \mathbb{X}_i, \mathbb{W}_i) = 0. \quad (15)$$

It is assumed that the support of $(\mathbb{X}, \mathbb{Z}, \mathbb{W})$ is contained in $[0,1]^3$. This assumption can always be satisfied by, if necessary, carrying out monotone increasing transformations of $\mathbb{X}$, $\mathbb{Z}$, and $\mathbb{W}$. For example, one can replace $\mathbb{X}$, $\mathbb{Z}$, and $\mathbb{W}$ by

$\Phi(\mathbb{X})$, $\Phi(\mathbb{Z})$, and $\Phi(\mathbb{W})$, where Φ is the normal distribution function. We take $(\mathbb{Y}, \mathbb{X}, \mathbb{Z}, \mathbb{W}, \mathbb{U})$ to be a vector, where $\mathbb{Y}$ and $\mathbb{U}$ are scalars, $\mathbb{X}$ and $\mathbb{W}$ are supported on $[0,1]$, and $\mathbb{Z}$ is supported on $[0,1]$. The model is

$$\mathbb{Y}_i = g(\mathbb{X}_i, \mathbb{Z}_i) + \mathbb{U}_i, \qquad E(\mathbb{U}_i \mid \mathbb{Z}_i, \mathbb{W}_i) = 0, \tag{16}$$

where $(\mathbb{Y}_i, \mathbb{X}_i, \mathbb{Z}_i, \mathbb{W}_i, \mathbb{U}_i)$, for $i \geq 1$, are independent and identically distributed as $(\mathbb{Y}, \mathbb{X}, \mathbb{Z}, \mathbb{W}, \mathbb{U})$. Thus, $\mathbb{X}$ and $\mathbb{Z}$ are endogenous and exogenous explanatory variables, respectively. Data $(\mathbb{Y}_i, \mathbb{X}_i, \mathbb{Z}_i, \mathbb{W}_i, \mathbb{U}_i)$, for $1 \leq i \leq n$, are observed.

Let $f_{\mathbb{XZW}}$ denote the density of $(\mathbb{X}, \mathbb{Z}, \mathbb{W})$, write $f_{\mathbb{Z}}$ for the density of $\mathbb{Z}$, and, for each $x_1, x_2 \in [0,1]^p$, and put

$$t_z(x_1, x_2) = \int f_{\mathbb{XZW}}(x_1, z, w) f_{\mathbb{XZW}}(x_2, z, w)\, dw. \tag{17}$$

Define the operator T_z on $L_2[0,1]^p$ by

$$(T_z \psi)(x) = \int t_z(\xi, x) \psi(\xi)\, d\xi. \tag{18}$$

It may be proved that, for each z for which T_z^{-1} exists,

$$\begin{aligned} g(x,z) &= f_{\mathbb{Z}}(z) E_{\mathbb{W}|\mathbb{Z}} \\ &\quad \times \left\{ E(\mathbb{Y} \mid \mathbb{Z} = z, \mathbb{W}) \left(T_z^{-1} f_{\mathbb{XZW}}\right)(x, z, \mathbb{W}) \mid \mathbb{Z} = z \right\}, \end{aligned} \tag{19}$$

where $E_{\mathbb{W}|\mathbb{Z}}$ denotes the expectation with respect to the distribution of $\mathbb{W}$ conditional on $\mathbb{Z}$. In this formulation, $(T_z^{-1} f_{\mathbb{XZW}})(x, z, \mathbb{W})$ denoted the result of applying T_z^{-1} to the function $f_{\mathbb{XZW}}(\cdot, z, \mathbb{W})$ and evaluating the resulting function at x.

To construct an estimator of $g(x,z)$, given $h > 0$ and $x = x^{(1)}$ and $\xi = \xi^{(1)}$, let $K_h(x,\xi) = K_h(x^{(j)}, \xi^{(j)})$, put $K_h(z,\xi)$ analogously for z and ξ, let $h_x, h_z > 0$, and define

$$\begin{aligned} &\widehat{f}_{\mathbb{XZW}}(x,z,w) \\ &\quad = \frac{1}{n h_x^2 h_z} \sum_{i=1}^{n} K_{h_x}(x - \mathbb{X}_i, x) K_{h_z}(z - \mathbb{Z}_i, z) K_{h_x}(w - \mathbb{W}_i, w), \\ &\widehat{f}^{-i}_{\mathbb{XZW}}(x,z,w) \\ &\quad = \frac{1}{(n-1) h_x^2 h_z} \sum_{1 \leq j \leq n: j \neq i} K_{h_x}(x - \mathbb{X}_j, x) \\ &\qquad\qquad \times K_{h_z}(z - \mathbb{Z}_j, z) K_{h_z}(w - \mathbb{W}_j, w), \\ &\widehat{t}_z(x_1, x_2) = \int \widehat{f}_{\mathbb{XZW}}(x_1, z, w) \widehat{f}_{\mathbb{XZW}}(x_2, z, w)\, dw, \\ &\left(\widehat{T}_z \psi\right)(x, z, w) = \int \widehat{t}_z(\xi, x) \psi(\xi, z, w)\, d\xi, \end{aligned} \tag{20}$$

where ψ is a function from R^3 to a real line. Then the estimator of $g(x,z)$ is

$$\widehat{g}(x,z) = \frac{1}{n} \sum_{i=1}^{n} \left(\widehat{T}_z^{+} \widehat{f}^{-i}_{\mathbb{XZW}}\right)(x, z, \mathbb{W}_i) Y_i K_{h_z}(z - \mathbb{Z}_i, z). \tag{21}$$

3. Asymptotic Results

In this section, we study the asymptotic properties of our proposed estimators. All proofs are presented in Appendix.

3.1. Asymptotic results of $\widehat{Z}_s$. To give the asymptotic results of $\widehat{Z}_s$, we need the following conditions.

(a) $X_1, \ldots, X_n$ are ρ-mixing dependent; namely, the ρ-mixing coefficients $\rho(l)$ satisfy $\rho(l) \to 0$ as $l \to \infty$, where

$$\rho(l) = \sup_{E(X_{i+l}X_i) - E(X_{i+l})E(X_i) \neq 0} \frac{\left|E(X_{i+l}X_i) - E(X_{i+l})E(X_i)\right|}{\sqrt{\operatorname{Var}(X_{i+l}) \operatorname{Var}(X_i)}} \tag{22}$$

with $X_i = X(t_i)$.

(b) $|Z_i| \leq C$ (a. s.) uniformly for $i = 1, \ldots, n$, where C is a positive constant and $Z_i = Z(t_i)$.

(c) The continuous kernel function $K(\cdot)$ is symmetric about 0, with a support of interval $[-1,1]$, and

$$\begin{gathered} \int_{-1}^{1} K(u)\, du = 1, \qquad \sigma_K^2 = \int_{-1}^{1} u^2 K(u)\, du \neq 0, \\ \int_{-1}^{1} |u|^j K^k(u)\, du < \infty \quad \text{for } j \leq k = 1, 2. \end{gathered} \tag{23}$$

Condition (a) is commonly used for weakly dependent process; see, for example, Kolmogorov and Rozanov [12], Bradley and Bryc [13], Lu and Lin [14], and Su and Lin [6]. Condition (b) is also reasonable because, as is shown by (10), Z_t can be regarded as the deviation between the adjacent two observations. Condition (c) is standard for nonparametric kernel function.

Theorem 1. *Besides conditions (a), (b), and (c), let $\{X_1, \ldots, X_n\}$ be an observation sequence on a stationary ρ-mixing Markov process with the ρ-mixing coefficients satisfying $\rho(l) = \rho^l$ for $0 < \rho < 1$. Furthermore, $X_1, \ldots, X_n$ have a common and probability density $p(x)$, and for each interior point x_0 in the support of $p(\cdot)$, $p(x_0) > 0$, $Z^2(x_0) > 0$, the functions $p(x)$ and $Z(x)$ have continuous two derivatives in neighborhood of x_0. As $n \to \infty$, such that $nh \to \infty$, $nh^5 \to 0$, and $nh\Delta^2 \to 0$, then*

$$\sqrt{nh}\left(\widehat{Z}^2(x_0) - Z^2(x_0)\right) \xrightarrow{d} \left(0, \frac{Z^4(x_0) J_K}{p(x_0)}\right), \tag{24}$$

where $J_K = \int_{-1}^{1} K^2(u) du < \infty$.

The asymptotic result in Theorem 1 is standard for nonparametric kernel estimator and here undersmoothing is used to eliminate asymptotic bias.

3.2. Asymptotic results of $\widehat{g}(x,z)$. This section gives conditions under which the HH estimator of the generator

g is asymptotically distributed as $N(0, I)$. The following additional notations are used.

Define $\mathbb{U}_i = \mathbb{Y}_i - g(\mathbb{X}_i, \mathbb{Z}_i)$, $S_{n1}(x,z) = n^{-1}\sum_{i=1}^{n} \mathbb{U}_i \widehat{T}^{+} \widehat{f}_{\mathbb{XZW}}^{(-i)}(x,z,\mathbb{W}_i) K_{q,h_z}(z-\mathbb{Z}_i, z)$, and $S_{n2}(x,z) = n^{-1}\sum_{i=1}^{n} g(\mathbb{X}_i,\mathbb{Z}_i)\widehat{T}^{+}\widehat{f}_{\mathbb{XZW}}^{(-i)}(x,z,\mathbb{W}_i)K_{q,h_z}(z - \mathbb{Z}_i, z)$. Then, $\widehat{g}(x,z) = S_{n1}(x,z) + S_{n2}(x,z)$. Define $T^{+} = (T + a_n I)^{-1}$. Write

$$\begin{aligned} S_{n1}(x,z) &= n^{-1}\sum_{i=1}^{n} \mathbb{U}_i \left(T^{+} f_{\mathbb{XZW}}\right)(x,z,\mathbb{W}_i) K_{q,h_z}(z-\mathbb{Z}_i,z) \\ &\quad + n^{-1}\sum_{i=1}^{n} \mathbb{U}_i \left(\widehat{T}^{+}\widehat{f}_{\mathbb{XZW}}^{(-i)} - T^{+} f_{\mathbb{XZW}}\right) \\ &\qquad \times (x,z,\mathbb{W}_i) K_{q,h_z}(z-\mathbb{Z}_i,z) \\ &= S_{n11}(x,z) + S_{n12}(x,z). \end{aligned} \tag{25}$$

Define $V_n(x,z) = n^{-1}\operatorname{Var}[\mathbb{U}(T^{+} f_{\mathbb{XZW}})(x,z,\mathbb{W})]$. It follows from a triangular array version of the Lindeberg-Levy central limit theorem that $S_{n11}(x,z)/\sqrt{V_n(x,z)} \longrightarrow^{d} N(0,1)$ as $n \to \infty$. Therefore, $[\widehat{g}(x,z) - g(x,z)]/\sqrt{V_n(x,z)} \longrightarrow^{d} N(0,1)$ if $[S_{n12}(x,z) + S_{n2}(x,z) - g(x,z)]/\sqrt{V_n(x,z)} = o_p(1)$.

Assumption 2. The data $\mathbb{Y}_i$, $\mathbb{X}_i$, $\mathbb{Z}_i$, $\mathbb{W}_i$ are independently and identically distributed as $(\mathbb{Y},\mathbb{X},\mathbb{X},\mathbb{W})$, where $(\mathbb{X},\mathbb{Z},\mathbb{W})$ is supported on $[0,1]^3$ and $E[\mathbb{Y} - g(\mathbb{X},\mathbb{Z}) \mid \mathbb{W},\mathbb{Z}] = 0$.

Assumption 3. The distribution of $(\mathbb{X},\mathbb{Z},\mathbb{W})$ has a density $f_{\mathbb{XZW}}$ with respect to Lebesgue measure. Moreover, $f_{\mathbb{XZW}}$ is r times differentiable with respect to any combination of its arguments, where derivatives at the boundary of $[0,1]^3$ are defined as one sided derivatives. The derivatives are bounded in absolute value by C. In addition, g is r times differentiable on $[0,1]^2$ with derivatives at 0 and 1 defined as one sided. The derivatives of g are bounded in absolute value by C. In addition, $E[\mathbb{Y}^2 \mid \mathbb{X},\mathbb{Z},\mathbb{W}] \le C$ and $E[\mathbb{Y}^2 \mid \mathbb{X},\mathbb{Z},\mathbb{W}] \le C$, and $E[\mathbb{U}^2 \mid \mathbb{Z},\mathbb{W}] \ge C_U$ for some finite constant C_U.

Assumption 4. The constants α and β satisfy $\alpha > 1$, $\beta > 1/2$, and $\beta - 1/2 \le \alpha < 2\beta$. Moreover, $b_j \le Cj^{-\beta}$, $j^{-\alpha} \le C\lambda_j$, and $\sum_{k=1}^{\infty} |d_{zjk}| \le Cj^{-\alpha/2}$ for all $j \ge 1$. In addition, there are finite strictly positive constants, $C_{\lambda 1}$ and $C_{\lambda 2}$, such that $C_{\lambda 1} \le \lambda_j \le C_{\lambda 2} j^{-\alpha}$ for all $j \ge 1$.

Assumption 5. The tuning parameters a_n and h satisfy $a_n \asymp n^{-(\rho\alpha)/(2\beta+\alpha)}$ and $h \asymp n^{-1}$, where $r \in [A_2', A_3']$.

Assumption 6. K_h denotes a generalized kernel function, with the properties $K_h(u,t) = 0$ if $u > t$ or $u < t-1$, for all $t \in [0,1]h^{-(j+1)}\int_{t-1} tu^j K_h(u,t)du = 1$ if $j = 0$, else 0 if $1 \le j \le r-1$. For each $\xi \in [0,1]$, $K_h(h,\xi)$ is supported on $[(\xi-1)/h, \xi/h] \cap \kappa$, where κ is a compact interval not depending on ξ. Moreover,

$$\sup_{h>0,\xi\in[0,1],u\in\kappa} K_h(hu,\xi) \mid < \infty. \tag{26}$$

Assumption 7. Consider $E_{\mathbb{W}}[T^{+} f_{\mathbb{XZW}}(x,z,\mathbb{W})]^2 \asymp E_{\mathbb{W}}[T^{+} f_{\mathbb{XZW}}(\cdot,\cdot,\mathbb{W})]^2$ and $E_{\mathbb{W}}[T^{+} f_{\mathbb{XZW}}(\cdot,\cdot,\mathbb{W})]^2 \asymp \int_0^1 \|T^{+} f_{\mathbb{XZW}}(\cdot,\cdot,\mathbb{W})\|^2 dw$.

Theorem 8. *Let Assumptions 2–7 hold. Then*

$$\frac{\widehat{g}(x,z) - g(x,z)}{\sqrt{V_n(x,z)}} \longrightarrow^{d} N(0,I) \tag{27}$$

holds except, possibly, on a set of (x,z) values whose Lebesgue is 0.

Corollary 9. *Let Assumptions 2–7 hold. And if $V_n(x,z)$ is replaced with the consistent estimator,*

$$\widehat{V}_n(x,z) = n^{-1}\sum_{i=1}^{n} \widehat{\mathbb{U}}_i^2 \left[\widehat{T}^{+}\widehat{f}_{xw}^{-i}(z,\mathbb{W}_i) K_{q,h_z}(z-\mathbb{Z}_i,z)\right]^2, \tag{28}$$

where $\widehat{\mathbb{U}}_i = \mathbb{Y}_i - \widehat{g}(\mathbb{X}_i,\mathbb{Z}_i)$. This yields the Studentized statistic $[\widehat{g}(x,z) - g(x,z)]/\sqrt{\widehat{V}_n(x,z)}$. Then

$$\frac{\widehat{g}(x,z) - g(x,z)}{\sqrt{\widehat{V}_n(x,z)}} \longrightarrow^{d} N(0,I) \tag{29}$$

holds except, possibly, on a set of (x,z) values whose Lebesgue is 0.

As was shown in the remark given in the previous section, even the conditional mean of error of the model is nonzero, and the newly proposed estimation is consistent because of the mixing dependency; for details see the proof of Theorem 8. Furthermore, because of the terminal condition, the asymptotic variance is larger than that without the use of the terminal condition.

4. Simulation Studies

In this section, we investigate the finite-sample behaviors by simulation.

Example 10. We consider a simple FBSDE as

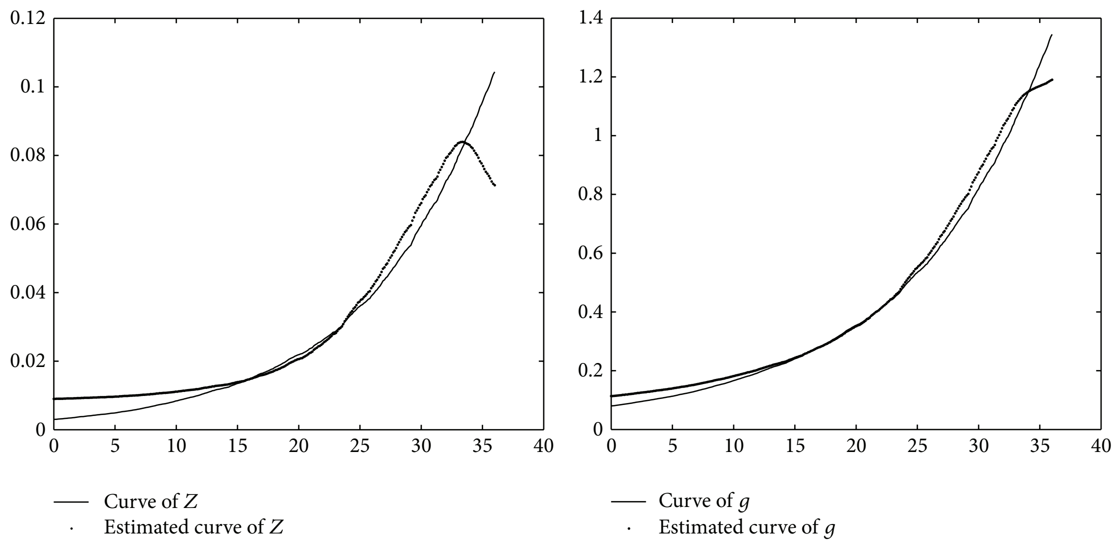
$$\begin{aligned} dY_t &= \left(\frac{\mu - r}{\sigma} Z_t + rY_t\right) dt + Z_t dB_t \\ &\triangleq \left(bY_t + cZ_t\right) + Z_t dB_t; \quad Y_T = \xi, \end{aligned} \tag{30}$$

where X_t is Geometric Brownian motion for modeling stock price satisfying

$$dX_t = \mu X_t dt + \sigma X_t dB_t; \quad X_0 = x, \tag{31}$$

while the riskless asset is the same as formula (31); $dP_t = rP_0 dt$.

Firstly, let $\mu = 0.1$, $\sigma = 0.01$, $\Delta = 0.12$, $n = 300$, $T = 36.6$, and $n_0 = n_1 = 10$. Obviously, $Z_t = n_1\sigma X_t$. We adopt Epanechnikov kernel defined by $K(u) = 3/4(1-u^2)I(|u| \le 1)$,



(a)

(b)

Figure 1: The real lines are the true curves of Z_t and function $g(t)$, respectively, and the dashed ones are estimated curves for them in Example 10.

where $I(\cdot)$ is the indicator function. For bandwidth selection, various data-driven techniques have been developed, such as cross-validation, the plug-in method, and the empirical bias method. However, these useful tools require additional computation intensiveness. In our simulation, we simply apply the rule of thumb bandwidth selector. For bandwidth selection, bandwidth $h = \text{std}(x)n^{-1/5}$. The values of the tuning parameters are $a_n = 0.05$, $\alpha = 1.2$, $\beta = 1$. Figure 1 presents the estimated curves for diffusion Z_t and drift g by one simulation.

Example 11. According to the theory of mathematical finance, we represent a European call option by the following FBSDEs model:

$$\begin{aligned} dX_s &= bX_s ds + \sigma X_s dW_s, \\ dY_s &= \left[rY_s + (b-r)\sigma^{-1}Z_s\right]ds + Z_s dW_s, \\ X_0 &= x, \quad Y_T = (X_T - K)^+, \quad s \in [0,T]. \end{aligned} \tag{32}$$

Here $\{X_s\}_{0\le s\le T}$ and $\{Y_s\}_{0\le s\le T}$ are the price processes of the stock and the option, respectively, and K is the striking price at the expiration date T. $\{X_s\}_{0\le s\le T}$ follows the geometric Brownian motion as

$$\begin{aligned} dX_s &= bX_s ds + \sigma X_s dW_s, \\ X_0 &= x, \quad s \in [0,T]. \end{aligned} \tag{33}$$

We use the Euler scheme to generate the price series of the stock as

$$X_{i+1} - X_i = bX_i\Delta + \sigma X_i \Delta^{1/2}\epsilon_i, \quad i = 0,\dots,n-1, \tag{34}$$

where $\{\epsilon_i\}_{i=0}^{n-1}$ is an i.i.d. series with standard normality.

The price series by Black Scholes formula is part of the solution of the FBSDEs above at discrete time points; that is,

$$Y_i = X_i N\left(d_+^i\right) - e^{-r(n-i)\Delta} K N\left(d_-^i\right), \tag{35}$$

which, together with

$$Z_i = \sigma X_i N\left(d_+^i\right), \tag{36}$$

gives us data generating formulae, where

$$N(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} dx \tag{37}$$

is a cumulative normal function, and

$$d_\pm^i = \frac{\ln(X_i/K) + \left(r \pm \sigma^2/2\right)((n-i)\Delta)}{\sigma\sqrt{(n-i)\Delta}}. \tag{38}$$

We produce the true curve of the drift coefficient by

$$g_i = -rY_i - (b-r)\sigma^{-1}Z_i. \tag{39}$$

We first apply formulas (21) and (11) to estimate g_i and Z_i^2, respectively. We adopt Epanechnikov kernel defined by $K(u) = 3/4(1-u^2)I(|u| \le 1)$, where $I(\cdot)$ is the indicator function. For bandwidth selection, we simply apply the rule of thumb bandwidth selector:

$$h = \text{constant} \times \text{std}\left(Y_0,\dots,Y_{n-1}\right) n^{-1/5} \tag{40}$$

to implement the estimation.

Let $K = 110$, $X_0 = 100$, $b = 0.1$, $\sigma = 0.18$, $r = 0.08$, $T = 60$, and $\Delta = 1/100$. The bandwidth parameters $h = 6.06$ and $h = 0.67$ are used for estimation of g_s and Z_s, respectively. The values of the tuning parameters are $a_n = 0.05$, $\alpha = 1.2$, and $\beta = 1$. To see the performance of our estimation method, the simulated and the estimated curves of the two coefficients of the backward equation are displayed in Figures 2 and 3.

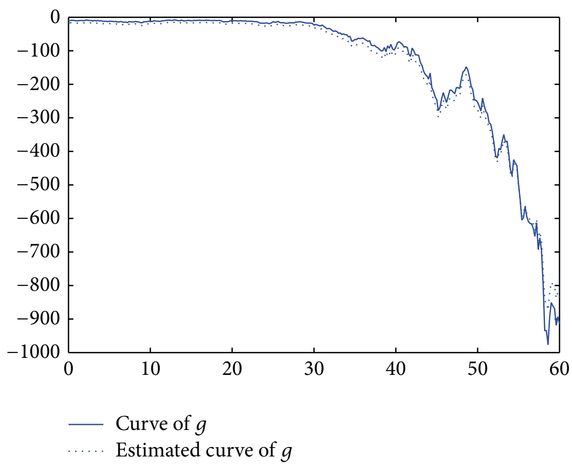


FIGURE 2: The simulated curve and the estimated curves of g_s in Example 11.

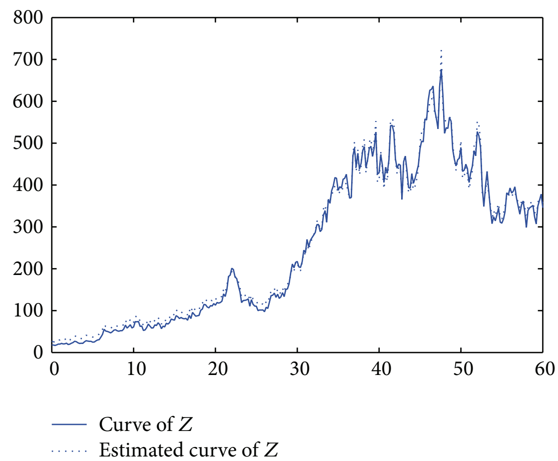


FIGURE 3: The simulated curve and the estimated curves of Z_s in Example 11.

Appendix

A. Proofs

Proof of Theorem 1. Denote $\mathscr{C} = \{X_1, \ldots, X_n, \ldots\}$. By the Taylor expansion and formula (8), we have

$$\begin{aligned} E\left(\widehat{Z}^2(x_0) \mid \mathscr{C}\right) &= \frac{\sum_{i=1}^{n-1} \Delta^{-1} K_h(X_i - x_0) E\left((Y_{i+1} - Y_i)^2 \mid \mathscr{C}\right)}{\sum_{i=1}^{n-1} K_h(X_i - x_0)} \\ &= \frac{\sum_{i=1}^{n-1} K_h(X_i - x_0)\left(Z_i^2 + O(\Delta)\right)}{\sum_{i=1}^{n-1} K_h(X_i - x_0)} \\ &= \frac{\int K_h(X_i - x_0)\left(Z^2(x) + O(\Delta)\right) p(x)\, dx \left(1 + O_p(nh)^{-1/2}\right)}{\int K_h(X_i - x_0)\, p(x)\, dx \left(1 + O_p(nh)^{-1/2}\right)} \\ &= \Big(\left(Z^2(x_0) + O(\Delta)\right) \\ &\quad \times \left(p(x_0) + (1/2) h^2 p^{(2)}(x_0) \sigma_K^2 + o(h^2)\right) \\ &\quad \times \left(1 + O_p(nh)^{-1/2}\right) \Big) \\ &\quad \times \Big(\left(p(x_0) + (1/2) h^2 p^{(2)}(x_0) \sigma_K^2 + o(h^2)\right) \\ &\qquad \times \left(1 + O_p(nh)^{-1/2}\right) \Big)^{-1} \\ &= Z^2(x_0) + \frac{p^{(2)}(x_0)}{2p(x_0)} h^2 Z^2(x_0) \sigma_K^2 + o(h^2) + O(\Delta). \end{aligned} \tag{A.1}$$

Furthermore,

$$\begin{aligned} \operatorname{Var}\left(\widehat{Z}^2(x_0) \mid \mathscr{C}\right) &= \frac{1}{\sum_{i=1}^{n-1} K_h^2(X_i - x_0)} \\ &\quad \times \Big\{ \sum_{i=1}^{n-1} \Delta^{-2} K_h^2(X_i - x_0) \operatorname{Var}\left((Y_{i+1} - Y_i)^2 \mid \mathscr{C}\right) \\ &\qquad + \sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \Delta^{-2} \operatorname{cov}\left(K_h(X_i - x_0)(Y_{i+1} - Y_i),\right. \\ &\qquad\qquad \left. K_h(X_{i+k} - x_0)(Y_{i+k+1} - Y_{i+k}) \mid \mathscr{C}\right) \Big\}. \end{aligned} \tag{A.2}$$

From the conditions of Markov process and ρ-mixing coefficient,

$$\begin{aligned} &\Big| \sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \Delta^{-2} \operatorname{cov}\Big(K_h(X_i - x_0)(Y_{i+1} - Y_i), \\ &\qquad K_h(X_{i+k} - x_0)(Y_{i+k+1} - Y_{i+k}) \Big) \Big| \\ &= \frac{1}{(n-1)^2} \sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \Big| E\Big((\Delta)^{-2} (Y_{i+1} - Y_i)^2 (Y_{i+k+1} - Y_{i+k})^2 \\ &\qquad \times \left(K_h(X_i - x_0) - E\left(K_h(X_i - x_0)\right)\right) \\ &\qquad \times \left(K_h(X_{i+k} - x_0) - E\left(K_h(X_{i+k} - x_0)\right)\right) \Big) \Big| \\ &= \frac{1}{(n-1)^2} \Big| E\Big(Z_i^2 Z_{i+l}^2 \left(K_h(X_i - x_0) - E\left(K_h(X_i - x_0)\right)\right) \\ &\qquad \times \left(K_h(X_{i+k} - x_0) - E\left(K_h(X_{i+k} - x_0)\right)\right) \Big) \Big| \\ &\quad + O(\Delta) \\ &\le \frac{C}{(n-1)^2 h} \sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \rho^k = O\left(\frac{1}{nh}\right) = o(1). \end{aligned} \tag{A.3}$$

Note that $(Y_{i+1} - Y_i)/\sqrt{\Delta} = g(t_i, Y_i, Z_i)\sqrt{\Delta} + Z_i\eta_i$, where $E(\eta_i) = 0$, $\mathrm{Var}(\eta_i) = 1$. Thus $\mathrm{Var}((Y_{i+1} - Y_i)/\sqrt{\Delta}) = Z_i^4 + O(\Delta)$ and furthermore

$$\begin{aligned}\mathrm{Var}&\left(\widehat{Z}^2(x_0) \mid \mathscr{C}\right)\\ &= \frac{\sum_{i=1}^{n-1} \Delta^{-2} K_h^2(X_i - x_0)\,\mathrm{Var}\left((Y_{i+1} - Y_i)^2 \mid \mathscr{C}\right)}{\sum_{i=1}^{n-1} K_h^2(X_i - x_0)} + O_p(1)\\ &= \frac{\sum_{i=1}^{n-1} K_h^2(X_i - x_0)\left(Z^4(x_0) + O\left(\sqrt{\Delta}\right)\right)}{\sum_{i=1}^{n-1} K_h^2(X_i - x_0)} + O_p(1)\\ &= \frac{Z^4(x_0) J_K + O\left(\sqrt{\Delta}\right)}{nhp(x_0)}\left(1 + O_p(nh)^{-1/2}\right).\end{aligned} \tag{A.4}$$

To our interest, both the conditional expectation and variance are independent on $\mathscr{C}$, so the condition could be erased.

From Lemma A.1 of Politis and Romano [15] and the relation between the α-mixing condition and the ρ-mixing condition (e.g., Theorem 1.1.1 of Lu and Lin [14]), we can ensure that $\{(Y_{i+1} - Y_i)^2,\ i = 1, \ldots, n-1\}$ is a ρ-mixing dependent process and the mixing coefficient, denoted by $\rho_Y(l)$, satisfies

$$\sum_{k=1}^{\infty} \rho_Y\left(2^k\right) \le C\sum_{k=1}^{\infty} \rho\left(2^k\right) = \sum_{k=1}^{\infty} \rho^{2^k} < \infty, \tag{A.5}$$

where C is a positive constant. Finally, we use the central limit theorems for ρ-mixing dependent process (e.g., Theorem 4.0.1 of Lu and Lin [14]) to complete this proof. □

Proof of Theorem 8. Theorem 8 follows from proving that $S_{n1}(x,z)/\sqrt{V_n(x,z)} \to^d N(0, I_2)$ and $[S_{n2}(x,z) - g(x,z)]/\sqrt{V_n(x,z)} = o_p(1)$ except, possibly, if (x,z) belongs to a set of Lebesgue measure 0. The first result is established in Lemma A.1, and the second is established in Lemma A.2. Throughout this Appendix, "for almost every (x,z)" means "for every $(x,z) \in [0,1]^2$ except, possibly, a set of Lebesgue measure 0." We make repeated use of the fact that if $E\|\psi\|^2 = O(n^{-s})$ for some $s > 0$, then $\psi(x,z) = o_p(n^{-s})$ for almost every (x,z). □

Lemma A.1 (asymptotic normality of $S_{n1}(x,z)/\sqrt{V_n(x,z)}$). *Let Assumptions 2–7 hold. Then $S_{n1}(x,z)/\sqrt{V_n(x,z)} \to^d N(0, I_2)$ for almost every (x,z).*

Proof. Define $S_{n11}(x,z) = n^{-1}\sum_{i=1}^{n} \mathbb{U}_i(T^+ f_{\mathbb{XZW}})(x,z,\mathbb{W}_i)$,

$$\begin{aligned}A_{n2}&(x,z)\\ &= n^{-1}\sum_{i=1}^{n}\mathbb{U}_i\left[T^+\left(\widehat{f}_{\mathbb{XZW}}^{(-i)} - f_{\mathbb{XZW}}\right)\right](x,z,\mathbb{W}_i),\\ A_{n3}&(x,z)\\ &= n^{-1}\sum_{i=1}^{n}\mathbb{U}_i\left[\left(\widehat{T}^+ - T^+\right) f_{\mathbb{XZW}}\right](x,z,\mathbb{W}_i),\\ A_{n4}&(x,z)\\ &= n^{-1}\sum_{i=1}^{n}\mathbb{U}_i\left[\left(\widehat{T}^+ - T^+\right)\left(\widehat{f}_{\mathbb{XZW}}^{(-i)} - f_{\mathbb{XZW}}\right)\right](x,z,\mathbb{W}_i).\end{aligned} \tag{A.6}$$

Then $S_{n1}(x,z) = S_{n11}(x,z) + A_{n2}(x,z) + A_{n3}(x,z) + A_{n4}(x,z)$. $S_{n11}(x,z)/\sqrt{V_n(x,z)} \to^d N(0, I_2)$ by a triangular array version of the Lindeberg-Levy central limit theorem. The proof of the triangular-array version of the theorem is identical to the proof of the ordinary Lindeberg-Levy theorem. The lemma follows if we can prove that $A_{nj}(x,z)/\sqrt{V_n(x,z)} = o_p(1)$ for $j = 2, 3, 4$ and almost every $(x,z) \in [0,1]^2$.

Assumption 7 and arguments like those leading to (6.2) of HH [11] show that

$$\iint_0^1 V_n(x,z)\,dx\,dz \asymp n^{-[2\beta+\alpha-\rho(\alpha+1)]/(2\beta+\alpha)}. \tag{A.7}$$

It follows from the Cauchy-Schwartz inequality, $E(\widehat{f}_{\mathbb{XZW}}^{(-i)} - f_{\mathbb{XZW}}) = O(h^r)$, and $\mathrm{Var}(\widehat{f}_{\mathbb{XZW}}^{(-i)}) = O[1/(nh^2)]$ that

$$E\|A_{n2}\|^2 = O\left(\frac{1}{n^2h^2a_n^2} + \frac{h^{2r}}{na_n^2}\right). \tag{A.8}$$

Therefore, it follows from Assumptions 5 and 7 that $A_{n2}(x,z)/\sqrt{V_n(x,z)} = o_p(1)$ for almost every (x,z). Now consider $A_{n3}(x,z)$. Define the operator $\Delta = \widehat{T} - T$. Then

$$A_{n3}(x,z) = -\left(\widehat{T} + a_n I\right)\Delta A_{n1}(x,z). \tag{A.9}$$

Therefore, the Cauchy-Schwartz inequality gives

$$\begin{aligned}E\|A_{n2}\|^2 &\le E\left\|\left(\widehat{T} + a_n I\right)\Delta\right\|^2 E\|S_{n11}\|^2\\ &= E\left\|\left(\widehat{T} + a_n I\right)\Delta\right\|^2 \iint_0^1 V_n(x,z)\,dx\,dz.\end{aligned} \tag{A.10}$$

HH show that

$$E\left\|\left(\widehat{T} + a_n I\right)\Delta\right\|^2 = O\left(\frac{1}{nha_n^2} + \frac{h^{2r}}{a_n^2}\right). \tag{A.11}$$

Therefore, it follows from Assumptions 5 and 7 that $A_{n3}(x,z)/\sqrt{V_n(x,z)} = o_p(1)$ for almost every (x,z). Finally, some algebra shows that

$$A_{n4}(x,z) = -\left(\widehat{T} + a_n I\right)^{-1}\Delta A_{n2}(x,z). \tag{A.12}$$

Therefore, $A_{n4}(x,z)/\sqrt{V_n(x,z)} = o_p(1)$ for almost every (x, z) follows from (A.11) and $A_{n2}(x,z)/\sqrt{V_n(x,z)} = o_p(1)$. □

Lemma A.2 (asymptotic negligibility of $S_{n2}(x,z) - g(x,z)$). *Let Assumptions 2–7 hold. Then $S_{n2}(x,z) - g(x,z)/\sqrt{V_n(x,z)} = o_p(1)$ for almost every (x,z).*

Proof. Define

$$D_n(x,z) = \iiint_0^1 g(\theta,\eta) f_{\mathbb{XZW}}(\theta,\eta,w) T^+ \times \left(\hat{f}_{\mathbb{XZW}} - f_{\mathbb{XZW}}\right)(x,z,w)\, d\theta\, d\eta\, dw,$$

$$A_{n1}(x,z) = n^{-1}\sum_{i=1}^{n} g(\mathbb{X}_i, Z_i)(T^+ f_{\mathbb{XZW}})(x,z,\mathbb{W}_i). \tag{A.13}$$

Redefine

$$A_{n2}(x,z) = n^{-1}\sum_{i=1}^{n} g(\mathbb{X}_i,\mathbb{Z}_i)\left[T^+\left(\hat{f}^{(-i)}_{\mathbb{XZW}} - f_{\mathbb{XZW}}\right)\right](x,z,\mathbb{W}_i) - D_n(x,z),$$

$$A_{n3}(x,z) = n^{-1}\sum_{i=1}^{n} g(\mathbb{X}_i,\mathbb{Z}_i)\left[\left(\hat{T}^+ - T^+\right) f_{\mathbb{XZW}}\right](x,z,\mathbb{W}_i) + D_n(x,z),$$

$$A_{n4}(x,z) = n^{-1}\sum_{i=1}^{n} g(\mathbb{X}_i,\mathbb{Z}_i)\left[\left(\hat{T}^+ - T^+\right)\left(\hat{f}^{(-i)}_{\mathbb{XZW}} - f_{\mathbb{XZW}}\right)\right] \times (x,z,\mathbb{W}_i). \tag{A.14}$$

Then $S_{n2}(x,z) = \sum_{j=1}^{4} A_{nj}(x,z)$. Arguments identical to those used to derive (6.2) and (6.3) of HH [11] show that $\|EA_{n1} - g\|^2 = O[n^{-\rho(2\beta_1)/(2\beta+\alpha)}]$ and

$$\iint_0^1 \operatorname{Var}[A_{n1}(x,z)]\, dx\, dz = On^{-[2\beta+\alpha-\rho(\alpha+1)]/(2\beta+\alpha)}. \tag{A.15}$$

Therefore, it follows from Assumptions 5 and 7 that

$$\frac{[EA_{n1}(x,z) - g(x,z)]}{\sqrt{V_n(x,z)}} = o(1), \tag{A.16}$$

$$V_n^{-1}(x,z)\iint_0^1 \operatorname{Var}[A_{n1}(x,z)]\, dx\, dz = O(1) \tag{A.17}$$

for almost every (x,z).

Now consider $A_{n2}(x,z)$. Define

$$D_{ni}(x,z) = \iiint_0^1 g(\theta,\eta) f_{\mathbb{XZW}}(\theta,\eta,w) T^+ \times \left(\hat{f}^{(-i)}_{\mathbb{XZW}} - f_{\mathbb{XZW}}\right)(x,z,w)\, d\theta\, d\eta\, dw,$$

$$A_{n21}(x,z) = n^{-1}\sum_{i=1}^{n} g(\mathbb{X}_i,\mathbb{Z}_i)\left[T^+\left(\hat{f}^{(-i)}_{\mathbb{XZW}} - f_{\mathbb{XZW}}\right)\right] \times (x,z,\mathbb{W}_i) - D_{ni}(x,z), \tag{A.18}$$

and $A_{n22}(x,z) = n^{-1}\sum_{i=1}^{n}[D_{ni}(x,z) - D_n(x,z)]$. HH show that $\|EA_{n21}\|^2 = O((h^{2r}/na_n^2) + (1/n^2h^2a_n^2))$ and $\|EA_{n22}\|^2 = O(1/n^2a_n^2)$. Therefore, it follows from Assumptions 5 and 7 that

$$\frac{A_{n2}(x,z)}{\sqrt{V_n(x,z)}} = o_p(1) \tag{A.19}$$

for almost every (x,z). Now consider $A_{n3}(x,z)$. Write

$$A_{n3}(x,z) = A_{n31}(x,z) + A_{n32}(x,z), \tag{A.20}$$

where $A_{n31}(x,z) = -(I + T^+\Delta)^{-1}T^+\Delta g(x,z) + D_n(x,z)$ and $A_{n32}(x,z) = -(\hat{T}^+ + a_n I)^{-1}\Delta(A_{n1} - g)(x,z)$. It follows from (A.11)-(A.16) and (A.20) that

$$\frac{A_{n32}(x,z)}{\sqrt{V_n(x,z)}} = o_p(1) \tag{A.21}$$

for almost every (x,z).

To analyze $A_{n31}(x,z)$, define

$$\begin{aligned}
B_{n1}(x,z) &= \iiint_0^1 \left[\hat{f}_{\mathbb{XZW}}(x,z,w) - f_{\mathbb{XZW}}(x,z,w)\right] \times f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw,\\
B_{n2}(x,z) &= \iiint_0^1 \left[\hat{f}_{\mathbb{XZW}}(x,z,w) - f_{\mathbb{XZW}}(x,z,w)\right] \times f_{\mathbb{XZW}}(x,x,w)\, g(x,x)\, dx\, dz\, dw,\\
B_{n3}(x,z) &= \iiint_0^1 \left[\hat{f}_{\mathbb{XZW}}(x,z,w) - f_{\mathbb{XZW}}(x,z,w),\ \hat{f}_{\mathbb{XZW}}(x,z,w) - f_{\mathbb{XZW}}(x,z,w)\right] \times g(x,z)\, dx\, dz\, dw,\\
B_{n11}(x,z) &= \iiint_0^1 \left[E\hat{f}_{\mathbb{XZW}}(x,z,w) - f_{\mathbb{XZW}}(x,z,w)\right] \times f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw,\\
B_{n12}(x,z) &= \iiint_0^1 \left[\hat{f}_{\mathbb{XZW}}(x,z,w) - E\hat{f}_{\mathbb{XZW}}(x,w)\right] \times f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw,\\
B_{n21}(x,z) &= \iiint_0^1 \left[E\hat{f}_{\mathbb{XZW}}(x,z,w) - f_{\mathbb{XZW}}(x,z,w)\right] \times f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw,\\
B_{n22}(x,z) &= \iiint_0^1 \left[\hat{f}_{\mathbb{XZW}}(x,z,w) - E\hat{f}_{\mathbb{XZW}}(x,z,w)\right] \times f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw.
\end{aligned} \tag{A.22}$$

Define $\delta = h^{2r} + (nh)^{-1}$. HH show that

$$A_{n31}(x,z) = -\left(I + T^{+}\Delta\right)^{-1} T^{+}\left(B_{n11} + B_{n12} + B_{n3}\right)(x,z) + \left(I + T^{+}\Delta\right)^{-1} T^{+}\Delta T^{+}\left(B_{n21} + B_{n22}\right)(x,z). \tag{A.23}$$

Define

$$\widetilde{A}_{n31}(x,z) = -\left(I + T^{+}\Delta\right)^{-1} T^{+}\left(B_{n11} + B_{n12} + B_{n3}\right)(x,z) + \left(I + T^{+}\Delta\right)^{-1} T^{+}\Delta T^{+} B_{n21}. \tag{A.24}$$

Then

$$E\left\|A_{n31}\right\|^{2} \le \text{const.}\left[E\left\|\widetilde{A}_{n31}\right\|^{2} + E\left\|(I + T\Delta)^{-1} T^{+}\Delta T^{+} B_{n22}\right\|^{2}\right], \tag{A.25}$$

$$E\left\|\widetilde{A}_{n31}\right\|^{2} \le \text{const.}\left(\left\|T^{+}B_{n11}\right\|^{4} + E\left\|T^{+}B_{n12}\right\|^{4} + E\left\|T^{+}\Delta T^{+}B_{n21}\right\|^{4} + E\left\|T^{+}B_{n3}\right\|^{4}\right)^{1/2}. \tag{A.26}$$

HH show that

$$\left\|T^{+}B_{n11}\right\| = O\left(\frac{h^{r}}{a_n}\right), \tag{A.27}$$

$$\left(E\left\|T^{+}\Delta T^{+}B_{n21}\right\|^{4}\right)^{1/2} = O\left(\frac{\delta h^{2r}}{a_n}\right), \tag{A.28}$$

$$\left(E\left\|T^{+}B_{n3}\right\|^{4}\right)^{1/2} = O\left(\frac{\delta^{2}}{a_n^{2}}\right). \tag{A.29}$$

See (6.11), (6.13), (6.14), and (6.15) of HH [11]. Moreover,

$$E\left\|(I + T\Delta)^{-1} T^{+}\Delta T^{+} B_{n22}\right\|^{2} = O\left(\frac{h^{2r-1}}{n a_n^{2+(\alpha+1)/\alpha}} + \frac{1}{n^{3}h^{5}a_n^{4}} + \frac{h^{4r}}{nha_n^{2}}\right). \tag{A.30}$$

See the arguments leading to (6.24) in HH [11] and the analogous result for their equation (6.24) in HH [11] and the analogous result for their quantity $E\|H_{n2}\|^{2}$. Combining (A.25)–(A.30) with Assumptions 5 and 7 yields the result that

$$\frac{A_{n4}(x,z)}{\sqrt{V_n(x,z)}} = \frac{-\left(I + T^{+}\Delta\right)^{-1} T^{+} B_{n12}}{\sqrt{V_n(x,z)}} + o_p(1). \tag{A.31}$$

Now consider $-(I + T^{+}\Delta)^{-1} T^{+} B_{n12}$. Standard calculations for kernel estimators show that

$$\iiint_{0}^{1} \widehat{f}_{\mathbb{XZW}}(x,z,w)\, f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw = n^{-1}\sum_{i=1}^{n} f_{\mathbb{XZW}}(x,z,\mathbb{W}_i)\, g(\mathbb{X}_i,\mathbb{Z}_i) + O(h^{r}). \tag{A.32}$$

Therefore,

$$T^{+}\iiint_{0}^{1} \widehat{f}_{\mathbb{XZW}}(x,z,w)\, f_{\mathbb{XZW}}(x,z,w)\, g(x,z)\, dx\, dz\, dw = A_{n1}(x,z) + o\left(\frac{h^{r}}{a_n}\right), \tag{A.33}$$

$$T^{+}B_{n12}(x,z) = A_{n1}(x,z) - EA_{n1}(x,z) + o\left(\frac{h^{r}}{a_n}\right). \tag{A.34}$$

But

$$\begin{aligned}\left(I + T^{+}\Delta\right)^{-1} T^{+} B_{n12}(x,z) &= T^{+}B_{n12} + \left[\left(I + T^{+}\Delta\right)^{-1} - I\right] T^{+} B_{n12} \\ &= T^{+}B_{n12} + \left(\widehat{T} + a_n I\right)^{-1} \Delta T^{+} B_{n12}.\end{aligned} \tag{A.35}$$

Therefore, it follows, by combining Assumption 7 and equations (A.11), (A.17), and (A.34), that

$$\left(I + T^{+}\Delta\right)^{-1} T^{+} B_{n12}(x,z) = A_{n1}(z) - EA_{n1}(x,z) + r_n, \tag{A.36}$$

where $E\|r_n\|^{2}/\sqrt{V_n(x,z)} = o(1)$ for almost every (x,z). Combining this result with (A.21) and (A.31) gives

$$\frac{A_{n3}(x,z)}{\sqrt{V_n(x,z)}} = \frac{-\left[A_{n1}(x,z) - EA_{n1}(x,z)\right]}{\sqrt{V_n(x,z)}} + o_p(1) \tag{A.37}$$

for almost every (x,z).

Now consider $A_{n4}(x,z)$. HH show that

$$A_{n4}(x,z) = -\left(I + T^{+}\Delta\right)^{-1} T^{+}\Delta\left(A_{n2} - T^{+}B_{n2}\right)(x,z). \tag{A.38}$$

Therefore, it follows from (A.19) and (A.30) that

$$\frac{A_{n4}(x,z)}{\sqrt{V_n(x,z)}} = o_p(1) \tag{A.39}$$

for almost every (x,z).

Now combine (A.19), (A.37), and (A.39) to obtain

$$\frac{S_{n2}(x,z)}{\sqrt{V_n(x,z)}} = \frac{\sum_{j=1}^{4} A_{nj}(x,z)}{\sqrt{V_n(x,z)}} = \frac{EA_{n1}(x,z)}{\sqrt{V_n(x,z)}} + o_p(1) \tag{A.40}$$

for almost every (x,z).The lemma follows by combining this result with (A.16).

This completes the proof. □

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

[1] F. Black and M. Scholes, "The pricing of options corporate liabilities," *Journal of Political Economy*, vol. 81, pp. 637–659, 1973.

[2] R. C. Merton, "Theory of rational option pricing," *Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.

[3] F. Antonelli, "Backward-forward stochastic differential equations," *The Annals of Applied Probability*, vol. 3, no. 3, pp. 777–793, 1993.

[4] H. Wang, W. Li, and X. Wang, "Asymptotic stabilization by state feedback for a class of stochastic nonlinear systems with time-varying coefficients," *Mathematical Problems in Engineering*, vol. 2014, Article ID 258093, 6 pages, 2014.

[5] W. Zhang and G. Li, "Discrete-time indefinite stochastic linear quadratic optimal control with second moment constraints," *Mathematical Problems in Engineering*, vol. 2014, Article ID 278142, 9 pages, 2014.

[6] Y. Su and L. Lin, "Semi-parametric estimation for forward-backward stochastic differential equations," *Communications in Statistics: Theory and Methods*, vol. 38, no. 11, pp. 1759–1775, 2009.

[7] X. Chen and L. Lin, "Nonparametric estimation for FBSDEs models with applications in finance," *Communications in Statistics—Theory and Methods*, vol. 39, no. 14, pp. 2492–2514, 2010.

[8] L. Lin, F. Li, and L. X. Zhu, "On regression with variance built-in mean regression function: a new financial model," Manuscript, 2009.

[9] Q. Zhang and L. Lin, "Terminal-dependent statistical inferences for FBSDE," *Stochastic Analysis and Applications*, vol. 32, pp. 128–151, 2014.

[10] C. Hsiao, *Analysis of Panel Data*, vol. 36 of *Econometric Society Monographs*, Cambridge University Press, Cambridge, UK, 2nd edition, 2003.

[11] P. Hall and J. L. Horowitz, "Nonparametric methods for inference in the presence of instrumental variables," *The Annals of Statistics*, vol. 33, no. 6, pp. 2904–2929, 2005.

[12] A. N. Kolmogorov and U. A. Rozanov, "On the strong mixing conditions of a stationary Gaussian process," *Theory of Probability and Its Applications*, vol. 2, pp. 222–227, 1960.

[13] R. C. Bradley and W. Bryc, "Multilinear forms and measures of dependence between random variables," *Journal of Multivariate Analysis*, vol. 16, no. 3, pp. 335–367, 1985.

[14] C. R. Lu and Z. Y. Lin, *Limit Theories for Mixing Dependent Variables*, Science Press, Beijing, China, 1997.

[15] D. N. Politis and J. P. Romano, "A general resampling scheme for triangular arrays of α-mixing random variables with application to the problem of spectral density estimation," *The Annals of Statistics*, vol. 20, no. 4, pp. 1985–2007, 1992.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 546107, 8 pages
http://dx.doi.org/10.1155/2014/546107

Research Article

Moving State Marine SINS Initial Alignment Based on High Degree CKF

Yong-Gang Zhang, Yu-Long Huang, Zhe-Min Wu, and Ning Li

College of Automation, Harbin Engineering University, No. 145 Nantong Street, Nangang District, Harbin 150001, China

Correspondence should be addressed to Yu-Long Huang; heuedu@163.com

Received 11 April 2014; Revised 3 June 2014; Accepted 5 June 2014; Published 24 June 2014

Academic Editor: Wuquan Li

A new moving state marine initial alignment method of strap-down inertial navigation system (SINS) is proposed based on high-degree cubature Kalman filter (CKF), which can capture higher order Taylor expansion terms of nonlinear alignment model than the existing third-degree CKF, unscented Kalman filter and central difference Kalman filter, and improve the accuracy of initial alignment under large heading misalignment angle condition. Simulation results show the efficiency and advantage of the proposed initial alignment method as compared with existing initial alignment methods for the moving state SINS initial alignment with large heading misalignment angle.

1. Introduction

It is well known that the attitude update of strap-down inertial navigation system (SINS) is achieved based on numerical integration [1]. Therefore, it is necessary to know initial navigation parameters including position, velocity, and attitude for navigation calculation. The procedure of estimating initial navigation parameters is initial alignment, and the accuracy of estimation of these initial navigation parameters, especially the estimation accuracy of attitude, is very important to subsequent navigation operation, since the initial attitude errors (or misalignment angles) will seriously degrade the performance of SINS and cause positioning and attitude errors [2]. Thus, it is important to estimate initial attitude and reduce misalignment angles. Initial alignment of SINS is usually accomplished in stationary mode [3]. However, a moving state initial alignment is necessary to maintain high navigation accuracy. Generally, after initial alignment, the resulting navigation state errors grow up because of the initialization errors and cumulative sensor inaccuracies [4]. Consequently, in large navigation errors, due to the growing sensor error and the poor orientation, SINS often requires to be realigned, and the initialization needs the ship to stop at the initial position for at least 5 to 10 minutes [4, 5]. However, it is inconvenient and impractical that there is not enough time to stop at the initial position. Therefore, a moving state initial alignment of SINS is necessary to enable the ship to start instantly [5]. Besides, in some applications, such as carrier-launched aircraft, it is necessary to achieve an accurate moving state (or in-motion) initial alignment of host SINS [6]. As the host carrier may be sailing while aligning the SINS of a carrier-launched aircraft, aiding information from host SINS will be used to accomplish the alignment, then a moving state alignment should be implemented to realign SINS for vessel in sail [6].

In moving state initial alignment of SINS, heading misalignment angle may be large since there is no reference to indicate current heading, especially for integrated alignment, and error model of SINS with large heading misalignment angle is nonlinear, which means linear estimation methods are not suitable for SINS initial alignment with large heading misalignment angle [7]. In order to solve the problem of moving state initial alignment with large heading misalignment angle, Kong et al. proposed an initial alignment method based on extended Kalman filter (EKF) [8]. However, it has low alignment accuracy and slow alignment speed. In order to improve the alignment accuracy and alignment speed, Zhou et al. proposed an initial alignment method based on unscented Kalman filter (UKF), which can at least capture the posterior mean and covariance to the second order of the Taylor series of any nonlinearity [9]. To improve the computational efficiency of UKF method, Chang et al. proposed an initial alignment method based on marginalized UKF [10]. To further improve the accuracy of UKF method, Long et al. proposed an initial alignment method based on central difference Kalman filter (CDKF), which can provide better covariance estimation than UKF [11]. To improve the numerical stability of UKF method, Sun proposed an initial alignment method based on cubature Kalman filter (CKF) [12], which is a special case of UKF with better numerical stability [13].

However, all moving state initial alignment methods mentioned above have limited alignment accuracy and alignment speed because they cannot capture the fifth order Taylor expansion terms of nonlinear alignment model. In order to improve alignment accuracy and alignment speed, a new moving state initial alignment method based on the fifth-degree CKF (5th-CKF) is proposed in this paper. For moving state initial alignment of SINS with large heading misalignment angle, the 5th-CKF addresses the strong nonlinearity problem better than existing methods because it can capture the fifth order Taylor expansion terms of nonlinear alignment model. As will be seen in our simulation results, the proposed initial alignment method outperforms existing initial alignment methods in terms of alignment accuracy and alignment speed.

The remainder of this paper is organized as follows. The nonlinear error model of moving state marine SINS initial alignment is presented in Section 2. The 5th-CKF method is formulated in Section 3. Section 4 focuses on the application of the 5th-CKF to the nonlinear estimation problem of moving state initial alignment of SINS and compares the proposed initial alignment method with existing initial alignment methods for the moving state SINS initial alignment with large heading misalignment angle. Concluding remarks are drawn in Section 5.

2. Marine SINS Initial Alignment Nonlinear Error Model

Initial alignment is a process to precisely determine initial values of strap-down matrix between the vehicle's body frame and the reference frame so that the navigation computer can start with exact initial conditions. Initial alignment is a key technique in SINS. The alignment accuracy and alignment speed will influence the performance of SINS navigation. Next we will firstly introduce nomenclature used in inertial technology and then formulate marine SINS nonlinear error model in moving state initial alignment.

2.1. SINS Nonlinear Error Model for Moving State Marine Initial Alignment. In this paper, we choose $\mathbf{x} = \left[\delta\varphi\ \delta\lambda\ \delta v_x\ \delta v_y\ \phi_x\ \phi_y\ \phi_z\ \nabla_x^b\ \nabla_y^b\ \varepsilon_x^b\ \varepsilon_y^b\ \varepsilon_z^b\right]^T$ as state vector in initial alignment, where $\delta\varphi$ and $\delta\lambda$ are errors of latitude and longitude (note we ignore the altitude error for marine application), δv_x and δv_y are velocity errors in east and north directions, ϕ_x, ϕ_y, and ϕ_z are rolling, pitching, and heading misalignment angles, respectively, ∇_x^b and ∇_y^b are constant bias of specific force in b frame, and ε_x^b, ε_y^b, and ε_z^b are constant drifts of gyro in b frame. If we denote vectors $\boldsymbol{\delta}\mathbf{v}^n = [\delta v_x\ \delta v_y\ \delta v_z]^T$, $\boldsymbol{\phi} = [\phi_x\ \phi_y\ \phi_z]^T$, $\boldsymbol{\nabla}^b = \left[\nabla_x^b\ \nabla_y^b\ \nabla_z^b\right]^T$, and $\boldsymbol{\varepsilon}^b = \left[\varepsilon_x^b\ \varepsilon_y^b\ \varepsilon_z^b\right]^T$, we will have the following SINS nonlinear error model for moving state marine initial alignment [10]:

$$\begin{aligned}
\delta\dot{\varphi} &= \frac{\delta v_y}{R_m}, \\
\delta\dot{\lambda} &= \frac{\delta v_x \sec\varphi}{R_n} + \frac{\delta\varphi v_x \sec\varphi \tan\varphi}{R_n}, \\
\boldsymbol{\delta}\dot{\mathbf{v}}^n &= \left[\mathbf{I} - \left(\mathbf{C}_n^{n\prime}\right)^T\right]\mathbf{C}_b^{n\prime}\hat{\mathbf{f}}_{\mathrm{sf}}^b + \left(\mathbf{C}_n^{n\prime}\right)^T\mathbf{C}_b^{n\prime}\boldsymbol{\delta}\mathbf{f}_{\mathrm{sf}}^b \\
&\quad - \left(2\hat{\boldsymbol{\omega}}_{ie}^n + \hat{\boldsymbol{\omega}}_{en}^n\right) \times \boldsymbol{\delta}\mathbf{v}^n \\
&\quad - \left(2\boldsymbol{\delta\omega}_{ie}^n + \boldsymbol{\delta\omega}_{en}^n\right) \times \left(\hat{\mathbf{v}}^n - \boldsymbol{\delta}\mathbf{v}^n\right), \\
\dot{\boldsymbol{\phi}} &= \mathbf{C}_w^{-1}\left[\left(\mathbf{I} - \mathbf{C}_n^{n\prime}\right)\hat{\boldsymbol{\omega}}_{in}^n + \mathbf{C}_n^{n\prime}\boldsymbol{\delta\omega}_{in}^n - \mathbf{C}_b^{n\prime}\boldsymbol{\delta\omega}_{ib}^b\right], \\
\dot{\boldsymbol{\nabla}}^b &= 0, \\
\dot{\boldsymbol{\varepsilon}}^b &= 0,
\end{aligned} \tag{1}$$

with

$$\mathbf{C}_w^{-1} = \frac{1}{\cos\phi_x}\begin{bmatrix} \cos\phi_y\cos\phi_x & 0 & \sin\phi_y\cos\phi_x \\ \sin\phi_y\sin\phi_x & \cos\phi_x & -\cos\phi_y\sin\phi_x \\ -\sin\phi_y & 0 & \cos\phi_y \end{bmatrix}, \tag{2}$$

where R_m is the meridian radius of curvature and R_n is the transverse radius of curvature, φ is the computed geographic latitude, and $\mathbf{C}_n^{n\prime}$ is the transformation matrix from true navigation frame (n frame) to erroneously computed navigation frame (n' frame) which is formulated as

$$\mathbf{C}_n^{n\prime} = \begin{bmatrix} \cos\phi_y\cos\phi_z - \sin\phi_y\sin\phi_x\sin\phi_z & \cos\phi_y\sin\phi_z + \sin\phi_y\sin\phi_x\cos\phi_z & -\sin\phi_y\cos\phi_x \\ -\cos\phi_x\sin\phi_z & \cos\phi_x\cos\phi_z & \sin\phi_x \\ \sin\phi_y\cos\phi_z + \cos\phi_y\sin\phi_x\sin\phi_z & \sin\phi_y\sin\phi_z - \cos\phi_y\sin\phi_x\cos\phi_z & \cos\phi_y\cos\phi_x \end{bmatrix}, \tag{3}$$

where $\mathbf{C}_b^{n'}$ is the computed attitude matrix, $\widehat{\mathbf{f}}_{\mathrm{sf}}^b$ is the specific force measured by accelerometers in the body frame, and $\boldsymbol{\delta}\mathbf{f}_{\mathrm{sf}}^b$ is the specific force error vector in the body frame, which can be formulated as

$$\boldsymbol{\delta}\mathbf{f}_{\mathrm{sf}}^b = \mathbf{w}_a^b + \boldsymbol{\nabla}^b, \tag{4}$$

where $\mathbf{w}_a^b = \begin{bmatrix} w_{ax}^b & w_{ay}^b & w_{az}^b \end{bmatrix}^T$ is a random white noise vector in the body frame and $\boldsymbol{\nabla}^b = \begin{bmatrix} \nabla_x^b & \nabla_y^b & \nabla_z^b \end{bmatrix}^T$ is a constant bias vector in the body frame. $\boldsymbol{\delta\omega}_{ib}^b$ is the gyro error vector in the body frame, which can be formulated as

$$\boldsymbol{\delta}\mathbf{w}_{ib}^b = \mathbf{w}_g^b + \boldsymbol{\varepsilon}^b, \tag{5}$$

where $\mathbf{w}_g^b = \begin{bmatrix} w_{gx}^b & w_{gy}^b & w_{gz}^b \end{bmatrix}^T$ is a random white noise vector in the body frame and $\boldsymbol{\varepsilon}^b = \begin{bmatrix} \varepsilon_x^b & \varepsilon_y^b & \varepsilon_z^b \end{bmatrix}^T$ is a constant drift vector in the body frame. $\boldsymbol{\omega}_{ie}^n$ is the angular rate of the earth frame with respect to the inertial frame. $\widehat{\boldsymbol{\omega}}_{ie}^n$ is the computed value of $\boldsymbol{\omega}_{ie}^n$, and $\boldsymbol{\delta\omega}_{ie}^n$ is the computational error of $\boldsymbol{\omega}_{ie}^n$. $\boldsymbol{\omega}_{ie}^n$, $\boldsymbol{\delta\omega}_{ie}^n$, and $\widehat{\boldsymbol{\omega}}_{ie}^n$ can be formulated as follows:

$$\begin{aligned} \boldsymbol{\omega}_{ie}^n &= \begin{bmatrix} 0 & \omega_{ie}\cos\varphi & \omega_{ie}\sin\varphi \end{bmatrix}^T, \\ \boldsymbol{\delta\omega}_{ie}^n &= \begin{bmatrix} 0 & -\omega_{ie}\sin\varphi\delta\varphi & \omega_{ie}\cos\varphi\delta\varphi \end{bmatrix}^T, \\ \widehat{\boldsymbol{\omega}}_{ie}^n &= \boldsymbol{\omega}_{ie}^n + \boldsymbol{\delta\omega}_{ie}^n, \end{aligned} \tag{6}$$

where $\boldsymbol{\omega}_{en}^n$ is the angular rate of the navigation frame with respect to the earth frame. $\widehat{\boldsymbol{\omega}}_{en}^n$ is the computed value of $\boldsymbol{\omega}_{en}^n$, and $\boldsymbol{\delta\omega}_{en}^n$ is the computational error of $\boldsymbol{\omega}_{en}^n$. $\boldsymbol{\omega}_{en}^n$, $\boldsymbol{\delta\omega}_{en}^n$, and $\widehat{\boldsymbol{\omega}}_{en}^n$ can be formulated as

$$\begin{aligned} \boldsymbol{\omega}_{en}^n &= \begin{bmatrix} -\dfrac{v_y}{R_m} & \dfrac{v_x}{R_n} & \dfrac{v_x\tan\varphi}{R_n} \end{bmatrix}^T, \\ \boldsymbol{\delta\omega}_{en}^n &= \begin{bmatrix} -\dfrac{\delta v_y}{R_m} & \dfrac{\delta v_x}{R_n} & \dfrac{\delta v_x\tan\varphi}{R_n} + \dfrac{v_x\sec\varphi^2\delta\varphi}{R_n} \end{bmatrix}^T, \\ \widehat{\boldsymbol{\omega}}_{en}^n &= \boldsymbol{\omega}_{en}^n + \boldsymbol{\delta\omega}_{en}^n, \end{aligned} \tag{7}$$

where v_x, v_y, v_z are true velocity values in east, north, and up direction.

According to the definitions of $\boldsymbol{\omega}_{in}^n$, $\boldsymbol{\delta\omega}_{in}^n$, and $\widehat{\boldsymbol{\omega}}_{in}^n$, they can be formulated as

$$\begin{aligned} \boldsymbol{\omega}_{in}^n &= \boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n, \\ \boldsymbol{\delta\omega}_{in}^n &= \boldsymbol{\delta\omega}_{ie}^n + \boldsymbol{\delta\omega}_{en}^n, \\ \widehat{\boldsymbol{\omega}}_{in}^n &= \boldsymbol{\omega}_{in}^n + \boldsymbol{\delta\omega}_{in}^n. \end{aligned} \tag{8}$$

We choose the velocity and position differences between SINS and external sensors, such as GPS or other higher accuracy SINSs as measurement vector $\mathbf{z}$, which can be formulated as

$$\mathbf{z} = \begin{bmatrix} \varphi_{\mathrm{INS}} - \varphi_{\mathrm{ref}} \\ \lambda_{\mathrm{INS}} - \lambda_{\mathrm{ref}} \\ v_{\mathrm{INS},x} - v_{\mathrm{ref},x} \\ v_{\mathrm{INS},y} - v_{\mathrm{ref},y} \end{bmatrix}, \tag{9}$$

where φ_{ref}, λ_{ref}, $v_{\mathrm{ref},x}$, and $v_{\mathrm{ref},y}$ are measured latitude, longitude, velocity in east, and north directions, respectively.

Note that the process model of moving state initial alignment introduced in (1) is a continuous model and we must transform it into discrete form. Given the sample time T, the propagations of position error, velocity error, and misalignment angles are discretized by using the fourth-degree Runge-Kutta method, and all the parts related to noise are discretized by using first-degree Runge-Kutta method. Based on (1) and (9), the discrete state equation and observation equation for state estimation can be formulated as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}\left(\mathbf{x}_{k-1}\right) + \mathbf{W}_{k-1}, \\ \mathbf{z}_k &= \mathbf{h}\left(\mathbf{x}_k\right) + \mathbf{V}_k, \end{aligned} \tag{10}$$

where $\mathbf{h}(\mathbf{x}_k) = \begin{bmatrix} \mathbf{I}_{4\times 4} & 0_{4\times 8} \end{bmatrix}\mathbf{x}_k$, $\mathbf{W}_{k-1}$ is the Gaussian random process noise with mean $\mathbf{0}$ and covariance $\mathbf{Q}_k$ and $\mathbf{V}_k$ is the Gaussian random measurement noise with mean $\mathbf{0}$ and covariance $\mathbf{R}_k$. Equation (10) formulates the nonlinear error model for moving state marine SINS initial alignment.

It is clear to see from (1) that the state equation of the error model of moving state marine SINS initial alignment is typically nonlinear. Thus, nonlinear filtering algorithms are necessary to estimate the state vector from which misalignment angles can be obtained to finish initial alignment. Next we will introduce high degree CKF method.

3. High Degree CKF

3.1. Brief Introduction of CKF. The heart of Gaussian filter is to compute multidimensional Gaussian-weighted integral [13, 14]. Different Gaussian approximate filters can be obtained when different integral rules are used. The third-degree CKF (3rd-CKF) is obtained when the third-degree spherical-radial cubature rule is used, and the third-degree spherical-radial cubature rule can be formulated as [13]

$$\begin{aligned} &\int_{\mathbf{R}^n} \mathbf{g}(\mathbf{x})\, N\left(\mathbf{x}; \widehat{\mathbf{x}}, \mathbf{P}_x\right) d\mathbf{x} \\ &\quad = \frac{1}{2n}\sum_{j=1}^{n}\left[\mathbf{g}\left(\sqrt{n\mathbf{P}_x}e_j + \widehat{\mathbf{x}}\right) + \mathbf{g}\left(-\sqrt{n\mathbf{P}_x}e_j + \widehat{\mathbf{x}}\right)\right], \end{aligned} \tag{11}$$

where $\mathbf{x}$ is an n-dimensional Gaussian random vector with mean $\widehat{\mathbf{x}}$ and covariance $\mathbf{P}_x$ and $\sqrt{\mathbf{P}_x}$ is the square root matrix of $\mathbf{P}_x$; that is, $\sqrt{\mathbf{P}_x}\sqrt{\mathbf{P}_x}^T = \mathbf{P}_x$, and $e_j = [0, 0, \ldots, \underset{j}{1}, \ldots, 0]^T$ denotes a unit vector to the direction of coordinate axis j.

The heart of the 3rd-CKF is the third-degree spherical-radial cubature rule in (11), which makes it possible to numerically compute multivariate moment integrals encountered in nonlinear Bayesian filter. The 3rd-CKF provides a systematic solution for high-dimensional nonlinear filtering problems. In addition, the 3rd-CKF is more stable and more principled in mathematics than sigma point approaches [13]. However, the accuracy of the 3rd-CKF is limit. To improve the accuracy of the 3rd-CKF, the 5th-CKF is proposed, which can capture higher order Taylor expansion terms of nonlinear function than the 3rd-CKF, thus higher accuracy can be obtained [14]. Next we will introduce the 5th-CKF method.

3.2. 5th-CKF Method. CKF is a recursive filtering method. We assume the posterior probability density of $\mathbf{x}_{k-1}$ has been already known in the previous update $p(\mathbf{x}_{k-1}) = \mathbf{N}(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$. Firstly we calculate the Cholesky decomposition of $\mathbf{P}_{k-1|k-1}$ as follows:

$$\mathbf{P}_{k-1|k-1} = \mathbf{S}_{k-1|k-1}\mathbf{S}_{k-1|k-1}^{T}. \tag{12}$$

The first class cubature-point and its weight are calculated as follows:

$$\mathbf{X}_{0i,k-1|k-1} = \hat{\mathbf{x}}_{k-1|k-1}, \qquad w_0 = \frac{2}{n+2}. \tag{13}$$

The second class cubature-points and their weights are calculated as follows:

$$\begin{aligned} \mathbf{X}_{1i,k-1|k-1} &= \sqrt{(n+2)}\mathbf{S}_{k-1|k-1}\mathbf{e}_i + \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{X}_{2i,k-1|k-1} &= -\sqrt{(n+2)}\mathbf{S}_{k-1|k-1}\mathbf{e}_i + \hat{\mathbf{x}}_{k-1|k-1}, \\ w_1 &= \frac{4-n}{2(n+2)^2}, \\ &\qquad (i = 1, 2, \ldots, n), \end{aligned} \tag{14}$$

where $\mathbf{e}_i$ denotes a unit vector to the direction of coordinate axis i.

The third class cubature-points and their weights are calculated as follows:

$$\begin{aligned} \mathbf{X}_{3i,k-1|k-1} &= \sqrt{(n+2)}\mathbf{S}_{k-1|k-1}\mathbf{s}_i^{+} + \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{X}_{4i,k-1|k-1} &= -\sqrt{(n+2)}\mathbf{S}_{k-1|k-1}\mathbf{s}_i^{+} + \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{X}_{5i,k-1|k-1} &= \sqrt{(n+2)}\mathbf{S}_{k-1|k-1}\mathbf{s}_i^{-} + \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{X}_{6i,k-1|k-1} &= -\sqrt{(n+2)}\mathbf{S}_{k-1|k-1}\mathbf{s}_i^{-} + \hat{\mathbf{x}}_{k-1|k-1}, \\ w_2 &= \frac{1}{(n+2)^2}, \\ &\qquad \left(i = 1, 2, \ldots, \frac{n(n-1)}{2}\right), \end{aligned} \tag{15}$$

where

$$\begin{aligned} \mathbf{s}_i^{+} &= \left\{\sqrt{\frac{1}{2}}\left(\mathbf{e}_j + \mathbf{e}_l\right) : j < l, j, l = 1, 2, \ldots, n\right\}, \\ \mathbf{s}_i^{-} &= \left\{\sqrt{\frac{1}{2}}\left(\mathbf{e}_j - \mathbf{e}_l\right) : j < l, j, l = 1, 2, \ldots, n\right\}. \end{aligned} \tag{16}$$

Sample points are obtained by propagating the above cubature-points through state equation in (10) as follows:

$$\begin{aligned} \mathbf{X}_{0i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{0i,k-1|k-1}\right), \\ \mathbf{X}_{1i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{1i,k-1|k-1}\right), \\ \mathbf{X}_{2i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{2i,k-1|k-1}\right), \\ \mathbf{X}_{3i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{3i,k-1|k-1}\right), \\ \mathbf{X}_{4i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{4i,k-1|k-1}\right), \\ \mathbf{X}_{5i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{5i,k-1|k-1}\right), \\ \mathbf{X}_{6i,k|k-1}^{*} &= \mathbf{f}\left(\mathbf{X}_{6i,k-1|k-1}\right). \end{aligned} \tag{17}$$

One-step state prediction $\hat{\mathbf{x}}_{k|k-1}$ is then obtained as weighted linear combination of sample points

$$\begin{aligned} &\hat{\mathbf{x}}_{k|k-1} \\ &= w_0\mathbf{X}_{0i,k|k-1}^{*} + w_1\sum_{j=1}^{n}\left(\mathbf{X}_{1i,k|k-1}^{*} + \mathbf{X}_{2i,k|k-1}^{*}\right) \\ &\quad + w_2\sum_{j=1}^{n(n-1)/2}\left(\mathbf{X}_{3i,k|k-1}^{*} + \mathbf{X}_{4i,k|k-1}^{*} + \mathbf{X}_{5i,k|k-1}^{*} + \mathbf{X}_{6i,k|k-1}^{*}\right). \end{aligned} \tag{18}$$

One-step state prediction error covariance $\mathbf{P}_{k|k-1}$ is updated as follows:

$$\begin{aligned} \mathbf{P}_{k|k-1} = {} & w_0\mathbf{X}_{0i,k|k-1}^{*}\mathbf{X}_{0i,k|k-1}^{*T} \\ & + w_1\sum_{j=1}^{n}\left(\mathbf{X}_{1i,k|k-1}^{*}\mathbf{X}_{1i,k|k-1}^{*T} + \mathbf{X}_{2i,k|k-1}^{*}\mathbf{X}_{2i,k|k-1}^{*T}\right) \\ & + w_2\sum_{j=1}^{n(n-1)/2}\left(\mathbf{X}_{3i,k|k-1}^{*}\mathbf{X}_{3i,k|k-1}^{*T} + \mathbf{X}_{4i,k|k-1}^{*}\mathbf{X}_{4i,k|k-1}^{*T}\right. \\ & \qquad \left. + \mathbf{X}_{5i,k|k-1}^{*}\mathbf{X}_{5i,k|k-1}^{*T} + \mathbf{X}_{6i,k|k-1}^{*}\mathbf{X}_{6i,k|k-1}^{*T}\right) \\ & - \hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}_{k|k-1}^{T} + \mathbf{Q}_{k-1}. \end{aligned} \tag{19}$$

Next the measurement update is performed. Cholesky decomposition of $\mathbf{P}_{k|k-1}$ is performed firstly:

$$\mathbf{P}_{k|k-1} = \mathbf{S}_{k|k-1}\mathbf{S}_{k|k-1}^{T}. \tag{20}$$

The first class cubature-point and its weight are calculated as follows:

$$\mathbf{X}_{0i,k|k-1} = \hat{\mathbf{x}}_{k|k-1}, \qquad w_0 = \frac{2}{n+2}. \tag{21}$$

Then the second class cubature-points and their weights are calculated as follows:

$$\begin{aligned} \mathbf{X}_{1i,k|k-1} &= \sqrt{(n+2)}\mathbf{S}_{k|k-1}\mathbf{e}_i + \hat{\mathbf{x}}_{k|k-1}, \\ \mathbf{X}_{2i,k|k-1} &= -\sqrt{(n+2)}\mathbf{S}_{k|k-1}\mathbf{e}_i + \hat{\mathbf{x}}_{k|k-1}, \\ w_1 &= \frac{4-n}{2(n+2)^2}, \\ &\qquad (i = 1, 2, \ldots, n). \end{aligned} \tag{22}$$

The third class cubature-points and their weights are calculated as follows:

$$\begin{aligned}
\mathbf{X}_{3i,k|k-1} &= \sqrt{(n+2)}\mathbf{S}_{k|k-1}\mathbf{s}_i^{+} + \hat{\mathbf{x}}_{k|k-1},\\
\mathbf{X}_{4i,k|k-1} &= -\sqrt{(n+2)}\mathbf{S}_{k|k-1}\mathbf{s}_i^{+} + \hat{\mathbf{x}}_{k|k-1},\\
\mathbf{X}_{3i,k|k-1} &= \sqrt{(n+2)}\mathbf{S}_{k|k-1}\mathbf{s}_i^{-} + \hat{\mathbf{x}}_{k|k-1},\\
\mathbf{X}_{4i,k|k-1} &= -\sqrt{(n+2)}\mathbf{S}_{k|k-1}\mathbf{s}_i^{-} + \hat{\mathbf{x}}_{k|k-1},\\
w_2 &= \frac{1}{(n+2)^2},\\
&\left(i = 1, 2, \ldots, \frac{n(n-1)}{2}\right).
\end{aligned} \tag{23}$$

Sample points are obtained by propagating the above cubature-points through observation equation as follows:

$$\begin{aligned}
\mathbf{Z}_{0i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{0i,k|k-1}\right),\\
\mathbf{Z}_{1i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{1i,k|k-1}\right),\\
\mathbf{Z}_{2i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{2i,k|k-1}\right)\\
\mathbf{Z}_{3i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{3i,k|k-1}\right),\\
\mathbf{Z}_{4i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{4i,k|k-1}\right),\\
\mathbf{Z}_{5i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{5i,k|k-1}\right),\\
\mathbf{Z}_{6i,k|k-1} &= \mathbf{h}\left(\mathbf{X}_{6i,k|k-1}\right).
\end{aligned} \tag{24}$$

One-step measurement prediction $\hat{\mathbf{z}}_{k|k-1}$ is then obtained as weighted linear combination of sample points:

$$\begin{aligned}
\hat{\mathbf{z}}_{k|k-1} &= w_0\mathbf{Z}_{0i,k|k-1} + w_1\sum_{j=1}^{n}\left(\mathbf{Z}_{1i,k|k-1} + \mathbf{Z}_{2i,k|k-1}\right)\\
&\quad + w_2\sum_{j=1}^{n(n-1)/2}\left(\mathbf{Z}_{3i,k|k-1}+\mathbf{Z}_{4i,k|k-1}+\mathbf{Z}_{5i,k|k-1}+\mathbf{Z}_{6i,k|k-1}\right).
\end{aligned} \tag{25}$$

Autocorrelation covariance matrix $\mathbf{P}_{\mathbf{zz},k|k-1}$ is obtained as follows:

$$\begin{aligned}
\mathbf{P}_{\mathbf{zz},k|k-1} &= w_0\mathbf{Z}_{0i,k|k-1}\mathbf{Z}_{0i,k|k-1}^T\\
&\quad + w_1\sum_{j=1}^{n}\left(\mathbf{Z}_{1i,k|k-1}\mathbf{Z}_{1i,k|k-1}^T + \mathbf{Z}_{2i,k|k-1}\mathbf{Z}_{2i,k|k-1}^T\right)\\
&\quad + w_2\sum_{j=1}^{n(n-1)/2}\left(\mathbf{Z}_{3i,k|k-1}\mathbf{Z}_{3i,k|k-1}^T + \mathbf{Z}_{4i,k|k-1}\mathbf{Z}_{4i,k|k-1}^T\right.\\
&\qquad\qquad \left. + \mathbf{Z}_{5i,k|k-1}\mathbf{Z}_{5i,k|k-1}^T + \mathbf{Z}_{6i,k|k-1}\mathbf{Z}_{6i,k|k-1}^T\right)\\
&\quad - \hat{\mathbf{z}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T + \mathbf{R}_k.
\end{aligned} \tag{26}$$

Cross-correlation covariance matrix $\mathbf{P}_{\mathbf{xz},k|k-1}$ is calculated as follows:

$$\begin{aligned}
\mathbf{P}_{\mathbf{xz},k|k-1} &= w_0\mathbf{X}_{0i,k|k-1}\mathbf{Z}_{0i,k|k-1}^T\\
&\quad + w_1\sum_{j=1}^{n}\left(\mathbf{X}_{1i,k|k-1}\mathbf{Z}_{1i,k|k-1}^T + \mathbf{X}_{2i,k|k-1}\mathbf{Z}_{2i,k|k-1}^T\right)\\
&\quad + w_2\sum_{j=1}^{n(n-1)/2}\left(\mathbf{X}_{3i,k|k-1}\mathbf{Z}_{3i,k|k-1}^T + \mathbf{X}_{4i,k|k-1}\mathbf{Z}_{4i,k|k-1}^T\right.\\
&\qquad\qquad \left. + \mathbf{X}_{5i,k|k-1}\mathbf{Z}_{5i,k|k-1}^T + \mathbf{X}_{6i,k|k-1}\mathbf{Z}_{6i,k|k-1}^T\right)\\
&\quad - \hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T.
\end{aligned} \tag{27}$$

The Kalman filter gain is calculated as follows:

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{xz},k|k-1}\mathbf{P}_{\mathbf{zz},k|k-1}^{-1}. \tag{28}$$

State estimation $\hat{\mathbf{x}}_{k|k}$ is calculated as follows:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}\right). \tag{29}$$

The state estimation error covariance $\mathbf{P}_{k|k}$ is calculated as follows:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{P}_{\mathbf{zz},k|k-1}\mathbf{K}_k^T. \tag{30}$$

$\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ will be used in the next iteration. From the estimated state vector $\hat{\mathbf{x}}_{k|k}$ we can obtain estimated misalignment angles $\hat{\boldsymbol{\phi}} = \left[\hat{\phi}_x \ \hat{\phi}_y \ \hat{\phi}_z\right]^T$, with which the strap-down matrix between vehicle's body frame and the reference frame $\mathbf{C}_b^n$ can be determined, and the navigation computer can start with exact initial conditions. $\mathbf{P}_{k|k}$ can be used to evaluate the accuracy of estimation. Next simulations will be performed to show the advantage of the proposed initial alignment method based on the 5th-CKF as compared with existing methods in marine initial alignment.

4. Simulations

Three simulations are performed with different parameter sets under different moving states of ship. In the first simulation, the ship is on the mooring. In the second simulation, the ship sails with constant speed $v_x = 2\,\text{m/s}$ and $v_y = 2\,\text{m/s}$. In the third simulation, the ship accelerates with $a_x = 0.02\,\text{m/s}^2$ and $a_y = 0.02\,\text{m/s}^2$ and initial velocity of $v_x = 2\,\text{m/s}$ and $v_y = 2\,\text{m/s}$. In addition, initial values of process noise covariance matrix and state and measurement noise covariance matrix in simulations are set as $\text{diag}\{0_{2\times1}(0.001^\circ/\text{h})^2(0.001^\circ/\text{h})^2(1\,\mu\text{g})^2(1\,\mu\text{g})^2(1\,\mu\text{g})^2 0_{5\times1}\}$, $0_{12\times1}$, $\text{diag}\{(0.1\,\text{m/s})^2(0.1\,\text{m/s})^2(10/\,\text{Re})^2(10/\,\text{Re})^2\}$, respectively. Other parameters used in simulations are shown in Table 1.

To compare the performance of existing initial alignment methods based on the 3rd-CKF, UKF, CDKF, and the proposed initial alignment method based on the 5th-CKF,

Table 1: Parameters used for simulations.

Swing amplitude (deg)	Swing period (s)			Initial error values of pitch and roll (deg)	Initial error values of heading (deg)	White noise bias stability of gyro (deg/h)	Bias of gyro (deg/h)	White noise error of accelerometer (μg)	Bias error of accelerometer (μg)
	Roll	Pitch	Yaw						
10	8	10	6	1	30	0.003	0.01	3.16	10

Table 2: Absolute value of steady state estimation error of misalignment angles when the ship is on the mooring with heading misalignment angle of 30°.

Initial alignment methods	3rd-CKF	UKF	CDKF	5th-CKF
Absolute value of steady state estimation error of rolling misalignment (arc mins)	0.75	0.45	0.36	0.15
Absolute value of steady state estimation error of pitching misalignment (arc mins)	0.58	0.37	0.32	0.16
Absolute value of steady state estimation error of heading misalignment (arc mins)	16	5	3.5	2.45

Table 3: Absolute value of steady state estimation error of misalignment angles when the ship sails with constant speed v_x = 2 m/s and v_y = 2 m/s and heading misalignment angle of 30°.

Initial alignment methods	3rd-CKF	UKF	CDKF	5th-CKF
Absolute value of steady state estimation error of rolling misalignment (arc mins)	0.78	0.47	0.4	0.16
Absolute value of steady state estimation error of pitching misalignment (arc mins)	0.6	0.4	0.34	0.18
Absolute value of steady state estimation error of heading misalignment (arc mins)	16	5.3	4.1	2.5

Table 4: Absolute value of steady state estimation error of misalignment angles when the ship accelerates with heading misalignment angle of 30°.

Initial alignment methods	3rd-CKF	UKF	CDKF	5th-CKF
Absolute value of steady state estimation error of rolling misalignment (arc mins)	0.64	0.33	0.25	0.09
Absolute value of steady state estimation error of pitching misalignment (arc mins)	0.63	0.4	0.35	0.15
Absolute value of steady state estimation error of heading misalignment (arc mins)	14	4.0	3.0	2.2

we choose the absolute value of estimation error of misalignment angles as performance metric. For a fair comparison, we make 500 independent Monte Carlo runs. Simulation results of existing methods and the proposed method are shown in Figures 1, 2, and 3 and Tables 2, 3, and 4, which corresponds to simulation 1, simulation 2, and simulation 3, respectively.

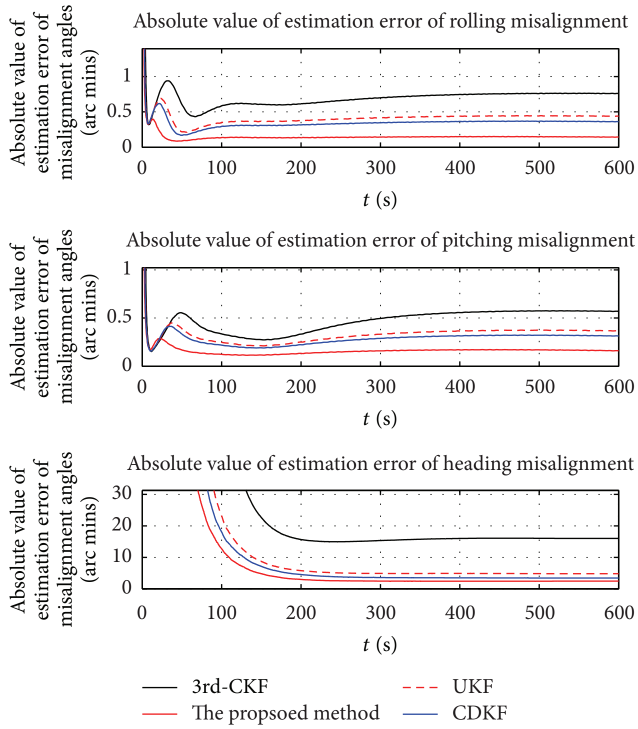


Figure 1: Absolute value of estimation error of misalignment angles based on existing methods and the proposed method when the ship is on the mooring with heading misalignment angle of 30°.

Besides, a comparison of computational complexity between the proposed method and existing methods is shown in Table 5.

It is seen from Figures 1–3 that the proposed initial alignment method has faster alignment speed than existing initial alignment methods under large heading misalignment angle conditions. From Tables 2–4, we also can see that the proposed initial alignment method outperforms existing initial alignment methods in terms of alignment accuracy under large heading misalignment angle conditions. As shown in Table 5, although the proposed initial alignment method has higher computational complexity than existing initial alignment methods, its computation requirement is acceptable for practical marine navigation application.

Theoretically, as discussed in Section 2, the initial alignment model is nonlinear for the case of large heading misalignment angle, and all nonlinear filtering algorithms only can achieve suboptimal estimation of initial misalignment angles. However, the 5th-CKF can capture higher order Taylor expansion terms of nonlinear initial alignment model than the 3rd-CKF, UKF, and CDKF. Thus, the proposed initial alignment method based on the 5th-CKF is superior

Table 5: Comparison of computational complexity.

Initial alignment methods	3rd-CKF	UKF	CDKF	5th-CKF
Computational complexity	$O(n^3)$ $(n = 12)$	$O(n^3)$ $(n = 12)$	$O(n^3)$ $(n = 12)$	$O(n^4)$ $(n = 12)$

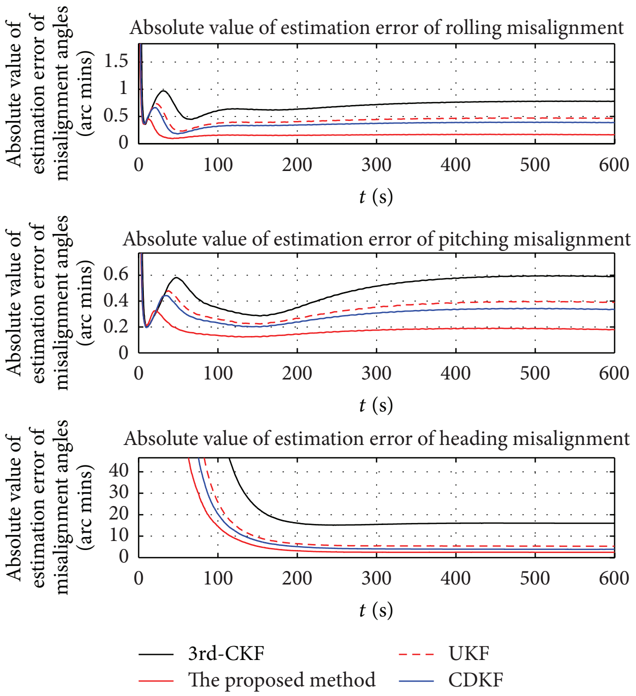


Figure 2: Absolute value of estimation error of misalignment angles based on existing methods and the proposed method when the ship sails with constant speed $v_x = 2$ m/s and $v_y = 2$ m/s and heading misalignment angle of 30°.

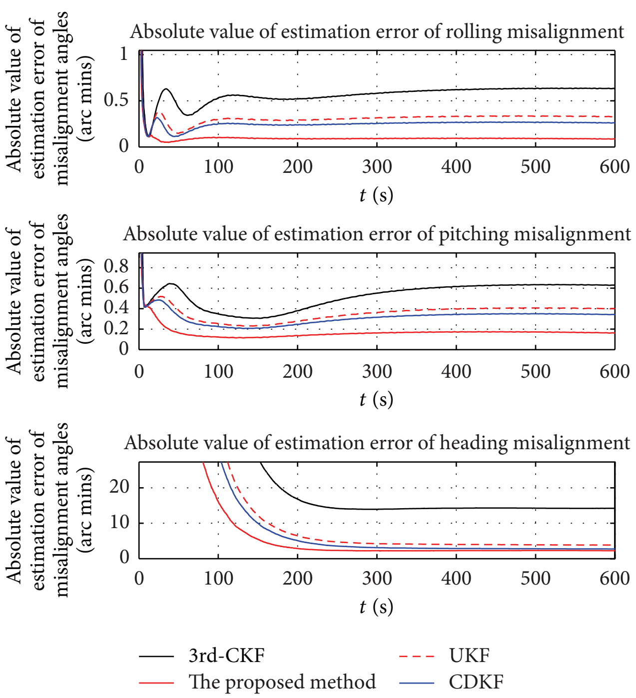


Figure 3: Absolute value of estimation error of misalignment angles based on existing methods and the proposed method when the ship accelerates with heading misalignment angle of 30°.

to existing methods based on the 3rd-CKF, UKF, and CDKF in terms of alignment accuracy and alignment speed under large heading misalignment angle. Theoretical analysis agrees with simulation results.

5. Conclusion

In this paper, a new moving state initial alignment method is proposed based on the 5th-CKF. Three simulations are performed for marine SINS initial alignment under different conditions, including mooring, moving with constant speed, and moving with constant acceleration. Simulation results show that the proposed marine SINS initial alignment method is superior to existing methods in terms of alignment accuracy and alignment speed for the moving state SINS initial alignment with large heading misalignment angle. It is more suitable for applications where fast and accurate alignment is necessary.

Nomenclatures

i Frame:	Inertial frame
e Frame:	Earth frame
n Frame:	True navigation frame ("east-north-up" local geographic frame)
b Frame:	Frame fixed to the vehicle (right-front-up)
n' Frame:	Erroneously computed navigation frame
Misalignment angle vector ϕ:	Euler angles between n frame and n' frame
$\mathbf{C}_m^n$:	Direction cosine matrix from m frame to n frame.

Conflict of Interests

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 61001154, 61201409 and 61371173, China Postdoctoral Science Foundation no. 2013M530147, Heilongjiang Postdoctoral Fund LBH-Z13052 and the Fundamental Research Funds for the Central Universities of Harbin Engineering University no. HEUCFX41307.

References

[1] Y. Wang, F. Sun, Y. Zhang, H. Liu, and H. Min, "Central difference particle filter applied to transfer alignment for SINS

on missiles," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 375–387, 2012.

[2] C. Sun and Z. Deng, "Transfer alignment of shipborne inertial-guided weapon systems," *Journal of Systems Engineering and Electronics*, vol. 20, no. 2, pp. 348–353, 2009.

[3] J. Fang and D. Wan, "A fast initial alignment method for strapdown inertial navigation system on stationary base," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 4, pp. 1501–1505, 1996.

[4] J. Ali and M. Ushaq, "A consistent and robust Kalman filter design for in-motion alignment of inertial navigation system," *Measurement*, vol. 42, no. 4, pp. 577–582, 2009.

[5] Y. Kubo, S. Fujioka, M. Nishiyama, and S. Sugimoto, "Nonlinear filtering methods for the INS/GPS in-motion alignment and navigation," *International Journal of Innovative Computing Information and Control*, vol. 2, no. 5, pp. 1137–1151, 2006.

[6] Y. Zhang, "An approach of DVL-aided SDINS alignment for in-motion vessel," *Optik*, vol. 124, no. 23, pp. 6270–6275, 2013.

[7] S. E. Dmitriyev, O. A. Stepanov, and S. V. Shepel, "Nonlinear filtering methods application in INS alignment," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 1, pp. 260–272, 1997.

[8] X. Kong, E. M. Nebot, and H. Durrant-Whyte, "Development of a non-linear psi-angle model for large misalignment errors and its application in INS alignment and calibration," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '99)*, pp. 1430–1435, Detroit, Mich, USA, May 1999.

[9] Z. Zhou, Y. Gao, and J. Chen, "Unscented Kalman filter for SINS alignment," *Journal of Systems Engineering and Electronics*, vol. 18, no. 2, pp. 327–333, 2007.

[10] L. Chang, B. Hu, A. Li, and F. Qin, "Strapdown inertial navigation system alignment based on marginalised unscented kalman filter," *IET Science, Measurement and Technology*, vol. 7, no. 2, pp. 128–138, 2013.

[11] R. Long, Y. Qin, and J. Xia, "Applying CDKF (central difference Kalman filter) to initial alignment of SINS for large misalignment angles," *Journal of Northwestern Polytechnical University*, vol. 28, no. 3, pp. 364–368, 2010 (Chinese).

[12] F. Sun and L. Tang, "Initial alignment of large azimuth misalignment angle in SINS based on CKF," *Chinese Journal of Scientific Instrument*, vol. 33, no. 2, pp. 327–333, 2012 (Chinese).

[13] I. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[14] B. Jia, M. Xin, and Y. Cheng, "High-degree cubature Kalman filter," *Automatica*, vol. 49, no. 2, pp. 510–518, 2013.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 489408, 8 pages
http://dx.doi.org/10.1155/2014/489408

Research Article

Robust H_∞ Control for Linear Stochastic Partial Differential Systems with Time Delay

Xisheng Dai,[1] Feiqi Deng,[2] and Jianxiang Zhang[1]

[1] *The School of Electrical and Information Engineering, Guangxi University of Science and Technology, Liuzhou 545006, China*
[2] *The School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China*

Correspondence should be addressed to Xisheng Dai; mathdxs@163.com

Received 23 March 2014; Accepted 2 June 2014; Published 23 June 2014

Academic Editor: Ming Gao

This paper investigates the problems of robust stochastic mean square exponential stabilization and robust H_∞ for stochastic partial differential time delay systems. Sufficient conditions for the existence of state feedback controllers are proposed, which ensure mean square exponential stability of the resulting closed-loop system and reduce the effect of the disturbance input on the controlled output to a prescribed level of H_∞ performance. A linear matrix inequality approach is employed to design the desired state feedback controllers. An illustrative example is provided to show the usefulness of the proposed technique.

1. Introduction

The H_∞ control, since it was first formulated by [1], has been extensively studied in the past years, and a great number of results on this subject have been reported in the literature; see, for example, [2–7] and the references therein. Recently, a great deal of attention has been paid regarding the study of partial differential systems (PDSs) [8–15]. Many phenomena in science and engineering have been modeled by deterministic partial differential systems, such as the control for elastic oscillating systems, the control for temperature field [13], the control for nuclear reactor, the robot with flexible connecting rod [15], population dynamics [16], neurophysiology, and biodynamics [17]. On the other hand, since most of the phenomena have spatiotemporal uncertainties due to the existence of different stochastic fluctuations, for a more accurate representation of the behavior, a stochastic partial differential system is an ideal model [16–21]. The control problem for SPDSs has been widely studied, including stability [16, 20, 21], stabilization [22], boundary and point adaptive control [23], optimal control [24–26], and parameter estimation [27]. However, it should be noted that up to now, there is little corresponding work on the robust H_∞ control for SPDSs.

In this paper, we focus on the robust H_∞ control problem of linear SPDSs with time delay. For robust H_∞ control of deterministic PDSs, the main research method is operator semigroup (see [8, 10, 12]), which is associated with solving operator equation or linear operator inequality [14]. However, general methods for solving linear operator inequality have not been developed yet, which makes most existing results difficult to be applied in practice. Later,the linear matrix inequality (LMI) is extended to uncertain distributed parameter systems [28, 29], and are used in stability analysis and H_∞ control [13, 14], respectively. Very recently, [17] studied robust filter for SPDSs by using LMI, which presented an explicit expression for the robust H_∞ filter. Motivated by these facts, our main purpose in this paper is to examine stochastic exponential stabilization and robust H_∞ control for linear SPDSs with time delay under Dirichlet boundary and Robin boundary conditions, respectively. The time delay is assumed to be unknown but bounded. First, we consider the problem of stochastic exponential stabilization for which a state feedback controller is designed such that the resulting closed-loop system is mean-square exponentially stable. Then, the problem of robust H_∞ control is addressed for which a state feedback controller is designed, for which not only is the resulting closed-loop system mean-square exponentially stable, but also is a prescribed H_∞ performance level satisfied. In terms of LMIs, sufficient conditions for the solvability of the above problems are obtained and explicit expressions of the desired state feedback controllers are presented. Here, the main method is constructing Lyapunov functional and using linear matrix inequality. Finally, an example is given to demonstrate the applicability and effectiveness of the developed theoretic results.

For convenience, we adopt the following basic notations in this paper. $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote, respectively, the n-dimensional Euclidean space and the set of all $n \times m$ real matrices. The superscript "T" denotes the transpose and the notation $X \geqslant Y$ (resp., $X > Y$) where X and Y are symmetric matrices, meaning that $X - Y$ is positive semidefinite (resp., positive definite). $\lambda_{\max}(P)$, $\lambda_{\min}(P)$ denote, respectively, maximum and minimum eigenvalue of a real symmetric matrix P. $L^2(\mathcal{O})$ is Lebergue square integrable function space defined on $\mathcal{O}$. For a scalar real value function $h(x) \in L^2(\mathcal{O})$, its $L^2(\mathcal{O})$ norm $\|h\|^2 = \int_{\mathcal{O}} h^2(x)dx$; if $h(x)$ is a vector, that is, $h(x) = (h_1(x), h_2(x), \ldots, h_n(x))^T$, then $\|h\|^2 = \int_{\Omega} h^T(x)h(x)dx$. $\mathscr{L}_2(\mathcal{O} \times [0,\infty); \mathbb{R}^{n_f})$ denotes the family of measurable function $f(x,t) : \mathcal{O} \times [0,T] \to \mathbb{R}^{n_f}$ such that $\mathbb{E}\int_0^t \int_{\mathcal{O}} f(x,s)^T f(x,s)dx\,dt := \int_{\Omega}\int_0^t \int_{\mathcal{O}} f(x,s,\omega)^T f(x,s,\omega)dx\,dt\,Pd\omega < \infty$, where $\mathbb{E}$ represents the mathematical expectation and $\{f(x,t), t \in [0,T]\}$ is stochastic process at the space location $x \in \mathcal{O}$ and a function of three arguments, that is, $f(x,t,\omega)$, $x \in \mathcal{O}$, $t \in [0,T]$, $\omega \in \Omega$.

2. Problem Statement and Preliminaries

Consider the following linear stochastic partial differential system with time delay:

$$\begin{aligned} dy(x,t) = {} & [D\Delta y(x,t) + Ay(x,t) + A_1 y(x,t-\tau) \\ & + Bu(x,t) + B_v v(x,t)]\,dt \\ & + [Cy(x,t) + C_1 y(x,t-\tau) + C_v v(x,t)]\,dW \\ & \times (x,t), \\ z(x,t) = {} & Ly(x,t), \end{aligned} \tag{1}$$

where $(x,t) \in \mathcal{O} \times \mathbb{R}^+$ and $\mathcal{O} = \{x, 0 \leqslant |x| \leqslant l < +\infty\} \subset \mathbb{R}^m$ is the bounded domain with smooth boundary $\partial\mathcal{O}$. $D > 0$ is constant. The symbol $\Delta = \sum_{i=1}^m (\partial^2/\partial x_i^2)$ is Laplace operator defined on $\mathcal{O}$, $y(x,t) = [y_1(x,t), y_2(x,t), \ldots, y_n(x,t)]^T$ is the systems state variable, and x and t are the space and time variables, respectively. $u(x,t) \in \mathscr{L}_2(\mathcal{O} \times [0,\infty); \mathbb{R}^{n_u})$ is admissible control and $z(x,t) \in \mathscr{L}_2(\mathcal{O} \times [0,\infty); \mathbb{R}^{n_z})$ is measured output. $v(x,t) \in \mathscr{L}_2(\mathcal{O} \times [0,\infty); \mathbb{R}^{n_v})$ is the vector of the random external disturbance and $B_v, C_v \in \mathbb{R}^{n\times n_v}$ are the disturbance influence matrix. A, A_1, C, C_1, L are known real constant matrix of appropriate dimension. The scalar $\tau > 0$ is an unknown but bounded time delay of the system. $W(x,t)$ is Wiener random field (see [18]) with covariance operator $\mathscr{R}$ in $L^2(\mathcal{O})$, that is, $W(x,t) = \sum_{i=1}^{\infty} \sqrt{\mu_i} w_i(t) e_i(x)$, where $\{w_i(t)\}$ is a sequence of independent, identically distributed standard Brownian motions defining a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$, and the set $\{e_i(x)\}$ is a complete orthonormal basis on $L^2(\mathcal{O})$. Then $\mathbb{E}W(x,t) = 0$, $\mathbb{E}[W(x_1,s)W(x_2,t)] = r(x_1,x_2)(t \wedge s)$, where $r(x_1,x_2)$ is symmetric kernel of operator $\mathscr{R}$. For a continuous adapted random process $\sigma(x,t)$, $\mathbb{E}\{\int_0^t \int_{\mathcal{O}} \sigma(x,s)dW_s dx\}^2 = \mathbb{E}\int_0^t \int_{\mathcal{O}} r(x,x)\sigma(x,s)^2 dx\,dt$. In this paper, we assume covariance function $r(x_1,x_2)$ is bounded, that is, $r(x_1,x_2) \leqslant r_0$, $\forall x_1, x_2 \in \mathcal{O}$.

Initial value and boundary value conditions of (1) satisfied

$$y(x,s) = \phi(x,s), \quad s \in [-\tau, 0], \tag{2}$$

$$y(x,t) = 0, \quad (x,t) \in \partial\mathcal{O} \times [-h,\infty) \tag{3}$$

or

$$\frac{\partial y(x,t)}{\partial \nu} + Ny(x,t) = 0, \quad (x,t) \in \partial\mathcal{O} \times [-h,\infty), \tag{4}$$

where ν is the unit outward normal vector of $\partial\mathcal{O}$ and N is positive constant. $\phi(x,s)$ is continuous adapted random process and $\mathbb{E}\|\phi\|_C^2 = \mathbb{E}\{\sup_{-\tau\leqslant s\leqslant 0}\|\phi(\cdot,s)\|^2\} < \infty$.

Definition 1. The equilibrium point $y(0,t) \equiv 0$ of the system (1) is said to be mean-square exponentially stable with a decay rate $\delta > 0$ if there exist positive constants c such that $\mathbb{E}\|y(\cdot,t)\|^2 \leqslant ce^{-2\delta t}\mathbb{E}\|\phi\|_C^2$, $t \geqslant 0$.

In this paper, our aim is to develop techniques of robust stochastic stabilization and robust H_∞ control for stochastic partial differential time delay systems (1). More specifically, we are concerned with the following two problems.

(1) Stochastic exponential stabilization problem: design a state feedback controller

$$u(x,t) = Ky(x,t) \tag{5}$$

for systems (1) with initial boundary value condition (2) and (3) or (4) with $v(x,t) = 0$ such that the resulting closed-loop system is mean-square exponentially stable.

(2) Robust H_∞ control problem: given a constant scalar $\gamma > 0$, design a state feedback controller in the form of (5) such that the resulting closed-loop system is mean square exponentially stable, and for any nonzero $v(x,t) \in \mathscr{L}_2(\mathcal{O} \times [0,\infty); \mathbb{R}^{n_v})$, $\phi(x,s) = 0$, $s \in [-\tau, 0]$, and we have

$$\begin{aligned} & \mathbb{E}\int_0^{\infty}\int_{\mathcal{O}} z^T(x,t)z(x,t)\,dx\,dt \\ & \quad \leqslant \gamma^2 \int_0^{\infty}\int_{\mathcal{O}} v^T(x,t)v(x,t)\,dx\,dt. \end{aligned} \tag{6}$$

We conclude this section by recalling the following lemmas which will be used in the proof of our main results.

Lemma 2 (Schur complement [29]). *Given constant matrices S_1, S_2, S_3 where $S_1 = S_1^T$, and $0 < S_2 = S_2^T$, then $S_1 + S_3^T S_2^{-1} S_3 < 0$ if and only if*

$$\begin{bmatrix} S_1 & S_3^T \\ S_3 & -S_2 \end{bmatrix} < 0 \quad or \quad \begin{bmatrix} -S_2 & S_3 \\ S_3^T & S_1 \end{bmatrix} < 0. \tag{7}$$

3. Mean Square Exponential Stabilization

In this section, an LMI approach is developed to solve the problem of exponential stabilization formulated in the previous section. The main result is given in the following theorem.

Theorem 3. *Consider the stochastic time-delay partial differential system*

$$\begin{aligned} dy(x,t) &= [D\Delta y(x,t) + Ay(x,t) \\ &\quad + A_1 y(x,t-\tau) + Bu(x,t)]\, dt \\ &\quad + [Cy(x,t) + C_1 y(x,t-\tau)]\, dW(x,t), \end{aligned} \tag{8}$$

with (2) *and* (3) *being its initial and boundary value conditions, respectively. Then system* (8) *is exponential stabilizable in mean square with decay rate δ if there exist matrices $Y \in \mathbb{R}^{n_u \times n}$ and $X > 0$, $Z > 0$, such that the following LMI holds:*

$$\Gamma = \begin{bmatrix} \Gamma_{11} & A_1 Z & \sqrt{r_0} X C^T & X \\ * & \Gamma_{22} & \sqrt{r_0} Z C_1^T & 0 \\ * & * & -X & 0 \\ * & * & * & -Z \end{bmatrix} < 0, \tag{9}$$

where

$$\begin{aligned} \Gamma_{11} &= 2\delta X + AX + XA^T + BY + Y^T B^T, \\ \Gamma_{22} &= \Pi_{22} = -e^{-2\delta\tau} Z. \end{aligned} \tag{10}$$

In this case, a stabilizing state feedback controller can be chosen by $u(x,t) = YX^{-1} y(x,t)$.

Proof. Applying the controller in (5) to system (8), we obtain the closed-loop system as

$$\begin{aligned} dy(x,t) &= [D\Delta y(x,t) + Ay(x,t) + A_1 y(x,t-\tau) + BKy(x,t)]\, dt \\ &\quad + [Cy(x,t) + C_1 y(x,t-\tau)]\, dW(x,t). \end{aligned} \tag{11}$$

For given decay rate δ, choose a Lyapunov functional candidate for system (11) as

$$\begin{aligned} V(t,y) &= V_1 + V_2 \\ &= e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) Py(x,t)\, dx \\ &\quad + \int_{t-\tau}^{t} \int_{\mathcal{O}} e^{2\delta s} y^T(x,s) Qy(x,s)\, dx\, ds, \end{aligned} \tag{12}$$

where P, Q are a pair of positive symmetric matrices. Then, by Itô's formula, the stochastic differential $dV(t,y)$ along (11) can be obtained as (see, e.g., [16, 18, 20])

$$\begin{aligned} dV(t,y) &= LV(t,y)\, dt \\ &\quad + 2 \int_{\mathcal{O}} y^T(x,t) \\ &\qquad \times [Cy(x,t) + C_1 y(x,t-\tau)]\, dW(x,t), \end{aligned} \tag{13}$$

where $LV(t,y) = LV_1(t,y) + LV_2(t,y)$.

We can deduce that

$$\begin{aligned} LV_1(t,y) &\leqslant 2\delta e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) Py(x,t)\, dx \\ &\quad + e^{2\delta t} \int_{\mathcal{O}} [D\Delta y(x,t) + Ay(x,t) + A_1 y(x,t_h) \\ &\qquad + BKy(x,t)]^T Py(x,t)\, dx \\ &\quad + e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) P [D\Delta y(x,t) + Ay(x,t) \\ &\qquad + A_1 y(x,t_h) \\ &\qquad + BKy(x,t)]\, dx \\ &\quad + r_0 e^{2\delta t} \int_{\mathcal{O}} [Cy(x,t) + C_1 y(x,t-\tau)]^T \\ &\qquad \times P [Cy(x,t) + C_1 y(x,t_h)]\, dx \\ &= 2e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) Py(x,t)\, dx \\ &\quad + e^{2\delta t} \int_{\mathcal{O}} \left[(\Delta y(x,t))^T \cdot DPy(x,t) \right. \\ &\qquad \left. + y^T(x,t) PD\Delta y(x,t) \right] dx \\ &\quad + e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) \left[A^T P + PA + (BK)^T P + PBK \right. \\ &\qquad \left. + r_0 C^T PC \right] y(x,t)\, dx \\ &\quad + e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) \left(A_1^T P + PA_1 \right. \\ &\qquad \left. + r_0 C^T PC_1 + r_0 C_1^T PC \right) \\ &\qquad \times y(x,t-\tau)\, dx \\ &\quad + r_0 e^{2\delta t} \int_{\mathcal{O}} y^T(x,t-\tau) C_1^T PC_1 y(x,t-\tau)\, dx, \end{aligned} \tag{14}$$

$$\begin{aligned} LV_2(t,y) &= e^{2\delta t} \int_{\mathcal{O}} y^T(x,t) Qy(x,t)\, dx \\ &\quad - e^{2\delta(t-\tau)} \int_{\mathcal{O}} y^T(x,t-\tau) Qy(x,t_h)\, dx \\ &= e^{2\delta t} \left(\int_{\mathcal{O}} y^T(x,t) Qy(x,t)\, dx \right. \\ &\qquad \left. - e^{2\delta\tau} \int_{\mathcal{O}} y^T(x,t-\tau) Qy(x,t-\tau)\, dx \right). \end{aligned} \tag{15}$$

Considering Dirichlet boundary condition (3) and using Green formula, we have

$$\begin{aligned} &\int_{\mathcal{O}} (\Delta y(x,t))^T DPy(x,t)\, dx \\ &\quad = \int_{\partial\mathcal{O}} y(x,t)^T DP \frac{\partial y(x,t)}{\partial \nu} dx \\ &\qquad - \int_{\mathcal{O}} (\nabla y(x,t))^T DP \nabla y(x,t)\, dx \\ &\quad = -D \int_{\mathcal{O}} (\nabla y(x,t))^T P \nabla y(x,t)\, dx. \end{aligned} \tag{16}$$

Because $D > 0$ and P is positive definite matrix, then by (16) we have

$$e^{2\delta t} \int_{\mathcal{O}} \left[(\Delta y(x,t))^T DPy(x,t) + y^T(x,t) PD\Delta y(x,t) \right] dx \leqslant 0 \quad (17)$$

which together with (14)~(17) yields

$$LV(t,y) \leqslant e^{2\delta t} \int_{\mathcal{O}} \eta(x,t)^T \overline{\Gamma} \eta(x,t)\, dx, \quad (18)$$

where $\eta^T(x,t) = [y^T(x,t), y^T(x,t_h)]$,

$$\begin{aligned} \overline{\Gamma} &= \begin{bmatrix} \overline{\Gamma}_{11} & \overline{\Gamma}_{12} \\ * & \overline{\Gamma}_{22} \end{bmatrix}, \\ \overline{\Gamma}_{11} &= 2\delta P + PA + A^T P + PBK + (BK)^T P + Q + r_0 C^T PC, \\ \overline{\Gamma}_{12} &= PA_1 + r_0 C^T PC_1, \\ \overline{\Gamma}_{22} &= -e^{-2\delta\tau} Q + r_0 C_1^T PC_1. \end{aligned} \quad (19)$$

So if $\overline{\Gamma} < 0$, then (18) implies $LV(t,y) < 0$. By Lemma 2, $\overline{\Gamma} < 0$ is equivalent to

$$\widetilde{\Gamma} = \begin{bmatrix} \widetilde{\Gamma}_{11} & PA_1 & \sqrt{r_0} C^T P \\ * & \widetilde{\Gamma}_{22} & \sqrt{r_0} C_1^T P \\ * & * & -P \end{bmatrix} < 0, \quad (20)$$

where

$$\begin{aligned} \widetilde{\Gamma}_{11} &= 2\delta P + PA + A^T P + PBK + (BK)^T P + Q, \\ \widetilde{\Gamma}_{22} &= -e^{-2\delta\tau} Q. \end{aligned} \quad (21)$$

Then, pre- and postmultiplying the LMI in (20) by diag $\{P^{-1}, Q^{-1}, P^{-1}\}$ and let $X = P^{-1}$, $Z = Q^{-1}$, $Y = KX$, we have

$$\widehat{\Gamma} = \begin{bmatrix} \widehat{\Gamma}_{11} & A_1 Z & \sqrt{r_0} X C^T \\ * & \widehat{\Gamma}_{22} & \sqrt{r_0} Z C_1^T \\ * & * & -X \end{bmatrix} < 0, \quad (22)$$

where

$$\begin{aligned} \widehat{\Gamma}_{11} &= 2\delta X + AX + XA^T + BY + Y^T B^T + XZ^{-1} X, \\ \widehat{\Gamma}_{22} &= \Pi_{22} = -e^{-2\delta\tau} Z. \end{aligned} \quad (23)$$

Therefore, by Lemma 2 again, if the matrix inequality in (9) holds, then the inequality (22) or, equivalently, (20) holds, which leads to $\overline{\Gamma} < 0$. Hence $LV(t,y) < 0$.

Now we prove system (8) to be mean square exponentially stable. Integrating both sides of (13) and taking expectation, we have

$$\begin{aligned} \mathbb{E}V(t, y(x,t)) &= \mathbb{E}V(0, y(x,0)) + \mathbb{E}\Big\{ \int_0^t LV(s, y(x,s))\, ds \\ &\quad + 2\int_0^t \int_{\mathcal{O}} y(x,s)^T P \times (Cy(x,s) + C_1 y(x,s_h))\, dW_s dx \Big\} \\ &\leqslant \mathbb{E}V(0, y(x,0)) + \mathbb{E}\int_0^t LV(s, y(x,s))\, ds \leqslant \mathbb{E}V(0, y_0), \end{aligned} \quad (24)$$

$$\begin{aligned} \mathbb{E}V(0, y(x,0)) &= \mathbb{E}\int_{\mathcal{O}} y(x,0)^T P y(x,0)\, dx + \mathbb{E}\int_{-\tau}^0 \int_{\mathcal{O}} e^{2\delta s} y(x,s)^T Q y(x,s)\, dx\, ds \\ &\leqslant \lambda_{\max}(P)\mathbb{E}\|\phi\|_C^2 + \lambda_{\max}(Q)\mathbb{E}\|\phi\|_C^2 \int_{-\tau}^0 e^{2\delta s} ds \leqslant c_1 \mathbb{E}\|\phi\|_C^2, \end{aligned} \quad (25)$$

where constant $c_1 = [\lambda_{\max}(P) + \lambda_{\max}(Q)((1 - e^{-2\delta\tau})/2\delta)]$.

On the other hand, by (12), we have

$$V(t, y(x,t)) \geqslant e^{2\delta t} \int_{\mathcal{O}} y(x,t)^T P y(x,t)\, dx, \quad (26)$$

and, hence, (24) and (25) yield

$$\begin{aligned} \lambda_{\min}(P) e^{2\delta t} \mathbb{E}\int_{\mathcal{O}} y^T(x,t) y(x,t)\, dx &\leqslant e^{2\delta t} \mathbb{E}\int_{\mathcal{O}} y^T(x,t) P y(x,t)\, dx \leqslant \mathbb{E}V(t, y(x,t)) \\ &\leqslant \mathbb{E}V(0, y_0). \end{aligned} \quad (27)$$

Then from (25) and (27), we have

$$\mathbb{E}\|y(\cdot,t)\|^2 = \mathbb{E}\int_{\mathcal{O}} y(x,t)^T y(x,t)\, dx \leqslant c\mathbb{E}\|\phi\|_C^2 e^{-2\delta t}, \quad (28)$$

where $c = c_1/\lambda_{\min}(P)$ therefore, by Definition 1, system (8) is mean square exponentially stable with decay rate δ. The proof of Theorem 3 is complete. □

If the boundary value condition of system (8) is replaced by Robin boundary condition (4), then

$$\begin{aligned}
&\int_{\mathscr{O}} (\Delta y(x,t))^T DPy(x,t)\,dx \\
&= \int_{\partial\mathscr{O}} (y(x,t))^T DP\frac{\partial y(x,t)}{\partial \nu}dx \\
&\quad - \int_{\mathscr{O}} (\nabla y(x,t))^T DP\nabla y(x,t)\,dx \\
&\leqslant \int_{\partial\mathscr{O}} (y(x,t))^T DP\frac{\partial y(x,t)}{\partial \nu}dx \\
&= -D\int_{\partial\mathscr{O}} y^T(x,t)\,PNy(x,t)\,dx \leqslant 0.
\end{aligned} \tag{29}$$

Substituting (29) into (14), similar to the proof of Theorem 3, we can obtain the following.

Theorem 4. *Consider the stochastic partial differential system* (8) *whose initial condition is* (2) *and boundary value* (4)*. Then the system is mean square exponentially stabilizable if there exist matrices $Y \in \mathbb{R}^{n_u\times n}$ and $X > 0$, $Z > 0$, such that the following LMI holds:*

$$\Phi = \begin{bmatrix} \Phi_{11} & A_1Z & \sqrt{r_0}XC^T & X \\ * & \Phi_{22} & \sqrt{r_0}ZC_1^T & 0 \\ * & * & -X & 0 \\ * & * & * & -Z \end{bmatrix} < 0, \tag{30}$$

where

$$\begin{aligned}
\Phi_{11} &= 2\delta X + AX + XA^T + BY + Y^TB^T, \\
\Phi_{22} &= \Pi_{22} = -e^{-2\delta\tau}Z.
\end{aligned} \tag{31}$$

In this case, a stabilizing state feedback controller can be chosen by $u(x,t) = YX^{-1}y(x,t)$.

4. Robust H_∞ Control

In this section, a sufficient condition for the solvability of the robust H_∞ control problem is proposed and an LMI approach for designing the desired state feedback controllers is developed. Now, we are ready to give our main result in this paper as follows.

Theorem 5. *Given a scalar $\gamma > 0$, then under initial boundary value conditions* (2) *and* (3)*, the stochastic partial differential system* (1) *is robust mean-square exponentially stabilizable with disturbance attenuation $\gamma > 0$ if there exist matrices $Y \in \mathbb{R}^{n_u\times n}$ and $X > 0$, $Z > 0$, such that the following LMI holds:*

$$\Upsilon = \begin{bmatrix} \Upsilon_{11} & A_1Z & B_\upsilon & \sqrt{r_0}XC^T & XL^T & X \\ * & \Upsilon_{22} & 0 & \sqrt{r_0}ZC_1^T & 0 & 0 \\ * & * & -\gamma^2 I & \sqrt{r_0}C_\upsilon^T & 0 & 0 \\ * & * & * & -X & 0 & 0 \\ * & * & * & * & -I & 0 \\ * & * & * & * & * & -Z \end{bmatrix} < 0, \tag{32}$$

where

$$\begin{aligned}
\Upsilon_{11} &= 2\delta X + AX + XA^T + BY + Y^TB^T, \\
\Upsilon_{22} &= -e^{-2\delta\tau}Z.
\end{aligned} \tag{33}$$

Then a suitable robust H_∞ controller can be chosen by $u(x,t) = YX^{-1}y(x,t)$.

Proof. Obviously, by Lemma 2, if (32) holds, then (9) also holds. Therefore, by Theorem 3, system (1) is mean square exponentially stabilizable if $\upsilon(x,t) = 0$. Next, we shall show that under the zero-initial condition, (6) holds for nonzero $\upsilon(x,t) \in \mathscr{L}_2(\mathscr{O}\times[0,\infty);\mathbb{R}^{n_\upsilon})$.

We consider the Lyapunov functional $V(t,y)$ in (12), by Itô's formula [16, 18, 20],

$$\begin{aligned}
dV(t,y) &= L_\upsilon V(t,y)\,dt \\
&\quad + 2\int_{\mathscr{O}} y^T(x,t)\left[Cy(x,t) + C_1y(x,t_h)\right. \\
&\qquad \left. + C_\upsilon \upsilon(x,t)\right] dW(x,t).
\end{aligned} \tag{34}$$

Integrating $dV(t,y)$ from 0 to t and taking expectation, we can obtain that

$$\mathbb{E}V(t,y) = \mathbb{E}\int_0^t L_\upsilon V(s,y)\,ds. \tag{35}$$

We can calculate that

$$\begin{aligned}
L_\upsilon V(t,y) &\leqslant LV_1(t,y) + LV_2(t,y) \\
&\quad + r_0e^{2\delta t}\int_{\mathscr{O}} y(x,t)^T\left(PB_\upsilon + B_\upsilon^TP\right)\upsilon(x,t)\,dx \\
&\quad + r_0e^{2\delta t}\int_{\mathscr{O}} y(x,t)^T\left(C^TPC_\upsilon + C_\upsilon^TPC\right) \\
&\qquad \times \upsilon(x,t)\,dx \\
&\quad + r_0e^{2\delta t}\int_{\mathscr{O}} y(x,t_h)^T\left(C_1^TPC_\upsilon + C_\upsilon^TPC_1\right) \\
&\qquad \times \upsilon(x,t)\,dx \\
&\quad + r_0e^{2\delta t}\int_{\mathscr{O}} \upsilon(x,t)^TC_\upsilon^TPC_\upsilon\upsilon(x,t)\,dx,
\end{aligned} \tag{36}$$

where $LV_1(t,y)$ and $LV_2(t,y)$ satisfy (14) and (15), respectively. Then

$$L_\upsilon V(t,y) \leqslant e^{2\delta t}\int_{\mathscr{O}} \zeta(x,t)^T\overline{\Upsilon}\zeta(x,t)\,dx, \tag{37}$$

where $\zeta^T(x,t) = [y^T(x,t)\ \ y^T(x,t_h)\ \ \upsilon^T(x,t)]$, and

$$\overline{\Upsilon} = \begin{bmatrix} \overline{\Upsilon}_{11} & \overline{\Upsilon}_{12} & r_0\left(PB_\upsilon + C^TPC_\upsilon\right) \\ * & \overline{\Upsilon}_{22} & r_0C_1^TPC_\upsilon \\ * & * & r_0C_\upsilon^TPC_\upsilon \end{bmatrix}, \tag{38}$$

$$\overline{\Upsilon}_{ij} = \overline{\Gamma}_{ij}, \quad i,j = 1,2. \tag{39}$$

Let

$$J=\mathbb{E}\left\{\int_0^t\int_{\mathscr{O}}\left(z^T(x,t)\,z(x,t)\right.\right.$$
$$\left.\left.-\gamma^2 v^T(x,t)\,v(x,t)\right)dx\,dt\right\}, \tag{40}$$

and then (6) is equivalent to $J<0$. Moreover, by (35)~(38), it follows that

$$\begin{aligned}J&=\mathbb{E}\left\{\int_0^t\int_{\mathscr{O}}\left(z^T(x,t)\,z(x,t)-\gamma^2 v^T(x,t)\,v(x,t)\right)dx\,dt\right\}\\&\leqslant \mathbb{E}\left\{\int_0^t\int_{\mathscr{O}}\left(z^T(x,t)\,z(x,t)\,dx\,dt-\gamma^2 v^T(x,t)\,v(x,t)\right)dx\,dt\right\}\\&\quad+e^{-2\delta t}\left(\mathbb{E}\int_0^t LV_v(s,y)\,ds-\mathbb{E}V(t,y)\right)\\&\leqslant \mathbb{E}\left\{\int_0^t\int_{\mathscr{O}}\left(z^T(x,t)\,z(x,t)\,dx\,dt-\gamma^2 v^T(x,t)\,v(x,t)\right)dx\,dt\right\}\\&\quad+e^{-2\delta t}\mathbb{E}\int_0^t LV_v(s,y)\,ds\\&\leqslant \mathbb{E}\left\{\int_0^t\int_{\mathscr{O}}\zeta(x,t)^T\Upsilon\zeta(x,t)\,dx\right\},\end{aligned} \tag{41}$$

where

$$\widehat{\Upsilon}=\begin{bmatrix}\widehat{\Upsilon}_{11} & \widehat{\Upsilon}_{12} & r_0PB_v+r_0C^TPC_v\\ * & \widehat{\Upsilon}_{22} & r_0C_1^TPC_v\\ * & * & -\gamma^2I+r_0C_v^TPC_v\end{bmatrix} \tag{42}$$

with $\widehat{\Upsilon}_{ij}=\overline{\Upsilon}_{ij}=\overline{\Gamma}_{ij}$, $i,j=1,2$.

According to Lemma 2, $\widehat{\Upsilon}<0$ is equivalent to

$$\widetilde{\Upsilon}=\begin{bmatrix}\widetilde{\Upsilon}_{11} & PA_1 & PB_v & \sqrt{r_0}C^TP\\ * & \widetilde{\Upsilon}_{22} & 0 & \sqrt{r_0}C_1^TP\\ * & * & -\gamma^2I & \sqrt{r_0}C_v^TP\\ * & * & * & -P\end{bmatrix}<0, \tag{43}$$

where

$$\begin{aligned}\widetilde{\Upsilon}_{11}&=2\delta P+PA+A^TP\\&\quad+PBK+(BK)^TP+Q+L^TL,\\ \widetilde{\Upsilon}_{22}&=-e^{-2\delta\tau}Q.\end{aligned} \tag{44}$$

In (43), pre- and postmultiplying the LMI by diag $\{P^{-1},Q^{-1},I,P^{-1}\}$ and letting $X=P^{-1}$, $Z=Q^{-1}$, $Y=KX^{-1}$ and following the same line as in the proof of Theorem 3, we can deduce that (32) is equivalent to (43) and $\widehat{\Upsilon}<0$, which together with (41) implies that $J<0$. Therefore, the inequality (6) holds. This completes the proof. □

Remark 6. Similar to the proof of Theorem 3, in order to calculate the L_vV, the proof of Theorem 5 has been used in Itô's formula of infinite dimensional version (see [18]).

If boundary value condition becomes (4), then similar to the proof of Theorem 5, we have the following.

Theorem 7. *Given a scalar $\gamma>0$, then under initial boundary value conditions (2) and (4), the stochastic partial differential system (1) is robust stabilizable with disturbance attenuation γ if there exist matrices $Y\in\mathbb{R}^{n_u\times n}$ and $X>0$, $Z>0$, such that the following LMI holds:*

$$\Lambda=\begin{bmatrix}\Lambda_{11} & A_1Z & B_v & \sqrt{r_0}XC^T & XL^T & X\\ * & \Lambda_{22} & 0 & \sqrt{r_0}ZC_1^T & 0 & 0\\ * & * & -\gamma^2I & \sqrt{r_0}C_v^T & 0 & 0\\ * & * & * & -X & 0 & 0\\ * & * & * & * & -I & 0\\ * & * & * & * & * & -Z\end{bmatrix}<0, \tag{45}$$

where

$$\begin{aligned}\Lambda_{11}&=2\delta X+AX+XA^T+BY+Y^TB^T,\\ \Lambda_{22}&=-e^{-2\delta\tau}Z.\end{aligned} \tag{46}$$

Then a suitable robust H_∞ controller can be chosen by $u(x,t)=YX^{-1}y(x,t)$.

5. An Illustrative Example

In this section, we provide an illustrative example to demonstrate the effectiveness of the proposed method.

Consider the stochastic partial differential system with time delay in (1) under initial boundary value conditions (2) and (3). We let $\mathscr{O}=[0,1]$, that is, $0\leqslant x\leqslant 1$, $D=1$, $y(x,t)=[y_1(x,t),y_2(x,t)]^T$. Then $l=1$, $m=1$, $n=2$, $d=1$. Let time delay $\tau=0.1$, decay rate $\delta=0.1$, and the upper bound of covariance function is $r_0=1$,

$$\begin{aligned}A&=\begin{bmatrix}-0.8 & 0.5\\ 0.2 & -0.5\end{bmatrix}, & A_1&=\begin{bmatrix}-0.4 & 0.1\\ 0.1 & -0.6\end{bmatrix},\\ B&=\begin{bmatrix}-0.5 & 0.2\\ 0.2 & 0.4\end{bmatrix}, & B_v&=\begin{bmatrix}0.4 & 0.1\\ 0.2 & -0.5\end{bmatrix},\\ C&=\begin{bmatrix}0.6 & 0.4\\ 0.1 & -0.5\end{bmatrix}, & C_1&=\begin{bmatrix}-0.5 & -0.2\\ 0 & 0.4\end{bmatrix},\\ L&=\begin{bmatrix}-0.5 & 0\\ 0 & 0.1\end{bmatrix}, & C_v&=\begin{bmatrix}-0.2 & 0.1\\ 0.1 & 0.2\end{bmatrix}.\end{aligned} \tag{47}$$

In this example, attention is focused on the design of a state feedback controller, the resulting closed-loop system is robustly stochastically mean square exponential stable with disturbance attenuation $\gamma = 0.8$. For this purpose, we use the Matlab LMI Control Toolbox to solve the LMI (32) and obtain the solution as follows:

$$X = \begin{bmatrix} 2.8437 & -0.4194 \\ -0.4194 & 2.7987 \end{bmatrix},$$
$$Y = \begin{bmatrix} -8.6142 & 77.4013 \\ -94.8986 & 0.9777 \end{bmatrix}, \quad (48)$$
$$Z = \begin{bmatrix} 1.2837 & -0.4050 \\ -0.4050 & 1.4857 \end{bmatrix}.$$

Therefore, by Theorem 5, it can be seen that the robust H_∞ control problem is solvable and a desired state feedback control law can be chosen as

$$u(x,t) = \begin{bmatrix} -7.2557 & -28.7426 \\ -34.0280 & -4.7460 \end{bmatrix} y(x,t). \quad (49)$$

6. Conclusions

In this paper, the problems of robust stochastic exponential stabilization and robust H_∞ control for liner stochastic partial differential systems with time delay have been studied under Dirichlet and Robin boundary, respectively. An LMI approach has been developed to design state feedback controllers, which not only guarantees mean square exponential stability of the closed-loop system but also reduces the effect of the disturbance input on the controlled output to a prescribed level. A numerical example has been given to show the effectiveness of the proposed method.

Conflict of Interests

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 61364006 and 61273126), PhD Start-up Fundation of Guangxi University of Science and Technology (no. 03081520), and Guangxi Higher Education Science Research Projection (no. 201203YB125).

References

[1] G. Zames, "Feedback and optimal sensitivity: model reference transformation, multiplicative seminorms, and approximative inverses," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 301–320, 1981.

[2] W. Zhang, B. Chen, L. Sheng, and M. Gao, "Robust H_2/H_∞ filter design for a class of nonlinear stochastic systems with state-dependent noise," *Mathematical Problems in Engineering*, vol. 2012, Article ID 750841, 16 pages, 2012.

[3] J. Li, L. Wei, C. Liu, and H. Yu, "H_∞ control of network-based systems with packet disordering and packet loss compensation," *Mathematical Problems in Engineering*, vol. 2013, Article ID 836314, 9 pages, 2013.

[4] C. Tseng and B. Chen, "H_∞ fuzzy estimation for a class of nonlinear discrete-time dynamic systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2605–2619, 2001.

[5] W. Zhang, B. Chen, and C. Tseng, "Robust H_∞ filtering for nonlinear stochastic systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 2 I, pp. 589–598, 2005.

[6] B. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[7] S. Xu, P. Shi, Y. Chu, and Y. Zou, "Robust stochastic stabilization and H_∞ control of uncertain neutral stochastic time-delay systems," *Journal of Mathematical Analysis and Applications*, vol. 314, no. 1, pp. 1–16, 2006.

[8] R. Curtain and H. Zwart, *An Introduction to Infinite Dimensional Linear Systems*, Springer, Heidelberg, Germany, 1995.

[9] J. Wu, *Theory and Applications of Partial Functional Differential Equations*, Springer, New York, NY, USA, 1996.

[10] A. Bensoussan, G. da Prato, M. Delfour, and S. Mitter, *Representation and Control of Infinite Dimensional Systems: Volume 2*, Birkhauser, Boston, Mass, USA, 1993.

[11] O. A. Ladyzhenskaya and N. N. Uraltseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, London, UK, 1968.

[12] C. Foias, H. Ozbay, and A. Tannenbaum, *Robust Control of Infinite Dimensional Systems: Frequency Domain Methods*, Springer, London, UK, 1996.

[13] E. Fridman and Y. Orlov, "An LMI approach to H_∞ boundary control of semilinear parabolic and hyperbolic systems," *Automatica*, vol. 45, no. 9, pp. 2060–2066, 2009.

[14] E. Fridman and Y. Orlov, "Exponential stability of linear distributed parameter systems with time-varying delays," *Automatica*, vol. 45, no. 1, pp. 194–201, 2009.

[15] C. Lin and B. Chen, "Robust observer-based control of large flexible structures," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 116, no. 4, pp. 713–722, 1994.

[16] K. Liu, *Stability of Infinite Dimensional Stochastic Differential Equations with Applications*, vol. 135 of *Monographs and Surveys in Pure and Applied Mathematics*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2006.

[17] B. Chen, W. Chen, and W. Zhang, "Robust filter for nonlinear stochastic partial differential systems in sensor signal processing: fuzzy approach," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 5, pp. 957–970, 2012.

[18] P. L. Chow, *Stochastic Partial Differential Equations*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2007.

[19] G. da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, vol. 44 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, Cambridge, UK, 1992.

[20] K. Liu and A. Truman, "Moment and almost sure Lyapunov exponents of mild solutions of stochastic evolution equations with variable delays via approximation approaches," *Journal of Mathematics of Kyoto University*, vol. 41, no. 4, pp. 749–768, 2002.

[21] T. Taniguchi, "The exponential stability for stochastic delay partial differential equations," *Journal of Mathematical Analysis and Applications*, vol. 331, no. 1, pp. 191–205, 2007.

[22] P. Li and N. U. Ahmed, "Feedback stabilization of some nonlinear stochastic systems on Hilbert space," *Nonlinear Analysis*, vol. 17, no. 1, pp. 31–43, 1991.

[23] T. E. Duncan, B. Maslowski, and B. Pasik-Duncan, "Adaptive boundary and point control of linear stochastic distributed parameter systems," *SIAM Journal on Control and Optimization*, vol. 32, no. 3, pp. 648–672, 1994.

[24] X. Y. Zhou, "On the necessary conditions of optimal controls for stochastic partial differential equations," *SIAM Journal on Control and Optimization*, vol. 31, no. 6, pp. 1462–1478, 1993.

[25] S. G. Tzafestas and J. M. Nightingale, "Optimal control of a class linear stochastic distributed parameter systems," *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 8, pp. 1213–1220, 1968.

[26] B. Øksendal, "Optimal control of stochastic partial differential equations," *Stochastic Analysis and Applications*, vol. 23, no. 1, pp. 165–179, 2005.

[27] Y. T. Kim, "Parameter estimation in infinite-dimensional stochastic differential equations," *Statistics and Probability Letters*, vol. 45, no. 3, pp. 195–204, 1999.

[28] K. Ikeda, T. Azuma, and K. Uchida, "Infinite-dimensional LMI approach to analysis and synthesis for linear time-delay systems," *Kybernetika*, vol. 37, no. 4, pp. 505–520, 2001.

[29] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequality in Systems and Control Theory*, SIAM Frontier Series, SIAM, 1994.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 948134, 10 pages
http://dx.doi.org/10.1155/2014/948134

Research Article

Nonlinear Stochastic H_∞ Control with Markov Jumps and (x, u, v)-Dependent Noise: Finite and Infinite Horizon Cases

Li Sheng,[1] Meijun Zhu,[1] Weihai Zhang,[2] and Yuhong Wang[1]

[1] *College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China*
[2] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Li Sheng; victory8209@163.com

Received 8 April 2014; Accepted 1 June 2014; Published 18 June 2014

Academic Editor: Xuejun Xie

This paper is concerned with the H_∞ control problem for nonlinear stochastic Markov jump systems with state, control, and external disturbance-dependent noise. By means of inequality techniques and coupled Hamilton-Jacobi inequalities, both finite and infinite horizon H_∞ control designs of such systems are developed. Two numerical examples are provided to illustrate the effectiveness of the proposed design method.

1. Introduction

H_∞ control is one of the most important robust control approaches, which can efficiently eliminate the effect of the exogenous disturbance [1, 2]. Since Hinrichsen and Pritchard introduced H_∞ control to linear stochastic systems [3], the nonlinear stochastic H_∞ control and filtering problems have received considerable attention in both theory and practical applications [4–9]. In [4], the nonlinear stochastic H_∞ designs were first developed by solving a second-order nonlinear Hamilton-Jacobi inequality. The H_∞ filtering problems for general nonlinear continuous-time and discrete-time stochastic systems were discussed in [6] and [7], respectively. In [8], the quantized H_∞ control problem for a class of nonlinear stochastic time-delay network-based systems with probabilistic data missing is studied.

On the other hand, Itô stochastic systems with Markov jumps have attracted increasing attention due to their powerful modeling ability in many fields [10, 11]. For linear stochastic systems with Markov jumps, many important issues have been studied, such as stability and stabilization [12, 13], observability and detectability [14], optimal control [15], and H_2/H_∞ control [16]. The control issues for nonlinear stochastic Markov jump systems (NSMJSs) have also been widely investigated. In [17], the notion of exponential dissipativity of NSMJSs was introduced, and it was used to estimate the possible variations of the output feedback control. In [18], the stabilization of nonlinear Markov jump systems with partly unknown transition probabilities was studied via fuzzy control. In [19], the H_∞ control problems of NSMJSs were studied, which extended the results of [4] to stochastic systems with Markovian jump parameters.

Most of the existing literature was concerned with stochastic Markov jump systems with state-dependent noise or both state and disturbance dependent noise ((x, v)-dependent noise for short) [16, 19]. However, for most natural phenomena described by Itô stochastic systems, not only state but also control input or external disturbance maybe corrupted by noise. By introducing three coupled Hamilton-Jacobi equations (HJEs), the finite and infinite horizon H_∞ control problems were solved for Itô stochastic systems with all system state, control, and disturbance-dependent noise ((x, u, v)-dependent noise for short) in [20] and [21], respectively. In [22], the finite/infinite horizon H_∞ control of nonlinear stochastic systems with (x, u, v)-dependent noise was solved by means of a Hamilton-Jacobi inequality (HJI) instead of three coupled Hamilton-Jacobi equations (HJEs). However, the control problems of nonlinear stochastic systems with Markov jumps and (x, u, v)-dependent noise have never been tackled and deserved further research.

In this paper, the finite and infinite horizon H_∞ control problems are studied for nonlinear stochastic Markov jumps systems with (x, u, v)-dependent noise. Firstly, a very useful elementary identity is proposed. Then, by using the technique of squares completion, a sufficient condition for finite/infinite horizon H_∞ control of NSMJSs is presented based on a set of coupled HJIs. By means of linear matrix inequalities (LMIs), a sufficient condition for infinite horizon H_∞ control of linear stochastic Markov jump systems is derived. Finally, two numerical examples are provided to show the effectiveness of our obtained results.

For conveniences, we make use of the following notations throughout this paper. $\mathscr{R}^n$ is the n-dimensional Euclidean space. $\mathscr{R}^{n\times m}$ is the set of all $n \times m$ real matrices. $A > 0$ ($A \geq 0$): A is a positive definite (positive semi-definite) symmetric matrix. A' is the transpose of a matrix A. I is the identity matrix. $\|x\|$ is the Euclidean norm of a vector x. $L^2_{\mathscr{F}}([0,T];\mathscr{R}^l)$ (resp., $L^2_{\mathscr{F}}(\mathscr{R}_+;\mathscr{R}^l)$) is the space of nonanticipative stochastic processes $y(t) \in \mathscr{R}^l$ with respect to increasing σ-algebras $\mathscr{F}_t$ ($t \geq 0$) satisfying $\|y(t)\|_{L^2_{\mathscr{F}}([0,T];\mathscr{R}^l)} = (E\int_0^T \|y(t)\|^2 dt)^{1/2} < \infty$ (resp., $\|y(t)\|_{L^2_{\mathscr{F}}(\mathscr{R}_+;\mathscr{R}^l)} = (E\int_0^\infty \|y(t)\|^2 dt)^{1/2} < \infty$).

2. Definitions and Preliminaries

Consider the following time-varying nonlinear stochastic Markov jump systems with (x, u, v)-dependent noise:

$$\begin{aligned} dx(t) &= \left[f(t,x,r_t) + g(t,x,r_t)u(t) + h(t,x,r_t)v(t)\right]dt \\ &\quad + \left[l(t,x,r_t) + q(t,x,r_t)u(t) \right. \\ &\quad \left. + s(t,x,r_t)v(t)\right]dw(t), \\ z(t) &= \operatorname{col}(m(t,x,r_t), u(t)) := \begin{bmatrix} m(t,x,r_t) \\ u(t) \end{bmatrix}, \\ x(0) &= x_0 \in \mathscr{R}^n, \end{aligned} \tag{1}$$

where $x(t) \in \mathscr{R}^n$, $u(t) \in \mathscr{R}^{n_u}$, $v(t) \in \mathscr{R}^{n_v}$, and $z(t) \in \mathscr{R}^{n_z}$ represent the system state, control input, exogenous input, and regulated output, respectively. $w(t)$ is the one-dimensional standard Wiener process defined on the complete probability space $(\Omega, \mathscr{F}, P)$, with the natural filter $\mathscr{F}_t$ generated by $w(\cdot)$ and $r(\cdot)$ up to time t. The jumping process $\{r_t, t \geq 0\}$ is a continuous-time discrete-state Markov process taking values in a finite set $\mathscr{T} = \{1, \ldots, N\}$. The transition probabilities for the process r_t are defined as

$$P(r_{t+h} = j \mid r_t = k) = \begin{cases} \pi_{kl}h + o(h), & \text{if } k \neq j, \\ 1 + \pi_{kk}h + o(h), & \text{if } k = j, \end{cases} \tag{2}$$

where $h > 0$, $\lim_{h\to 0}(o(h)/h) = 0$, and $\pi_{kl} \geq 0$ ($k, j \in \mathscr{T}, k \neq j$) denotes the switching rate from mode k at time t to mode j at time $t + h$ and $\pi_{kk} = -\sum_{j=1, j\neq k} \pi_{kj}$ for all $k \in \mathscr{T}$. In this paper, the processes r_t and $w(t)$ are assumed to be independent. For every $r_t = k \in \mathscr{T}$, $f(t,x,k)$, $g(t,x,k)$, $h(t,x,k)$, $l(t,x,k)$, $q(t,x,k)$, $s(t,x,k)$, and $m(t,x,k)$ are Borel measurable functions of suitable dimensions, which guarantee that system (1) has a unique strong solution [23].

The finite horizon H_∞ control for system (1) is defined as follows.

Definition 1. For given $\gamma > 0$, the state feedback control $u_T^*(t) \in L^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_u})$ is called a finite horizon H_∞ control of system (1), if for the zero initial state and any nonzero $v(t) \in L^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_v})$, we have $\|\mathscr{L}_T\| < \gamma$ with

$$\begin{aligned} \|\mathscr{L}_T\| &= \sup_{\substack{v(t)\in\mathscr{L}^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_v}),\\ v(t)\neq 0, u(t)=u_T^*(t), x_0=0}} \frac{\|z(t)\|_{L^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_z})}}{\|v(t)\|_{L^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_v})}} \\ &= \sup_{\substack{v(t)\in\mathscr{L}^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_v}),\\ v(t)\neq 0, u(t)=u_T^*(t), x_0=0}} E\left\{\int_0^T \left(\|m(t,x,r_t)\|^2 \right.\right. \\ &\qquad \left.\left. + \|u_T^*(t)\|^2\right) dt \mid r_0 = k\right\}^{1/2} \\ &\qquad \times \left(E\left\{\int_0^T \|v(t)\|^2 dt \mid r_0 \right.\right. \\ &\qquad \left.\left. = k\right\}^{1/2}\right)^{-1}, \\ &\qquad i \in \mathscr{T}, \end{aligned} \tag{3}$$

where $\mathscr{L}_T$ is an operator associated with system (1) which is defined as

$$\begin{aligned} &\mathscr{L}_T : \mathscr{L}^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_v}) \longmapsto \mathscr{L}^2_{\mathscr{F}}([0,T];\mathscr{R}^{n_z}), \\ &\mathscr{L}_T(v(t)) = z(t)|_{x_0=0}, \quad t \in [0,T]. \end{aligned} \tag{4}$$

Consider the time-invariant nonlinear stochastic Markov jump systems with (x, u, v)-dependent noise

$$\begin{aligned} dx(t) &= \left[f(x,r_t) + g(x,r_t)u(t) + h(x,r_t)v(t)\right]dt \\ &\quad + \left[l(x,r_t) + q(x,r_t)u(t) + s(x,r_t)v(t)\right]dw(t), \\ z(t) &= \operatorname{col}(m(x,r_t), u(t)) := \begin{bmatrix} m(x,r_t) \\ u(t) \end{bmatrix}, \\ x(0) &= x_0 \in \mathscr{R}^n. \end{aligned} \tag{5}$$

The infinite horizon H_∞ control for system (5) is defined as follows.

Definition 2. For given $\gamma > 0$, the control $u_\infty^*(t) \in L^2_{\mathscr{F}}(\mathscr{R}_+;\mathscr{R}^{n_u})$ is called an infinite horizon H_∞ control of system (1), if the following is considered.

(i) For the zero initial state and any nonzero $v(t) \in L^2_{\mathscr{F}}(\mathscr{R}_+; \mathscr{R}^{n_v})$, we have $\|\mathscr{L}_\infty\| < \gamma$ with

$$\begin{aligned}\|\mathscr{L}_\infty\| &= \sup_{\substack{v(t)\in\mathscr{L}^2_{\mathscr{F}}(\mathscr{R}_+;\,\mathscr{R}^{n_v}),\\ v(t)\neq 0, u(t)=u^*_\infty(t), x_0=0}} \frac{\|z(t)\|_{L^2_{\mathscr{F}}(\mathscr{R}_+;\,\mathscr{R}^{n_z})}}{\|v(t)\|_{L^2_{\mathscr{F}}(\mathscr{R}_+;\,\mathscr{R}^{n_v})}} \\ &= \sup_{\substack{v(t)\in\mathscr{L}^2_{\mathscr{F}}(\mathscr{R}_+;\mathscr{R}^{n_v}),\\ v(t)\neq 0, u(t)=u^*_\infty(t), x_0=0}} E\left\{\int_0^\infty \left(\|m(x,r_t)\|^2 + \|u^*_\infty(t)\|^2\right) dt \mid r_0 = k\right\}^{1/2} \\ &\quad \times \left(E\left\{\int_0^\infty \|v(t)\|^2 dt \mid r_0 = k\right\}^{1/2}\right)^{-1}, \quad i \in \mathscr{T},\end{aligned} \tag{6}$$

where $\mathscr{L}_\infty$ is an operator associated with system (5) which is defined as

$$\begin{gathered}\mathscr{L}_\infty : \mathscr{L}^2_{\mathscr{F}}(\mathscr{R}_+;\mathscr{R}^{n_v}) \longmapsto \mathscr{L}^2_{\mathscr{F}}(\mathscr{R}_+;\mathscr{R}^{n_z}), \\ \mathscr{L}_\infty(v(t)) = z(t)|_{x_0=0}, \quad t \in \mathscr{R}_+.\end{gathered} \tag{7}$$

(ii) System (5) is internally stable; that is, the following system

$$\begin{aligned}dx(t) &= [f(x,r_t) + g(x,r_t)u^*_\infty(x,r_t)]\,dt \\ &\quad + [l(x,r_t) + q(x,r_t)u^*_\infty(x,r_t)]\,dw(t),\end{aligned} \tag{8}$$

is globally asymptotically stable in probability [10].

To give our main results, we need the following lemmas.

Lemma 3 (see [10]). *(Generalized Itô formula). Let $\alpha(t,x,k)$ and $\beta(t,x,k)$ be given $\mathscr{R}^n$-valued, $\mathscr{F}_t$-adapted process, $k \in \mathscr{T}$, and $dx(t) = \alpha(t,x(t),r_t)dt + \beta(t,x(t),r_t)dw(t)$. Then for given $V(t,x,k) \in \mathscr{C}^{1,2}([0,T);\mathscr{R}^n)$, $k \in \mathscr{T}$, we have*

$$\begin{aligned}&E\{V(T,x(T),r_T) - V(s,x(s),r_s) \mid r_s = k\} \\ &\quad = E\left\{\int_s^T \mathscr{L}_T V(t,x(t),r_t)\,dt \mid r_s = k\right\},\end{aligned} \tag{9}$$

where

$$\begin{aligned}\mathscr{L}_T V(t,x,k) &= \frac{\partial V(t,x,k)}{\partial t} + \frac{\partial V'(t,x,k)}{\partial x}\alpha(t,x,k) \\ &\quad + \sum_{j=1}^N \pi_{kj} V(t,x,k) \\ &\quad + \frac{1}{2}\beta'(t,x,k)\frac{\partial V^2(t,x,k)}{\partial x^2}\beta(t,x,k).\end{aligned} \tag{10}$$

Lemma 4. *If $x, b \in \mathscr{R}^n$, $A \in \mathscr{R}^{n\times n}$ is a symmetric matrix and A^{-1} exists, we have*

$$x'Ax + x'b + b'x = \left(x + A^{-1}b\right)' A\left(x + A^{-1}b\right) - b'A^{-1}b. \tag{11}$$

Proof. This lemma is very easily proved by using completing squares technique, so the proof is omitted. □

3. Main Results

3.1. Finite Horizon Case. The following sufficient condition is presented for the finite horizon H_∞ control of system (1). For convenience, denote $(\cdot)_k := (\cdot)(t,x,k)$ in this subsection.

Theorem 5. *Assume that there exists a set of nonnegative functions $V(t,x,k) \in \mathscr{C}^{1,2}([0,T]\times\mathscr{R}^n\times\mathscr{T};\mathscr{R})$, $V(0,0,k) = 0$, and $\partial^2 V(t,x,k)/\partial x^2 \geq 0$ for all nonzero $x \in \mathscr{R}^n$, $k \in \mathscr{T}$. If $V(t,x,k)$ solves the following coupled HJIs:*

$$\begin{aligned}\Delta_k &= \frac{\partial V_k}{\partial t} + \frac{\partial V'_k}{\partial x} f_k + \frac{1}{2} l'_k \frac{\partial^2 V_k}{\partial x^2} l_k + m'_k m_k + \sum_{j=1}^N \pi_{kj} V_j \\ &\quad + \frac{1}{4}\left(l'_k \frac{\partial^2 V_k}{\partial x^2} s_k + \frac{\partial V'_k}{\partial x} h_k\right)\left(\gamma^2 I - s'_k \frac{\partial^2 V_k}{\partial x^2} s_k\right)^{-1} \\ &\quad \times \left(s'_k \frac{\partial^2 V_k}{\partial x^2} l_k + h'_k \frac{\partial V_k}{\partial x}\right) \\ &\quad - \frac{1}{4}\left(l'_k \frac{\partial^2 V_k}{\partial x^2} q_k + \frac{\partial V'_k}{\partial x} g_k\right)\left(I + q'_k \frac{\partial^2 V_k}{\partial x^2} q_k\right)^{-1} \\ &\quad \times \left(q'_k \frac{\partial^2 V_k}{\partial x^2} l_k + g'_k \frac{\partial V_k}{\partial x}\right) < 0, \\ &\quad \gamma^2 I - s'_k \frac{\partial^2 V_k}{\partial x^2} s_k > 0, \quad V(T,x,k) = 0,\end{aligned} \tag{12}$$

then

$$u^*_T(t,x,k) = -\frac{1}{2}\left(I + q'_k \frac{\partial^2 V_k}{\partial x^2} q_k\right)^{-1}\left(q'_k \frac{\partial^2 V_k}{\partial x^2} l_k + g'_k \frac{\partial V_k}{\partial x}\right) \tag{13}$$

is a finite horizon H_∞ control of system (1).

Proof. For any $T > 0$ and the initial state $x_0 = 0$, $r_0 = k$, applying Lemma 3, we have

$$\begin{aligned}&E[V(x(T),r_T) - V(0,r_0) \mid r_0 = k] \\ &\quad = E\left\{\int_0^T \left[\frac{\partial V_{r_t}}{\partial t} + \frac{\partial V'_{r_t}}{\partial x}\left(f_{r_t} + g_{r_t}u + h_{r_t}v\right)\right.\right. \\ &\qquad + \frac{1}{2}\left(l_{r_t} + q_{r_t}u + s_{r_t}v\right)'\frac{\partial^2 V_{r_t}}{\partial x^2}\left(l_{r_t} + q_{r_t}u + s_{r_t}v\right) \\ &\qquad + \left\|m_{r_t}\right\|^2 + \|u\|^2 - \gamma^2\|v\|^2 - \|z\|^2 \\ &\qquad \left.\left. + \gamma^2\|v\|^2 + \sum_{j=1}^N \pi_{kj} V_j\right] dt \mid r_0 = k\right\}\end{aligned}$$

$$= E\left\{\int_0^T \left[\frac{\partial V_{r_t}}{\partial t} - \|z\|^2 + \gamma^2\|v\|^2 + \sum_{j=1}^{N}\pi_{kj}V_j + \Theta_1(t,v,x,r_t) + \Theta_2(t,x,r_t) + \Theta_3(t,u,x,r_t) + \frac{1}{2}\left(u' q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}v + v' s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}u\right)\right]dt \mid r_0 = k\right\}, \tag{14}$$

where $(\cdot)_{r_t}$ denotes $(\cdot)(t,x,r_t)$ and

$$\begin{aligned}
\Theta_1(t,v,x,r_t) &= v'\left(-\gamma^2 I + \frac{1}{2}s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right)v + \frac{1}{2}\left(l'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t} + \frac{\partial V'_{r_t}}{\partial x}h_{r_t}\right)v + \frac{1}{2}v'\left(s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + h'_{r_t}\frac{\partial V_{r_t}}{\partial x}\right),\\
\Theta_2(t,x,r_t) &= \frac{\partial V_{r_t}}{\partial t} + \frac{\partial V'_{r_t}}{\partial x}f_{r_t} + \frac{1}{2}l'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + m'_{r_t}m_{r_t},\\
\Theta_3(t,u,x,r_t) &= u'\left(I + \frac{1}{2}q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}\right)u + \frac{1}{2}\left(l'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t} + \frac{\partial V'_{r_t}}{\partial x}g_{r_t}\right)u + \frac{1}{2}u'\left(q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + g'_{r_t}\frac{\partial V_{r_t}}{\partial x}\right).
\end{aligned} \tag{15}$$

Since $\partial^2 V_k/\partial x^2 \geq 0$, $k \in \mathcal{T}$, we have

$$\frac{1}{2}\left(-u' q'_{r_t} + v' s'_{r_t}\right)\frac{\partial^2 V_{r_t}}{\partial x^2}\left(-q_{r_t}u + s_{r_t}v\right) \geq 0, \tag{16}$$

which means that

$$\frac{1}{2}\left(u' q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}v + v' s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}u\right) \leq \frac{1}{2}u' q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}u + \frac{1}{2}v' s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}v. \tag{17}$$

Considering the above inequality and (14), we have

$$\begin{aligned}
&E\left[V(x(T), r_T) - V(0, r_0) \mid r_0 = k\right]\\
&\leq E\left\{\int_0^T \left[\Theta_1(t,v,x,r_t) + \Theta_2(t,x,r_t) + \Theta_3(t,u,x,r_t) + \sum_{j=1}^{N}\pi_{kj}V(x,j) - \|z\|^2 + \gamma^2\|v\|^2 + \frac{1}{2}\left(u' q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}v + v' s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}u\right)\right]dt \mid r_0 = k\right\}\\
&= E\left[\int_0^T \left(\overline{\Theta}_1(t,v,x,r_t) + \Theta_2(t,x,r_t) + \overline{\Theta}_3(t,u,x,r_t) + \sum_{j=1}^{N}\pi_{kj}V(x,j) - \|z\|^2 + \gamma^2\|v\|^2\right)dt \mid r_0 = k\right],
\end{aligned} \tag{18}$$

where

$$\begin{aligned}
\overline{\Theta}_1(v,x,r_t) &= v'\left(-\gamma^2 I + s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right)v + \frac{1}{2}\left(l'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t} + \frac{\partial V'_{r_t}}{\partial x}h_{r_t}\right)v + \frac{1}{2}v'\left(s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + h'_{r_t}\frac{\partial V_{r_t}}{\partial x}\right).\\
\overline{\Theta}_3(t,u,x,r_t) &= u'\left(I + q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}\right)u + \frac{1}{2}\left(l'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t} + \frac{\partial V'_{r_t}}{\partial x}g_{r_t}\right)u + \frac{1}{2}u'\left(q'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + g'_{r_t}\frac{\partial V_{r_t}}{\partial x}\right).
\end{aligned} \tag{19}$$

Applying Lemma 4 to $\overline{\Theta}_1(t,v,x,r_t)$ and $\overline{\Theta}_3(t,u,x,r_t)$, we have

$$\begin{aligned}
\overline{\Theta}_1(t,v,x,r_t) &= (v+\Gamma_1)'\left(-\gamma^2 I + s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right)(v+\Gamma_1)\\
&\quad - \frac{1}{4}\left(l'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t} + \frac{\partial V'_{r_t}}{\partial x}h_{r_t}\right) \times \left(-\gamma^2 I + s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right)^{-1} \times \left(s'_{r_t}\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + h'_{r_t}\frac{\partial V_{r_t}}{\partial x}\right),
\end{aligned}$$

$$\overline{\Theta}_3(t,u,x,r_t) = (u+\Gamma_2)'\left(I + q_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}\right)(u+\Gamma_2) - \frac{1}{4}\left(l_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t} + \frac{\partial V_{r_t}'}{\partial x}g_{r_t}\right) \times \left(I + q_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}\right)^{-1} \times \left(q_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + g_{r_t}'\frac{\partial V_{r_t}}{\partial x}\right), \tag{20}$$

where

$$\Gamma_1 = \frac{1}{2}\left(\gamma^2 I + s_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right)^{-1}\left(s_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + h_{r_t}'\frac{\partial V_{r_t}}{\partial x}\right),$$
$$\Gamma_2 = \frac{1}{2}\left(I + q_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}\right)^{-1}\left(q_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}l_{r_t} + g_{r_t}'\frac{\partial V_{r_t}}{\partial x}\right). \tag{21}$$

Substituting (20) into (18), and considering (12), it yields

$$E\left[V(x(T),r_T) - V(0,r_0) \mid r_0 = k\right] < E\left\{\int_0^T \left[(v+\Gamma_1)'\left(-\gamma^2 I + s_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right)(v+\Gamma_1) + (u+\Gamma_2)'\left(I + q_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}q_{r_t}\right)(u+\Gamma_2) + \gamma^2\|v\|^2 - \|z\|^2\right] dt \mid r_0 = k\right\}. \tag{22}$$

Taking $u = u_T^* = -\Gamma_2$ and considering the second item of (12), (22) leads to

$$E\left(\int_0^T \|z(t)\|^2 dt \mid r_0 = k\right) < \gamma^2 E\left(\int_0^T \|v(t)\|^2 dt \mid r_0 = k\right) - E\left[\int_0^T (v(t)+\Gamma_1)' \times \left(\gamma^2 I - s_{r_t}'\frac{\partial^2 V_{r_t}}{\partial x^2}s_{r_t}\right) \times (v(t)+\Gamma_1)\, dt \mid r_0 = k\right], \tag{23}$$

which means $\|\mathscr{L}_T\| < \gamma$ in Definition 1. The theorem is proved. □

Remark 6. The proof of Theorem 5 is based on an elementary identity (11), which avoids using stochastic dissipative theory as done in [4, 5]. We believe that the identity (11) will have many other applications in system analysis and synthesis.

3.2. Infinite Horizon Case. In contrast to the finite horizon case, the infinite horizon H_∞ control exhibits more complexity due to the additional requirement of stabilizing the closed-loop system internally. The following sufficient condition is derived for the infinite horizon H_∞ control of system (5). In this subsection, denote $(\cdot)_k := (\cdot)(x,k)$.

Theorem 7. *Assume that there exists a set of nonnegative functions $V(x,k) \in \mathscr{C}^2(\mathscr{R}^n \times \mathscr{T};\mathscr{R})$, $V(0,k) = 0$, and $\partial^2 V(x,k)/\partial x^2 \geq 0$ for all nonzero $x \in \mathscr{R}^n$, $k \in \mathscr{T}$. If $V(x,k)$ solves the following coupled HJIs:*

$$\Lambda_k = \frac{\partial V_k'}{\partial x}f_k + \frac{1}{2}l_k'\frac{\partial^2 V_k}{\partial x^2}l_k + m_k'm_k + \sum_{j=1}^N \pi_{kj}V_j + \frac{1}{4}\left(l_k'\frac{\partial^2 V_k}{\partial x^2}s_k + \frac{\partial V_k'}{\partial x}h_k\right)\left(\gamma^2 I - s_k'\frac{\partial^2 V_k}{\partial x^2}s_k\right)^{-1} \times \left(s_k'\frac{\partial^2 V_k}{\partial x^2}l_k + h_k'\frac{\partial V_k}{\partial x}\right) - \frac{1}{4}\left(l_k'\frac{\partial^2 V_k}{\partial x^2}q_k + \frac{\partial V_k'}{\partial x}g_k\right)\left(I + q_k'\frac{\partial^2 V_k}{\partial x^2}q_k\right)^{-1} \times \left(q_k'\frac{\partial^2 V_k}{\partial x^2}l_k + g_k'\frac{\partial V_k}{\partial x}\right) < 0,$$
$$\gamma^2 I - s_k'\frac{\partial^2 V_k}{\partial x^2}s_k > 0, \tag{24}$$

then

$$u_\infty^*(x,k) = -\frac{1}{2}\left(I + q_k'\frac{\partial^2 V_k}{\partial x^2}q_k\right)^{-1}\left(q_k'\frac{\partial^2 V_k}{\partial x^2}l_k + g_k'\frac{\partial V_k}{\partial x}\right) \tag{25}$$

is an infinite horizon H_∞ control of system (5).

Proof. Similar to the proof of Theorem 5, it is easy to show $\|\mathscr{L}_\infty\| < \gamma$ under condition (24). Next, we need to prove system (8) to be globally asymptotically stable in probability. Let $u_k^* = u_\infty^*(x,k)$ and let $\mathscr{L}_\infty$ be the infinitesimal generator of the system (8) which is similar to $\mathscr{L}_T$ in Lemma 3; then

$$\mathscr{L}_\infty V_k = \frac{\partial V_k'}{\partial x}(f_k + g_k u_k^*) + \sum_{j=1}^N \pi_{kj}V_j + \frac{1}{2}(l_k + q_k u_k^*)'\frac{\partial^2 V_k}{\partial x^2}(l_k + q_k u_k^*)$$
$$= \frac{\partial V_k'}{\partial x}f_k + \frac{\partial V_k'}{\partial x}g_k u_k^* + \sum_{j=1}^N \pi_{kj}V_j + \frac{1}{2}l_k'\frac{\partial^2 V_k}{\partial x^2}l_k + \frac{1}{2}l_k'\frac{\partial^2 V_k}{\partial x^2}q_k u_k^*$$

$$
\begin{aligned}
&+ \frac{1}{2}(q_k u_k^*)' \frac{\partial^2 V_k}{\partial x^2} l_k + \frac{1}{2}(q_k u_k^*)' \frac{\partial^2 V_k}{\partial x^2} (q_k u_k^*) \\
&= \frac{\partial V_k'}{\partial x} f_k + \frac{1}{2} l_k' \frac{\partial^2 V_k}{\partial x^2} l_k + \sum_{j=1}^{N} \pi_{kj} V_j + \Upsilon_{1k} + \Upsilon_{2k},
\end{aligned}
\tag{26}
$$

where

$$
\begin{aligned}
\Upsilon_{1k} &= \frac{\partial V_k'}{\partial x} g_k u_k^* + \frac{1}{2} l_k' \frac{\partial^2 V_k}{\partial x^2} q_k u_k^* + \frac{1}{2}(q_k u_k^*)' \frac{\partial^2 V_k}{\partial x^2} l_k \\
&= -\frac{1}{2} \frac{\partial V_k'}{\partial x} g_k \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \\
&\quad \times \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right) - \frac{1}{4} l_k' \frac{\partial^2 V_k}{\partial x^2} q_k \\
&\quad \times \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right) \\
&\quad - \frac{1}{4} \left(\frac{\partial V_k'}{\partial x} g_k + l_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right) \\
&\quad \times \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} q_k' \frac{\partial^2 V_k}{\partial x^2} l_k, \\
\Upsilon_{2k} &= \frac{1}{2}(q_k u_k^*)' \frac{\partial^2 V_k}{\partial x^2} (q_k u_k^*) \\
&= \frac{1}{8} \left(\frac{\partial V_k'}{\partial x} g_k + l_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right) \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \\
&\quad \times q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \\
&\quad \times \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right).
\end{aligned}
\tag{27}
$$

It can be checked that

$$
\begin{aligned}
\Upsilon_{1k} &= -\frac{1}{2} \left(l_k' \frac{\partial^2 V_k}{\partial x^2} q_k + \frac{\partial V_k'}{\partial x} g_k \right) \\
&\quad \times \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right),
\end{aligned}
\tag{28}
$$

$$
\begin{aligned}
\Upsilon_{2k} &\leq \frac{1}{8} \left(l_k' \frac{\partial^2 V_k}{\partial x^2} q_k + \frac{\partial V_k'}{\partial x} g_k \right) \\
&\quad \times \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right).
\end{aligned}
\tag{29}
$$

The following inequality is used during the calculation of (29):

$$
\begin{aligned}
(I+X)^{-1} X (I+X)^{-1} &\leq (I+X)^{-1} (I+X) (I+X)^{-1} \\
&= (I+X)^{-1}, \quad X > 0.
\end{aligned}
\tag{30}
$$

Implementing (28) and (29) into (26) and considering (24), it yields

$$
\begin{aligned}
\mathscr{L}_\infty V_k &\leq \frac{\partial V_k'}{\partial x} f_k + \frac{1}{2} l_k' \frac{\partial^2 V_k}{\partial x^2} l_k + \sum_{j=1}^{N} \pi_{kj} V_j \\
&\quad - \frac{3}{8} \left(l_k' \frac{\partial^2 V_k}{\partial x^2} q_k + \frac{\partial V_k'}{\partial x} g_k \right) \\
&\quad \times \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right) \\
&\leq \frac{\partial V_k'}{\partial x} f_k + \frac{1}{2} l_k' \frac{\partial^2 V_k}{\partial x^2} l_k + \sum_{j=1}^{N} \pi_{kj} V_j \\
&\quad - \frac{1}{4} \left(l_k' \frac{\partial^2 V_k}{\partial x^2} q_k + \frac{\partial V_k'}{\partial x} g_k \right) \\
&\quad \times \left(I + q_k' \frac{\partial^2 V_k}{\partial x^2} q_k \right)^{-1} \left(q_k' \frac{\partial^2 V_k}{\partial x^2} l_k + g_k' \frac{\partial V_k}{\partial x} \right) \\
&< -m_k' m_k - \frac{1}{4} \left(l_k' \frac{\partial^2 V_k}{\partial x^2} s_k + \frac{\partial V_k'}{\partial x} h_k \right) \\
&\quad \times \left(\gamma^2 I - s_k' \frac{\partial^2 V_k}{\partial x^2} s_k \right)^{-1} \left(s_k' \frac{\partial^2 V_k}{\partial x^2} l_k + h_k' \frac{\partial V_k}{\partial x} \right) \\
&\leq 0
\end{aligned}
\tag{31}
$$

which implies that (8) is globally asymptotically stable in probability from [10]. This theorem is completed. □

Remark 8. The methods proposed in [19, 24] cannot be applied to study the H_∞ problem of NSMJSs with (x, u, v)-dependent noise, although they are suitable for NSMJSs with (x, v)-dependent noise. One reason for this is that u and v are no longer separable in the conditions, when they enter the diffusion term simultaneously. In the proofs of Theorems 5 and 7, we resort to Lemma 4 to solve this problem.

Remark 9. When there is no Markov jump parameters, system (1)/(5) will reduce to general nonlinear stochastic systems with (x, u, v)-dependent noise, which was studied in [20, 21]. However, all the conditions of H_∞ control design for nonlinear stochastic systems in [20, 21] were given in terms of three coupled HJEs, which were difficult to be solved. According to this paper, the sufficient condition for H_∞ control of nonlinear stochastic systems can be derived by means of a single set of coupled HJIs, which is easier to be verified than three coupled HJEs in [20, 21].

From Theorem 7, the following nonlinear stochastic bounded real lemma will be derived for Markov jump system:

$$
\begin{aligned}
dx(t) &= \left[f(x,r_t) + h(x,r_t)v\right]dt \\
&\quad + \left[l(x,r_t) + s(x,r_t)v\right]dw, \qquad (32)\\
z(t) &= m(x,r_t), \qquad x(0) = x_0 \in \mathscr{R}^n.
\end{aligned}
$$

Lemma 10. *For a prescribed $\gamma > 0$, system (32) is internally stable and $\|\mathscr{L}_\infty\|_{u(t)\equiv 0} < \gamma$, if there exists a set of nonnegative functions $V(x,k) \in \mathscr{C}^2(\mathscr{R}^n \times \mathscr{T};\mathscr{R})$, $V(0,k) = 0$, and $\partial^2 V(x,k)/\partial x^2 \geq 0$ for all nonzero $x \in \mathscr{R}^n$, $k \in \mathscr{T}$, satisfying the following coupled HJIs:*

$$
\begin{gathered}
\frac{\partial V_k'}{\partial x} f_k + \frac{1}{2} l_k' \frac{\partial^2 V_k}{\partial x^2} l_k + m_k' m_k + \sum_{j=1}^{N} \pi_{kj} V_j \\
+ \frac{1}{4}\left(l_k' \frac{\partial^2 V_k}{\partial x^2} s_k + \frac{\partial V_k'}{\partial x} h_k\right) \\
\times \left(\gamma^2 I - s_k' \frac{\partial^2 V_k}{\partial x^2} s_k\right)^{-1} \left(s_k' \frac{\partial^2 V_k}{\partial x^2} l_k + h_k' \frac{\partial V_k}{\partial x}\right) \\
< 0, \\
\gamma^2 I - s_k' \frac{\partial^2 V_k}{\partial x^2} s_k > 0, \quad k \in \mathscr{T}.
\end{gathered} \tag{33}
$$

Proof. Letting $g_k \equiv 0$, $q_k \equiv 0$, and $u \equiv 0$ in Theorem 7, we obtain (33) easily. □

Next, we present a sufficient condition for the following linear stochastic Markov jump systems with (x,u,v)-dependent noise:

$$
\begin{aligned}
dx(t) &= \left[A(r_t)x + B(r_t)u + C(r_t)v\right]dt \\
&\quad + \left[A_1(r_t)x + B_1(r_t)u + C_1(r_t)v\right]dw, \qquad (34)\\
z(t) &= \operatorname{col}\left(D(r_t)x, u\right), \qquad x(0) = x_0 \in \mathscr{R}^n.
\end{aligned}
$$

Corollary 11. *System (34) is internally stable and $\|\mathscr{L}_\infty\| < \gamma$ for given $\gamma > 0$, if there exist matrices $X_k > 0$, Y_k $k \in \mathscr{T}$ satisfying the following LMIs:*

$$
\begin{bmatrix}
\Sigma_k & C_k & X_k D_k' & X_k A_{1k}' + Y_k' B_{1k}' & 0 & \psi_k'(X) \\
* & -\gamma^2 I & 0 & C_{1k}' & C_{1k}' & 0 \\
* & * & -I & 0 & 0 & 0 \\
* & * & * & -X_k & 0 & 0 \\
* & * & * & * & -X_k & 0 \\
* & * & * & * & * & -\phi_k(X)
\end{bmatrix} < 0, \tag{35}
$$

where

$$
\begin{aligned}
\Sigma_k &= A_k X_k + B_k Y_k + X_k A_k' + Y_k' B_k' + \pi_{kk} X_k, \\
\psi_k(X) &= \left[\sqrt{\pi_{k1}} X_k, \dots, \sqrt{\pi_{kk-1}} X_k, \right. \\
&\qquad \left. \sqrt{\pi_{kk+1}} X_k, \dots, \sqrt{\pi_{kN}} X_k\right]', \\
\phi_k(X) &= \operatorname{diag}\left\{X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_N\right\}.
\end{aligned} \tag{36}
$$

Moreover, the state feedback gain matrices are given by $K_k = Y_k X_k^{-1}$.

Proof. Firstly, consider the following unforced linear stochastic Markov jump systems:

$$
\begin{gathered}
dx(t) = \left[A(r_t)x + C(r_t)v\right]dt + \left[A_1(r_t)x + C_1(r_t)v\right]dw, \\
z(t) = D(r_t)x, \qquad x(0) = x_0 \in \mathscr{R}^n.
\end{gathered} \tag{37}
$$

Let $V_k = x' P_k x$, $P_k > 0$, $k \in \mathscr{T}$; (33) in Lemma 10 can be written as follows:

$$
\begin{aligned}
&2x' P_k A_k x + x' A_{1k}' P_k A_{1k} x + x' D_k' D_k x + x' \sum_{j=1}^{N} \pi_{kj} P_j x \\
&\quad + \left(x' A_{1k}' P_k C_{1k} + x' P_k C_k\right) \\
&\quad \times \left(\gamma^2 I - 2C_{1k}' P_k C_{1k}\right)^{-1} \left(C_{1k}' P_k A_{1k} x + C_k' P_k x\right) \\
&= x' \left[P_k A_k + A_k' P_k + A_{1k}' P_k A_{1k} + D_k' D_k + \sum_{j=1}^{N} \pi_{kj} P_j \right. \\
&\quad + \left(A_{1k}' P_k C_{1k} + P_k C_k\right)\left(\gamma^2 I - 2C_{1k}' P_k C_{1k}\right)^{-1} \\
&\quad \left. \times \left(C_{1k}' P_k A_{1k} + C_k' P_k\right)\right] x < 0,
\end{aligned} \tag{38}
$$

$$
\gamma^2 I - 2C_{1k}' P_k C_{1k} > 0. \tag{39}
$$

By Schur complement, (38) is equivalent to

$$
\begin{bmatrix}
\left\{ P_k A_k + A_k' P_k + A_{1k}' P_k A_{1k} + \sum_{j=1}^{N} \pi_{kj} P_j \right\} & A_{1k}' P_k C_{1k} + P_k C_k & D_k' \\
* & -\gamma^2 I + 2C_{1k}' P_k C_{1k} & 0 \\
* & * & -I
\end{bmatrix} < 0, \tag{40}
$$

which also implies (39). Moreover, (40) can be rewritten as

$$
\begin{bmatrix}
P_k A_k + A_k' P_k + \sum_{j=1}^{N} \pi_{kj} P_j & P_k C_k & D_k' \\
* & -\gamma^2 I & 0 \\
* & * & -I
\end{bmatrix}
+ \begin{bmatrix} A_{1k}' \\ C_{1k}' \\ 0 \end{bmatrix} P_k \begin{bmatrix} A_{1k} & C_{1k} & 0 \end{bmatrix}
+ \begin{bmatrix} 0 \\ C_{1k}' \\ 0 \end{bmatrix} P_k \begin{bmatrix} 0 & C_{1k} & 0 \end{bmatrix} < 0,
\tag{41}
$$

which yields

$$\begin{bmatrix} P_kA_k + A'_kP_k + \sum_{j=1}^{N}\pi_{kj}P_j & P_kC_k & D'_k & A'_{1k} & 0 \\ * & -\gamma^2 I & 0 & C'_{1k} & C'_{1k} \\ * & * & -I & 0 & 0 \\ * & * & * & -P_k^{-1} & 0 \\ * & * & * & * & -P_k^{-1} \end{bmatrix} < 0, \tag{42}$$

according to Schur complement. Pre- and postmultiplying (42) by $\operatorname{diag}\{P_k^{-1}\ I\ I\ I\ I\}$ and denoting $X_k = P_k^{-1}$, we have

$$\begin{bmatrix} A_kX_k + X_kA'_k + \pi_{kk}X_k & C_k & X_kD'_k & X_kA'_{1k} & 0 & \psi'_k(X) \\ * & -\gamma^2 I & 0 & C'_{1k} & C'_{1k} & 0 \\ * & * & -I & 0 & 0 & 0 \\ * & * & * & -X_k & 0 & 0 \\ * & * & * & * & -X_k & 0 \\ * & * & * & * & * & -\phi_k(X) \end{bmatrix} < 0, \tag{43}$$

where $\psi_k(X)$ and $\phi_k(X)$ are defined as in (35).

Considering closed-loop system (34) with state feedback control $u(t) = K_kx(t)$, $k \in \mathcal{T}$. Replacing A_k, A_{1k} by $A_k + B_kK_k$, $A_{1k} + B_{1k}K_k$, respectively, and setting $Y_k = K_kX_k$ in (43) yield (35). Therefore, the state feedback gain matrices can be obtained by $K_k = Y_kX_k^{-1}$. □

Remark 12. Although HJIs (12) in Theorem 5 or (24) in Theorem 7 can be solved by trial and error in some simple cases (see Example 13 in Section 4), they are difficult to be dealt with for high-dimensional systems. In order to avoid solving the HJIs, the Taylor series approach [25] or fuzzy approach based on Takagi-Sugeno model [24, 26] can be considered to design the nonlinear stochastic H_∞ controller.

4. Numerical Example

In this section, two numerical examples are provided to illustrate the effectiveness of the developed results.

Example 13. Consider one-dimensional two-mode time-invariant nonlinear stochastic Markov jump systems with generator matrix $\Pi = \left[\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, and the two subsystems are as follows:

$$\begin{aligned} &\mathrm{I}: \begin{cases} dx = \left(x^3 + 4xu + xv\right)dt + \left(x^2 + u + v\right)dw, \\ z = \operatorname{col}\left(x^2, u\right), \quad x(0) = x_0 \in \mathscr{R}^n, \end{cases} \\ &\mathrm{II}: \begin{cases} dx = \left(\frac{1}{2}x^3 + 3xu + \frac{1}{2}xv\right)dt + \left(x^2 + u + v\right)dw, \\ z = \operatorname{col}\left(\frac{1}{2}x^2, u\right), \quad x(0) = x_0 \in \mathscr{R}^n. \end{cases} \end{aligned} \tag{44}$$

Set $V(x) = p_ix^2$, $i = 1, 2$, with $p_1 > 0$ and $p_2 > 0$ to be determined; then HJIs (24) become

$$\begin{aligned} &2p_1x\cdot x^3 + \frac{1}{2}x^2\cdot 2p_1\cdot x^2 + x^2\cdot x^2 \\ &\quad + \frac{1}{4}\left(x^2\cdot 2p_1 + 2p_1x\cdot x\right)^2\left(\gamma^2 - 2p_1\right)^{-1} \\ &\quad - \frac{1}{4}\left(x^2\cdot 2p_1 + 2p_1x\cdot 4x\right)^2\left(1 + 2p_1\right)^{-1} \\ &\quad - p_1x^2 + p_2x^2 < 0, \\ &\gamma^2 - 2p_1 > 0, \\ &2p_2x\cdot\frac{1}{2}x^3 + \frac{1}{2}x^2\cdot 2p_2\cdot x^2 + \frac{1}{2}x^2\cdot\frac{1}{2}x^2 \\ &\quad + \frac{1}{4}\left(x^2\cdot 2p_2 + 2p_2x\cdot\frac{1}{2}x\right)^2\left(\gamma^2 - 2p_2\right)^{-1} \\ &\quad - \frac{1}{4}\left(x^2\cdot 2p_2 + 2p_2x\cdot 3x\right)^2\left(1 + 2p_2\right)^{-1} \\ &\quad + p_1x^2 - p_2x^2 < 0, \\ &\gamma^2 - 2p_2 > 0. \end{aligned} \tag{45}$$

For given $\gamma = \sqrt{3}$, the above inequalities have solutions $p_1 = 1$ and $p_2 = 1$. According to Theorem 7, the infinite H_∞ controllers of system (44) are $u_\infty^*(x,1) = -(5/3)x^2$ and $u_\infty^*(x,2) = -(4/3)x^2$.

Example 14. Consider two-dimensional two-mode linear stochastic Markov jump systems (34) with the following parameters:

$$\begin{aligned} &A_1 = \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix}, \qquad B_1 = C_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ &A_{11} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B_{11} = C_{11} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ &A_2 = \begin{bmatrix} 2 & 1 \\ 0 & -2 \end{bmatrix}, \qquad B_2 = C_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ &A_{12} = \begin{bmatrix} 2 & 0 \\ 1 & 11 \end{bmatrix}, \qquad B_{12} = C_{12} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ &D_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad D_2 = \begin{bmatrix} 1 & 1 \end{bmatrix}, \\ &\Pi = \begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix}. \end{aligned} \tag{46}$$

With the choice of $\gamma = \sqrt{2}$, a possible solution of LMIs (35) in Corollary 11 can be found by using Matlab LMI control toolbox:

$$\begin{aligned} &X_1 = \begin{bmatrix} 0.8879 & -0.4074 \\ -0.4074 & 3.6974 \end{bmatrix}, \qquad X_2 = \begin{bmatrix} 0.5026 & 0.0464 \\ 0.0464 & 5.4520 \end{bmatrix}, \\ &Y_1 = \begin{bmatrix} -3.1690 & -5.3082 \end{bmatrix}, \qquad Y_2 = \begin{bmatrix} -4.5025 & -8.5671 \end{bmatrix}. \end{aligned} \tag{47}$$

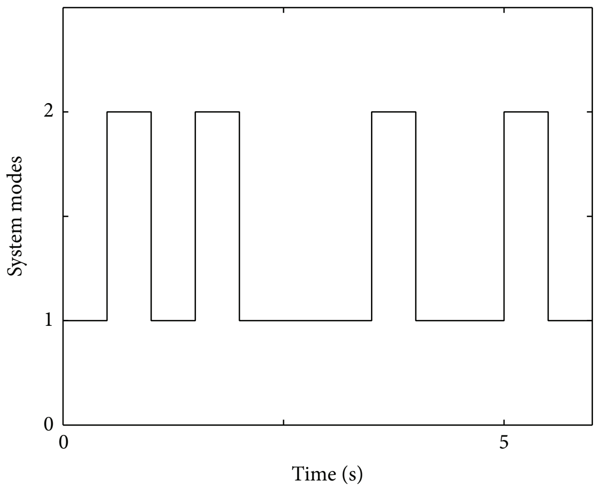


Figure 1: Result of the changing between modes.

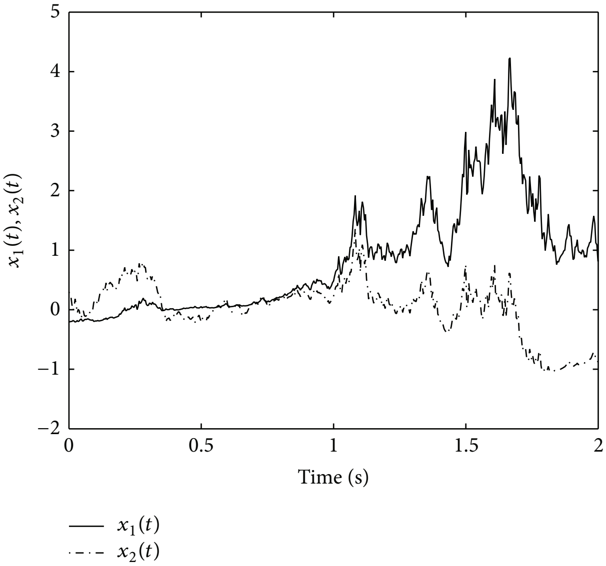


Figure 2: The state responses of unforced system (37).

Then, the H_∞ control gain matrices of system (34) are as follows:

$$K_1 = Y_1 X_1^{-1} = [-4.4531 \quad -1.9263], \qquad (48)$$
$$K_2 = Y_2 X_2^{-1} = [-8.8209 \quad -1.4963].$$

Figure 1 shows the result of the changing between modes during the simulation with the initial mode at mode 1. The initial condition is chosen as $x_0 = [-0.2 \quad 0.3]'$ and exogenous input $v(t) = e^{-t}$. By means of Euler-Maruyama method [27], the state responses of unforced system (37) and controlled system (34) are shown in Figures 2 and 3, respectively. From Figure 3, one can find that the controlled system (34) can achieve stability and attenuation performance in the sense of mean square by using the proposed H_∞ controller.

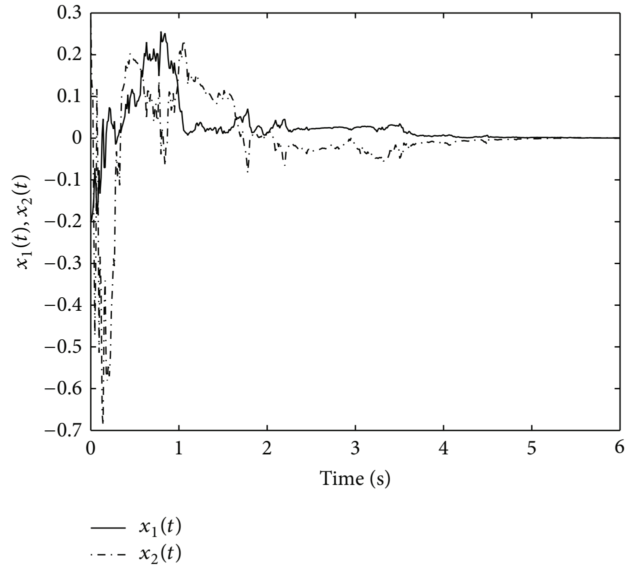


Figure 3: The state responses of controlled system (34).

5. Conclusions

In this paper, we have studied the H_∞ control problem for NSMJSs with (x, u, v)-dependent noise. A sufficient condition for finite/infinite horizon H_∞ control has been derived in terms of a set of coupled HJIs. It can be found that Lemma 4 plays an essential role during the proof of Theorems 5 and 7. The validity of the obtained results has been verified by two examples.

Conflict of Interests

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

Acknowledgments

This work is supported by National Natural Science Foundation of China (nos. 61203053 and 61174078), China Postdoctoral Science Foundation (no. 2013M531635), Research Fund for the Doctoral Program of Higher Education of China (no. 20120133120014), Special Funds for Postdoctoral Innovative Projects of Shandong Province (no. 201203096), Fundamental Research Fund for the Central Universities (nos. 11CX04042A, 12CX02010A, and 14CX02093A), Research Fund for the Taishan Scholar Project of Shandong Province of China, and SDUST Research Fund (no. 2011KYTD105).

References

[1] G. Zames, "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 301–320, 1981.

[2] I. R. Petersen, V. A. Ugrinovskii, and A. V. Savkin, *Robust Control Design Using* H_∞*-Methods*, Springer, Berlin, Germany, 2000.

[3] D. Hinrichsen and A. J. Pritchard, "Stochastic H_∞," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1504–1538, 1998.

[4] W. Zhang and B.-S. Chen, "State feedback H_∞ control for a class of nonlinear stochastic systems," *SIAM Journal on Control and Optimization*, vol. 44, no. 6, pp. 1973–1991, 2006.

[5] N. Berman and U. Shaked, "H_∞-like control for nonlinear stochastic systems," *Systems & Control Letters*, vol. 55, no. 3, pp. 247–257, 2006.

[6] W. Zhang, B.-S. Chen, and C.-S. Tseng, "Robust H_∞ filtering for nonlinear stochastic systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 589–598, 2005.

[7] B. Shen, Z. Wang, H. Shu, and G. Wei, "On nonlinear H_∞ filtering for discrete-time stochastic systems with missing measurements," *IEEE Transactions on Automatic Control*, vol. 53, no. 9, pp. 2170–2180, 2008.

[8] Z. Wang, B. Shen, H. Shu, and G. Wei, "Quantized H_∞ control for nonlinear stochastic time-delay systems with missing measurements," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1431–1444, 2012.

[9] B. Shen, S. X. Ding, and Z. Wang, "Finite-horizon H_∞ fault estimation for linear discrete time-varying systems with delayed measurements," *Automatica*, vol. 49, no. 1, pp. 293–296, 2013.

[10] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, UK, 2006.

[11] V. Dragan, T. Morozan, and A.-M. Stoica, *Mathematical Methods in Robust Control of Discrete-Time Linear Stochastic Systems*, Springer, New York, NY, USA, 2010.

[12] Z.-Y. Li, B. Zhou, Y. Wang, and G.-R. Duan, "On eigenvalue sets and convergence rate of Itô stochastic systems with Markovian switching," *IEEE Transactions on Automatic Control*, vol. 56, no. 5, pp. 1118–1124, 2011.

[13] L. Sheng, M. Gao, and W. Zhang, "Spectral characterisation for stability and stabilisation of linear stochastic systems with Markovian switching and its applications," *IET Control Theory & Applications*, vol. 7, no. 5, pp. 730–737, 2013.

[14] L. Shen, J. Sun, and Q. Wu, "Observability and detectability of discrete-time stochastic systems with Markovian jump," *Systems & Control Letters*, vol. 62, no. 1, pp. 37–42, 2013.

[15] O. L. V. Costa and A. de Oliveira, "Optimal mean-variance control for discrete-time linear systems with Markovian jumps and multiplicative noises," *Automatica*, vol. 48, no. 2, pp. 304–315, 2012.

[16] L. Sheng, W. Zhang, and M. Gao, "Relationship between Nash equilibrium strategies and H_2/H_∞ control of stochastic Markov jump systems with multiplicative noise," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, 2014.

[17] P. V. Pakshin, "Exponential dissipativeness of the random-structure diffusion processes and problems of robust stabilization," *Automation and Remote Control*, vol. 68, no. 10, pp. 1852–1870, 2007.

[18] L. Sheng and M. Gao, "Stabilization for Markovian jump nonlinear systems with partly unknown transition probabilities via fuzzy control," *Fuzzy Sets and Systems*, vol. 161, no. 21, pp. 2780–2792, 2010.

[19] Z. Lin, Y. Lin, and W. Zhang, "A unified design for state and output feedback H_∞ control of nonlinear stochastic Markovian jump systems with state and disturbance-dependent noise," *Automatica*, vol. 45, no. 12, pp. 2955–2962, 2009.

[20] W. Zhang, H. Zhang, and B.-S. Chen, "Stochastic H_2/H_∞ control with (x, u, v)-dependent noise: finite horizon case," *Automatica*, vol. 42, no. 11, pp. 1891–1898, 2006.

[21] W. Zhang and G. Feng, "Nonlinear stochastic H_2/H_∞ control with (x, u, v)-dependent noise: infinite horizon case," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1323–1328, 2008.

[22] W. Zhang, B.-S. Chen, H. Tang, L. Sheng, and M. Gao, "Some remarks on general nonlinear stochastic H_∞ control with state, control, and disturbance-dependent noise," *IEEE Transactions on Automatic Control*, vol. 59, no. 1, pp. 237–242, 2014.

[23] V. Dragan, T. Morozan, and A.-M. Stoica, *Mathematical Methods in Robust Control of Linear Stochastic Systems*, Springer, New York, NY, USA, 2006.

[24] Z. Lin, Y. Lin, and W. Zhang, "H_∞ stabilisation of non-linear stochastic active fault-tolerant control systems: fuzzy-interpolation approach," *IET Control Theory & Applications*, vol. 4, no. 10, pp. 2003–2017, 2010.

[25] M. Abu-Khalaf, J. Huang, and F. L. Lewis, *Nonlinear* $\mathrm{H}_2/\mathrm{H}_\infty$ *Constrained Feedback Control*, Springer, London, UK, 2006.

[26] L. Sheng, M. Gao, and W. Zhang, "Dissipative control for Markov jump non-linear stochastic systems based on T-S fuzzy model," *International Journal of Systems Science*, vol. 45, no. 5, pp. 1213–1224, 2014.

[27] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM Review*, vol. 43, no. 3, pp. 525–546, 2001.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 958474, 13 pages
http://dx.doi.org/10.1155/2014/958474

Research Article

Covariance-Based Estimation from Multisensor Delayed Measurements with Random Parameter Matrices and Correlated Noises

R. Caballero-Águila,[1] A. Hermoso-Carazo,[2] and J. Linares-Pérez[2]

[1] *Departamento de Estadística, Universidad de Jaén, Paraje Las Lagunillas, 23071 Jaén, Spain*
[2] *Departamento de Estadística, Universidad de Granada, Avenida, Fuentenueva, 18071 Granada, Spain*

Correspondence should be addressed to J. Linares-Pérez; jlinares@ugr.es

Received 18 March 2014; Revised 30 April 2014; Accepted 19 May 2014; Published 16 June 2014

Academic Editor: Suiyang Khoo

The optimal least-squares linear estimation problem is addressed for a class of discrete-time multisensor linear stochastic systems subject to randomly delayed measurements with different delay rates. For each sensor, a different binary sequence is used to model the delay process. The measured outputs are perturbed by both random parameter matrices and one-step autocorrelated and cross correlated noises. Using an innovation approach, computationally simple recursive algorithms are obtained for the prediction, filtering, and smoothing problems, without requiring full knowledge of the state-space model generating the signal process, but only the information provided by the delay probabilities and the mean and covariance functions of the processes (signal, random parameter matrices, and noises) involved in the observation model. The accuracy of the estimators is measured by their error covariance matrices, which allow us to analyze the estimator performance in a numerical simulation example that illustrates the feasibility of the proposed algorithms.

1. Introduction

In the past decades, the development of network technologies has promoted the study of the estimation problem in multisensor systems, where the observations provided by all the sensor networks are transmitted to a fusion center for being processed, thus obtaining the whole available information on the signal. This kind of systems with multiple sensors is becoming an interesting research topic due to its broad scope of application as they can provide more information than traditional communication systems with a single sensor. This form of transmission has several advantages, such as low cost or simple installation and maintenance; however, due to the imperfection of the communication channels, during the transmission process, there exist often random sensor delays and/or multiple packet dropouts. Standard observation models are not appropriate under these random uncertainties, and classical estimation algorithms, where the measurements generated by the system are available in real time, cannot be applied directly. Therefore, new algorithms are needed and, recently, the estimation problem in multisensor systems with some of the aforementioned random uncertainties has become a research topic of growing interest (see, e.g., [1–6] and references therein).

There are many current applications, for example, networked multiple sensor systems with measurement-based output feedback, where the measurements may be randomly delayed due to network congestion or random failures in the transmission mechanism. Several modifications of the standard estimation algorithms have been proposed to incorporate the effects of randomly delayed measurements, in both linear and nonlinear systems. Assuming full knowledge of the state-space model of the signal process to be estimated we can mention [7–10] and using covariance information, [11, 12], among others. Although all papers above mentioned involve systems with randomly delayed sensors, their major handicap is that all the sensors are assumed to have the same delay characteristics. Nevertheless, such an assumption is not realistic in many practical situations, where the information is gathered by an array of heterogeneous sensors, and the delay probability at each individual sensor can be different from the others. In recent years, this approach has been generalized considering multiple delayed sensors with different delay characteristics (see, e.g., [13, 14], using the state-space model, and [15, 16], using covariance information).

Furthermore, in many sensor network applications the measured outputs present uncertainties which cannot be described only by the usual additive disturbances, and multiplicative noises must be included in the observation equations to model such uncertainties (see, e.g., [17, 18]). Also, in the context of missing and fading measurements, the observation equations include multiplicative noises described by scalar random variables with arbitrary discrete probability distribution over the interval $[0, 1]$ (see, e.g., [19–21]). The above systems are a special case of systems with random parameter matrices, which have important practical significance and arise in areas such as digital control of chemical processes, systems with human operators, economic systems, and stochastically sampled digital control systems [22].

In [22, 23], the optimal linear filtering problem in systems with independent random state transition and measurement matrices is addressed by transforming the original system into one with deterministic parameter matrices and state-dependent process and measurement noises, to which the Kalman filter is applied. Although in [22] the Kalman filter is applied without providing any theoretical justification, in [23] it is shown that, under mild conditions, the transformed system satisfies the Kalman filter requirements and, hence, optimal linear estimators are obtained for systems with independent random parameter matrices. In [24], systems with deterministic transition matrices and one-step correlated measurement matrices are considered, and the optimal recursive state estimation problem is addressed by converting the observation equation into one with deterministic measurement matrices and applying the optimal Kalman filter for the case of one-step correlated measurement noise. In the above-mentioned papers, although the noises of the transformed system with deterministic matrices depend on the system state and therefore can be correlated, the original system noises are assumed to be independent white processes. This assumption can be restrictive in many real world problems in which correlation and cross-correlation of the noises may be present. For this reason, the estimation problem in systems with correlated and cross-correlated noises is becoming an active research topic (see [25–29] for systems with deterministic matrices and [30, 31] for systems with random parameter matrices, among others). In [30] a locally optimal filter in the class of Kalman-type recursive filters is presented and, in [31], the optimal least-squares linear filter is derived.

Motivated by the above analysis, in this paper we address the signal estimation problem from measurements coming from multiple sensors which are randomly delayed by one sampling time with different delay characteristics, under the assumption that the measured outputs are perturbed by both random parameter matrices and one-step autocorrelated and cross-correlated observation noises. The main contributions of this paper can be highlighted as follows: (1) the observation model considers simultaneously random delayed measurements and both random parameter matrices and correlated noises (one-step autocorrelation and also one-step cross-correlations between different sensor noises are considered) in the measured outputs; (2) optimal LS linear recursive filtering and smoothing algorithms are obtained without requiring signal augmentation approach thus avoiding the expensive computational cost; (3) the proposed algorithms are obtained without requiring full knowledge of the state-space model generating the signal process; and (4) the innovation technique is used, simplifying substantially the derivation of the algorithms since the innovation process is a white noise.

The rest of the paper is organized as follows. In Section 2, we present the delayed measurement model to be considered and the assumptions and properties under which the LS linear estimation problem is addressed. The innovation approach which, as mentioned above, yields straightforward derivation of the estimation algorithms is given in Section 3. The recursive filtering and smoothing algorithms are derived in Sections 4 and 5, respectively. In Section 6, the performance of the proposed filtering algorithms is illustrated by a numerical simulation example where the signal of a first-order autoregressive model is estimated from delayed observations coming from two sensors with different delay characteristics, considering two kinds of measured outputs with correlated noises. The paper concludes with some final comments in Section 7.

Notation. The notation used throughout the paper is standard. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space and $\mathbb{R}^{m\times n}$ is the set of all $m \times n$ real matrices. A^T and A^{-1} denote the transpose and inverse of a matrix A, respectively. The shorthand $\mathrm{Diag}(a_1, \ldots, a_m)$ denotes a diagonal matrix whose diagonal entries are $a_1, \ldots, a_m$. $\mathbf{1} = (1, \ldots, 1)^T$ denotes the all-ones vector and I the identity matrix. If the dimensions of matrices are not explicitly stated, they are assumed to be compatible with algebraic operations. The notation $\circ$ denotes the Hadamard product ($[C \circ D]_{ij} = C_{ij}D_{ij}$). δ_{k-s} represents the Kronecker delta function, which is equal to one if $k = s$ and zero otherwise. Moreover, for arbitrary random vectors, α and β, we will denote $\mathrm{Cov}[\alpha, \beta] = E[(\alpha - E[\alpha])(\beta - E[\beta])^T]$, where $E[\cdot]$ stands for the mathematical expectation operator.

2. Problem Formulation

The aim of this paper is to find recursive algorithms for the optimal least-squares (LS) linear filtering and smoothing problems of an n-dimensional discrete-time random signal z_k using measurements perturbed by random observation matrices and correlated additive noises, which are transmitted by multiple sensors where one-step random delays with different rates may occur during the transmission process.

The estimation problem is addressed under the assumption that the evolution model of the signal to be estimated is unknown and only information about its mean and covariance functions is available; this information is specified in the following assumption.

Assumption 1. The n-dimensional signal process $\{z_k;\ k \geq 1\}$ has zero mean and its autocovariance function is expressed in a separable form, $E[z_k z_j^T] = A_k B_j^T$, $j \leq k$, where A and B are known $n \times M$ matrix functions.

Remark 2. Although Assumption 1 might seem restrictive, it covers many practical situations; for example, when the system matrix Φ in the state-space model of a stationary signal is available, the signal autocovariance function is $E[z_k z_j^T] = \Phi^{k-j} E[z_j z_j^T]$, $j \leq k$, and Assumption 1 is clearly satisfied, taking $A_k = \Phi^k$ and $B_j = E[z_j z_j^T](\Phi^{-j})^T$. Also, processes with finite-dimensional, possibly time-variant, state-space models have semiseparable covariance functions, $E[z_k z_j^T] = \sum_{i=1}^r a_k^i b_j^{iT}$, $j \leq k$ (see [32]), and this structure is a particular case of that assumed, just taking $A_k = (a_k^1, a_k^2, \ldots, a_k^r)$ and $B_j = (b_j^1, b_j^2, \ldots, b_j^r)$. Consequently, the structural assumption on the signal autocovariance function covers both stationary and nonstationary signals.

Next, the observation model with one-step random delays is described and the assumptions under which the LS linear estimation problem will be addressed are presented.

2.1. Delayed Observation Model. Let $\{z_k;\ k \geq 1\}$ be the signal process satisfying Assumption 1 and consider m sensors which provide scalar measurements of the signal according to the following model:

$$\widetilde{y}_k^i = H_k^i z_k + \widetilde{v}_k^i, \quad k \geq 1,\ i = 1, 2, \ldots, m, \tag{1}$$

where $\widetilde{y}_k^i \in \mathbb{R}$ is the measurement provided by the ith sensor at time k (actual output); $\{H_k^i;\ k \geq 1\}$ are $1 \times n$ random parameter matrices; $\{\widetilde{v}_k^i;\ k \geq 1\}$ are measurement noises. The following assumptions are established on this model.

Assumption 3. For $i = 1, 2, \ldots, m$, $\{H_k^i;\ k \geq 1\}$ are $1 \times n$ random parameter matrices with known means, $E[H_k^i] = \overline{H}_k^i$; H_k^i and H_s^j are independent for $k \neq s$; the covariances and cross-covariances at the same time, $\mathrm{Cov}[h_{i,p}^k, h_{j,q}^k]$, are also known ($h_{i,p}^k$ denotes the pth entry of H_k^i, for $p = 1, 2, \ldots, n$).

Assumption 4. The additive measurement noises $\{\widetilde{v}_k^i;\ k \geq 1\}$, $i = 1, 2, \ldots, m$, are zero-mean processes with $\mathrm{Cov}[\widetilde{v}_k^i, \widetilde{v}_s^j] = \widetilde{R}_{k,k}^{ij} \delta_{k-s} + \widetilde{R}_{k,s}^{ij} \delta_{k-s+1} + \widetilde{R}_{k,s}^{ij} \delta_{k-s-1}$, for $i, j = 1, 2, \ldots, m$.

Remark 5. From Assumption 4, the measurement noises of any two sensors are correlated at the same sampling time and at consecutive sampling times and uncorrelated otherwise; the cross-covariances of $\widetilde{v}_k^i$ with $\widetilde{v}_k^j$, $\widetilde{v}_{k-1}^j$, and $\widetilde{v}_{k+1}^j$ are $\widetilde{R}_{k,k}^{ij}$, $\widetilde{R}_{k,k-1}^{ij}$, and $\widetilde{R}_{k,k+1}^{ij}$, respectively.

It is assumed that, at any sampling time, the outputs are transmitted from the m different sensors to a data processing center producing the signal estimation and, as a consequence of possible failures during the transmission process, one-step delays may occur randomly in the measurements used for estimation. These measurement delays are modelled by introducing different sequences of Bernoulli variables whose values, zero or one, indicate whether the current measurement is up-to-date or delayed, respectively. Specifically, assume that, at initial time $k = 1$, the actual outputs, $\widetilde{y}_1^i$, are always available for the estimation but, at any time $k > 1$, the available measurements coming from each sensor may be randomly delayed by one sampling time according to different delay rates. Therefore, if $\{\gamma_k^i;\ k > 1\}$, $i = 1, 2, \ldots, m$, denote sequences of Bernoulli random variables, the available measurements from the ith sensor are described by

$$\begin{aligned} y_k^i &= \left(1 - \gamma_k^i\right) \widetilde{y}_k^i + \gamma_k^i \widetilde{y}_{k-1}^i, \quad k > 1; \\ y_1^i &= \widetilde{y}_1^i, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{2}$$

Remark 6. Model (2) is commonly used to describe measurements coming from multiple sensors which are one-step randomly delayed with different delay rates (see, e.g., [13] using the state-space model and [15] using covariance information). From (2) it is clear that if $\gamma_k^i = 1$, which occurs with a certain probability p_k^i, then $y_k^i = \widetilde{y}_{k-1}^i$ and the measurement from the ith sensor is delayed by one sampling period; otherwise, $\gamma_k^i = 0$ and $y_k^i = \widetilde{y}_k^i$, which means that the measurement is up-to-date with probability $1 - p_k^i$. Therefore, the variables $\{\gamma_k^i;\ k > 1\}$ model the random delays of the ith sensor and the following assumption is made.

Assumption 7. For $i = 1, 2, \ldots, m$, the process $\{\gamma_k^i; k > 1\}$ is a sequence of independent Bernoulli random variables with known probabilities $P[\gamma_k^i = 1] = p_k^i$, $\forall k > 1$. For $i, j = 1, 2, \ldots, m$ the variables γ_k^i and γ_s^j are independent for $k \neq s$, and $\mathrm{Cov}[\gamma_k^i, \gamma_k^j]$ are known.

Note that this assumption is more general than that considered in [13, 15] where the processes $\{\gamma_k^i;\ k > 1\}$, for $i = 1, 2, \ldots, m$, are assumed to be mutually independent.

Finally, the following independence hypothesis is also assumed.

Assumption 8. For $i = 1, 2, \ldots, m$, the signal process, $\{z_k;\ k \geq 1\}$, and the processes $\{H_k^i;\ k \geq 1\}$, $\{\widetilde{v}_k^i;\ k \geq 1\}$, and $\{\gamma_k^i;\ k > 1\}$ are mutually independent.

To address the optimal LS linear estimation problem of the signal based on the measurements coming from all the sensors, $\{y_1^i, y_2^i, \ldots, y_L^i, i = 1, 2, \ldots, m\}$, $L \geq k$, the centralized fusion method will be used. For this purpose, the observation equations of the different sensors (1) and (2) are combined yielding the following vectorial observation model:

$$\begin{aligned} \widetilde{y}_k &= H_k z_k + \widetilde{v}_k, \quad k \geq 1, \\ y_k &= \left(I - \Gamma_k\right) \widetilde{y}_k + \Gamma_k \widetilde{y}_{k-1}, \quad k > 1;\ y_1 = \widetilde{y}_1, \end{aligned} \tag{3}$$

where $\widetilde{y}_k = (\widetilde{y}_k^1, \dots, \widetilde{y}_k^m)^T$, $H_k = (H_k^{1T}, \dots, H_k^{mT})^T$, $\widetilde{v}_k = (\widetilde{v}_k^1, \dots, \widetilde{v}_k^m)^T$, and $\Gamma_k = \mathrm{Diag}(\gamma_k^1, \dots, \gamma_k^m)$.

Hence, the problem is to obtain the LS linear estimator of the signal, z_k, based on the randomly delayed observations $\{y_1, \dots, y_L\}$ given in (3). Next, we present the statistical properties of the processes involved in observation model (3), from which the LS linear filtering and fixed-point smoothing algorithms of the signal z_k will be derived; these properties are easily inferred from the model Assumptions 3–8 previously established.

(i) $\{H_k; k \geq 1\}$ are $m \times n$ independent random parameter matrices with known means, $\overline{H}_k \equiv E[H_k] = (\overline{H}_k^{1T}, \dots, \overline{H}_k^{mT})^T$, and known covariances, $\mathrm{Cov}[h_{i,p}^k, h_{j,q}^s]$, where $h_{i,p}^k$ denotes the (i, p)th entry of matrix, H_k, for $i = 1, 2, \dots, m$ and $p = 1, 2, \dots, n$.

(ii) $\{\widetilde{v}_k;\ k \geq 1\}$ is a zero-mean process with $\mathrm{Cov}[\widetilde{v}_k, \widetilde{v}_s] = \widetilde{R}_{k,k}\delta_{k-s} + \widetilde{R}_{k,s}\delta_{k-s+1} + \widetilde{R}_{k,s}\delta_{k-s-1}$, where $\widetilde{R}_{k,s} = (\widetilde{R}_{k,s}^{ij})_{i,j=1,2,\dots,m}$.

(iii) The random matrices $\{\Gamma_k;\ k > 1\}$ are independent or, equivalently, the m-dimensional process $\{\gamma_k;\ k > 1\}$, where $\gamma_k = (\gamma_k^1, \dots, \gamma_k^m)^T$, is a white sequence. The first- and second-order moments of these processes are known, and the following notation will be used:

$$\begin{aligned}
&\overline{\Gamma}_k \equiv E[\Gamma_k] = \mathrm{Diag}\left(p_k^1, \dots, p_k^m\right), \\
&K_k^{\gamma} \equiv E\left[\gamma_k \gamma_k^T\right], \\
&K_k^{\mathbf{1}-\gamma} \equiv E\left[(\mathbf{1} - \gamma_k)(\mathbf{1} - \gamma_k)^T\right], \\
&K_k^{\gamma,\mathbf{1}-\gamma} \equiv E\left[\gamma_k(\mathbf{1} - \gamma_k)^T\right].
\end{aligned} \tag{4}$$

(iv) The signal process, $\{z_k;\ k \geq 1\}$, and the processes $\{H_k;\ k \geq 1\}$, $\{\widetilde{v}_k;\ k \geq 1\}$, and $\{\Gamma_k;\ k > 1\}$ are mutually independent.

Remark 9. From the above properties, the following ones, which will be frequently used in the derivation of the algorithms, are obtained.

(a) The covariances of vectors $\widetilde{H}_k z_k$, with $\widetilde{H}_k = H_k - \overline{H}_k$, are given by

$$K_k^{\widetilde{H}z} \equiv E\left[\widetilde{H}_k z_k z_k^T \widetilde{H}_k^T\right] = E\left[\widetilde{H}_k A_k B_k^T \widetilde{H}_k^T\right], \quad k \geq 1. \tag{5}$$

This identity is easily obtained from the conditional expectation properties, using the independence of z_k and $\widetilde{H}_k$, and Assumption 1:

$$\begin{aligned}
E\left[\widetilde{H}_k z_k z_k^T \widetilde{H}_k^T\right] &= E\left[E\left[\widetilde{H}_k z_k z_k^T \widetilde{H}_k^T \mid \widetilde{H}_k\right]\right] \\
&= E\left[\widetilde{H}_k E\left[z_k z_k^T \mid \widetilde{H}_k\right] \widetilde{H}_k^T\right] \\
&= E\left[\widetilde{H}_k E\left[z_k z_k^T\right] \widetilde{H}_k^T\right].
\end{aligned} \tag{6}$$

From (5), these matrices are known and their entries are given by

$$\left(K_k^{\widetilde{H}z}\right)_{ij} = \sum_{p=1}^{n}\sum_{q=1}^{n} \mathrm{Cov}\left[h_{i,p}^k, h_{j,q}^k\right]\left(A_k B_k^T\right)_{pq}, \quad i, j = 1, 2, \dots, m. \tag{7}$$

(b) The covariances of the vectors $H_k z_k$ are given by

$$E\left[H_k z_k z_k^T H_k^T\right] = K_k^{\widetilde{H}z} + \overline{H}_k A_k B_k^T \overline{H}_k^T, \quad k \geq 1. \tag{8}$$

(c) If G is a $m \times m$ random matrix independent of $\{\Gamma_k,\ k > 1\}$, the Hadamard product properties and (iii) lead to

$$\begin{aligned}
E\left[\Gamma_k G \Gamma_k\right] &= K_k^{\gamma} \circ E[G], \\
E\left[\Gamma_k G (I - \Gamma_k)\right] &= K_k^{\gamma,\mathbf{1}-\gamma} \circ E[G], \\
E\left[(I - \Gamma_k) G (I - \Gamma_k)\right] &= K_k^{\mathbf{1}-\gamma} \circ E[G].
\end{aligned} \tag{9}$$

To simplify future formulas and expressions, the observation model (3) will be written equivalently as follows:

$$\begin{aligned}
&y_k = x_k + v_k, \quad k \geq 1, \\
&x_k = (I - \Gamma_k) H_k z_k + \Gamma_k H_{k-1} z_{k-1}, \quad k \geq 2; \quad x_1 = H_1 z_1, \\
&v_k = (I - \Gamma_k)\widetilde{v}_k + \Gamma_k \widetilde{v}_{k-1}, \quad k \geq 2; \quad v_1 = \widetilde{v}_1.
\end{aligned} \tag{10}$$

Taking into account the model properties and those specified in Remark 9, the first- and second-order properties of the processes $\{x_k;\ k \geq 1\}$ and $\{v_k;\ k \geq 1\}$ and, consequently, those of the observation process, $\{y_k;\ k \geq 1\}$, are easily obtained. They are established in the following lemmas.

Lemma 10. *The process $\{x_k;\ k \geq 1\}$ has zero mean and $K_k^x \equiv \mathrm{Cov}[x_k, x_k]$ is given by*

$$\begin{aligned}
K_k^x = {} & K_k^{1-\gamma} \circ \left(K_k^{\widetilde{H}z} + \overline{H}_k A_k B_k^T \overline{H}_k^T\right) \\
& + K_k^{1-\gamma,\gamma} \circ \left(\overline{H}_k A_k B_{k-1}^T \overline{H}_{k-1}^T\right) \\
& + K_k^{\gamma,1-\gamma} \circ \left(\overline{H}_{k-1} B_{k-1} A_k^T \overline{H}_k^T\right) \\
& + K_k^{\gamma} \circ \left(K_{k-1}^{\widetilde{H}z} + \overline{H}_{k-1} A_{k-1} B_{k-1}^T \overline{H}_{k-1}^T\right), \quad k \geq 2; \\
K_1^x = {} & K_1^{\widetilde{H}z} + \overline{H}_1 A_1 B_1^T \overline{H}_1^T.
\end{aligned} \tag{11}$$

Lemma 11. *$\{v_k;\ k \geq 1\}$ has zero mean and $Cov[v_k, v_s] = R_{k,k}\delta_{k-s} + R_{k,s}\delta_{k-s-1} + R_{k,s}\delta_{k-s-2}$, $s \leq k$, where*

$$\begin{aligned}
R_{k,k} &= K_k^{1-\gamma} \circ \widetilde{R}_{k,k} + K_k^{1-\gamma,\gamma} \circ \widetilde{R}_{k,k-1} \\
&\quad + K_k^{\gamma,1-\gamma} \circ \widetilde{R}_{k-1,k} + K_k^{\gamma} \circ \widetilde{R}_{k-1,k-1}, \quad k > 1; \\
R_{k,k-1} &= \left(I - \overline{\Gamma}_k\right) \widetilde{R}_{k,k-1} \left(I - \overline{\Gamma}_{k-1}\right) + \overline{\Gamma}_k \widetilde{R}_{k-1,k-1} \left(I - \overline{\Gamma}_{k-1}\right) \\
&\quad + \Gamma_k \widetilde{R}_{k-1,k-2} \overline{\Gamma}_{k-1}, \quad k > 2; \\
R_{k,k-2} &= \overline{\Gamma}_k \widetilde{R}_{k-1,k-2} \left(I - \overline{\Gamma}_{k-2}\right), \quad k > 3; \\
R_{1,1} &= \widetilde{R}_{1,1}, \quad R_{2,1} = \left(I - \overline{\Gamma}_2\right) \widetilde{R}_{2,1} + \overline{\Gamma}_2 \widetilde{R}_{1,1}, \quad R_{3,1} = \overline{\Gamma}_3 \widetilde{R}_{2,1}.
\end{aligned} \tag{12}$$

Lemma 12. *The processes $\{x_k;\ k \geq 1\}$ and $\{v_k;\ k \geq 1\}$ are uncorrelated and, consequently,*

$$K_k^y \equiv Cov[y_k, y_k] = K_k^x + R_{k,k}, \quad k \geq 1, \tag{13}$$

with K_k^x and $R_{k,k}$ given in (11) *and* (12)*, respectively.*

3. Innovation Approach to the LS Linear Estimation Problem

To obtain a recursive algorithm for the LS linear estimator, $\widehat{z}_{k/L}$, of the signal, z_k, based on the randomly delayed observations, $\{y_1, \ldots, y_L\}$, an innovation approach will be used [32]. This approach consists of transforming the observation process $\{y_k;\ k \geq 1\}$ into an equivalent one (innovation process) of orthogonal vectors $\{\mu_k;\ k \geq 1\}$, defined by $\mu_k = y_k - \widehat{y}_{k/k-1}$, where $\widehat{y}_{k/k-1}$ is the orthogonal projection of y_k into the linear space generated by $\{\mu_1, \ldots, \mu_{k-1}\}$. The orthogonality property of the new process allows us to simplify the estimators' expressions (which also simplifies the algorithms derivation) in comparison to those obtained when the estimators are expressed directly as linear combination of the observations.

Specifically, if w_k denotes a random vector to be estimated, the LS linear estimator of w_k based on the observations $\{y_1, \ldots, y_L\}$ (which will be denoted as $\widehat{w}_{k/L}$) agrees with that based on the innovations $\{\mu_1, \ldots, \mu_L\}$ or, equivalently, with the orthogonal projection of w_k onto the linear space generated by $\{\mu_1, \ldots, \mu_L\}$. Hence,

$$\widehat{w}_{k/L} = \sum_{j=1}^{L} N_{k,j} \mu_j, \tag{14}$$

and the impulse-response function, $N_{k,j}$, $j = 1, \ldots, L$, is calculated from the orthogonality property, $E[(w_k - \widehat{w}_{k/L})\mu_s^T] = 0$, $s \leq L$, which leads to the Wiener-Hopf equation, taking into account that $E[\mu_j \mu_s^T] = 0$ for $j \neq s$,

$$E\left[w_k \mu_s^T\right] = N_{k,s} E\left[\mu_s \mu_s^T\right], \quad s \leq L. \tag{15}$$

Consequently, by denoting $\Pi_j = E[\mu_j \mu_j^T]$, the following general expression for the LS linear estimators of w_k is obtained:

$$\widehat{w}_{k/L} = \sum_{j=1}^{L} E\left[w_k \mu_j^T\right] \Pi_j^{-1} \mu_j. \tag{16}$$

3.1. Innovation Process. As indicated above, the innovation at time k is defined as $\mu_k = y_k - \widehat{y}_{k/k-1}$, where $\widehat{y}_{k/k-1}$, the orthogonal projection of y_k onto the linear space generated by $\{\mu_1, \ldots, \mu_{k-1}\}$, is the LS one-stage linear predictor of y_k. From (10) and the orthogonal projection lemma, this estimator can be expressed by

$$\begin{aligned}
\widehat{y}_{k/k-1} &= \widehat{x}_{k/k-1} + \widehat{v}_{k/k-1}, \quad k \geq 2, \\
\widehat{y}_{1/0} &= 0,
\end{aligned} \tag{17}$$

so we need the one-stage predictors $\widehat{x}_{k/k-1}$ and $\widehat{v}_{k/k-1}$ which, by using the general expression (16) for the LS linear estimators, are given by

$$\begin{aligned}
\widehat{x}_{k/k-1} &= \sum_{j=1}^{k-1} E\left[x_k \mu_j^T\right] \Pi_j^{-1} \mu_j, \\
\widehat{v}_{k/k-1} &= \sum_{j=1}^{k-1} E\left[v_k \mu_j^T\right] \Pi_j^{-1} \mu_j.
\end{aligned} \tag{18}$$

(1) From the independence property (iv), it is clear that

$$E\left[x_k \mu_j^T\right] = \left(I - \overline{\Gamma}_k\right) \overline{H}_k E\left[z_k \mu_j^T\right] + \overline{\Gamma}_k \overline{H}_{k-1} E\left[z_{k-1} \mu_j^T\right], \quad j \leq k-2, \tag{19}$$

and hence, for $k > 2$,

$$\begin{aligned}
\widehat{x}_{k/k-1} &= \left(I - \overline{\Gamma}_k\right) \overline{H}_k \sum_{j=1}^{k-2} E\left[z_k \mu_j^T\right] \Pi_j^{-1} \mu_j \\
&\quad + \overline{\Gamma}_k \overline{H}_{k-1} \sum_{j=1}^{k-2} E\left[z_{k-1} \mu_j^T\right] \Pi_j^{-1} \mu_j \\
&\quad + E\left[x_k \mu_{k-1}^T\right] \Pi_{k-1}^{-1} \mu_{k-1};
\end{aligned} \tag{20}$$

then, from (16) for $\widehat{z}_{k/k-2}$ and $\widehat{z}_{k-1/k-2}$, we obtain

$$\begin{aligned}
\widehat{x}_{k/k-1} &= \left(I - \overline{\Gamma}_k\right) \overline{H}_k \widehat{z}_{k/k-2} + \overline{\Gamma}_k \overline{H}_{k-1} \widehat{z}_{k-1/k-2} \\
&\quad + E\left[x_k \mu_{k-1}^T\right] \Pi_{k-1}^{-1} \mu_{k-1}, \quad k > 2, \\
\widehat{x}_{2/1} &= E\left[x_2 \mu_1^T\right] \Pi_1^{-1} \mu_1.
\end{aligned} \tag{21}$$

(2) The uncorrelation of $\widetilde{v}_k$ and $\widetilde{v}_{k-1}$ with $\widetilde{v}_1, \ldots, \widetilde{v}_{k-3}$ and the independence property (iv) guarantee that

$E[v_k\mu_j^T] = 0, j = 1,\ldots,k-3$, and $E[v_k\mu_{k-2}^T] = E[v_k y_{k-2}^T]$, and this last expectation is equal to $E[v_k v_{k-2}^T]$ from the uncorrelation of v_k and x_{k-2}; hence,

$$\widehat{v}_{k/k-1} = R_{k,k-2}\Pi_{k-2}^{-1}\mu_{k-2} + E\left[v_k\mu_{k-1}^T\right]\Pi_{k-1}^{-1}\mu_{k-1}, \quad k > 2,$$
$$\widehat{v}_{2/1} = E\left[v_2\mu_1^T\right]\Pi_1^{-1}\mu_1. \tag{22}$$

Now, from (21) and (22), by denoting $T_{k,k-1} = E[y_k\mu_{k-1}^T]$, it is immediately clear that the innovation at time k can be expressed as

$$\mu_k = y_k - \left(I - \overline{\Gamma}_k\right)\overline{H}_k\widehat{z}_{k/k-2} - \overline{\Gamma}_k\overline{H}_{k-1}\widehat{z}_{k-1/k-2}$$
$$- R_{k,k-2}\Pi_{k-2}^{-1}\mu_{k-2} - T_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k > 2; \tag{23}$$
$$\mu_2 = y_2 - T_{2,1}\Pi_1^{-1}\mu_1; \qquad \mu_1 = y_1,$$

and, hence, its determination requires that of the linear signal predictors, $\widehat{z}_{k/L}$, $L = k-1, k-2$. The derivation of the linear predictors is analogous to that of the filter, $\widehat{z}_{k/k}$, so both are obtained simultaneously in the following section.

4. Prediction and Filtering Recursive Algorithm

The following theorem presents a recursive algorithm for the signal LS linear predictor and filter based on the delayed observation model given in Section 2.

Theorem 13. *The signal predictors $\widehat{z}_{k/L}$, $L < k$, and the signal filter, $\widehat{z}_{k/k}$, are obtained as*

$$\widehat{z}_{k/L} = A_k O_L, \quad L < k; \quad \widehat{z}_{k/k} = A_k O_k, \tag{24}$$

where the vectors O_k are recursively calculated from

$$O_k = O_{k-1} + J_k\Pi_k^{-1}\mu_k, \quad k \geq 1; \quad O_0 = 0. \tag{25}$$

The matrix function J is given by

$$J_k = G_{B_k}^T - r_{k-2}G_{A_k}^T - J_{k-2}\Pi_{k-2}^{-1}R_{k,k-2}^T - J_{k-1}\Pi_{k-1}^{-1}T_{k,k-1}^T, \quad k > 2;$$
$$J_2 = G_{B_2}^T - J_1\Pi_1^{-1}T_{2,1}^T, \qquad J_1 = B_1^T\overline{H}_1^T, \tag{26}$$

with $r_k = E[O_k O_k^T]$ recursively obtained from

$$r_k = r_{k-1} + J_k\Pi_k^{-1}J_k^T, \quad k \geq 1; \quad r_0 = 0. \tag{27}$$

The innovation, μ_k, satisfies

$$\mu_k = y_k - G_{A_k}O_{k-2} - R_{k,k-2}\Pi_{k-2}^{-1}\mu_{k-2} - T_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k > 2;$$
$$\mu_2 = y_2 - T_{2,1}\Pi_1^{-1}\mu_1, \qquad \mu_1 = y_1, \tag{28}$$

where $T_{k,k-1} = E[y_k\mu_{k-1}^T]$ is recursively obtained from

$$T_{k,k-1} = G_{A_k}J_{k-1} + \overline{\Gamma}_k K_{k-1}^{\overline{H}z}\left(I - \overline{\Gamma}_{k-1}\right) + R_{k,k-1}$$
$$- R_{k,k-2}\Pi_{k-2}^{-1}T_{k-1,k-2}^T, \quad k > 2; \tag{29}$$
$$T_{2,1} = G_{A_2}B_1^T\overline{H}_1^T + \overline{\Gamma}_2 K_1^{\overline{H}z} + R_{2,1}.$$

The innovation covariance matrix, Π_k, is given by

$$\Pi_k = K_k^y - G_{A_k}r_{k-2}G_{A_k}^T - R_{k,k-2}\Pi_{k-2}^{-1}R_{k,k-2}^T$$
$$- T_{k,k-1}\Pi_{k-1}^{-1}T_{k,k-1}^T - G_{A_k}J_{k-2}\Pi_{k-2}^{-1}R_{k,k-2}^T$$
$$- R_{k,k-2}\Pi_{k-2}^{-1}J_{k-2}^T G_{A_k}^T, \quad k > 2; \tag{30}$$
$$\Pi_2 = K_2^y - T_{2,1}\Pi_1^{-1}T_{2,1}^T, \qquad \Pi_1 = K_1^y.$$

The matrices $K_k^{\overline{H}z}$ and K_k^y are given in (5) *and* (13)*, respectively. $R_{k,s}$, for $s = k, k-1, k-2$, are given in* (12)*. Finally, the matrices G_{A_k} and G_{B_k} are defined by*

$$G_{\Psi_k} = \left(I - \overline{\Gamma}_k\right)\overline{H}_k\Psi_k + \overline{\Gamma}_k\overline{H}_{k-1}\Psi_{k-1}, \quad \Psi = A, B. \tag{31}$$

Proof. From the general expression (16), $\widehat{z}_{k/L} = \sum_{j=1}^{L} E[z_k\mu_j^T]\Pi_j^{-1}\mu_j$, for $L \leq k$, and the coefficients $S_{k,j} = E[z_k\mu_j^T]$, $j \leq k$, must be calculated in order to determine the predictors and filter of z_k. Using expression (23) for μ_j, we obtain

$$S_{k,j} = E\left[z_k y_j^T\right] - E\left[z_k\widehat{z}_{j/j-2}^T\right]\overline{H}_j^T\left(I - \overline{\Gamma}_j\right)$$
$$- E\left[z_k\widehat{z}_{j-1/j-2}^T\right]\overline{H}_{j-1}^T\overline{\Gamma}_j$$
$$- S_{k,j-2}\Pi_{j-2}^{-1}R_{j,j-2}^T - S_{k,j-1}\Pi_{j-1}^{-1}T_{j,j-1}^T, \quad 2 < j \leq k,$$
$$S_{k,2} = E\left[z_k y_2^T\right] - S_{k,1}\Pi_1^{-1}T_{2,1}^T,$$
$$S_{k,1} = E\left[z_k y_1^T\right]. \tag{32}$$

(a) On the one hand, from (10) and independence hypotheses, we have

$$E\left[z_k y_j^T\right] = E\left[z_k x_j^T\right]$$
$$= E\left[z_k z_j^T\right]\overline{H}_j^T\left(I - \overline{\Gamma}_j\right) + E\left[z_k z_{j-1}^T\right]\overline{H}_{j-1}^T\overline{\Gamma}_j$$
$$= A_k G_{B_j}^T, \quad j \geq 2,$$
$$E\left[z_k y_1^T\right] = E\left[z_k x_1^T\right] = E\left[z_k z_1^T\right]\overline{H}_1^T = A_1 B_1^T\overline{H}_1^T, \tag{33}$$

where Assumption 1 and expression (31) for G_{B_j} have been used.

(b) On the other hand, using again (16) for $\widehat{z}_{j/j-2}$ and $\widehat{z}_{j-1/j-2}$ and $E[z_k\mu_i^T] = S_{k,i}$, we have

$$E\left[z_k\widehat{z}_{h/j-2}^T\right] = E\left[z_k\left(\sum_{i=1}^{j-2}S_{h,i}\Pi_i^{-1}\mu_i\right)^T\right] = \sum_{i=1}^{j-2}S_{k,i}\Pi_i^{-1}S_{h,i}^T, \quad h = j, j-1, \quad j > 2. \tag{34}$$

Therefore,

$$\begin{aligned}
S_{k,j} &= A_k G_{B_j}^T - \sum_{i=1}^{j-2}S_{k,i}\Pi_i^{-1}S_{j,i}^T\overline{H}_j^T\left(I-\overline{\Gamma}_j\right) \\
&\quad - \sum_{i=1}^{j-2}S_{k,i}\Pi_i^{-1}S_{j-1,i}^T\overline{H}_{j-1}^T\overline{\Gamma}_j \\
&\quad - S_{k,j-2}\Pi_{j-2}^{-1}R_{j,j-2}^T - S_{k,j-1}\Pi_{j-1}^{-1}T_{j,j-1}^T, \quad 2 < j \le k, \\
S_{k,2} &= A_k G_{B_2}^T - S_{k,1}\Pi_1^{-1}T_{2,1}^T, \\
S_{k,1} &= A_k B_1^T\overline{H}_1^T,
\end{aligned} \tag{35}$$

and this expression guarantees that

$$S_{k,j} = A_k J_j, \quad 1 \le j \le k, \tag{36}$$

where J is a function satisfying

$$\begin{aligned}
J_j &= G_{B_j}^T - \sum_{i=1}^{j-2}J_i\Pi_i^{-1}S_{j,i}^T\overline{H}_j^T\left(I-\overline{\Gamma}_j\right) - \sum_{i=1}^{j-2}J_i\Pi_i^{-1}S_{j-1,i}^T\overline{H}_{j-1}^T\overline{\Gamma}_j \\
&\quad - J_{j-2}\Pi_{j-2}^{-1}R_{j,j-2}^T - J_{j-1}\Pi_{j-1}^{-1}T_{j,j-1}^T, \quad j > 2, \\
J_2 &= G_{B_2}^T - J_1\Pi_1^{-1}T_{2,1}^T, \\
J_1 &= B_1^T\overline{H}_1^T.
\end{aligned} \tag{37}$$

Hence, denoting

$$O_k = \sum_{i=1}^{k}J_i\Pi_i^{-1}\nu_i, \qquad O_0 = 0, \tag{38}$$

which obviously satisfies (25), expression (24) for the predictors and filter is proved.

Next, taking into account (36) and denoting

$$r_k = E\left[O_k O_k^T\right] = \sum_{j=1}^{k}J_j\Pi_j^{-1}J_j^T, \quad k \ge 1; \quad r_0 = 0, \tag{39}$$

expression (26) for J_k is easily derived just making $j = k$ in (37). The recursive formula (27) for r_k is immediately clear from (39).

Expression (28) for μ_k is derived by substituting $\widehat{z}_{k/k-2} = A_k O_{k-2}$ and $\widehat{z}_{k-1/k-2} = A_{k-1}O_{k-2}$ in (23) and considering expression (31) for G_{A_k}.

To prove recursive expression (29) for $T_{k,k-1} = E[y_k\mu_{k-1}^T] = E[x_k\mu_{k-1}^T] + E[v_k\mu_{k-1}^T]$, $k \ge 2$, we calculate both expectations as follows.

(1) From expression (10) for x_k, using the independence properties and $E[z_h\mu_{k-1}^T] = S_{h,k-1} = A_h J_{k-1}$, for $h = k, k-1$, we have

$$E\left[x_k\mu_{k-1}^T\right] = G_{A_k}J_{k-1} + \overline{\Gamma}_k E\left[\widetilde{H}_{k-1}z_{k-1}\mu_{k-1}^T\right], \quad k > 2, \tag{40}$$

and since

$$\begin{aligned}
E\left[\widetilde{H}_{k-1}z_{k-1}\mu_{k-1}^T\right] &= E\left[\widetilde{H}_{k-1}z_{k-1}y_{k-1}^T\right] = E\left[\widetilde{H}_{k-1}z_{k-1}x_{k-1}^T\right] \\
&= E\left[\widetilde{H}_{k-1}z_{k-1}z_{k-1}^T\widetilde{H}_{k-1}^T\right]\left(I-\overline{\Gamma}_{k-1}\right),
\end{aligned} \tag{41}$$

it is clear that

$$E\left[x_k\mu_{k-1}^T\right] = G_{A_k}J_{k-1} + \overline{\Gamma}_k K_{k-1}^{\widetilde{H}z}\left(I-\overline{\Gamma}_{k-1}\right), \quad k > 2. \tag{42}$$

Analogously, $E[x_2\mu_1^T] = G_{A_2}J_1 + \overline{\Gamma}_2 K_1^{\widetilde{H}z}$.

(2) Using that $\widehat{y}_{k-1/k-2} = \sum_{j=1}^{k-2}T_{k-1,j}\Pi_j^{-1}\mu_j$ and $E[v_k\mu_j^T] = 0$, $j = 1, \ldots, k-3$, we have

$$\begin{aligned}
E\left[v_k\mu_{k-1}^T\right] &= E\left[v_k y_{k-1}^T\right] - E\left[v_k\widehat{y}_{k-1/k-2}^T\right] \\
&= E\left[v_k y_{k-1}^T\right] - E\left[v_k\mu_{k-2}^T\right]\Pi_{k-2}^{-1}T_{k-1,k-2}^T, \quad k > 2,
\end{aligned} \tag{43}$$

and since $E[v_k y_{k-1}^T] = E[v_k v_{k-1}^T]$ and $E[v_k\mu_{k-2}^T] = E[v_k y_{k-2}^T] = E[v_k v_{k-2}^T]$, we obtain

$$E\left[v_k\mu_{k-1}^T\right] = R_{k,k-1} - R_{k,k-2}\Pi_{k-2}^{-1}T_{k-1,k-2}^T, \quad k > 2. \tag{44}$$

Clearly, $E[v_2\mu_1^T] = E[v_2 y_1^T] = E[v_2 v_1^T] = R_{2,1}$.

So expression (29) is proved.

Finally, formula (30) for the innovation covariance matrices is obtained by writing $\Pi_k = E[y_k y_k^T] - E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$, using the expression for the observation predictor, $\widehat{y}_{k/k-1}$, and taking into account that $E[O_k O_k^T] = r_k$ and $E[O_k\mu_k^T] = J_k$. □

4.1. Filtering Error Covariance Matrices. The performance of the LS estimators $\widehat{z}_{k/L}$, $L \le k$, is measured by the covariance matrices of the estimation errors, $\Sigma_{k/L} = E[(z_k - \widehat{z}_{k/L})(z_k - \widehat{z}_{k/L})^T]$. Since the error of a LS linear estimator is orthogonal to the estimator, using Assumption 1, these matrices are given by

$$\Sigma_{k/L} = A_k B_k^T - E\left[\widehat{z}_{k/L}\widehat{z}_{k/L}^T\right], \quad L \le k. \tag{45}$$

Then, by using expression (24) and taking into account that $r_L = E[O_L O_L^T]$, we obtain the following expressions for the prediction and filtering error covariance matrices:

$$\Sigma_{k/L} = A_k \left[B_k^T - r_L A_k^T\right], \quad L \leq k. \tag{46}$$

Note that the computation of the prediction and filtering error covariance matrices does not depend on the current set of observations, as it only needs the matrices A_k and B_k, which are known, and the matrices r_L, which are recursively calculated from (27); hence, the error covariance matrices provide a measure of the estimator performance even before we get any observed data.

5. Fixed-Point Smoothing Algorithm

In this section, we present a recursive algorithm for the LS linear fixed-point smoothers, $\widehat{z}_{k/L}$, $L > k$, where k is fixed and recursions for increasing L are proposed. By starting from the general expression for the LS linear estimator of the signal, $\widehat{z}_{k/L} = \sum_{j=1}^{L} S_{k,j}\Pi_j^{-1}\mu_j$, where $S_{k,j} = E[z_k \mu_j^T]$ and $\Pi_j = E[\mu_j \mu_j^T]$, it is clear that the linear fixed-point smoothers, $\widehat{z}_{k/L}$, $L > k$, can be recursively calculated as

$$\widehat{z}_{k/L} = \widehat{z}_{k/L-1} + S_{k,L}\Pi_L^{-1}\mu_L, \quad L > k, \tag{47}$$

with the linear filter, $\widehat{z}_{k/k}$, as initial condition.

Hence, to calculate the fixed-point smoothing estimators, $\widehat{z}_{k/L}$, for $L > k$ (k fixed), we need a recursive relation in L for $S_{k,L} = E[z_k \mu_L^T] = E[z_k y_L^T] - E[z_k \widehat{y}_{L/L-1}^T]$, $L > k$.

On the one hand, as in the proof of Theorem 13, using (10) and taking into account the independence hypotheses, together with Assumption 1 and expression (31) for G_{A_L}, we have

$$\begin{aligned} E\left[z_k y_L^T\right] &= E\left[z_k x_L^T\right] \\ &= E\left[z_k z_L^T\right]\overline{H}_L^T\left(I - \overline{\Gamma}_L\right) + E\left[z_k z_{L-1}^T\right]\overline{H}_{L-1}^T\overline{\Gamma}_j \\ &= B_k G_{A_L}^T, \quad L > k. \end{aligned} \tag{48}$$

On the other hand, the expression of $\widehat{y}_{L/L-1}^T$ obtained from (28) for $k = L$ yields

$$\begin{aligned} E\left[z_k \widehat{y}_{L/L-1}^T\right] &= E\left[z_k O_{L-2}^T\right] G_{A_L}^T + S_{k,L-2}\Pi_{L-2}^{-1}R_{L,L-2}^T \\ &\quad + S_{k,L-1}\Pi_{L-1}^{-1}T_{L,L-1}^T, \quad L > k, \ (L > 2), \\ E\left[z_1 \widehat{y}_{2/1}^T\right] &= S_{1,1}\Pi_1^{-1}T_{2,1}^T = A_1 J_1 \Pi_1^{-1} T_{2,1}^T. \end{aligned} \tag{49}$$

Therefore, defining the function $E_{k,L} = E[z_k O_L^T]$, the following expression holds:

$$\begin{aligned} S_{k,L} &= \left[B_k - E_{k,L-2}\right] G_{A_L}^T - S_{k,L-2}\Pi_{L-2}^{-1}R_{L,L-2}^T \\ &\quad - S_{k,L-1}\Pi_{L-1}^{-1}T_{L,L-1}^T, \quad L > k, \ (L > 2), \\ S_{1,2} &= B_1 G_{A_2}^T - S_{1,1}\Pi_1^{-1}T_{2,1}^T, \end{aligned} \tag{50}$$

with initial conditions given by $S_{k,k-1} = A_k J_{k-1}$ and $S_{k,k} = A_k J_k$, from (36).

Finally, we need a recursive expression for $E_{k,L}$, $L > k - 2$. Taking into account that, from the orthogonality property, $E_{k,k-1} = E[z_k O_{k-1}^T] = E[\widehat{z}_{k/k-1} O_{k-1}^T]$ and $E_{k,k} = E[z_k O_k^T] = E[\widehat{z}_{k/k} O_k^T]$, using (24), and that $r_k = E[O_k O_k^T]$, we have that $E_{k,k-1} = A_k r_{k-1}$ and $E_{k,k} = A_k r_k$. Now, using (25) for O_L, the following formula is immediately deduced:

$$E_{k,L} = E_{k,L-1} + S_{k,L}\Pi_L^{-1}J_L^T, \quad L > k. \tag{51}$$

Summarizing these results, the following recursive *fixed-point smoothing algorithm* is obtained.

Theorem 14. *The fixed-point smoother $\widehat{z}_{k/L}$, with, $L > k$, of the signal z_k is calculated as*

$$\widehat{z}_{k/L} = \widehat{z}_{k/L-1} + S_{k,L}\Pi_L^{-1}\mu_L, \quad L > k, \tag{52}$$

with initial condition given by the filter, $\widehat{z}_{k/k}$, and

$$\begin{aligned} S_{k,L} &= \left[B_k - E_{k,L-2}\right] G_{A_L}^T - S_{k,L-2}\Pi_{L-2}^{-1}R_{L,L-2}^T \\ &\quad - S_{k,L-1}\Pi_{L-1}^{-1}T_{L,L-1}^T, \quad L > k, \ (L > 2), \\ S_{1,2} &= B_1 G_{A_2}^T - S_{1,1}\Pi_1^{-1}T_{2,1}^T, \end{aligned} \tag{53}$$

with $S_{k,k-1} = A_k J_{k-1}$ and $S_{k,k} = A_k J_k$.

The matrices $E_{k,L}$ satisfy the following recursive formula:

$$\begin{aligned} E_{k,L} &= E_{k,L-1} + S_{k,L}\Pi_L^{-1}J_L^T, \quad L > k; \\ E_{k,k-1} &= A_k r_{k-1}, \qquad E_{k,k} = A_k r_k. \end{aligned} \tag{54}$$

The filter $\widehat{z}_{k/k}$, the matrices G_{A_L}, $T_{L,L-1}$, and J_L, and the innovations ν_L and their covariance matrices Π_L are obtained from the linear filtering algorithm given in Theorem 13.

Using the recursive formula of the fixed-point smoother, the following recursive expression for the *fixed-point smoothing error covariance matrices*, $\Sigma_{k/L} = E[(z_k - \widehat{z}_{k/L})(z_k - \widehat{z}_{k/L})^T]$, $L > k$, is immediately deduced:

$$\Sigma_{k/L} = \Sigma_{k/L-1} - S_{k,L}\Pi_L^{-1}S_{k/L}^T, \quad L > k, \tag{55}$$

with the filtering error covariance matrix, $\Sigma_{k/k}$, as initial condition.

6. Numerical Simulation Example

In this section, the applicability of the proposed prediction, filtering, and fixed-point smoothing algorithms is shown by a numerical simulation example with two kinds of measured outputs. For this purpose, the signal values and their observations have been simulated in MATLAB and the signal estimates have been calculated, as well as the corresponding error variances, which provide a measure of the estimation accuracy.

It is assumed that $\{z_k;\ k \geq 1\}$ is a zero-mean scalar signal with autocovariance function $E[z_k z_j] = 1.025641 \times 0.95^{k-j}$,

$j \le k$, which is factorizable according to Assumption 1 just taking, for example, $A_k = 1.025641 \times 0.95^k$ and $B_k = 0.95^{-k}$. For the simulations, the signal is assumed to be generated by an autoregressive model, $z_{k+1} = 0.95 z_k + w_k$, where $\{w_k;\ k \ge 1\}$ is a zero-mean white Gaussian noise with variance 0.1, for all k.

Measurements coming from two sensors are considered and, according to the proposed observation model, it is assumed that, at any sampling time $k \ge 2$, the measured output from the ith sensor, $\tilde{y}_k^i$, can be randomly delayed by one sampling period during network transmission; that is,

$$y_k^i = \left(1 - \gamma_k^i\right) \tilde{y}_k^i + \gamma_k^i \tilde{y}_{k-1}^i, \quad k \ge 2; \qquad y_1^i = \tilde{y}_1^i, \quad i = 1, 2, \tag{56}$$

where $\{\gamma_k^i;\ k > 1\}$, $i = 1, 2$, are independent sequences of independent Bernoulli random variables with $P[\gamma_k^1 = 1] = p^i$, $\forall k > 1$.

Case 1 (systems with observation multiplicative noises). Consider measurements coming from two sensors,

$$\begin{aligned} \tilde{y}_k^1 &= \left(1 + 0.1\epsilon_k^1\right) z_k + \tilde{v}_k^1, \quad k \ge 1, \\ \tilde{y}_k^2 &= \left(0.5 + 0.1\epsilon_k^2\right) z_k + \tilde{v}_k^2, \quad k \ge 1, \end{aligned} \tag{57}$$

where the multiplicative noises $\{\epsilon_k^i;\ k \ge 1\}$, $i = 1, 2$, are independent zero-mean Gaussian white processes with unit variance, and the additive noises $\{\tilde{v}_k^i;\ k \ge 1\}$, $i = 1, 2$, are defined by $\tilde{v}_k^i = c_i(\eta_k + \eta_{k+1})$, $i = 1, 2$, with $c_1 = 1$, $c_2 = 0.5$, and $\{\eta_k;\ k \ge 1\}$ a zero-mean Gaussian white process with variance 0.5. Clearly, according to Assumption 4, the additive noises $\{\tilde{v}_k^i;\ k \ge 1\}$ are one-step autocorrelated with

$$\begin{aligned} \tilde{R}_{k,k}^{ii} &= c_i^2, \qquad \tilde{R}_{k,k+1}^{ii} = 0.5c_i^2, \\ \tilde{R}_{k,k}^{ij} &= c_i c_j, \qquad \tilde{R}_{k,k+1}^{ij} = 0.5 c_i c_j. \end{aligned} \tag{58}$$

Firstly, to compare the performance of the predictor, $\hat{z}_{k/k-1}$, filter, $\hat{z}_{k/k}$, and fixed-point smoothers, $\hat{z}_{k/L}$, with $L = k+1, k+2, k+3$, the corresponding error variances are calculated considering constant delay probabilities, $p^1 = 0.1$ and $p^2 = 0.3$. The results are displayed in Figure 1 which shows that the error variances corresponding to the fixed-point smoother are less than those of the filter and the filtering error variances are smaller than the prediction ones, thus confirming that the smoother has the best performance while the predictor has the worst performance. This figure also shows that the performance of the fixed-point smoothers improves as the number of available observations increases. Analogous results are obtained for other values of the probabilities p^i, $i = 1, 2$.

Next, we study the filtering error variances, $\Sigma_{k/k}$, when the delay probabilities p^1 and p^2 are varied from 0.1 to 0.9. In all the cases, the filtering error variances present insignificant variation from the 10th iteration onwards and, consequently, only the error variances at a specific iteration are shown here. Figure 2(a) displays the filtering error variances at $k = 50$

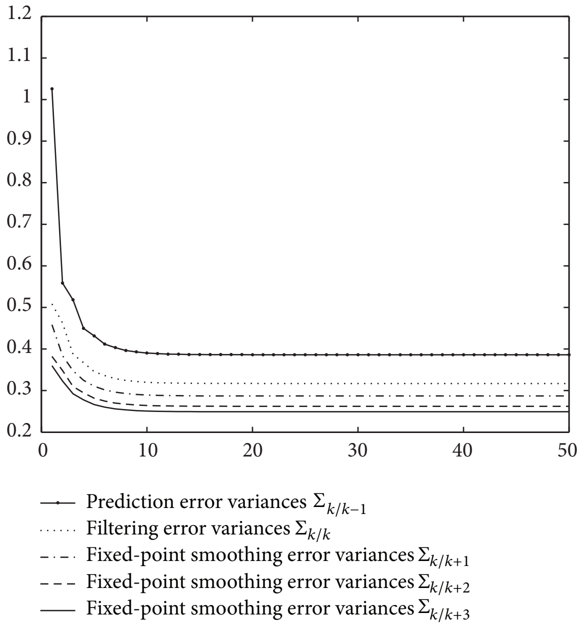


Figure 1: Prediction, filtering, and smoothing error variances, when $p^1 = 0.1$ and $p^2 = 0.3$.

versus p^1 (for constant values of p^2) and Figure 2(b) shows these variances versus p^2 (for constant values of p^1).

From these figures it is concluded that the performance of the filter improves as the delay probabilities, p^i, $i = 1, 2$, decrease. Consequently, more accurate estimations are obtained as p^i comes nearer to 0, a case in which all the observations arrive on time.

Case 2 (systems with missing measurements). As in [28], consider missing measurements from two sensors, with different missing characteristics and noise correlation:

$$\tilde{y}_k^i = \theta_k^i z_k + \tilde{v}_k^i, \quad k \ge 1,\ i = 1, 2, \tag{59}$$

where the noise processes $\{\tilde{v}_k^i;\ k \ge 1\}$, $i = 1, 2$, are the same as those in Example 1. Two different independent sequences of random variables $\{\theta_k^i;\ k \ge 1\}$, $i = 1, 2$, with a probability distribution over the interval $[0, 1]$ are used to model the missing phenomenon: $\{\theta_k^1;\ k \ge 1\}$ is a sequence of independent variables with $P[\theta_k^1 = 0] = 0.1$, $P[\theta_k^1 = 0.5] = 0.5$, and $P[\theta_k^1 = 1] = 0.4$, and $\{\theta_k^2;\ k \ge 1\}$ is a sequence of independent Bernoulli variables with $P[\theta_k^2 = 1] = \overline{\theta}$. For all k, the means and variances of these variables are $E[\theta_k^1] = 0.65$, $E[\theta_k^2] = \overline{\theta}$, $\mathrm{Var}[\theta_k^1] = 0.1025$, and $\mathrm{Var}[\theta_k^2] = \overline{\theta}(1 - \overline{\theta})$.

For different values of the missing probability $\overline{\theta}$ and the delay probabilities p^1 and p^2, a comparative analysis, similar to that carried out in Case 1, based on the estimation error variances of the predictor, filter, and smoother was performed. For all values, the results were similar to those given in Figure 1, showing that the fixed-point smoothing error variances are less than the filtering ones which, in

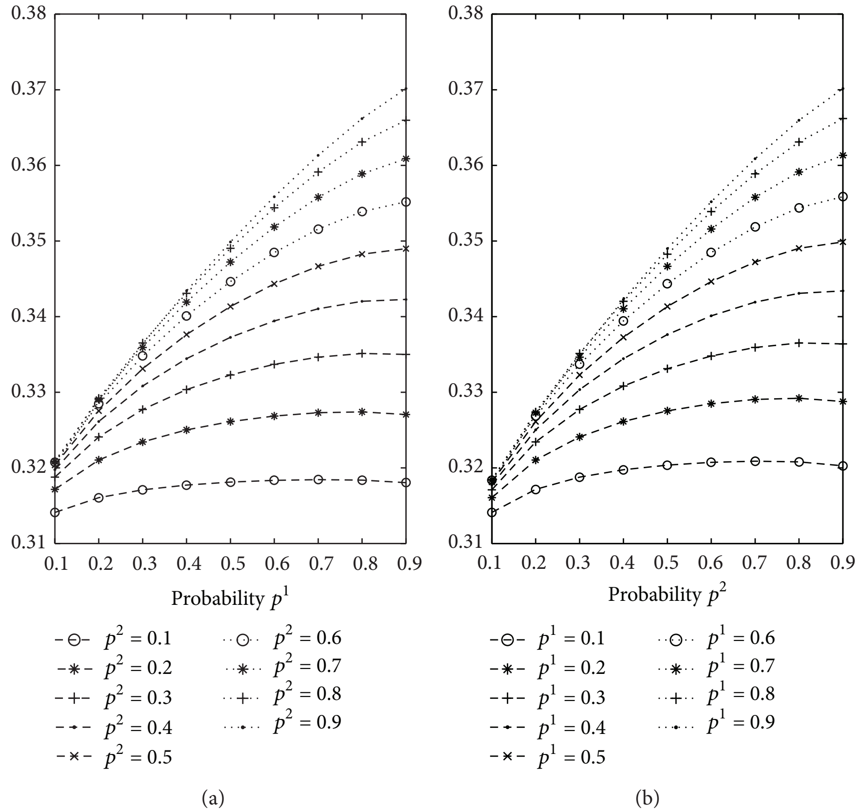


FIGURE 2: (a) Filtering error variances, $\Sigma_{50/50}$, versus p_1 (for constant values of p_2). (b) Filtering error variances $\Sigma_{50/50}$ versus p_2 (for constant values of p_1).

turn, are smaller than the prediction error variances, thus confirming the comments on Figure 1.

Next, considering a fixed value of $\overline{\theta}$, namely, $\overline{\theta} = 0.5$, the filtering error variances have been calculated for different values of the delay probabilities p^1 and p^2. Specifically, the values $p^1 = 0.1, 0.3, 0.4$ and $p^2 = 0.1, 0.3, 0.4, 0.5$ have been used. The results are given in Figure 3 which shows that, as the delay probability p^1 or p^2 increases, the filtering error variances become greater and, consequently, worse estimations are obtained. Also, a similar study to that performed in Figure 2 has been carried out in this case; specifically, for fixed values of $\overline{\theta}$ and fixed delay probability in one of the sensors, the filtering error variances have been analyzed for different delay probabilities in the other sensor. The results are omitted as they are completely analogous to those displayed in Figure 2.

Also, to analyze the performance of the proposed estimators versus the probability $\overline{\theta}$ that the signal is present in the measurements of the second sensor, the filtering error variances have been calculated for $p^1 = 0.1$, $p^2 = 0.3$, and $\overline{\theta}$ varying from $\overline{\theta} = 0.3$ to $\overline{\theta} = 0.8$. The results are displayed in Figure 4; this figure shows that, as $\overline{\theta}$ increases, the filtering error variances become smaller and, hence, better estimations are obtained. Analogous conclusions are deduced for other values of p^1, p^2, and $\overline{\theta}$.

Finally, we present a comparative analysis of the proposed filter and the following filters:

(a) the suboptimal Kalman-type filter [13] for systems with uncorrelated white noises and one-step random delays,

(b) the optimal linear filter based on covariance information [15] for the same class of systems considered in [13],

(c) the centralized Kalman-type filter [26] for systems with correlated and cross-correlated noises,

(d) the optimal centralized filter [28] for systems with missing measurements and correlated and cross-correlated noises.

Considering the values $\overline{\theta} = 0.75$, $p^1 = 0.4$, and $p^2 = 0.5$ and using one thousand independent simulations, the different filtering estimates were compared using the mean square error (MSE) at each time instant k, which is calculated as $\mathrm{MSE}_k = (1/1000)\sum_{s=1}^{1000}(x_k^{(s)} - \widehat{x}_{k/k}^{(s)})^2$, where $\{x_k^{(s)};\ 1 \le k \le 50\}$ denote the sth set of artificially simulated data and $\widehat{x}_{k/k}^{(s)}$ is the filter at the sampling time k in the sth simulation run. The results are displayed in Figure 5, which shows that (a) the proposed filtering algorithm provides better estimations than the other four filtering algorithms; (b) the performance

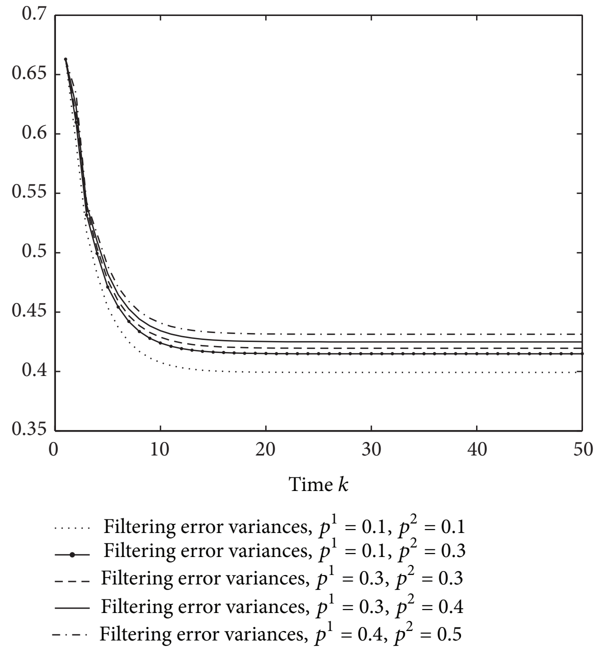


Figure 3: Filtering error variances for different values of p^1 and p^2, when $\overline{\theta} = 0.5$.

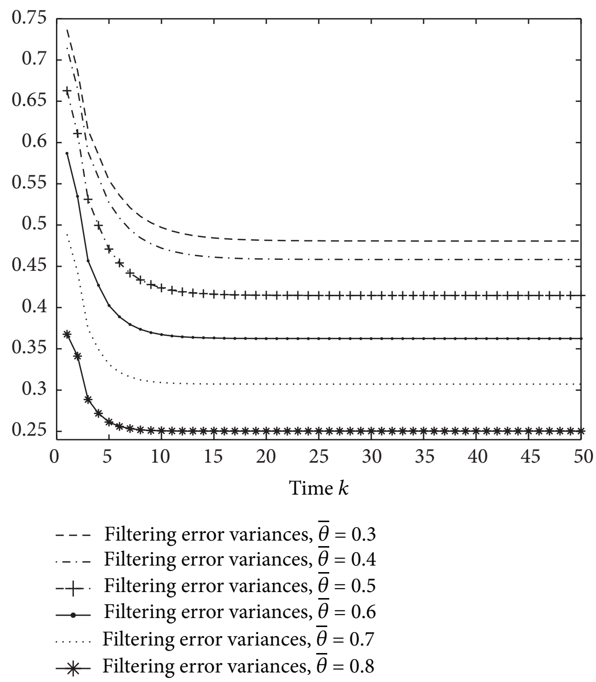


Figure 4: Filtering error variances for $\overline{\theta} = 0.3$ to $\overline{\theta} = 0.8$, when $p^1 = 0.1$ and $p^2 = 0.3$.

of the optimal filter [15] is better than that of the suboptimal filter [13]; (c) the performance of the filters [13, 15] is better than that of the filters [26, 28] since these filters ignore any delay assumption; (d) the filtering algorithm in [26] provides the worst estimations as this filter considers correlated and cross-correlated noises, but neither missing observations nor delayed measurements are taken into account.

7. Conclusions

The optimal least-squares linear estimation problem from randomly delayed measurements has been investigated for discrete-time multisensor linear stochastic systems with both random parameter matrices and correlated noises in the measured outputs. The main contributions are summarized as follows.

(1) The current multisensor observation model considers simultaneously one-step random delayed measurements with different delay rates and both random parameter matrices and correlated noises in the measured outputs. This observation model covers those situations where the sensor noises are one-step autocorrelated and also one-step cross-correlations between different sensor noises are considered. This correlation assumption is valid in a wide spectrum of applications, for example, in target tracking systems where a target is observed by multiple sensors and all of them operate in the same noisy environment. A similar study to that performed in this paper would allow us to generalize the current results by considering more general situations in which the signal and the observation noises are correlated. This extension would cover systems where the sensor and process noises are correlated and would constitute an interesting research topic.

(2) The random delay in each sensor is modelled by a sequence of independent Bernoulli random variables, whose parameters represent the delay probabilities. Another interesting future direction would be to complement the current study considering randomly delayed measurements correlated at consecutive sampling times, thus covering situations where two successive observations cannot be delayed. This kind of delay is usual in situations such as network congestion, random failures in the transmission mechanism, or data inaccessibility at certain times.

(3) Using covariance information, recursive optimal LS linear prediction, filtering, and smoothing algorithms, with a simple computational procedure, are derived by an innovation approach without requiring full knowledge of the state-space model generating the signal process.

(4) The applicability of the proposed algorithms is illustrated by a numerical simulation example, where a scalar state process generated by a first-order autoregressive model is estimated from delayed measurements coming from two sensors, in the following cases: (1) systems with observation multiplicative noises and (2) systems with missing measurements, both with correlated observation noises.

Conflict of Interests

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

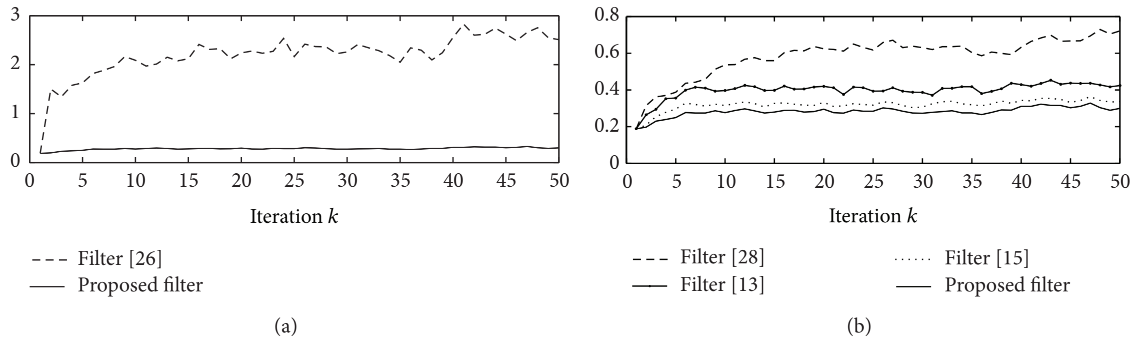


Figure 5: Filtering mean square errors, when $\overline{\theta} = 0.75$, $p^1 = 0.4$, and $p^2 = 0.5$.

Acknowledgment

This research is supported by Ministerio de Educación y Ciencia (Grant no. MTM2011-24718).

References

[1] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "A new estimation algorithm from measurements with multiple-step random delays and packet dropouts," *Mathematical Problems in Engineering*, vol. 2010, Article ID 258065, 18 pages, 2010.

[2] J. Ma and S. Sun, "Information fusion estimators for systems with multiple sensors of different packet dropout rates," *Information Fusion*, vol. 12, no. 3, pp. 213–222, 2011.

[3] S. L. Sun, "Optimal linear estimation for networked systems with one-step random delays and multiple packet dropouts," *Acta Automatica Sinica*, vol. 38, no. 3, pp. 349–356, 2012.

[4] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Recursive least-squares quadratic smoothing from measurements with packet dropouts," *Signal Processing*, vol. 92, no. 4, pp. 931–938, 2012.

[5] S. Sun, "Optimal linear filters for discrete-time systems with randomly delayed and lost measurements with/without time stamps," *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1551–1556, 2013.

[6] S. Sun and W. Xiao, "Optimal linear estimators for systems with multiple random measurement delays and packet dropouts," *International Journal of Systems Science*, vol. 44, no. 2, pp. 358–370, 2013.

[7] Z. Wang, D. W. C. Ho, and X. Liu, "Robust filtering under randomly varying sensor delay with variance constraints," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 51, no. 6, pp. 320–326, 2004.

[8] A. Hermoso-Carazo and J. Linares-Pérez, "Unscented filtering from delayed observations with correlated noises," *Mathematical Problems in Engineering*, vol. 2009, Article ID 681593, 9 pages, 2009.

[9] Y. H. Yang, M. Y. Fu, and H. S. Zhang, "State estimation subject to random communication delays," *Acta Automatica Sinica*, vol. 39, no. 3, pp. 237–243, 2013.

[10] X. Wang, Y. Liang, Q. Pan, C. Zhao, and F. Yang, "Design and implementation of Gaussian filter for nonlinear system with randomly delayed measurements and correlated noises," *Applied Mathematics and Computation*, vol. 232, pp. 1011–1024, 2014.

[11] A. Hermoso-Carazo and J. Linares-Pérez, "Linear and quadratic least-squares estimation using measurements with correlated one-step random delay," *Digital Signal Processing*, vol. 18, no. 3, pp. 450–464, 2008.

[12] R. Caballero-Águila, A. Hermoso-Carazo, J. D. Jiménez-López, J. Linares-Pérez, and S. Nakamori, "Recursive estimation of discrete-time signals from nonlinear randomly delayed observations," *Computers & Mathematics with Applications*, vol. 58, no. 6, pp. 1160–1168, 2009.

[13] F. O. Hounkpevi and E. E. Yaz, "Minimum variance generalized state estimators for multiple sensors with different delay rates," *Signal Processing*, vol. 87, no. 4, pp. 602–613, 2007.

[14] J. Wen, L. Pen, and S. K. Nguang, "Asynchronous H∞ filtering of switched time-delay systems with network induced random occurrences," *Signal Processing*, vol. 98, pp. 62–73, 2014.

[15] R. Caballero-Águila, A. Hermoso-Carazo, J. D. Jiménez-López, J. Linares-Pérez, and S. Nakamori, "Signal estimation with multiple delayed sensors using covariance information," *Digital Signal Processing*, vol. 20, no. 2, pp. 528–540, 2010.

[16] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Linear estimation based on covariances for networked systems featuring sensor correlated random delays," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1233–1244, 2013.

[17] M. J. García-Ligero, A. Hermoso-Carazo, J. Linares-Pérez, and S. Nakamori, "Derivation of linear estimation algorithms from measurements affected by multiplicative and additive noises," *Journal of Computational and Applied Mathematics*, vol. 234, no. 3, pp. 794–804, 2010.

[18] F. Peng and S. Sun, "Distributed fusion estimation for multisensor multirate systems with stochastic observation multiplicative noises," *Mathematical Problems in Engineering*, vol. 2014, Article ID 373270, 8 pages, 2014.

[19] G. Wei, Z. Wang, and H. Shu, "Robust filtering with stochastic nonlinearities and multiple missing measurements," *Automatica*, vol. 45, no. 3, pp. 836–841, 2009.

[20] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, "Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements," *Automatica*, vol. 48, no. 9, pp. 2007–2015, 2012.

[21] R. Caballero-Águila, I. García-Garrido, and J. Linares-Pérez, "Information fusion algorithms for state estimation in multisensor systems with correlated missing measurements," *Applied Mathematics and Computation*, vol. 226, pp. 548–563, 2014.

[22] W. L. De Koning, "Optimal estimation of linear discrete-time systems with stochastic parameters," *Automatica*, vol. 20, no. 1, pp. 113–115, 1984.

[23] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song, and D. Wang, "Globally optimal multisensor distributed random parameter matrices kalman filtering fusion with applications," *Sensors*, vol. 8, no. 12, pp. 8086–8103, 2008.

[24] X. Shen, Y. Zhu, and Y. Luo, "Optimal state estimation of linear discrete-time systems with correlated random parameter matrices," in *Proceedings of the 30th Chinese Control Conference (CCC '11)*, pp. 1488–1493, July 2011.

[25] J. Feng and M. Zeng, "Optimal distributed Kalman filtering fusion for a linear dynamic system with cross-correlated noises," *International Journal of Systems Science*, vol. 43, no. 2, pp. 385–398, 2012.

[26] J. Feng, Z. Wang, and M. Zeng, "Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises," *Information Fusion*, vol. 14, no. 1, pp. 78–86, 2013.

[27] L. Yan, X. Rong Li, Y. Xia, and M. Fu, "Optimal sequential and distributed fusion for state estimation in cross-correlated noise," *Automatica*, vol. 49, no. 12, pp. 3607–3612, 2013.

[28] R. Caballero-Águila, I. García-Garrido, and J. Linares-Pérez, "Optimal fusion filtering in multisensor stochastic systems with missing measurements and correlated noises," *Mathematical Problems in Engineering*, vol. 2013, Article ID 418678, 14 pages, 2013.

[29] P. Cui, H. Zhang, J. Lam, and L. Ma, "Real-time Kalman filtering based on distributed measurements," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 14, pp. 1597–1608, 2013.

[30] J. Hu, Z. Wang, and H. Gao, "Recursive filtering with random parameter matrices, multiple fading measurements and correlated noises," *Automatica*, vol. 49, no. 11, pp. 3440–3448, 2013.

[31] J. Linares-Pérez, R. Caballero-Águila, and I. García-Garrido, "Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications," *International Journal of Systems Science*, 2014.

[32] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 2000.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 516291, 6 pages
http://dx.doi.org/10.1155/2014/516291

Research Article

An Upper Bound of Large Deviations for Capacities

Xiaomin Cao

School of Mathematics, Shandong University, Jinan 250000, China

Correspondence should be addressed to Xiaomin Cao; caoxiaomin319320@163.com

Received 23 April 2014; Accepted 22 May 2014; Published 15 June 2014

Academic Editor: Xuejun Xie

Up to now, most of the academic researches about the large deviation and risk theory are under the framework of the classical linear expectations. But motivated by problems of model uncertainties in statistics, measures of risk, and superhedging in finance, sublinear expectations are extensively studied. In this paper, we obtain a type of large deviation principle under the sublinear expectation. This result is a new expression of the Gärtner-Ellis theorem under the sublinear expectations which is in the classical theory of large deviations. In addition, we introduce a new process under the sublinear expectations, that is, the G-Poisson process. We give an application of our result and obtain the rate function of the compound G-Poisson process in the upper bound of large deviations for capacities. The application of our result opens a new field for the research of risk theory in the future.

1. Introduction

Large deviation theory is one of the key techniques of modern probability, a role which is emphasized by the recent award of the Abel prize to S.R.S. Varadhan, one of the pioneers of the subject. The large deviation principle characterizes the limiting behavior as $\varepsilon \to \infty$ of a family of probability measures μ_ε in terms of a rate function. Also Cramér's theorem has been widely known for a long time as a fundamental result in large deviations. It is very useful in many fields. But Cramér's theorem is limited to the i.i.d. case. However, a glance at the proof should be enough to convince the reader that some extension to the non-i.i.d. case is possible. As described in [1], Gärtner-Ellis theorem is a generalization of Cramér's theorem in non-i.i.d situation to conclusions.

Motivated by problems of model uncertainties in statistics, measures of risk, and superhedging in finance, sublinear expectations are extensively studied [2]. Since the paper [3] on coherent risk measures, authors have been more and more interested in sublinear expectations [4, 5]. By Peng [6], we know that a sublinear expectation $\widehat{E}$ can be represented as the upper expectation of a set of linear expectations $\{E_\theta : \theta \in \Theta\}$; that is, $\widehat{E}[\cdot] = \sup_{\theta\in\Theta} E_\theta[\cdot]$. In most cases, this set is often treated as an uncertain model of probabilities $\{P_\theta : \theta \in \Theta\}$ and the notion of sublinear expectation provides a robust way to measure a risk loss X. In fact, nonlinear expectation theory provides many rich, flexible, and elegant tools and plays an important role in many aspects. In particular, its important application in stochastic dominance, stochastic differential game, financial mathematics, economics, and partial differential equations attracted a large number of mathematicians, economists, and financial experts to join the research, for instance, the application of nonlinear expectation in the dynamic measurement and dominance of financial risk, backward stochastic differential equation theory and its application in financial products innovation, pricing, and so forth. We can see its recent developments from the following literature [7–12].

In this paper, we are interested in

$$\overline{E}[\cdot] = \sup_{P\in\mathscr{P}} E_P[\cdot], \tag{1}$$

where $\mathscr{P}$ is a set of probability measures, especially set $\overline{V}(A) = \overline{E}[I_A] = \sup_{P\in\mathscr{P}} E_P[I_A]$, $\forall A \in \mathscr{F}$. Obviously, $\overline{V}$ is a capacity. Under the sublinear expectation, the upper bound of Cramér's theorem has come to a conclusion similar to the linear expectation (see [13]). On this basis, additionally, the main aim of this paper is to obtain Gärtner-Ellis's upper bound for the capacity $\overline{V}$.

This paper is organized as follows. In Section 2, we give some notions and lemmas that are useful in this paper. In Section 3, we give the main result including the proof. In

Section 4, we give a brief application of our result in the classical risk model.

2. Preliminaries

We present some preliminaries in the theory of sublinear expectations. More details of this section can be found in Peng [6, 14, 15].

Definition 1. Let Ω be a given set and let $\mathscr{H}$ be a linear space of real valued functions defined on Ω. We assume that all constants are in $\mathscr{H}$ and that $X \in \mathscr{H}$ implies $|X| \in \mathscr{H}$. $\mathscr{H}$ is considered as the space of our "random variables." A nonlinear expectation $\widehat{E}$ on $\mathscr{H}$ is a functional $\widehat{E} : \mathscr{H} \mapsto \mathbb{R}$ satisfying the following properties: for all $X, Y \in \mathscr{H}$, one has

(a) monotonicity: if $X \geqslant Y$, then $\widehat{E}[X] \geqslant \widehat{E}[Y]$;

(b) constant preserving: $\widehat{E}[c] = c$, $c \in \mathbb{R}$.

The triple $(\Omega, \mathscr{H}, \widehat{E})$ is called a nonlinear expectation space (compare with a probability space $(\Omega, \mathscr{H}, P)$). We are mainly concerned with sublinear expectation where the expectation $\widehat{E}$ satisfies also

(c) subadditivity: $\widehat{E}[X] - \widehat{E}[Y] \leqslant \widehat{E}[X - Y]$;

(d) positive homogeneity: $\widehat{E}[\lambda X] = \lambda \widehat{E}[X]$, $\forall \lambda \geqslant 0$.

If only (c) and (d) are satisfied, $\widehat{E}$ is called a sublinear functional.

The following representation theorem for sublinear expectations is very useful (see Peng [6, 15] for the proof).

Lemma 2. *Let $\widehat{E}$ be a sublinear functional defined on $(\Omega, \mathscr{H})$; that is, (c) and (d) hold for $\widehat{E}$. Then there exists a family $\{E_\theta : \theta \in \Theta\}$ of linear functionals on $(\Omega, \mathscr{H})$ such that*

$$\widehat{E}[X] = \max_{\theta \in \Theta} E_\theta [X]. \tag{2}$$

If (a) and (b) also hold, then E_θ are linear expectations for $\theta \in \Theta$. If we make, furthermore, the following assumption.

(H1) *For each sequence $\{X_n\}_{n=1}^{\infty} \subset \mathscr{H}$ such that $X_n(\omega) \downarrow 0$ for ω, one has $\widehat{E}[X_n] \downarrow 0$.*

Then for each $\theta \in \Theta$, there exists a unique (σ-additive) probability measure P_θ defined on $(\Omega, \sigma(\mathscr{H}))$ such that

$$E_\theta [X] = \int_\Omega X(\omega)\, dP_\theta(\omega), \quad X \in \mathscr{H}. \tag{3}$$

In this paper, we are interested in the following sublinear expectation:

$$\overline{E}[\cdot] = \sup_{P \in \mathscr{P}} E_P [\cdot], \tag{4}$$

where $\mathscr{P}$ is a set of probability measures. Let Ω be a given set and let $\mathscr{F}$ be a σ-algebra. Define $\overline{V}(A) := \overline{E}[I_A] = \sup_{P \in \mathscr{P}} E_P[I_A]$, $\forall A \in \mathscr{F}$; then $\overline{V}$ is a capacity.

Let $C(\mathbb{R}^n)$ denote the space of continuous functions defined on $\mathbb{R}^n$.

Now we recall some important notions of sublinear expectations distributions (see Peng [6, 14, 15]).

Definition 3. Let X_1 and X_2 be two random variables in a sublinear expectation space $(\Omega, \mathscr{F}, \overline{E})$. They are called identically distributed, denoted by $X_1 \sim X_2$, if for $\varphi \in C(\mathbb{R})$, $\overline{E}[\varphi(X_1)]$ and $\overline{E}[\varphi(X_2)]$ exist; then one has

$$\overline{E}[\varphi(X_1)] = \overline{E}[\varphi(X_2)]. \tag{5}$$

Definition 4. In a sublinear expectation space $(\Omega, \mathscr{F}, \overline{E})$, a random vector $Y = (Y_1, \ldots, Y_n)$ is said to be independent of another random vector $X = (X_1, \ldots, X_m)$, if for $\varphi \in C(\mathbb{R}^{m+n})$, $\overline{E}[\varphi(X, Y)]$ and $\overline{E}[\overline{E}[\varphi(x, Y)]_{x=X}]$ exist; then one has

$$\overline{E}[\varphi(X, Y)] = \overline{E}\left[\overline{E}[\varphi(x, Y)]_{x=X}\right]. \tag{6}$$

We conclude this section with some notations on large deviations under a sublinear expectation [16].

Let S be a topology space and $\mathscr{S}$ be a σ-algebra on S. Let $(U_n, n \geq 1)$ be a family of measurable maps from Ω into S and $b(n)$, $n \geq 1$ be a positive function satisfying $b(n) \to \infty$ as $n \to \infty$. A nonnegative function I on S is called a (good) rate function if $\{I \leq l\}$ is (compact) closed for all $0 \leq l < \infty$.

$(\overline{V}(U_n \in \cdot), n \geq 1)$ is said to satisfy large deviation principle (LDP) with speed $b(n)$ and with rate function $I(x)$ if for any measurable closed subset $F \subset S$,

$$\limsup_{n \to \infty} \frac{1}{b(n)} \log \overline{V}(U_n \in F) \leq -\inf_{x \in F} I(x) \tag{7}$$

and for any measurable open subset $O \subset S$,

$$\liminf_{n \to \infty} \frac{1}{b(n)} \log \overline{V}(U_n \in O) \geq -\inf_{x \in O} I(x). \tag{8}$$

Equations (7) and (8) are referred, respectively, to as upper bound of large deviations (ULD) and lower bound of large deviations (LLD).

$(\overline{V}(U_n \in \cdot), n \geq 1)$ is said to be exponentially tight if for any $L > 0$, there exists a compact set $K_L \subset S$ such that

$$\limsup_{n \to \infty} \frac{1}{b(n)} \log \overline{V}(U_n \in K_L^c) \leq -L. \tag{9}$$

$(\overline{V}(U_n \in \cdot), n \geq 1)$ is said to satisfy w-upper bound of large deviations with speed $b(n)$ and with rate function $I(x)$ if (7) for any compact subset $F \subset S$.

It is known that if $(\overline{V}(U_n \in \cdot), n \geq 1)$ satisfies w-large deviation principle with speed $b(n)$ and with rate function I and is exponentially tight, then it satisfies large deviation principle with speed $b(n)$ and with rate function I.

Definition 5. For any rate function I and any $\delta > 0$, the δ-rate function is defined as

$$I^\delta(x) \triangleq \min\left\{I(x) - \delta, \frac{1}{\delta}\right\}. \tag{10}$$

While in general I^δ is not a rate function, its usefulness stems from the fact that for any set Γ,

$$\lim_{\delta\to 0}\inf_{x\in\Gamma} I^\delta(x) = \inf_{x\in\Gamma} I(x). \tag{11}$$

Consequently, the upper bound in (7) is equivalent to the statement that for any $\delta > 0$ and for any measurable set Γ,

$$\limsup_{n\to\infty}\frac{1}{b(n)}\log\overline{V}(U_n\in\Gamma) \le -\inf_{x\in\overline{\Gamma}} I^\delta(x). \tag{12}$$

3. Main Result

In this section, firstly let us present some notations and assumptions that are used in the following details.

Consider a sequence of random vectors $X_n \in \mathbb{R}^d$; let $\{X_n; n \ge 1\}$ be identically distributed under $\overline{E}[\cdot]$, where X_n possesses logarithmic moment generating function $\overline{\Lambda}_n(\lambda) := \log\overline{E}[e^{\langle\lambda,X_n\rangle}]$, $\forall\lambda \in \mathbb{R}^d$. We also assume that each X_{n+1} is independent of $(X_1,\dots,X_n)$ for $n = 1,2,\dots$ under $\overline{E}[\cdot]$. Denote $\overline{S}_n = (1/n)\sum_{i=1}^n X_i$.

Specifically, the following assumption is imposed throughout this section.

Assumption 6. For each $\lambda \in \mathbb{R}^d$, the logarithmic moment generating function, defined as the limit

$$\overline{\Lambda}(\lambda) := \lim_{n\to\infty}\frac{1}{n}\overline{\Lambda}_n(n\lambda) \tag{13}$$

exists as an extended real number. Furthermore, the origin belongs to the interior of $\mathcal{D}_{\overline{\Lambda}} := \{\lambda : \overline{\Lambda}(\lambda) < \infty\}$.

Define

$$\begin{gathered}\underline{x} := -\overline{E}[-X];\\ \overline{\Lambda^*}(x) := \sup_{\lambda\in R}\left[\lambda x - \overline{\Lambda}(\lambda)\right], \quad \forall x\in\mathbb{R}^d,\end{gathered} \tag{14}$$

where $\overline{\Lambda^*}(\cdot)$ is the Fenchel-Legendre transform of $\overline{\Lambda}(\cdot)$, with $\mathcal{D}_{\overline{\Lambda^*}} := \{\lambda : \overline{\Lambda^*}(\lambda) < \infty\}$.

We always assume that

(H2) if $A_n \uparrow \Omega$, then $\overline{V}(A_n)\uparrow 1$.

Lemma 7. *Let N be a fixed integer. Then, for every $a_\varepsilon^i \ge 0$,*

$$\limsup_{\varepsilon\to 0}\varepsilon\log\left(\sum_{i=1}^N a_\varepsilon^i\right) = \max_{i=1}^N\limsup_{\varepsilon\to 0}\varepsilon\log a_\varepsilon^i. \tag{15}$$

Proof. First note that for all ε,

$$0 \le \varepsilon\log\left(\sum_{i=1}^N a_\varepsilon^i\right) - \max_{i=1}^N\varepsilon\log a_\varepsilon^i \le \varepsilon\log N. \tag{16}$$

Since N is fixed, $\varepsilon\log N \to 0$ as $\varepsilon \to 0$ and

$$\limsup_{\varepsilon\to 0}\max_{i=1}^N\varepsilon\log a_\varepsilon^i = \max_{i=1}^N\limsup_{\varepsilon\to 0}\varepsilon\log a_\varepsilon^i. \tag{17}$$

□

The next lemma describes a property of $\overline{\Lambda^*}(\cdot)$, which will be used to give a more accurate expression of the rate function.

Lemma 8. *If $\log\overline{E}e^{\lambda X_1} < +\infty$ for some $\lambda\in\mathbb{R}^+$, then for any $x \ge \underline{x}$, $\overline{\Lambda^*}(x) = \sup_{\lambda\ge 0}\{\lambda x - \log\overline{E}e^{\lambda X_1}\}$ and for any $x \le \underline{x}$, $\overline{\Lambda^*}(x) = \sup_{\lambda\le 0}\{\lambda x - \log\overline{E}e^{\lambda X_1}\}$.*

Here, we omit the proof of Lemma 8 (refer to [13] or [17]).

The following theorem is the main result of this paper.

Theorem 9. *Let Assumption 6 hold. Then we have for any closed set $F\subset\mathbb{R}^d$,*

$$\limsup_{n\to\infty}\frac{1}{n}\log\overline{V}\left(\overline{S}_n\in F\right) \le -\inf_{x\in F}\overline{\Lambda^*}(x), \tag{18}$$

where $\overline{\Lambda^*}(\cdot)$ is a convex rate function.

Proof. As mentioned in Section 2, establishing the upper bound is equivalent to proving that for every $\delta > 0$ and every closed set $F\subset\mathbb{R}^d$,

$$\limsup_{n\to\infty}\frac{1}{n}\log\overline{V}\left(\overline{S}_n\in F\right) \le \delta - \inf_{x\in F} I^\delta(x), \tag{19}$$

where I^δ is the δ-rate function associated with $\overline{\Lambda^*}$. Fix a compact set $\Gamma\subset\mathbb{R}^d$. For every $q\in\Gamma$, choose $\lambda_q\in\mathbb{R}^d$ such that

$$\left\langle\lambda_q, q\right\rangle - \overline{\Lambda}\left(\lambda_q\right) \ge I^\delta(q). \tag{20}$$

This is feasible on account of the definitions of $\overline{\Lambda^*}$ and I^δ. For each q, choose $\rho_q > 0$ such that $\rho_q|\lambda_q| \le \delta$ and let $B_{q,\rho_q} = \{x : |x-q| < \rho_q\}$ be the ball with center at q and radius ρ_q. Observe for every n, $\lambda\in\mathbb{R}^d$, and measurable $G\subset\mathbb{R}^d$ that

$$\begin{aligned}\overline{V}\left(\overline{S}_n\in G\right) &= \overline{E}\left[1_{\overline{S}_n\in G}\right]\\ &\le \overline{E}\left[\exp\left(n\left\langle\lambda,\overline{S}_n\right\rangle - \inf_{x\in G}\{n\langle\lambda,x\rangle\}\right)\right].\end{aligned} \tag{21}$$

In particular, for each n and $q\in\Gamma$,

$$\begin{aligned}&\overline{V}\left(\overline{S}_n\in B_{q,\rho_q}\right)\\ &\quad\le \overline{E}\left[\exp\left(n\left\langle\lambda_q,\overline{S}_n\right\rangle\right)\right]\exp\left(-\inf_{x\in B_{q,\rho_q}}\left\{n\left\langle\lambda_q,x\right\rangle\right\}\right).\end{aligned} \tag{22}$$

Also, for any $q\in\Gamma$,

$$-\inf_{x\in B_{q,\rho_q}}\left\langle\lambda_q,x\right\rangle \le \rho_q\left|\lambda_q\right| - \left\langle\lambda_q,q\right\rangle \le \delta - \left\langle\lambda_q,q\right\rangle, \tag{23}$$

and therefore,

$$\begin{aligned}\frac{1}{n}\log\overline{V}\left(\overline{S}_n\in B_{q,\rho_q}\right) &\le -\inf_{x\in B_{q,\rho_q}}\left\langle\lambda_q,x\right\rangle + \overline{\Lambda}\left(\lambda_q\right)\\ &\le \delta - \left\langle\lambda_q,q\right\rangle + \overline{\Lambda}\left(\lambda_q\right).\end{aligned} \tag{24}$$

Since Γ is compact, one may extract from the open covering $\bigcup_{q\in\Gamma} B_{q,\rho_q}$ of Γ a finite covering that consists of $N = N(\Gamma,\delta) < \infty$ such balls with centers $q_1,\ldots,q_N$ in Γ. By the union of events bound and the preceding inequality,

$$\frac{1}{n}\log\overline{V}\left(\overline{S}_n\in\Gamma\right)\leq\frac{1}{n}\log N+\delta - \min_{i=1,\ldots,N}\left\{\left\langle\lambda_{q_i},q_i\right\rangle-\overline{\Lambda}\left(\lambda_{q_i}\right)\right\}. \tag{25}$$

Hence, by our choice of λ_q,

$$\limsup_{n\to\infty}\frac{1}{n}\log\overline{V}\left(\overline{S}_n\in\Gamma\right)\leq\delta-\min_{i=1,\ldots,N}I^{\delta}\left(q_i\right). \tag{26}$$

Since $q_i \in \Gamma$, the upper bound (7) is established for all compact sets.

As described earlier in Section 2, the upper bound of large deviations is extended to all closed subsets of $\mathbb{R}^d$ by showing that $\overline{V}(\overline{S}_n\in\cdot)$ is an exponentially tight family of probability measures. Let $H_\rho \triangleq [-\rho,\rho]^d$. Since $H_\rho^c=\bigcup_{j=1}^d\{x:|x^j|>\rho\}$, the union of events bound yields

$$\overline{V}\left(\overline{S}_n\in H_\rho^c\right)\leq\sum_{j=1}^d\overline{V}\left(\overline{S}_n^j\in[\rho,\infty)\right)+\sum_{j=1}^d\overline{V}\left(\overline{S}_n^j\in(-\infty,-\rho]\right), \tag{27}$$

where $\overline{S}_n^j$, $j=1,\ldots,d$ are the coordinates of the random vector $\overline{S}_n$; namely, $(\overline{V}(\overline{S}_n^j\in\cdot))$ are the laws governing $(1/n)\sum_{i=1}^n X_i^j$. Let $\mathbf{e}_j$ denote the jth unit vector in $\mathbb{R}^d$ for $j=1,\ldots,d$. Since $0\in\mathscr{D}_{\overline{\Lambda}}^0$ (refer to [13, Lemma 3.1]), there exist $\theta_j>0$ and $\eta_j>0$ such that $\overline{\Lambda}(\theta_j\mathbf{e}_j)<\infty$ and $\overline{\Lambda}(-\eta_j\mathbf{e}_j)<\infty$ for $j=1,\ldots,d$. By Chebyshev's inequality, we have

$$\begin{gathered}\overline{V}\left(\overline{S}_n^j\in(-\infty,-\rho]\right)\leq\exp\left(-n\eta_j\rho+\overline{\Lambda}_n\left(-n\eta_j\mathbf{e}_j\right)\right),\\ \overline{V}\left(\overline{S}_n^j\in[\rho,\infty)\right)\leq\exp\left(-n\theta_j\rho+\overline{\Lambda}_n\left(n\theta_j\mathbf{e}_j\right)\right),\\ j=1,\ldots,d.\end{gathered} \tag{28}$$

Hence, for $j=1,\ldots,d$,

$$\begin{gathered}\lim_{\rho\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\overline{V}\left(\overline{S}_n^j\in(-\infty,-\rho]\right)=-\infty,\\ \lim_{\rho\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\overline{V}\left(\overline{S}_n^j\in[\rho,\infty)\right)=-\infty.\end{gathered} \tag{29}$$

Consequently, by the union of events bound and Lemma 7,

$$\lim_{\rho\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\overline{V}\left(\overline{S}_n\in H_\rho^c\right)=-\infty. \tag{30}$$

Therefore, $\overline{V}(\overline{S}_n\in\cdot)$ is an exponentially tight family of probability measures, since the hypercubes H_ρ are compact. □

Remark 10. Since $\overline{E}$ is not linear, Cramér's method is not useful for lower bound of large deviations. This is consistent with the conclusion of [13]. In the paper [13], the author gives a counter example to illustrate that under the sublinear expectation, the lower bound of Cramér's theorem is not obtained. Since Gärtner-Ellis theorem be a generalization of Cramér's theorem in non-i.i.d situation to conclusions, we see under the assumptions of theorem, the lower bound of Gärtner-Ellis theorem does not hold.

4. Application

In this section, we consider the classical risk process under sublinear expectation $\overline{E}[\cdot]$. The classical risk process $(R^x(t))_{t\geq0}$ is defined by

$$R^x(t)=x+pt-S(t),\quad t\geq0, \tag{31}$$

where $x>0$ is the initial capital and $p>0$ is the (constant) premium rate, and the aggregate claims process $(S(t))_{t\geq0}$ is a compound Poisson process. More precisely we have $S(t)=\sum_{i=1}^{N(t)}X_i$ and $N(t)=\sum_{i\geq1}1_{\{T_i\leq t\}}$, where $\{X_i\}_{i\geq1}$ is a sequence of positive random variables, $(N(t))_{t\geq0}$ is a counting process with points $\{T_i\}_{i\geq1}$, $\{X_i\}_{i\geq1}$, and $(N(t))_{t\geq0}$ independent, the X_i, $i=1,2,\ldots$, are independent and identically distributed, and where $(N(t))_{t\geq0}$ is a Poisson process with intensity μ under linear expectation. Now we consider $(N(t))_{t\geq0}$ is a G-Poisson process (its definition refers to [18]) under sublinear expectation $\overline{E}[\cdot]$. Then, $(S(t))_{t\geq0}$ is a compound G-Poisson process correspondingly.

We also assume the following superexponential condition holds for the random variables $\{X_i\}_{i\geq1}$ under sublinear expectation $\overline{E}[\cdot]$.

Assumption 11. $\overline{E}[e^{\lambda X_1}]<\infty$ for all $\lambda\in\mathbb{R}$.

Let $\varphi_{N(t)}(\lambda)$ be the moment generating function (m.g.f.) of $(N(t))_{t\geq0}$; that is,

$$\varphi_{N(t)}(\lambda)=\overline{E}e^{\lambda N(t)},\quad t\geq0,\ \forall\lambda\in\mathbb{R}. \tag{32}$$

Then its logarithmic moment generating function is expressed as follows:

$$\Lambda_{N(t)}(\lambda)\triangleq\log\overline{E}e^{\lambda N(t)}=\log\varphi_{N(t)}(\lambda). \tag{33}$$

Let $\psi_{N(t)}(\lambda)$ be the limit of the normalized logarithmic moment generating function of $(N(t))_{t\geq0}$; that is,

$$\psi_{N(t)}(\lambda)=\lim_{t\to\infty}\frac{1}{t}\log\overline{E}e^{\lambda N(t)}. \tag{34}$$

In order to obtain the m.g.f. of the process $(S(t))_{t\geq0}$, firstly we introduce a lemma which plays a role in the next lemma. We omit its proof which can be found in [19, Lemma 1.1].

Lemma 12. *If a sequence of d-dimensions random variables $\{X_i,i=1,\ldots,m\}$ under sublinear expectation space $(\Omega,\mathscr{H},\overline{E})$ satisfies for any $i=1,\ldots,m-1$, X_{i+1} is independent of $(X_1,\ldots,X_i)$, then the following conclusions are established.*

(1) *If* $\varphi_1, \ldots, \varphi_m$ *are lower semicontinuous functions in* $\mathbb{R}^d \to [0, +\infty)$*, one has*

$$\overline{E}(\varphi_1(X_1)\cdots\varphi_m(X_m)) = \overline{E}(\varphi_1(X_1))\cdots\overline{E}(\varphi_m(X_m)). \tag{35}$$

(2) *If* $\psi_1, \ldots, \psi_m$ *are upper semicontinuous functions in* $\mathbb{R}^d \to [0, +\infty)$ *and there exists a continuous function* Ψ *such that* $\Psi(X_i) \in \mathbb{L}_c^1 \triangleq \{X \in \mathbb{L}^1; X$ *is quasicontinuous and* $\lim_{n\to\infty}\overline{E}[|X|I_{|X|>n}] = 0\}$*, and* $\psi_i \le \Psi$*, for any* $i = 1, \ldots, m$*, one has*

$$\overline{E}(\psi_1(X_1)\cdots\psi_m(X_m)) = \overline{E}(\psi_1(X_1))\cdots\overline{E}(\psi_m(X_m)). \tag{36}$$

Lemma 13. *If* $(N(t))_{t\ge0}$ *and* $\{X_i\}_{i\ge1}$ *are independent, then for each* $\lambda \in \mathbb{R}$

$$\varphi_{S(t)}(\lambda) = \varphi_{N(t)}(\Lambda_{X_1}(\lambda)), \quad t \ge 0. \tag{37}$$

Proof. By Lemma 12, we have

$$\begin{aligned}\varphi_{S(t)}(\lambda) &= \overline{E}\left[e^{\lambda S(t)}\right] = \overline{E}\left[e^{\lambda\sum_{i=1}^{N(t)}X_i}\right] \\ &= \overline{E}\left[\overline{E}\left[e^{\lambda\sum_{i=1}^{k}X_i}\right]_{k=N(t)}\right] = \overline{E}\left[\left[\overline{E}\left[e^{\lambda X_1}\right]\right]^k_{k=N(t)}\right] \\ &= \overline{E}\left[\left(\overline{E}e^{\lambda X_1}\right)^{N(t)}\right] = \overline{E}\left[e^{N(t)\log\overline{E}[e^{\lambda X_1}]}\right] \\ &= \varphi_{N(t)}(\Lambda_{X_1}(\lambda)), \quad t \ge 0.\end{aligned} \tag{38}$$

This completes the proof. □

Then we can see

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t}\log\overline{E}\left[e^{\lambda S(t)}\right] &= \lim_{t\to\infty}\frac{1}{t}\log\varphi_{N(t)}(\Lambda_{X_1}(\lambda)) \\ &= \psi_{N(t)}(\Lambda_{X_1}(\lambda)).\end{aligned} \tag{39}$$

That is to say, the normalized logarithmic moment generating function of $(S(t))_{t\ge0}$ has a limit. By Theorem 9 that we obtained in Section 3, we can say $(S(t))_{t\ge0}$ satisfies the upper bound of Gärtner-Ellis theorem with rate function I defined by

$$I(x) = \sup_{\lambda\in\mathbb{R}}\left\{\lambda x - \psi_{N(t)}(\Lambda_{X_1}(\lambda))\right\}, \quad \forall x \in \mathbb{R}. \tag{40}$$

Next, we give you a brief description about G-Poisson process in Ren's Ph.D. thesis [17]. Let $(N(t))_{t\ge0}$ be G-Poisson process under sublinear expectation $\overline{E}[\cdot]$. Then, $\overline{E}[\varphi(x + N(t))]$ satisfies the following one-dimensional equation:

$$\begin{aligned}\partial_t u(t,x) - G(u(t,x+1) - u(t,x)) &= 0, \\ u(0,x) = \varphi(x), \quad 0 \le t &\le 1,\end{aligned} \tag{41}$$

where $G(a) = \mu_2 a^+ - \mu_1 a^-$, $0 \le \mu_1 \le \mu_2$. Referring to [18], we know, for any increasing function ϕ,

$$\overline{E}[\phi(x + N(t))] = \sum_{i=0}^{\infty}\phi(x+i)\frac{(\mu_2 t)^i}{i!}e^{-\mu_2 t}, \tag{42}$$

and for any decreasing function ϕ,

$$\overline{E}[\phi(x + N(t))] = \sum_{i=0}^{\infty}\phi(x+i)\frac{(\mu_1 t)^i}{i!}e^{-\mu_1 t}. \tag{43}$$

Then we have $\log\overline{E}e^{\lambda N(t)} = \mu_2 t(e^\lambda - 1)$, for any $\lambda \ge 0$, and $\log\overline{E}e^{\lambda N(t)} = \mu_1 t(e^\lambda - 1)$, for any $\lambda \le 0$.

Since X_i represents the amount claimed, we know $X_i \ge 0$ and $i \ge 1$; and so $\Lambda_{X_1}(\lambda) = \log\overline{E}[e^{\lambda X_1}] \ge 0$. In the above formula (42), we set $x = 0$,

$$\begin{aligned}\varphi_{N(t)}(\Lambda_{X_1}(\lambda)) &= \overline{E}\left[e^{N(t)\Lambda_{X_1}(\lambda)}\right] \\ &= \sum_{i=0}^{\infty}e^{i\Lambda_{X_1}(\lambda)}\frac{(\mu_2 t)^i}{i!}e^{-\mu_2 t} \\ &= \sum_{i=0}^{\infty}\frac{\left(\mu_2 t e^{\Lambda_{X_1}(\lambda)}\right)^i}{i!}e^{-\mu_2 t} \\ &= e^{\mu_2 t(e^{\Lambda_{X_1}(\lambda)}-1)}.\end{aligned} \tag{44}$$

Substituting this result into (39), we have

$$\lim_{t\to\infty}\frac{1}{t}\log\overline{E}\left[e^{\lambda S(t)}\right] = \mu_2\left(e^{\Lambda_{X_1}(\lambda)} - 1\right). \tag{45}$$

By (40) and Lemma 8, we get the rate function of $(S(t))_{t\ge0}$ in the upper bound of Gärtner-Ellis theorem. As described below, for $x \ge \underline{x}$,

$$\begin{aligned}I(x) &= \sup_{\lambda\in\mathbb{R}}\left\{\lambda x - \psi_{N(t)}(\Lambda_{X_1}(\lambda))\right\} \\ &= \sup_{\lambda\ge0}\left\{\lambda x - \mu_2\left(e^{\Lambda_{X_1}(\lambda)} - 1\right)\right\}.\end{aligned} \tag{46}$$

The above result can be calculated further, specifically according to the different distributions of the amount claimed.

Conflict of Interests

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

References

[1] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, New York, NY, USA, 2nd edition, 1998.

[2] S. Peng, "G-Expectation, G-Brownian motion and related stochastic calculus of Itô's type," in *Proceedings of the 2005 Abel Symposium*, vol. 2, pp. 541–567, Springer, 2006.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[4] H. Föllmer and A. Schied, "Convex measures of risk and trading constraints," *Finance and Stochastics*, vol. 6, no. 4, pp. 429–447, 2002.

[5] M. Frittelli and E. Rossaza Gianin, "Dynamic convex risk measures," in *New Riak Measures for the 21st Century*, G. Szegö, Ed., pp. 227–248, John Wiley & Sons, New York, NY, USA, 2004.

[6] S. Peng, "Survey on normal distributions, central limit theorem, Brownian motion and the related stochastic calculus under sublinear expectations," *Science in China A: Mathematics*, vol. 52, no. 7, pp. 1391–1411, 2009.

[7] L. G. Epstein and S. Ji, "Ambiguous volatility, possibility and utility in continuous time," *Journal of Mathematical Economics*, vol. 50, pp. 269–282, 2014.

[8] L. G. Epstein and S. Ji, "Ambiguous volatility and asset pricing in continuous time," *Review of Financial Studies*, vol. 26, no. 7, pp. 1740–1786, 2013.

[9] M. Hu, S. Ji, S. Peng, and Y. Song, "Backward stochastic differential equations driven by *G*-Brownian motion," *Stochastic Processes and Their Applications*, vol. 124, no. 1, pp. 759–784, 2014.

[10] M. Hu, S. Ji, S. Peng, and Y. Song, "Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by *G*-Brownian motion," *Stochastic Processes and Their Applications*, vol. 124, no. 2, pp. 1170–1195, 2014.

[11] M. Hu, S. Ji, and S. Yang, "A stochastic recursive optimal control problem under the *G*-expectation framework," *Applied Mathematics & Optimization*, 2014.

[12] X. Xiao, "Stochastic dominance under the nonlinear expected utilities," . In press.

[13] F. Hu, "On Cramér's theorem for capacities," *Comptes Rendus de l'Académie des Sciences I*, vol. 348, no. 17-18, pp. 1009–1013, 2010.

[14] S. Peng, "*G*-Brownian motion and dynamic risk measure under volatility uncertainty," http://arxiv.org/abs/0711.2834.

[15] S. Peng, "A new central limit theorem under sublinear expectations," http://arxiv.org/abs/0803.2656.

[16] F. Gao and H. Jiang, "Large deviations for stochastic differential equations driven by *G*-Brownian motion," *Stochastic Processes and Their Applications*, vol. 120, no. 11, pp. 2212–2240, 2010.

[17] L. Ren, *G-Lévy processes in finite and infinite dimensional space and related topics [Ph.D. thesis]*, 2012.

[18] M. Hu and S. Peng, "*G*-Lévy processesunder sublinear expectations," http://arxiv.org/abs/0911.3533.

[19] F. Gao and M. Xu, "Large deviations and Moderate deviations of independent random variables under sublinear expectations," *Science China Mathematics*, vol. 41, no. 4, pp. 337–352, 2011 (Chinese).

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 272745, 8 pages
http://dx.doi.org/10.1155/2014/272745

Research Article

The Stability and Stabilization of Stochastic Delay-Time Systems

Gang Li[1,2] and Ming Chen[1]

[1] *School of Mathematics and System Science, Shandong University of Science and Technology, Qingdao 265590, China*
[2] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 265590, China*

Correspondence should be addressed to Gang Li; ligangccm@163.com

Received 17 December 2013; Revised 17 April 2014; Accepted 1 May 2014; Published 3 June 2014

Academic Editor: Wuquan Li

The aim of this paper is to investigate the stability and the stabilizability of stochastic time-delay deference system. To do this, we use mainly two methods to give a list of the necessary and sufficient conditions for the stability and stabilizability of the stochastic time-delay deference system. One way is in term of the operator spectrum and H-representation; the other is by Lyapunov equation approach. In addition, we introduce the notion of unremovable spectrum of stochastic time-delay deference system, describe the PBH criterion of the unremovable spectrum of time-delay system, and investigate the relation between the unremovable spectrum and the stabilizability of stochastic time-delay deference system.

1. Introduction

The stochastic time-delay system is one of the fundamental research branches in the theory of control systems, which is usually applied in the fields of electronics, machinery, chemicals, life sciences, economics, and so on. As is well known, the stability is an essential concept in linear system theory, which is relative to the system matrix root-clustering in subregions of the complex plane, and also the spectral operator approach is effective in the study of the eigenvalue placement of a matrix (see [1]). Since two classic books [2, 3] appeared, stochastic stability and stabilization of Itô differential systems have been investigated by many researchers for several decades; we refer the reader to [4–6] and the references therein. More specifically, for linear time-invariant stochastic (LTIS) systems, most work is concentrated on the investigation of mean square stabilization, which has important applications in system analysis and design. Some necessary and sufficient conditions for the mean square stabilization of LTIS systems have been obtained in terms of generalized algebraic Riccati equation (GARE) or linear matrix inequality (LMI) in [7–14] or spectra of some operators in [5, 9, 15]. For the stochastic delay-time systems, the present results were mainly obtained by Lyapunov functional approach. We concentrate our attention upon the stability and stabilization of stochastic systems by the operator spectrum.

The structure of this paper is as follows. In Section 2, with the aid of the operator spectrum, H-representation, and Lyapunov equation approach, some necessary and sufficient conditions are given for the stability and the stabilizability of stochastic delay-time systems. In Section 3, the unremovable spectrum of stochastic delay-time systems is introduced, and PBH criterion of the stabilizability of stochastic delay-time systems is presented.

For convenience, we adopt the following traditional notations. S^n: the set of all symmetric matrices, whose components may be complex; $N = \{0, 1, 2, \ldots\}$; $A'(\mathrm{Ker}(A))$: the transpose (kernel space) of a matrix A; $A \geq 0$ $(A > 0)$ is a positive semidefinite (positive-definite) symmetric matrix A; I: identity matrix; $\sigma(L)$: spectral set of the operator or matrix L; $C^-(C^{-0})$: the open left (closed left) hand side complex plane. $D(0,\alpha) = \{\lambda \mid \|\lambda\| < \alpha\}$; $\|\cdot\|$ is the l_2-norm; $L^2_{\mathcal{F}_t}(R^+, R^{n_x})$: space of nonanticipative stochastic processes $x(t) \in R^{n_x}$ with respect to an increasing σ-algebra $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying $E\|x(t)\|^2 < \infty$. Finally, we make the assumption throughout this paper that all systems have real coefficients.

2. The Stability of Stochastic Delay-Time Systems

In this section, we will investigate the stability and stabilizability of the stochastic time-delay deference system using the spectrum of operator and Lyapunov equation approach. At first, we introduce a Lyapunove operator. Consider the following linear difference system with constant delays:

$$x(t+1) = F_0 x(t) + G_0 x(t) w(t) + \sum_{j=1}^{m} \left[F_j x(t-j) + G_j x(t-j) w(t) \right], \quad t \in N, \tag{1}$$

with the initial condition

$$x(k) = \varphi(k), \quad k = 0, -1, -2, \ldots, -m. \tag{2}$$

Here, $x \in R^n$ is a column vector, $F_j, G_j \in R^{n\times n}$, $j = 0, 1, \ldots, m$, are constant coefficient matrices, $\varphi(k) \in R^n$ is a deterministic initial condition, $\{w(t) \in R, t \in N\}$ is a sequence of real random variables defined on a complete probability space $\{\Omega, \mathcal{F}, \mathcal{F}_t, \mu\}$ which is a wide sense stationary, second-order process with $E(w(t)) = 0$ and $E(w(t)w(s)) = \delta_{s,t}$, where $\delta_{s,t}$ is the Kronecker delta with $\mathcal{F}_t = \sigma\{w(s) : 0 \le s \le t\}$.

Definition 1. The trivial stationary solution $x(t) = 0$ of the system (1) is called mean square stable if, for any arbitrarily small number $\varepsilon > 0$, there exists a number $\delta > 0$, when $\|\varphi\| < \delta$, such that

$$E\|x(t)\|^2 < \varepsilon, \tag{3}$$

for a solution $x(t) = x(t, \varphi)$ of (1).

Definition 2. The trivial stationary solution $x(t) = 0$ of the system (1) is called asymptotically mean square stable if it is stable in the sense of Definition 1 and, moreover, any solution $x(t) = x(t, \varphi)$ of (1) satisfies

$$\lim_{t\to+\infty} E\|x(t)\|^2 = 0. \tag{4}$$

We consider the problem of finding criteria and sufficient conditions for the mean square asymptotic stability of the trivial stationary solution $x(t) = 0$ by operator spectra. Since the stochastic system (1) is a system with time-delays, it seems impossible to construct the operator directly for (1) like the operator in [15]. So, first of all, we introduce the following column vector $\tilde{x}(t)$ of new variables of dimension $n(m+1)$:

$$\tilde{x}(t) = \left[x'(t), x'(t-1), \ldots, x'(t-m) \right]'. \tag{5}$$

The stochastic system (1) with time-delays can now be written in the form of an equivalent stochastic system of dimension $n(m+1)$ without delay; namely,

$$\tilde{x}(t+1) = [F + G\omega(t)]\, \tilde{x}(t), \tag{6}$$

where F and G denote the following $n(m+1) \times n(m+1)$ matrices:

$$F = \begin{pmatrix} F_0 & F_1 & \cdots & F_{m-1} & F_m \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{pmatrix}, \quad G = \begin{pmatrix} G_0 & G_1 & \cdots & G_{m-1} & G_m \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{7}$$

If we set $X(t) = E\tilde{x}(t)\tilde{x}'(t)$, $X(t)$ satisfies the following difference equation:

$$X(t+1) = FX(t)F' + GX(t)G'. \tag{8}$$

Motivated by (8), we introduce the following linear Lyapunov operator:

$$\mathcal{L}_{F,G} : X \in S^{n(m+1)} \longmapsto FX(t)F' + GX(t)G' \in S^{n(m+1)}. \tag{9}$$

With the use of the Kronecker matrix product, the matrix equation (8) can be rewritten in the vector matrix form as follows:

$$\vec{X}(t+1) = \widehat{F}\vec{X}, \tag{10}$$

where $\vec{X}(t)$ denotes the $n^2(m+1)^2$-dimensional column vector

$$\vec{X}(t) = \left[X_{1,1}(t), \ldots, X_{1,n}(t), \ldots, X_{1,n(m+1)}(t), \ldots, X_{n(m+1),n(m+1)}(t) \right]' \tag{11}$$

and $\widehat{F} \in R^{n^2(m+1)^2 \times n^2(m+1)^2}$ has the form $\widehat{F} = F \otimes F + G \otimes G$.

Now, we are in a position to give a spectral description for the stability of system (1) by H-representation in [14].

Lemma 3. *Let $H_{n(m+1)}$ be a $n^2(m+1)^2 \times (n(m+1)[n(m+1)+1]/2)$ matrix and $\operatorname{rank}(H_{n(m+1)}) = (n(m+1)[n(m+1)+1])/2$; then $H'_{n(m+1)}H_{n(m+1)}$ is invertible.*

Theorem 4. *The trivial solution $x(t) = 0$ of system (1) is asymptotically mean square stable if and only if $\sigma(\mathcal{L}_{F,G}) \subset D(0,1)$.*

Proof. If we set $X(t) = E\tilde{x}(t)\tilde{x}'(t)$, $X(t)$ satisfies

$$\begin{gathered} X(t+1) = FX(t)F' + GX(t)G', \\ X(k) = \tilde{x}(k)\tilde{x}'(k) \in S^{n(m+1)}, \\ k = 0, -1, -2, \ldots, -m, \quad t \in N. \end{gathered} \tag{12}$$

Since $X(\cdot)$ is real symmetric, (12) is a linear matrix equation with $n(m+1)[n(m+1)+1]/2$ different variables; that is, it is in fact an $n(m+1)[n(m+1)+1]/2$ th-order linear system. We define a map $\widetilde{\mathcal{L}}$ from $S^{n(m+1)}$ to $C^{n(m+1)[n(m+1)+1]/2}$ as follows.

For any $Y = (Y_{ij})_{n(m+1)\times n(m+1)} \in S^{n(m+1)}$, set

$$\begin{aligned}\widetilde{Y} = \widetilde{\mathscr{L}}(Y) = (&Y_{11},\ldots,Y_{1,n(m+1)},\ldots,Y_{n(m+1)-1,n(m+1)-1},\\ &Y_{n(m+1)-1,n(m+1)},\ldots,Y_{n(m+1),n(m+1)})';\end{aligned} \tag{13}$$

then there exists an unique matrix $\theta(H_{n(m+1)}) \in R^{(n(m+1)[n(m+1)+1]/2)\times(n(m+1)[n(m+1)+1]/2)}$, by H-representation of [14], such that (12) is equivalent to

$$\begin{aligned}&\widetilde{X}(t+1) = \widetilde{\mathscr{L}}(\mathscr{L}_{F,G}(X)) = \theta(H_{n(m+1)})\widetilde{X}(t),\\ &\widetilde{X}(k) = \left[H'_{n(m+1)}H_{n(m+1)}\right]^{-1}H'_{n(m+1)}\\ &\qquad\times\vec{X}(k) \in R^{n(m+1)[n(m+1)+1]/2},\\ &k = 0,-1,-2,\ldots,-m,\quad t\in N,\end{aligned} \tag{14}$$

where $\theta(H_{n(m+1)}) = [H'_{n(m+1)}H_{n(m+1)}]^{-1}H'_{n(m+1)}[F\otimes F + G\otimes G]H_{n(m+1)}$, $\widetilde{X} \in R^{n(m+1)[n(m+1)+1]/2}$. Obviously, since the system of (14) for moments is deterministic, the proof of the theorem is carried out by the standard method for deterministic difference equations. Seeking the general solution of system (14) in the exponential form $\widetilde{X}(t) = c\lambda^t$, where $c,\lambda =$ const, we arrive at the characteristic equation $\det(\lambda I - \theta(H_{n(m+1)})) = 0$. That is, $\lim_{t\to+\infty} E\tilde{x}(t)\tilde{x}'(t) = \lim_{t\to+\infty} X(t) = 0 \Leftrightarrow \lim_{t\to+\infty}\widetilde{X}(t) = 0 \Leftrightarrow \sigma(\theta(H_{n(m+1)})) \subset D(0,1)$.

By (9) and (14), for any eigenvalue λ and its corresponding eigenvector $Y = (Y_{ij})_{n\times n} \in S^n$ of $\mathscr{L}_{F,G}$, from $\mathscr{L}_{F,G}(Y) = \lambda Y$, we have $\widetilde{\mathscr{L}}(\mathscr{L}_{F,G}(Y)) = \theta(H_{n(m+1)})\widetilde{Y} = \lambda\widetilde{Y}$, which yields $\sigma(\mathscr{L}_{F,G}) = \sigma(\theta(H_{n(m+1)}))$. The above discussion concludes the proof of Theorem 4. The proof of Theorem 4 is complete. □

Remark 5. In Theorem 4, a necessary and sufficient condition for the asymptotically mean square stability of system (1) via the spectrum of $\mathscr{L}_{F,G}$ is presented, which can be called "spectral criterion."

Theorem 6. *The trivial solution $x = 0$ of system (1) is asymptotically mean square stable if and only if, for any $Q \in S^{n(m+1)}$ with $Q > 0$, there exists a $P \in S^{n(m+1)}$ such that $P > 0$ and P is a solution of the following Lyapunov equation:*

$$P - \mathscr{L}_{F,G}(P) = Q. \tag{15}$$

Proof. We introduce an $n^2(m+1)^2$-parameter stochastic Lyapunov function as a quadratic form:

$$V\left(\vec{X}(t)\right) = \vec{X}'P\vec{X},\quad P\in S^{n(m+1)}. \tag{16}$$

The role of parameters is played by $n^2(m+1)^2$ elements of the positive-definite matrix, which should be determined. The statement of the theorem can be established in a way that is standard for the method of Lyapunov functions for stochastic difference equations. So the trivial solution $\tilde{x}(t) = 0$ of system (6) is asymptotically mean square stable if and only if for any $Q > 0$, the Lyapunov equation (15) has a solution $P > 0$. By the proof of Theorem 4, the trivial solution $x = 0$ of system (1) is asymptotically mean square stable if and only if the trivial solution $\tilde{x}(t) = 0$ of system (6) is asymptotically mean square stable. The proof of Theorem 6 is complete. □

From the proof of Theorem 4 and the method of Lyapunov functions for difference equations, we immediately get the following result.

Theorem 7. *The trivial solution $x = 0$ of system (1) is asymptotically mean square stable if and only if, for any $Q \in S^{n(m+1)[n(m+1)+1]/2}$ with $Q > 0$, there exists a $P \in S^{n(m+1)[n(m+1)+1]/2}$ such that $P > 0$ and P is a solution of the following Lyapunov equation:*

$$P - \widehat{F}'P\widehat{F} = Q. \tag{17}$$

Corollary 8. *If $\sigma(\mathscr{L}_{F,G}) \subset D(0,1)$, then $\sigma(F) \subset D(0,1)$.*

Proof. By Theorems 4 and 6, $\sigma(\mathscr{L}_{F,G}) \subset D(0,1)$ holds if and only if there is a matrix $P \in S^{n(m+1)}$ such that $P > 0$ and P is a solution of the following Lyapunov equation:

$$P - \mathscr{L}_{F,G}(P) = Q, \tag{18}$$

for any $Q > 0$. So, there exists a $P \in P \in S^{n(m+1)}$ such that $P > 0$ and P is a solution of the following Lyapunov equation:

$$P - F'PF > 0, \tag{19}$$

which is equivalent to $\sigma(F) \subset D(0,1)$; that is, the system

$$\begin{aligned}&x(t+1) = F_0x(t) + \sum_{1}^{j=m} F_jx(t-j),\\ &x(k) = \varphi(k),\quad k = 0,-1,-2,\ldots,-m,\ t\in N,\end{aligned} \tag{20}$$

is asymptotically Lyapunov stable. The proof of Corollary 8 is complete. □

Now, we present some results about mean square stability of system (1). From the process of Theorems 4–7, we easily obtain the following Theorems 9–10, so we omit their proofs.

Theorem 9. *If the trivial stationary solution $x = 0$ of the system (1) is mean square stable, then $\sigma(\mathscr{L}_{F,G}) \subset \overline{D}(0,1)$.*

Theorem 10. *$\sigma(\mathscr{L}_{F,G}) \subset \overline{D}(0,1)$ if and only if one of the following conditions holds.*

(1) *For any $\varepsilon > 0$ and $Q > 0$, the following Lyapunov equation*

$$P - \mathscr{L}_{e^{-\varepsilon}F,e^{-\varepsilon}G}(P) = Q \tag{21}$$

has a positive-definite solution P.

(2) $\sigma(\mathscr{L}_{e^{-\varepsilon}F,e^{-\varepsilon}G}) \subset D(0,1)$.

Corollary 11. *If $\sigma(\mathscr{L}_{F,G}) \subset \overline{D}(0,1)$, then $\sigma(F) \subset \overline{D}(0,1)$.*

Proof. Since $\sigma(\mathscr{L}_{F,G}) \subset \overline{D}(0,1)$, we have, for any sufficient small $\varepsilon > 0$, $\sigma(\mathscr{L}_{e^{-\varepsilon}F,e^{-\varepsilon}G}) \subset D(0,1)$. By Theorem 4, $(e^{-\varepsilon}F, e^{-\varepsilon}G)$ is mean square stable, which implies $\sigma(\mathscr{L}_{e^{-\varepsilon}F}) \subset D(0,1)$ by Corollary 8. Let $\varepsilon \to 0$, we have $\sigma(F) \subset \overline{D}(0,1)$ by continuity of spectrum [16]. The proof of Corollary 11 is complete. □

Remark 12. $\sigma(F) \subset \overline{D}(0,1)$ does not imply $\sigma(\mathscr{L}_{F,G}) \subset \overline{D}(0,1)$, which is one of the essential differences between stochastic system and deterministic system.

Theorem 13. *The trivial solution $x = 0$ of system* (1) *is mean square stable if and only if, for any $Q \geq 0$, there exists a $P \in S^{n(m+1)}$ such that $P > 0$ and P is a solution of the following Lyapunov equation:*

$$P - \mathscr{L}_{F,G}(P) = Q. \tag{22}$$

Proof. We introduce an $n^2(m+1)^2$-parameter stochastic Lyapunov function as a quadratic form:

$$V\left(\vec{X}(t)\right) = \vec{X}'P\vec{X}, \quad P \in R^{n(m+1)\times n(m+1)}. \tag{23}$$

By the method of Lyapunov functions for stochastic difference equations, we can get the result. The proof of Theorem 13 is complete. □

Now, we give an example to show how to solve the spectrum of stochastic time-delay deference system by H-representation.

Example 14. Consider the following stochastic system:

$$x(t+1) = ax(t) + bx(t-1) + cx(t)\omega(t) + dx(t-1)\omega(t). \tag{24}$$

Letting $\overline{x}(t) = (x(t), x(t-1))'$,

$$\overline{x}(t+1) = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}\overline{x}(t) + \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix}\overline{x}(t)\omega(t). \tag{25}$$

Letting $X(t) = E\,\overline{x}(t)\overline{x}'(t)$,

$$X(t+1) = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}X(t)\begin{pmatrix} a & 1 \\ b & 0 \end{pmatrix} + \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix}X(t)\begin{pmatrix} c & 0 \\ d & 0 \end{pmatrix}. \tag{26}$$

Letting $F = \left(\begin{smallmatrix} a & b \\ 1 & 0 \end{smallmatrix}\right)$, $G = \left(\begin{smallmatrix} c & d \\ 0 & 0 \end{smallmatrix}\right)$, $\vec{X}(t) = (x_{11}, x_{12}, x_{12}, x_{22})'$,

$$\begin{aligned}\vec{X}(t+1) &= (F\otimes F + G\otimes G)\vec{X}(t) \\ &= \begin{pmatrix} a^2+c^2 & ab+cd & ab+cd & b^2+d^2 \\ a & 0 & b & 0 \\ a & b & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\vec{X}(t).\end{aligned} \tag{27}$$

Choose

$$H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{28}$$

Let $\widetilde{X}(t) = (x_{11}, x_{12}, x_{22})'$; then $\vec{X}(t) = H_2\widetilde{X}(t)$ and

$$\widetilde{X}(t+1) = \left(H_2'H_2\right)^{-1}H_2'(F\otimes F + G\otimes G)H_2\widetilde{X}(t). \tag{29}$$

So $\sigma((H_2'H_2)^{-1}H_2'(F\otimes F + G\otimes G)H_2) = \sigma(\mathscr{L})$ if we choose $a = 1$, $b = 0$, $c = 0$, $d = 1/4$, and $\sigma(\mathscr{L}) = \{0, 1/2, 1/2\}$.

For a state feedback control law $u(t) = Kx(t)$, we introduce a linear operator $\mathscr{L}_K$ associated with the closed-loop system:

$$\begin{aligned} x(t+1) = {} & F_0x(t) + M_0u(t) + \left(G_0x(t) + N_0u(t)\right)\omega(t) \\ & + \sum_{j=1}^{m}\left[F_jx(t-j) + M_ju(t-j)\right. \\ & \left. + \left(G_jx(t-j) + N_ju(t-j)\right)\omega(t)\right], \\ x(k) = {} & \varphi(k) \in R^n, \quad k = 0, -1, \ldots, -m, \ t \in N, \end{aligned} \tag{30}$$

where $x \in R^n$ is a column vector, $F_j, G_j \in R^{n\times n}$, $j = 0, 1, \ldots, m$ are constant coefficient matrices, $\varphi(k)$ is a deterministic initial condition, and $u(t) \in R^n$ is a control input.

Definition 15. The trivial stationary solution $x(t) = 0$ of the system (30) is called mean square stabilization if there exists an input feedback K such that, for any arbitrarily small number $\varepsilon > 0$, one can find a number $\delta > 0$, when $\|\varphi\| < \delta$, satisfying

$$E\|x(t)\|^2 < \varepsilon, \tag{31}$$

for a solution $x(t) = x(t, \varphi)$ of (30).

Definition 16. The trivial stationary solution $x(t) = 0$ of the system (30) is called asymptotically mean square stabilization if it is stable in the sense of Definition 15 and, moreover,

$$\lim_{t\to+\infty} E\|x(t)\|^2 = 0. \tag{32}$$

We introduce the following column vectors $\overline{x}(t)$ and $\overline{u}(t)$ of new variables of dimension $n(m+1)$:

$$\begin{aligned}\overline{x}(t) &= \left[x'(t), x'(t-1), \ldots, x'(t-m)\right]', \\ \overline{u}(t) &= \left[u'(t), u'(t-1), \ldots, u'(t-m)\right]'.\end{aligned} \tag{33}$$

The stochastic system (30) with time-delays can now be written in the form of an equivalent stochastic system of dimension $n(m+1)$ without time-delay; namely,

$$\overline{x}(t+1) = \overline{F}\overline{x}(t) + \overline{M}\overline{u}(t) + \left(\overline{G}\overline{x}(t) + \overline{N}\overline{u}(t)\right)\omega(t), \tag{34}$$

where

$$\overline{F}=\begin{pmatrix} F_0 & F_1 & \cdots & F_{m-1} & F_m \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{pmatrix},\qquad \overline{G}=\begin{pmatrix} G_0 & G_1 & \cdots & G_{m-1} & G_m \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{35}$$

$$\overline{M}=\begin{pmatrix} M_0 & M_1 & \cdots & M_{m-1} & M_m \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & 0 & 0 \end{pmatrix},\qquad \overline{N}=\begin{pmatrix} N_0 & N_1 & \cdots & N_{m-1} & N_m \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{36}$$

Take a control input $\overline{u}(t)=\overline{K}\overline{x}(t)$ with

$$\overline{K}=\begin{pmatrix} K & 0 & \cdots & 0 & 0 \\ 0 & K & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & K \end{pmatrix}, \tag{37}$$

and set $X(t)=E\overline{x}(t)\overline{x}'(t)$; $X(t)$ satisfies the following difference equation:

$$\begin{aligned} X(t+1) &= \left(\overline{F}+\overline{MK}\right)X(t)\left(\overline{F}+\overline{MK}\right)' \\ &\quad + \left(\overline{G}+\overline{NK}\right)X(t)\left(\overline{G}+\overline{NK}\right)'. \end{aligned} \tag{38}$$

Motivated by (38), we introduce the following linear Lyapunov operator:

$$\begin{aligned} \mathscr{L}_{\overline{K}}: X\in S^{n(m+1)} &\longmapsto \left(\overline{F}+\overline{MK}\right)X(t)\left(\overline{F}+\overline{MK}\right)' \\ &\quad + \left(\overline{G}+\overline{NK}\right)X(t)\left(\overline{G}+\overline{NK}\right)' \in S^{n(m+1)}. \end{aligned} \tag{39}$$

With the use of the Kronecker matrix product, the matrix equation (38) can be rewritten in the vector matrix form as follows:

$$\vec{X}(t+1)=\widehat{A}\vec{X}, \tag{40}$$

where $\vec{X}(t)$ denotes the $n^2(m+1)^2$-dimensional column vector

$$\begin{aligned} \vec{X}(t) &= \left[X_{1,1}(t),\ldots,X_{1,n}(t),\ldots,\right. \\ &\qquad \left. X_{1,n(m+1)}(t),\ldots,X_{n(m+1),n(m+1)}(t)\right]' \end{aligned} \tag{41}$$

and $\widehat{A}=(\overline{F}+\overline{MK})\otimes(\overline{F}+\overline{MK})+(\overline{G}+\overline{NK})\otimes(\overline{G}+\overline{NK})\in S^{n^2(m+1)^2}$.

From Theorems 4–13, we can easily obtain the following results.

Corollary 17. *The trivial solution $x=0$ of system (30) is asymptotical mean square stabilizaton if and only if $\sigma(\mathscr{L}_{\overline{K}})\subset D(0,1)$.*

Corollary 18. *The trivial solution $x(t)=0$ of system (30) is asymptotically mean square stable if and only if, for any $Q\in S^{n(m+1)}$ with $Q>0$, there exists a $P\in S^{n(m+1)}$ such that $P>0$ and P is a solution of the following Lyapunov equation:*

$$P-\mathscr{L}_{\overline{F}+\overline{MK},\overline{G}+\overline{NK}}(P)=Q. \tag{42}$$

Corollary 19. *The trivial solution $x=0$ of system (30) is asymptotically mean square stable if and only if, for any $Q\in S^{n(m+1)[n(m+1)+1]/2}$ with $Q>0$, there exists a $P\in S^{n(m+1)[n(m+1)+1]/2}$ such that $P>0$ and P is a solution of the following Lyapunov equation:*

$$P-\theta\left(H^K_{n(m+1)}\right)'P\theta\left(H^K_{n(m+1)}\right)=Q, \tag{43}$$

where

$$\begin{aligned} \theta\left(H^K_{n(m+1)}\right) &= \left[H'_{n(m+1)}H_{n(m+1)}\right]^{-1}H'_{n(m+1)} \\ &\quad \times\left[\left(\overline{F}+\overline{MK}\right)\otimes\left(\overline{F}+\overline{MK}\right)\right. \\ &\quad \left. +\left(\overline{G}+\overline{NK}\right)\otimes\left(\overline{G}+\overline{NK}\right)\right]H_{n(m+1)}. \end{aligned} \tag{44}$$

Corollary 20. *The trivial solution $x=0$ of system (30) is asymptotically mean square stable if and only if, for any $Q\in S^{n^2(m+1)^2}$, $Q>0$, there exists a $P\in S^{n^2(m+1)^2}$ such that $P>0$ and P is a solution of the following Lyapunov equation:*

$$P-\widehat{A}'P\widehat{A}=Q. \tag{45}$$

Corollary 21. *The trivial solution $x=0$ of system (30) is mean square stable if and only if, for any $Q\in S^{n(m+1)}$ with $Q\geq 0$, there exists a $P\in S^{n(m+1)}$ such that $P>0$ and P is a solution of the following Lyapunov equation:*

$$P-\mathscr{L}_{\overline{F}+\overline{MK},\overline{G}+\overline{NK}}(P)=Q. \tag{46}$$

Corollary 22. *The trivial solution $x=0$ of system (1) is mean square stable if and only if, for any $Q\in S^{n(m+1)[n(m+1)+1]/2}$ with $Q\geq 0$, there exists a $P\in S^{n(m+1)[n(m+1)+1]/2}$ such that $P>0$ and P is a solution of the following Lyapunov equation:*

$$P-\theta\left(H^K_{n(m+1)}\right)'P\theta\left(H^K_{n(m+1)}\right)=Q, \tag{47}$$

where

$$\begin{aligned} \theta\left(H^K_{n(m+1)}\right) &= \left[H'_{n(m+1)}H_{n(m+1)}\right]^{-1} \\ &\quad \times H'_{n(m+1)}\left[F\otimes F+G\otimes G\right]H_{n(m+1)}. \end{aligned} \tag{48}$$

Corollary 23. *The trivial solution $x = 0$ of system* (1) *is mean square stable if and only if, for any $Q \geq 0$, there exists a $P \in S^{n^2(m+1)^2}$ such that $P > 0$ and P is a solution of the following Lyapunov equation:*

$$P - \widehat{A}' P \widehat{A} = Q. \tag{49}$$

3. Popov-Belevith-Hautus Criterion of the Stabilizability

In this section, we will investigate the properties of unremovable spectrum of time-delay deference system and the relation between unremovable spectrum and the stabilizability of time-delay deference system. Consider the following linear stochastic system with time-delays:

$$\begin{aligned} x(t+1) &= F_0 x(t) + M_0 u(t) + \left(G_0 x(t) + N_0 u(t)\right) w(t) \\ &\quad + \sum_{j=1}^{m} \left[F_j x(t-j) + M_j u(t-j) \right. \\ &\quad \left. + \left(G_j x(t-j) + N_j u(t-j)\right) w(t)\right], \\ x(k) &= \varphi(k) \in R^n, \quad k = 0, -1, -2, \ldots, -m, \ t \in N. \end{aligned} \tag{50}$$

Definition 24. We say that λ is an unremovable spectrum of system (50) with state feedback if there exists $Z \neq 0 \in S^{n(m+1)}$, such that,for any $\overline{K} \in R^{n(m+1)\times n(m+1)}$,

$$\begin{aligned} \mathscr{L}^*_{\overline{K}}(Z) &= (\overline{F} + \overline{MK})' Z \left(\overline{F} + \overline{MK}\right) \\ &\quad + \left(\overline{G} + \overline{NK}\right)' Z \left(\overline{G} + \overline{NK}\right) = \lambda Z \end{aligned} \tag{51}$$

holds.

Remark 25. It is easy to see that the operator $\mathscr{L}^*_{\overline{K}}$ is the adjoint operator of the operator $\mathscr{L}_{\overline{K}}$ with the inner product $\langle Z, Y\rangle = \operatorname{trace}(Z^*, Y)$ for any $Z, Y \in S^{n(m+1)}$. As we restrict the coefficients to real matrices, $\sigma(\mathscr{L}^*_{\overline{K}}) = \sigma(\mathscr{L}_{\overline{K}})$. By Corollaries 17–23, we know that any one of them can characterize the stabilizability of system (50).

Obviously, if λ is an unremovable spectrum, then it can be regarded as an uncontrollable mode as in deterministic systems. We give a theorem with respect to the unremovable spectrum below.

Theorem 26. *(Stochastic PBH criterion) λ is an unremovable spectrum of system* (50) *if and only if there exists $Z \neq 0 \in S^{n(m+1)}$, such that the following three equalities hold simultaneously:*

$$\begin{aligned} \overline{F}' Z \overline{F} + \overline{G}' Z \overline{G} &= \lambda Z, \\ \overline{F}' Z \overline{M} + \overline{G}' Z \overline{N} &= 0, \\ \overline{N}' Z \overline{N} &= -\overline{M}' Z \overline{M}. \end{aligned} \tag{52}$$

Proof. Note that (51) can be written as

$$\begin{aligned} &\overline{F}' Z \overline{F} + \overline{K}' \overline{M}' Z \overline{MK} + \overline{G}' Z \overline{G}' + \overline{K}' \overline{N}' Z \overline{NK} \\ &\quad + \left(\overline{F}' Z \overline{M} + \overline{G}' Z \overline{N}\right) \overline{K} + \overline{K}' \left(\overline{M}' Z \overline{F} + \overline{N}' Z \overline{G}\right) = \lambda Z, \end{aligned} \tag{53}$$

so if (52) holds, then (53) automatically holds. So the sufficiency is proved.

To prove the necessity, we first take $\overline{K} = 0$ in (51), then

$$\overline{F}' Z \overline{F} + \overline{G} Z \overline{G}' = \lambda Z \tag{54}$$

holds. Again, from (53), it follows that

$$\begin{aligned} &\overline{K}' \overline{M} Z \overline{M}' \overline{K} + \overline{KN} Z \overline{N}' \overline{K}' \\ &\quad + \left(\overline{F}' Z \overline{M} + \overline{G}' Z \overline{N}\right) \overline{K} + \overline{K}' \left(\overline{M}' Z \overline{F} + \overline{N}' Z \overline{G}\right) = 0. \end{aligned} \tag{55}$$

Let $\mathscr{F} = \overline{F}' Z \overline{M} + \overline{G}' Z \overline{N}$, $\mathscr{M} = \overline{M} Z \overline{M}'$, $\mathscr{N} = \overline{N} Z \overline{N}'$; then (55) becomes

$$\mathscr{F} \overline{K} + \overline{K}' \mathscr{F}' = -\overline{K}' \mathscr{M} \overline{K} - \overline{K}' \mathscr{N} \overline{K}. \tag{56}$$

Since the left-hand side in (56) is linear with respect to $\overline{K}$, we must have $\mathscr{M} + \mathscr{N} = 0$. In fact, due to the linearity of the following equation

$$\begin{aligned} &\left(\overline{K} + \overline{K}\right)' (\mathscr{M} + \mathscr{N}) \left(\overline{K} + \overline{K}\right) \\ &\quad = 4\overline{K}' (\mathscr{M} + \mathscr{N}) \overline{K} = 2\overline{K}' (\mathscr{M} + \mathscr{N}) \overline{K}, \end{aligned} \tag{57}$$

$\overline{K}'(\mathscr{M} + \mathscr{N})\overline{K} = 0$. Because of the arbitrariness of $\overline{K}$, it is necessary that $\mathscr{M} + \mathscr{N} = 0$; that is, $\overline{M} Z \overline{M}' = -\overline{N} Z \overline{N}'$. To prove $\mathscr{F} = 0$ or $\overline{F}' Z \overline{M} + \overline{G}' Z \overline{N} = 0$, we note that (56) becomes $\overline{K}' \mathscr{F}' = -\mathscr{F} \overline{K}$. Denote $\mathscr{F} = (f_{ij})_{n(m+1)\times n(m+1)}$, and take

$$\overline{K} = \overline{K}_{ij} = (k_{ls})_{n(m+1)\times n(m+1)} = \begin{cases} 1, & l = i, s = j, \\ 0, & \text{otherwise}. \end{cases} \tag{58}$$

From $\overline{K}' \mathscr{F}' = -\mathscr{F} \overline{K}$, one knows that $f_{ij} = 0$, $i, j = 1, 2, \ldots, n(m+1)$; that is, $\mathscr{F} = 0$. The proof of Theorem 26 is complete. □

Theorem 27. *If system* (50) *is asymptotically mean square stabilizable, then all unremovable spectra of system* (50) *must belong to $D(0, 1)$.*

Proof. If there is an unremovable spectrum μ of (50) with $|\mu| \geq 1$, then, by Theorem 26, there exists $Z \neq 0 \in S^{n(m+1)}$, such that the following three equalities hold simultaneously:

$$\begin{aligned} \overline{F}' Z \overline{F} + \overline{G}' Z \overline{G} &= \mu Z, \\ \overline{F}' Z \overline{M} + \overline{G}' Z \overline{N} &= 0, \\ \overline{N}' Z \overline{N} &= -\overline{M}' Z \overline{M}. \end{aligned} \tag{59}$$

So, for any feedback gain $\overline{K}$, we obtain $\mathscr{L}_{\overline{K}}(Z) = \mu Z, \mu \notin D(0,1)$. The proof of Theorem 27 is complete. It follows that, for any $\overline{K}$, $\mu \in \sigma(\mathscr{L}_{\overline{K}})$, which contradicts the asymptotical mean square stabilization of system (50).

By Theorems 26 and 27, the deterministic Popov-Belevith-Hautus Criterion can be stated in another form as follows. □

Corollary 28. *Assume that* $G_j = N_j = 0$, $j = 0, 1, \ldots, m$; *then system* (50) *is asymptotically stabilizable if and only if all of the unremovable spectra of system* (50) *belong to* $D(0,1)$; *that is, there does not exist a nonzero* $Z \in S^{n(m+1)}$, *and* $\lambda \notin D(0,1)$ *satisfying*

$$\overline{F}' Z \overline{F} = \lambda Z, \qquad Z\overline{M} = 0. \tag{60}$$

Proof. The necessity is obvious. To prove the sufficiency part, we note that if (F_0, M_0) is not stabilizable, then, by Popov-Belevith-Hautus Criterion, there exists a nonzero $\xi \in C^{n(m+1)}$, $\lambda \notin D(0,1)$ satisfying $\xi'\overline{F} = \lambda\xi'$, $\xi'\overline{M} = 0$. Take $Z = \xi\xi'$, then (60) holds, which contradicts the given condition. Corollary 28 is proved. □

Remark 29. Corollary 28 indicates that there is no difference between unremovable spectrum and uncontrollable mode for deterministic systems.

Theorem 30. *If system* (50) *is mean square stabilizable, then all the existing unremovable spectra of* (50) *must belong to* $\overline{D}(0,1)$.

Proof. If there is an unremovable spectrum μ of system (50) with $\mu \notin \overline{D}(0,1)$, then, by Theorem 26, there exists $Z \neq 0 \in S^{n(m+1)}$, such that the following three equalities hold simultaneously:

$$\begin{aligned} \overline{F}' Z \overline{F} + \overline{G}' Z \overline{G} &= \mu Z, \\ \overline{F}' Z \overline{M} + \overline{G}' Z \overline{N} &= 0, \\ \overline{N}' Z \overline{N} &= -\overline{M}' Z \overline{M}. \end{aligned} \tag{61}$$

So for any feedback gain $\overline{K}$, we obtain $\mathscr{L}^*_{\overline{K}}(Z) = \mu Z$, $\mu \notin \overline{D}(0,1)$. It follows that, for any $\overline{K}$, $\mu \in \sigma(\mathscr{L}_{\overline{K}})$, which contradicts mean square stabilization of system (50). The proof of Theorem 30 is complete. □

4. Conclusion

In this paper, we investigate the stability and stabilizability of stochastic delay-time systems. By H-representation, present the spectral criteria of the stability and stabilizability. By generalized Lyapunov equation approach, the equivalent conditions of mean square stabilizability and asymptotically mean square stabilizability of system (50) are given. We introduce the notion of unremovable spectrum of stochastic time-delay deference system, present the PBH criterion of the unremovable spectrum of stochastic time-delay system, and investigate the relation between the unremovable spectrum and the stabilizability of stochastic time-delay deference system.

Conflict of Interests

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

Acknowledgment

This work was supported by the National Natural Science Foundation of China under Grant nos. 61174078 and 61201430.

References

[1] A. G. Mazko, "The Lyapunov matrix equation for a certain class of regions bounded by algebraic curves," *Soviet Automatic Control*, vol. 42, pp. 12–17, 1980.

[2] R. Z. Has'minskii, *Stochastic Stability of Differential Equations*, vol. 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*, Sijtjoff and Noordhoff, Alphen, The Netherlands, 1980.

[3] H. J. Kushner, *Stochastic Stability and Control*, vol. 33 of *Mathematics in Science and Engineering*, Academic Press, New York, NY, USA, 1967.

[4] M. A. Rami and X. Y. Zhou, "Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1131–1143, 2000.

[5] T. Damm, *Rational Matrix Equations in Stochastic Control*, vol. 297 of *Lecture Notes in Control and Information Sciences*, Springer, New York, NY, USA, 2004.

[6] X. Mao, *Stochastic Differential Equations and Their Applications*, Horwood, Chichester, UK, 1997.

[7] Y. Liu, *Backward stochastic differential equation and stochastic control system [Ph.D. dissertation]*, Shandong University, Jinan, China, 1999.

[8] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15, SIAM, Philadelphia, Pa, USA, 1994.

[9] M. D. Fragoso, O. L. V. Costa, and C. E. de Souza, "A new approach to linearly perturbed Riccati equations arising in stochastic control," *Applied Mathematics and Optimization*, vol. 37, no. 1, pp. 99–126, 1998.

[10] W. M. Wonham, "On a matrix Riccati equation of stochastic control," *SIAM Journal on Control*, vol. 6, no. 4, pp. 681–697, 1968.

[11] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, vol. 44 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1983.

[12] W. Zhang, H. Zhang, and B.-S. Chen, "Generalized Lyapunov equation approach to state-dependent stochastic stabilization/detectability criterion," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1630–1642, 2008.

[13] W. Zhang and B. S. Chen, "Some properties of generalized Lyapunov equations," in *Proceedings of the Chinese Control and Decision Conference (CCDC '11)*, pp. 3137–3141, Mianyang, China, May 2011.

[14] W. Zhang and B. S. Chen, "H-representation and applications to generalized Lyapunov equations and linear stochastic systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3009–3022, 2012.

[15] W. Zhang and B. S. Chen, "On stabilizability and exact observability of stochastic systems with their applications," *Automatica*, vol. 40, no. 1, pp. 87–94, 2004.

[16] D. G. Korenevskii, "Stability criteria for solutions of systems of linear deterministic or stochastic delay difference equations with continuous time," *Mathematical Notes*, vol. 70, no. 1-2, pp. 192–205, 2001.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 421687, 8 pages
http://dx.doi.org/10.1155/2014/421687

Research Article

Backward Stochastic H_2/H_∞ Control: Infinite Horizon Case

Zhen Wu[1] and Qixia Zhang[2]

[1] *School of Mathematics, Shandong University, Jinan 250100, China*
[2] *School of Mathematical Sciences, University of Jinan, Jinan 250022, China*

Correspondence should be addressed to Qixia Zhang; zhangqixia110@163.com

Received 4 March 2014; Accepted 1 May 2014; Published 22 May 2014

Academic Editor: Weihai Zhang

The mixed H_2/H_∞ control problem is studied for systems governed by infinite horizon backward stochastic differential equations (BSDEs) with exogenous disturbance signal. A necessary and sufficient condition for the existence of a unique solution to the H_2/H_∞ control problem is derived. The equivalent feedback solution is also discussed. Contrary to deterministic or stochastic forward case, the feedback solution is no longer feedback of the current state; rather, it is feedback of the entire history of the state.

1. Introduction

H_∞ control is one of the most important robust control approaches in which control law is sought to efficiently eliminate the effect of the exogenous disturbance in the practical system. We refer the reader to [1–3] and the references therein. If the purpose is to select control not only to restrain the exogenous disturbance, but also to minimize a cost function when the worst case disturbance d^* is implemented, this is the so-called mixed H_2/H_∞ control problem. Mixed H_2/H_∞ control problem has attracted much attention and has been widely applied to various fields. Please refer to [4, 5] for more information.

It should be pointed out that the above-mentioned works are concerned only with the forward stochastic systems. The case of systems governed by backward stochastic differential equations with exogenous disturbance signal, to our best knowledge, seems to be open. The objective of this paper is to develop an H_2/H_∞ control theory for infinite horizon backward stochastic systems.

A BSDE is an Itô stochastic differential equation (SDE) for which a random terminal condition on the state has been specified. Since BSDEs are well-defined dynamic systems, it is very natural and appealing to study the control problems involving BSDEs as well as their applications in lots of different fields, especially in finance, economics, insurance, and so forth. Please refer to [6–12] for more details. This paper is concerned with mixed H_2/H_∞ control of backward systems governed by infinite horizon linear BSDEs, namely, an infinite horizon backward stochastic H_2/H_∞ control problem. This means that our purpose is to study mixed H_2/H_∞ backward stochastic control problem in infinite horizon which presents more robust and stable sense in practise. For that, as preliminaries, we first need to review some results on infinite horizon BSDEs in Section 2. Chen and Wang [13] gave an existence and uniqueness result under a kind of Lipschitz condition suitable for one-dimensional infinite horizon BSDEs. Wu [14] generalized the result of [13] into the poisson jump process case in unbounded stopping time duration and obtained the corresponding comparison theorem. In this section, under this frame, we get the existence and uniqueness result for the infinite horizon matrix-valued BSDEs.

In Section 3, similar to the deterministic or stochastic forward case, we formulate the infinite horizon backward stochastic H_2/H_∞ control problem. In Section 4, a necessary and sufficient condition for the existence of a unique solution to the H_2/H_∞ control problem is derived. It is shown that the existence of a unique solution to the control problem is equivalent to the corresponding uncontrolled perturbed system to have a $\mathbb{L}_2$-gain less than or equal to γ and the resulting solution is characterized by the solution of an uncontrolled forward backward stochastic differential equation (FBSDE).

Under some monotone assumptions, Hu and Peng [15] and Peng and Wu [16] obtained the existence and uniqueness results in an arbitrarily prescribed time duration. Wu and Xu [17] gave some comparison theorems for FBSDEs. Riccati equation plays an important role to get the feedback form of the optimal control; please refer to Yong and Zhou [18] for the details. Section 5 gives the equivalent linear feedback solution by virtue of the solution of a Riccati-type equation. As it turns out, the infinite horizon backward stochastic H_2/H_∞ control can no longer be expressed as a linear feedback of the current state like that in deterministic or stochastic forward case. Rather, it depends, in general, on the entire past history of the state pair $(x(\cdot), z(\cdot))$.

2. Notations and Preliminary Results of Infinite Horizon BSDEs

To treat the infinite horizon backward stochastic H_2/H_∞ control problem, we need the following preliminary results of infinite horizon BSDEs.

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a completed filtering probability space; let $(W_t)_{t\geq 0}$ be a standard one-dimensional Wiener process (our assumption that $W(\cdot)$ is scalar-valued is for the sake of simplicity; no essential difficulties are encountered when extending our analysis to the case of vector-valued Wiener process). $\{\mathcal{F}_t\}_{t\geq 0}$ is the natural filtration generated by this Wiener process $W(\cdot)$ up to time t, where $\mathcal{F}_0$ contains all P-null sets of $\mathcal{F}$ and $\mathcal{F}_\infty = \vee_{t\geq 0}\mathcal{F}_t$.

Throughout this paper, we adopt the following conventional notations. S^n: the set of symmetric $n \times n$ matrices with real elements; A^T: the transpose of the matrix A; $A \geq 0$ $(A > 0)$: A is positive semidefinite (positive definite) real matrix; I: identity matrix; $\|x\| := x^T x = (\sum_{i=1}^n |x_i|^2)^{1/2}$ for n-dimensional vector $x = (x_1, \ldots, x_n)^T$; $\|A\| := \max_{x\in R^n, \|x\|=1} \|Ax\|$ for $A \in R^{n\times n}$; $N(\cdot) >$ (≥ 0): $N(t) > (\geq) 0$ for a.s. $t \in R^+$; $M(\cdot) > (\geq) N(\cdot)$: $M(\cdot) - N(\cdot) > (\geq) 0$; X: a given Hilbert space;

$$L^2_{\mathcal{F}}(R^+; X) =: \left\{ f : R^+ \times \Omega \longrightarrow X \text{ is an } \mathcal{F}_t - \text{adapted process such that } \mathbb{E}\int_0^\infty \|f(t)\|^2 dt < \infty \right\};$$

$$\mathcal{S}^2 =: \left\{ v_t,\ 0 \leq t \leq \infty, \text{ is an } \mathcal{F}_t - \text{adapted process such that } \mathbb{E}\left[\sup_{0\leq t\leq\infty} \|v_t\|^2\right] < \infty \right\};$$

$$L^2 =: \left\{ \xi,\ \xi \text{ is a vector-valued } \mathcal{F}_\infty - \text{measurable random variable such that } \mathbb{E}\|\xi\|^2 < \infty \right\}. \tag{1}$$

We consider the infinite horizon BSDE:

$$x_t = \xi + \int_t^\infty f(s, x_s, z_s)\, ds - \int_t^\infty z_s dW_s, \quad t \in [0, \infty]; \tag{2}$$

(x, z) take value in $R^n \times R^n$, $\xi \in L^2$, and f is a map from $\Omega \times [0, \infty] \times R^n \times R^n$ onto R^n which satisfies the following.

(H2.1) For all $(x, z) \in R^n \times R^n$, $f(\cdot, x, z)$ is progressively measurable and

$$\mathbb{E}\left(\int_0^\infty \|f(s, 0, 0)\|\, ds\right)^2 < \infty. \tag{3}$$

(H2.2) There exist two positive deterministic functions $u_1(t)$ and $u_2(t)$ such that, for all $(x_i, z_i) \in R^n \times R^n$, $i = 1, 2$,

$$\begin{aligned} &\|f(t, x_1, z_1) - f(t, x_2, z_2)\| \\ &\quad \leq u_1(t)\|x_1 - x_2\| + u_2(t)\|z_1 - z_2\|, \quad t \in [0, \infty), \\ &\int_0^\infty u_1(t)\, dt < \infty, \qquad \int_0^\infty u_2^2(t)\, dt < \infty. \end{aligned} \tag{4}$$

Then we have the following.

Theorem 1 (see Wu [14]). *There exists a unique solution $(x, z) \in \mathcal{S}^2 \times L^2_{\mathcal{F}}(R^+; R^n)$ satisfying the BSDE (2).*

Let us again consider a function F, which will be in the sequel the generator of the BSDE, defined on $\Omega\times[0,\infty]\times S^n\times S^n$, with values in S^n, such that the process $(F(t, y, z))_{t\in[0,\infty]}$ is a progressively measurable process for each $(y, z) \in S^n \times S^n$.

Along the line of Chen and Wang [13] or Wu [14] combined with that in Peng [19] for matrixed-valued BSDEs result in finite horizon, we get the following existence and uniqueness theorem for infinite horizon matrix-valued BSDEs.

Theorem 2. *Suppose that F satisfies the following.*

(H2.1′) *For all $(y, z) \in S^n \times S^n$, $F(\cdot, y, z)$ is progressively measurable and*

$$\mathbb{E}\left(\int_0^\infty \|F(s, 0, 0)\|\, ds\right)^2 < \infty. \tag{5}$$

(H2.2′) *There exist two positive deterministic functions $u_1(t)$ and $u_2(t)$ such that, for all $(y_i, z_i) \in R^n \times R^n$, $i = 1, 2$,*

$$\begin{aligned} &\|F(t, y_1, z_1) - F(t, y_2, z_2)\| \\ &\quad \leq u_1(t)\|y_1 - y_2\| + u_2(t)\|z_1 - z_2\|, \quad t \in [0, \infty), \end{aligned} \tag{6}$$

and $\int_0^\infty u_1(t)dt < \infty$, $\int_0^\infty u_2^2(t)dt < \infty$, ξ *is a given* S^n*-valued random variable, and* $\xi \in L^2$. *Then, the following matrix-valued infinite horizon BSDE*

$$Y_t = \xi + \int_t^\infty F(s, Y_s, Z_s)\,ds - \int_t^\infty Z_s dW_s \tag{7}$$

admits a unique solution $(Y, Z) \in \mathcal{S}^2 \times L^2_{\mathcal{F}}(R^+; R^n)$.

3. Problem Statement

Now, we consider the following stochastic control system governed by an infinite horizon linear BSDE:

$$\begin{aligned} x(t) = \xi - \int_t^\infty [&A(s)x(s) + B(s)u(s) \\ &+ C(s)d(s) + D(s)z(s)]\,ds \\ &- \int_t^\infty z(s)\,dW(s). \end{aligned} \tag{8}$$

$Z \in \mathcal{R}^{n_Z}$ is the penalty output, and the energy of the output signal Z is given by

$$\|Z\|_2^2 = x_0^T H x_0 + \mathbb{E}\int_0^\infty \left[x_t^T Q_t x_t + z_t^T S_t z_t + u_t^T u_t\right] dt, \tag{9}$$

where H is a nonnegative symmetric constant matrix and $Q_t(\omega)$ and $S_t(\omega)$ are nonnegative symmetric bounded progressively measurable matrix-valued processes. u and d stand for the control input and exogenous disturbance signal, respectively. The energy of the disturbances is

$$\|d\|_2^2 = \mathbb{E}\int_0^\infty d_t^T d_t dt. \tag{10}$$

Later, we will state assumptions on the coefficients $A(\cdot)$, $B(\cdot)$, $C(\cdot)$, $D(\cdot)$, $Q(\cdot)$, $S(\cdot)$ so as to guarantee the existence of a unique solution pair $(x(\cdot), z(\cdot)) \in \mathcal{S}^2 \times L^2_{\mathcal{F}}(R^+; R^n)$ of BSDE (8) for any $u \in L^2_{\mathcal{F}}(R^+; R^{n_u})$, $d \in L^2_{\mathcal{F}}(R^+; R^{n_d})$, and $\xi \in L^2$. We refer to such a four-tuple $(x(\cdot), z(\cdot); u(\cdot), d(\cdot))$ as an admissible triple.

Now, we first define the infinite horizon backward stochastic H_2/H_∞ control as follows.

Definition 3 (backward stochastic H_2/H_∞ control). For given $\gamma > 0$ and $d \in L^2_{\mathcal{F}}(R^+; R^{n_d})$, find, if possible, a control $u = u^* \in L^2_{\mathcal{F}}(R^+; R^{n_u})$, such that

(i) the trajectory of the closed-loop system (8) with $\xi = 0$ satisfies

$$\|Z\|_2^2 \le \gamma^2 \|d\|_2^2, \quad \forall d \ne 0 \in L^2_{\mathcal{F}}(R^+; R^{n_d}) \text{ and} \tag{11}$$

(ii) when the worst case disturbance ([4]) $d^* \in L^2_{\mathcal{F}}(R^+; R^{n_d})$, if existing, is implemented in (8), u^* minimizes the quadratic performance $\|Z\|_2^2$ simultaneously.

If we define

$$\begin{aligned} J_1(u, d) &= \|Z\|_2^2 - \gamma^2 \|d\|_2^2, \\ J_2(u, d) &= \|Z\|_2^2 \end{aligned} \tag{12}$$

then the mixed H_2/H_∞ control problem is equivalent to find the Nash equilibria (u^*, d^*) defined as

$$J_1(u^*, d^*) \ge J_1(u^*, d), \quad \forall d \in L^2_{\mathcal{F}}(R^+; R^{n_d}), \tag{13}$$

$$J_2(u, d^*) \ge J_2(u^*, d^*), \quad \forall u \in L^2_{\mathcal{F}}(R^+; R^{n_u}), \tag{14}$$

$$J_1(u^*, d) \le 0, \quad \forall d \ne 0 \in L^2_{\mathcal{F}}(R^+; R^{n_d}), \quad \xi = 0. \tag{15}$$

Obviously, inequality (15) is associated with the H_∞ performance. The first Nash inequality (13) is to keep that d^* is the worst case disturbance, while the second one (14) is related with the H_2 performance. Clearly, if the Nash equilibria (u^*, d^*) exist and satisfy inequality (15), then u^* is our desired H_2/H_∞ controller and d^* is the worst case disturbance. In this case, we also say that the infinite horizon backward stochastic H_2/H_∞ control admits a solution (u^*, d^*).

Throughout this paper, we assume the following.

(A1) All matrices mentioned in this paper are bounded progressively measurable processes.

(A2)

$$\begin{aligned} &\mathbb{E}\int_0^\infty \|A(t)\|\,dt < \infty, \\ &\mathbb{E}\int_0^\infty \|D(t)\|^2 dt < \infty. \end{aligned} \tag{16}$$

(A3)

$$\begin{gathered} \gamma > 0, \\ Q \ge 0, \qquad \mathbb{E}\int_0^\infty \|Q(t)\|\,dt < \infty, \\ BB^T(\cdot) > \frac{CC^T(\cdot)}{\gamma^2}, \\ \mathbb{E}\int_0^\infty \left\| B(t)B(t)^T - \frac{C(t)C(t)^T}{\gamma^2} \right\| dt < \infty. \end{gathered} \tag{17}$$

From Theorem 2, we obtain that assumption (A2) is sufficient to guarantee the existence of a unique solution pair $(x(\cdot), z(\cdot)) \in \mathcal{S}^2 \times L^2_{\mathcal{F}}(R^+; R^n)$ of BSDE (8) for any $u \in L^2_{\mathcal{F}}(R^+; R^{n_u})$ and $d \in L^2_{\mathcal{F}}(R^+; R^{n_d})$.

4. The Necessary and Sufficient Condition

In this section, we give a necessary and sufficient condition for the existence of a unique solution to the backward stochastic H_2/H_∞ control problem. We begin our presentation with some preliminaries.

Consider the following uncontrolled stochastic perturbed system:

$$dx_t = \left[A(t)x_t + C(t)d_t + D(t)z_t\right]dt + z_t dB(t), \quad x(\infty) = \xi, \quad t \in [0,\infty). \tag{18}$$

Let Z be the to-be-controlled output. For any $0 < T < \infty$, define the perturbation operator $\mathbb{L} : L^2_{\mathcal{F}}(R^+; R^{n_d}) \to L^2_{\mathcal{F}}(R^+; R^{n_Z})$ as

$$\mathbb{L}(d) = Z|_{x_\infty=0}, \quad t \geq 0, \ d \in L^2_{\mathcal{F}}(R^+; R^{n_d}), \tag{19}$$

with its norm

$$\begin{aligned}\|\mathbb{L}\|_2 &:= \sup_{d\in L^2_{\mathcal{F}}(R^+;R^{n_d}),\, d\neq 0,\, x_\infty=0} \frac{\|\mathbb{L}(d)\|_2}{\|d\|_2} \\ &= \sup_{d\in L^2_{\mathcal{F}}(R^+;R^{n_d}),\, d\neq 0,\, x_\infty=0} \frac{\|Z\|_2}{\|d\|_2},\end{aligned} \tag{20}$$

where

$$\|Z\|_2^2 = x_0^T H x_0 + \mathbb{E}\int_0^\infty \left[x_t^T Q_t x_t + y_t^T S_t y_t\right]dt. \tag{21}$$

Obviously, $\mathbb{L}$ is a nonlinear operator.

Definition 4. Let $\gamma > 0$; system (18) is said to have $\mathbb{L}_2$-gain less than or equal to γ if for any nonzero $d \in L^2_{\mathcal{F}}(R^+; R^{n_d})$, $\|\mathbb{L}\|_2 \leq \gamma$.

Proposition 5. *For system* (18) *and given disturbance attenuation $\gamma > 0$, if there exists a function $\mathbf{P}(\cdot)$, satisfying the following SDE (the variables t and ω are suppressed):*

$$\begin{aligned}dP &= \left[-A^T P - PA - Q - \frac{PCC^T P}{\gamma^2}\right]dt - D^T P dW(t), \\ &P + S \leq 0, \quad P(0) = -H, \quad t \in [0,\infty),\end{aligned} \tag{22}$$

then $\|\mathbb{L}\|_2 \leq \gamma$.

Proof. It only needs to note the following identity:

$$\begin{aligned}&\|Z\|_2^2 - \gamma^2\|d\|_2^2 \\ &\quad = \|Z\|_2^2 - \gamma^2\|d\|_2^2 + \mathbb{E}\int_0^\infty d\left(x^T P x\right) - x_0^T H x_0 \\ &\quad = -\gamma^2\left\|v - \frac{C^T P x}{\gamma^2}\right\|_2^2 + \mathbb{E}\int_0^\infty y^T (P+S)\, y\, dt \leq 0.\end{aligned} \tag{23}$$

□

The following theorem is a necessary and sufficient condition for the existence of a unique solution to the infinite horizon backward stochastic H_2/H_∞ control problem.

Theorem 6. *For system* (8)*, the backward stochastic H_2/H_∞ control problem admits a solution if and only if the corresponding uncontrolled system* (18) *has $\mathbb{L}_2$-gain less than or equal to γ. Moreover, if the backward stochastic H_2/H_∞ control problem admits a solution, then the solution is unique with*

$$u^* = \frac{B^T p^*}{2}, \qquad d^* = -\frac{C^T p^*}{2\gamma^2}, \tag{24}$$

where (p^*, x^*, z^*) is the solution of the following FBSDE:

$$\begin{aligned}dp_t^* &= \left[2Qx_t^* - A^T p_t^*\right]dt + \left[2Sz_t^* - D^T p_t^*\right]dB(t), \\ dx_t^* &= \left[Ax_t^* + \frac{BB^T p_t^*}{2} - \frac{CC^T p_t^*}{2\gamma^2} + Dz_t^*\right]dt + z_t^* dB(t), \\ p_0^* &= 2Hx_0^*, \quad x_\infty^* = \xi, \quad t \in [0,\infty).\end{aligned} \tag{25}$$

Proof. (1) The Sufficient Condition. To show that the backward stochastic H_2/H_∞ control problem admits a unique solution (u^*, d^*) if the corresponding uncontrolled system (18) has $\mathbb{L}_2$-gain less than or equal to γ, we will show that (u^*, d^*) is a solution firstly.

Look at the above FBSDE; from [16], the FBSDE (25) has a unique solution (p_t^*, x_t^*, z_t^*). Now, we try to prove that d^* is the worst case disturbance. For any given $d \in L^2_{\mathcal{F}}(R^+; R^{n_d})$, suppose that x^d is the trajectory corresponding to $(u^*, d) \in L^2_{\mathcal{F}}(R^+; R^{n_u}) \times L^2_{\mathcal{F}}(R^+; R^{n_d})$. It is easy to see the trajectory corresponding to

$$(0, d^* - d) \tag{26}$$

is $x^* - x^d$ with initial state $x_0^* - x_0^d$ and terminal state 0. Hence, $(x^* - x^d, z^* - z^d)$ is the solution corresponding to $d^* - d$ for system (18) with terminal state 0. Since system (18) has $\mathbb{L}_2$-gain less than or equal to γ, then

$$\begin{aligned}&\mathbb{E}\left[\int_0^\infty \left[-\left(x^* - x^d\right)^T Q\left(x^* - x^d\right) - \left(z^* - z^d\right)^T S\left(z^* - z^d\right) \right.\right. \\ &\qquad \left.\left. + \gamma^2 (d^* - d)^T (d^* - d)\right]dt\right] \\ &\qquad - \left(x_0^* - x_0^d\right)^T H\left(x_0^* - x_0^d\right) \geq 0,\end{aligned} \tag{27}$$

$$\begin{aligned}&J_2(u^*, d^*) - J_2(u^*, d) \\ &\quad = \mathbb{E}\left[\int_0^\infty \left[x^{*T} Q x^* - x^{dT} Q x^d + z^{*T} S z^* - z^{dT} S z^d \right.\right. \\ &\qquad \left.\left. - \gamma^2 d^{*T} d^* + \gamma^2 d^T d\right]dt\right] \\ &\qquad + x_0^{*T} H x_0^* - x_0^{dT} H x_0^d\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[\int_0^\infty \left[\left(x^* - x^d\right)^T Q\left(x^* - x^d\right)\right.\right. \\
&\qquad - 2x^{dT} Q\left(x^d - x^*\right) + \left(z^* - z^d\right)^T S\left(z^* - z^d\right) \\
&\qquad \left.\left. -2z^{dT} S\left(z^d - z^*\right) - \gamma^2 d^{*T} d^* + \gamma^2 d^T d\right] dt\right] \\
&\quad + \left(x_0^* - x_0^d\right)^T H\left(x_0^* - x_0^d\right) - 2x_0^{dT} H\left(x_0^d - x_0^*\right).
\end{aligned} \tag{28}
$$

Applying Itô's formula to $p^{*T}(x^d - x^*)$,

$$
\begin{aligned}
&- 2x_0^{*T} H\left(x_0^d - x_0^*\right) \\
&= \mathbb{E}\left[\int_0^\infty d\left[p^{*T}\left(x^d - x^*\right)\right]\right] \\
&= \mathbb{E}\int_0^\infty \left[2x^{*T} Q\left(x^d - x^*\right)\right. \\
&\qquad \left. +2z^{*T} S\left(z^d - z^*\right) + \left(C^T p^*\right)^T (d - d^*)\right] dt\Big] \\
&= \mathbb{E}\left[\int_0^\infty \left[2x^{*T} Q\left(x^d - x^*\right)\right.\right. \\
&\qquad \left.\left. +2z^{*T} S\left(z^d - z^*\right) - 2\gamma^2 v^{*T} (d - d^*)\right] dt\right].
\end{aligned} \tag{29}
$$

Substituting $2x_0^{*T} H(x_0^d - x_0^*)$ into (28), we get

$$
\begin{aligned}
&J_2\left(u^*, d^*\right) - J_2\left(u^*, d\right) \\
&\quad = \mathbb{E}\left[\int_0^\infty \left[-\left(x^* - x^d\right)^T Q\left(x^* - x^d\right)\right.\right. \\
&\qquad - \left(z^* - z^d\right)^T S\left(z^* - z^d\right) \\
&\qquad \left.\left. + \gamma^2 (d^* - d)^T (d^* - d)\right] dt\right] \\
&\qquad - \left(x_0^* - x_0^d\right)^T H\left(x_0^* - x_0^d\right).
\end{aligned} \tag{30}
$$

From (27), then

$$
J_2\left(u^*, d^*\right) - J_2\left(u^*, d\right) \geq 0. \tag{31}
$$

So d^* is the worst case disturbance. Moreover, for $x_\infty = 0$, the FBSDE (25) admits a unique solution $(p^*, x^*, z^*) = (0, 0, 0)$; then

$$
J_2\left(u^*, d\right) \leq J_2\left(u^*, d^*\right) = 0. \tag{32}
$$

Hence, u^* restrains the exogenous disturbance. In the following, we will show that u^* also minimizes that cost function when the worst case disturbance d^* is implemented into system (8).

For any $u \in L^2_{\mathcal{F}}(\mathbb{R}^{n_u})$, let x_t^u be the trajectory of the system (8) corresponding to (u, d^*). Let us first consider

$$
J_1\left(u^*, d^*\right) - J_1\left(u, d^*\right) = I_1, \tag{33}
$$

where

$$
\begin{aligned}
I_1 &= -\mathbb{E}\left[\int_0^\infty \left[x^{*T} Q(t) x^* - x^{uT} Q x^u + z^{*T} S(t) z^*\right.\right. \\
&\qquad \left.\left. - z^{uT} S z^u + u^{*T} u^* - u^T u\right] dt\right] \\
&\quad + x_0^{*T} H x_0^* - x_0^{uT} H x_0^u \\
&= \mathbb{E}\left[\int_0^\infty \left[(x^* - x^u)^T Q (x^* - x^u)\right.\right. \\
&\qquad + (z^* - z^u)^T S (z^* - z^u) + (u^* - u)^T (u^* - u) \\
&\qquad + 2x^{*T} Q(t) (x^u - x^*) + 2y^{*T} S(t) (z^u - z^*) \\
&\qquad \left.\left. + 2u^{*T} (u - u^*)\right] dt\right] \\
&\quad + (x_0^* - x_0^u)^T H (x_0^* - x_0^u) + 2x_0^{*T} H (x_0^u - x_0^*).
\end{aligned} \tag{34}
$$

From $p_0^* = 2Hx_0^*$, we use Itô's formula to $p_t^{*T}(x_t^u - x_t^*)$ and get

$$
\begin{aligned}
&2x_0^{*T} H (x_0^u - x_0^*) \\
&\quad = -\mathbb{E}\left[\int_0^\infty \left[2x^{*T} Q(t) (x^u - x^*)\right.\right. \\
&\qquad \left.\left. + 2z^{*T} S (z^u - z^*) + 2u^{*T} (u - u^*)\right] dt\right].
\end{aligned} \tag{35}
$$

Then because of Q, S, and H being nonnegative, we have

$$
J_1\left(u^*, v^*\right) - J_1\left(u, v^*\right) = I_1 \geq 0. \tag{36}
$$

Therefore, d^* minimizes the cost function when the worst case disturbance d^* is implemented into system (8).

So, $(u^*, d^*) = (B^T p^*/2, -C^T p^*/2\gamma^2)$ is a solution of the backward stochastic H_2/H_∞ control problem.

We are now in a position to prove the uniqueness of the solution. Assume that the backward stochastic H_2/H_∞ control has a solution (u^1, d^1), (x^1, z^1) is the corresponding solution for (8), and p^1 is the solution of the following BSDE:

$$
\begin{aligned}
dp^1 &= \left[2Qx^1 - A^T p^1\right] dt + \left[2Sz^1 - D^T p^1\right] dB(t), \\
p_0^1 &= 2Hx_0^1.
\end{aligned} \tag{37}
$$

Implementing d^1, having

$$
\inf_{u \in L^2_{\mathcal{F}}(R^{n_u})} J_1\left(u, d^1\right), \tag{38}
$$

is a standard LQ optimal control problem. By uniqueness, $u^1 = B^T p^1/2$.

Let x be the trajectory corresponding to $(u^1, d) = (u^1, -C^T p^1/\gamma^2)$; then

$$\begin{aligned}
0 &\geq J_2(u^1, d) - J_2(u^1, d^1) \\
&= \mathbb{E}\left[\int_0^\infty \left[x^T Q x - x^{1T} Q x^1 + z^T S z - z^{1T} S z^1 \right.\right. \\
&\qquad \left.\left. -\gamma^2 d^T d + \gamma^2 d^{1T} d^1\right] dt\right] \\
&\quad + x_0^T H x_0 - x_0^{1T} H x_0^1 \\
&= \mathbb{E}\left[\int_0^\infty \left[(x^1 - x)^T Q (x^1 - x) \right.\right. \\
&\qquad - 2x^{1T} Q (x^1 - x) + (z^1 - z)^T S (z^1 - z) \\
&\qquad \left.\left. -2z^{1T} S (z^1 - z) + \gamma^2 d^{1T} d^1 - \gamma^2 d^T d\right] dt\right] \\
&\quad + (x_0^1 - x_0)^T H (x_0^1 - x_0) - 2x_0^{1T} H (x_0^1 - x_0).
\end{aligned} \tag{39}$$

Applying Itô's formula to $p^{1T}(x^1 - x)$,

$$\begin{aligned}
&2x_0^{1T} H (x_0^1 - x_0) \\
&\quad = \mathbb{E}\int_0^\infty d\left[p^{1T}(x^1 - x)\right] \\
&\quad = \mathbb{E}\left[\int_0^\infty \left[2x^{1T} Q (x^1 - x) \right.\right. \\
&\qquad \left.\left. + (C^T z^1)^T (d^1 - d) + 2z^{1T} S (z^1 - z)\right] dt\right] \\
&\quad = \mathbb{E}\left[\int_0^\infty \left[2x^{1T} Q (x^1 - x) \right.\right. \\
&\qquad \left.\left. + 2z^{1T} S (z^1 - z) - 2\gamma^2 d^T (d^1 - d)\right] dt\right].
\end{aligned} \tag{40}$$

Substituting $-2x_0^{1T} H (x_0^1 - x_0)$ into (39), then

$$\begin{aligned}
0 &\geq J_2(\gamma, x_0; u^1, d) - J_2(\gamma, x_0; u^1, d^1) \\
&= \mathbb{E}\left[\int_0^\infty \left[(x^1 - x)^T Q (x^1 - x) \right.\right. \\
&\qquad + (z^1 - z)^T S (z^1 - z) \\
&\qquad \left.\left. + \gamma^2 (z - z^1)^T (z - z^1)\right] dt\right].
\end{aligned} \tag{41}$$

Because of Q, S, and M being nonnegative, we get $d^1 = v = C^T z^1/\gamma^2$.

Therefore, $(u^1, d^1) = (u^*, d^*)$.

(2) The Necessary Condition. Here we assume that a solution exists; then from the uniqueness of the solution, we get that (u^*, d^*) is the unique solution and we will show that system (18) has $\mathbb{L}_2$-gain less than or equal to γ.

For $x_T = 0$, the FBSDE (25) has a unique solution $(p^*, x^*, z^*) = (0, 0, 0)$; then $(u^*, d^*) = (0, 0)$ and

$$J_2(u^*, d) \leq J_2(u^*, d^*) = 0, \quad \forall d \in L^2_{\mathcal{F}}(\mathcal{R}^{n_d}). \tag{42}$$

Therefore, system (18) has $\mathbb{L}_2$-gain less than or equal to γ. □

5. The Linear Feedback Solution

The main result of this section gives the equivalent linear feedback solution. For the purpose of this section the coefficients A_t, B_t, C_t, D_t, E_t, Q_t, and S_t are assumed deterministic functions; (18) has $\mathbb{L}_2$-gain less than or equal to γ.

Let (p, x, z) be the solution of (25); we first give the relations between p, x, and z using the undetermined coefficients method. Now, we introduce the following generalized matrix-valued Riccati equation (the variables t are suppressed):

$$\begin{aligned}
&\dot{K} - AK - KA^T + 2KQK - \frac{BB^T}{2} \\
&\quad + \frac{CC^T}{2\gamma^2} + D(I - 2KS)^{-1} K D^T = 0, \qquad K(\infty) = 0.
\end{aligned} \tag{43}$$

Similar to the line developed by Lim and Zhou [6], we can prove that (43) admits a unique solution $K(\cdot)$. Letting $K(\cdot)$ be the solution to (43), we define the following equations:

$$\begin{aligned}
&dh = \left[Ah - 2KQh + D(I - 2KS)^{-1}\eta\right] dt + \eta dB(t), \\
&h(\infty) = \xi.
\end{aligned} \tag{44}$$

Equation (44) is a linear BSDE and admits a unique solution (h, η).

Theorem 7. *Suppose that $(p(\cdot), x(\cdot), z(\cdot))$, $K(\cdot)$, and $(h(\cdot), \eta(\cdot))$ are the solutions of (25), (43), and (44), respectively; then the following relations are satisfied:*

$$\begin{aligned}
&x(t) = K(t) p(t) + h(t), \\
&z(t) = (I - 2K(t)S(t))^{-1} \left(\eta(t) - K(t) D(t)^T p(t)\right), \\
&x(0) = (I - 2K(0)H)^{-1} h(0).
\end{aligned} \tag{45}$$

Proof. Let $x(t) = K(t)p(t) + h(t)$. We apply Itô's formula to $x(t)$, $K(t)p(t) + h(t)$, respectively, and it is easy to check that $K(t)$ and $h(t)$ satisfy (43) and (44), respectively. □

From Theorem 7, we know that $x(\cdot)$ can be written to the functions of $K(\cdot)$, $p(\cdot)$, and $h(\cdot)$. Now we would like to derive the feedback solution using the undetermined coefficients method. First, we introduce the generalized matrix-valued Riccati equation and a linear SDE:

$$\begin{aligned}
&\dot{\Sigma} + \Sigma A + A^T \Sigma \\
&\quad + \Sigma \left[\frac{BB^T}{2} - \frac{CC^T}{2\gamma^2} - D(I - 2KS)^{-1} KD^T \right] \\
&\quad \times \Sigma - 2Q = 0, \\
&\Sigma(0) = 2H,
\end{aligned} \tag{46}$$

$$\begin{aligned}
dr &= \left[-A^T r - \frac{\Sigma BB^T r}{2} + \frac{\Sigma CC^T r}{2\gamma^2} + \Sigma D(I - 2KS)^{-1} KD^T r \right. \\
&\quad \left. - \Sigma D(I - 2KS)^{-1} \eta \right] dt \\
&\quad + \left[(2S - \Sigma)(I - 2KS)^{-1} \left[\eta - KD^T (I - \Sigma K)^{-1} (\Sigma h + r) \right] \right. \\
&\quad \left. - D^T (I - \Sigma K)^{-1} (\Sigma h + r) \right] dB(t), \\
r(0) &= 0.
\end{aligned} \tag{47}$$

Similar to the line developed by Lim and Zhou [6], we can prove that (46) admits a unique solution $\Sigma(\cdot)$. Equation (47) is a linear SDE and has a unique solution $r(\cdot)$.

Theorem 8. *The backward stochastic H_2/H_∞ control problem has a feedback solution (u^*, v^*),*

$$u^* = \frac{B^T(\Sigma x + r)}{2}, \qquad v^* = -\frac{C^T(\Sigma x + r)}{2\gamma^2}. \tag{48}$$

Proof. Let $p(t) = \Sigma(t)x(t) + r(t)$. We apply Itô's formula to $p(t)$ and $\Sigma(t)x(t) + r(t)$, respectively, and it is easy to check that $\Sigma(t)$ and $r(t)$ satisfy (46) and (47). □

Remark 9. From Theorem 8, we see that the solution involves an additional random nonhomogeneous term $r(\cdot)$. This addition disqualifies (48) from a feedback control of the current state, contrary to the deterministic or stochastic forward H_2/H_∞ (see [4, 5]) cases. The reason is because $r(\cdot)$ depends on $(h(\cdot), \eta(\cdot))$, which in turn depends on ξ, the terminal condition of part of the state variable, $x(\cdot)$. This is one of the major distinctive features of the backward stochastic H_2/H_∞ problem.

Finally, it is important to recognize that the expressions for the backward stochastic H_2/H_∞ control, as presented in Theorems 6 and 8, are equivalent expressions of the same process; that is, this does not contradict the uniqueness of the solution.

We present an example to illustrate the above theoretical results as follows.

Example 10. Consider the backward stochastic H_2/H_∞ control problem of the following one-dimensional system:

$$\begin{aligned}
x(t) &= \xi - \int_t^\infty \left[2e^{-s} x(s) + \sqrt{4e^{-s} + 2}\, u(s) \right. \\
&\qquad \left. + \sqrt{e^{-s} + 1}\, d(s) + 2e^{-s/2} z(s) \right] ds \\
&\quad - \int_t^\infty z(s)\, dW(s),
\end{aligned} \tag{49}$$

with controlled output energy

$$\|Z\|_2^2 = \mathbb{E} \int_0^\infty \left[\frac{e^{-t} x_t^2}{2} + u_t^2 \right] dt. \tag{50}$$

If we take $\gamma = \sqrt{2}/2$, then the Riccati equation (43) specializes to

$$K(t) = \int_t^\infty \left[e^{-s} \left(K(s)^2 - 1 \right) \right] ds. \tag{51}$$

Solving it yields $K(t) = (1 - e^{2e^{-t}})/(1 + e^{2e^{-t}})$. Equation (44) specializes to

$$\begin{aligned}
h(t) &= \xi - \int_t^\infty \left[\left\{ 2e^{-s} - e^{-s} \frac{1 - e^{2e^{-s}}}{1 + e^{2e^{-s}}} \right\} h(s) + 2e^{-s/2} \eta(t) \right] dt \\
&\quad - \int_t^\infty \eta(t)\, dW(t).
\end{aligned} \tag{52}$$

Then, from Theorem 7, we get a unique solution

$$\begin{aligned}
(u^*, d^*) &= \left(\sqrt{\frac{2e^{-t} + 1}{2}} \cdot \frac{1 + e^{2e^{-t}}}{1 - e^{2e^{-t}}} \cdot [x(t) - h(t)], \right. \\
&\qquad \left. - \sqrt{e^{-t} + 1} \cdot \frac{1 + e^{2e^{-t}}}{1 - e^{2e^{-t}}} \cdot [x(t) - h(t)] \right),
\end{aligned} \tag{53}$$

of the backward H_2/H_∞ control problem.

Conflict of Interests

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

Acknowledgments

This work was supported by the National Natural Science Foundation (11221061 and 61174092), 111 project (B12023), the National Science Fund for Distinguished Young Scholars of China (no. 11125102), the Science and Technology Project of Shandong Province (2013GRC32201), and the Doctoral Foundation of University of Jinan (XBS1213).

References

[1] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 34, no. 8, pp. 831–847, 1989.

[2] D. Hinrichsen and A. J. Pritchard, "Stochastic H_∞," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1504–1538, 1998.

[3] G. Zames, "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 26, no. 2, pp. 301–320, 1981.

[4] B.-S. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[5] D. J. N. Limebeer, B. D. O. Anderson, and B. Hendel, "A Nash game approach to mixed H_2/H_∞ control," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 39, no. 1, pp. 69–82, 1994.

[6] A. E. B. Lim and X. Y. Zhou, "Linear-quadratic control of backward stochastic differential equations," *SIAM Journal on Control and Optimization*, vol. 40, no. 2, pp. 450–474, 2001.

[7] J.-M. Bismut, "An introductory approach to duality in optimal stochastic control," *SIAM Review*, vol. 20, no. 1, pp. 62–78, 1978.

[8] D. Duffie and L. G. Epstein, "Stochastic differential utility," *Econometrica*, vol. 60, no. 2, pp. 353–394, 1992.

[9] N. El Karoui, S. Peng, and M. C. Quenez, "Backward stochastic differential equations in finance," *Mathematical Finance*, vol. 7, no. 1, pp. 1–71, 1997.

[10] É. Pardoux and S. G. Peng, "Adapted solution of a backward stochastic differential equation," *Systems & Control Letters*, vol. 14, no. 1, pp. 55–61, 1990.

[11] G. Wang and Z. Yu, "A Pontryagin's maximum principle for non-zero sum differential games of BSDEs with applications," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 55, no. 7, pp. 1742–1747, 2010.

[12] G. Wang and Z. Yu, "A partial information non-zero sum differential game of backward stochastic differential equations with applications," *Automatica*, vol. 48, no. 2, pp. 342–352, 2012.

[13] Z. Chen and B. Wang, "Infinite time interval BSDEs and the convergence of g-martingales," *Australian Mathematical Society. Journal A: Pure Mathematics and Statistics*, vol. 69, no. 2, pp. 187–211, 2000.

[14] Z. Wu, "Fully coupled FBSDE with Brownian motion and Poisson process in stopping time duration," *Journal of the Australian Mathematical Society*, vol. 74, no. 2, pp. 249–266, 2003.

[15] Y. Hu and S. Peng, "Solution of forward-backward stochastic differential equations," *Probability Theory and Related Fields*, vol. 103, no. 2, pp. 273–283, 1995.

[16] S. Peng and Z. Wu, "Fully coupled forward-backward stochastic differential equations and applications to optimal control," *SIAM Journal on Control and Optimization*, vol. 37, no. 3, pp. 825–843, 1999.

[17] Z. Wu and M. Xu, "Comparison theorems for forward backward SDEs," *Statistics & Probability Letters*, vol. 79, no. 4, pp. 426–435, 2009.

[18] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, vol. 43, Springer, New York, NY, USA, 1999.

[19] S. Peng, "Backward stochastic differential equations and applications to optimal control," *Applied Mathematics and Optimization*, vol. 27, no. 2, pp. 125–144, 1993.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 635758, 7 pages
http://dx.doi.org/10.1155/2014/635758

Research Article

The H_∞ Control for Bilinear Systems with Poisson Jumps

Rui Zhang,[1] Weihai Zhang,[2] and Xiangyun Lin[3]

[1] *College of Electronic Communication and Physics, Shandong University of Science and Technology, Qingdao 266590, China*
[2] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China*
[3] *College of Mathematics and System Science, Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Xiangyun Lin; lxy9393@sina.com

Received 3 April 2014; Accepted 5 May 2014; Published 21 May 2014

Academic Editor: Xuejun Xie

This paper discusses the state feedback H_∞ control problem for a class of bilinear stochastic systems driven by both Brownian motion and Poisson jumps. By completing square method, we obtain the H_∞ control by solutions of the corresponding Hamilton-Jacobi equations (HJE). By the tensor power series method, we also shift such HJEs into a kind of Riccati equations, and the H_∞ control is represented with the form of tensor power series.

1. Introduction

The main purpose of H_∞ control design is to find the law to efficiently eliminate the effect of the disturbance [1, 2]. Theoretically, study of H_∞ control first starts from the deterministic linear systems, and the derivation of the state-space formulation of the standard H_∞ control leads to a breakthrough, which can be found in the paper [3]. In recent years, stochastic H_∞ control systems, such as Markovian jump systems [4–6], H_∞ Gaussian control design [7], and Itô differential systems [8–13], have received a great deal of attention. However, up to now, most of the work on stochastic H_∞ control is confined to Itô type or Markovian jump systems. Yet, there are still many systems which contain Poisson jumps in economics and natural science. In 1970s, Boel and Varaiya [14] and Rishel [15] considered the optimal control problem with random Poisson jumps, and many basic results have been made. From then on, many scholars and economists also study the system and its applications; for further reference, we refer to [16–20] and their references. But those results mostly concentrate on optimal control and its application in financial market or corresponding theories. Of course, such model still can be disturbed by exogenous disturbance and its robustness is also an important problem. The objective of this paper is to develop an H_∞-type theory over infinite time horizon for the disturbance attenuation of stochastic bilinear systems with Poisson jumps by dynamic state feedback.

Generally, the key of H_∞ control design is to solve a general Hamilton-Jacobi equation (HJE). However, up to now, there is still no effective algorithm to solve such a general HJE. In order to solve the HJE given in this paper, we extend a tensor power series approach which is used in [21] and also give the simulation of the trajectory of output z under H_∞ control. This paper will follow along the lines of [22] to study the stochastic H_∞ control with infinite horizons and finite horizon for a class of nonlinear stochastic differential systems with Poisson jumps. The paper is organized as follows.

In Section 2, we review Itô's theories about the system driven by Brownian motion and Poisson jumps. In Section 3, we obtain the H_∞ by solving the HJE which is proved by the completing square method. In Section 4, we discuss the problem of finite horizon H_∞ control with jumps, and using the tensor power series approach, we discuss the approximating H_∞ control given in the paper. For convenience, we adopt the following notation.

$\mathcal{S}_n(\mathcal{R})$ denotes the set of all real $n\times n$ symmetric matrices; A' is the transpose of the corresponding matrix A; $A > 0$ $(A \geq 0)$ is the positive definite (semidefinite) matrix A; I is the identity matrix; $\mathbb{E}\xi$ is the expectation of random variable ξ; $\|x\|$ is the Euclidean norm of vector $x \in \mathcal{R}^{n_x}$ and n_x is the dimension of x; $\mathcal{L}^2([0,T],\mathcal{R}^{n_y})$ is the set of n_y-dimensional stochastic process y defined on interval $[0,T]$ (T can take ∞), taking values in $\mathscr{R}^{n_y}$, with norm

$$\|y\|_{\mathscr{L}^2([0,T],\mathscr{R}^{n_y})} = \left(E\int_0^T \|y(t)\|^2 dt\right)^{1/2} < \infty; \quad (1)$$

$C^{1,2}(\mathscr{R}_+,\mathscr{R}^{n_x})$ is the class of function $V(t,x)$ twice continuously differential with respect to $x \in \mathscr{R}^{n_x}$ and once continuously differential with respect to t; $\langle x,y\rangle$ is the inner product of two vectors $x, y \in \mathscr{R}^n$.

2. Preliminaries

For a given complete probability space $(\Omega,\mathscr{F},\mathbb{P})$, let W_t and μ be the Brownian motion and the Poisson random measure, respectively, which are mutually independent:

(i) a 1-dimensional standard Brownian motion $\{W_t\}_{t\geq 0}$;

(ii) a Poisson random measure μ on $\mathbb{R}_+\times E$, where $E \subset \mathbb{R}^l$ is a nonempty open set equipped with its Borel field $\mathscr{B}(E)$, with the compensator $\widehat{\mu}(de,dt) = \lambda(de)dt$, such that $\{\widetilde{\mu}((0,t]\times A) = (\mu-\widehat{\mu})((0,t]\times A)\}_{t\geq 0}$ is a martingale for all $A \in \mathscr{B}(E)$ satisfying $\lambda(A) < \infty$. Here λ is an arbitrarily given σ-finite Lévy measure on $(E,\mathscr{B}(E))$, that is, a measure on $(E,\mathscr{B}(E))$ with the property that $\int_E (1\wedge |e|^2)\lambda(de) < \infty$. We also let

$$\begin{aligned}\mathscr{F}_t = \sigma&\left[\iint\int_{(0,s]\times A}\widetilde{\mu}(de,ds) : s\leq t, A\in\mathscr{B}(E)\right]\vee\sigma\\ &\times[W_s : s\leq t]\vee\mathscr{N},\end{aligned} \quad (2)$$

where $\mathscr{N}$ denotes the totality of P-null sets.

In order to discuss the systems driven by Brownian motion and Poisson jumps, we first review the theorem about Itô's formula for such stochastic processes.

Theorem 1. *Let M_t be a square integral continuous martingale; A_t is a continuous adapted process with finite variance. $\gamma(s,e)$ is locally square integral due to μ and P; $x(t)$ satisfies the following Itô type stochastic process:*

$$x(t) = x(0) + M_t + A_t + \int_0^t\int_E \gamma(s,e)\,\widetilde{\mu}(de,ds). \quad (3)$$

Then for $F(t,x) \in C^{1,2}(\mathscr{R}_+,\mathscr{R}^{n_x})$, we have (see [23] Chapter I, §3, Theorem 11)

$$\begin{aligned}dF(t,x_t) &= F_t(t,x_t)\,dt + F_x(t,x_t)\,d(M_t+A_t)\\ &\quad + \frac{1}{2}F_{xx}(t,x_t)\,d\langle M\rangle_t + \int_E\left[F(t,x_t+\gamma(t,e)) - F(t,x_t)\right.\\ &\qquad\left. - F_x(t,x_t)\,\gamma(t,e)\,\right]\lambda(de)\,dt\\ &\quad + \int_E\left[F(t,x_t+\gamma(t,e)) - F(t,x_t)\right]\widetilde{\mu}(de,dt),\end{aligned} \quad (4)$$

where $\langle M\rangle$ denotes the predictable compensator of martingale M.

In the paper, for convenience, x_t is shorten as x. Furthermore, for $F \in C^{1,2}(\mathscr{R}_+,\mathscr{R}^{n_x})$, if using Itô formula to $F(t,x_t)$ and integrating from s to t $(0\leq s<t)$, then taking expectation with both sides

$$\begin{aligned}&\mathbb{E}F(t,x_t) - \mathbb{E}F(s,x_s)\\ &\quad= \int_s^t \mathbb{E}F_t(r,x_r)\,dr\\ &\qquad + \frac{1}{2}\int_s^t \mathbb{E}\left[\sigma(r,x_r)'F_{xx}(r,x_r)\,\sigma(r,x_r)\right]dr\\ &\qquad + \int_s^t \mathbb{E}\left\{\int_E\left[F(r,x_r+\gamma(r,x_{r-},e))\right.\right.\\ &\qquad\qquad \left. - F(r,x_r) - \langle F_x(r,x_r),\gamma(r,e)\rangle\right]\\ &\qquad\qquad \left.\times\lambda(de)\right\}dr,\end{aligned} \quad (5)$$

we can see that $\mathbb{E}F(t,x_t)$ is continuous with respect to time t. Since we mainly use the results of expectations of some well functions on x_t and those expectations are continuous with respect to time t, so, for briefness, in the rest of this paper the sign x_{t-} under integration $\int_E$ is also shortened as x.

3. The H_∞ Control for Bilinear Systems with Jumps

We consider the following bilinear system driven by Poisson jumps:

$$\begin{aligned}dx &= (Ax + Bxu + Kv)\,dt + CxdW + \int_E G(e)\,x\widetilde{\mu}(de,dt),\\ z &= \begin{bmatrix} Mx \\ u\end{bmatrix},\end{aligned} \quad (6)$$

where $v \in \mathscr{L}^2([0,T],\mathscr{R}^{n_d})$ represents the exogenous disturbance, A, B, and C are constant $n_x\times n_x$ matrices, $K \in \mathscr{R}K \in \mathscr{R}^{n_x\times n_d}$, and $G(e) \in \mathscr{R}^{n_x\times n_x}$ only depends on e. If there exists an $u_T^* \in \mathscr{L}^2([0,T],\mathscr{R}^{n_u})$ such that for any given $\gamma > 0$ and all $v \in \mathscr{L}^2([0,T],\mathscr{R}^{n_d})$, $x(0) = 0$, the closed-loop system satisfies

$$\|z\|_{\mathscr{L}^2([0,T],\mathscr{R}^{n_z})} \leq \gamma\|v\|_{\mathscr{L}^2([0,T],\mathscr{R}^{n_d})}, \quad (7)$$

we call u_T^* the H_∞ control of (6).

Theorem 2. *Suppose there exists a nonnegative solution $V \in C^{1,2}([0,T],\mathscr{R}^{n_x})$ to the HJE*

$$\begin{aligned}\mathscr{H}^1_T(V_T(t,x)) :=& \frac{\partial V_T}{\partial t}+\frac{\partial V_T}{\partial x}'Ax+\frac{1}{2}x'M'Mx\\&+\frac{1}{2\gamma^2}\frac{\partial V_T}{\partial x}'KK'\frac{\partial V_T}{\partial x}\\&+\frac{1}{2}x'C'\frac{\partial^2 V_T}{\partial x^2}Cx-\frac{1}{2}\frac{\partial V_T}{\partial x}'Bxx'B'\frac{\partial V_T}{\partial x}\\&+\int_E\Big[V_T(t,x+G(e)x)-V_T(t,x)\\&-\frac{\partial V}{\partial x}'G(e)x\Big]\lambda(de)=0,\\ V_T(T,x)=0,\quad & V_T(t,0)=0,\quad \forall(t,x)\in[0,T]\times\mathscr{R}^{n_x}.\end{aligned}\tag{8}$$

Then $u^*_T=-x'B'(\partial V_T/\partial x)$ is an H_∞ control for system (6).

Proof. Applying Itô's formula to $V(t,x)$, we have

$$\begin{aligned}&V(T,x_T)-V(0,0)\\&=\int_0^T\Big\{V(t,x)+V'_x(t,x)(Ax+Bxu+Kv)\\&\quad+\frac{1}{2}x'C'V_{xx}Cx\\&\quad+\int_E[V(t,x+G(e)x)-V(t,x)\\&\quad-V_x(t,x)G(e)x]\lambda(de)\Big\}dt\\&\quad+\int_0^T V'_x(t,x)CxdW_t\\&\quad+\int_0^T[V(t,x+G(e)x)-V(t,x)]\tilde{\mu}(de,dt).\end{aligned}\tag{9}$$

Taking expectation with both sides and applying $V(T,x)=0$ and $V(t,0)=0$, we obtain

$$\begin{aligned}0=&\int_0^T\mathbb{E}\Big\{V(t,x)+V'_x(t,x)(Ax+Bxu+Kv)\\&+\frac{1}{2}x'C'V_{xx}Cx\\&+\int_E[V(t,x+G(e)x)-V(t,x)\\&-V_x(t,x)G(e)x]\lambda(de)\\&+x'_tM'Mx+u'_tu-\gamma^2v'_tv\Big\}dt\\&-\|z\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_z})}+\gamma^2\|v\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_d})}.\end{aligned}\tag{10}$$

Completing square for u and v, respectively, we have

$$\begin{aligned}0=&\int_0^T\mathbb{E}\left\{\mathscr{H}^1_T(V(t,x))+\|u-u^*_T\|^2-\frac{1}{\gamma^2}\|v-v^*_T\|^2\right\}dt\\&-\|z\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_z})}+\gamma^2\|v\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_d})},\end{aligned}\tag{11}$$

where

$$u^*_T=-x'B'\frac{\partial V_T}{\partial x},\qquad v^*_T=\frac{1}{\gamma^2}\left(K'\frac{\partial V_T}{\partial x}+K'\frac{\partial^2 V_T}{\partial x^2}Cx\right).\tag{12}$$

By HJE (8) and let $u=u^*_T$, we have

$$\begin{aligned}&\|z\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_z})}-\gamma^2\|v\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_d})}\\&\quad=-\frac{1}{\gamma^2}\int_0^T\mathbb{E}\left\{\|v-v^*_T\|^2\right\}dt.\end{aligned}\tag{13}$$

So, the following inequality is true:

$$\|z\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_z})}\le\gamma^2\|v\|^2_{\mathscr{L}^2([0,T],\mathscr{R}^{n_d})}.\tag{14}$$

This proves that u^*_T is an H_∞ control for system (6). □

Remark 3. From the proof of Theorems 2 and (13), we can see that (u^*_T, v^*_T) given by (12) is a saddle point for the following stochastic game problem:

$$\min_{u\in\mathscr{L}^2([0,T],\mathscr{R}^{n_u})}\ \max_{d\in\mathscr{L}^2([0,T],\mathscr{R}^{n_d})}\mathbb{E}\int_0^T\left(\|z\|^2-\gamma^2\|d\|^2\right)dt.\tag{15}$$

4. The Tensor Power Series Representation of H_∞ Control

Generally speaking, it is very hard to solve HJE (8). Now we use an approximation algorithm which is called tensor power series approach to solve a special case of HJE (8). In what follows, suppose $V_T(t,x)$ satisfying (8) has the following form:

$$V_T(t,x)=\sum_{i=1}^{\infty}\langle\otimes_i x, P_i(t)\otimes_i x\rangle,\tag{16}$$

where $P_i(t)$, $i\ge1$, are symmetrically and continuously differential matrix-valued functions on $[0,T]$, $\otimes$ is the Kronecker product of matrix (or vectors), and $\otimes_i x=x\otimes\cdots\otimes x$ is i times Kronecker product of x.

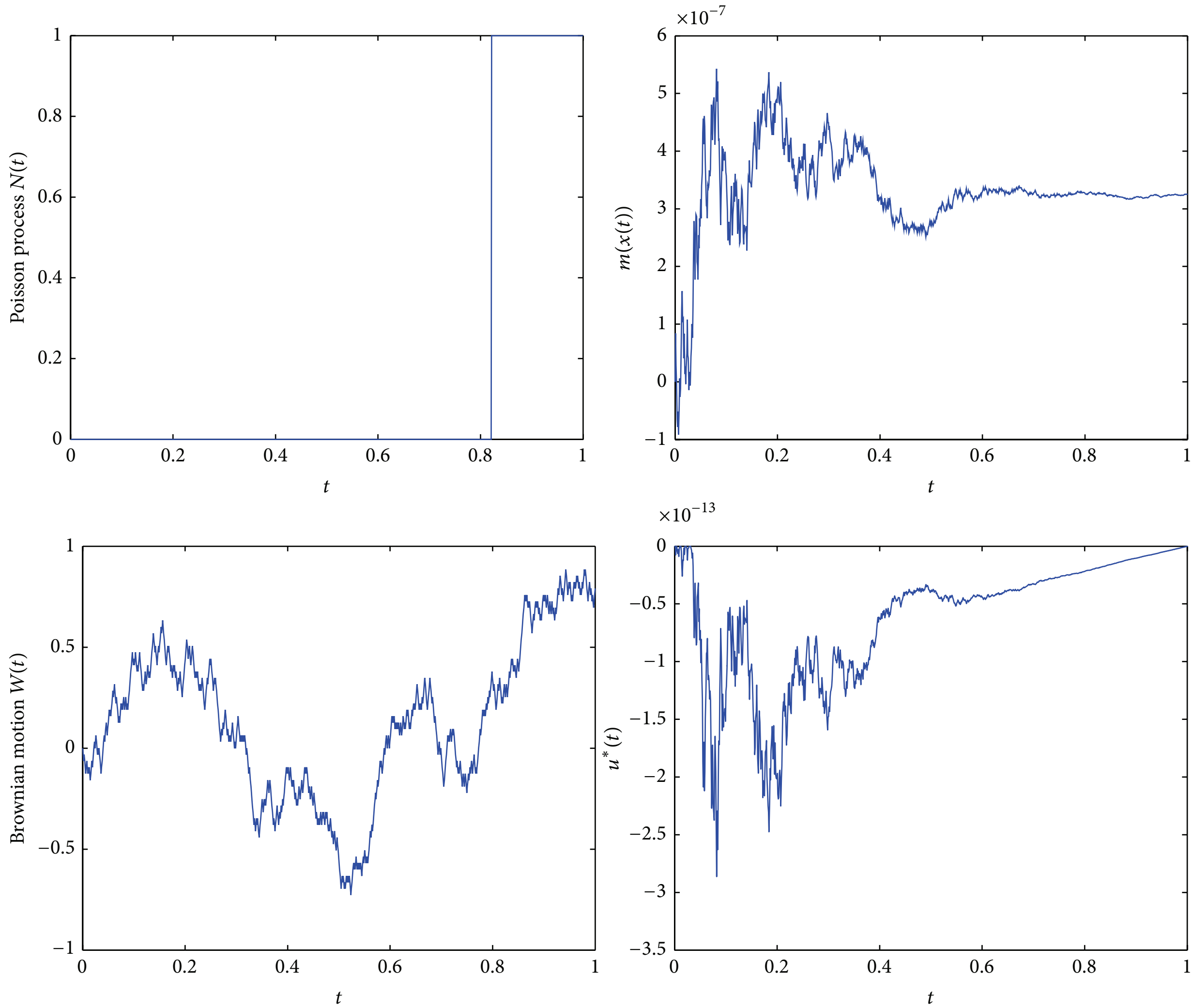


FIGURE 1: Tracking performance of Example 10.

Theorem 4. *For given $\gamma > 0$, suppose $P_i(t)$ $(i = 1, 2, \ldots)$ satisfy the following Riccati equations:*

$$\dot{P}_1 + A^{(1)}P_1 + P_1A^{(1)'} + \frac{1}{2}M'M + \frac{2}{\gamma^2}\left(Q^{(1)}P_1\right)\otimes\left(P_1Q^{(1)'}\right)$$
$$+ R^{(1)}\left(P_1\right) + \int_E \Big[\left(I_{n_x} + G(e)\right)'P_1\left(I_{n_x} + G(e)\right) - P_1$$
$$-G^{(1)}(e)P_1 - P_1G^{(1)'}(e)\Big]\lambda(de) = 0,$$

$$\dot{P}_i + A^{(i)}P_i + P_iA^{(i)'} + \frac{2}{\gamma^2}\sum_{r+j=i+1}\left(K^{(r)}P_r\right)\otimes\left(P_jK^{(j)'}\right)$$
$$+ R^{(i)}\left(P_i\right) - 2\sum_{r+j=i}\left(B^{(r)}P_r\right)\otimes\left(P_jB^{(j)'}\right)$$
$$+ \int_E \Big[\left(\otimes_i\left(I_{n_x} + G(e)\right)\right)'P_i\left(\otimes_i\left(I_{n_x} + G(e)\right)\right)$$
$$-P_i - G^{(i)}(e)P_i - P_iG^{(i)'}(e)\Big]\lambda(de) = 0,$$

$$P_i(T) = 0, \quad i = 1, 2, \ldots. \tag{17}$$

Then the H_∞ control u_T^* for system (6) *can be given by*

$$u_T^* = -2\sum_{i=1}^{\infty}\left\langle \otimes_i x, \left(B^{(i)}P_i\right)\otimes_i x\right\rangle, \tag{18}$$

where $B^{(i)} = \sum_{j=1}^{n_x} B_j' \otimes D^{(i,j)}$ and $D^{(i,j)}$ is given by following Lemma 6.

In order to prove Theorem 4, we need the following lemmas, and Lemmas 5–8 are given without proofs.

Lemma 5. *For any $x \in \mathcal{R}^{n_x}$, $y \in \mathcal{R}^{n_y}$, $u \in \mathcal{R}^{n_u}$, $v \in \mathcal{R}^{n_v}$, $P \in \mathcal{R}^{n_x \times n_y}$, and $Q \in \mathcal{R}^{n_u \times n_v}$ we have*

$$\langle x, Py\rangle\langle u, Qv\rangle = \left\langle x\otimes u, (P\otimes Q)(y\otimes v)\right\rangle. \tag{19}$$

Lemma 6. *For any matrix $P \in \mathcal{S}_{n_x^i}(\mathcal{R})$, $K \in \mathcal{R}^{n_k}$, and integer i, we have*

$$\frac{\partial \langle \otimes_i x, P \otimes_i x\rangle'}{\partial x} K = 2\left\langle \otimes_{i-1} x, \left(K^{(i)} P\right) \otimes_i x\right\rangle, \tag{20}$$

where $K^{(i)} = \sum_{j=1}^{n_x} k_j D^{(i,j)}$, $D^{(i,j)} = \sum_{l=1}^{i} D_l^{(i,j)}$, and $D_l^{(i,j)} = I_{n_x^{l-1}} \otimes e_j' \otimes I_{n_x^{i-l}}$.

Lemma 7. *Let $V_T(t,x) = \sum_{i=1}^{\infty} \langle \otimes_i x, P_i(t) \otimes_i x\rangle$ exist. We have*

$$\begin{aligned}&\frac{\partial V_T}{\partial x}' K K' \frac{\partial V_T}{\partial x}\\ &\quad = 4\sum_{m=1}^{\infty}\left\langle \otimes_m x, \sum_{i+j=m+1}\left(K^{(i)} P_i\right) \otimes \left(P_j K^{(j)'}\right) \otimes_m x\right\rangle.\end{aligned} \tag{21}$$

Lemma 8. *For any matrix $P \in \mathcal{S}_{n_x^i}(\mathcal{R})$, $x \in \mathcal{R}^{n_x}$, and integer i, we have*

$$\frac{\partial \left\langle \otimes_i x, P \otimes_i x\right\rangle'}{\partial x} A x = 2\left\langle \otimes_i x, \left(A^{(i)} P\right) \otimes_i x\right\rangle, \tag{22}$$

where $A^{(i)} = \sum_{j=1}^{n_x} A_j' \otimes D^{(i,j)}$, and A_j is the jth row vector of matrix A.

Lemma 9. *For any matrix $P \in \mathcal{S}_{n_x^i}(\mathcal{R})$ and integer i, we have*

$$x' C' \frac{\partial^2 \left\langle \otimes_i x, P \otimes_i x\right\rangle}{\partial x^2} C x = 2\left\langle \otimes_i x, R^{(i)}(P) \otimes_i x\right\rangle, \tag{23}$$

where $R^{(i)}(P) = C^{(i)} P C^{(i)'} + Q^{(i)} P$, $Q^{(i)} = \sum_{s=1}^{n_x}\sum_{t=1}^{n_x} C_s' \otimes C_t' \otimes (D^{(i-1,s)} D^{(i,t)})$, C_s is the sth row vector of matrix C, and $C^{(i)}$ is determined as $A^{(i)}$ in Lemma 8.

Proof. Let $K = Cx = (k_1, k_2, \ldots, k_{n_x})'$; then we have

$$\begin{aligned}
&x' C' \frac{\partial^2 \left\langle \otimes_i x, P \otimes_i x\right\rangle}{\partial x^2} C x\\
&\quad = \sum_{s=1}^{n_x}\sum_{t=1}^{n_x} \frac{\partial^2 \left\langle \otimes_i x, P \otimes_i x\right\rangle}{\partial x_s \partial x_t} k_s k_t\\
&\quad = 2\sum_{s=1}^{n_x}\sum_{t=1}^{n_x} k_s k_t \frac{\partial}{\partial x_s}\left\langle \otimes_{i-1} x, \sum_{l=1}^{i} D_l^{(i,t)} P\right\rangle\\
&\quad = 2\sum_{s=1}^{n_x}\sum_{t=1}^{n_x} k_s k_t \left[\sum_{m=1}^{i-1}\left\langle \otimes_{m-1} x \otimes e_s \otimes \otimes_{i-m-1} x,\right.\right.\\
&\qquad\qquad \left.\sum_{l=1}^{i} D_l^{(i,t)} P \otimes_i x\right\rangle\\
&\qquad \left.+ \sum_{m=1}^{i}\left\langle \otimes_{i-1} x, \sum_{l=1}^{i} D_l^{(i,t)} P \otimes_{m-1} x \otimes e_s \otimes \otimes_{i-m} x\right\rangle\right]\\
&\quad = 2\sum_{s=1}^{n_x}\sum_{t=1}^{n_x} k_s k_t \left[\left\langle \otimes_{i-2} x, \sum_{m=1}^{i-1} D_m^{i-1,s}\sum_{l=1}^{i} D_l^{(i,t)} P \otimes_i x\right\rangle\right.\\
&\qquad \left. + \left\langle \otimes_{i-1} x, \sum_{m=1}^{i}\sum_{l=1}^{i} D_l^{(i,t)} P D_m^{(i,t)'} \otimes_{i-1}\right\rangle\right]\\
&\quad = 2\sum_{s=1}^{n_x}\sum_{t=1}^{n_x}\left[\left\langle x, C_s'\right\rangle \left\langle x, C_t'\right\rangle\right.\\
&\qquad \times \left\langle \otimes_{i-2} x, \sum_{m=1}^{i-1} D_m^{i-1,s}\sum_{l=1}^{i} D_l^{(i,t)} P \otimes_i x\right\rangle\\
&\qquad + \left\langle x, C_s'\right\rangle \left\langle \otimes_{i-1} x, \sum_{m=1}^{i}\sum_{l=1}^{i} D_l^{(i,t)} P D_m^{(i,t)'} \otimes_{i-1}\right\rangle\\
&\qquad \left.\times \left\langle C_t', x\right\rangle\right].
\end{aligned} \tag{24}$$

By Lemma 5,

$$\begin{aligned}
&x' C' \frac{\partial^2 \left\langle \otimes_i x, P \otimes_i x\right\rangle}{\partial x^2} C x\\
&\quad = 2\sum_{s=1}^{n_x}\sum_{t=1}^{n_x}\left[\left\langle \otimes_i x, C_s' \otimes C_t' \otimes \sum_{m=1}^{i-1} D_m^{i-1,s}\right.\right.\\
&\qquad\qquad \left. \times \sum_{l=1}^{i} D_l^{(i,t)} P \otimes_i x\right\rangle\\
&\qquad + \left\langle \otimes_i x, \left(C_s' \otimes \sum_{m=1}^{i}\sum_{l=1}^{i} D_l^{(i,t)} P D_m^{(i,t)'}\right.\right.\\
&\qquad\qquad \left.\left.\left. \otimes C_t\right) \otimes_i x\right\rangle\right].
\end{aligned} \tag{25}$$

So we can obtain (23). □

Proof of Theorem 4. Applying Lemmas 6–9, we have

$$\begin{aligned}
V_T(t, x + G(e)x) = \sum_{i=1}^{\infty}&\left\langle \otimes_i x, \left(\otimes_i \left(I_{n_x} + G(e)\right)\right)' P_i(t)\right.\\
&\left.\times \left(\otimes_i \left(I_{n_x} + G(e)\right)\right) \otimes_i x\right\rangle,
\end{aligned} \tag{26}$$

$$\frac{\partial V_T}{\partial t} = \sum_{i=1}^{\infty}\left\langle \otimes_i x, \dot{P}_i \otimes_i x\right\rangle, \tag{27}$$

$$\frac{\partial V_T}{\partial x}' Ax = \sum_{i=1}^{\infty} \left\langle \otimes_i x, \left(A^{(i)} P_i\right) \otimes_i x \right\rangle + \sum_{i=1}^{\infty} \left\langle \otimes_i x, \left(P_i A^{(i)'}\right) \otimes_i x \right\rangle, \tag{28}$$

$$\frac{\partial V}{\partial x}' G(e)\, x = \sum_{i=1}^{\infty} \left\langle \otimes_i x, \left(D^{(i)}(e) P_i\right) \otimes_i x \right\rangle + \sum_{i=1}^{\infty} \left\langle \otimes_i x, \left(P_i D^{(i)'}\right) \otimes_i x \right\rangle, \tag{29}$$

$$\frac{\partial V_T}{\partial x}' Bxx' B' \frac{\partial V_T}{\partial x} = 4\sum_{m=2}^{\infty} \left\langle \otimes_m x, \sum_{i+j=m} \left(B^{(i)} P_i\right) \otimes \left(P_j B^{(j)'}\right) \otimes_m x \right\rangle, \tag{30}$$

$$x' C' \frac{\partial^2 V_T}{\partial x^2} Cx = 2\sum_{i=1}^{\infty} \left\langle \otimes_i x, R^{(i)}\left(P_i\right) \otimes_i x \right\rangle. \tag{31}$$

Substituting (26)–(31) and (21) into (8) with terminal conditions $P_i(T) = 0$ $(i = 1, 2, \ldots)$, we can prove that $V_T(t, x)$ satisfies HJE (8). By Theorem 2, the H_∞ control for system (6) can be given as

$$u_T^* = -x' B' \frac{\partial V_T}{\partial x}. \tag{32}$$

Similar to (29), we prove that the H_∞ control u_T^* can be represented with the form of (18). □

By Theorem 4, we can obtain the approximation of H_∞ control for system (6).

Now we apply the result of tensor power approach to an example.

Example 10. Consider the system (6) with coefficients

$$A = \begin{bmatrix} 0.04 & 0.02 \\ 0.02 & 0.04 \end{bmatrix}, \qquad B = \begin{bmatrix} 0.02 & 0.02 \\ 0.02 & 0.02 \end{bmatrix},$$
$$K = \begin{bmatrix} 0.02 \\ 0.02 \end{bmatrix}, \qquad C = \begin{bmatrix} -0.02 & 0.04 \\ 0.02 & 0.02 \end{bmatrix}, \tag{33}$$
$$G = \begin{bmatrix} 0.02 & -0.02 \\ 0.02 & 0.02 \end{bmatrix}, \qquad M = \begin{bmatrix} -0.04 & 0.02 \end{bmatrix}.$$

$N(t)$ is Poisson measure with parameter $\lambda = 2$; $W(t)$ is 1-dimensional Brownian motion and $\gamma = 1$. Here the approximation of u_T^* is given by

$$u^* = -2\sum_{i=1}^{6} \left\langle \otimes_i x, \left(B^{(i)} P_i\right) \otimes_i x \right\rangle, \tag{34}$$

and Figure 1 is the simulation of $m(x) = Mx$ and H_∞ control u^*, where u^* is the approximation of u_T^* of system (6). For the theories of simulation, we will discuss them in another paper. Here we only give the results of simulation.

5. Concluding Remarks

We have discussed the state feedback H_∞ control for a class of bilinear stochastic system where both Brownian motion and Poisson process are present. In order to solve the HJE given in the paper, we also discuss the method of tensor power series approach.

Conflict of Interests

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

Acknowledgment

This work was supported by the NSF of China, Grant nos. 61174078, 11271007, and 61170054.

References

[1] C. S. Wu and B. S. Chen, "Unified design for H_2, H_∞, and mixed control of spacecraft," *Journal of Guidance, Control, and Dynamics*, vol. 22, pp. 884–896, 1999.

[2] C. S. Wu and B. S. Chen, "Adaptive attitude control of spacecraft: Mixed H_2/ H_∞ approach," *Journal of Guidance, Control, and Dynamics*, vol. 24, pp. 755–766, 2001.

[3] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Transactions on Automatic Control*, vol. 34, no. 8, pp. 831–847, 1989.

[4] M. D. S. Aliyu and E. K. Boukas, "H_∞ control for Markovian jump nonlinear systems," in *Proceedings of the 37th IEEE Conference on Decision and Control*, pp. 766–771, Tampa, Fla, USA, 1998.

[5] E. K. Boukas and Z. K. Liu, "Robust H_∞ control of discrete-time Markovian jump linear systems with mode-dependent time-delays," *IEEE Transactions on Automatic Control*, vol. 46, no. 12, pp. 1918–1924, 2001.

[6] D. P. de Farias, J. C. Geromel, J. B. R. do Val, and O. L. V. Costa, "Output feedback control of Markov jump linear systems in continuous-time," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 944–949, 2000.

[7] X. Chen and K. Zhou, "Multiobjective H_2/H_∞ control design," *SIAM Journal on Control and Optimization*, vol. 40, no. 2, pp. 628–660, 2001.

[8] D. Hinrichsen and A. J. Pritchard, "Stochastic H_∞," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1504–1538, 1998.

[9] T. Damm, "State-feedback H_∞-type control of linear systems with time-varying parameter uncertainty," *Linear Algebra and Its Applications*, vol. 351-352, pp. 185–210, 2002.

[10] B.-S. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[11] V. A. Ugrinovskii, "Robust H_∞ control in the presence of stochastic uncertainty," *International Journal of Control*, vol. 71, no. 2, pp. 219–237, 1998.

[12] V. Dragan, A. Halanay, and A. Stoica, "The γ-attenuation problem for systems with state dependent noise," *Stochastic Analysis and Applications*, vol. 17, no. 3, pp. 395–404, 1999.

[13] V. Dragan and T. Morozan, "Global solutions to a game-theoretic Riccati equation of stochastic control," *Journal of Differential Equations*, vol. 138, no. 2, pp. 328–350, 1997.

[14] R. Boel and P. Varaiya, "Optimal control of jump processes," *SIAM Journal on Control and Optimization*, vol. 15, no. 1, pp. 92–119, 1977.

[15] R. Rishel, "A minimum principle for controlled jump processes," in *Control Theory, Numerical Methods and Computer Systems Modelling*, vol. 107 of *Lecture Notes in Economics and Mathematical Systems*, pp. 493–508, Springer, 1975.

[16] S. J. Tang and X. J. Li, "Necessary conditions for optimal control of stochastic systems with random jumps," *SIAM Journal on Control and Optimization*, vol. 32, no. 5, pp. 1447–1475, 1994.

[17] Z. Wu, "Fully coupled FBSDE with Brownian motion and Poisson process in stopping time duration," *Journal of the Australian Mathematical Society*, vol. 74, no. 2, pp. 249–266, 2003.

[18] I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*, Springer, 1972.

[19] C. Graham, "McKean-Vlasov Itô-Skorohod equations, and nonlinear diffusions with discrete jump sets," *Stochastic Processes and Their Applications*, vol. 40, no. 1, pp. 69–82, 1992.

[20] Z. Wu, "A general maximum principle for optimal control of forward-backward stochastic systems," *Automatica*, vol. 49, no. 5, pp. 1473–1480, 2013.

[21] W. Zhang and B.-S. Chen, "Stochastic affine quadratic regulator with applications to tracking control of quantum systems," *Automatica*, vol. 44, no. 11, pp. 2869–2875, 2008.

[22] W. Zhang and B.-S. Chen, "State feedback H_∞ control for a class of nonlinear stochastic systems," *SIAM Journal on Control and Optimization*, vol. 44, no. 6, pp. 1973–1991, 2006.

[23] I. I. Gihman and A. V. Skorohod, *The Theory of Stochastic Processes III*, Springer, 1979.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 249860, 11 pages
http://dx.doi.org/10.1155/2014/249860

Research Article

Dissipative Delay-Feedback Control for Nonlinear Stochastic Systems with Time-Varying Delay

Guici Chen,[1,2] Jianzhong Zhou,[1] and Yongchuan Zhang[1]

[1] *School of Hydropower & Information Engineering, Huazhong University of Science & Technology, Wuhan 430074, China*
[2] *Hubei Province Key Laboratory of Systems Science in Metallurgical Process, Wuhan University of Science & Technology, Wuhan 430081, China*

Correspondence should be addressed to Guici Chen; gcichen@163.com

Received 26 January 2014; Accepted 2 May 2014; Published 21 May 2014

Academic Editor: Weihai Zhang

The dissipative delay-feedback control problems for nonlinear stochastic delay systems (NSDSs) based on dissipativity analysis are studied in this paper. Based on the Lyapunov stability theory and stochastic analysis technique, both delay-independent and delay-dependent dissipativity criteria are established as linear matrix inequalities- (LMIs-) based feasibility tests. The obtained results in this paper for the nominal systems include the available results on H_∞ approach and passivity for stochastic delay systems as special cases. The delay-dependent feedback controller is designed by considering the relationship among the time-varying delay, its lower and upper bound, and its differential without ignoring any terms, which effectively reduces the conservative. A numerical example is given to illustrate the theoretical developments.

1. Introduction

The stochastic differential systems appear as a natural description of many observed phenomena of real world, which have been come to play an important role in many fields including population dynamics, macroeconomics, chemical reactor control, communication network, image processes, and mobile robot localization. Therefore, the stability and stabilizability of nonlinear stochastic differential systems affine in the control have been studied in the past years by means of the stochastic Lyapunov theory [1–4]. As we all know, time delay in a control loop is one of the main sources of instability, oscillation, and poor performance and naturally encountered in a number of engineering control problems and physical systems. Therefore, time-delay systems [5, 6] and stochastic time-delay systems [7–12] have attracted many researchers' attention and have been extensively studied. The problems include stability analysis [5, 6], stabilization problems [7, 8], and robust controller design [9–12].

On the other hand, studying the dissipativity analysis and synthesis problems has a strong motivation due to their simplicity and effectiveness in dealing with robust and nonlinear systems. Since the notation of dissipative dynamical system was introduced by Willems [13], dissipative systems have been of particular interest to researchers in areas of systems, circuits, networks and control, and so forth. Moreover, passivity of a certain system in a feedback interconnected system will ensure the overall stability of that feedback system if uncertainties or nonlinearities can be characterized by a strict passive system. Hence, dissipative theory has wide ranging implications and applications in control theory. For instance, dissipativity was crucially used in the stability analysis of nonlinear system [14]; the theory of dissipative systems generalizes basic tools including the passivity theory, bounded real lemma, Kalman Yakubovich lemma, and the circle criterion [15]. Among the relevant topics are the dissipativity analysis and synthesis for time-delay systems [16, 17]. These results show that the dissipativity-based methods are highly effective in design the robust controller.

Due to what is above mentioned, we believe that time delay is often harmful factor of systems. However, time delay is also surprising since plenty of studies have shown that time delay can also benefit the control, such as time-delay control. As mentioned previously, many results have been published about the control of systems with state delays but without input delays, which is called memoryless controllers, or to more general, memoryless controllers only includes an instantaneous feedback term. The time-delay control is an approach which gives a small delay h in the controller design, so as to reduce the effect of instability factor and exogenous disturbance. See, for example [18–20], and the references therein. Rather than adjusting control gains or identifying model parameters, its essential idea is to use past observations regarding both the control input and system response, which is an open problem now. In this paper, dissipative delay-feedback control problems for nonlinear stochastic systems with time-varying delay are studied based on dissipativity. The delay-dependent feedback controller is designed by considering the relationship among the time-varying delay, its lower and upper bound, and its differential without ignoring any terms, which effectively reduces the conservative.

2. Problems Statement and Preliminaries

In this paper, we consider the following nonlinear stochastic delayed systems (NSDSs) defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$:

$$\begin{aligned} \mathrm{d}x(t) = \{&F(t, x(t), x(t-\tau(t))) + Bu(t, t-\tau(t)) \\ &+Dv(t)\}\,\mathrm{d}t \\ &+ G(t, x(t), x(t-\tau(t)))\,\mathrm{d}\omega(t), \\ z(t) &= Cx(t) + C_d x(t-\tau(t)), \\ x(t) &= \varphi(t), \quad t \in [-\tau_2, 0), \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector; $u(t, t-\tau(t)) \in \mathbb{R}^m$ is the control input, which depends on not only the real time but also the delay; we call the controller memory controller; $z(t) \in \mathbb{R}^q$ is the control output; $v(t) \in \mathbb{R}^p$ is the exogenous disturbance input, which satisfies $v(t) \in L_2([0,\infty), \mathbb{R}^p)$, where $L_2([0,\infty), \mathbb{R}^p)$ is the space of nonanticipatory square-summable stochastic process with respect to $(\mathcal{F}_t)_{t>0}$ with the following norm: $\|v(t)\|_2^2 = \mathbb{E}\int_0^\infty \|v(t)\|^2 \mathrm{d}t$. $\omega(t) \in \mathbb{R}^l$ is a scalar Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}[\mathrm{d}\omega(t)] = 0$, $\mathbb{E}[\mathrm{d}^2\omega(t)] = \mathrm{d}t$.

In the sequel, we seek to study the problems of dissipative analysis and delay-feedback control for the two cases of time delay.

Case 1. Time delay is a constant τ.

Case 2. $\tau(t)$ is the time-varying delay, which is a differential function satisfying

$$0 \le \tau_1 \le \tau(t) \le \tau_2, \qquad \dot{\tau}(t) \le \tau_d \le 1, \tag{2}$$

where τ_1, τ_2, and τ_d are nonnegative constants.

Remark 1. Obviously, when $\tau_d = 0$, $\tau(t) = \tau_1 = \tau_2$, that means the time delay is a constant; this case has been extensively studied. On the other hand, the time-varying delay $\tau(t) \ge \tau_1$; here τ_1 is equal to or greater than 0, which has less conservativeness than $\tau(t) > 0$.

Assumption 2. $F(\cdot,\cdot,\cdot)$ is a nonlinear vector function which can be decomposed as follows:

$$\begin{aligned} F&(t, x(t), x(t-\tau(t))) \\ &= A_0 x(t) + f(t, x(t)) \\ &\quad + A_d x(t-\tau(t)) + f_d(t, x(t-\tau(t))), \end{aligned} \tag{3}$$

where f, f_d are vector-valued functions; we assume

$$\|f(t, x(t))\| \le \beta \|x(t)\|, \tag{4}$$

$$\|f_d(t, x(t-\tau(t)))\| \le \beta_d \|x(t-\tau(t))\|, \tag{5}$$

where β, β_d are known real positive constants.

Obviously, we know that

$$f(0,0) = 0, \qquad f_d(0,0) = 0. \tag{6}$$

Equivalently stated, condition (4) implies that there exists a scalar $\kappa > 0$ such that

$$\kappa\left(\beta^2 x^T(t)x(t) - f^T(t, x(t)) f(t, x(t))\right) \ge 0. \tag{7}$$

Similarly, condition (5) implies that there exists a scalar $\kappa_d > 0$ such that

$$\begin{aligned} \kappa_d\Big(&\beta_d^2 x^T(t-\tau(t)) x(t-\tau(t)) \\ &-f_d^T(t, x(t-\tau(t))) f_d(t, x(t-\tau(t)))\Big) \ge 0. \end{aligned} \tag{8}$$

Assumption 3. $G(\cdot,\cdot,\cdot)$ is a nonlinear vector function which satisfies

$$\begin{aligned} \mathrm{Trace}&\left(G^T(t, x(t), x(t-\tau(t)))\, G(t, x(t), x(t-\tau(t)))\right) \\ &\le x^T(t)\Theta_1^T\Theta_1 x(t) + x^T(t-\tau(t))\Theta_2^T\Theta_2 x(t-\tau(t)), \end{aligned} \tag{9}$$

where Θ_1, Θ_2 are known real matrices.

Hence, nonlinear stochastic delay systems (NSDSs) (1) can be rewritten as

$$\begin{aligned} \mathrm{d}x(t) = \{&Ax(t) + f(t, x(t)) + A_d x(t-\tau(t)) \\ &+ f_d(t, x(t-\tau(t))) \\ &+ Bu(t, t-\tau(t)) + Dv(t)\}\,\mathrm{d}t \\ &+ G(t, x(t), x(t-\tau(t)))\,\mathrm{d}\omega(t), \\ z(t) &= Cx(t) + C_d x(t-\tau(t)), \\ x(t) &= \varphi(t), \quad t \in [-\tau_2, 0). \end{aligned} \tag{10}$$

Definition 4 (see [21]). Given matrices $Q^T = Q \le 0$, $R^T = R \ge 0$, and S, nonlinear stochastic delay systems (NSDSs) (10) are called (Q, S, R)-dissipative if, for some real function $\eta(\cdot)$, $\eta(0) = 0$,

$$\begin{aligned} \mathbb{E}\int_0^T &\left[z^T(s)Qz(s) + 2v^T(s)Sz(s) + v^T(s)Rv(s)\right]\mathrm{d}s \\ &+ \eta(x_0) \ge 0, \quad \forall T \ge 0. \end{aligned} \tag{11}$$

Furthermore, if, for a scalar $\alpha > 0$,

$$\mathbb{E}\int_0^T \left[z^T(s)Qz(s) + 2v^T(s)Sz(s) + v^T(s)Rv(s)\right]\mathrm{d}s + \eta(x_0) \geq \mathbb{E}\int_0^T \alpha v^T(s)v(s)\,\mathrm{d}s, \quad \forall T \geq 0, \tag{12}$$

NSDSs (10) are called strictly (Q, S, R)-dissipative.

Lemma 5 (see [22]). *Given three constant matrices S_1, S_2, and S_3, where $S_3 = S_3^T < 0$ and $S_1 = S_1^T < 0$, then $S_1 - S_2^T S_3^{-1} S_2 < 0$ holds if and only if $\begin{pmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{pmatrix} < 0$ or $\begin{pmatrix} S_3 & S_2 \\ S_2^T & S_1 \end{pmatrix} < 0$.*

Lemma 6 (see [23]). *For given positive symmetric matrix $M = M^T > 0$, two scalars a and b satisfying $a < b$, and vector function $x(t) : [a, b] \to \mathbb{R}^n$, then*

$$\left[\int_a^b x(s)\,ds\right]^T M \left[\int_a^b x(s)\,ds\right] \leq (b-a)\int_a^b x^T(s)Mx(s)\,ds. \tag{13}$$

3. Dissipativity Analysis for NSDSs

In this section, our primary purpose is to develope delay-independent and delay-dependent stochastically stability and dissipativity criteria for NSDSs (10) based on Definition 4.

3.1. Delay-Independent Dissipativity. In this sequel, we consider the time delay as unknown constant pertaining to Case 1 and hence the results developed hereinafter will be independent of the size of delay. Without regard to the control input, setting $u(t, t-\tau(t)) = 0$, then (10) can be rewritten as

$$\begin{aligned} \mathrm{d}x(t) &= \{Ax(t) + f(t, x(t)) + A_d x(t-\tau) \\ &\quad + f_d(t, x(t-\tau)) + Dv(t)\}\,\mathrm{d}t \\ &\quad + G(t, x(t), x(t-\tau))\,\mathrm{d}\omega(t), \\ z(t) &= Cx(t) + C_d x(t-\tau), \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0). \end{aligned} \tag{14}$$

Theorem 7. *Consider the NSDSs* (14). *Given some scalars $\alpha > 0$, $\beta > 0$, and $\beta_d > 0$ and matrices $Q = Q^T \leq 0$, $R = R^T > 0$, and S, suppose there exist matrices $P = P^T > 0$, $W = W^T > 0$ and positive scalars $\kappa > 0$, $\kappa_d > 0$ such that the following LMI holds:*

$$\begin{pmatrix} \Sigma_1 & PA_d & \Sigma_2 & P & P & C^TQ \\ * & \Sigma_3 & -C_d^TS & 0 & 0 & C_d^TQ \\ * & * & -R_\alpha & 0 & 0 & 0 \\ * & * & * & -\kappa I & 0 & 0 \\ * & * & * & * & -\kappa_d I & 0 \\ * & * & * & * & * & Q \end{pmatrix} < 0; \tag{15}$$

then the NSDSs (14) *are strictly (Q, S, R)-dissipative independent of delay, where $\Sigma_1 = A^TP + PA + W + \Theta_1^T P\Theta_1 + \kappa\beta^2 I$, $\Sigma_2 = PD - C^TS$, $\Sigma_3 = \Theta_2^T P\Theta_2 - W + \kappa_d\beta_d^2 I$, and $R_\alpha = (R - \alpha I)$.*

Proof. At first we introduce the following Lyapunov-Krasovskii functional (LKF):

$$V(x_t) = x^T(t)Px(t) + \int_{t-\tau}^t x^T(s)Wx(s)\,\mathrm{d}s. \tag{16}$$

Evaluating the Itô derivative of $V(x_t)$ along the solution of NSDSs (14), we have

$$\begin{aligned} \mathscr{L}V(x_t) &= 2x^TP[Ax(t) + f(t, x(t)) + A_d x(t-\tau) \\ &\quad + f_d(t, x(t-\tau)) + Dv(t)] \\ &\quad + x^TWx - x^T(t-\tau)Wx(t-\tau) \\ &\quad + G^T(t, x(t), x(t-\tau))PG(t, x(t), x(t-\tau)). \end{aligned} \tag{17}$$

Noting (7)–(9), we obtain

$$\begin{aligned} &\mathscr{L}V(x_t) \\ &\leq \mathscr{L}V(x_t) \\ &\quad + \kappa\left(\beta^2 x^T(t)x(t) - f^T(t, x(t))f(t, x(t))\right) \\ &\quad + \kappa_d\big(\beta_d^2 x^T(t-\tau(t))x(t-\tau(t)) \\ &\quad - f_d^T(t, x(t-\tau(t))) \\ &\quad \times f_d(t, x(t-\tau(t)))\big) \\ &\leq \xi^T\Sigma\xi, \end{aligned} \tag{18}$$

where

$$\begin{aligned} \xi &= \big(x^T(t), x^T(t-\tau), v^T(t), f^T(t, x(t)), \\ &\quad f_d^T(t, x(t-\tau))\big)^T, \\ \Sigma &= \begin{pmatrix} \Sigma_1 & PA_d & PD & P & P \\ * & \Sigma_3 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & -\kappa I & 0 \\ * & * & * & * & -k_d I \end{pmatrix}. \end{aligned} \tag{19}$$

Hence, we have

$$\begin{aligned} &\mathscr{L}V(x_t) - z^T(s)Qz(s) - 2v^T(s)Sz(s) \\ &\quad - v^T(s)(R - \alpha I)v(s) \leq \xi^T\widetilde{\Sigma}\xi, \end{aligned} \tag{20}$$

where

$$\widetilde{\Sigma} = \Sigma - \begin{pmatrix} C^TQC & C^TQC_d^T & C^TS & 0 & 0 \\ * & C_d^TQC_d & C_d^TS & 0 & 0 \\ * & * & (R-\alpha I) & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & 0 \end{pmatrix}. \tag{21}$$

According to Lemma 5 and applying the congruent transformation, we know that

$$\mathscr{L}V(x_t) - z^T(s) Qz(s) - 2v^T(s) Sz(s) - v^T(s)(R-\alpha I) v(s) \le 0. \tag{22}$$

Then, integrating (22) from 0 to T and taking mathematical expectation, we obtain that (12) holds which completes the proof. □

Remark 8. When $Q = -I$, $S = 0$, and $R_\alpha = \gamma^2 I$, strictly (Q, S, R)-dissipativity reduces to the H_∞ performance level. When $Q = 0$, $S = I$, and $R_\alpha = \gamma I$, strictly (Q, S, R)-dissipativity reduces to the strictly passivity.

So the following corollaries stand out as special cases.

Corollary 9. *Consider the NSDSs* (14). *Given some scalars $\gamma > 0$, $\beta > 0$, and $\beta_d > 0$, suppose there exist matrices $P = P^T > 0$, $W = W^T > 0$ and positive scalars $\kappa > 0$, $\kappa_d > 0$ such that the following LMI holds:*

$$\begin{pmatrix} \Sigma_1 & PA & PD & P & P & -C^T \\ * & \Sigma_3 & 0 & 0 & 0 & -C_d^T \\ * & * & -\gamma^2 I & 0 & 0 & 0 \\ * & * & * & -\kappa I & 0 & 0 \\ * & * & * & * & -\kappa_d I & 0 \\ * & * & * & * & * & -I \end{pmatrix} < 0; \tag{23}$$

then the NSDSs (14) *are stochastically asymptotically stable and independent of delay with disturbance level γ.*

Corollary 10. *Consider the NSDSs* (14). *Given some scalars $\gamma > 0$, $\beta > 0$, and $\beta_d > 0$, suppose there exist matrices $P = P^T > 0$, $W = W^T > 0$ and positive scalars $\kappa > 0$, $\kappa_d > 0$ such that the following LMI holds:*

$$\begin{pmatrix} \Sigma_1 & PA & PD - C^T & P & P \\ * & \Sigma_3 & -C_d^T & 0 & 0 \\ * & * & -\gamma I & 0 & 0 \\ * & * & * & -\kappa I & 0 \\ * & * & * & * & -\kappa_d I \end{pmatrix} < 0; \tag{24}$$

then the NSDSs (14) *are strictly passive independent of delay.*

3.2. Delay-Dependent Dissipativity. We now direct attention to the type of dissipativity which depends on the time-varying delay, which pertains to Case 2. Without consideration of the control input, and defining a new state variable

$$y(t) = Ax(t) + f(t, x(t)) + A_d x(t-\tau(t)) + f_d(t, x(t-\tau(t))) + Dv(t), \tag{25}$$

then, the NSDSs (10) can be rewritten as

$$\begin{aligned} \mathrm{d}x(t) &= y(t)\,\mathrm{d}t + G(t, x(t), x(t-\tau(t)))\,\mathrm{d}\omega(t), \\ z(t) &= Cx(t) + C_d x(t-\tau(t)), \\ x(t) &= \varphi(t), \quad t \in [-\tau_2, 0). \end{aligned} \tag{26}$$

Theorem 11. *Consider the NSDSs* (26). *For the given scalars $\alpha > 0$, $\tau_2 > \tau_1 \ge 0$, $\tau_d > 0$, $\beta > 0$, and $\beta_d > 0$, the NSDSs* (26) *are strictly (Q, S, R)-dissipative for all time-varying delays if there exist symmetric positive-definite matrices P, V_1, V_2, V_3, W_1, W_2, Z_1, and Z_2; any appropriately dimensioned matrices N, M, F, and H; and two positive scalars $\kappa > 0$, $\kappa_d > 0$ such that the following LMIs hold:*

$$\begin{gathered} \begin{pmatrix} \Xi & -\tau_2 \widehat{M} \\ * & -\tau_2 W_1 \end{pmatrix} < 0, \\ \begin{pmatrix} \Xi & -(\tau_2 - \tau_1)\widehat{N} \\ * & -(\tau_2-\tau_1)(W_1 + W_2) \end{pmatrix} < 0, \\ \begin{pmatrix} \Xi & -(\tau_2 - \tau_1)\widehat{F} \\ * & -(\tau_2-\tau_1) W_2 \end{pmatrix} < 0, \end{gathered} \tag{27}$$

where

$$\begin{gathered} \Xi = \begin{pmatrix} \Xi_{11} & \Xi_{12} & \Xi_{13} & \Xi_{14} & A^T H & \Xi_{16} \\ * & \Xi_{22} & -C_d^T S & 0 & A_d^T H & \Xi_{26} \\ * & * & -R_\alpha & 0 & D^T H & 0 \\ * & * & * & \Xi_{44} & \Xi_{45} & 0 \\ * & * & * & * & \Xi_{55} & 0 \\ * & * & * & * & * & \Xi_{66} \end{pmatrix}, \\ \Xi_{11} = PA + A^T P + V_1 + V_2 + V_3 + \tau_2 Z_1 + (\tau_2 - \tau_1) Z_2 + M + M^T + \kappa\beta^2 I, \\ \Xi_{12} = N - F - M + PA_d, \qquad \Xi_{13} = PD - C^T S, \\ \Xi_{14} = \begin{pmatrix} F & -N & P & P \end{pmatrix}, \\ \Xi_{16} = \begin{pmatrix} C^T Q & \Theta_1 P & 0 \end{pmatrix}, \qquad \Xi_{22} = -(1-\tau_d) V_2 + \kappa_d \beta_d^2 I, \\ \Xi_{26} = \begin{pmatrix} C_d^T Q & 0 & \Theta_2 P \end{pmatrix}, \\ \Xi_{44} = \operatorname{diag}\{-V_1, -V_3, -\kappa I, -\kappa_d I\}, \\ \Xi_{55} = \tau_2 W_1 + (\tau_2 - \tau_1) W_2 - H - H^T, \\ \Xi_{66} = \operatorname{diag}\{Q, -P, -P\}, \qquad \Xi_{45} = \begin{pmatrix} 0 & 0 & H^T & H^T \end{pmatrix}^T, \\ \widehat{N} = \left(N^T, 0_{1\times 10}\right)^T, \qquad \widehat{M} = \left(M^T, 0_{1\times 10}\right)^T, \\ \widehat{F} = \left(F^T, 0_{1\times 10}\right)^T. \end{gathered} \tag{28}$$

Proof. We construct a LKF as follows:

$$V(x_t) = V_1(x_t) + V_2(x_t) + V_3(x_t) + V_4(x_t), \tag{29}$$

where

$$V_1(x_t) = x^T(t) P x(t),$$
$$V_2(x_t) = \int_{t-\tau_1}^{t} x^T(s) V_1 x(s)\,\mathrm{d}s + \int_{t-\tau(t)}^{t} x^T(s) V_2 x(s)\,\mathrm{d}s + \int_{t-\tau_2}^{t} x^T(s) V_3 x(s)\,\mathrm{d}s,$$
$$V_3(x_t) = \int_{-\tau_2}^{0}\int_{t+\theta}^{t} y^T(s) W_1 y(s)\,\mathrm{d}s + \int_{-\tau_2}^{-\tau_1}\int_{t+\theta}^{t} y^T(s) W_2 y(s)\,\mathrm{d}s,$$
$$V_4(x_t) = \int_{-\tau_2}^{0}\int_{t+\theta}^{t} x^T(s) Z_1 x(s)\,\mathrm{d}s + \int_{-\tau_2}^{-\tau_1}\int_{t+\theta}^{t} x^T(s) Z_2 x(s)\,\mathrm{d}s. \tag{30}$$

Then, the weak infinitesimal operator $\mathscr{L}$ of the stochastic process x_t along the evolution of $V(x_t)$ is given by

$$\mathscr{L}V(x_t) = \mathscr{L}V_1(x_t) + \mathscr{L}V_2(x_t) + \mathscr{L}V_3(x_t) + \mathscr{L}V_4(x_t), \tag{31}$$

where

$$\begin{aligned}\mathscr{L}V_1(x_t) \le{}& 2x^T(t) P [Ax(t) + f(t, x(t)) + A_d x(t-\tau(t)) + f_d(t, x(t-\tau(t))) + Dv(t)] \\ &+ x^T(t)\Theta_1^T P \Theta_1 x(t) + x(t-\tau(t))\Theta_2^T P\Theta_2 x(t-\tau(t)),\end{aligned} \tag{32}$$

$$\begin{aligned}\mathscr{L}V_2(x_t) \le{}& x^T(t)(V_1+V_2+V_3)x(t) - x^T(t-\tau_1)V_1 x(t-\tau_1) \\ &- x^T(t-\tau_2)V_3 x(t-\tau_2) - x^T(t-\tau(t))(1-\tau_d)V_2 x(t-\tau(t)),\end{aligned} \tag{33}$$

$$\begin{aligned}\mathscr{L}V_3(x_t) ={}& \tau_2 y^T(t) W_1 y(t) - \int_{t-\tau_2}^{t} y^T(s) W_1 y(s)\,\mathrm{d}s \\ &+ (\tau_2-\tau_1) y^T(t) W_2 y(t) - \int_{t-\tau_2}^{t-\tau_1} y^T(s) W_2 y(s)\,\mathrm{d}s.\end{aligned} \tag{34}$$

By using Lemma 6, we get the following inequality:

$$\begin{aligned}\mathscr{L}V_4(x_t) \le{}& \tau_2 x^T(t) Z_1 x(t) + (\tau_2-\tau_1) x^T(t) Z_2 x(t) \\ &- \frac{1}{\tau_2}\left(\int_{t-\tau(t)}^{t} x(s)\mathrm{d}s\right)^T Z_1 \left(\int_{t-\tau(t)}^{t} x(s)\,\mathrm{d}s\right) \\ &- \frac{1}{\tau_2-\tau_1}\left(\int_{t-\tau_2}^{t-\tau(t)} x(s)\mathrm{d}s\right)^T Z_2 \left(\int_{t-\tau_2}^{t-\tau(t)} x(s)\,\mathrm{d}s\right) \\ &- \frac{1}{\tau_2-\tau_1}\left(\int_{t-\tau(t)}^{t-\tau_1} x(s)\mathrm{d}s\right)^T Z_2 \left(\int_{t-\tau(t)}^{t-\tau_1} x(s)\,\mathrm{d}s\right),\end{aligned} \tag{35}$$

setting $\xi(t) = (x^T(t), x^T(t-\tau(t)), v^T(t), x^T(t-\tau_1), x^T(t-\tau_2), f^T(x(t)), f_d^T(x(t-\tau(t))), y^T(t), (\int_{t-\tau(t)}^{t} x(s)\mathrm{d}s)^T, (\int_{t-\tau_2}^{t-\tau(t)} x(s)\mathrm{d}s)^T, (\int_{t-\tau(t)}^{t-\tau_1} x(s)\mathrm{d}s)^T)^T$. And we introduce the following four zero equations:

$$2x^T(t) M\left[x(t) - x(t-\tau(t)) - \int_{t-\tau(t)}^{t} \mathrm{d}x(s)\right] = 0, \tag{36}$$

$$2x^T(t) N\left[x(t-\tau(t)) - x(t-\tau_2) - \int_{t-\tau_2}^{t-\tau(t)} \mathrm{d}x(s)\right] = 0, \tag{37}$$

$$2x^T(t) F\left[x(t-\tau_1) - x(t-\tau(t)) - \int_{t-\tau(t)}^{t-\tau_1} \mathrm{d}x(s)\right] = 0, \tag{38}$$

$$\begin{aligned}&2y^T(t) H [Ax(t) + f(t, x(t)) + A_d x(t-\tau(t)) \\ &\quad + f_d(t, x(t-\tau(t))) + Dv(t) - y(t)] = 0.\end{aligned} \tag{39}$$

Summing up (31)–(39), we obtain

$$\begin{aligned}\mathscr{L}V(x_t) \le{}& \mathscr{L}V(x_t) \\ &+ \kappa\left(\beta^2 x^T(t) x(t) - f^T(t, x(t)) f(t, x(t))\right) \\ &+ \kappa_d\left(\beta_d^2 x^T(t-\tau(t)) x(t-\tau(t)) - f_d^T(t, x(t-\tau(t))) f_d(t, x(t-\tau(t)))\right) \\ &+ 2x^T(t) M\left[x(t) - x(t-\tau(t)) - \int_{t-\tau(t)}^{t} y(s)\,\mathrm{d}s\right] \\ &- 2x^T(t) M \int_{t-\tau(t)}^{t} G(x_s)\,\mathrm{d}\omega(s) \\ &+ 2x^T(t) N\left[x(t-\tau(t)) - x(t-\tau_2) - \int_{t-\tau_2}^{t-\tau(t)} y(s)\,\mathrm{d}s\right]\end{aligned}$$

$$\begin{aligned}
&- 2x^T(t) N \int_{t-\tau_2}^{t-\tau(t)} G(x_s)\,\mathrm{d}\omega(s) \\
&+ 2x^T(t) F \Big[x(t-\tau_1) - x(t-\tau(t)) \\
&\qquad - \int_{t-\tau(t)}^{t-\tau_1} y(s)\,\mathrm{d}s \Big] \\
&- 2x^T(t) F \int_{t-\tau(t)}^{t-\tau_1} G(x_s)\,\mathrm{d}\omega(s) \\
&+ 2y^T(t) H \big[Ax(t) + f(t,x(t)) \\
&\qquad + A_d x(t-\tau(t)) \\
&\qquad + f_d(t,x(t-\tau(t))) \\
&\qquad + Dv(t) - y(t) \big],
\end{aligned} \tag{40}$$

where M, N, F, and H are matrices with appropriate dimensions. Hence,

$$\begin{aligned}
&\mathscr{L}V(x_t) - z^T(s) Qz(s) - 2v^T(s) Sz(s) \\
&\quad - v^T(s)(R-\alpha I) v(s) \\
&\le \xi^T(t) \Phi \xi(t) - 2x^T(t) M \\
&\quad \times \Big[\int_{t-\tau(t)}^{t} y(s)\,\mathrm{d}s + \int_{t-\tau(t)}^{t} G(x_s)\,\mathrm{d}\omega(s) \Big] \\
&\quad - 2x^T(t) N \Big[\int_{t-\tau_2}^{t-\tau(t)} y(s)\,\mathrm{d}s + \int_{t-\tau_2}^{t-\tau(t)} G(x_s)\,\mathrm{d}\omega(s) \Big] \\
&\quad - 2x^T(t) F \Big[\int_{t-\tau(t)}^{t-\tau_1} y(s)\,\mathrm{d}s + \int_{t-\tau(t)}^{t-\tau_1} G(x_s)\,\mathrm{d}\omega(s) \Big] \\
&\quad - \int_{t-\tau_2}^{t} y^T(s) W_1 y(s)\,\mathrm{d}s - \int_{t-\tau_2}^{t-\tau_1} y^T(s) W_2 y(s)\,\mathrm{d}s,
\end{aligned} \tag{41}$$

where

$$\Phi = \begin{pmatrix}
\Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & A^T H & 0 \\
* & \Phi_{22} & -C_d^T S & 0 & A_d^T H & 0 \\
* & * & -R_\alpha & 0 & D^T H & 0 \\
* & * & * & \Phi_{44} & \Phi_{45} & 0 \\
* & * & * & * & \Phi_{55} & 0 \\
* & * & * & * & * & \Phi_{66}
\end{pmatrix},$$

$$\begin{aligned}
\Phi_{11} &= PA + A^T P + \Theta_1^T P \Theta_1 + V_1 + V_2 + V_3 + \tau_2 Z_1 \\
&\quad + (\tau_2 - \tau_1) Z_2 + M + M^T - C^T QC + \kappa\beta^2 I, \\
\Phi_{12} &= N - F - M + PA_d - C^T QC_d, \qquad \Phi_{13} = PD - C^T S, \\
\Phi_{14} &= \begin{pmatrix} F & -N & P & P \end{pmatrix}, \\
\Phi_{22} &= \Theta_2^T P \Theta_2 - C_d^T QC_d - (1-\tau_d) V_2 + \kappa_d \beta_d^2 I, \\
\Phi_{44} &= \operatorname{diag}\{-V_1, -V_3, -\kappa I, -\kappa_d I\}, \\
\Phi_{45} &= \begin{pmatrix} 0 & 0 & H^T & H^T \end{pmatrix}^T, \\
\Phi_{55} &= \tau_2 W_1 + (\tau_2 - \tau_1) W_2 - H - H^T, \\
\Phi_{66} &= \operatorname{diag}\Big\{ -\frac{1}{\tau_2} Z_1, -\frac{1}{\tau_2 - \tau_1} Z_2, -\frac{1}{\tau_2 - \tau_1} Z_2 \Big\}.
\end{aligned} \tag{42}$$

So it easy to obtain that

$$\begin{aligned}
&\mathscr{L}V(x_t) - z^T(s) Qz(s) - 2v^T(s) Sz(s) - v^T(s)(R-\alpha I) v(s) \\
&\le \frac{1}{\tau_2} \int_{t-\tau(t)}^{t} \eta^T(t,s) \begin{pmatrix} \Phi & -\tau_2 \widehat{M} \\ * & -\tau_2 W_1 \end{pmatrix} \eta(t,s)\,\mathrm{d}s \\
&\quad + \frac{1}{\tau_2 - \tau_1} \int_{t-\tau(t)}^{t} \eta^T(t,s) \\
&\qquad \times \begin{pmatrix} \Phi & -(\tau_2 - \tau_1) \widehat{N} \\ * & -(\tau_2 - \tau_1)(W_1 + W_2) \end{pmatrix} \eta(t,s)\,\mathrm{d}s \\
&\quad + \frac{1}{\tau_2 - \tau_1} \int_{t-\tau(t)}^{t} \eta^T(t,s) \begin{pmatrix} \Phi & -(\tau_2 - \tau_1) \widehat{F} \\ * & -(\tau_2 - \tau_1) W_2 \end{pmatrix} \eta(t,s)\,\mathrm{d}s \\
&\quad - 2x^T(t) M \int_{t-\tau(t)}^{t} G(x_s)\,\mathrm{d}\omega(s) \\
&\quad - 2x^T(t) N \int_{t-\tau_2}^{t-\tau(t)} G(x_s)\,\mathrm{d}\omega(s) \\
&\quad - 2x^T(t) F \int_{t-\tau(t)}^{t-\tau_1} G(x_s)\,\mathrm{d}\omega(s),
\end{aligned} \tag{43}$$

where $\eta^T(t,s) = [\xi^T(t), y^T(s)]$.

By Lemma 5 and applying the congruent transformation to (27), it follows that

$$\begin{aligned}
&\mathscr{L}V(x_t) - z^T(s) Qz(s) - 2v^T(s) Sz(s) \\
&\quad - v^T(s)(R-\alpha I) v(s) \\
&\quad \le -2x^T(t) M \int_{t-\tau(t)}^{t} G(x_s)\,\mathrm{d}\omega(s) \\
&\quad - 2x^T(t) N \int_{t-\tau_2}^{t-\tau(t)} G(x_s)\,\mathrm{d}\omega(s) \\
&\quad - 2x^T(t) F \int_{t-\tau(t)}^{t-\tau_1} G(x_s)\,\mathrm{d}\omega(s).
\end{aligned} \tag{44}$$

Then, integrating both sides of (44) from 0 to T and taking mathematical expectation, we obtain that (12) holds, which completes the proof. □

Similarly, as the special case, we can easily obtain the following corollaries.

Corollary 12. *Consider the NSDSs* (26). *For the given scalars* $\gamma > 0$, $\tau_2 > \tau_1 \geq 0$, $\tau_d > 0$, $\beta > 0$, *and* $\beta_d > 0$, *the NSDSs* (26) *are stochastically asymptotically stable with* H_∞ *performance level for all time-varying delays if there exist symmetric positive-definite matrices* P, V_1, V_2, V_3, W_1, W_2, Z_1, *and* Z_2; *any appropriately dimensioned matrices* N, M, F, *and* H; *and two positive scalars* $\kappa > 0$, $\kappa_d > 0$ *such that the following LMIs hold:*

$$\begin{pmatrix} \Xi & -\tau_2\widehat{M} \\ * & -\tau_2 W_1 \end{pmatrix} < 0,$$

$$\begin{pmatrix} \Xi & -(\tau_2-\tau_1)\widehat{N} \\ * & -(\tau_2-\tau_1)(W_1+W_2) \end{pmatrix} < 0, \tag{45}$$

$$\begin{pmatrix} \Xi & -(\tau_2-\tau_1)\widehat{F} \\ * & -(\tau_2-\tau_1)W_2 \end{pmatrix} < 0,$$

where

$$\Xi = \begin{pmatrix} \Xi_{11} & \Xi_{12} & PD & \Xi_{14} & A^T H & \Xi_{16} \\ * & \Xi_{22} & 0 & 0 & A_d^T H & \Xi_{26} \\ * & * & -\gamma^2 I & 0 & D^T H & 0 \\ * & * & * & \Xi_{44} & \Xi_{45} & 0 \\ * & * & * & * & \Xi_{55} & 0 \\ * & * & * & * & * & \Xi_{66} \end{pmatrix}, \tag{46}$$

$$\Xi_{16} = \begin{pmatrix} -C^T & \Theta_1 P & 0 \end{pmatrix}, \qquad \Xi_{26} = \begin{pmatrix} -C_d^T & 0 & \Theta_2 P \end{pmatrix},$$

$$\Xi_{66} = \operatorname{diag}\{-I, -P, -P\},$$

$$\Xi_{11},\ \Xi_{12},\ \Xi_{14},\ \Xi_{22},\ \Xi_{44},\ \Xi_{45},\ \Xi_{55},$$

defined in Theorem 11.

Corollary 13. *Consider the NSDSs* (26). *For the given scalars* $\gamma > 0$, $\tau_2 > \tau_1 \geq 0$, $\tau_d > 0$, $\beta > 0$, *and* $\beta_d > 0$, *the NSDSs* (26) *are strictly* (Q, S, R)*-passive for all time-varying delays if there exist symmetric positive-definite matrices* P, V_1, V_2, V_3, W_1, W_2, Z_1, Z_2; *any appropriately dimensioned matrices* N, M, F, *and* H; *and two positive scalars* $\kappa > 0$, $\kappa_d > 0$ *such that the following LMIs hold:*

$$\begin{pmatrix} \Xi & -\tau_2\widehat{M} \\ * & -\tau_2 W_1 \end{pmatrix} < 0,$$

$$\begin{pmatrix} \Xi & -(\tau_2-\tau_1)\widehat{N} \\ * & -(\tau_2-\tau_1)(W_1+W_2) \end{pmatrix} < 0, \tag{47}$$

$$\begin{pmatrix} \Xi & -(\tau_2-\tau_1)\widehat{F} \\ * & -(\tau_2-\tau_1)W_2 \end{pmatrix} < 0,$$

where

$$\Xi = \begin{pmatrix} \Xi_{11} & \Xi_{12} & PD - C^T & \Xi_{14} & A^T H & \Xi_{16} \\ * & \Xi_{22} & -C_d^T & 0 & A_d^T H & \Xi_{26} \\ * & * & -\gamma I & 0 & D^T H & 0 \\ * & * & * & \Xi_{44} & \Xi_{45} & 0 \\ * & * & * & * & \Xi_{55} & 0 \\ * & * & * & * & * & \Xi_{66} \end{pmatrix}, \tag{48}$$

$$\Xi_{16} = \begin{pmatrix} \Theta_1 P & 0 \end{pmatrix}, \qquad \Xi_{26} = \begin{pmatrix} 0 & \Theta_2 P \end{pmatrix},$$

$$\Xi_{66} = \operatorname{diag}\{-P, -P\}, \qquad \widehat{N} = \left(N^T, 0_{1\times 9}\right)^T,$$

$$\widehat{M} = \left(M^T, 0_{1\times 9}\right)^T, \qquad \widehat{F} = \left(F^T, 0_{1\times 9}\right)^T,$$

$$\Xi_{11},\ \Xi_{12},\ \Xi_{14},\ \Xi_{22},\ \Xi_{44},\ \Xi_{45},\ \Xi_{55}$$

defined in Theorem 11.

4. Dissipative Delay-Feedback Control for NSDSs

Extending on the results of the foregoing section, our aim is to develope an LMIs-based solution to the problem of designing a delay-feedback controller as

$$u(t, t-\tau(t)) = K_0 x(t) + K_1 x(t-\tau(t)), \tag{49}$$

which will render the NSDSs (10) strictly (Q, S, R)-dissipative. The closed-loop systems is now described by

$$\begin{aligned} dx(t) &= \left\{\widetilde{A}x(t) + f(t, x(t)) + \widetilde{A}_d x(t-\tau(t)) \right. \\ &\quad \left. + f_d(t, x(t-\tau(t))) + Dv(t)\right\} dt \\ &\quad + G(t, x(t), x(t-\tau(t)))\, d\omega(t), \\ z(t) &= Cx(t) + C_d x(t-\tau(t)), \\ x(t) &= \varphi(t), \quad t \in [-\tau_2, 0), \end{aligned} \tag{50}$$

where $\widetilde{A} = A + BK_0$, $\widetilde{A}_d = A_d + BK_1$.

Applying Theorem 7, together with Lemma 5 and congruent transformation, we can get the following theorem without detailed proofs.

Theorem 14. *Consider the NSDSs* (50). *Given some scalars* $\alpha > 0$, $\beta > 0$, $\beta_d > 0$, $\kappa > 0$, *and* $\kappa_d > 0$ *and matrices* $Q = Q^T \leq 0$, $R = R^T > 0$, *and* S, *suppose there exist matrices* $X = X^T > 0$, $W = W^T > 0$, Y, *and* Y_d *such that the following LMI holds:*

$$\begin{pmatrix} \Gamma_1 & A_d X + BY_d & D - XC^T S & \Gamma_2 \\ * & -\widetilde{W} & -XC_d^T S & \Gamma_3 \\ * & * & -R_\alpha & 0 \\ * & * & * & \Gamma_4 \end{pmatrix} < 0; \tag{51}$$

then the NSDSs (50) *are strictly* (Q, S, R)*-dissipative independent of delay; the feedback gain is* $K_0 = YX^{-1}$, $K_1 = Y_d X^{-1}$, *where* $\Gamma_1 = XA^T + AX + Y^T B^T + BY + \widetilde{W}$, $\Gamma_2 = (I\ \ I\ \ XC^T Q\ \ X\Theta_1\ \ 0\ \ \kappa\beta X\ \ 0)$, $\Gamma_3 = (0\ \ 0\ \ XC_d^T Q\ \ 0\ \ X\Theta_2\ \ 0\ \ \kappa_d\beta_d X)$, *and* $\Gamma_4 = \operatorname{diag}(-\kappa I, -\kappa_d I, Q, -X, -X, -\kappa I, -\kappa_d I)$.

Similarly, applying Theorem 11, together with Lemma 5 and congruent transformation, we can obtain the following theorem without detailed proofs.

Theorem 15. *Consider the NSDSs* (50). *For the given scalars* $\alpha > 0$, $\tau_2 > \tau_1 \geq 0$, $\tau_d > 0$, $\kappa > 0$, $\kappa_d > 0$, $\beta > 0$, *and* $\beta_d > 0$, *the NSDSs* (50) *are strictly* (Q, S, R)*-dissipative for all time-varying delays and the feedback gain is* $K_0 = YX^{-1}$, $K_1 = Y_d X^{-1}$, *if there exist symmetric positive-definite matrices* X, $\widetilde{V}_1$, $\widetilde{V}_2$, $\widetilde{V}_3$, $\widetilde{W}_1$, $\widetilde{W}_2$, $\widetilde{Z}_1$, *and* $\widetilde{Z}_2$ *and any appropriately dimensioned matrices* $\widetilde{N}$, $\widetilde{M}$, $\widetilde{F}$, Y, *and* Y_d *such that the following LMIs hold:*

$$\begin{pmatrix} \Delta & -\tau_2 \check{M} \\ * & -\tau_2 \widetilde{W}_1 \end{pmatrix} < 0,$$

$$\begin{pmatrix} \Delta & -(\tau_2 - \tau_1)\check{N} \\ * & -(\tau_2 - \tau_1)\left(\widetilde{W}_1 + \widetilde{W}_2\right) \end{pmatrix} < 0, \quad (52)$$

$$\begin{pmatrix} \Delta & -(\tau_2 - \tau_1)\check{F} \\ * & -(\tau_2 - \tau_1)\widetilde{W}_2 \end{pmatrix} < 0,$$

where

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} & \Delta_{14} & \Delta_{15} & \Delta_{16} \\ * & \Delta_{22} & -XC_d^T S & 0 & \Delta_{25} & \Delta_{26} \\ * & * & -R_\alpha & 0 & D^T & 0 \\ * & * & * & \Delta_{44} & \Delta_{45} & 0 \\ * & * & * & * & \Delta_{55} & 0 \\ * & * & * & * & * & \Delta_{66} \end{pmatrix},$$

$$\Delta_{11} = AX + BY + XA^T + Y^T B^T + \widetilde{V}_1 + \widetilde{V}_2 + \widetilde{V}_3 + \tau_2 \widetilde{Z}_1 + (\tau_2 - \tau_1)\widetilde{Z}_2 + \widetilde{M} + \widetilde{M}^T,$$

$$\Delta_{12} = \widetilde{N} - \widetilde{F} - \widetilde{M} + A_d X + BY_d, \qquad \Delta_{13} = D - XC^T S,$$

$$\Delta_{14} = \begin{pmatrix} \widetilde{F} & -\widetilde{N} & I & I \end{pmatrix}, \qquad \Delta_{15} = Y^T B^T + XA^T,$$

$$\Delta_{16} = \begin{pmatrix} XC^T Q & X\Theta_1 & 0 & \kappa\beta X & 0 \end{pmatrix},$$

$$\Delta_{22} = -\ (1 - \tau_d)\widetilde{V}_2,$$

$$\Delta_{25} = Y_d^T B^T + XA_d^T,$$

$$\Delta_{26} = \begin{pmatrix} XC_d^T Q & 0 & X\Theta_2 & 0 & \kappa_d\beta_d X \end{pmatrix},$$

$$\Delta_{44} = \operatorname{diag}\left\{-\widetilde{V}_1, -\widetilde{V}_3, -\kappa I, -\kappa_d I\right\},$$

$$\Delta_{45} = \begin{pmatrix} 0 & 0 & I & I \end{pmatrix}^T,$$

$$\Delta_{55} = \tau_2 \widetilde{W}_1 + (\tau_2 - \tau_1)\widetilde{W}_2 - X - X^T,$$

$$\Delta_{66} = \operatorname{diag}\left\{Q, -X, -X, -\kappa I, -\kappa_d I\right\},$$

$$\check{N} = \left(\widetilde{N}^T, 0_{1\times 12}\right)^T, \qquad \check{M} = \left(\widetilde{M}^T, 0_{1\times 12}\right)^T,$$

$$\check{F} = \left(\widetilde{F}^T, 0_{1\times 12}\right)^T. \quad (53)$$

5. Numerical Example with Simulation

In this section, we will give an example to show the correctness of the derived results and the effectiveness of the designed controller. Consider the following nonlinear stochastic delay systems:

$$\begin{aligned} dx(t) = \Bigg[& \begin{pmatrix} -0.1 & 1 & 0 & 1 \\ 2 & -1 & 2.5 & 0 \\ -1 & -1.5 & -5 & 0 \\ 0 & 0 & 2 & -5 \end{pmatrix} x(t) + \begin{pmatrix} 0.1x_1 \sin(x_3) \\ 0.4x_2 \cos(x_4) \\ 0.3x_3 \sin(x_1 x_4) \\ 0.5x_4 \cos(x_2 x_3) \end{pmatrix} \\ & + \begin{pmatrix} -1 & 1 & 0 & 1 \\ 2 & -3 & 2.5 & 0 \\ -1 & -2 & -3 & 0 \\ 0 & 0 & 2 & 3 \end{pmatrix} x(t-\tau) \\ & + \begin{pmatrix} 0.1x_1(t-\tau)\sin(x_3(t-\tau)) \\ 0.2x_2(t-\tau)\cos(x_4(t-\tau)) \\ 0.3x_3(t-\tau)\sin(x_1(t-\tau)x_4(t-\tau)) \\ 0.3x_4(t-\tau)\cos(x_2(t-\tau)x_3(t-\tau)) \end{pmatrix} \\ & + \begin{pmatrix} 0 & -1 \\ -0.5 & 0 \\ -0.2 & 0 \\ 0 & -0.2 \end{pmatrix} u(t, t-\tau) + \begin{pmatrix} 0.5 & 0.1 & 0.1 & 0 \\ 0.1 & 0.1 & 0.5 & 0.2 \\ 0 & 0.1 & 0 & 1 \\ 0 & 0 & 1 & 0.4 \end{pmatrix} v(t) \Bigg] dt \\ & + \begin{pmatrix} 0.35x_1 \sin(x_1) + 0.35x_1(t-\tau) \times \cos(x_2(t-\tau)) \\ \dfrac{0.71t}{1+t}(x_2 + x_2(t-\tau)) \\ 0.35x_3 \sin(x_1 x_3) - 0.35x_3(t-\tau) \times \sin\left(x_3^2(t-\tau)\right) \\ \dfrac{14t}{7+100t}(x_4 - x_4(t-\tau)) \end{pmatrix} d\omega(t), \end{aligned}$$

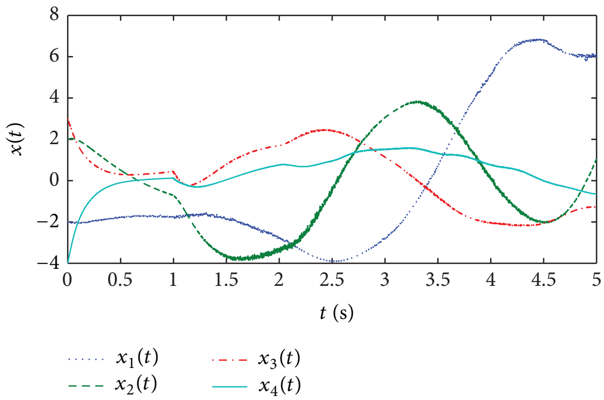


Figure 1: The states curves of open-loop NSDSs in (54) without the control with initial state $(-2, 2, 3, -4)^T$ and time delay $\tau = 1$.

$$z(t) = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0.2 & 1 \end{pmatrix} x(t) + \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0.5 & 0.5 & -1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & -1 \end{pmatrix} x(t-\tau). \tag{54}$$

For Case 1, when the initial condition $x(0) = [-2, 2, 3, -4]^T$ is used and $v(t)$ is a random vector of zero mean and 0.3 standard deviation, we can see that the uncontrolled NSDSs (54) are not stable from Figure 1. Hence the design of dissipative delay-feedback controller is necessary. From (54), we can get $\Theta_1 = \Theta_2 = \mathrm{diag}(0.5, 1, 0.5, 2)$, and $\beta = 0.5$, $\beta_d = 0.3$, for given $S = I$, $R = 5I$, and $Q = -0.2I$; applying Theorem 14 to this example, we have

$$X = \begin{pmatrix} 38.2649 & -9.3804 & -3.6924 & 7.8006 \\ -9.3804 & 25.9896 & 10.5274 & -1.9189 \\ -3.6924 & 10.5274 & 4.5464 & -0.7337 \\ 7.8006 & -1.9189 & -0.7337 & 2.2634 \end{pmatrix},$$
$$Y = 1.0e+004 * \begin{pmatrix} -0.0254 & 1.5597 & 0.6249 & -0.0044 \\ 0.8881 & -0.0450 & -0.0175 & 0.1783 \end{pmatrix}, \tag{55}$$
$$Y_d = 1.0e+003 * \begin{pmatrix} 1.1951 & 0.5087 & 0.2100 & 0.2313 \\ 0.0308 & 0.2403 & 0.0994 & -0.0020 \end{pmatrix}.$$

So the delay-feedback controller parameters can be calculated as follows:

$$K = \begin{pmatrix} 149.4677 & 780.3725 & -306.7689 & 27.4064 \\ 258.0621 & 100.2942 & -67.4539 & -38.6640 \end{pmatrix},$$
$$K_d = \begin{pmatrix} 43.1863 & 33.7078 & 0.3173 & -17.9504 \\ 5.8953 & 7.6550 & 6.9022 & -12.4934 \end{pmatrix}. \tag{56}$$

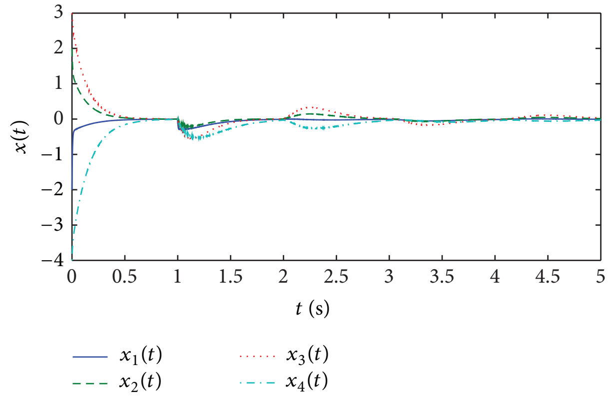


Figure 2: The states curves of closed-loop NSDSs in (54) under the delay-feedback control with initial state $(-2, 2, 3, -4)^T$ and time delay $\tau = 1$.

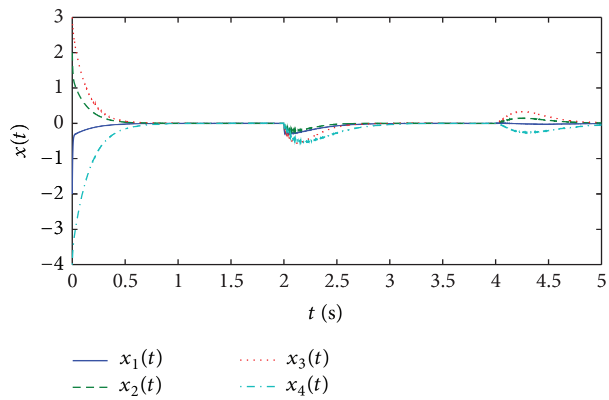


Figure 3: The states curves of closed-loop NSDSs in (54) under the delay-feedback control with initial state $(-2, 2, 3, -4)^T$ and time delay $\tau = 2$.

The states curves and the output curves of closed-loop NSDSs in (54) can be seen in Figures 2, 3, and 4; from Figure 5, we can see that (12) holds. Hence, the closed-loop NSDSs are strictly (Q, S, R)-dissipative; we can also see that the delay-feedback controller is delay-independent.

6. Conclusions

The dissipative delay-feedback control problems for nonlinear stochastic delay systems (NSDSs) have been investigated based on dissipativity analysis. The systems are subjected to stochastic disturbance, nonlinear disturbance, and two cases time-delay effects, which often exist in a wide variety of industrial processes and are the main sources of instability. Based on the Lyapunov stability theory and stochastic analysis technique, both delay-independent and delay-dependent dissipativity criteria have been established in terms of linear matrix inequalities (LMIs). The available results on H_∞ approach and passivity for stochastic delay systems as special cases of the developed results also have been given in this paper. The delay-dependent feedback controller has been

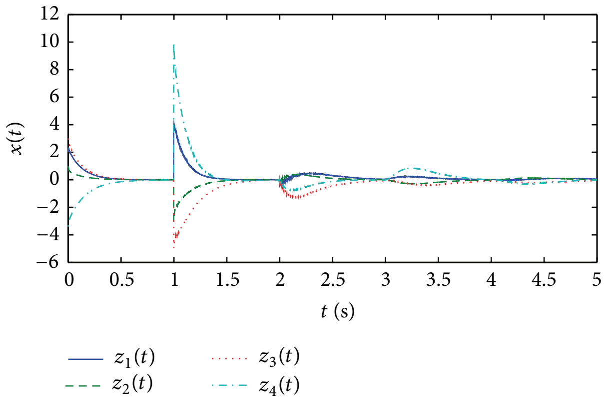


FIGURE 4: The output curves of closed-loop NSDSs in (54) under the delay-feedback control with initial state $(-2, 2, 3, -4)^T$ and time delay $\tau = 1$.

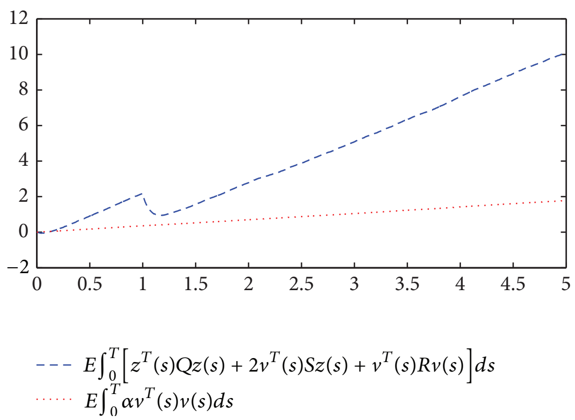


FIGURE 5: The dissipativity level of closed-loop NSDSs in (54) under the delay-feedback control with initial state $(-2, 2, 3, -4)^T$ and time delay $\tau = 1$.

designed by considering the relationship among the time-varying delay, its lower and upper bound, and its difference without ignoring any terms, which effectively reduces the conservative. A numerical example also has been given to verify the effectiveness of the proposed methods.

Conflict of Interests

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

Acknowledgments

This work is supported by the National Science Foundation of China (NSFC) under Grants nos. 61104127, 51079057, and 61134012; China Postdoctoral Science Foundation under Grant no. 2012M521428; and Hubei Province Key Laboratory of Systems Science in Metallurgical Process (Wuhan University of Science and Technology) under Grant no. Y201320.

References

[1] R. Z. Khaminskii, *Stochastic Stability of Differential Equations*, vol. 7, Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands, 1980.

[2] Y. Shen and J. Wang, "An improved algebraic criterion for global exponential stability of recurrent neural networks with time-varying delays," *IEEE Transactions on Neural Networks*, vol. 19, no. 3, pp. 528–531, 2008.

[3] Y. Shen and J. Wang, "Robustness analysis of global exponential stability of recurrent neural networks in the presence of time delays and random disturbances," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 1, pp. 87–96, 2012.

[4] G. Chen and Y. Shen, "Robust reliable H_∞ control for nonlinear stochastic Markovian jump systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 431576, 16 pages, 2012.

[5] Y. Shen and J. Wang, "Noise-induced stabilization of the recurrent neural networks with mixed time-varying delays and Markovian-switching parameters," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1857–1862, 2007.

[6] Y. Shen and J. Wang, "Almost sure exponential stability of recurrent neural networks with Markovian switching," *IEEE Transactions on Neural Networks*, vol. 20, no. 5, pp. 840–855, 2009.

[7] Q. Luo, X. Mao, and Y. Shen, "New criteria on exponential stability of neutral stochastic differential delay equations," *Systems & Control Letters*, vol. 55, no. 10, pp. 826–834, 2006.

[8] C. Wang and Y. Shen, "Improved delay-dependent robust stability criteria for uncertain time delay systems," *Applied Mathematics and Computation*, vol. 218, no. 6, pp. 2880–2888, 2011.

[9] W. Zhang and L. Xie, "Interval stability and stabilization of linear stochastic systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 810–815, 2009.

[10] W. Zhang, B.-S. Chen, and C.-S. Tseng, "Robust H_∞ filtering for nonlinear stochastic systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, part 1, pp. 589–598, 2005.

[11] B.-S. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[12] W. Zhang, Y. Huang, and H. Zhang, "Stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise," *Automatica*, vol. 43, no. 3, pp. 513–521, 2007.

[13] J. C. Willems, "Dissipative dynamical systems. I: general theory," *Archive for Rational Mechanics and Analysis*, vol. 45, pp. 321–351, 1972.

[14] D. Hill and P. Moylan, "The stability of nonlinear dissipative systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 5, pp. 708–711, 1976.

[15] Z. Q. Tan, Y. C. Soh, and L. Xie, "Dissipative control for linear discrete-time systems," *Automatica*, vol. 35, no. 9, pp. 1557–1564, 1999.

[16] M. S. Mahmoud, "Delay-dependent dissipativity analysis and synthesis of switched delay systems," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 1, pp. 1–20, 2011.

[17] M. S. Mahmoud, Y. Shi, and F. M. AL-Sunni, "Dissipativity analysis and synthesis of a class of nonlinear systems with time-varying delays," *Journal of the Franklin Institute: Engineering and Applied Mathematics*, vol. 346, no. 6, pp. 570–592, 2009.

[18] K. Youcef-Toumi and O. Itô, "A time delay controller for systems with unknown dynamics," *Journal of Dynamic Systems, Measurement, and Control*, vol. 112, no. 1, pp. 133–142, 1990.

[19] V. Kapila, A. Tzes, and Q. Yan, "Closed-loop input shaping for flexible structures using time-delay control," *Journal of Dynamic Systems, Measurement and Control*, vol. 122, no. 3, pp. 454–460, 2000.

[20] X. Mao, J. Lam, and L. Huang, "Stabilisation of hybrid stochastic differential equations by delay feedback control," *Systems & Control Letters*, vol. 57, no. 11, pp. 927–935, 2008.

[21] D. J. Hill and P. J. Moylan, "Dissipative dynamical systems: basic input-output and state properties," *Journal of the Franklin Institute*, vol. 309, no. 5, pp. 327–357, 1980.

[22] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15, SIAM, Philadelphia, Pa, USA, 1994.

[23] K. Gu, "An integral inequality in the stability problem of time-delay systems," in *Proceedings of the 39th IEEE Confernce on Decision and Control*, pp. 2805–2810, Sydney, Australia, December 2000.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 278142, 9 pages
http://dx.doi.org/10.1155/2014/278142

Research Article

Discrete-Time Indefinite Stochastic Linear Quadratic Optimal Control with Second Moment Constraints

Weihai Zhang[1] and Guiling Li[2]

[1] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China*
[2] *College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Weihai Zhang; w_hzhang@163.com

Received 1 April 2014; Accepted 29 April 2014; Published 18 May 2014

Academic Editor: Xuejun Xie

This paper studies the discrete-time stochastic linear quadratic (LQ) problem with a second moment constraint on the terminal state, where the weighting matrices in the cost functional are allowed to be indefinite. By means of the matrix Lagrange theorem, a new class of generalized difference Riccati equations (GDREs) is introduced. It is shown that the well-posedness, and the attainability of the LQ problem and the solvability of the GDREs are equivalent to each other.

1. Introduction

LQ control, initiated by Kalman [1] and extended to stochastic systems by Wonham [2], is one of the most important classes of optimal control issues from both theory and application point of view; we refer the reader to [2–8] for representative work in this area. Different from the classical LQ in modern control theory, it was found in [9, 10] that a stochastic LQ problem with indefinite control weighting matrices can still be well-posed, which evoked a series of subsequent researches; see, for example, [11, 12].

It is well known that in practical engineering, the system state and control input are always subject to various constraints, so how to solve the constrained stochastic LQ issue is a more attractive topic; we refer the reader to [13–19]. Reference [14] presented a tractable approach for LQ controller design of the system with additive noise. Reference [16] was about the constrained LQ of deterministic systems with state equality constraints. Reference [13] studied the parametrization of the solutions of finite-horizon constrained LQ control. Reference [15] was devoted to a stochastic LQ optimal control and an application to portfolio selection, where the control variable is confined to a cone, and all the coefficients of the state equation are random processes. Reference [19] studied the indefinite stochastic LQ control problem of continuous-time Itô systems with a linear equality constraint $Mx(T) = \xi$ on the terminal state and gave a necessary condition for the existence of an optimal controller. Reference [20] generalized the results of [19] to discrete-time stochastic systems.

In this paper, different from [19, 20] on the constraint conditions, we would like to deal with stochastic LQ control of discrete-time multiplicative noise systems with a second moment constraint $E[x(T)^T x(T)] = c$ and such constraints are often encountered in H_∞ filtering design; see [21, 22]. By means of Lagrange theorem, we present a necessary condition for the existence of an optimal linear state feedback control with the second moment constraint on the terminal states. It is proved that the solvability of GDRE is necessary and sufficient for the existence of an optimal control under either of the state feedback case or of the open-loop forms. Moreover, we show that the well-posedness and the attainability of the constrained LQ problem, the feasibility of the LMI, and the solvability of the GDRE are equivalent to each other. The novel contribution of this paper is to consider a constrained discrete-time LQ optimal stochastic control, which includes some results of [23] as special cases. A new class of generalized difference Riccati equations (GDREs) is first introduced.

The remainder of the paper is organized as follows. Section 2 gives some definitions and preliminaries. In Section 3, the optimal state feedback control is studied using the matrix Lagrange theorem. We give a necessary and sufficient condition for the well-posedness of the constrained LQ control in Section 4. Section 5 shows the equivalence among the well-posedness and the attainability of the LQ problem, the feasibility of the LMI, and the solvability of the GDRE. The set of all optimal controls is determined. We conclude the paper in Section 6.

Throughout the paper, the following notations are adopted: M^T denotes the transpose of M. $M > 0$ ($M \geq 0$): M is a positive definite (positive semidefinite) symmetric matrix. $\mathrm{tr}(M)$: the trace of a square matrix M. $R^{m\times n}$: the space of all $m \times n$ matrices. S^n: the space of all $n \times n$ symmetric matrices.

2. Problem Setting

Consider the following constrained discrete-time stochastic LQ control problem.

Problem 1. Consider

$$\begin{aligned}
\min_{u(t_0),\ldots,u(T-1)} \quad & J\left(t_0, x_0; u\left(t_0\right), u\left(t_0+1\right), \ldots, u(T-1)\right) \\
\text{s.t.} \quad & x(t+1) = [A(t)x(t) + B(t)u(t)] \\
& \qquad + [C(t)x(t) + D(t)u(t)]\, w(t), \\
& x\left(t_0\right) = x_0, \\
& E\left\{x(T)^T x(T)\right\} = c, \\
& t \in \left\{t_0, t_0+1, \ldots, T-1\right\},
\end{aligned} \tag{1}$$

where the state $x(t) \in R^n$, the control input $u(t) \in R^m$, and the noise $w(t) \in R^1$, $t \in \{t_0, t_0+1, \ldots, T-1\}$,

$$\begin{aligned}
& J\left(t_0, x_0; u\left(t_0\right), u\left(t_0+1\right), \ldots, u(T-1)\right) \\
& := E\left\{x(T)^T Q(T) x(T)\right. \\
& \qquad \left. + \sum_{t=t_0}^{T-1}\left[x(t)^T Q(t) x(t) + u(t)^T R(t) u(t)\right]\right\}.
\end{aligned} \tag{2}$$

The process $\{w(t_0), w(t_0+1), \ldots, w(T-1)\}$ is a sequence of second-order stationary random variables defined on a complete probability space $(\Omega, \mathscr{F}, \mathscr{P})$. Without loss of generality, we assume that

$$E\{w(s)\} = 0, \qquad E\{w(s)w(t)\} = \delta_{st}, \tag{3}$$

where δ_{st} is the Kronecker delta, $s, t \in \{t_0, t_0+1, \ldots, T-1\}$. $c \geq 0$ is a constant, $A(t), B(t), C(t), D(t), Q(t)$, and $R(t)$ are matrices having appropriate dimensions determined from context, and $Q(t)$ and $R(t)$ are real symmetric indefinite matrices. x_0 is a given deterministic vector.

Definition 2. Problem 1 is called well-posed, if $\forall x_0 \in R^n$,

$$\begin{aligned}
V\left(x_0\right) = & \inf_{u(t_0), u(t_0+1), \ldots, u(T-1)} J\left(t_0, x_0; u\left(t_0\right), \ldots, u(T-1)\right) \\
& > -\infty.
\end{aligned} \tag{4}$$

Definition 3. Problem 1 is called attainable, if $\forall x_0 \in R^n$, there exists a sequence $\{u^*(t_0), u^*(t_0+1), \ldots, u^*(T-1)\}$, such that $V(x_0) = J(t_0, x_0; u^*(t_0), u^*(t_0+1), \ldots, u^*(T-1))$. In this case, $\{u_0^*, u_1^*, \ldots, u_{N-1}^*\}$ is called an optimal control sequence.

Now, let us consider a mathematical programming (MP) problem in a matrix space:

$$\begin{aligned}
\min \quad & f(X) \\
\text{s.t.} \quad & \mathbf{h}(X) = \mathbf{0}.
\end{aligned} \tag{5}$$

Definition 4. Let X^* be a point satisfying

$$\mathbf{h}\left(X^*\right) = \left(h_1\left(X^*\right), \ldots, h_p\left(X^*\right)\right)^T = \mathbf{0}, \tag{6}$$

and then X^* is said to be a constraint *regular point* if the gradient vectors $\nabla h_j(X^*)$, $j = 1, \ldots, p$, are linearly independent.

Lemma 5 (Lagrange theorem [24]). *Assume that the functions f, $h_1, \ldots, h_p$, are twice continuously differentiable. If a regular point X^* is also a relative minimum point for the original MP, then there exists a vector $\lambda \in R^p$ such that*

$$\nabla_X L\left(X^*, \lambda^*\right) = \mathbf{0}, \tag{7}$$

where the Lagrangian function $L(X, \lambda) := f(X) + \lambda^T \mathbf{h}(X)$.

3. A Necessary Condition for State Feedback Control

In this section, by the matrix Lagrange theorem, we present a necessary condition for Problem 1 based on a new type of GDREs.

Let $X(t) = E[x(t)x(t)^T]$. Through a simple calculation, the following deterministic optimal control Problem 6 is equivalent to the original Problem 1 under the state feedback $u(t) = K(t)x(t)$ for $t \in \{t_0, t_0+1, \ldots, T-1\}$.

Problem 6. Consider

$$\begin{aligned}
\min_{K(t_0),\dots,K(T-1)} \quad & J\left(t_0, x_0; K\left(t_0\right) x_0, \dots, K\left(T-1\right) x\left(T-1\right)\right) \\
\text{s.t.} \quad & X\left(t+1\right) = \left(A\left(t\right) + B\left(t\right) K\left(t\right)\right) \\
& \qquad \times X\left(t\right)\left(A\left(t\right) + B\left(t\right) K\left(t\right)\right)^T \\
& \qquad + \left(C\left(t\right) + D\left(t\right) K\left(t\right)\right) \\
& \qquad \times X\left(t\right)\left(C\left(t\right) + D\left(t\right) K\left(t\right)\right)^T, \\
& t = t_0, t_0 + 1, \dots, T-1, \\
& X\left(t_0\right) = X_0 = x_0 x_0^T, \\
& \operatorname{tr}\left[X\left(T\right)\right] = c
\end{aligned} \tag{8}$$

with

$$\begin{aligned}
J&\left(t_0, x_0; K\left(t_0\right) x_0, \dots, K\left(T-1\right) x\left(T-1\right)\right) \\
&= \sum_{t=t_0}^{T-1} \operatorname{tr}\left\{\left[Q\left(t\right) + K(t)^T R\left(t\right) K\left(t\right)\right] X\left(t\right)\right\} \\
&\quad + \operatorname{tr}\left[Q\left(T\right) X\left(T\right)\right].
\end{aligned} \tag{9}$$

Remark 7. If Problem 1 has a linear feedback optimal control solution $u^*(t) = K^*(t)x(t)$, $t \in \{t_0, t_0+1, \dots, T-1\}$, then $K^*(t)$, $t \in \{t_0, t_0+1, \dots, T-1\}$, are the optimal solution of Problem 6.

Theorem 8. *If Problem 1 is attainable by $u(t) = K^*(t)x(t)$, and the regular point $(K^*(t), X^*(t))$ is a locally optimal solution of Problem 6, then there exist symmetric matrices $P(t), t \in \{t_0, t_0+1, \dots, T-1\}$, and $\lambda \in R^1$ solving the following GDRE:*

$$\begin{aligned}
P\left(t\right) &= A(t)^T P\left(t+1\right) A\left(t\right) + C(t)^T P\left(t+1\right) C\left(t\right) \\
&\quad + Q\left(t\right) - H(t)^T G(t)^{\dagger} H\left(t\right), \\
H\left(t\right) &= B(t)^T P\left(t+1\right) A\left(t\right) + D(t)^T P\left(t+1\right) C\left(t\right), \\
G\left(t\right) &= R\left(t\right) + B(t)^T P\left(t+1\right) B\left(t\right) \\
&\quad + D(t)^T P\left(t+1\right) D\left(t\right) \geq \mathbf{0}, \\
G\left(t\right) G(t)^{\dagger} H\left(t\right) &= H\left(t\right), \quad t \in \left\{t_0, t_0+1, \dots, T-1\right\}, \\
P\left(T\right) &= Q\left(T\right) + \lambda I,
\end{aligned} \tag{10}$$

where $G^{\dagger}$ is the Moore-Penrose generalized inverse of G. Moreover,

$$K^*\left(t\right) = -G(t)^{\dagger} H\left(t\right) + Y\left(t\right) - G(t)^{\dagger} G\left(t\right) Y\left(t\right) \tag{11}$$

with $Y(t) \in R^{m\times n}$, $t = t_0, t_0+1, \dots, T-1$, being any given real matrices:

$$\begin{aligned}
V\left(x_0\right) &= J\left(t_0, x_0; u^*\left(t_0\right), u^*\left(t_0+1\right), \dots, u^*\left(T-1\right)\right) \\
&= x_0^T P\left(t_0\right) x_0 - c\lambda.
\end{aligned} \tag{12}$$

To prove Theorem 8, we mainly use Lemma 5 to Problem 6 together with the following lemma to obtain GDRE (10) and then apply the technique of completing squares to show (12).

Lemma 9 (see [12]). *Let A, B, C be given matrices with appropriate sizes; then the matrix equation*

$$AXB = C \tag{13}$$

has a solution X if and only if

$$AA^{\dagger}CB^{\dagger}B = C. \tag{14}$$

Moreover, any solution to $AXB = C$ can be represented by

$$X = A^{\dagger}CB^{\dagger} + Y - A^{\dagger}AYBB^{\dagger}, \tag{15}$$

where Y is any matrix with appropriate size.

Proof. According to Remark 7, $K^*(t)$ is also the optimal solution of Problem 6. Problem 6 is a typical MP problem about $X(t)$ and $K(t)$ as follows:

$$\begin{aligned}
\min \quad & f\left[X\left(t\right), K\left(t\right)\right] \\
\text{s.t.} \quad & h_{t+1}\left[X\left(t\right), K\left(t\right)\right] = 0, \quad t \in \left\{t_0, t_0+1, \dots, T-1\right\}, \\
& h\left[X\left(T\right)\right] = 0,
\end{aligned} \tag{16}$$

where

$$\begin{aligned}
f\left[X\left(t\right), K\left(t\right)\right] &= \sum_{t=t_0}^{T-1} \operatorname{tr}\left\{\left[Q\left(t\right) + K(t)^T R\left(t\right) K\left(t\right)\right] X\left(t\right)\right\} \\
&\quad + \operatorname{tr}\left[Q\left(T\right) X\left(T\right)\right], \\
h_{t+1}\left[X\left(t\right), K\left(t\right)\right] &= A\left(t\right) X\left(t\right) A(t)^T \\
&\quad + A\left(t\right) X\left(t\right) K(t)^T B(t)^T \\
&\quad + B\left(t\right) K\left(t\right) X\left(t\right) A(t)^T \\
&\quad + B\left(t\right) K\left(t\right) X\left(t\right) K(t)^T B(t)^T \\
&\quad + C\left(t\right) X\left(t\right) C(t)^T \\
&\quad + C\left(t\right) X\left(t\right) K(t)^T D(t)^T \\
&\quad + D\left(t\right) K\left(t\right) X\left(t\right) C(t)^T \\
&\quad + D\left(t\right) K\left(t\right) X\left(t\right) K(t)^T D(t)^T \\
&\quad - X\left(t+1\right), \\
h\left[X\left(T\right)\right] &= \operatorname{tr}\left[X\left(T\right)\right] - c.
\end{aligned} \tag{17}$$

Let matrices $P(t+1)$, $t \in \{t_0, t_0+1, \dots, T-1\}$, be the Lagrangian multipliers of

$$h_{t+1}\left[X\left(t\right), K\left(t\right)\right] = 0, \quad t \in \left\{t_0, t_0+1, \dots, T-1\right\}, \tag{18}$$

and let $\lambda \in R^1$ be the Lagrangian multiplier of $h[X(T)] = 0$; then the Lagrangian function

$$\begin{aligned}\mathscr{L} = & f[X(t), K(t)] \\ & + \sum_{t=t_0}^{T-1} \operatorname{tr}\{P(t+1) h_{t+1}[X(t), K(t)]\} + \lambda h[X(T)].\end{aligned} \tag{19}$$

According to the the matrix Lagrange theorem, we obtain

$$\frac{\partial \mathscr{L}}{\partial (K_t)} = 0, \quad t = t_0, t_0+1, \ldots, T-1, \tag{20}$$

$$\frac{\partial \mathscr{L}}{\partial (X_t)} = 0 \quad t = t_0, t_0+1, \ldots, T. \tag{21}$$

Based on the partial rule of gradient matrices, (20) can be transformed into

$$\begin{aligned}&\left[R(t) + B(t)^T P(t+1) B(t) + D(t)^T P(t+1) D(t)\right] K(t) \\ &\quad + B(t)^T P(t+1) A(t) + D(t)^T P(t+1) C(t) = 0.\end{aligned} \tag{22}$$

Let

$$\begin{aligned}G(t) &= R(t) + B(t)^T P(t+1) B(t) + D(t)^T P(t+1) D(t), \\ H(t) &= B(t)^T P(t+1) A(t) + D(t)^T P(t+1) C(t).\end{aligned} \tag{23}$$

Then we obtain

$$\begin{aligned}G(t) G(t)^\dagger H(t) &= H(t), \\ G(t) K(t) + H(t) &= 0.\end{aligned} \tag{24}$$

Applying Lemma 9, we have

$$\begin{aligned}K^*(t) &= -G(t)^\dagger H(t) + Y(t) - G(t)^\dagger G(t) Y(t), \\ Y(t) &\in R^{m\times n}, \quad t = t_0, t_0+1, \ldots, T-1.\end{aligned} \tag{25}$$

Equation (21) yields

$$\begin{aligned}P(T) &= Q(T) + \lambda I, \\ P(t) &= Q(t) + A(t)^T P(t+1) A(t) + C(t)^T P(t+1) C(t) \\ &\quad + K(t)^T \left[R(t) + B(t)^T P(t+1) B(t) \right.\\ &\qquad \left. + D(t)^T P(t+1) D(t)\right] K(t) \\ &\quad + K(t)^T \left[D(t)^T P(t+1) C(t) \right.\\ &\qquad \left. + B(t)^T P(t+1) A(t)\right] \\ &\quad + \left[A(t)^T P(t+1) B(t) \right.\\ &\qquad \left. + C(t)^T P(t+1) D(t)\right] K(t).\end{aligned} \tag{26}$$

Substituting $K^*(t) = -G(t)^\dagger H(t) + Y(t) - G(t)^\dagger G(t) Y(t)$ into (26), it follows

$$\begin{aligned}P(t) = & A(t)^T P(t+1) A(t) + C(t)^T P(t+1) C(t) \\ & + Q(t) - H(t)^T G(t)^\dagger H(t).\end{aligned} \tag{27}$$

Without loss of generality, we can assume that P is symmetric. Otherwise, we can take $\overline{P} = (P^T + P)/2$. The objective functional

$$\begin{aligned}&J(t_0, x_0; u(t_0), u(t_0+1), \ldots, u(T-1)) \\ &\quad = E\left\{\sum_{t=t_0}^{T-1} \left[x(t)^T Q(t) x(t) + u(t)^T R(t) u(t)\right]\right\} \\ &\qquad + \operatorname{tr}[X(T) Q(T)] \\ &\quad = E\left\{\sum_{t=t_0}^{T-1} \left[x(t)^T Q(t) x(t) + u(t)^T R(t) u(t) \right.\right.\\ &\qquad\qquad + x(t+1)^T P(t+1) x(t+1) \\ &\qquad\qquad \left.\left. - x(t)^T P(t) x(t)\right]\right\} \\ &\qquad + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\} \\ &\quad = E\left\{\sum_{t=t_0}^{T-1} \left[x(t)^T Q(t) x(t) + u(t)^T R(t) u(t) \right.\right.\\ &\qquad\qquad \left. - x(t)^T P(t) x(t)\right] \\ &\qquad\qquad + \left[x(t)^T A(t)^T + u(t)^T B(t)^T\right] \\ &\qquad\qquad \times P(t+1) [A(t) x(t) + B(t) u(t)] \\ &\qquad\qquad + \left[x(t)^T C(t)^T + u(t)^T D(t)^T\right] \\ &\qquad\qquad \left. \times P(t+1) [C(t) x(t) + D(t) u(t)]\right\} \\ &\qquad + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\}.\end{aligned} \tag{28}$$

A completion of square implies

$$\begin{aligned}&J(t_0, x_0; u(t_0), u(t_0+1), \ldots, u(T-1)) \\ &\quad = E\left\{\sum_{t=t_0}^{T-1} \left[u(t) + G(t)^\dagger H(t) x(t)\right]^T \right.\\ &\qquad\qquad \left. \times G(t) \left[u(t) + G(t)^\dagger H(t) x(t)\right]\right\} \\ &\qquad + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\}.\end{aligned} \tag{29}$$

We assert that $P(t+1)$ must satisfy

$$G(t) = R(t) + B(t)^T P(t+1) B(t) + D(t)^T P(t+1) D(t) \geq 0. \tag{30}$$

If it is not so, there is $G(l)$ for $l \in \{t_0, \dots, T-1\}$ with a negative eigenvalue $\lambda < 0$. Denote the unitary eigenvector with respect to λ by v_λ. Let $\delta \neq 0$ be an arbitrary scalar; we construct a control sequence as follows:

$$\tilde{u}(t) = \begin{cases} -G(t)^\dagger H(t) x(t), & t \neq l, \\ \delta|\lambda|^{-1/2} v_\lambda - G(t)^\dagger H(t) x(t), & t = l. \end{cases} \tag{31}$$

The associated cost functional becomes

$$\begin{aligned} & J(t_0, x_0; \tilde{u}(t_0), \tilde{u}(t_0+1), \dots, \tilde{u}(T-1)) \\ &= E\left\{\sum_{t=t_0}^{T-1} \left[\tilde{u}(t) + G(t)^\dagger H(t) x(t)\right]^T \times G(t)\left[\tilde{u}(t) + G(t)^\dagger H(t) x(t)\right]\right\} \\ &\quad + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\} \\ &= \left[\delta|\lambda|^{-1/2} v_\lambda\right]^T G(l) \left[\delta|\lambda|^{-1/2} v_\lambda\right] \\ &\quad + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\} \\ &= -\delta^2 + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\}. \end{aligned} \tag{32}$$

Let $\delta \to \infty$; then $J(t_0, x_0; \tilde{u}(t_0), \tilde{u}(t_0+1), \dots, \tilde{u}(T-1)) \to -\infty$, which contradicts the attainability of Problem 1. So (30) holds.

In view of (29) and (30), (11) and (12) are easily derived. The proof is completed. □

Remark 10. In Theorem 8, in order to apply matrix Lagrange theorem, we assume the optimal solution $(K_*(t), X_*(t))^T$ is a regular point. Generally speaking, for a given LQ control, it is easy to examine the regular condition.

Below, we present a numerical example to illustrate the effectiveness of Theorem 8.

Example 11. In Problem 1, we set

$$\begin{gathered} E\left\{x(2)^T x(2)\right\} = c = 73, \qquad x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ A_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \\ B_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ C_0 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \qquad C_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \\ D_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \end{gathered} \tag{33}$$

The state and control weighting matrices are as

$$\begin{gathered} Q_0 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad Q_1 = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \\ Q_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad R_0 = -1, \qquad R_1 = -7. \end{gathered} \tag{34}$$

By the relationship between Problems 1 and 6, we know

$$X_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \operatorname{tr}[X(2)] = 73. \tag{35}$$

Applying Theorem 8, we obtain

$$\begin{gathered} X(0) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad X(1) = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}, \\ X(2) = \begin{bmatrix} 34 & -26 \\ -26 & 39 \end{bmatrix}, \\ P(2) = \lambda I = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \lambda = 2. \end{gathered} \tag{36}$$

Stage 2. Consider

$$\begin{gathered} G(1) = 1 > 0, \qquad H(1) = (2, 0), \\ P(1) = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \qquad K(1) = (-2, 0). \end{gathered} \tag{37}$$

Stage 1. Consider

$$\begin{gathered} G(0) = 1 > 0, \qquad H(0) = (2, -1), \\ K(0) = (-2, 1), \qquad P(0) = \begin{bmatrix} 5 & -3 \\ -3 & 0 \end{bmatrix}. \end{gathered} \tag{38}$$

The optimal cost value of Problem 1 is

$$V(x_0) = E\left[x_0^T P(t_0) x_0 - c\lambda\right] = -146. \tag{39}$$

We are able to test the regular condition of $(K^*(t), X^*(t))^T$ as follows. In Problem 6,

$$\begin{aligned} h_{(t+1)}(X(t), K(t)) = {} & A(t) X(t) A(t)^T \\ & + A(t) X(t) K(t)^T B(t)^T \\ & + B(t) K(t) X(t) A(t)^T \\ & + B(t) K(t) X(t) K(t)^T B(t)^T \\ & + C(t) X(t) C(t)^T \\ & + C(t) X(t) K(t)^T D(t)^T \\ & + D(t) K(t) X(t) C(t)^T \\ & + D(t) K(t) X(t) K(t)^T D(t)^T \\ & - X(t+1), \quad t = 0, 1, \end{aligned} \tag{40}$$

which is linear about $X(t)$ and quadratic about $K(t)$, while $h(X(T)) = \operatorname{tr}[X(T)] - c$ is linear about $X(T)$. By simple calculations, $\nabla h_1(K^*(0), X^*(0))$, $\nabla h_2(K^*(1), X^*(1))$, and $\nabla h(X^*(T))$ are all nonzero vectors and hence are linearly independent.

4. Well-Posedness

In this section, we first establish the link between the well-posedness of Problem 1 and the feasibility of some LMIs and then prove that the solvability of GDRE (10) is not only necessary but also sufficient to the well-posedness of Problem 1. Moreover, the well-posedness and the attainability of Problem 1, the feasibility of some LMIs, and the solvability of GDRE (10) are equivalent to each other.

Theorem 12. *Problem 1 is well-posed if there exist symmetric matrices $P(t), t \in \{t_0, t_0+1, \ldots, T-1\}$, and $\lambda \in R^1$ solving the following LMIs:*

$$\overline{M}_t := \begin{bmatrix} A(t)^T P(t+1) A(t) - P(t) + C(t)^T P(t+1) C(t) + Q(t) & H(t)^T \\ H(t) & G(t) \end{bmatrix} \geq 0, \tag{41}$$

$$P(T) \leq Q(T) + \lambda I, \tag{42}$$

where

$$\begin{aligned} H(t) &= B(t)^T P(t+1) A(t) + D(t)^T P(t+1) C(t), \\ G(t) &= R(t) + B(t)^T P(t+1) B(t) + D(t)^T P(t+1) D(t). \end{aligned} \tag{43}$$

Proof. Note that

$$\begin{aligned} &J(t_0, x_0; u(t_0), u(t_0+1), \ldots, u(T-1)) \\ &\quad = E\left\{\sum_{t=t_0}^{T-1}\left[x(t)^T Q(t) x(t) + u(t)^T R(t) u(t)\right] + x(T)^T Q(T) x(T)\right\} \\ &\quad = E\left\{\sum_{t=t_0}^{T-1}\left[x(t)^T Q(t) x(t) + u(t)^T R(t) u(t) + x(t+1)^T P(t+1) x(t+1) - x(t)^T P(t) x(t)\right] \right. \\ &\qquad \left. + x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\} \\ &\quad = E\left\{\sum_{t=t_0}^{T-1}\begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \overline{M}_t \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} + x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\}. \end{aligned} \tag{44}$$

By (41), it is easy to deduce that the cost functional is bounded from below by

$$\begin{aligned} &E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\} \\ &\quad \geq x_0^T P(t_0) x_0 - c\lambda. \end{aligned} \tag{45}$$

Hence, Problem 1 is well-posed. □

Remark 13. Theorem 12 tells us that any symmetric matrices $P(t), t \in \{t_0, t_0+1, \ldots, T-1\}$, and $\lambda \in R^1$ satisfying LMIs (41)-(42) provide a lower bound

$$x_0^T P(t_0) x_0 - c\lambda \tag{46}$$

for the cost function. In what follows, we will show that this lower bound is an exact optimal cost value if $P(t)$ and $\lambda \in R^1$ solve GDRE (10).

We have shown that if the LMIs (41)-(42) are satisfied, then the constrained LQ Problem 1 is well-posed. Below, we further show some other equivalent conditions.

Lemma 14 (extended Schur's lemma [25]). *Let the matrices $M = M^T$, H, $G = G^T$ be given with appropriate sizes. Then, the following three conditions are equivalent:*

$$\begin{aligned} &(1)\ M - HG^{\dagger}H^T \geq 0, \qquad G \geq 0, \qquad H\left(I - GG^{\dagger}\right) = 0. \\ &(2)\ \begin{bmatrix} M & H \\ H^T & G \end{bmatrix} \geq 0. \\ &(3)\ \begin{bmatrix} G & H^T \\ H & M \end{bmatrix} \geq 0. \end{aligned} \tag{47}$$

Theorem 15. *Problem 1 is well-posed if and only if there exist symmetric matrices $P(t), t \in \{t_0, t_0+1, \ldots, T-1\}$, and $\lambda \in R^1$ satisfying GDRE (10). Furthermore, the optimal cost is*

$$\begin{aligned} V(x_0) &= x_0^T P(t_0) x_0 - E\left\{x(T)^T [Q(T) - P(T)] x(T)\right\} \\ &= x_0^T P(t_0) x_0 - c\lambda. \end{aligned} \tag{48}$$

A key to prove Theorem 15 is the necessity part, where the stochastic optimization principle is used.

Proof. Necessity. For $t_0 \le l \le T-1$, define

$$\begin{aligned} V^l&[x(l)] \\ &= \inf_{u(l),\ldots,u(T-1)} E\left\{\sum_{t=l}^{T-1}\left[x(t)^T Q(t)x(t) + u(t)^T R(t)u(t)\right]\right. \\ &\qquad\left. + x(T)^T Q(T)x(T)\right\}. \end{aligned} \tag{49}$$

By the stochastic optimization principle, when $V^{l_1}[x(l_1)]$ is finite, then so is $V^{l_2}[x(l_2)]$ for any $l_1 \le l_2$. Since Problem 1 is assumed to be well-posed at t_0, $V^l[x(l)]$ is finite at any stage $0 \le l \le T-1$. Now let us start with $l = T-1$, and let $P(T) = Q(T) - \lambda I$, and we have

$$\begin{aligned} V^{T-1}&[x(T-1)] - E\left\{x(T)^T[Q(T)-P(T)]x(T)\right\} \\ &= \inf_{u(T-1)}\left\{E\left[x(T-1)^T Q(T-1)x(T-1)\right.\right. \\ &\qquad\left.\left. + u(T-1)^T R(T-1)u(T-1)\right]\right\} \\ &\quad + E\left[x(T)^T P(T)x(T)\right] \\ &= \inf_{u(T-1)} E\left\{x(T-1)^T\right. \\ &\qquad\times\left[Q(T-1) + A(T-1)^T P(T)A(T-1)\right. \\ &\qquad\quad\left. + C(T-1)^T P(T)C(T-1)\right]x(T-1) \\ &\qquad\times 2x(T-1)^T \\ &\qquad\times\left[B(T-1)^T P(T)A(T-1)\right. \\ &\qquad\quad\left. + D(T-1)^T P(T)C(T-1)\right]u(T-1) \\ &\qquad + u(T-1)^T \\ &\qquad\times\left[R(T-1) + B(T-1)^T P(T)B(T-1)\right. \\ &\qquad\quad\left.\left. + D(T-1)^T P(T)D(T-1)\right]u(T-1)\right\}. \end{aligned} \tag{50}$$

Since $V^{T-1}[x(T-1)]$ is finite, using Lemma 4.3 of [23], there exists a symmetric matrix $P(T-1)$ such that

$$\begin{aligned} V^{T-1}&[x(T-1)] - E\left\{x(T)^T[Q(T)-P(T)]x(T)\right\} \\ &= E\left[x(T-1)^T P(T-1)x(T-1)\right], \\ P(T-1) &= A(T-1)^T P(T)A(T-1) \\ &\quad + C(T-1)^T P(T)C(T-1) \\ &\quad + Q(T-1) - H(T-1)^T \\ &\quad \times G(T-1)^{\dagger} H(T-1), \\ H(T-1) &= B(T-1)^T P(T)A(T-1) \\ &\quad + D(T-1)^T P(T)C(T-1), \\ G(T-1) &= R(T-1) + B(T-1)^T P(T)B(T-1) \\ &\quad + D(T-1)^T P(T)D(T-1) \ge \mathbf{0}. \end{aligned} \tag{51}$$

The obtained solution sequence of symmetric matrices $P(t)$, $t = l, l+1, \ldots, T-1$, and $\lambda \in R^1$ to GDRE (10) satisfy

$$\begin{aligned} V^l[x(l)] &- E\left\{x(T)^T[Q(T)-P(T)]x(T)\right\} \\ &= E\left[x(l)^T P(l)x(l)\right]. \end{aligned} \tag{52}$$

Then by the stochastic optimality principle, the following holds:

$$\begin{aligned} V^{l-1}&[x(l-1)] \\ &= \inf_{u(l-1)} E\left\{x(l-1)^T Q(l-1)x(l-1)\right. \\ &\qquad + u(l-1)^T R(l-1)u(l-1) \\ &\qquad\left. + V^l[x(l)]\right\} \\ &= \inf_{u(l-1)} E\left[x(l-1)^T Q(l-1)x(l-1)\right. \\ &\qquad + u(l-1)^T R(l-1)u(l-1) \\ &\qquad\left. + x(l)^T P(l)x(l)\right] \\ &\quad + E\left\{x(T)^T[Q(T)-P(T)]x(T)\right\} \\ &= \inf_{u(l-1)} E\left\{x(l-1)^T\right. \\ &\qquad\times\left[Q(l-1) + A(l-1)^T P(l)A(l-1)\right. \\ &\qquad\quad\left. + C(l-1)^T P(l)C(l-1)\right]x(l-1) \\ &\qquad + 2x(l-1)^T \\ &\qquad\times\left[B(l-1)^T P(l)A(l-1)\right. \\ &\qquad\quad\left. + D(l-1)^T P(l)C(l-1)\right]u(l-1) \\ &\qquad + u(l-1)^T \\ &\qquad\times\left[R(l-1) + B(l-1)^T P(l)B(l-1)\right. \\ &\qquad\quad\left.\left. + D(l-1)^T P(l)D(l-1)\right]u(l-1)\right\}. \end{aligned} \tag{53}$$

Lemma 4.3 of [23] provides necessary and sufficient conditions for the finiteness of $V^{l-1}[x(l-1)]$:

$$\begin{aligned} P(l-1) &= A(l-1)^T P(l) A(l-1) + C(l-1)^T P(l) C(l-1) \\ &\quad + Q(l-1) - H(l-1)^T G(l-1)^{\dagger} H(l-1), \\ H(l-1) &= B(l-1)^T P(l) A(l-1) \\ &\quad + D(l-1)^T P(l) C(l-1), \\ G(l-1) &= R(l-1) + B(l-1)^T P(l) B(l-1) \\ &\quad + D(l-1)^T P(l) D(l-1) \geq \mathbf{0}, \\ G(l-1) & G(l-1)^{\dagger} H(l-1) - H(l-1) = 0. \end{aligned} \tag{54}$$

Moreover,

$$\begin{aligned} V^{l-1}[x(l-1)] &- E\left\{x(T)^T [Q(T) - P(T)] x(T)\right\} \\ &= x(l)^T P(l) x(l). \end{aligned} \tag{55}$$

The above proves the necessity part via mathematical induction.

Sufficiency. From the proof of Theorem 8, if GDRE (10) admits a solution $P(t)$ and λ, Problem 1 is not only well-posed, but also attainable. The proof of this theorem is complete. □

5. Other Equivalent Conditions

In this section, we present some other equivalent conditions for Problem 1.

Theorem 16. *For the constrained LQ Problem 1, the following are equivalent:*

(i) *Problem 1 is well-posed.*

(ii) *Problem 1 is attainable.*

(iii) *The LMIs* (41)-(42) *are feasible.*

(iv) *The GDRE* (10) *is solvable.*

Furthermore, when any one of the above conditions is satisfied, Problem 1 is attainable by

$$\begin{aligned} u(t) &= \left[R(t) + B(t)^T P(t+1) B(t) + D(t)^T P(t+1) D(t)\right]^{\dagger} \\ &\quad \cdot \left[B(t)^T P(t+1) A(t) + D(t)^T P(t+1) C(t)\right] x(t), \end{aligned} \tag{56}$$

where $P(t), t \in \{t_0, t_0+1, \ldots, T-1\}$, *are solutions to the GDRE* (10).

Proof. Applying Theorems 12–15, (i) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv). (ii) $\Rightarrow$ (iv) is shown by Theorem 8. The rest is to prove (iv) $\Rightarrow$ (ii) and (56). Let $P(t), t \in \{t_0, t_0+1, \ldots, T-1\}$, solve the GDRE (10). In view of

$$\begin{aligned} \sum_{t=t_0}^{T-1} & \left[x(t+1)^T P(t+1) x(t+1) - x(t)^T P(t) x(t)\right] \\ &= E\left[x(T)^T P(T) x(T) - x_0^T P(t_0) x_0\right], \end{aligned} \tag{57}$$

a completion of squares yields

$$\begin{aligned} J & (t_0, x_0; u(t_0), u(t_0+1), \ldots, u(T-1)) \\ &= E\left\{\sum_{t=t_0}^{T-1} \left[u(t) + G(t)^{\dagger} H(t) x(t)\right]^T \right. \\ &\qquad \left. \times G(t) \left[u(t) + G(t)^{-1} H(t) x(t)\right]\right\} \\ &\quad + E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\}, \end{aligned} \tag{58}$$

which shows

$$\begin{aligned} V(x_0) &= E\left\{x(T)^T [Q(T) - P(T)] x(T) + x_0^T P(t_0) x_0\right\}, \\ u^*(t) &= -G(t)^{\dagger} H(t) x(t), \quad t \in \{t_0, t_0+1, \ldots, T-1\}. \end{aligned} \tag{59}$$

Finally, we present a general expression for the optimal control set based on the solution to GDRE (10). □

Theorem 17. *Assume that the GDRE* (10) *admits a solution. Then the set of all optimal controls is determined by*

$$\begin{aligned} u(t)^{[Y(t), Z(t)]} &= -\left[G(t)^{\dagger} H(t) + Y(t) - G(t)^{\dagger} G(t) Y(t)\right] x(t) \\ &\quad + Z(t) - G(t)^{\dagger} G(t) Z(t), \end{aligned} \tag{60}$$

where $Y(t) \in R^{m\times n}$ *and* $Z(t) \in R^m$ *are arbitrary random variables defined on the probability space* $(\Omega, \mathcal{F}, \mathcal{P})$. *Moreover, the optimal cost value is uniquely given by*

$$V(x_0) = x_0^T P(t_0) x_0 - c\lambda, \tag{61}$$

where $P(t), t \in \{t_0, t_0+1, \ldots, T-1\}$, *and* λ *are the solution to the GDRE* (10).

Proof. This theorem can be proved by repeating the same procedure as in Theorem 5.1 of [23]. □

6. Conclusion

In this paper, we have investigated a class of indefinite stochastic LQ control problems with second moment constraints on the terminal state. By the matrix Lagrange theorem, we have established a new GDRE (10) associated with the constrained optimization Problem 1. In addition, by introducing LMIs (41)-(42), we show that the well-posedness and the attainability of Problem 1, the feasibility of the LMIs (41)-(42), and the solvability of GDRE (10) are equivalent to each other.

Conflict of Interests

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61174078), the Research Fund for the Taishan Scholar Project of Shandong Province of China and SDUST Research Fund (no. 2011KYTD105), and State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (Grant no. LAPS13018).

References

[1] R. E. Kalman, "Contributions to the theory of optimal control," *Boletín de la Sociedad Matemática Mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[2] W. M. Wonham, "On a matrix Riccati equation of stochastic control," *SIAM Journal on Control and Optimization*, vol. 6, no. 4, pp. 681–697, 1968.

[3] C. E. de Souza and M. D. Fragoso, "On the existence of maximal solution for generalized algebraic Riccati equations arising in stochastic control," *Systems and Control Letters*, vol. 14, no. 3, pp. 233–239, 1990.

[4] B. S. Chen and W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[5] W. Zhang and B. S. Chen, "On stabilizability and exact observability of stochastic systems with their applications," *Automatica*, vol. 40, no. 1, pp. 87–94, 2004.

[6] B. D. O. Anderson and J. B. Moore, *Optimal Control Linear Quadratic Mathods*, Prentice-Hall, New York, NY, USA, 1989.

[7] F. L. Lewis, *Optimal Control*, John Wiley and Sons, New York, NY, USA, 1986.

[8] Y. Huang, W. Zhang, and H. Zhang, "Infinite horizon linear quadratic optimal control for discrete-time stochastic systems," *Asian Journal of Control*, vol. 10, no. 5, pp. 608–615, 2008.

[9] S. Chen, X. J. Li, and X. Y. Zhou, "Stochastic linear quadratic regulators with indefinite control weight costs," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1685–1702, 1998.

[10] M. A. Rami and X. Y. Zhou, "Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1131–1143, 2000.

[11] D. D. Yao, S. Zhang, and X. Y. Zhou, "Stochastic linear-quadratic control via semidefinite programming," *SIAM Journal on Control and Optimization*, vol. 40, no. 3, pp. 801–823, 2001.

[12] M. Ait Rami, J. B. Moore, and X. Y. Zhou, "Indefinite stochastic linear quadratic control and generalized differential Riccati equation," *SIAM Journal on Control and Optimization*, vol. 40, no. 4, pp. 1296–1311, 2001.

[13] A. Ferrante, G. Marro, and L. Ntogramatzidis, "A parametrization of the solutions of the finite-horizon LQ problem with general cost and boundary conditions," *Automatica*, vol. 41, no. 8, pp. 1359–1366, 2005.

[14] D. Bertsimas and D. B. Brown, "Constrained stochastic LQC: a tractable approach," *IEEE Transactions on Automatic Control*, vol. 52, no. 10, pp. 1826–1841, 2007.

[15] Y. Hu and X. Y. Zhou, "Constrained stochastic LQ control with random coefficients, and application to portfolio selection," *SIAM Journal on Control and Optimization*, vol. 44, no. 2, pp. 444–466, 2005.

[16] S. Ko and R. R. Bitmead, "Optimal control for linear systems with state equality constraints," *Automatica*, vol. 43, no. 9, pp. 1573–1582, 2007.

[17] A. Kojima and M. Morari, "LQ control for constrained continuous-time systems," *Automatica*, vol. 40, no. 7, pp. 1143–1155, 2004.

[18] Z. Zhou and C. Randy, "An algorithm for state constrained stochastic linear-quadratic control," in *Proceedings of the American Control Conference*, pp. 1476–1481, San Francisco, Calif, USA, June 2011.

[19] Y. L. Huang and W. H. Zhang, "Study on stochastic linear quadratic optimal control with constraint," *Acta Automatica Sinica*, vol. 32, no. 2, pp. 246–254, 2006 (Chinese).

[20] G. Li and W. Zhang, "Discrete-time indefinite stochastic linear quadratic optimal control with equality constraints," in *Proceeding of the Chinese Control and Decision Conference*, pp. 4999–5004, Guiyang, China, May 2013.

[21] H. Dong, Z. Wang, D. W. C. Ho, and H. Gao, "Variance-constrained H_∞ filtering for a class of nonlinear time-varying systems with multiple missing measurements: the finite-horizon case," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2534–2543, 2010.

[22] L. Ma, Y. Bo, Y. Zhou, and Z. Guo, "Error variance-constrained H_∞ filtering for a class of nonlinear stochastic systems with degraded measurements: the finite horizon case," *International Journal of Systems Science*, vol. 43, no. 12, pp. 2361–2372, 2012.

[23] M. Ait Rami, X. Chen, and X. Y. Zhou, "Discrete-time indefinite LQ control with state and control dependent noises," *Journal of Global Optimization*, vol. 23, no. 3-4, pp. 245–265, 2002.

[24] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Athena Scientific, Belmont, Mass, USA, 1996.

[25] A. Albert, "Conditions for positive and nonnegative definiteness in terms of pseudoinverses," *SIAM Journal on Applied Mathematics*, vol. 17, no. 2, pp. 434–440, 1969.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 835254, 8 pages
http://dx.doi.org/10.1155/2014/835254

Research Article

Research on Multiprincipals Selecting Effective Agency Mode in the Student Loan System

Libo Ding,[1,2] Bangyi Li,[1] and Suling Feng[2]

[1] *College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*
[2] *School of Economics, University of Jinan, Jinan 250022, China*

Correspondence should be addressed to Libo Ding; uao_dinglb@ujn.edu.cn

Received 19 February 2014; Accepted 9 April 2014; Published 8 May 2014

Academic Editor: Ming Gao

An effective agency mode is the key to solve incentive problems in Chinese student loan system. Principal-agent frameworks are considered in which two principals share one common agent that is performing one single task but each prefers the different aspect of the task. Three models are built and decision mechanisms are given. The studies show that the three modes have different effects. Exclusive dealing mode is not good for long-term effect because sometimes it guides agent ignoring repayment. If effort proportionality coefficient and observability are both unchanged, principals all prefer common agency, but independent contracting mode may be more efficient in reality because not only the total outputs under that mode are larger than those under cooperation one, but also preferring independent contracting mode can stimulate the bank participating in the game.

1. Introduction

An effective management structure is a necessary condition for the student loans operation. Different countries have different structures of loan management system, such as bank, state agency, and other types of organization. In China, the student loans are operated by the most basic level agency called county-student financial assistance center which is regulated by government, and the funds are provided by policy bank. In this structure, one agent faces two or more principals; namely, various principals share one common agent. In these situations, conflicts typically arise among principals when the agent uses its time and effort to different principals; moreover, the agent's moral hazard and adverse selection problems can make the conflicts complicated. Usually, incentives must be provided to induce optimal performance when the agent's effort or his ability is unobservable, but the incentives provided by different principals could affect each other, which can decide how to use its time and effort alternatively by the agent.

Traditional principal-agent theory has offered many techniques dealing with optimal performance in principal-agent problems; some new techniques dealing with optimization problems in ambiguity environment are discussed by a study group [1–4], and backward stochastic differential equations are used in their important works in this field to deal with more complex problems [5–8]. In our study, multiprincipals sharing one agent which was called common agent and how to select an optimal agent mode are the core. Different principals sharing a common agent were first developed in the seminal paper of Bernheim and Whinston [9, 10]. In their studies, different principals simultaneously and independently influence a common agent. While complete and incomplete information were both contained in the studies, they show that implementation is always efficient and that noncooperative behavior induces an efficient action choice if and only if collusion among the principals would implement the first-best action at the first-best level of cost. They also investigate the existence of equilibria, the distribution of net rewards among principals, the characteristics of actions chosen in inefficient equilibria, and potential institutional remedies for welfare losses induced in noncooperative behavior. Subsequently, the studies about common agency are blooming so that more and more scholars focus on incentives in common agency, among which Martimort's series of work [11–17] forms a study framework of multiprincipals;

others also contribute to characteristics of common agency game equilibrium [18–22]. Some researchers are interested in designing incentive mechanism [14, 16, 23] and pay attention to agent facing multitasks [24, 25], and others pay more attention to the cooperation and competition among multiprincipals [24, 26, 27]. In the field of application, in addition to common sales agency problems, financial and insurance market, tax competition, and auction, researchers focus on multiprincipals problems of regulation or organizational design [12, 13, 28–31].

In Chinese current student loan system, the government is not only the regulator, but also a principal, who designs the management structure and selects the bank which takes part in the student loan system. So in this current paper, we consider a principal-agent framework, in which the model has multiple principals (basic level government and policy bank) and one single agent (staff of county-student financial assistance center) performing one single task, but the two principals have different preference in the same task's different aspects. From the government's point of view, the objective of the loan policy is to achieve the maximum of social welfare. The government hopes that students, as many as possible, from families with financial difficulties could be able to obtain loans to solve education problem. In the long run, the government's concern should not be the number of students who obtain loans but the students' repayment in order to facilitate the repeated game and obtain the long-term cooperation with banks. But in reality, the government often pays more attention to the short-term effect, which is manifested as its excessive emphasis on the quantity of students accepted by the agent, but does not pay enough attention to the effort of agent urging borrowers' repayment. In other words, the government prefers the agent paying more effort to handling more loan contracts. On the contrary, from the bank's point of view, more loan agreements often mean more benefits along with more risks; the bank pays more attention to the repayment of those students, so the agent's effort to urge the borrowers to repay on time is the key to the bank. In other words, the bank's preference is the effort to urge students to repay the loans. Resolving the conflicts of different preferences of principals is the key to guarantee the effective implementation of the student loans policy. In our hypothesis, the task's two aspects are regarded as two alternative tasks because the agent must reduce the effort and time in one aspect when he wants to take another aspect seriously. So incentive contracts offered by principals are the key to solve the conflict and meanwhile maximize their own profit.

In Mezzetti's model [27], the single agent performs related tasks for different principals who are horizontally differentiated and each principal requires that a task should be performed. The equilibrium under cooperation between two principals, exclusive dealing, and independent contracting are discussed in Mezzetti's article. Firstly, the principals offer the common agent an incentive contract that maximizes their joint payoff under cooperation. Secondly, each principal chooses an incentive contract noncooperatively and cannot contract on the agent's output for the other principal under independent contracting. Thirdly, each principal makes contract with a different, but ex ante identical, agent under exclusive dealing. In our paper, ideas are borrowed from Mezzetti [27] to discuss the incentive contracts offered by government and bank (principals) to county-student financial assistance center staff (agent) and help the principals having different preferences select effective agency mode. In any kind of agency mode, the agent will select the optimal effort level to maximize his expected utility when his effort or ability is unobservable.

2. Major Assumptions and Variable Declaration

(1) Two principals i ($i = 1, 2$; 1 is the government; 2 is the bank) contract with a common agent (county-student financial assistance center staff) to perform the student-loan-management task. Government prefers the agent paying more effort to handling more loan contracts; the bank prefers more the agent's efforts to urge students' repayment. The principals are all risk neutral whose expected utility is equal to their expected return. The agent is risk averse: his utility function has the characteristics of constantly absolute risk aversion and $\rho = -u''/u' > 0$ is the parameter of risk aversion degree.

(2) Principal's utility function is v_i; the agent's corresponding utility function is u_1, u_2; reservation wage $\overline{\omega}_i > 0, \overline{\omega}_2 > \overline{\omega}_1 > 0$, means the agent's opportunity income obtaining from the bank is higher than that from the government.

(3) The effort level a_1, a_2, agent working for different principals' preference, is unobservable. Let $k_i > 0$ be the proportionality coefficient between agent's effort and his output for two principals. θ_i is private information of the agent and as a random variable, normally distributed in $[0, \sigma_i^2]$: variance $\sigma_1^2 < \sigma_2^2$ means the bank's preference is more difficult than government's preference to be completed. Thus, the agent's output on principals' task is $\pi_i = k_i a_i + \theta_i$.

(4) Let α_i and β_i be the flat fee and the incentive coefficient, respectively, that each principal pays to the agent. The principals offer incentives contracts to the agent, and the agent's payoff is

$$s(\pi_i) = \alpha_i + \beta_i \pi_i, \quad (0 \le \beta_i \le 1). \tag{1}$$

(5) The common agent's effort cost in different tasks is alternative; let t be the alternative coefficient; $t = 1$ means the maximum alternative. The cost function is $C(a_1, a_2) = a_1^2/2 + a_2^2/2 - t a_1 a_2, 0 \le t \le 1$. The cost function in exclusive dealing mode is $C(a_i) = a_i^2/2$.

3. Exclusive Dealing Mode

Under exclusive dealing mode, the optimal incentive contracts offered by two principals exclusively are similar to different principals selecting different agents and offering his agent exclusive contract, which is a model containing the single principal and single agent. Thus, the agent's real income is

$$\omega_i = s(\pi_i) - c(a_i) = \alpha_i + \beta_i k_i (a_i + \theta) - \frac{a_i^2}{2}, \quad (2)$$

and the agent's certainty equivalence wealth (CEW) is

$$\overline{\omega}_i = E\omega_i - \frac{\rho\beta_i^2\sigma_i^2}{2} = \alpha_i + \beta_i k_i a_i - \frac{a_i^2}{2} - \frac{\rho\beta_i^2\sigma_i^2}{2}. \quad (3)$$

Under exclusive dealing incentive contract, each agent, using his reservation wage as a benchmark, performs his task maximizing his own certainty equivalence wealth.

The agent's incentive compatibility constraint (IC) is

$$\max_{a_i}\left(\alpha_i + \beta_i k_i a_i - \frac{a_i^2}{2} - \frac{\rho\beta_i^2\sigma_i^2}{2}\right), \quad (4)$$

and the agent's individual rationality constraint (IR) is

$$\alpha_i + \beta_i k_i a_i - \frac{a_i^2}{2} - \frac{\rho\beta_i^2\sigma_i^2}{2} \geq \overline{\omega}_i. \quad (5)$$

Each risk-neutral principal's expected utility, equal to his expected return, is

$$\begin{aligned} Ev_i[\pi_i - s(\pi_i)] &= v_i\{E[\pi_i - s(\pi_i)]\} \\ &= v_i[-\alpha_i + (1-\beta_i)k_i a_i] \quad (6) \\ &= -\alpha_i + (1-\beta_i)k_i a_i, \end{aligned}$$

and each principal will select the optimal incentive scheme (α_i, β_i), to maximize his own expected income.

The model is

$$\begin{aligned} &\max_{\alpha_i,\beta_i} \quad [-\alpha_i + (1-\beta_i)k_i a_i] \\ &\text{s.t.} \quad (\text{IR})\, \alpha_i + \beta_i k_i a_i - \frac{a_i^2}{2} - \frac{\rho\beta_i^2\sigma_i^2}{2} \geq \overline{\omega}_i \\ &\qquad (\text{IC})\, \max_{a_i} CE_A = \overline{\omega}_i = \left(\alpha_i + \beta_i k_i a_i - \frac{a_i^2}{2} - \frac{\rho\beta_i^2\sigma_i^2}{2}\right). \end{aligned} \quad (7)$$

Under each optimal incentive scheme (α_i, β_i), the agent's IC should ensure maximizing his CEW, $\overline{\omega}_i$, and the first-order condition is

$$\frac{\partial\overline{\omega}_i}{\partial a_i} = \beta_i k_i - a_i = 0, \quad \text{thus,} \;\; a_i = \beta_i k_i. \quad (8)$$

We denote by $\beta_E^*, a_{Ei}^*, \alpha_E^*$ (subscript E on behalf of the exclusive dealing situation) the second-best solution when feeding IC, IR, and formula (8) to objective function. The second-best solution is

$$\begin{aligned} &\beta_E^* = \frac{k_i^2}{k_i^2 + \rho\sigma_i^2}, \qquad a_{Ei}^* = \frac{k_i^3}{k_i^2 + \rho\sigma_i^2}, \\ &\alpha_E^* = \overline{\omega}_i + \frac{\rho\sigma_i^2 k_i^4 + k^6 - 2k^9}{2(k_i^2 + \rho\sigma_i^2)^2}. \end{aligned} \quad (9)$$

Proposition 1. *Under exclusive dealing mode, the decision mechanism of principal is to determine the second-optimal incentive coefficient which satisfies the following:*

$$\beta_{Ei}^* = \frac{k_i^2}{k_i^2 + \rho\sigma_i^2}. \quad (10)$$

In order to obtain the agent's optimal response

$$a_{Ei}^* = \frac{k_i^3}{k_i^2 + \rho\sigma_i^2}. \quad (11)$$

The incentive coefficient was determined by the agent's risk aversion degree, variances, and proportionality coefficient.

4. Independent Contracting Mode

Under independent contracting mode, each principal designs incentive contract to common agent noncooperatively meanwhile maximizing his own profit:

$$s(\pi_i) = \alpha_i + \beta_i \pi_i, \quad (0 \leq \beta_i \leq 1), \quad (12)$$

and the agent's effort costs in two principals' preference are correlative. In two principals' separate incentive mechanism, the agent's response selects the optimal effort level to adapt to the incentive contracts; meanwhile its IC should ensure that its separate real income is not less than the separate $\overline{\omega}_i$, and the IR should ensure maximizing agent's own total CEW:

$$CE_A = \alpha_1 + \alpha_2 + \beta_1 k_1 a_1 + \beta_2 k_2 a_2 - \frac{\rho\beta^T\Sigma\beta}{2} - C(a_1, a_2). \quad (13)$$

Principals will determine their separate optimal incentive scheme (α_i, β_i), and their maximization problems can be written as follows:

$$\begin{aligned} &\max_{\alpha_i,\beta_i} \quad [-\alpha_i + (1-\beta_i)k_i a_i] \\ &\text{s.t.} \quad (\text{IR})\, \alpha_i + \beta_i k_i a_i - \frac{\rho\beta_i^2\sigma_i^2}{2} - C(a_1, a_2) \geq \overline{\omega}_i \\ &\qquad (\text{IC})\, \max_{a_1,a_2} CE_A = \alpha_1 + \alpha_2 + \beta_1 k_1 a_1 + \beta_2 k_2 a_2 \\ &\qquad\qquad - \frac{\rho\beta^T\Sigma\beta}{2} - C(a_1, a_2) \end{aligned} \quad (14)$$

and the results of calculating the partial derivative of CEW about a_1, a_2 are

$$\begin{aligned} \frac{\partial CE_A}{\partial a_1} &= \beta_1 k_1 - a_1 + t a_2 \\ \frac{\partial CE_A}{\partial a_2} &= \beta_2 k_2 - a_2 + t a_1, \\ a_1 &= \frac{\beta_1 k_1 + t\beta_2 k_2}{1 - t^2} \\ a_2 &= \frac{\beta_2 k_2 + t\beta_1 k_1}{1 - t^2}. \end{aligned} \quad (15)$$

Feed a_1, a_2 into IC, and then get results as follows:

$$\alpha_1 = \overline{\omega}_1 - \beta_1 k_1 a_1 + \frac{\rho\beta_1^2\sigma_1^2}{2} + \frac{a_1^2}{2} + \frac{a_2^2}{2} - t a_1 a_2,$$
$$\alpha_2 = \overline{\omega}_2 - \beta_2 k_2 a_2 + \frac{\rho\beta_2^2\sigma_2^2}{2} + \frac{a_1^2}{2} + \frac{a_2^2}{2} - t a_1 a_2. \quad (16)$$

Feed α_i into two principals' separate objective function (subscript I on behalf of the independent contracting situation):

$$\max v_{I1} = -\overline{\omega}_1 - \frac{\rho\beta_1^2\sigma_1^2}{2} - \frac{a_1^2}{2} - \frac{a_2^2}{2} + t a_1 a_2 + k_1 a_1$$
$$= -\overline{\omega}_1 - \frac{\rho\beta_1^2\sigma_1^2}{2} - \left(\left(\beta_1^2 k_1 + \beta_2^2 k_2^2\right) + 2t\left(\beta_1 k_1 + \beta_2 k_2\right) - 2k_1\left(\beta_1 k_1 + t\beta_2 k_2\right)\right)\left(2\left(1-t^2\right)\right)^{-1},$$
$$\max v_{I2} = -\overline{\omega}_2 - \frac{\rho\beta_2^2\sigma_2^2}{2} - \frac{a_1^2}{2} - \frac{a_2^2}{2} + t a_1 a_2 + k_2 a_2$$
$$= -\overline{\omega}_2 - \frac{\rho\beta_2^2\sigma_2^2}{2} - \left(\left(\beta_1^2 k_1 + \beta_2^2 k_2^2\right) + 2t\left(\beta_1 k_1 + \beta_2 k_2\right) - 2k_2\left(\beta_2 k_2 + t\beta_1 k_1\right)\right)\left(2\left(1-t^2\right)\right)^{-1}. \quad (17)$$

Calculate the partial derivative of the previous two formulas about β_1, β_2. We have

$$\beta_1 = \frac{k_1^2 - t k_1 k_2 \beta_2}{\rho\sigma_1^2\left(1-t^2\right) + k_1^2},$$
$$\beta_2 = \frac{k_2^2 - t k_1 k_2 \beta_1}{\rho\sigma_2^2\left(1-t^2\right) + k_2^2}. \quad (18)$$

We denote by $\beta_{I1}^*, \beta_{I2}^*$ the second-best solutions under independent contracting mode of simultaneous equations (15) and (18). Thus,

$$\beta_{I1}^* = \frac{\rho\sigma_2^2\left(1-t^2\right)k_1^2 + k_1^2 k_2^2 - t k_1 k_2^3}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2},$$
$$\beta_{I2}^* = \frac{\rho\sigma_1^2\left(1-t^2\right)k_2^2 + k_1^2 k_2^2 - t k_2 k_1^3}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2},$$
$$a_{I1}^* = \frac{\rho\left[\sigma_2^2 k_1^3 + t\sigma_1^2 k_2^3\right] + k_1^3 k_2^2}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2},$$
$$a_{I2}^* = \frac{\rho\left[\sigma_1^2 k_2^3 + t\sigma_2^2 k_1^3\right] + k_2^3 k_1^2}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2}. \quad (19)$$

Proposition 2. *Under independent contracting mode, the different incentive coefficients given by different principal are as follows:*

$$\beta_{I1}^* = \frac{\rho\sigma_2^2\left(1-t^2\right)k_1^2 + k_1^2 k_2^2 - t k_1 k_2^3}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2},$$
$$\beta_{I2}^* = \frac{\rho\sigma_1^2\left(1-t^2\right)k_2^2 + k_1^2 k_2^2 - t k_2 k_1^3}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2}, \quad (20)$$

which are determined jointly by the agent's risk aversion degree, variances, alternative coefficient, and proportionality coefficient. The best corresponding responses of agent are

$$a_{I1}^* = \frac{\rho\left[\sigma_2^2 k_1^3 + t\sigma_1^2 k_2^3\right] + k_1^3 k_2^2}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2},$$
$$a_{I2}^* = \frac{\rho\left[\sigma_1^2 k_2^3 + t\sigma_2^2 k_1^3\right] + k_2^3 k_1^2}{\left[\rho\sigma_1^2\left(1-t^2\right)+k_1^2\right]\left[\rho\sigma_2^2\left(1-t^2\right)+k_2^2\right] - t^2 k_1^2 k_2^2}. \quad (21)$$

5. Cooperation between Principals Mode

Under cooperation mode, two principals offer common incentive contract (α, β) to common agent in order to maximize their joint profit:

$$s\left(\pi_1, \pi_2\right) = \alpha + \beta\left(\pi_1 + \pi_2\right), \quad (0 \le \beta \le 1), \quad (22)$$

and the total expected return of two principals is

$$E\left(v_1 + v_2\right) = Ev\left[\pi_1 + \pi_2 - s\left(\pi_1 + \pi_2\right)\right]$$
$$= v\left\{E\left[\pi_1 + \pi_2 - s\left(\pi_1 + \pi_2\right)\right]\right\}$$
$$= v\left[-\alpha + (1-\beta)\left(k_1 a_1 + k_2 a_2\right)\right]$$
$$= -\alpha + (1-\beta)\left(k_1 a_1 + k_2 a_2\right). \quad (23)$$

Under cooperation, we consider that (α, β) must satisfy IC with the sum of reservation wages of two principals' separate contract in order for incentive agent to perform the tasks, and IR is to maximize agent's CEW:

$$CE_A = \alpha + \beta\left(k_1 a_1 + k_2 a_2\right) - \frac{\rho\beta^2\left(\sigma_1^2 + \sigma_2^2\right)}{2} - C\left(a_1, a_2\right). \quad (24)$$

We can write principals' maximization problem as follows:

$$\max_{\alpha,\beta} \quad \left[-\alpha + (1-\beta)\left(k_1 a_1 + k_2 a_2\right)\right]$$
$$\text{s.t.} \quad (\text{IR})\ \alpha + \beta k_1 a_1 + \beta k_2 a_2 - \frac{\rho\beta^2\left(\sigma_1^2 + \sigma_2^2\right)}{2} - C\left(a_1, a_2\right) \ge \overline{\omega}_1 + \overline{\omega}_2$$
$$(\text{IC})\ \max_{a_1,a_2} CE_A = \alpha + \beta k_1 a_1 + \beta k_2 a_2 - \frac{\rho\beta^2\left(\sigma_1^2 + \sigma_2^2\right)}{2} - C\left(a_1, a_2\right). \quad (25)$$

The calculation process and results are as follows:

$$\frac{\partial CE_A}{\partial a_1} = k_1\beta - a_1 + ta_2$$
$$\frac{\partial CE_A}{\partial a_2} = k_2\beta - a_2 + ta_1, \quad (26)$$
$$a_1 = \frac{(k_1 + tk_2)\beta}{1-t^2} \qquad a_2 = \frac{(k_2 + tk_1)\beta}{1-t^2}.$$

Feed a_1, a_2 into IC separately. Then,

$$\alpha = \overline{\omega}_1 + \overline{\omega}_2 - \left(\frac{2k_1^2 + 3tk_1k_2 + 2k_2^2}{2(1-t^2)} - \frac{\rho(\sigma_1^2 + \sigma_2^2)}{2}\right)\beta^2. \quad (27)$$

Feed α into principals' joint objective function (subscript C on behalf of the cooperation situation):

$$\max_{\alpha,\beta} v_C = -(\overline{\omega}_1 + \overline{\omega}_2) + \left(\frac{(k_1^2 + tk_1k_2 + k_2^2)}{2(1-t^2)} - \frac{\rho(\sigma_1^2 + \sigma_2^2)}{2}\right)\beta^2 + \frac{(k_1^2 + k_2^2 + 2tk_1k_2)}{1-t^2}\beta, \quad (28)$$

and the first-order condition is

$$\left(\frac{k_1^2 + tk_1k_2 + k_2^2}{1-t^2} - \rho(\sigma_1^2 + \sigma_2^2)\right)\beta + \frac{k_1^2 + k_2^2 + 2tk_1k_2}{1-t^2} = 0. \quad (29)$$

We denote by β_C^* the second-best solutions under cooperation; the results are

$$\beta_C^* = \frac{k_1^2 + k_2^2 + 2tk_1k_2}{(k_1^2 + tk_1k_2 + k_2^2) - \rho(\sigma_1^2 + \sigma_2^2)(1-t^2)},$$
$$a_{C1}^* = \frac{(k_1 + tk_2)(k_1^2 + k_2^2 + 2tk_1k_2)}{(1-t^2)[\rho(\sigma_1^2 + \sigma_2^2)(1-t^2) - (k_1^2 + tk_1k_2 + k_2^2)]}$$
$$a_{C2}^* = \frac{(k_2 + tk_1)(k_1^2 + k_2^2 + 2tk_1k_2)}{(1-t^2)[\rho(\sigma_1^2 + \sigma_2^2)(1-t^2) - (k_1^2 + tk_1k_2 + k_2^2)]}. \quad (30)$$

Proposition 3. *Under cooperation contracting mode, the joint decision mechanism of two principals is*

$$\beta_C^* = \frac{k_1^2 + k_2^2 + 2tk_1k_2}{(k_1^2 + tk_1k_2 + k_2^2) - \rho(\sigma_1^2 + \sigma_2^2)(1-t^2)}. \quad (31)$$

The best effort responses of common agent to different tasks are

$$a_{C1}^* = \frac{(k_1 + tk_2)(k_1^2 + k_2^2 + 2tk_1k_2)}{(1-t^2)[\rho(\sigma_1^2 + \sigma_2^2)(1-t^2) - (k_1^2 + tk_1k_2 + k_2^2)]},$$
$$a_{C2}^* = \frac{(k_2 + tk_1)(k_1^2 + k_2^2 + 2tk_1k_2)}{(1-t^2)[\rho(\sigma_1^2 + \sigma_2^2)(1-t^2) - (k_1^2 + tk_1k_2 + k_2^2)]}. \quad (32)$$

6. Numerical Analysis and Discussions

Numerical analysis is discussed in this section in order to illustrate the decision mechanism of both sides and compare the incentive efficient further in different modes.

Firstly, parameters are set according to their ranges in the models' assumption as follows:

$$\sigma_1^2 = 0.1, \qquad \sigma_2^2 = 1, \qquad \rho = 0.005,$$
$$k_1 = 1, \qquad k_2 3, \qquad \overline{\omega}_1 = 1, \qquad \overline{\omega}_2 = 2. \quad (33)$$

The results of β, a, π in three modes are compared when $t = 0.1, 0.3, 0.5, 0.8$; the influence of alternative coefficient on principals and the agent's decision mechanism are illustrated in Table 1.

Under the condition of unchangeable alternative coefficient the following can be drawn from Table 1.

(1) Under exclusive dealing mode, $\beta_{E1}^* > \beta_{E2}^*$; namely, the principals offer greater incentive on the easy supervision task. Under independent contracting mode, $\beta_{I2}^* > \beta_{I1}^*$; namely, the principal offers greater incentive on the difficult supervision task.

(2) Both principals prefer to select common agency who only considers the influence of alternative coefficient and $\beta_C^* > \beta_{I1}^*$ and $\beta_C^* > \beta_{E1}^*$ mean principal with easy supervision task prefers to select cooperation mode, but principal with difficult supervision task will select cooperation mode when alternative coefficient (t) is small; otherwise independent mode will be selected when t gradually becomes larger and $\pi_{I(\mathrm{sum})}^* > \pi_{C(\mathrm{sum})}^*$ and $\pi_{I(\mathrm{sum})}^* > \pi_{E(\mathrm{sum})}^*$ mean that total outputs under independent mode are always larger than those under the other two modes. The changing of alternative coefficient (t) will not influence the incentive under exclusive dealing mode, but it can influence that in common agency. That is to say, under cooperation mode, the incentive will change in the same direction with alternative coefficient. And under independent contracting mode, it will change still in the same direction on the difficult supervision task but change inversely on the easy one.

(3) Consider that $a_{E2}^* > a_{E1}^*$, $a_{I2}^* > a_{I1}^*$, and $a_{C2}^* > a_{C1}^*$ mean that agent makes more efforts on the difficult supervision task under any agency mode because of the principal's different incentives in different mode. When other conditions remain unchanged, the effort becomes greater, while the alternative coefficient gets larger. When other conditions remain unchanged, the efforts on two tasks both become greater gradually with the difficult supervision task's variance getting larger under cooperation mode. On the contrary, the effort becomes smaller under independent mode.

Secondly, parameters are set according to its range in models assumption as follows:

$$\overline{\omega}_1 = 1, \qquad \overline{\omega}_2 = 2, \qquad \rho = 0.005,$$
$$t = 0.3, \qquad k_1 = 1, \qquad k_2 = 3. \quad (34)$$

Table 1: Different outputs under three modes when alternative coefficient (t) changes.

Common agency										Exclusive dealing	
Independent contracting mode					Cooperation mode						
t	0.1	0.3	0.5	0.8	t	0.1	0.3	0.5	0.8		
β_{I1}^*	0.7072	0.1103	−0.6658	−3.8863	β_C^*	1.0296	1.0830	1.1308	1.1937	β_{E1}^*	1.0000
β_{I2}^*	0.9759	0.9885	1.1105	2.0358						β_{E2}^*	0.9994
a_{I1}^*	1.2345	2.0407	3.9995	24.9909	a_{C1}^*	1.3520	2.2613	3.7694	11.2741	a_{E1}^*	1.0000
a_{I2}^*	3.7017	6.1193	11.9940	74.9600	a_{C2}^*	3.2241	3.9275	5.2771	12.6004	a_{E2}^*	2.9983
π_{I1}^*	1.2345	2.0407	3.9995	24.9909	π_C^*	4.5761	6.1887	9.0465	23.8745	π_{E1}^*	1.0000
π_{I2}^*	11.1051	18.3580	35.9821	224.8799						π_{E2}^*	8.9950
$\pi_{I(\text{sum})}^*$	12.3396	20.3987	39.9821	249.8708	$\pi_{C(\text{sum})}^*$	4.5761	6.1887	9.0465	23.8745	$\pi_{E(\text{sum})}^*$	9.9950

Table 2: Different outputs under three modes when one variance (σ_2^2) changes.

Common agency								Exclusive dealing			
Independent contracting mode				Cooperation mode							
σ_2^2	100	36	9	σ_2^2	100	36	9	σ_2^2	100	36	9
β_{I1}^*	0.1490	0.1231	0.1116	β_C^*	1.1317	1.1010	1.0885	β_{E1}^*	0.9804	0.9804	0.9804
β_{I2}^*	0.9356	0.9681	0.9826					β_{E2}^*	0.9474	0.9804	0.9950
a_{I1}^*	2.0226	2.0295	2.0325	a_{C1}^*	2.3629	2.2987	2.2726	a_{E1}^*	0.9804	0.9804	0.9804
a_{I2}^*	5.8193	6.0027	6.0840	a_{C2}^*	4.1040	3.9925	3.9472	a_{E2}^*	2.8421	2.9412	2.9851
π_{I1}^*	2.0226	2.0295	2.0325	π_C^*	6.4669	6.2912	6.2199	π_{E1}^*	0.9804	0.9804	0.9804
π_{I2}^*	17.4578	18.0082	18.2520					π_{E2}^*	8.5263	8.8235	8.9552
$\pi_{I(\text{sum})}^*$	19.4804	20.0377	20.2845	$\pi_{C(\text{sum})}^*$	6.4669	6.2912	6.2199	$\pi_{E(\text{sum})}^*$	9.5067	9.8039	9.9366

The results of β, a, π in three modes are compared when

$$\begin{gathered} \sigma_1^2 = 4, \qquad \sigma_2^2 = 100, \qquad k_1^2\sigma_2^2 > k_2^2\sigma_1^2, \\ \sigma_1^2 = 4, \qquad \sigma_2^2 = 36, \qquad k_1^2\sigma_2^2 = k_2^2\sigma_1^2, \\ \sigma_1^2 = 4, \qquad \sigma_2^2 = 9, \qquad k_1^2\sigma_2^2 < k_2^2\sigma_1^2. \end{gathered} \tag{35}$$

The influence of variance on principals and the agent's decision mechanism are illustrated in Table 2.

According to Table 2, if the influence of task's variance was considered merely it can be obtained as follows.

$\beta_C^* > \beta_{Ii}^*$ and $\beta_C^* > \beta_{Ei}^*$ mean that principals always prefer cooperation mode. When other conditions remain unchanged, with the difficult supervision task's variance getting larger, the incentive offered by the principal whose task is difficult to be supervised becomes smaller gradually under exclusive dealing mode. The incentive becomes greater on the easy supervision task, but it becomes smaller on the difficult supervision task under independent contracting mode, while the incentive becomes greater gradually under cooperation mode.

Through the above analysis, implications and suggestions on how to select the effective agency mode can be got as follows.

(1) Because the government's ultimate goal is to realize the maximum social welfare, it should think highly of urging borrowers' repayment rather than merely consider the quantity of loan contracts just like what they do in reality. Because the principal whose task is easy to be supervised prefers to offer more incentives under exclusive dealing mode, selecting exclusive dealing mode will lead the staff to pay more attention to sign more loan contracts but ignore to urge repayment, which is not good for the long-term effect of national student financial aid policy.

(2) If the effort proportionality coefficient and variance are both unchanged, both principals prefer to select common agency, but each principal's preference degree of selecting cooperation or independent mode is different according to the difficulty degree of the task. We consider that the government prefers cooperation mode, although under it the total output is less than that under independent mode. In order to stimulate the bank participating in the policy, the government should select the mode that the bank prefers.

(3) The study shows that although principals offer different incentives in different modes, the agent always offers more effort to the difficult supervision task under any mode, which not only gives enlightenment that the student loans repayment is the key in financial aid policy, but also warns that incentive mechanism designing absolutely according to the study results may lead us to ignore the quantity of student loans which is the base to realize national policy objective. So in the practical mechanism designing, the government that is not just a principal but more importantly a regulator should comprehensively consider more

affecting factors such as total output, bank and staff's enthusiasm, and the continuity of policy.

7. Conclusions

The research on multiprincipals and how to select effective agency mode in the student loan system has been carried out. Three models of cooperation between principals, exclusive dealing, and independent contracting have been investigated and discussed. Decision mechanisms are given and efficiencies among three modes are contrasted by numerical analysis. Under the condition of unchangeable alternative coefficient three main conclusions were obtained and discussed. Under exclusive dealing mode and independent contracting mode the principals offer greater incentive on the easy supervision task and difficult supervision task, respectively. And both principals prefer to select common agency who only considers the influence of alternative coefficient. Considering the influence of task's variance principals always prefer cooperation mode. The studies show that exclusive dealing mode is not good for student financial aid policy's long-term effect because it sometimes guides agent ignoring repayment; if effort proportionality coefficient and observability are both unchanged, both principals prefer common agency, but independent contracting mode may be more efficient in reality because not only the total outputs under it are larger than those under cooperation mode, but also preferring independent contracting mode could stimulate the bank participating in the game; the conclusion that agent always offers more efforts to the difficult supervision task under any mode warns that incentive mechanisms designing absolutely according to the study results may lead us to ignore loans quantity, so the government, which is not just a principal but more importantly a regulator, should consider comprehensively more affecting factors in practice.

Conflict of Interests

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

Acknowledgment

The research was supported by Scientific Research Projects of Education Department of Shandong Province (no. J13WH119).

References

[1] L. G. Epstein and S. L. Ji, "Ambiguous volatility and asset pricing in continuous time," *The Review of Financial Studies*, vol. 26, no. 7, pp. 1740–1786, 2013.

[2] S. Ji and Q. Wei, "A maximum principle for fully coupled forward-backward stochastic control systems with terminal state constraints," *Journal of Mathematical Analysis and Applications*, vol. 407, no. 2, pp. 200–210, 2013.

[3] S. Ji and X. Y. Zhou, "A maximum principle for stochastic optimal control with terminal state constraints, and its applications," *Communications in Information and Systems*, vol. 6, no. 4, pp. 321–337, 2006.

[4] L. G. Epstein and S. Ji, "Ambiguous volatility, possibility and utility in continuous time," *Journal of Mathematical Economics*, vol. 50, pp. 269–282, 2014.

[5] M. Hu, S. Ji, S. Peng, and Y. Song, "Backward stochastic differential equations driven by G-Brownian motion," *Stochastic Processes and Their Applications*, vol. 124, no. 1, pp. 759–784, 2014.

[6] S. Ji and X. Y. Zhou, "A generalized Neyman-Pearson lemma for g-probabilities," *Probability Theory and Related Fields*, vol. 148, no. 3-4, pp. 645–669, 2010.

[7] M. Hu, S. Ji, S. Peng, and Y. Song, "Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion," *Stochastic Processes and Their Applications*, vol. 124, no. 2, pp. 1170–1195, 2014.

[8] S. Ji and X. Y. Zhou, "The Neyman-Pearson lemma under G-probability," *Comptes Rendus Mathématique. Académie des Sciences. Paris*, vol. 346, no. 3-4, pp. 209–212, 2008.

[9] B. D. Bernheim and M. D. Whinston, "Common marketing agency as a device for facilitating collusion," *The RAND Journal of Economics*, vol. 16, no. 2, pp. 269–281, 1985.

[10] B. D. Bernheim and M. D. Whinston, "Common agency," *Econometrica*, vol. 54, no. 4, pp. 923–942, 1986.

[11] D. Martimort, "Exclusive dealing, common agency, and multiprincipals incentive theory," *The RAND Journal of Economics*, vol. 27, no. 1, pp. 1–31, 1996.

[12] D. Martimort, "The multiprincipal nature of government," *European Economic Review*, vol. 40, no. 3–5, pp. 673–685, 1996.

[13] D. Martimort, "Renegotiation design with multiple regulators," *Journal of Economic Theory*, vol. 88, no. 2, pp. 261–293, 1999.

[14] D. Martimort and L. Stole, "Common agency equilibria with discrete mechanisms and discrete types," Cesifo Working Paper 572, 2001.

[15] D. Martimort and L. Stole, "Contractual externalities and common agency equilibria," *Advances in Theoretical Economics*, vol. 3, no. 1, article 4, 2003.

[16] D. Martimort and L. Stole, "The revelation and delegation principles in common agency games," *Econometrica*, vol. 70, no. 4, pp. 1659–1673, 2002.

[17] D. Martimort and L. Stole, "Market participation in delegated and intrinsic common-agency games," *Journal of Economics*, vol. 40, no. 1, pp. 78–102, 2009.

[18] A. Dixit, G. M. Grossman, and E. Helpman, "Common agency and coordination: general theory and application to government policy making," *The Journal of Political Economy*, vol. 105, no. 4, pp. 752–769, 1997.

[19] D. Laussel and M. Le Breton, "Conflict and cooperation. The structure of equilibrium payoffs in common agency," *Journal of Economic Theory*, vol. 100, no. 1, pp. 93–128, 2001.

[20] A. Semenov, "Common agency with risk-averse agent," *Journal of Mathematical Economics*, vol. 46, no. 1, pp. 38–49, 2010.

[21] C. d'Aspremont and R. Dos Santos Ferreira, "Oligopolistic competition as a common agency game," *Games and Economic Behavior*, vol. 70, no. 1, pp. 21–33, 2010.

[22] D. Martimort and L. Stole, "Selecting equilibria in common agency games," *Journal of Economic Theory*, vol. 144, no. 2, pp. 604–634, 2009.

[23] A. Attar, G. Piaser, and N. Porteiro, "A note on common agency models of moral hazard," *Economics Letters*, vol. 95, no. 2, pp. 278–284, 2007.

[24] L. Stole, *Mechanism Design under Common Agency*, Mimeo, University of Chicago, 1991.

[25] B. S. Desgagne, *Incentives in Common Agency*, CIRANO Scientific Series, 2001.

[26] G. Biglaiser and C. Mezzetti, "Principals competing for an agent in the presence of adverse selection and moral hazard," *Journal of Economic Theory*, vol. 61, no. 2, pp. 302–330, 1993.

[27] C. Mezzetti, "Common agency with horizontally differentiated principals," *The RAND Journal of Economics*, vol. 28, no. 2, pp. 323–345, 1997.

[28] E. Bond and T. Gersik, *Regulation of Multinational Firms with Two Active Governments: A Common Agency Approach*, Mimeo, Department of Economics, Penn State University, 1993.

[29] V. Carrasco, "Common agency, organizational design and the hold-up problem," *Economics Letters*, vol. 108, no. 3, pp. 264–268, 2010.

[30] R. Dur and H. Roelfsema, "Social exchange and common agency in organizations," *Journal of Socio-Economics*, vol. 39, no. 1, pp. 55–63, 2010.

[31] K. Siqueira, T. Sandler, and J. Cauley, "Common agency and state-owned enterprise reform," *China Economic Review*, vol. 20, no. 2, pp. 208–217, 2009.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 486052, 17 pages
http://dx.doi.org/10.1155/2014/486052

Research Article

Preserving Global Exponential Stability of Hybrid BAM Neural Networks with Reaction Diffusion Terms in the Presence of Stochastic Noise and Connection Weight Matrices Uncertainty

Yan Li[1,2] and Yi Shen[1]

[1] *Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China*
[2] *College of Science, Huazhong Agriculture University, Wuhan 430079, China*

Correspondence should be addressed to Yi Shen; yeeshen0912@gmail.com

Received 8 February 2014; Accepted 6 March 2014; Published 17 April 2014

Academic Editor: Weihai Zhang

We study the impact of stochastic noise and connection weight matrices uncertainty on global exponential stability of hybrid BAM neural networks with reaction diffusion terms. Given globally exponentially stable hybrid BAM neural networks with reaction diffusion terms, the question to be addressed here is how much stochastic noise and connection weights matrices uncertainty the neural networks can tolerate while maintaining global exponential stability. The upper threshold of stochastic noise and connection weights matrices uncertainty is defined by using the transcendental equations. We find that the perturbed hybrid BAM neural networks with reaction diffusion terms preserve global exponential stability if the intensity of both stochastic noise and connection weights matrices uncertainty is smaller than the defined upper threshold. A numerical example is also provided to illustrate the theoretical conclusion.

1. Introduction

The bidirectional associative memory (BAM) neural networks were first introduced by Kosko in which the neurons in one layer are fully interconnected to the neurons in the other layer, while there are no interconnection among the neurons in the same layers [1–3]. The BAM neural networks widely have applications in pattern recognition, robot, signal processing, associative memory, solving optimization problems, and automatic control engineering. For most successful applications of BAM neural networks, the stability analysis on BAM neural networks is usually a prerequisite. The exponential stability and periodic oscillatory solution of BAM neural networks with delays were studied by Cao et al. [4, 5]. Moreover, in BAM neural networks, diffusion phenomena can hardly be avoided when electrons are moving in asymmetric electromagnetic fields. The BAM neural networks with reaction diffusion terms described by partial differential equations were investigated by many authors [6–11]. Sometimes, it is necessary to assess the parameters of the neural network that may experience abrupt changes caused by certain phenomena such as component failure or repair, change of subsystem interconnection, and environmental disturbance. The continuous-time Markov chains have been used to model these parameter jumps [12–14]. These neural networks with Markov chains are usually called hybrid neural networks. The almost surely exponential stability, moment exponential stability, and stabilization of hybrid neural networks were also researched; see, for example, [15–17]. By making use of impulsive control, Zhu and Cao [18] considered the stability of hybrid neural networks with mixed delay.

For neural networks with stochastic noise, the system is usually described by stochastic differential equations. The stability of stochastic neural networks with delay or reaction diffusion terms was extensively analyzed by using the Itô formula and the linear matrix inequality (LMI) methods [18–22]. As is well known, stochastic noise is often the sources of instability and may destabilize the stable neural networks [23]. For stable hybrid BAM neural networks with reaction diffusion terms, it is interesting to determine how much noise the stochastic neural networks can tolerate while maintaining global exponential stability.

Moreover, the connection weights of neurons depend on certain resistance and capacitance values which include uncertainty. The robust stability about parameter matrices uncertainty in neural networks was investigated by many authors [24, 25]. If the uncertainty in connection weights matrices is too large, the neural networks may be unstable. Therefore, for stable hybrid BAM neural networks with reaction diffusion terms, it is also interesting to determine how much connection weights matrices uncertainty the neural networks can also tolerate while maintaining global exponential stability.

In this paper, we will study the impact of stochastic noise and connection weight matrices uncertainty of hybrid BAM neural networks with reaction diffusion terms. We give the upper threshold of stochastic noise and connection weights matrices uncertainty defined by using the transcendental equations. We find that the perturbed hybrid BAM neural networks with reaction diffusion terms preserve global exponential stability if the intensity of both stochastic noise and connection weights matrices uncertainty is smaller than the defined upper threshold.

The remainder of this paper is organized as follows. Some preliminaries are given in Section 2. Section 3 discusses the impact of the stochastic noise on global exponential stability of these neural networks. Section 4 discusses the impact of the connection weight matrices uncertainty and stochastic noise on global exponential stability of these neural networks. Finally, an example with numerical simulation is given to illustrate the effectiveness of the obtained results in Section 5.

2. Preliminaries

Throughout this paper, unless otherwise specified, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e., it is increasing and right continuous while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets). Let $W(t)$ be a scalar Brownian motion (Wiener process) defined on the probability space. Let A^T denote the transpose of A. If A is a matrix, its operator norm is denoted by $\|A\| = \sup\{|Ax| : |x| = 1\}$, where $|\cdot|$ is the Euclidean norm. Let $r(t)$, $t \geq 0$, be a right-continuous Markov chain on the probability space taking values in a finite state space $\mathbb{S} = \{1, 2, \ldots, N\}$ with the generator $\Gamma = (\gamma_{pq})_{N\times N}$ given by

$$\mathbb{P}\{r(t+\Delta) = q \mid r(t) = p\} = \begin{cases} \gamma_{pq}\Delta + o(\Delta) & \text{if } p \neq q \\ 1 + \gamma_{pp}\Delta + o(\Delta) & \text{if } p = q, \end{cases} \tag{1}$$

where $\Delta > 0$. Here, $\gamma_{pq} > 0$ is the transition rate from p to q if $p \neq q$ while

$$\gamma_{pp} = -\sum_{q\neq p}\gamma_{pq}. \tag{2}$$

We assume that the Markov chain $r(\cdot)$ is independent of the Brownian motion $W(\cdot)$. It is well known that almost every sample path of $r(\cdot)$ is a right-continuous step function with finite number of simple jumps in any finite subinterval of $\mathbb{R}_+ := [0, +\infty)$.

In this paper, we will consider the following hybrid BAM neural networks with reaction diffusion terms:

$$\begin{aligned} \frac{\partial \tilde{u}_i(t,x)}{\partial t} &= \sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(\overline{D}_{ik}(r(t))\frac{\partial \tilde{u}_i(t,x)}{\partial x_k}\right) - a_i(r(t))\tilde{u}_i(t,x) + \sum_{j=1}^{n} c_{ji}(r(t))\tilde{f}_j(\tilde{v}_j(t,x)) + I_i, \\ \frac{\partial \tilde{v}_j(t,x)}{\partial t} &= \sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(\overline{D}^{*}_{jk}(r(t))\frac{\partial \tilde{v}_j(t,x)}{\partial x_k}\right) - b_j(r(t))\tilde{v}_j(t,x) + \sum_{i=1}^{m} e_{ij}(r(t))\tilde{g}_i(\tilde{u}_i(t,x)) + J_j, \end{aligned} \tag{3}$$

where $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, $t \geq t_0 \geq 0$, $t_0 \in \mathbb{R}_+$, and the initial value $r(t_0) = i_0 \in \mathbb{S}$. Consider $x = (x_1, x_2, \ldots, x_l) \in \Omega_0 \subset \mathbb{R}^l$; Ω_0 is a compact set with smooth boundary $\partial\Omega_0$ in space $\mathbb{R}^l$, and $0 < \operatorname{mes}\Omega_0 < +\infty$. $\tilde{u}(t,x) = (\tilde{u}_1(t,x), \ldots, \tilde{u}_m(t,x)) \in \mathbb{R}^m$ and $\tilde{v}(t,x) = (\tilde{v}_1(t,x), \ldots, \tilde{v}_n(t,x)) \in \mathbb{R}^n$ $\tilde{u}_i(t,x)$, $\tilde{v}_j(t,x)$, are the state of the ith neurons and the jth neurons at times t and in space x, respectively. $\tilde{f}_j$ and $\tilde{g}_i$ denote the signal functions on the jth neurons and the ith neurons at times t and in space x, respectively. I_i and J_j denote the external input on the ith neurons and the jth neurons, respectively. $a_i(r(t)) > 0$ and $b_j(r(t)) > 0$ denote the rates with which the ith neurons and the jth neurons will reset its potential to the resting state in isolation when disconnected from the networks and external inputs, respectively. $c_{ji}(r(t))$ and $e_{ij}(r(t))$ denote the strength of the jth neurons on the ith neurons and the ith neurons on the jth neurons, respectively. Smooth functions $\overline{D}_{ik}(r(t)) := \overline{D}_{ik}(r(t), x, u) \geq 0$ and $\overline{D}^{*}_{jk}(r(t)) := \overline{D}^{*}_{jk}(r(t), x, u) \geq 0$ correspond to the transmission diffusion operator along the ith neurons and the jth neurons, respectively.

The initial conditions and boundary conditions are given by

$$\begin{aligned} &\tilde{u}_i(t_0, x) = \overline{\phi}_i(x), \quad x \in \Omega_0, t_0 \in \mathbb{R}_+, \ i = 1, 2, \ldots, m, \\ &\tilde{v}_j(t_0, x) = \overline{\psi}_j(x), \quad x \in \Omega_0, t_0 \in \mathbb{R}_+, \ j = 1, 2, \ldots, n, \\ &\left.\frac{\partial \tilde{u}_i(t,x)}{\partial \vec{n}}\right|_{\partial\Omega_0} = \left(\frac{\partial \tilde{u}_i(t,x)}{\partial x_1}, \ldots, \frac{\partial \tilde{u}_i(t,x)}{\partial x_l}\right)^T = 0, \end{aligned}$$

$$(t,x) \in [t_0, +\infty) \times \partial\Omega_0, \quad i = 1, 2, \ldots, m,$$
$$\left.\frac{\partial \tilde{v}_j(t,x)}{\partial \vec{n}}\right|_{\partial\Omega_0} = \left(\frac{\partial \tilde{v}_j(t,x)}{\partial x_1}, \ldots, \frac{\partial \tilde{v}_j(t,x)}{\partial x_l}\right)^T = 0,$$
$$(t,x) \in [t_0, +\infty) \times \partial\Omega_0, \quad j = 1, 2, \ldots, n. \tag{4}$$

The neuron activation functions $\tilde{f}$ and $\tilde{g}$ are global Lipschitz continuous; that is, there exist constants $K > 0$ and $L > 0$, such that

$$\begin{aligned} \left|\tilde{f}(\tilde{v}) - \tilde{f}(\tilde{v}^*)\right| &\le K\left|\tilde{v} - \tilde{v}^*\right|, \quad \forall \tilde{v}, \tilde{v}^* \in \mathbb{R}^n \\ \left|\tilde{g}(\tilde{u}) - \tilde{g}(\tilde{u}^*)\right| &\le L\left|\tilde{u} - \tilde{u}^*\right|, \quad \forall \tilde{u}, \tilde{u}^* \in \mathbb{R}^m. \end{aligned} \tag{5}$$

Then, the neural networks (3) have a unique state ($\tilde{u}(t,x;t_0,\overline{\phi}(x))$ and $\tilde{v}(t,x;t_0,\overline{\psi}(x))$) for any initial values ($\overline{\phi}(x), \overline{\psi}(x)$) (see [26, 27]).

In addition, we assume that the neural networks (3) have an equilibrium point $u^* = (u_1^*, \ldots, u_m^*) \in \mathbb{R}^m$, $v^* = (v_1^*, \ldots, v_n^*) \in \mathbb{R}^n$.

Let $u(t,x) = \tilde{u}(t,x) - u^*$, $v(t,x) = \tilde{v}(t,x) - v^*$, $f(v(t,x)) = \tilde{f}(v(t,x) + v^*) - \tilde{f}(v^*)$, $g(u(t,x)) = \tilde{g}(u(t,x) + u^*) - \tilde{f}(u^*)$, $D_{ik}(r(t)) = \overline{D}_{ik}(r(t), x, u(t,x) + u^*)$, and $D_{ik}^*(r(t)) = \overline{D}_{ik}^*(r(t), x, v(t,x) + v^*)$, and then (3) can be rewritten as

$$\begin{aligned} \frac{\partial u_i(t,x)}{\partial t} &= \sum_{k=1}^{l} \frac{\partial}{\partial x_k}\left(D_{ik}(r(t)) \frac{\partial u_i(t,x)}{\partial x_k}\right) \\ &\quad - a_i(r(t)) u_i(t,x) \\ &\quad + \sum_{j=1}^{n} c_{ji}(r(t)) f_j(v_j(t,x)) \\ \frac{\partial v_j(t,x)}{\partial t} &= \sum_{k=1}^{l} \frac{\partial}{\partial x_k}\left(D_{jk}^*(r(t)) \frac{\partial v_j(t,x)}{\partial x_k}\right) \\ &\quad - b_j(r(t)) v_j(t,x) \\ &\quad + \sum_{i=1}^{m} e_{ij}(r(t)) g_i(u_i(t,x)). \end{aligned} \tag{6}$$

The initial conditions and boundary conditions are given by

$$u_i(t_0,x) = \phi_i(x) = \overline{\phi}_i(x) - u_i^*,$$
$$x \in \Omega_0, t_0 \in \mathbb{R}_+, \quad i = 1, 2, \ldots, m,$$
$$v_j(t_0,x) = \psi_j(x) = \overline{\psi}_j(x) - v_j^*,$$
$$x \in \Omega_0, t_0 \in \mathbb{R}_+, \quad j = 1, 2, \ldots, n,$$
$$\left.\frac{\partial u_i(t,x)}{\partial \vec{n}}\right|_{\partial\Omega_0} = \left(\frac{\partial u_i(t,x)}{\partial x_1}, \ldots, \frac{\partial u_i(t,x)}{\partial x_l}\right)^T = 0,$$
$$(t,x) \in [t_0, +\infty) \times \partial\Omega_0, \quad i = 1, 2, \ldots, m,$$
$$\left.\frac{\partial v_j(t,x)}{\partial \vec{n}}\right|_{\partial\Omega_0} = \left(\frac{\partial v_j(t,x)}{\partial x_1}, \ldots, \frac{\partial v_j(t,x)}{\partial x_l}\right)^T = 0,$$
$$(t,x) \in [t_0, +\infty) \times \partial\Omega_0, \quad j = 1, 2, \ldots, n. \tag{7}$$

Hence, the origin is an equilibrium point of (6). The stability of the equilibrium point of (3) is equivalent to the stability of the origin of the state space of (6).

From (5), we give the assumption about activations functions f and g.

Assumption (H1). The neuron activation functions f and g are global Lipschitz continuous; that is, there exist constants $K > 0$ and $L > 0$, such that

$$\begin{aligned} \left|f(v) - f(v^*)\right| &\le K\left|v - v^*\right|, \quad \forall v, v^* \in \mathbb{R}^n, \ f(0) = 0, \\ \left|g(u) - g(u^*)\right| &\le L\left|u - u^*\right|, \quad \forall u, u^* \in \mathbb{R}^m, \ g(0) = 0. \end{aligned} \tag{8}$$

We consider the following function vector space:

$$U = \begin{cases} v(t,x) : [t_0, +\infty) \times \Omega_0 \longrightarrow \mathbb{R}^n, \\ v(t,x) \text{ is continuous on } t \text{ and} \\ \quad \text{twice continuous differentiable on } x. \end{cases} \tag{9}$$

For every pair of (v, z) in U and every given $t \in \mathbb{R}_+$, define inner product for v and z with

$$\langle v, z \rangle = \int_{\Omega_0} (v(\cdot, x))^T z(\cdot, x)\, dx \in \mathbb{R}_+. \tag{10}$$

Obviously, it satisfies inner product axiom, and the norm can be deduced by

$$\begin{aligned} \|v(\cdot,x)\|_2 &= \sqrt{\langle v(\cdot,x), v(\cdot,x)\rangle} \\ &= \sqrt{\int_{\Omega_0} |v(\cdot,x)|^2 dx} = \sqrt{\sum_{i=1}^{n} \int_{\Omega_0} |v_i(\cdot,x)|^2 dx}. \end{aligned} \tag{11}$$

Definition 1. The neural networks (6) are said to be global exponentially stable if for any ϕ, ψ, there exist $\alpha > 0$ and $\beta > 0$, such that

$$\begin{aligned} &\|u(t,x;t_0,\phi)\|_2^2 + \|v(t,x;t_0,\psi)\|_2^2 \\ &\quad \le \alpha\left(\|\phi\|_2^2 + \|\psi\|_2^2\right) \exp(-\beta(t - t_0)), \quad \forall t \ge t_0. \end{aligned} \tag{12}$$

For the purpose of simplicity, we rewrite (6) as follows:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot (D(r(t)) \circ \nabla u) - A(r(t)) u(t,x) \\ &\quad + C(r(t)) f(v(t,x)) \\ \frac{\partial v}{\partial t} &= \nabla \cdot (D^*(r(t)) \circ \nabla v) - B(r(t)) v(t,x) \\ &\quad + E(r(t)) g(u(t,x)). \end{aligned} \tag{13}$$

The initial conditions and boundary conditions are given by

$$\begin{aligned}
&u(t_0,x)=\phi(x)=\overline{\phi}(x)-u^*, \quad x\in\Omega_0, t_0\in\mathbb{R}_+,\\
&v(t_0,x)=\psi(x)=\overline{\psi}(x)-v^*, \quad x\in\Omega_0, t_0\in\mathbb{R}_+,\\
&\left.\frac{\partial u(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial u(t,x)}{\partial x_1},\ldots,\frac{\partial u(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad (t,x)\in[t_0,+\infty)\times\partial\Omega_0,\\
&\left.\frac{\partial v(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial v(t,x)}{\partial x_1},\ldots,\frac{\partial v(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad (t,x)\in[t_0,+\infty)\times\partial\Omega_0,
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
D(r(t))&=(D_{ik}(r(t),x,u))_{m\times l},\\
D^*(r(t))&=(D^*_{jk}(r(t),x,v))_{n\times l},\\
u(t,x)&=(u_1(t,x),\ldots,u_m(t,x))^T,\\
v(t,x)&=(v_1(t,x),\ldots,v_n(t,x))^T,\\
\nabla u=(\nabla u_1,\ldots,\nabla u_m)^T,&\qquad \nabla v=(\nabla v_1,\ldots,\nabla v_n)^T,\\
\nabla u_i=\left(\frac{\partial u_i}{\partial x_1},\ldots,\frac{\partial u_i}{\partial x_l}\right)^T,&\qquad \nabla v_j=\left(\frac{\partial v_j}{\partial x_1},\ldots,\frac{\partial v_j}{\partial x_l}\right)^T,\\
A(r(t))&=\operatorname{diag}(a_1(r(t)),\ldots,a_m(r(t))),\\
B(r(t))&=\operatorname{diag}(b_1(r(t)),\ldots,b_n(r(t))),\\
C(r(t))&=(c_{ji}(r(t)))_{n\times m},\\
E(r(t))&=(e_{ij}(r(t)))_{m\times n},\\
f(v)&=(f_1(v_1),\ldots,f_n(v_n))^T,\\
g(u)&=(g_1(u_1),\ldots,g_m(u_m))^T,\\
(D(r(t))\circ\nabla u)&=\left(D_{ik}(r(t))\frac{\partial u_i}{\partial x_k}\right),\\
(D^*(r(t))\circ\nabla v)&=\left(D^*_{jk}(r(t))\frac{\partial v_j}{\partial x_k}\right).
\end{aligned} \tag{15}$$

Here, $\circ$ denotes Hadamard product of matrix D and ∇u and D^* and ∇v.

3. Noise Impact on Stability

In this section, we consider the noise-induced neural networks (6) described by the stochastic partial differential equations

$$\begin{aligned}
d\bar{u}_i(t,x)=&\left\{\sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(D_{ik}(r(t))\frac{\partial\bar{u}_i(t,x)}{\partial x_k}\right)\right.\\
&-a_i(r(t))\bar{u}_i(t,x)\\
&\left.+\sum_{j=1}^{n}c_{ji}(r(t))f_j(\bar{v}_j(t,x))\right\}dt\\
&+\sigma\bar{u}_i(t,x)\,dW(t),\\
d\bar{v}_j(t,x)=&\left\{\sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(D^*_{jk}(r(t))\frac{\partial\bar{v}_j(t,x)}{\partial x_k}\right)\right.\\
&-b_j(r(t))\bar{v}_j(t,x)\\
&\left.+\sum_{i=1}^{m}e_{ij}(r(t))g_i(\bar{u}_i(t,x))\right\}dt\\
&+\sigma\bar{v}_j(t,x)\,dW(t).
\end{aligned} \tag{16}$$

The initial conditions and boundary conditions are given by

$$\begin{aligned}
&\bar{u}_i(t_0,x)=\phi_i(x), \quad x\in\Omega_0,\ t_0\in\mathbb{R}_+,\ i=1,2,\ldots,m,\\
&\bar{v}_j(t_0,x)=\psi_j(x), \quad x\in\Omega_0,\ t_0\in\mathbb{R}_+,\ j=1,2,\ldots,n,\\
&\left.\frac{\partial\bar{u}_i(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial\bar{u}_i(t,x)}{\partial x_1},\ldots,\frac{\partial\bar{u}_i(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad (t,x)\in[t_0,+\infty)\times\partial\Omega_0, \quad i=1,2,\ldots,m,\\
&\left.\frac{\partial\bar{v}_j(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial\bar{v}_j(t,x)}{\partial x_1},\ldots,\frac{\partial\bar{v}_j(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad (t,x)\in[t_0,+\infty)\times\partial\Omega_0, \quad j=1,2,\ldots,n,
\end{aligned} \tag{17}$$

where σ is the noise intensity.

We rewrite (16) as follows:

$$\begin{aligned}
d\bar{u}(t,x)=&\{\nabla\cdot(D(r(t))\circ\nabla\bar{u})-A(r(t))\bar{u}(t,x)\\
&+C(r(t))f(\bar{v}(t,x))\}dt+\sigma\bar{u}(t,x)\,dW(t),\\
d\bar{v}(t,x)=&\{\nabla\cdot(D^*(r(t))\circ\nabla\bar{v})-B(r(t))\bar{v}(t,x)\\
&+E(r(t))g(\bar{u}(t,x))\}dt+\sigma\bar{v}(t,x)\,dW(t).
\end{aligned} \tag{18}$$

For the globally exponentially stable neural networks (6), we will characterize how much stochastic noise the neural networks (16) can tolerate while maintaining global exponential stability.

Definition 2. The neural networks (16) are said to be almost surely globally exponentially stable, if for any ϕ and ψ the Lyapunov exponent

$$\limsup_{t\to\infty}\frac{\log\left(\left\|\bar{u}\left(t,x;t_0,\phi\right)\right\|_2+\left\|\bar{v}\left(t,x;t_0,\psi\right)\right\|_2\right)}{t}<0,\quad \text{a.s.} \tag{19}$$

Definition 3. The neural networks (16) are said to be mean square globally exponentially stable, if, for any ϕ and ψ, the Lyapunov exponent

$$\limsup_{t\to\infty}\frac{\log\mathbb{E}\left\{\left\|\bar{u}\left(t,x;t_0,\phi\right)\right\|_2^2+\left\|\bar{v}\left(t,x;t_0,\psi\right)\right\|_2^2\right\}}{t}<0, \tag{20}$$

where $(\bar{u}(t,x;t_0,\phi),\ \bar{v}(t,x;t_0,\psi))$ is the state of neural networks (16).

From the above definitions, it is clear that the almost sure global exponential stability of the neural networks (16) implies the mean square global exponential stability of the neural networks (16) (see [26, 27]) but not vice versa.

Theorem 4. *Under Assumption (H1), the mean square global exponential stability of neural networks* (16) *implies the almost sure global exponential stability of the neural networks* (16).

Proof. For any $(\phi(x),\psi(x))\not\equiv(0,0)$, we denote the state $(\bar{u}(t,x;t_0,\phi),\ \bar{v}(t,x;t_0,\psi))$ of (16) as $(\bar{u}(t,x),\bar{v}(t,x))$. By Definition 3, there exist $\lambda>0$ and $C>0$, such that

$$\begin{aligned}&\mathbb{E}\left\{\left\|\bar{u}\left(t,x\right)\right\|_2^2+\left\|\bar{v}\left(t,x\right)\right\|_2^2\right\}\\&\quad\le C\left(\|\phi\|_2^2+\|\psi\|_2^2\right)e^{-\lambda(t-t_0)},\quad t\ge t_0.\end{aligned} \tag{21}$$

Let $r(t)=p\in\mathbb{S}$. Construct average Lyapunov functional

$$\begin{aligned}&V\left(\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),p\right)\\&\quad=\int_{\Omega_0}\left|\bar{u}\left(t,x\right)\right|^2\mathrm{d}x+\int_{\Omega_0}\left|\bar{v}\left(t,x\right)\right|^2\mathrm{d}x\\&\quad=\int_{\Omega_0}\sum_{i=1}^{m}\bar{u}_i^2\left(t,x\right)\mathrm{d}x+\int_{\Omega_0}\sum_{j=1}^{n}\bar{v}_j^2\left(t,x\right)\mathrm{d}x.\end{aligned} \tag{22}$$

Let $n=1,2,\ldots$, by Itô formula and Assumption (H1), for $t_0+n-1\le t\le t_0+n$,

$$\begin{aligned}&V\left(\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),p\right)\\&=V\left(\bar{u}\left(t_0+n-1,x\right),\bar{v}\left(t_0+n-1,x\right),p\right)\\&\quad+\int_{t_0+n-1}^{t}\int_{\Omega_0}2\bar{u}^T\left(s,x\right)\left[\nabla\cdot\left(D\left(r\left(s\right)\right)\circ\nabla\bar{u}\right)-A\left(r\left(s\right)\right)\bar{u}\right.\\&\qquad\left.+C\left(r\left(s\right)\right)f\left(\bar{v}\right)\right]\mathrm{d}x\mathrm{d}s\\&\quad+\sigma^2\int_{t_0+n-1}^{t}\int_{\Omega_0}\left|\bar{u}\left(s,x\right)\right|^2\mathrm{d}x\mathrm{d}s\\&\quad+\int_{t_0+n-1}^{t}\int_{\Omega_0}2\bar{v}^T\left(s,x\right)\left[\nabla\cdot\left(D^*\left(r\left(s\right)\right)\circ\nabla\bar{v}\right)-B\left(r\left(s\right)\right)\bar{v}\right.\\&\qquad\left.+E\left(r\left(s\right)\right)g\left(\bar{u}\right)\right]\mathrm{d}x\mathrm{d}s\\&\quad+\sigma^2\int_{t_0+n-1}^{t}\int_{\Omega_0}\left|\bar{v}\left(s,x\right)\right|^2\mathrm{d}x\mathrm{d}s\\&\quad+2\sigma\int_{t_0+n-1}^{t}\int_{\Omega_0}\left|\bar{u}\left(s,x\right)\right|^2\mathrm{d}x\mathrm{d}W\left(s\right)\\&\quad+2\sigma\int_{t_0+n-1}^{t}\int_{\Omega_0}\left|\bar{v}\left(s,x\right)\right|^2\mathrm{d}x\mathrm{d}W\left(s\right)\\&\quad+\sum_{q=1}^{N}\gamma_{pq}\int_{t_0+n-1}^{t}\int_{\Omega_0}\left(\left|\bar{u}\left(s,x\right)\right|^2+\left|\bar{v}\left(s,x\right)\right|^2\right)\mathrm{d}x\mathrm{d}s.\end{aligned} \tag{23}$$

By boundary condition and Gauss formula, we get

$$\begin{aligned}&2\int_{\Omega_0}\bar{u}^T\left(s,x\right)\left[\nabla\cdot\left(D\left(r\left(s\right)\right)\circ\nabla\bar{u}\right)\right]\mathrm{d}x\\&\quad=2\sum_{i=1}^{m}\sum_{k=1}^{l}\int_{\Omega_0}\bar{u}_i\frac{\partial}{\partial x_k}\left(D_{ik}\left(r\left(s\right)\right)\frac{\partial\bar{u}_i}{\partial x_k}\right)\mathrm{d}x\\&\quad=2\sum_{i=1}^{m}\int_{\Omega_0}\nabla\cdot\left(\bar{u}_iD_{ik}\left(r\left(s\right)\right)\frac{\partial\bar{u}_i}{\partial x_k}\right)_{k=1}^{l}\mathrm{d}x\\&\qquad-2\sum_{i=1}^{m}\int_{\Omega_0}\left(D_{ik}\left(r\left(s\right)\right)\frac{\partial\bar{u}_i}{\partial x_k}\right)_{k=1}^{l}\cdot\nabla\bar{u}_i\mathrm{d}x\\&\quad=2\sum_{i=1}^{m}\int_{\partial\Omega_0}\left(\bar{u}_iD_{ik}\left(r\left(s\right)\right)\frac{\partial\bar{u}_i}{\partial x_k}\right)_{k=1}^{l}\mathrm{d}x\\&\qquad-2\sum_{i=1}^{m}\sum_{k=1}^{l}\int_{\Omega_0}D_{ik}\left(r\left(s\right)\right)\left(\frac{\partial\bar{u}_i}{\partial x_k}\right)^2\mathrm{d}x\\&\quad=-2\sum_{i=1}^{m}\sum_{k=1}^{l}\int_{\Omega_0}D_{ik}\left(r\left(s\right)\right)\left(\frac{\partial\bar{u}_i}{\partial x_k}\right)^2\mathrm{d}x,\end{aligned} \tag{24}$$

$$\begin{aligned}&2\int_{\Omega_0}\bar{v}^T\left(s,x\right)\left[\nabla\cdot\left(D^*\left(r\left(s\right)\right)\circ\nabla\bar{v}\right)\right]\mathrm{d}x\\&\quad=2\sum_{j=1}^{n}\sum_{k=1}^{l}\int_{\Omega_0}\bar{v}_j\frac{\partial}{\partial x_k}\left(D^*_{jk}\left(r\left(s\right)\right)\frac{\partial\bar{v}_j}{\partial x_k}\right)\mathrm{d}x\\&\quad=2\sum_{j=1}^{n}\int_{\Omega_0}\nabla\cdot\left(\bar{v}_jD^*_{jk}\left(r\left(s\right)\right)\frac{\partial\bar{v}_j}{\partial x_k}\right)_{k=1}^{l}\mathrm{d}x\\&\qquad-2\sum_{j=1}^{n}\int_{\Omega_0}\left(D^*_{jk}\left(r\left(s\right)\right)\frac{\partial\bar{v}_j}{\partial x_k}\right)_{k=1}^{l}\cdot\nabla\bar{v}_j\mathrm{d}x\end{aligned}$$

$$= 2\sum_{j=1}^{n}\int_{\partial\Omega_0}\left(\bar{v}_j D_{jk}^{*}(r(s))\frac{\partial\bar{v}_j}{\partial x_k}\right)_{k=1}^{l}\mathrm{d}x$$
$$-2\sum_{j=1}^{n}\sum_{k=1}^{l}\int_{\Omega_0} D_{jk}^{*}(r(s))\left(\frac{\partial\bar{v}_j}{\partial x_k}\right)^2\mathrm{d}x$$
$$= -2\sum_{j=1}^{n}\sum_{k=1}^{l}\int_{\Omega_0} D_{jk}^{*}(r(s))\left(\frac{\partial\bar{v}_j}{\partial x_k}\right)^2\mathrm{d}x. \tag{25}$$

By Hölder's inequality, we have

$$\int_{\Omega_0} 2\bar{u}(t,x)^T C(r(s))\, f(\bar{v}(t,x))\,\mathrm{d}x$$
$$\leq \max_{p\in\mathbb{S}}\|C(p)\|\left[\int_{\Omega_0}|\bar{u}(t,x)|^2\mathrm{d}x + K^2\int_{\Omega_0}|\bar{v}(t,x)|^2\mathrm{d}x\right], \tag{26}$$

$$\int_{\Omega_0} 2\bar{v}(t,x)^T E(r(s))\, g(\bar{u}(t,x))\,\mathrm{d}x$$
$$\leq \max_{p\in\mathbb{S}}\|E(p)\|\left[\int_{\Omega_0}|\bar{v}(t,x)|^2\mathrm{d}x + L^2\int_{\Omega_0}|\bar{u}(t,x)|^2\mathrm{d}x\right]. \tag{27}$$

Substituting (24)–(27) into (23), we get

$$V(\bar{u}(t,x),\bar{v}(t,x),p)$$
$$= V(\bar{u}(t_0+n-1,x),\bar{v}(t_0+n-1,x),p)$$
$$+\max_{p\in\mathbb{S}}\left[2\|A(p)\| + \|C(p)\| + \|E(p)\|L^2 + \sigma^2\right]$$
$$\times\int_{t_0+n-1}^{t}\int_{\Omega_0}|\bar{u}(s,x)|^2\mathrm{d}x\mathrm{d}s$$
$$+\max_{p\in\mathbb{S}}\left[2\|B(p)\| + \|E(p)\| + \|C(p)\|K^2 + \sigma^2\right]$$
$$\times\int_{t_0+n-1}^{t}\int_{\Omega_0}|\bar{v}(s,x)|^2\mathrm{d}x\mathrm{d}s$$
$$+2|\sigma|\int_{t_0+n-1}^{t}\int_{\Omega_0}\left(|\bar{u}(s,x)|^2 + |\bar{v}(s,x)|^2\right)\mathrm{d}x\mathrm{d}W(s), \tag{28}$$

where we use $\sum_{q=1}^{N}\gamma_{pq} = 0$.

From (28), we have

$$\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n} V(\bar{u}(t,x),\bar{v}(t,x),p)\right)$$
$$\leq V(\bar{u}(t_0+n-1,x),\bar{v}(t_0+n-1,x),p)$$
$$+C_1\int_{t_0+n-1}^{t_0+n}\mathbb{E}V(\bar{u}(s,x),\bar{v}(s,x),r(s))\,\mathrm{d}s$$
$$+2|\sigma|\,\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n}\int_{t_0+n-1}^{t} V(\bar{u}(s,x),\bar{v}(s,x),\right.$$
$$\left. r(s))\,\mathrm{d}W(s)\right), \tag{29}$$

where $C_1 = [2\|\widehat{A}\| + 2\|\widehat{B}\| + \|\widehat{C}\| + \|\widehat{C}\|K^2 + \|\widehat{E}\| + \|\widehat{E}\|L^2 + 2\sigma^2]$ and $\|\widehat{A}\| = \max_{p\in\mathbb{S}}\|A(p)\|$.

On the other hand, by the Burkholder-Davis-Gundy inequality [27] and $2\sqrt{ab} \leq (a/\varepsilon) + \varepsilon b(a > 0, b > 0, \varepsilon > 0)$, we have

$$2|\sigma|\,\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n}\int_{t_0+n-1}^{t} V(\bar{u}(s,x),\bar{v}(s,x),r(s))\,\mathrm{d}W(s)\right)$$
$$\leq 4\sqrt{2}\mathbb{E}\left(\int_{t_0+n-1}^{t_0+n} 4\sigma^2 V^2(\bar{u}(s,x),\bar{v}(s,x),r(s))\,\mathrm{d}s\right)^{1/2}$$
$$\leq 4\sqrt{2}\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n} V(\bar{u}(s,x),\bar{v}(s,x),r(s))\right.$$
$$\left.\times\int_{t_0+n-1}^{t_0+n} 4\sigma^2 V(\bar{u}(s,x),\bar{v}(s,x),r(s))\,\mathrm{d}s\right)^{1/2}$$
$$\leq \frac{1}{2}\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n} V(\bar{u}(t,x),\bar{v}(t,x),p)\right)$$
$$+64\sigma^2\int_{t_0+n-1}^{t_0+n}\mathbb{E}V(\bar{u}(s,x),\bar{v}(s,x),r(s))\,\mathrm{d}s. \tag{30}$$

Substituting the above inequality into (29), we get

$$\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n} V(\bar{u}(t,x),\bar{v}(t,x),p)\right)$$
$$\leq 2\mathbb{E}V(\bar{u}(t_0+n-1,x),\bar{v}(t_0+n-1,x),p) \tag{31}$$
$$+2\left[C_1 + 64\sigma^2\right]\int_{t_0+n-1}^{t_0+n}\mathbb{E}V(s)\,\mathrm{d}s.$$

By induction and the mean square global exponential stability of neural networks (16),

$$\mathbb{E}\left(\sup_{t_0+n-1\le t\le t_0+n} V(\bar{u}(t,x),\bar{v}(t,x),p)\right)$$
$$\leq C\left(\|\phi\|_2^2 + \|\psi\|_2^2\right)\left(2 + 2\left[C_1 + 64\sigma^2\right]\right)e^{-\lambda(n-1)}. \tag{32}$$

Let $\varepsilon \in (0,\lambda)$, by Chebyshev's inequality [27], it follows from (32) that

$$\mathbb{P}\left\{\sup_{t_0+n-1\le t\le t_0+n} V(\bar{u}(t,x),\bar{v}(t,x),p) > e^{-(\lambda-\varepsilon)(n-1)}\right\}$$

$$\leq e^{-(\lambda-\varepsilon)(n-1)}\mathbb{E}\left(\sup_{t_0+n-1\leq t\leq t_0+n} V\left(\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),p\right)\right)$$
$$\leq C\left(\|\phi\|_2^2+\|\psi\|_2^2\right)\left(2+2\left[C_1+64\sigma^2\right]\right)e^{-\varepsilon(n-1)}. \tag{33}$$

By Borel-Cantelli Lemma [27], for almost all $\omega \in \Omega$,

$$\sup_{t_0+n-1\leq t\leq t_0+n} 2V\left(\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),p\right)\leq 2e^{-(\lambda-\varepsilon)(n-1)} \tag{34}$$

holds for all but finitely many n. Hence, there exists an $n_0 = n_0(\omega)$, for all $\omega \in \Omega$, excluding a $\mathbb{P}$-null set, for the above inequality that holds whenever $n \geq n_0$. Consequently, for almost all $\omega \in \Omega$,

$$\frac{\log 2V\left(\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),p\right)}{t}\leq -\frac{\left(\lambda-\varepsilon\right)\left(n-1\right)}{t_0+n-1}+\frac{2}{t_0+n-1}, \tag{35}$$

if $t_0+n-1\leq t\leq t_0+n$. Therefore,

$$\limsup_{t\to\infty}\frac{\log\left(\|\bar{u}\left(t,x\right)\|_2+\|\bar{v}\left(t,x\right)\|_2\right)}{t}\leq -\frac{\left(\lambda-\varepsilon\right)}{2}\quad \text{a.s.} \tag{36}$$

□

Theorem 5. *Let Assumption (H1) hold and the neural networks* (6) *be globally exponentially stable. Then, the neural networks* (16) *is mean square globally exponentially stable and also almost surely globally exponentially stable, if there exist* $\mu_q > 0, (q \in \mathbb{S})$ *and* $|\sigma| < \bar{\sigma}$, *where* $\bar{\sigma}$ *is a unique positive solution of the transcendental equation*

$$\frac{4\bar{\sigma}^2\alpha\widehat{\mu}}{\beta}\exp\left\{\frac{2\Delta\left(\widehat{\mu}C_2+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\}+2\alpha\exp\left\{-\beta\Delta\right\}=1, \tag{37}$$

$$\Delta>\frac{\ln\left(2\alpha\right)}{\beta}>0, \tag{38}$$

where $C_2=[2\|\widehat{A}\|+2\|\widehat{B}\|+(1+K^2)\|\widehat{C}\|+(1+L^2)\|\widehat{E}\|+2\bar{\sigma}^2]$, $\|\widehat{A}\|=\max_{p\in\mathbb{S}}\|A(p)\|$, *and so forth and* $\widehat{\mu}=\max_{p\in\mathbb{S}}\mu_p$ *and* $\breve{\mu}=\min_{p\in\mathbb{S}}\mu_p$.

Proof. For any $(\phi(x),\psi(x))$, we denote the state $(\bar{u}(t,x;t_0,\phi),\bar{v}(t,x;t_0,\psi))$ of (16) as $(\bar{u}(t,x),\bar{v}(t,x))$ and the state $(u(t,x;t_0,\phi),v(t,x;t_0,\psi))$ of (6) as $(u(t,x),v(t,x))$.

From (6) and (18) and stochastic Fubini's Theorem, we have

$$\begin{aligned}
&\int_{\Omega_0}\left(u\left(t,x\right)-\bar{u}\left(t,x\right)\right)\mathrm{d}x+\int_{\Omega_0}\left(v\left(t,x\right)-\bar{v}\left(t,x\right)\right)\mathrm{d}x\\
&=\int_{t_0}^{t}\int_{\Omega_0}\nabla\cdot\left(D\left(r\left(s\right)\right)\circ\nabla\left(u-\bar{u}\right)\right)\mathrm{d}x\mathrm{d}s\\
&+\int_{t_0}^{t}\int_{\Omega_0}\left[-A\left(r\left(s\right)\right)\left(u\left(s,x\right)-\bar{u}\left(s,x\right)\right)\right.\\
&\qquad\left.+C\left(r\left(s\right)\right)\left(f\left(v\left(s,x\right)\right)-f\left(\bar{v}\left(s,x\right)\right)\right)\right]\mathrm{d}x\mathrm{d}s\\
&-\int_{t_0}^{t}\int_{\Omega_0}\sigma\bar{u}\left(s,x\right)\mathrm{d}x\mathrm{d}W\left(s\right)\\
&+\int_{t_0}^{t}\int_{\Omega_0}\nabla\cdot\left(D^*\left(r\left(s\right)\right)\circ\nabla\left(v-\bar{v}\right)\right)\mathrm{d}x\mathrm{d}s\\
&+\int_{t_0}^{t}\int_{\Omega_0}\left[-B\left(r\left(s\right)\right)\left(v\left(s,x\right)-\bar{v}\left(s,x\right)\right)\right.\\
&\qquad\left.+E\left(r\left(s\right)\right)\left(g\left(u\left(s,x\right)\right)-g\left(\bar{u}\left(s,x\right)\right)\right)\right]\mathrm{d}x\mathrm{d}s\\
&-\int_{t_0}^{t}\int_{\Omega_0}\sigma\bar{v}\left(s,x\right)\mathrm{d}x\mathrm{d}W\left(s\right).
\end{aligned} \tag{39}$$

Construct average Lyapunov functional

$$\begin{aligned}
&V\left(u\left(t,x\right),v\left(t,x\right),\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),r\left(t\right)\right)\\
&=\int_{\Omega_0}\mu_{r(t)}\left[\left|u\left(t,x\right)-\bar{u}\left(t,x\right)\right|^2+\left|v\left(t,x\right)-\bar{v}\left(t,x\right)\right|^2\right]\mathrm{d}x,
\end{aligned} \tag{40}$$

where $\mu_{r(t)}>0$.

By applying generalized Itô formula [27], we have

$$\begin{aligned}
&\mathrm{d}V\left(u\left(t,x\right),v\left(t,x\right),\bar{u}\left(t,x\right),\bar{v}\left(t,x\right),p\right)\\
&=\int_{\Omega_0}2\mu_p(u\left(t,x\right)-\bar{u}\left(t,x\right))^T\left(\nabla\cdot\left(D\left(p\right)\circ\nabla\left(u-\bar{u}\right)\right)\right)\mathrm{d}x\mathrm{d}t\\
&+\int_{\Omega_0}2\mu_p(u\left(t,x\right)-\bar{u}\left(t,x\right))^T\\
&\qquad\times\left[-A\left(p\right)\left(u\left(t,x\right)-\bar{u}\left(t,x\right)\right)\right.\\
&\qquad\left.+C\left(p\right)\left(f\left(v\left(t,x\right)\right)-f\left(\bar{v}\left(t,x\right)\right)\right)\right]\mathrm{d}x\mathrm{d}t\\
&+\int_{\Omega_0}\sigma^2\mu_p|\bar{u}\left(t,x\right)|^2\mathrm{d}x\mathrm{d}t\\
&-2\int_{\Omega_0}\sigma(u\left(t,x\right)-\bar{u}\left(t,x\right))^T\bar{u}\left(t,x\right)\mathrm{d}x\mathrm{d}W\left(t\right)\\
&+\int_{\Omega_0}2\mu_p(v\left(t,x\right)-\bar{v}\left(t,x\right))^T\\
&\qquad\times\left(\nabla\cdot\left(D^*\left(p\right)\circ\nabla\left(v-\bar{v}\right)\right)\right)\mathrm{d}x\mathrm{d}t\\
&+\int_{\Omega_0}2\mu_p(v\left(t,x\right)-\bar{v}\left(t,x\right))^T\\
&\qquad\times\left[-B\left(p\right)\left(v\left(t,x\right)-\bar{v}\left(t,x\right)\right)\right.\\
&\qquad\left.+E\left(p\right)\left(g\left(u\left(t,x\right)\right)-g\left(\bar{u}\left(t,x\right)\right)\right)\right]\mathrm{d}x\mathrm{d}t\\
&+\int_{\Omega_0}\sigma^2\mu_p|\bar{v}\left(t,x\right)|^2\mathrm{d}x\mathrm{d}t
\end{aligned}$$

$$-2\int_{\Omega_0}\sigma(v(t,x)-\bar{v}(t,x))^T\bar{v}(t,x)\,\mathrm{d}x\mathrm{d}W(t)$$
$$+\sum_{q=1}^{N}\gamma_{pq}\mu_q\int_{\Omega_0}\left[|u(t,x)-\bar{u}(t,x)|^2+|v(t,x)-\bar{v}(t,x)|^2\right]\mathrm{d}x. \tag{41}$$

By boundary condition and (24), we have

$$2\mu_p\int_{\Omega_0}(u(t,x)-\bar{u}(t,x))^T(\nabla\cdot(D(p)\circ\nabla(u-\bar{u})))\,\mathrm{d}x\mathrm{d}t = -2\mu_p\sum_{i=1}^{m}\sum_{k=1}^{l}\int_{\Omega_0}D_{ik}(p)\left(\frac{\partial(u_i-\bar{u}_i)}{\partial x_k}\right)^2\mathrm{d}x. \tag{42}$$

By boundary condition and (25), we have

$$2\mu_p\int_{\Omega_0}(v(t,x)-\bar{v}(t,x))^T(\nabla\cdot(D^*(p)\circ\nabla(v-\bar{v})))\,\mathrm{d}x\mathrm{d}t = -2\mu_p\sum_{j=1}^{n}\sum_{k=1}^{l}\int_{\Omega_0}D^*_{jk}(p)\left(\frac{\partial(v_j-\bar{v}_j)}{\partial x_k}\right)^2\mathrm{d}x. \tag{43}$$

By Hölder's inequality, we get

$$2\mu_p\int_{\Omega_0}(u(t,x)-\bar{u}(t,x))^TC(p)\times(f(v(t,x))-f(\bar{v}(t,x)))\,\mathrm{d}x \le \max_{p\in\mathbb{S}}(\mu_p\|C(p)\|)\left[\int_{\Omega_0}|u(t,x)-\bar{u}(t,x)|^2\mathrm{d}x+K^2\int_{\Omega_0}|v(t,x)-\bar{v}(t,x)|^2\mathrm{d}x\right], \tag{44}$$

$$2\mu_p\int_{\Omega_0}(v(t,x)-\bar{v}(t,x))^TE(p)\times(g(u(t,x))-g(\bar{u}(t,x)))\,\mathrm{d}x \le \max_{p\in\mathbb{S}}(\mu_p\|E(p)\|)\left[\int_{\Omega_0}|v(t,x)-\bar{v}(t,x)|^2\mathrm{d}x+L^2\int_{\Omega_0}|u(t,x)-\bar{u}(t,x)|^2\mathrm{d}x\right]. \tag{45}$$

From (42)–(45) and Assumption (H1), we obtain that

$$\begin{aligned}\mathrm{d}V(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),p) &\le \left(\hat{\mu}C_1+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)\times\int_{\Omega_0}\left(|u(t,x)-\bar{u}(t,x)|^2+|v(t,x)-\bar{v}(t,x)|^2\right)\mathrm{d}x\mathrm{d}t\\ &+2\sigma^2\hat{\mu}\int_{\Omega_0}\left(|u(t,x)|^2+|v(t,x)|^2\right)\mathrm{d}x\mathrm{d}t\\ &-2\int_{\Omega_0}\sigma(u(t,x)-\bar{u}(t,x))^T\bar{u}(t,x)\,\mathrm{d}x\mathrm{d}W(t)\\ &-2\int_{\Omega_0}\sigma(v(t,x)-\bar{v}(t,x))^T\bar{v}(t,x)\,\mathrm{d}x\mathrm{d}W(t).\end{aligned} \tag{46}$$

When $t\le t_0+2\Delta$, we have

$$\begin{aligned}\mathbb{E}V(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t)) &\le \left(\hat{\mu}C_1+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)\times\int_{t_0}^{t}\mathbb{E}\int_{\Omega_0}\left(|u(s,x)-\bar{u}(s,x)|^2+|v(s,x)-\bar{v}(s,x)|^2\right)\mathrm{d}x\mathrm{d}s\\ &+2\sigma^2\hat{\mu}\int_{t_0}^{t}\alpha\left(\|\phi\|_2^2+\|\psi\|_2^2\right)\exp(-\beta(s-t_0))\,\mathrm{d}s\\ &-2\sigma\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(u(s,x)-\bar{u}(s,x))^T\bar{u}(s,x)\,\mathrm{d}x\mathrm{d}W(s)\\ &-2\sigma\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(v(s,x)-\bar{v}(s,x))^T\bar{v}(s,x)\,\mathrm{d}x\mathrm{d}W(s).\end{aligned} \tag{47}$$

By stochastic Fubini's Theorem, we have

$$\begin{aligned}&\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(u(s,x)-\bar{u}(s,x))^T\bar{u}(s,x)\,\mathrm{d}x\mathrm{d}W(s)=0,\\ &\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(v(s,x)-\bar{v}(s,x))^T\bar{v}(s,x)\,\mathrm{d}x\mathrm{d}W(s)=0.\end{aligned} \tag{48}$$

By (47), one get

$$\begin{aligned}\mathbb{E}V(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t)) &\le \frac{\left(\hat{\mu}C_1+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\times\int_{t_0}^{t}\mathbb{E}V(u(s,x),v(s,x),\bar{u}(s,x),\bar{v}(s,x),r(s))\,ds\\ &+\frac{2\sigma^2\alpha\hat{\mu}\left(\|\phi\|_2^2+\|\psi\|_2^2\right)}{\beta}.\end{aligned} \tag{49}$$

When $t_0+\Delta\le t\le t_0+2\Delta$, by applying Gronwall's inequality, we have

$$\begin{aligned}\mathbb{E}\left(\|u(t,x)-\bar{u}(t,x)\|_2^2+\|v(t,x)-\bar{v}(t,x)\|_2^2\right) &= \mathbb{E}V(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t))\\ &\le \frac{2\sigma^2\alpha\hat{\mu}\left(\|\phi\|_2^2+\|\psi\|_2^2\right)}{\beta}\end{aligned}$$

$$\times \exp \frac{\left(\widehat{\mu}C_1 + \max_{p\in\mathbb{S}} \sum_{q=1}^{N} \gamma_{pq}\mu_q\right)}{\breve{\mu}} (t - t_0)$$
$$\leq \sup_{t_0\leq t\leq t_0+\Delta} \mathbb{E}\left(\|\bar{u}(t,x)\|^2 + \|\bar{v}(t,x)\|^2\right)$$
$$\times \frac{2\sigma^2\alpha\widehat{\mu}}{\beta} \exp \frac{2\Delta\left(\widehat{\mu}C_1 + \max_{p\in\mathbb{S}} \sum_{q=1}^{N} \gamma_{pq}\mu_q\right)}{\breve{\mu}}. \tag{50}$$

By the global exponential stability of (6), we have

$$\begin{aligned}
&\mathbb{E}\left(\|\bar{u}(t,x)\|_2^2 + \|\bar{v}(t,x)\|_2^2\right) \\
&\leq 2\mathbb{E}\left(\|u(t,x) - \bar{u}(t,x)\|_2^2 + \|v(t,x) - \bar{v}(t,x)\|_2^2\right) \\
&\quad + 2\mathbb{E}\left(\|u(t,x)\|_2^2 + \|v(t,x)\|_2^2\right) \\
&\leq \sup_{t_0\leq t\leq t_0+\Delta} \mathbb{E}\left(\|\bar{u}(t,x)\|^2 + \|\bar{v}(t,x)\|^2\right) \\
&\quad \times \frac{4\sigma^2\alpha\widehat{\mu}}{\beta} \exp \frac{2\Delta\left(\widehat{\mu}C_1 + \max_{p\in\mathbb{S}} \sum_{q=1}^{N} \gamma_{pq}\mu_q\right)}{\breve{\mu}} \\
&\quad + 2\alpha\left(\|\phi\|_2^2 + \|\psi\|_2^2\right) \exp\{-\beta(t - t_0)\}.
\end{aligned} \tag{51}$$

Moreover,

$$\begin{aligned}
&\mathbb{E}\left(\|\bar{u}(t,x)\|_2^2 + \|\bar{v}(t,x)\|_2^2\right) \\
&\leq \left\{\frac{4\sigma^2\alpha\widehat{\mu}}{\beta} \exp\left\{\frac{2\Delta\left(\widehat{\mu}C_1 + \max_{p\in\mathbb{S}} \sum_{q=1}^{N} \gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\}\right. \\
&\quad \left. + 2\alpha \exp\{-\beta\Delta\}\right\} \\
&\quad \times \sup_{t_0\leq t\leq t_0+\Delta} \mathbb{E}\left(\|\bar{u}(t,x)\|^2 + \|\bar{v}(t,x)\|^2\right).
\end{aligned} \tag{52}$$

From (37), when $|\sigma| < \bar{\sigma}$, we have

$$\frac{4\sigma^2\alpha\widehat{\mu}}{\beta} \exp\left\{\frac{2\Delta\left(\widehat{\mu}C_1 + \max_{p\in\mathbb{S}} \sum_{q=1}^{N} \gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\} + 2\alpha \exp\{-\beta\Delta\} < 1. \tag{53}$$

Let

$$\begin{aligned}
\gamma = &\left(-\log\left\{\frac{4\sigma^2\alpha\widehat{\mu}}{\beta}\right.\right. \\
&\quad \times \exp\left\{\frac{2\Delta\left(\widehat{\mu}C_1 + \max_{p\in\mathbb{S}} \sum_{q=1}^{N} \gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\} \\
&\quad \left.\left. + 2\alpha \exp\{-\beta\Delta\}\right\}\right) \\
&\quad \times (\Delta)^{-1} > 0.
\end{aligned} \tag{54}$$

By (52), we have

$$\begin{aligned}
&\sup_{t_0+\Delta\leq t\leq t_0+2\Delta} \mathbb{E}\left(\|\bar{u}(t,x)\|_2^2 + \|\bar{v}(t,x)\|_2^2\right) \\
&\quad \leq \exp(-\gamma\Delta)\left(\sup_{t_0\leq t\leq t_0+\Delta} \mathbb{E}\left(\|\bar{u}(t,x)\|_2^2 + \|\bar{v}(t,x)\|_2^2\right)\right).
\end{aligned} \tag{55}$$

For any positive integer $m = 1, 2, \ldots$, from the existence and uniqueness of the flow of (16) (see [28]), when $t \geq t_0 + (m - 1)\Delta$, we have

$$\begin{aligned}
&(\bar{u}(t,x;t_0,\phi), \bar{v}(t,x;t_0,\psi)) \\
&\quad = (\bar{u}(t,x;t_0 + (m-1)\Delta, \bar{u}(t_0 + (m-1)\Delta, x; t_0, \phi)), \\
&\qquad \bar{v}(t,x;t_0 + (m-1)\Delta, \bar{v}(t_0 + (m-1)\Delta, x; t_0, \psi))).
\end{aligned} \tag{56}$$

From (55) and (56),

$$\begin{aligned}
&\sup_{t_0+m\Delta\leq t\leq t_0+(m+1)\Delta} \mathbb{E}\left(\|\bar{u}(t,x;t_0,\phi)\|_2^2 + \|\bar{v}(t,x;t_0,\psi)\|_2^2\right) \\
&= \sup_{t_0+(m-1)\Delta+\Delta\leq t\leq t_0+(m-1)\Delta+2\Delta} \mathbb{E}\,\|\bar{u}(t,x;t_0 + (m-1)\Delta, \\
&\qquad \bar{u}(t_0 + (m-1)\Delta, x; t_0, \phi))\|_2^2 \\
&\quad + \sup_{t_0+(m-1)\Delta+\Delta\leq t\leq t_0+(m-1)\Delta+2\Delta} \mathbb{E}\,\|\bar{v}(t,x;t_0 + (m-1)\Delta, \\
&\qquad \bar{v}(t_0 + (m-1)\Delta, x; t_0, \psi))\|_2^2 \\
&\leq \exp(-\gamma\Delta)\left(\sup_{t_0+(m-1)\Delta\leq t\leq t_0+m\Delta} \mathbb{E}\left(\|\bar{u}(t,x;t_0,\phi)\|_2^2\right.\right. \\
&\qquad \left.\left. + \|\bar{v}(t,x;t_0,\psi)\|_2^2\right)\right) \\
&\qquad \vdots \\
&\leq \exp(-\gamma m\Delta)\left(\sup_{t_0\leq t\leq t_0+\Delta} \mathbb{E}\left(\|\bar{u}(t,x;t_0,\phi)\|_2^2\right.\right. \\
&\qquad \left.\left. + \|\bar{v}(t,x;t_0,\psi)\|_2^2\right)\right).
\end{aligned} \tag{57}$$

Hence, for any $t \geq t_0 + \Delta$, there exists a positive integer m, such that $t_0 + m\Delta \leq t \leq t_0 + (m + 1)\Delta$, and we have

$$\begin{aligned}
&\mathbb{E}\left(\|\bar{u}(t,x;t_0,\phi)\|_2^2+\|\bar{v}(t,x;t_0,\psi)\|_2^2\right)\\
&\le \exp(-\gamma m\Delta)\\
&\quad\times\left(\sup_{t_0\le t\le t_0+\Delta}\mathbb{E}\left(\|\bar{u}(t,x;t_0,\phi)\|_2^2+\|\bar{v}(t,x;t_0,\psi)\|_2^2\right)\right)\\
&\le \exp\{-\gamma t+\gamma t_0+\gamma\Delta\}\\
&\quad\times\left(\sup_{t_0\le t\le t_0+\Delta}\mathbb{E}\left(\|\bar{u}(t,x;t_0,\phi)\|_2^2+\|\bar{v}(t,x;t_0,\psi)\|_2^2\right)\right)\\
&\le C_3\exp\{\gamma\Delta\}\exp\{-\gamma(t-t_0)\},
\end{aligned}\tag{58}$$

where $C_3 = \sup_{t_0\le t\le t_0+\Delta}\mathbb{E}(\|\bar{u}(t,x;t_0,\phi)\|_2^2+\|\bar{v}(t,x;t_0,\psi)\|_2^2)$. The above inequality also holds for $t_0 \le t \le t_0+\Delta$.

Therefore, the neural networks (16) are mean square globally exponentially stable, and by Theorem 4, the neural networks (16) are also almost surely globally exponentially stable. □

4. Connection Weight Matrices Uncertainty and Noise Impact on Stability

In this section, we first consider the parameter uncertainty intensity which is added to the self-feedback matrix $(A,B)^T$ of the neural networks (16). Then, the neural networks (16) are changed as

$$\begin{aligned}
d\bar{u}_i(t,x) = &\left\{\sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(D_{ik}(r(t))\frac{\partial\bar{u}_i(t,x)}{\partial x_k}\right)\right.\\
&-(1+\lambda)a_i(r(t))\bar{u}_i(t,x)\\
&\left.+\sum_{j=1}^{n}c_{ji}(r(t))f_j(\bar{v}_j(t,x))\right\}dt\\
&+\sigma\bar{u}_i(t,x)\,dW(t)\\
d\bar{v}_j(t,x) = &\left\{\sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(D^*_{jk}(r(t))\frac{\partial\bar{v}_j(t,x)}{\partial x_k}\right)\right.\\
&-(1+\lambda)b_j(r(t))\bar{v}_j(t,x)\\
&\left.+\sum_{i=1}^{m}e_{ij}(r(t))g_i(\bar{u}_i(t,x))\right\}dt\\
&+\sigma\bar{v}_j(t,x)\,dW(t).
\end{aligned}\tag{59}$$

The initial conditions and boundary conditions are given by

$$\begin{aligned}
&\bar{u}_i(t_0,x)=\phi_i(x),\quad x\in\Omega_0,\ t_0\in\mathbb{R}_+,\ i=1,2,\ldots,m,\\
&\bar{v}_j(t_0,x)=\psi_j(x),\quad x\in\Omega_0,\ t_0\in\mathbb{R}_+,\ j=1,2,\ldots,n,\\
&\left.\frac{\partial\bar{u}_i(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial\bar{u}_i(t,x)}{\partial x_1},\ldots,\frac{\partial\bar{u}_i(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad(t,x)\in[t_0,+\infty)\times\partial\Omega_0,\quad i=1,2,\ldots,m,\\
&\left.\frac{\partial\bar{v}_j(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial\bar{v}_j(t,x)}{\partial x_1},\ldots,\frac{\partial\bar{v}_j(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad(t,x)\in[t_0,+\infty)\times\partial\Omega_0,\quad j=1,2,\ldots,n,
\end{aligned}\tag{60}$$

where λ is the self-feedback matrix $(A,B)^T$ uncertainty intensity and σ is the noise intensity.

We rewrite (59) as follows:

$$\begin{aligned}
d\bar{u}(t,x) = &\{\nabla\cdot(D(r(t))\circ\nabla\bar{u})-(1+\lambda)A(r(t))\bar{u}(t,x)\\
&+C(r(t))f(\bar{v}(t,x))\}dt+\sigma\bar{u}(t,x)\,dW(t)\\
d\bar{v}(t,x) = &\{\nabla\cdot(D^*(r(t))\circ\nabla\bar{v})-(1+\lambda)B(r(t))\bar{v}(t,x)\\
&+E(r(t))g(\bar{u}(t,x))\}dt+\sigma\bar{v}(t,x)\,dW(t).
\end{aligned}\tag{61}$$

For the global exponential stability of neural networks (6), we will characterize how much the intensity of both the self-feedback matrix $(A,B)^T$ uncertainty and stochastic noise the stochastic neural networks (59) can tolerate while maintaining global exponential stability.

Theorem 6. *Let Assumption (H1) hold and let the neural networks* (6) *be globally exponentially stable. Then, the neural networks* (59) *are mean square globally exponential stability and also almost sure globally exponential stability, if there exists* $\mu_q > 0, (q\in\mathbb{S})$, *and* (λ,σ) *is in the inner of the closed curve described by the following transcendental equation:*

$$\begin{aligned}
&\frac{4\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\widehat{A}\|+\|\widehat{B}\|\right)\right)\alpha}{\beta}\\
&\quad\times\exp\left\{\frac{2\Delta\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\check{\mu}}\right\},\\
&+2\alpha\exp\{-\beta\Delta\}=1,
\end{aligned}\tag{62}$$

$$\Delta > \frac{\ln(2\alpha)}{\beta} > 0,\tag{63}$$

where $C_4 = [(3+2\lambda^2)(\|\widehat{A}\|+\|\widehat{B}\|)+(1+K^2)\|\widehat{C}\|+(1+L^2)\|\widehat{E}\|+2\sigma^2]$, $\|\widehat{A}\| = \max_{p\in\mathbb{S}}\|A(p)\|$, *and so forth and* $\hat{\mu} = \max_{p\in\mathbb{S}}\mu_p$ *and* $\check{\mu} = \min_{p\in\mathbb{S}}\mu_p$.

Proof. For any $(\phi(x),\psi(x))$, we denote the state $(\bar{u}(t,x;t_0,\phi),\bar{v}(t,x;t_0,\psi))$ of (59) as $(\bar{u}(t,x),\bar{v}(t,x))$ and the state $(u(t,x;t_0,\phi),v(t,x;t_0,\psi))$ of (6) as $(u(t,x),v(t,x))$.

From (6) and (61) and stochastic Fubini's Theorem, we have

$$\begin{aligned}
&\int_{\Omega_0}(u(t,x)-\bar{u}(t,x))\,dx+\int_{\Omega_0}(v(t,x)-\bar{v}(t,x))\,dx\\
&=\int_{t_0}^{t}\int_{\Omega_0}\nabla\cdot(D(r(s))\circ\nabla(u-\bar{u}))\,dx\,ds
\end{aligned}$$

$$
\begin{aligned}
&+\int_{t_0}^{t}\int_{\Omega_0}\left[-A(r(s))(u(s,x)-\bar{u}(s,x))\right.\\
&\qquad\left.+C(r(s))\left(f(v(s,x))-f(\bar{v}(s,x))\right)\right]\mathrm{d}x\mathrm{d}s\\
&-\int_{t_0}^{t}\int_{\Omega_0}\sigma\bar{u}(s,x)\,\mathrm{d}x\mathrm{d}W(s)\\
&+\int_{t_0}^{t}\int_{\Omega_0}\lambda A(r(s))\bar{u}(s,x)\,\mathrm{d}x\mathrm{d}s\\
&+\int_{t_0}^{t}\int_{\Omega_0}\nabla\cdot\left(D^*(r(s))\circ\nabla(v-\bar{v})\right)\mathrm{d}x\mathrm{d}s\\
&+\int_{t_0}^{t}\int_{\Omega_0}\left[-B(r(s))(v(s,x)-\bar{v}(s,x))\right.\\
&\qquad\left.+E(r(s))\left(g(u(s,x))-g(\bar{u}(s,x))\right)\right]\mathrm{d}x\mathrm{d}s\\
&-\int_{t_0}^{t}\int_{\Omega_0}\sigma\bar{v}(s,x)\,\mathrm{d}x\mathrm{d}W(s)\\
&+\int_{t_0}^{t}\int_{\Omega_0}\lambda B(r(s))\bar{v}(s,x)\,\mathrm{d}x\mathrm{d}s.
\end{aligned}
\tag{64}
$$

Construct the average Lyapunov functional

$$
\begin{aligned}
&V(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t))\\
&\quad=\int_{\Omega_0}\mu_{r(t)}\left[|u(t,x)-\bar{u}(t,x)|^2+|v(t,x)-\bar{v}(t,x)|^2\right]\mathrm{d}x,
\end{aligned}
\tag{65}
$$

where $\mu_{r(t)}>0$.

By applying generalized Itô formula [27], we have

$$
\begin{aligned}
&\mathrm{d}V\left(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),p\right)\big|_{(21)}\\
&=\int_{\Omega_0}2\mu_p(u(t,x)-\bar{u}(t,x))^T\left(\nabla\cdot\left(D(p)\circ\nabla(u-\bar{u})\right)\right)\mathrm{d}x\mathrm{d}t\\
&\quad+\int_{\Omega_0}2\mu_p(u(t,x)-\bar{u}(t,x))^T\\
&\qquad\times\left[-A(p)(u(t,x)-\bar{u}(t,x))\right.\\
&\qquad\quad+C(p)\left(f(v(t,x))-f(\bar{v}(t,x))\right)\\
&\qquad\quad\left.+\lambda A(p)\bar{u}(t,x)\right]\mathrm{d}x\mathrm{d}t\\
&\quad+\int_{\Omega_0}\sigma^2\mu_p|\bar{u}(t,x)|^2\mathrm{d}x\mathrm{d}t\\
&\quad-2\int_{\Omega_0}\sigma(u(t,x)-\bar{u}(t,x))^T\bar{u}(t,x)\,\mathrm{d}x\mathrm{d}W(t)\\
&\quad+\int_{\Omega_0}2\mu_p(v(t,x)-\bar{v}(t,x))^T\\
&\qquad\times\left(\nabla\cdot\left(D^*(p)\circ\nabla(v-\bar{v})\right)\right)\mathrm{d}x\mathrm{d}t\\
&\quad+\int_{\Omega_0}2\mu_p(v(t,x)-\bar{v}(t,x))^T\\
&\qquad\times\left[-B(p)(v(t,x)-\bar{v}(t,x))\right.\\
&\qquad\quad+E(p)\left(g(u(t,x))-g(\bar{u}(t,x))\right)\\
&\qquad\quad\left.+\lambda B(p)\bar{v}(t,x)\right]\mathrm{d}x\mathrm{d}t\\
&\quad+\int_{\Omega_0}\sigma^2\mu_p|\bar{v}(t,x)|^2\mathrm{d}x\mathrm{d}t\\
&\quad-2\int_{\Omega_0}\sigma(v(t,x)-\bar{v}(t,x))^T\bar{v}(t,x)\,\mathrm{d}x\mathrm{d}W(t)\\
&\quad+\sum_{q=1}^{N}\gamma_{pq}\mu_q\int_{\Omega_0}\left[|u(t,x)-\bar{u}(t,x)|^2\right.\\
&\qquad\left.+|v(t,x)-\bar{v}(t,x)|^2\right]\mathrm{d}x.
\end{aligned}
\tag{66}
$$

By Hölder's inequality, we get

$$
\begin{aligned}
&\int_{\Omega_0}2\mu_p(u(t,x)-\bar{u}(t,x))^T\lambda A(p)\bar{u}(t,x)\,\mathrm{d}x\\
&\le\max_{p\in\mathbb{S}}\left\|\mu_pA(p)\right\|\left[\int_{\Omega_0}|u(t,x)-\bar{u}(t,x)|^2\mathrm{d}x\right.\\
&\qquad\left.+\lambda^2\int_{\Omega_0}|\bar{u}(t,x)|^2\mathrm{d}x\right]\\
&=\max_{p\in\mathbb{S}}\left\|\mu_pA(p)\right\|\left[\int_{\Omega_0}|u(t,x)-\bar{u}(t,x)|^2\mathrm{d}x\right.\\
&\qquad\left.+\lambda^2\int_{\Omega_0}|u(t,x)-\bar{u}(t,x)-u(t,x)|^2\mathrm{d}x\right]\\
&\le\max_{p\in\mathbb{S}}\left\|\mu_pA(p)\right\|\left[\left(1+2\lambda^2\right)\int_{\Omega_0}|u(t,x)-\bar{u}(t,x)|^2\mathrm{d}x\right.\\
&\qquad\left.+2\lambda^2\int_{\Omega_0}|u(t,x)|^2\mathrm{d}x\right],\\
&\int_{\Omega_0}2\mu_p(v(t,x)-\bar{v}(t,x))^T\lambda B(p)\bar{v}(t,x)\,\mathrm{d}x\\
&\le\max_{p\in\mathbb{S}}\left\|\mu_pB(p)\right\|\left[\int_{\Omega_0}|v(t,x)-\bar{v}(t,x)|^2\mathrm{d}x\right.\\
&\qquad\left.+\lambda^2\int_{\Omega_0}|\bar{v}(t,x)|^2\mathrm{d}x\right]\\
&=\max_{p\in\mathbb{S}}\left\|\mu_pB(p)\right\|\left[\int_{\Omega_0}|v(t,x)-\bar{v}(t,x)|^2\mathrm{d}x\right.\\
&\qquad\left.+\lambda^2\int_{\Omega_0}|v(t,x)-\bar{v}(t,x)-v(t,x)|^2\mathrm{d}x\right]\\
&\le\max_{p\in\mathbb{S}}\left\|\mu_pB(p)\right\|\left[\left(1+2\lambda^2\right)\int_{\Omega_0}|v(t,x)-\bar{v}(t,x)|^2\mathrm{d}x\right.\\
&\qquad\left.+2\lambda^2\int_{\Omega_0}|v(t,x)|^2\mathrm{d}x\right].
\end{aligned}
\tag{67}
$$

From (42), (43), and (67) and Assumption (H1), we obtain that

$$
\begin{aligned}
&\mathrm{d}V\left(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),p\right)\\
&\le\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)\\
&\quad\times\int_{\Omega_0}\left(|u(t,x)-\bar{u}(t,x)|^2\right.\\
&\qquad\left.+|v(t,x)-\bar{v}(t,x)|^2\right)\mathrm{d}x\mathrm{d}t\\
&\quad+2\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\hat{A}\|+\|\hat{B}\|\right)\right)\\
&\quad\times\int_{\Omega_0}\left(|u(t,x)|^2+|v(t,x)|^2\right)\mathrm{d}x\mathrm{d}t\\
&\quad-2\int_{\Omega_0}\sigma(u(t,x)-\bar{u}(t,x))^T\bar{u}(t,x)\,\mathrm{d}x\mathrm{d}W(t)\\
&\quad-2\int_{\Omega_0}\sigma(v(t,x)-\bar{v}(t,x))^T\bar{v}(t,x)\,\mathrm{d}x\mathrm{d}W(t).
\end{aligned}\tag{68}
$$

When $t\le t_0+2\Delta$, we have

$$
\begin{aligned}
&\mathbb{E}V\left(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t)\right)\\
&\le\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)\\
&\quad\times\int_{t_0}^{t}\mathbb{E}\int_{\Omega_0}\left(|u(s,x)-\bar{u}(s,x)|^2\right.\\
&\qquad\left.+|v(s,x)-\bar{v}(s,x)|^2\right)\mathrm{d}x\mathrm{d}s\\
&\quad+2\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\hat{A}\|+\|\hat{B}\|\right)\right)\\
&\quad\times\int_{t_0}^{t}\alpha\left(\|\phi\|_2^2+\|\psi\|_2^2\right)\exp\left(-\beta(s-t_0)\right)\mathrm{d}s\\
&\quad-2\sigma\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(u(s,x)-\bar{u}(s,x))^T\bar{u}(s,x)\,\mathrm{d}x\mathrm{d}W(s)\\
&\quad-2\sigma\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(v(s,x)-\bar{v}(s,x))^T\bar{v}(s,x)\,\mathrm{d}x\mathrm{d}W(s).
\end{aligned}\tag{69}
$$

By stochastic Fubini's Theorem, we have

$$
\begin{aligned}
&\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(u(s,x)-\bar{u}(s,x))^T\bar{u}(s,x)\,\mathrm{d}x\mathrm{d}W(s)=0,\\
&\mathbb{E}\int_{t_0}^{t}\int_{\Omega_0}(v(s,x)-\bar{v}(s,x))^T\bar{v}(s,x)\,\mathrm{d}x\mathrm{d}W(s)=0.
\end{aligned}\tag{70}
$$

By (69), one get

$$
\begin{aligned}
&\mathbb{E}V\left(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t)\right)\\
&\le\frac{\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\\
&\quad\times\int_{t_0}^{t}\mathbb{E}V\left(u(s,x),v(s,x),\bar{u}(s,x),\bar{v}(s,x),r(s)\right)ds\\
&\quad+\frac{2\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\hat{A}\|+\|\hat{B}\|\right)\right)\alpha\left(\|\phi\|_2^2+\|\psi\|_2^2\right)}{\beta}.
\end{aligned}\tag{71}
$$

When $t_0+\Delta\le t\le t_0+2\Delta$, by applying Gronwall's inequality, we have

$$
\begin{aligned}
&\mathbb{E}\left(\|u(t,x)-\bar{u}(t,x)\|_2^2+\|v(t,x)-\bar{v}(t,x)\|_2^2\right)\\
&\quad=\mathbb{E}V\left(u(t,x),v(t,x),\bar{u}(t,x),\bar{v}(t,x),r(t)\right)\\
&\quad\le\frac{2\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\hat{A}\|+\|\hat{B}\|\right)\right)\alpha\left(\|\phi\|_2^2+\|\psi\|_2^2\right)}{\beta}\\
&\qquad\times\exp\frac{\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}(t-t_0)\\
&\quad\le\sup_{t_0\le t\le t_0+\Delta}\mathbb{E}\left(\|\bar{u}(t,x)\|^2+\|\bar{v}(t,x)\|^2\right)\\
&\qquad\times\frac{2\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\hat{A}\|+\|\hat{B}\|\right)\right)\alpha}{\beta}\\
&\qquad\times\exp\frac{2\Delta\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}.
\end{aligned}\tag{72}
$$

By the global exponential stability of (6), we have

$$
\begin{aligned}
&\mathbb{E}\left(\|\bar{u}(t,x)\|_2^2+\|\bar{v}(t,x)\|_2^2\right)\\
&\le2\mathbb{E}\left(\|u(t,x)-\bar{u}(t,x)\|_2^2+\|v(t,x)-\bar{v}(t,x)\|_2^2\right)\\
&\quad+2\mathbb{E}\left(\|u(t,x)\|_2^2+\|v(t,x)\|_2^2\right)\\
&\le\sup_{t_0\le t\le t_0+\Delta}\mathbb{E}\left(\|\bar{u}(t,x)\|^2+\|\bar{v}(t,x)\|^2\right)\\
&\quad\times\frac{4\hat{\mu}\left(\sigma^2+\lambda^2\left(\|\hat{A}\|+\|\hat{B}\|\right)\right)\alpha}{\beta}\\
&\quad\times\exp\frac{2\Delta\left(\hat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\\
&\quad+2\alpha\left(\|\phi\|_2^2+\|\psi\|_2^2\right)\exp\left\{-\beta(t-t_0)\right\}.
\end{aligned}\tag{73}
$$

Moreover,

$$\begin{aligned}
&\mathbb{E}\left(\|\bar{u}(t,x)\|_2^2+\|\bar{v}(t,x)\|_2^2\right)\\
&\leq\left\{\frac{4\widehat{\mu}\left(\sigma^2+\lambda^2\left(\|\widehat{A}\|+\|\widehat{B}\|\right)\right)\alpha}{\beta}\right.\\
&\qquad\times\exp\left\{\frac{2\Delta\left(\widehat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\}\\
&\qquad\left.+2\alpha\exp\{-\beta\Delta\}\right\}\\
&\quad\times\sup_{t_0\leq t\leq t_0+\Delta}\mathbb{E}\left(\|\bar{u}(t,x)\|^2+\|\bar{v}(t,x)\|^2\right).
\end{aligned}\tag{74}$$

From (62), when (λ,σ) is in the inner of the closed curve described by the transcendental equation, we have

$$\begin{aligned}
&\frac{4\widehat{\mu}\left(\sigma^2+\lambda^2\left(\|\widehat{A}\|+\|\widehat{B}\|\right)\right)\alpha}{\beta}\\
&\quad\times\exp\left\{\frac{2\Delta\left(\widehat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\}\\
&\quad+2\alpha\exp\{-\beta\Delta\}<1.
\end{aligned}\tag{75}$$

Let

$$\begin{aligned}
\gamma=&\left(-\log\left\{\frac{4\widehat{\mu}\left(\sigma^2+\lambda^2\left(\|\widehat{A}\|+\|\widehat{B}\|\right)\right)\alpha}{\beta}\right.\right.\\
&\qquad\times\exp\left\{\frac{2\Delta\left(\widehat{\mu}C_4+\max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\}\\
&\qquad\left.\left.+2\alpha\exp\{-\beta\Delta\}\right\}\right)(\Delta)^{-1}>0.
\end{aligned}\tag{76}$$

By (74), we have

$$\begin{aligned}
&\sup_{t_0+\Delta\leq t\leq t_0+2\Delta}\mathbb{E}\left(\|\bar{u}(t,x)\|_2^2+\|\bar{v}(t,x)\|_2^2\right)\\
&\quad\leq\exp(-\gamma\Delta)\left(\sup_{t_0\leq t\leq t_0+\Delta}\mathbb{E}\left(\|\bar{u}(t,x)\|_2^2+\|\bar{v}(t,x)\|_2^2\right)\right).
\end{aligned}\tag{77}$$

Similar to the proof of Theorem 5, we can prove that the neural networks (59) are mean square globally exponentially stable and also almost surely globally exponentially stable. □

To continue, we consider the parameter uncertainty intensity which is added to the connection weight matrix $(C,E)^T$ of the neural networks (16). Then, the neural networks (16) are changed as

$$\begin{aligned}
d\bar{u}_i(t,x)=&\left\{\sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(D_{ik}(r(t))\frac{\partial\bar{u}_i(t,x)}{\partial x_k}\right)\right.\\
&-a_i(r(t))\bar{u}_i(t,x)\\
&\left.+\sum_{j=1}^{n}(1+\delta)c_{ji}(r(t))f_j(\bar{v}_j(t,x))\right\}dt\\
&+\sigma\bar{u}_i(t,x)\,dW(t),\\
d\bar{v}_j(t,x)=&\left\{\sum_{k=1}^{l}\frac{\partial}{\partial x_k}\left(D^*_{jk}(r(t))\frac{\partial\bar{v}_j(t,x)}{\partial x_k}\right)\right.\\
&-b_j(r(t))\bar{v}_j(t,x)\\
&\left.+\sum_{i=1}^{m}(1+\delta)e_{ij}(r(t))g_i(\bar{u}_i(t,x))\right\}dt\\
&+\sigma\bar{v}_j(t,x)\,dW(t).
\end{aligned}\tag{78}$$

The initial conditions and boundary conditions are given by

$$\begin{aligned}
&\bar{u}_i(t_0,x)=\phi_i(x),\quad x\in\Omega_0,\ t_0\in\mathbb{R}_+,\ i=1,2,\ldots,m,\\
&\bar{v}_j(t_0,x)=\psi_j(x),\quad x\in\Omega_0,\ t_0\in\mathbb{R}_+,\ j=1,2,\ldots,n,\\
&\left.\frac{\partial\bar{u}_i(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial\bar{u}_i(t,x)}{\partial x_1},\ldots,\frac{\partial\bar{u}_i(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad(t,x)\in[t_0,+\infty)\times\partial\Omega_0,\quad i=1,2,\ldots,m,\\
&\left.\frac{\partial\bar{v}_j(t,x)}{\partial\vec{n}}\right|_{\partial\Omega_0}=\left(\frac{\partial\bar{v}_j(t,x)}{\partial x_1},\ldots,\frac{\partial\bar{v}_j(t,x)}{\partial x_l}\right)^T=0,\\
&\qquad(t,x)\in[t_0,+\infty)\times\partial\Omega_0,\quad j=1,2,\ldots,n,
\end{aligned}\tag{79}$$

where δ is the connection weight matrix $(C,E)^T$ uncertainty intensity and σ is the noise intensity.

We rewrite (78) as follows:

$$\begin{aligned}
d\bar{u}(t,x)=&\left\{\nabla\cdot(D(r(t))\circ\nabla\bar{u})-A(r(t))\bar{u}(t,x)\right.\\
&\left.+(1+\delta)C(r(t))f(\bar{v}(t,x))\right\}dt
\end{aligned}$$

$$+ \sigma \bar{u}(t,x)\,\mathrm{d}W(t),$$
$$\mathrm{d}\bar{v}(t,x) = \{\nabla \cdot (D^*(r(t)) \circ \nabla \bar{v}) - B(r(t))\bar{v}(t,x) + (1+\delta)E(r(t))g(\bar{u}(t,x))\}\,\mathrm{d}t + \sigma \bar{v}(t,x)\,\mathrm{d}W(t). \tag{80}$$

For the global exponential stability of neural networks (6), we will characterize how much the intensity of both the connection weight matrix $(C,E)^T$ uncertainty and stochastic noise the stochastic neural networks (78) can tolerate while maintaining global exponential stability.

Theorem 7. *Let Assumption (H1) hold and let the neural networks* (6) *be global exponential stability. Then, the neural networks* (78) *are mean square globally exponentially stable and also almost surely globally exponentially stable, if there exists $\mu_q > 0$, $(q \in \mathbb{S})$, and (δ,σ) is in the inner of the closed curve described by the following transcendental equation:*

$$\frac{4\hat{\mu}\left(\sigma^2 + \delta^2\left(K^2\|\widehat{C}\| + L^2\|\widehat{E}\|\right)\right)\alpha}{\beta} \times \exp\left\{\frac{2\Delta\left(\hat{\mu}C_5 + \max_{p\in\mathbb{S}}\sum_{q=1}^{N}\gamma_{pq}\mu_q\right)}{\breve{\mu}}\right\} + 2\alpha\exp\{-\beta\Delta\} = 1,$$
$$\Delta > \frac{\ln(2\alpha)}{\beta} > 0, \tag{81}$$

where $C_5 = [2(\|\widehat{A}\| + \|\widehat{B}\|) + (2 + (1+2\delta^2)K^2)\|\widehat{C}\| + (2 + (1 + 2\delta^2)L^2)\|\widehat{E}\| + 2\sigma^2]$, $\|\widehat{A}\| = \max_{p\in\mathbb{S}}\|A(p)\|$, *and so forth and* $\hat{\mu} = \max_{p\in\mathbb{S}}\mu_p$ *and* $\breve{\mu} = \min_{p\in\mathbb{S}}\mu_p$.

The proof is similar to the proof of Theorem 6.

5. Illustrate Example

Example 1. Consider hybrid BAM neural networks with reaction diffusion terms

$$\frac{\partial u(t,x)}{\partial t} = D(r(t))\frac{\partial^2 u(t,x)}{\partial x^2} - a(r(t))u(t,x) + c(r(t))f(v(t,x)),$$
$$\frac{\partial v(t,x)}{\partial t} = D_k^*(r(t))\frac{\partial^2 v(t,x)}{\partial x^2} - b(r(t))v(t,x) + e(r(t))g(u(t,x)). \tag{82}$$

The initial conditions and boundary conditions are given by

$$u(0,x) = \sin(x), \quad x \in [-5,5],$$
$$v(0,x) = \cos(x) - 1, \quad x \in [-5,5],$$
$$\left.\frac{\partial u(t,x)}{\partial \vec{n}}\right|_{\partial\Omega_0} = \frac{\partial u(t,x)}{\partial x} = 0, \quad (t,x) \in [t_0,+\infty) \times \partial\Omega_0,$$
$$\left.\frac{\partial v(t,x)}{\partial \vec{n}}\right|_{\partial\Omega_0} = \frac{\partial v(t,x)}{\partial x} = 0, \quad (t,x) \in [t_0,+\infty) \times \partial\Omega_0, \tag{83}$$

where

$$\Gamma = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix},$$
$$\begin{pmatrix} D(1) & 0 \\ 0 & D^*(1) \end{pmatrix} = \begin{pmatrix} 0.002 & 0 \\ 0 & 0.003 \end{pmatrix},$$
$$\begin{pmatrix} D(2) & 0 \\ 0 & D^*(2) \end{pmatrix} = \begin{pmatrix} 0.001 & 0 \\ 0 & 0.002 \end{pmatrix},$$
$$\begin{pmatrix} a(1) & 0 \\ 0 & b(1) \end{pmatrix} = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.3 \end{pmatrix}, \tag{84}$$
$$\begin{pmatrix} 0 & c(1) \\ e(1) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0.2 \\ 0.2 & 0 \end{pmatrix},$$
$$\begin{pmatrix} a(2) & 0 \\ 0 & b(2) \end{pmatrix} = \begin{pmatrix} 0.3 & 0 \\ 0 & 0.2 \end{pmatrix},$$
$$\begin{pmatrix} 0 & c(2) \\ e(2) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0.2 \\ 0.2 & 0 \end{pmatrix},$$

and $f(v) = \sin(v)$, $g(u) = (1/2)(|u+1| - |u-1|)$, and $K = L = 1$. According to Theorem 1 in [9] and Theorem 1 in [29], the neural networks (82) are global exponential stability with $\alpha = 1$ and $\beta = 1$.

In the presence of stochastic noise and self-feedback matrix $(A,B)^T$ uncertainty, the neural networks (82) become

$$\mathrm{d}u(t,x) = \left\{D(r(t))\frac{\partial^2 u(t,x)}{\partial x^2} - (a(r(t)) + \lambda)u(t,x) + c(r(t))f(v(t,x))\right\}\mathrm{d}t + \sigma u(t,x)\,\mathrm{d}W(t),$$
$$\mathrm{d}v(t,x) = \left\{D^*(r(t))\frac{\partial^2 v(t,x)}{\partial x^2} - (b(r(t)) + \lambda)v(t,x) + e(r(t))g(u(t,x))\right\}\mathrm{d}t + \sigma v(t,x)\,\mathrm{d}W(t). \tag{85}$$

According to Theorem 6, let $\Delta = 0.7 > \log(2\alpha)/\beta = 0.6931$ and $\mu_1 = 1$ and $\mu_2 = 2$. From (62), we have

$$8\left(\sigma^2 + 0.6\lambda^2\right)\exp\left\{10.08 + 3.36\lambda^2 + 5.6\sigma^2\right\} + 2\exp\{-0.7\} = 1. \tag{86}$$

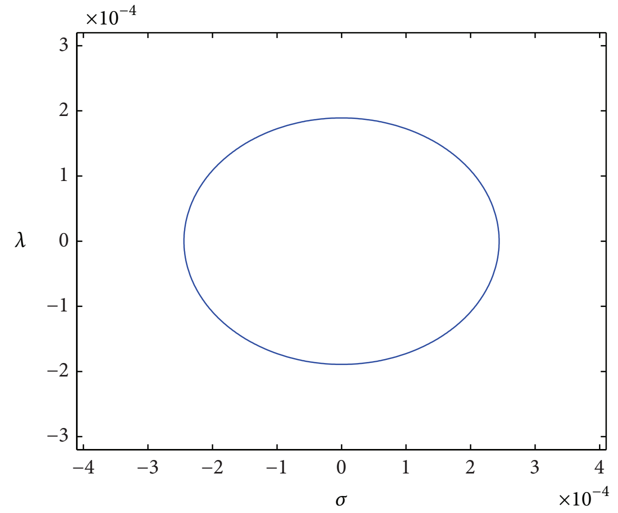


FIGURE 1: The stability region with (λ, σ) of the neural networks (85).

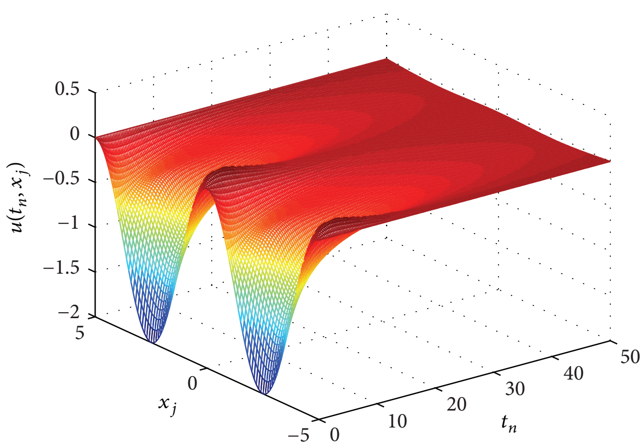


FIGURE 2: Surface curve of $u(t, x)$ of the neural networks (85) in model 1.

Then, we can obtain its closed curve for (λ, σ). Figure 1 depicts the stability region for (λ, σ) in (85).

Figures 2, 3, 4, and 5 depict the surface curves of the neural networks (85) with $(\lambda, \sigma) = (-10^{-4}, 10^{-4})$. It shows that the state of the neural networks (85) is mean square globally exponential stability and almost surely globally exponential stable, as the parameter (λ, σ) in the inner of the curve of Figure 1.

Figures 6, 7, 8 and 9 show the surface curves of the neural networks (85) with $(\lambda, \sigma) = (-0.1, 1.5)$. It shows that when the conditions in Theorem 6 do not hold, the neural networks (85) become unstable.

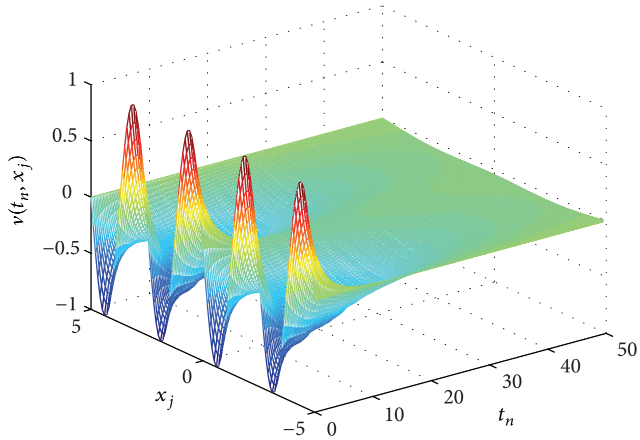


FIGURE 3: Surface curve of $v(t, x)$ of the neural networks (85) in model 1.

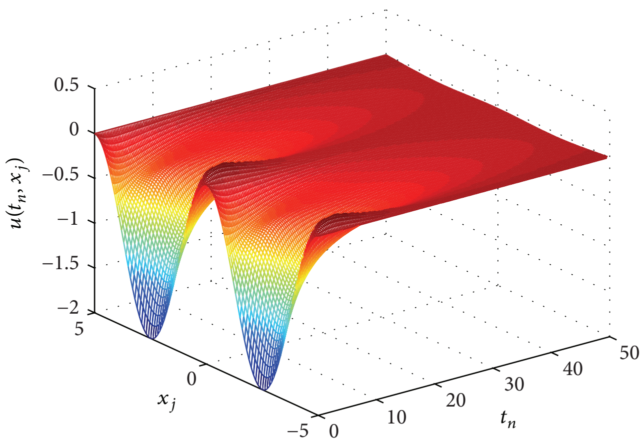


FIGURE 4: Surface curve of $u(t, x)$ of the neural networks (85) in model 2.

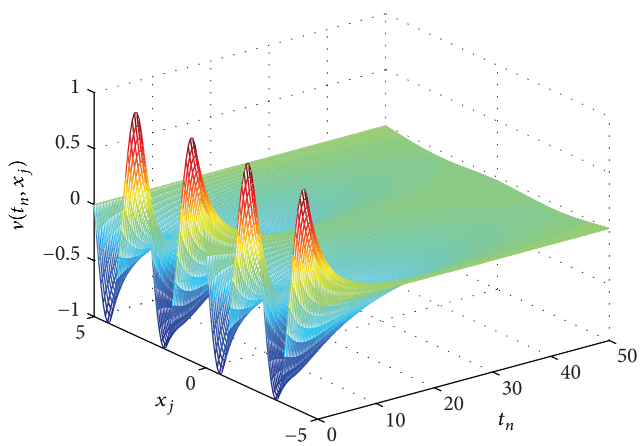


FIGURE 5: Surface curve of $v(t, x)$ of the neural networks (85) in model 2.

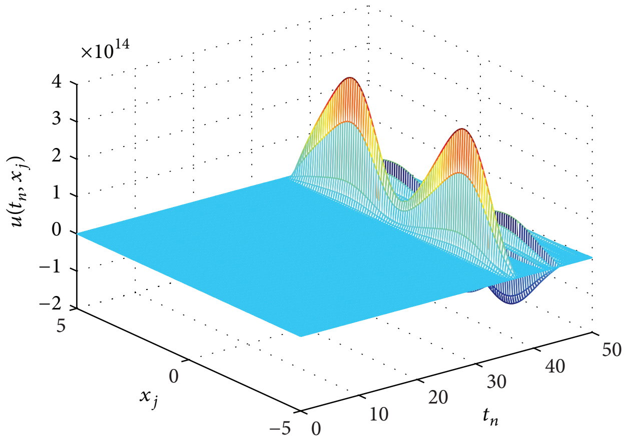


FIGURE 6: Surface curve of $u(t, x)$ of the neural networks (82) in model 1.

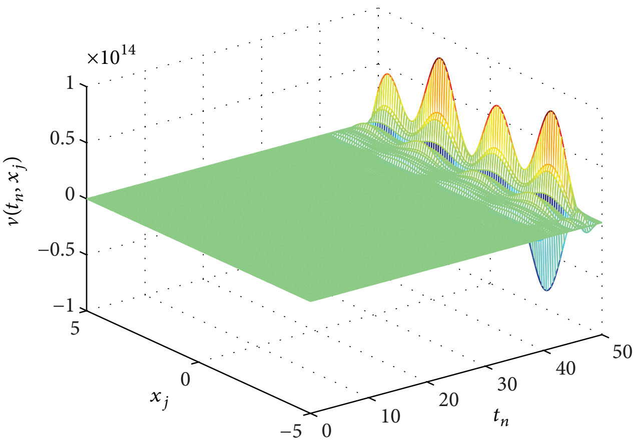


FIGURE 7: Surface curve of $v(t, x)$ of the neural networks (82) in model 1.

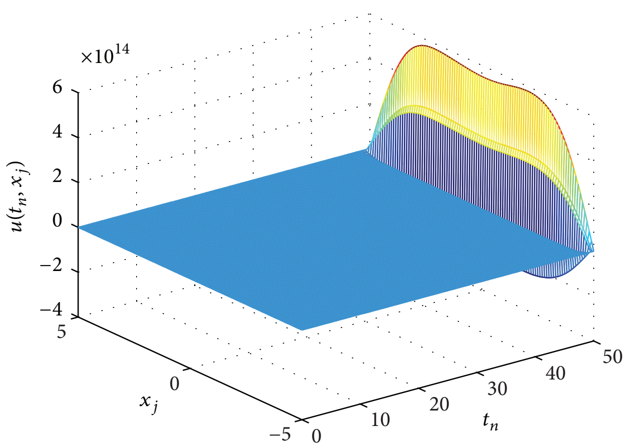


FIGURE 8: Surface curve of $u(t, x)$ of the neural networks (82) in model 2.

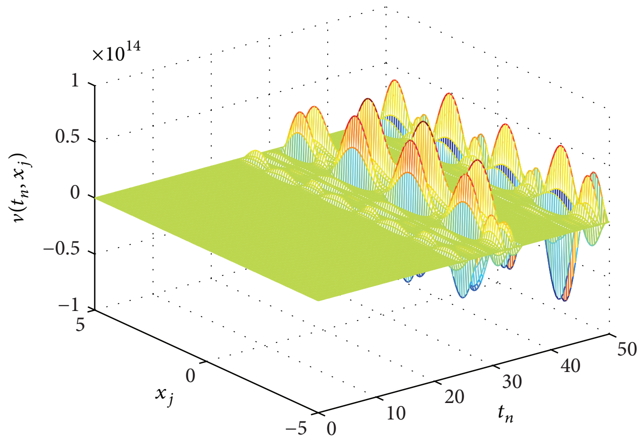


FIGURE 9: Surface curve of $v(t, x)$ of the neural networks (82) in model 2.

Conflict of Interests

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

Acknowledgments

This paper is partially supported by National Natural Science Foundation of China under Grants nos. 61304068, 61374085, and 61134012 and the Fundamental Research Funds for the Central Universities under Grant no. 2013QC019.

References

[1] B. Kosko, "Adaptive bidirectional associative memories," *Applied Optics*, vol. 26, no. 23, pp. 4947–4960, 1987.

[2] B. Kosko, "Bidirectional associative memories," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 49–60, 1988.

[3] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs, NJ, USA, 1992.

[4] J. Cao and L. Wang, "Exponential stability and periodic oscillatory solution in BAM networks with delays," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 457–463, 2002.

[5] A. Chen, J. Cao, and L. Huang, "Exponential stability of BAM neural networks with transmission delays," *Neurocomputing*, vol. 57, no. 1–4, pp. 435–454, 2004.

[6] P. Balasubramaniam and C. Vidhya, "Global asymptotic stability of stochastic BAM neural networks with distributed delays and reaction-diffusion terms," *Journal of Computational and Applied Mathematics*, vol. 234, no. 12, pp. 3458–3466, 2010.

[7] B. T. Cui and X. Y. Lou, "Global asymptotic stability of BAM neural networks with distributed delays and reaction-diffusion terms," *Chaos, Solitons & Fractals*, vol. 27, no. 5, pp. 1347–1354, 2006.

[8] X. Lou, B. Cui, and W. Wu, "On global exponential stability and existence of periodic solutions for BAM neural networks with distributed delays and reaction-diffusion terms," *Chaos, Solitons & Fractals*, vol. 36, no. 4, pp. 1044–1054, 2008.

[9] Q. Song and J. Cao, "Global exponential stability and existence of periodic solutions in BAM networks with delays and reaction-diffusion terms," *Chaos, Solitons & Fractals*, vol. 23, no. 2, pp. 421–430, 2005.

[10] Q. Song, Z. Zhao, and Y. Li, "Global exponential stability of BAM neural networks with distributed delays and reaction-diffusion terms," *Physics Letters A*, vol. 335, no. 2-3, pp. 213–225, 2005.

[11] Q. Zhu, X. Li, and X. Yang, "Exponential stability for stochastic reaction-diffusion BAM neural networks with time-varying and distributed delays," *Applied Mathematics and Computation*, vol. 217, no. 13, pp. 6078–6091, 2011.

[12] W. J. Anderson, *Continuous-Time Markov Chains: An Applications-Oriented Approach*, Springer Series in Statistics: Probability and Its Applications, Springer, New York, NY, USA, 1991.

[13] M. Athans, "Command-and-control (C2) theory: a challenge to control science," *IEEE Transactions on Automatic Control*, vol. 32, no. 4, pp. 286–293, 1987.

[14] M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Dekker, New York, NY, USA, 1990.

[15] H. Huang, D. W. C. Ho, and Y. Qu, "Robust stability of stochastic delayed additive neural networks with Markovian switching," *Neural Networks*, vol. 20, no. 7, pp. 799–809, 2007.

[16] Y. Shen and J. Wang, "Noise-induced stabilization of the recurrent neural networks with mixed time-varying delays and Markovian-switching parameters," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1857–1862, 2007.

[17] Y. Shen and J. Wang, "Almost sure exponential stability of recurrent neural networks with Markovian switching," *IEEE Transactions on Neural Networks*, vol. 20, no. 5, pp. 840–855, 2009.

[18] Q. Zhu and J. Cao, "Stability analysis of Markovian jump stochastic BAM neural networks with impulse control and mixed time delays," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 3, pp. 467–479, 2012.

[19] X. X. Liao and X. Mao, "Stability of stochastic neural networks," *Neural, Parallel & Scientific Computations*, vol. 4, no. 2, pp. 205–224, 1996.

[20] X. X. Liao and X. Mao, "Exponential stability and instability of stochastic neural networks," *Stochastic Analysis and Applications*, vol. 14, no. 2, pp. 165–185, 1996.

[21] Q. Luo, F. Deng, J. Bao, B. Zhao, and Y. Fu, "Stabilization of stochastic Hopfield neural network with distributed parameters," *Science in China. F. Information Sciences*, vol. 47, no. 6, pp. 752–762, 2004.

[22] Q. Luo, F. Deng, X. Mao, J. Bao, and Y. Zhang, "Theory and application of stability for stochastic reaction diffusion systems," *Science in China. F. Information Sciences*, vol. 51, no. 2, pp. 158–170, 2008.

[23] X. Mao, "Stability and stabilisation of stochastic differential delay equations," *IET Control Theory and Applications*, vol. 1, no. 6, pp. 1551–1566, 2007.

[24] H. Zhang, Z. Wang, and D. Liu, "Global asymptotic stability and robust stability of a class of Cohen-Grossberg neural networks with mixed delays," *IEEE Transactions on Circuits and Systems. I. Regular Papers*, vol. 56, no. 3, pp. 616–629, 2009.

[25] S. Zhu, Y. Shen, and G. Chen, "Exponential passivity of neural networks with time-varying delay and uncertainty," *Physics Letters A*, vol. 375, no. 2, pp. 136–142, 2010.

[26] K. Liu, *Stability of Infinite Dimensional Stochastic Differential Equations with Applications*, vol. 135 of *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2006.

[27] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, UK, 2006.

[28] D. J. Higham, X. Mao, and C. Yuan, "Preserving exponential mean-square stability in the simulation of hybrid stochastic differential equations," *Numerische Mathematik*, vol. 108, no. 2, pp. 295–325, 2007.

[29] X. Lou and B. Cui, "Stochastic exponential stability for Markovian jumping BAM neural networks with time-varying delays," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 37, no. 3, pp. 713–719, 2007.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 329437, 10 pages
http://dx.doi.org/10.1155/2014/329437

Research Article

Two Identification Methods for Dual-Rate Sampled-Data Nonlinear Output-Error Systems

Jing Chen[1] and Ruifeng Ding[2]

[1] *Wuxi Professional College of Science and Technology, Wuxi 214028, China*
[2] *School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China*

Correspondence should be addressed to Jing Chen; chenjing1981929@126.com

Received 4 February 2014; Accepted 20 March 2014; Published 10 April 2014

Academic Editor: Ming Gao

This paper presents two methods for dual-rate sampled-data nonlinear output-error systems. One method is the missing output estimation based stochastic gradient identification algorithm and the other method is the auxiliary model based stochastic gradient identification algorithm. Different from the polynomial transformation based identification methods, the two methods in this paper can estimate the unknown parameters directly. A numerical example is provided to confirm the effectiveness of the proposed methods.

1. Introduction

System identification plays an important part in many engineering applications [1–6]. Many identification methods assume that the input-output data at every sampling instant are available for linear systems [7–11] and nonlinear systems [12–20], which is usually not the case in practice. When the input and output signals of the systems have different sampling rates, these systems are usually called irregularly sampled-data systems [21–27], for example, dual-rate or multirate systems [28–30]. Dual-rate/multirate systems in which the input and the output are sampled at different frequencies arise widely in robust filtering and control [31–33], adaptive control [34–37], and system identification [38–43]. In the literature of dual-rate system identification, the so-called polynomial transformation technique is often used to transform the dual-rate model [44, 45].

As far as we know, the identification methods based on the polynomial transformation technique cannot directly estimate the parameters of the dual-rate system and the number of the unknown parameters to be estimated is more than the number of the unknown parameters of the original dual-rate system.

The nonlinear system consisting of a static nonlinear block followed by a linear dynamic system is called a Hammerstein system [46–49]. The nonlinearity of the Hammerstein system is usually expressed by some known basis functions [50, 51] or by a piece-wise polynomial function [52, 53]. When the Hammerstein system is a dual-rate system and has a preload nonlinearity, to the best of our knowledge, there is no work on identification of such systems. The main contributions of this paper are presenting the two methods directly for estimating the parameters of the dual-rate system. The proposed methods of this paper can combine the auxiliary model identification methods [54–57], the iterative identification methods [58–62], the multi-innovation identification methods [63–70], the hierarchical identification methods [71–83], and the two-stage or multistage identification methods [84, 85] to study identification problems for other linear systems [86–90] or nonlinear systems [91–97].

The rest of this paper is organized as follows. Section 2 introduces the dual-rate nonlinear output-error systems. Section 3 gives a missing output identification model based stochastic gradient algorithm. Section 4 provides an auxiliary model based stochastic gradient algorithm. Section 5 introduces an illustrative example. Finally, concluding remarks are given in Section 6.

2. Problem Formulation

Let "$A =: X$" or "$X := A$" stand for "A is defined as X," let the norm of a column vector X be $\|\mathbf{X}\|^2 := \operatorname{tr}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]$, and let the superscript T denote the matrix transpose.

Consider the following dual-rate nonlinear output-error system with colored noise:

$$y(t) = \frac{B(z)}{A(z)} f(u(t)) + v(t), \tag{1}$$

where $y(t)$ is the system output, $u(t)$ is the system input, $v(t)$ is a stochastic white noise with zero mean, $A(z)$ and $B(z)$ are the polynomials in the unit backward shift operator [$z^{-1}y(t) = y(t-1)$],

$$\begin{aligned} A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}, \\ B(z) &= b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n}, \end{aligned} \tag{2}$$

and $f(u(t))$ is a preload nonlinearity shown in Figure 1 and can be expressed as [98, 99]

$$f(u(t)) = \begin{cases} u(t) + m_1, & u(t) > 0, \\ 0, & u(t) = 0, \\ u(t) - m_2, & u(t) < 0, \end{cases} \tag{3}$$

where m_1 and $-m_2$ are two preload points.

For the dual-rate sampled-data system, all the input data $\{u(t), t = 0, 1, 2, \ldots\}$ and only the scarce output data $\{y(tq), t = 0, 1, 2, \ldots, (q \geqslant 2)\}$ are known. The intersample outputs or missing outputs $y(tq+j)$, $j = 1, 2, \ldots, q-1$ are unavailable.

Define a sign function

$$\operatorname{sgn}(u(t)) := \begin{cases} 1, & \text{if } u(t) > 0, \\ 0, & \text{if } u(t) = 0, \\ -1, & \text{if } u(t) < 0. \end{cases} \tag{4}$$

Then the function $f(u(t))$ can be expressed as

$$\begin{aligned} f(u(t)) = u(t) &+ \frac{m_1 + m_2}{2} \operatorname{sgn}(u(t)) \\ &+ \frac{m_1 - m_2}{2} \operatorname{sgn}\left(u^2(t)\right). \end{aligned} \tag{5}$$

Let

$$g_1 = \frac{m_1 + m_2}{2}, \qquad g_2 = \frac{m_1 - m_2}{2}. \tag{6}$$

Hence, we have

$$f(u(t)) = u(t) + g_1 \operatorname{sgn}(u(t)) + g_2 \operatorname{sgn}\left(u^2(t)\right). \tag{7}$$

Once g_1 and g_2 are estimated, the parameters m_1 and m_2 can be computed by $m_1 = g_1 + g_2$, $m_2 = g_1 - g_2$.

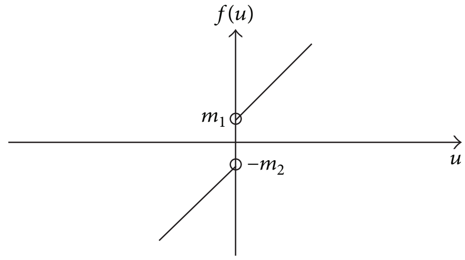


Figure 1: The preload characteristics.

3. The Missing Outputs Identification Model Based Stochastic Gradient Algorithm

Substituting (7) into (1) gets

$$\begin{aligned} A(z) y(t) = B(z) &\left(u(t) + g_1 \operatorname{sgn}(u(t)) + g_2 \operatorname{sgn}\left(u^2(t)\right)\right) \\ &+ A(z) v(t). \end{aligned} \tag{8}$$

Define the parameter vector $\boldsymbol{\theta}$ and information vector $\boldsymbol{\varphi}_1(t)$ as

$$\begin{aligned} \boldsymbol{\theta} := {} & [a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n, b_1 g_1, b_2 g_1, \ldots, \\ & b_n g_1, b_1 g_2, b_2 g_2, \ldots, b_n g_2]^{\mathrm{T}} \in \mathbb{R}^{4n}, \end{aligned} \tag{9}$$

$$\begin{aligned} \boldsymbol{\varphi}_1(t) := {} & \big[-y(t-1) + v(t-1), \\ & -y(t-2) + v(t-2), \ldots, -y(t-n) + v(t-n), \\ & u(t-1), u(t-2), \ldots, u(t-n), \\ & \operatorname{sgn}(u(t-1)), \operatorname{sgn}(u(t-2)), \ldots, \\ & \operatorname{sgn}(u(t-n)), \operatorname{sgn}\left(u^2(t-1)\right), \\ & \operatorname{sgn}\left(u^2(t-2)\right), \ldots, \operatorname{sgn}\left(u^2(t-n)\right)\big]^{\mathrm{T}} \in \mathbb{R}^{4n}. \end{aligned} \tag{10}$$

From (9) and (10), we get

$$y(t) = \boldsymbol{\varphi}_1^{\mathrm{T}}(t) \boldsymbol{\theta} + v(t) \tag{11}$$

or

$$y(tq) = \boldsymbol{\varphi}_1^{\mathrm{T}}(tq) \boldsymbol{\theta} + v(tq). \tag{12}$$

Let $\hat{\boldsymbol{\theta}}(t)$ be the estimate of $\boldsymbol{\theta}$. Defining and minimizing the cost function

$$J(\boldsymbol{\theta}) := \left[y(tq) - \boldsymbol{\varphi}_1^{\mathrm{T}}(tq)\boldsymbol{\theta}\right]^2 \tag{13}$$

give the following stochastic gradient (SG) algorithm for estimating $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}(tq) = \hat{\boldsymbol{\theta}}(tq-q) + \frac{\hat{\boldsymbol{\varphi}}_1(tq)}{r_1(tq)} e_1(tq), \tag{14}$$

$$\begin{aligned} \hat{\boldsymbol{\theta}}(tq-i) &= \hat{\boldsymbol{\theta}}(tq-q), \quad i = q-1,\ q-2, \ldots, 1, \\ e_1(tq) &= y(tq) - \hat{\boldsymbol{\varphi}}_1^{\mathrm{T}}(tq) \hat{\boldsymbol{\theta}}(tq-q), \end{aligned} \tag{15}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_1(tq) = \big[& -y(tq-1)+\hat{v}(tq-1), \\ & -y(tq-2)+\hat{v}(tq-2),\dots, \\ & -y(tq-n)+\hat{v}(tq-n), \\ & u(t-1),u(t-2),\dots,u(t-n), \\ & \operatorname{sgn}(u(t-1)),\operatorname{sgn}(u(t-2)),\dots, \\ & \operatorname{sgn}(u(t-n)),\operatorname{sgn}\left(u^2(t-1)\right), \\ & \operatorname{sgn}\left(u^2(t-2)\right),\dots,\operatorname{sgn}\left(u^2(t-n)\right)\big]^{\mathrm{T}},\end{aligned} \tag{16}$$

$$\hat{v}(tq-i) = y(tq-i) - \hat{\boldsymbol{\varphi}}_1^{\mathrm{T}}(tq-i)\hat{\boldsymbol{\theta}}(tq-i), \tag{17}$$

$$r_1(tq) = r_1(tq-q) + \|\hat{\boldsymbol{\varphi}}_1(tq)\|^2, \quad r(0)=1. \tag{18}$$

Since the information $\hat{\boldsymbol{\varphi}}_1(tq)$ on the right-hand sides of (16) contains the unknown variables $-y(tq-i)+\hat{v}(tq-i)$, $i = q-1, q-2,\dots,1$, the SG algorithm in (14)–(18) is impossible to implement. In this section, we use the missing outputs identification model (MOI) to overcome this difficulty; these unknown $-y(tq-i)+\hat{v}(tq-i)$ are replaced with the output estimates $-\hat{y}(tq-i)+\hat{v}(tq-i)$ of an MOI model,

$$\begin{aligned} -\hat{y}(tq-i)+\hat{v}(tq-i) &= -\hat{\boldsymbol{\varphi}}_1^{\mathrm{T}}(tq-i)\hat{\boldsymbol{\theta}}(tq-i), \\ & \qquad i = q-1,\ q-2,\dots,1, \\ \hat{\boldsymbol{\varphi}}_1(tq-i+1) & \\ = \big[& -\hat{y}(tq-i)+\hat{v}(tq-i), \\ & -\hat{y}(tq-i-1)+\hat{v}(tq-i-1),\dots, \\ & -\hat{y}(tq-q+1)+\hat{v}(tq-q+1), \\ & -y(tq-q)+\hat{v}(tq-q),\dots, \\ & -\hat{y}(tq-i+1-n)+\hat{v}(tq-i+1-n), \\ & u(tq-i),u(tq-i-1),\dots, \\ & u(tq-i+1-n),\operatorname{sgn}(u(tq-i)), \\ & \operatorname{sgn}(u(tq-i-1)),\dots, \\ & \operatorname{sgn}(u(tq-i+1-n)),\operatorname{sgn}\left(u^2(tq-i)\right), \\ & \operatorname{sgn}\left(u^2(tq-i-1)\right),\dots, \\ & \operatorname{sgn}\left(u^2(tq-i+1-n)\right)\big]^{\mathrm{T}}, \end{aligned} \tag{19}$$

where $-\hat{y}(tq-i)+\hat{v}(tq-i)$ represents the estimate of $-y(tq-i)+v(tq-i)$ at time $tq-i$, $\hat{\boldsymbol{\theta}}(tq-i)$ represents the estimate of $\boldsymbol{\theta}$ at time $tq-i$, and $\hat{\boldsymbol{\varphi}}_1(tq-i)$ represents the estimate of $\boldsymbol{\varphi}_1(q-i)$.

Thus, we have the following missing output estimates based SG (MOE-SG) algorithm for estimating the parameter vector $\boldsymbol{\theta}$ in (9):

$$\hat{\boldsymbol{\theta}}(tq) = \hat{\boldsymbol{\theta}}(tq-q) + \frac{\hat{\boldsymbol{\varphi}}_1(tq)}{r_1(tq)}e_2(tq), \tag{20}$$

$$\hat{\boldsymbol{\theta}}(tq-i) = \hat{\boldsymbol{\theta}}(tq-q), \quad i = q-1,\ q-2,\dots,1, \tag{21}$$

$$-\hat{y}(tq-i)+\hat{v}(tq-i) = -\hat{\boldsymbol{\varphi}}_1^{\mathrm{T}}(tq-i)\hat{\boldsymbol{\theta}}(tq-i), \tag{22}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_1(tq-i+1) & \\ = \big[& -\hat{y}(tq-i)+\hat{v}(tq-i), \\ & -\hat{y}(tq-i-1)+\hat{v}(tq-i-1),\dots, \\ & -\hat{y}(tq-q+1)+\hat{v}(tq-q+1), \\ & -y(tq-q)+\hat{v}(tq-q),\dots, \\ & -\hat{y}(tq-i+1-n)+\hat{v}(tq-i+1-n), \\ & u(tq-i),u(tq-i-1),\dots, \\ & u(tq-i+1-n),\operatorname{sgn}(u(tq-i)), \\ & \operatorname{sgn}(u(tq-i-1)),\dots,\operatorname{sgn}(u(tq-i+1-n)), \\ & \operatorname{sgn}\left(u^2(tq-i)\right),\operatorname{sgn}\left(u^2(tq-i-1)\right),\dots, \\ & \operatorname{sgn}\left(u^2(tq-i+1-n)\right)\big]^{\mathrm{T}},\end{aligned} \tag{23}$$

$$e_1(tq) = y(tq) - \hat{\boldsymbol{\varphi}}_1^{\mathrm{T}}(tq)\hat{\boldsymbol{\theta}}(tq-q), \tag{24}$$

$$r_1(tq) = r_1(tq-q) + \|\hat{\boldsymbol{\varphi}}_1(tq)\|^2, \quad r(0)=1. \tag{25}$$

The steps of computing the parameter estimate $\hat{\boldsymbol{\theta}}(tq)$ by the MOE-SG algorithm are listed as follows.

(1) Let $u(-j) = 0$, $y(-j) = 0$, $j = 0,1,2,\dots,n-1$, and give a small positive number ε.

(2) Let $t = 1$, $r(0) = 1$, and $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}/p_0$ with $\mathbf{1}$ being a column vector whose entries are all unity and $p_0 = 10^6$.

(3) Collect the input data $u(tq), u(tq-1),\dots,u(tq-n)$, and collect the output data $y(tq)$.

(4) Let $i = q-1$ and compute $-\hat{y}(tq-i)+\hat{v}(tq-i)$ by (22).

(5) Form $\hat{\boldsymbol{\varphi}}_1(tq-i+1)$ by (23).

(6) Decrease i by 1; if $i \geqslant 1$, go to step (4); otherwise, go to the next step.

(7) Compute $e_1(tq)$ and $r_1(tq)$ by (24) and (25), respectively.

(8) Update the parameter estimation vector $\hat{\boldsymbol{\theta}}(tq)$ by (20).

(9) Compare $\hat{\boldsymbol{\theta}}(tq)$ and $\hat{\boldsymbol{\theta}}(tq-q)$; if $\|\hat{\boldsymbol{\theta}}(tq)-\hat{\boldsymbol{\theta}}(tq-q)\| \leqslant \varepsilon$, then terminate the procedure and obtain the $\hat{\boldsymbol{\theta}}(tq)$; otherwise, increase t by 1 and go to step (3).

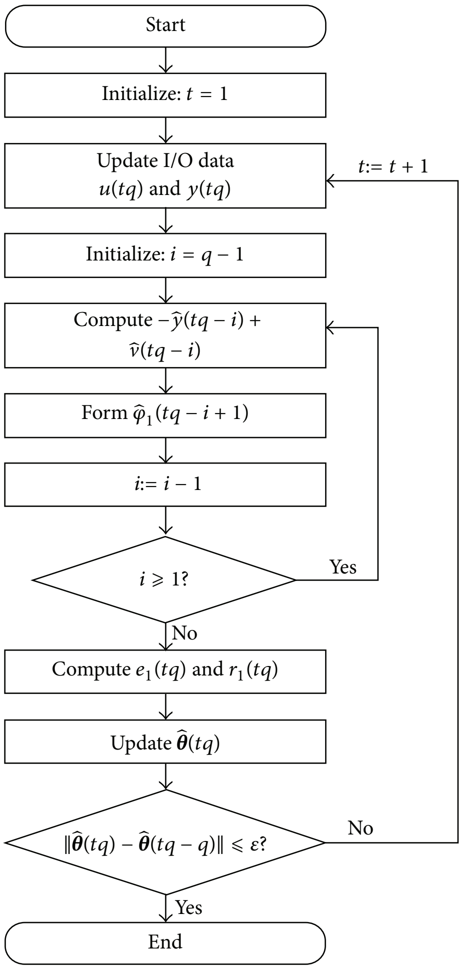


Figure 2: The flowchart of computing the estimate $\hat{\boldsymbol{\theta}}(tq)$.

The flowchart of computing the MOE-SG parameter estimate $\hat{\boldsymbol{\theta}}(tq)$ is shown in Figure 2.

4. The Auxiliary Model Based Stochastic Gradient Algorithm

Define

$$x(t) = \frac{B(z)}{A(z)}\left(u(t) + g_1 \operatorname{sgn}(u(t)) + g_2 \operatorname{sgn}\left(u^2(t)\right)\right). \tag{26}$$

From (8) and (26), we have

$$y(t) = x(t) + v(t). \tag{27}$$

Define the information vector $\boldsymbol{\varphi}_2(t)$ as

$$\begin{aligned}\boldsymbol{\varphi}_2(t) := \big[&-x(t-1), -x(t-2), \dots, -x(t-n),\\ &u(t-1), u(t-2), \dots, u(t-n),\\ &\operatorname{sgn}(u(t-1)), \operatorname{sgn}(u(t-2)), \dots,\\ &\operatorname{sgn}(u(t-n)),\\ &\operatorname{sgn}\left(u^2(t-1)\right), \operatorname{sgn}\left(u^2(t-2)\right), \dots,\\ &\operatorname{sgn}\left(u^2(t-n)\right)\big]^{\mathrm{T}} \in \mathbb{R}^{4n}.\end{aligned} \tag{28}$$

Then we get

$$x(t) = \boldsymbol{\varphi}_2^{\mathrm{T}}(t)\boldsymbol{\theta}, \tag{29}$$

$$y(t) = \boldsymbol{\varphi}_2^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t). \tag{30}$$

Assume t is an integer multiple of q and rewrite (30) as

$$y(tq) = \boldsymbol{\varphi}_2^{\mathrm{T}}(tq)\boldsymbol{\theta}(tq) + v(tq). \tag{31}$$

Let $\hat{\boldsymbol{\theta}}(t)$ be the estimate of $\boldsymbol{\theta}$. Defining and minimizing the cost function

$$J(\boldsymbol{\theta}) := \left[y(tq) - \boldsymbol{\varphi}_2^{\mathrm{T}}(tq)\boldsymbol{\theta}\right]^2 \tag{32}$$

give the following SG algorithm of estimating $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}(tq) = \hat{\boldsymbol{\theta}}(tq-q) + \frac{\boldsymbol{\varphi}_2(tq)}{r_2(tq)} e_2(tq), \tag{33}$$

$$e_2(tq) = y(tq) - \boldsymbol{\varphi}_2^{\mathrm{T}}(tq)\hat{\boldsymbol{\theta}}(tq-q), \tag{34}$$

$$\begin{aligned}\boldsymbol{\varphi}_2(tq) = \big[&-x(tq-1), -x(tq-2), \dots, -x(tq-n),\\ &u(t-1), u(t-2), \dots, u(t-n),\\ &\operatorname{sgn}(u(t-1)), \operatorname{sgn}(u(t-2)), \dots,\\ &\operatorname{sgn}(u(t-n)),\\ &\operatorname{sgn}\left(u^2(t-1)\right), \operatorname{sgn}\left(u^2(t-2)\right), \dots,\\ &\operatorname{sgn}(u^2(t-n))\big]^{\mathrm{T}},\end{aligned} \tag{35}$$

$$r_2(tq) = r_2(tq-q) + \|\boldsymbol{\varphi}_2(tq)\|^2, \quad r(0) = 1. \tag{36}$$

Because of the unknown variables $x(tq-i)$ in (33), the SG algorithm in (33)–(36) is impossible to implement. In this section, we use the auxiliary model; these unknown $x(tq-i)$ are replaced with the outputs $x_a(tq-i)$ of an auxiliary model,

$$x_a(tq-i) = \boldsymbol{\theta}_a^{\mathrm{T}}(tq-i)\boldsymbol{\varphi}_a(tq-i), \tag{37}$$

where $\boldsymbol{\theta}_a(tq-i)$ is the estimate $\hat{\boldsymbol{\theta}}(tq-i)$ of $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}_a(tq-i)$ is the estimate $\hat{\boldsymbol{\varphi}}_2(tq-i)$ of $\boldsymbol{\varphi}_2(tq-i)$. We can obtain an auxiliary model based stochastic gradient (AM-SG) algorithm:

$$\hat{\boldsymbol{\theta}}(tq) = \hat{\boldsymbol{\theta}}(tq-q) + \frac{\hat{\boldsymbol{\varphi}}_2(tq)}{r_2(tq)} e_2(tq), \tag{38}$$

$$\hat{\boldsymbol{\theta}}(tq-i) = \hat{\boldsymbol{\theta}}(tq-q), \quad i = q-1,\ q-2, \dots, 1, \tag{39}$$

$$x_a(tq-i) = \hat{\boldsymbol{\theta}}^{\mathrm{T}}(tq-i)\hat{\boldsymbol{\varphi}}_2(tq-i), \tag{40}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_2(tq-i+1) = \big[& -x_a(tq-i), -x_a(tq-i-1), \ldots, \\ & -x_a(tq-i+1-n), \\ & u(t-i), u(t-i-1), \ldots, \\ & u(t-i+1-n), \\ & \operatorname{sgn}(u(t-i)), \operatorname{sgn}(u(t-i-1)), \ldots, \\ & \operatorname{sgn}(u(t-i+1-n)), \\ & \operatorname{sgn}\left(u^2(t-i)\right), \\ & \operatorname{sgn}\left(u^2(t-i-1)\right), \ldots, \\ & \operatorname{sgn}(u^2(t-i+1-n))\big]^{\mathrm{T}},\end{aligned} \tag{41}$$

$$e_2(tq) = y(tq) - \hat{\boldsymbol{\varphi}}_2^{\mathrm{T}}(tq)\hat{\boldsymbol{\theta}}(tq-q), \tag{42}$$

$$r_2(tq) = r_2(tq-q) + \|\hat{\boldsymbol{\varphi}}_2(tq)\|^2, \quad r(0) = 1. \tag{43}$$

The steps of computing the parameter estimate $\hat{\boldsymbol{\theta}}(tq)$ by the AM-SG algorithm are listed as follows.

(1) Let $u(-j) = 0, y(-j) = 0, x(-j) = 0, j = 0, 1, 2, \ldots, n-1$, and give a small positive number ε.

(2) Let $t = 1$, $r(0) = 1$, and $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}/p_0$ with $\mathbf{1}$ being a column vector whose entries are all unity and $p_0 = 10^6$.

(3) Collect the input data $u(tq), u(tq-1), \ldots, u(tq-n)$, and collect the output data $y(tq)$.

(4) Let $i = q-1$ and compute $x_a(tq-i)$ by (40).

(5) Form $\hat{\boldsymbol{\varphi}}_2(tq-i+1)$ by (41).

(6) Decrease i by 1; if $i \geqslant 1$, go to step (4); otherwise, go to next step.

(7) Compute $e_2(tq)$ and $r_2(tq)$ by (42) and (43), respectively.

(8) Update the parameter estimation vector $\hat{\boldsymbol{\theta}}(tq)$ by (38).

(9) Compare $\hat{\boldsymbol{\theta}}(tq)$ and $\hat{\boldsymbol{\theta}}(tq-q)$; if $\|\hat{\boldsymbol{\theta}}(tq) - \hat{\boldsymbol{\theta}}(tq-q)\| \leqslant \varepsilon$, then terminate the procedure and obtain the $\hat{\boldsymbol{\theta}}(tq)$; otherwise, increase t by 1 and go to step (3).

The flowchart of computing the AM-SG parameter estimate $\hat{\boldsymbol{\theta}}(tq)$ is shown in Figure 3.

Remark 1. Compared with the polynomial transformation technique, the MOE-SG method and the AM-SG method can estimate the unknown parameters directly.

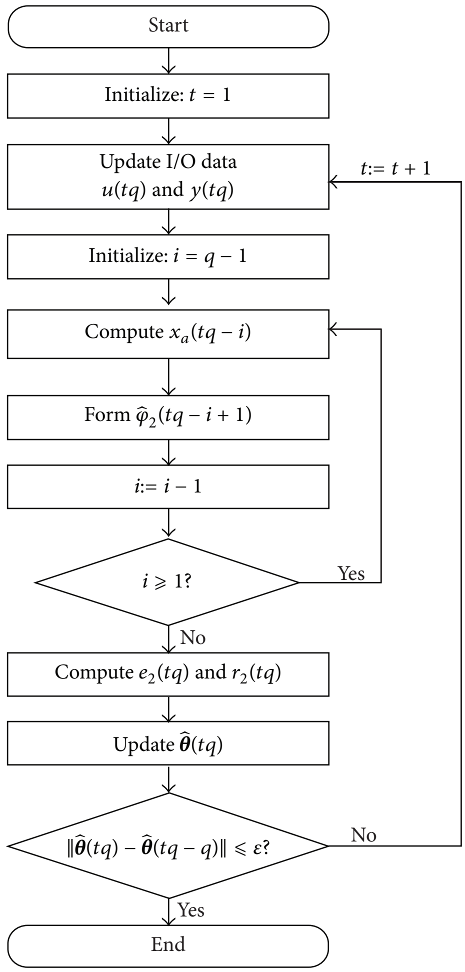


Figure 3: The flowchart of computing the estimate $\hat{\boldsymbol{\theta}}_2(tq)$.

5. Example

Consider the following nonlinear output-error system with the updating period $q = 2$:

$$\begin{aligned}y(t) &= \frac{B(z)}{A(z)} f(u(t)) + v(t), \\ A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} = 1 + 0.49z^{-1} - 0.2z^{-2}, \\ B(z) &= b_1 z^{-1} + b_2 z^{-2} = 0.2z^{-1} + 0.4z^{-2}, \\ f(u(t)) &= u(t) + \frac{m_1 + m_2}{2}\operatorname{sgn}(u(t)) \\ &\quad + \frac{m_1 - m_2}{2}\operatorname{sgn}\left(u^2(t)\right) \\ &= u(t) + \frac{0.5 + 0.3}{2}\operatorname{sgn}(u(t)) \\ &\quad + \frac{0.5 - 0.3}{2}\operatorname{sgn}\left(u^2(t)\right)\end{aligned}$$

Table 1: The MOE-SG algorithm estimates and errors.

t	1000	2000	3000	4000	5000	True values
a_1	0.30790	0.43409	0.48162	0.49513	0.49505	0.49000
a_2	−0.16601	−0.20319	−0.20626	−0.20656	−0.20341	−0.20000
b_1	0.19508	0.19548	0.19462	0.19665	0.19816	0.20000
b_2	0.36487	0.39043	0.39879	0.40105	0.39987	0.40000
b_1g_1	0.09729	0.09384	0.08995	0.08769	0.08705	0.08000
b_2g_1	0.13565	0.14818	0.15401	0.15931	0.15867	0.16000
b_1g_2	0.02161	0.02602	0.02558	0.02764	0.02770	0.02000
b_2g_2	0.02641	0.03181	0.03127	0.03378	0.03385	0.04000
δ (%)	26.70140	8.46344	2.72656	2.15284	1.91759	

Table 2: The AM-SG algorithm estimates and errors.

t	1000	2000	3000	4000	5000	True values
a_1	0.39201	0.46141	0.50310	0.49802	0.48917	0.49000
a_2	−0.18980	−0.19696	−0.19784	−0.20113	−0.20307	−0.20000
b_1	0.18974	0.19349	0.19872	0.20192	0.20281	0.20000
b_2	0.40122	0.41674	0.39648	0.40109	0.40350	0.40000
b_1g_1	0.09799	0.08924	0.08427	0.08475	0.08276	0.08000
b_2g_1	0.14716	0.15484	0.15489	0.16514	0.16040	0.16000
b_1g_2	0.02005	0.02781	0.02034	0.02761	0.02600	0.02000
b_2g_2	0.02674	0.03708	0.02712	0.03682	0.03467	0.04000
δ (%)	14.27547	5.08770	2.79209	1.91002	1.41209	

$$\begin{aligned} &= u(t) + g_1 \operatorname{sgn}(u(t)) + g_2 \operatorname{sgn}\left(u^2(t)\right) \\ &= u(t) + 0.4 \operatorname{sgn}(u(t)) + 0.1 \operatorname{sgn}\left(u^2(t)\right); \end{aligned} \tag{44}$$

the input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance and $\{v(t)\}$ is a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$. The unknown parameters are as follows:

$$\begin{aligned} \boldsymbol{\theta} &= [a_1, a_2, b_1, b_2, b_1g_1, b_2g_1, b_1g_2, b_2g_2]^{\mathrm{T}} \\ &= [0.49, -0.2, 0.2, 0.4, 0.08, 0.16, 0.02, 0.04]^{\mathrm{T}}. \end{aligned} \tag{45}$$

Applying the MOE-SG algorithm and the AM-SG algorithm to estimate the parameters, the parameter estimates and their errors based on the MOE-SG algorithm and the AM-SG algorithm are shown in Tables 1 and 2 and the parameter estimation errors $\delta := \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ versus t are shown in Figures 4 and 5.

From Tables 1 and 2 and Figures 4 and 5, we can draw the following conclusions.

(1) Both the MOE-SG algorithm and the AM-SG algorithm can estimate the unknown parameters directly.

(2) The parameter estimation errors become smaller and smaller and go to zero with t increasing.

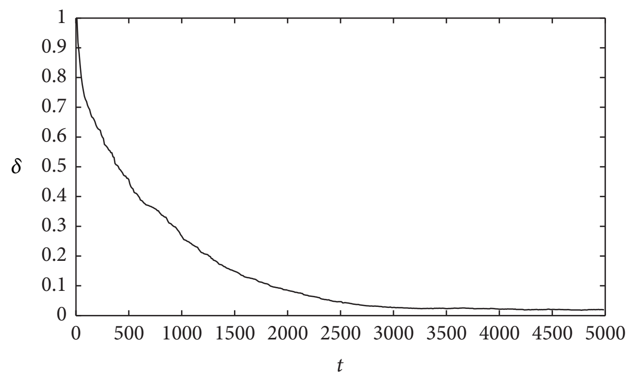


Figure 4: The parameter estimation errors δ versus t (MOE-SG).

6. Conclusions

Two identification methods for dual-rate nonlinear output-error systems are presented to estimate the unknown parameters directly and can avoid estimating more parameters than the original systems. Furthermore, the two methods can also be extended to other systems such as

$$\begin{aligned} y(t) &= \frac{B(z)}{A(z)} f(u(t)) + \frac{D(z)}{C(z)} v(t), \\ A(z)\, y(t) &= B(z)\, f(u(t)) + D(z)\, v(t). \end{aligned} \tag{46}$$

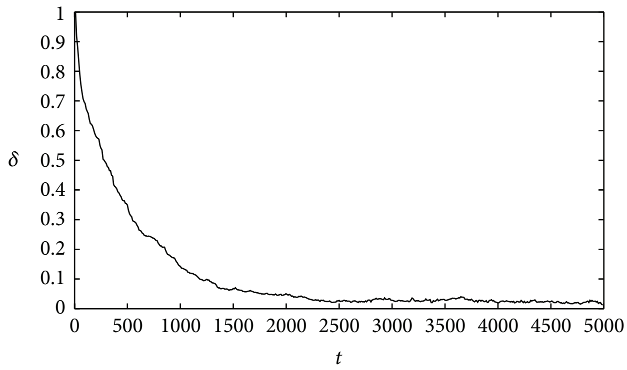


FIGURE 5: The parameter estimation errors δ versus t (AM-SG).

Conflict of Interests

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China and supported by the Natural Science Foundation of Jiangsu Province (no. BK20131109).

References

[1] F. Ding, *System Identification—New Theory and Methods*, Science Press, Beijing, China, 2013.

[2] F. Ding, *System Identification—Performances Analysis for Identification Methods*, Science Press, Beijing, China, 2014.

[3] Y. Liu, Y. Xiao, and X. Zhao, "Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model," *Applied Mathematics and Computation*, vol. 215, no. 4, pp. 1477–1483, 2009.

[4] Y. Liu, L. Xie, and F. Ding, "An auxiliary model based on a recursive least-squares parameter estimation algorithm for non-uniformly sampled multirate systems," *Proceedings of the Institution of Mechanical Engineers: Journal of Systems and Control Engineering*, vol. 223, no. 4, pp. 445–454, 2009.

[5] Y. Liu, J. Sheng, and R. F. Ding, "Convergence of stochastic gradient estimation algorithm for multivariable ARX-like systems," *Computers and Mathematics with Applications*, vol. 59, no. 8, pp. 2615–2627, 2010.

[6] Y. J. Liu and R. Ding, "Consistency of the extended gradient identification algorithm for multi-input multi-output systems with moving average noises," *International Journal of Computer Mathematics*, vol. 90, no. 9, pp. 1840–1852, 2013.

[7] F. Ding and T. Chen, "Performance bounds of forgetting factor least-squares algorithms for time-varying systems with finite measurement data," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 3, pp. 555–566, 2005.

[8] F. Ding, Y. Shi, and T. Chen, "Performance analysis of estimation algorithms of nonstationary ARMA processes," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1041–1053, 2006.

[9] F. Ding, T. Chen, and L. Qiu, "Bias compensation based recursive least-squares identification algorithm for MISO systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 5, pp. 349–353, 2006.

[10] F. Ding, H. Yang, and F. Liu, "Performance analysis of stochastic gradient algorithms under weak conditions," *Science in China F: Information Sciences*, vol. 51, no. 9, pp. 1269–1280, 2008.

[11] F. Ding, "Coupled-least-squares identification for multivariable systems," *IET Control Theory and Applications*, vol. 7, no. 1, pp. 68–79, 2013.

[12] F. Ding, X. G. Liu, and J. Chu, "Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle," *IET Control Theory and Applications*, vol. 7, pp. 176–184, 2013.

[13] D. Wang and F. Ding, "Least squares based and gradient based iterative identification for Wiener nonlinear systems," *Signal Processing*, vol. 91, no. 5, pp. 1182–1189, 2011.

[14] D. Q. Wang and F. Ding, "Hierarchical least squares estimation algorithm for Hammerstein-Wiener systems," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 825–828, 2012.

[15] D. Q. Wang, F. Ding, and Y. Y. Chu, "Data filtering based recursive least squares algorithm for Hammerstein systems using the key-term separation principle," *Information Sciences*, vol. 222, pp. 203–212, 2013.

[16] D. Q. Wang, F. Ding, and X. M. Liu, "Least squares algorithm for an input nonlinear system with a dynamic subspace state space model," *Nonlinear Dynamics*, vol. 75, no. 1-2, pp. 49–61, 2014.

[17] J. Chen, Y. Zhang, and R. F. Ding, "Auxiliary model based multi-innovation algorithms for multivariable nonlinear systems," *Mathematical and Computer Modelling*, vol. 52, no. 9-10, pp. 1428–1434, 2010.

[18] J. Chen, X. Wang, and R. F. Ding, "Gradient based estimation algorithm for Hammerstein systems with saturation and dead-zone nonlinearities," *Applied Mathematical Modelling*, vol. 36, no. 1, pp. 238–243, 2012.

[19] J. Chen and F. Ding, "Least squares and stochastic gradient parameter estimation for multivariable nonlinear Box-Jenkins models based on the auxiliary model and the multi-innovation identification theory," *Engineering Computations*, vol. 29, no. 8, pp. 907–921, 2012.

[20] J. Chen, Y. Zhang, and R. F. Ding, "Gradient-based parameter estimation for input nonlinear systems with ARMA noises based on the auxiliary model," *Nonlinear Dynamics*, vol. 72, no. 4, pp. 865–871, 2013.

[21] F. Ding, L. Qiu, and T. Chen, "Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems," *Automatica*, vol. 45, no. 2, pp. 324–332, 2009.

[22] F. Ding and J. Ding, "Least-squares parameter estimation for systems with irregularly missing data," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 7, pp. 540–553, 2010.

[23] F. Ding, P. X. Liu, and G. Liu, "Multiinnovation least-squares identification for system modeling," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 40, no. 3, pp. 767–778, 2010.

[24] Y. J. Liu, F. Ding, and Y. Shi, "Least squares estimation for a class of non-uniformly sampled systems based on the hierarchical identification principle," *Circuits, Systems and Signal Processing*, vol. 31, no. 6, pp. 1985–2000, 2012.

[25] F. Ding, G. Liu, and X. P. Liu, "Partially coupled stochastic gradient identification methods for non-uniformly sampled systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1976–1981, 2010.

[26] J. Ding, F. Ding, X. P. Liu, and G. Liu, "Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2677–2683, 2011.

[27] F. Ding, G. Liu, and X. P. Liu, "Parameter estimation with scarce measurements," *Automatica*, vol. 47, no. 8, pp. 1646–1655, 2011.

[28] J. Chen, "Several gradient parameter estimation algorithms for dual-rate sampled systems," *Journal of the Franklin Institute. Engineering and Applied Mathematics*, vol. 351, no. 1, pp. 543–554, 2014.

[29] J. Chen and R. Ding, "An auxiliary-model-based stochastic gradient algorithm for dual-rate sampled-data Box-Jenkins systems," *Circuits Systems and Signal Processing*, vol. 32, no. 5, pp. 2475–2485, 2013.

[30] J. Ding, Y. Shi, H. Wang, and F. Ding, "A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems," *Digital Signal Processing: A Review Journal*, vol. 20, no. 4, pp. 1238–1247, 2010.

[31] Y. Shi and B. Yu, "Output feedback stabilization of networked control systems with random delays modeled by Markov chains," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1668–1674, 2009.

[32] Y. Shi and H. Fang, "Kalman filter-based identification for systems with randomly missing measurements in a network environment," *International Journal of Control*, vol. 83, no. 3, pp. 538–551, 2010.

[33] Y. Shi and B. Yu, "Robust mixed $H2/H_{\infty}$ control of networked control systems with random time delays in both forward and backward communication links," *Automatica*, vol. 47, no. 4, pp. 754–760, 2011.

[34] F. Ding and T. Chen, "Least squares based self-tuning control of dual-rate systems," *International Journal of Adaptive Control and Signal Processing*, vol. 18, no. 8, pp. 697–714, 2004.

[35] F. Ding and T. Chen, "A gradient based adaptive control algorithm for dual-rate systems," *Asian Journal of Control*, vol. 8, no. 4, pp. 314–323, 2006.

[36] F. Ding, T. Chen, and Z. Iwai, "Adaptive digital control of Hammerstein nonlinear systems with limited output sampling," *SIAM Journal on Control and Optimization*, vol. 45, no. 6, pp. 2257–2276, 2007.

[37] J. Zhang, F. Ding, and Y. Shi, "Self-tuning control based on multi-innovation stochastic gradient parameter estimation," *Systems and Control Letters*, vol. 58, no. 1, pp. 69–75, 2009.

[38] Y. Shi, F. Ding, and T. Chen, "Multirate crosstalk identification in xDSL systems," *IEEE Transactions on Communications*, vol. 54, no. 10, pp. 1878–1886, 2006.

[39] Y. J. Liu, F. Ding, and Y. Shi, "An efficient hierarchical identification method for general dual-rate sampled-data systems," *Automatica*, vol. 50, no. 3, pp. 962–973, 2014.

[40] F. Ding and T. Chen, "Identification of dual-rate systems based on finite impulse response models," *International Journal of Adaptive Control and Signal Processing*, vol. 18, no. 7, pp. 589–598, 2004.

[41] F. Ding and T. Chen, "Combined parameter and output estimation of dual-rate systems using an auxiliary model," *Automatica*, vol. 40, no. 10, pp. 1739–1748, 2004.

[42] F. Ding and T. Chen, "Parameter estimation of dual-rate stochastic systems by using an output error method," *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1436–1441, 2005.

[43] F. Ding and T. Chen, "Hierarchical identification of lifted state-space models for general dual-rate systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 6, pp. 1179–1187, 2005.

[44] F. Ding, P. X. Liu, and Y. Shi, "Convergence analysis of estimation algorithms for dual-rate stochastic systems," *Applied Mathematics and Computation*, vol. 176, no. 1, pp. 245–261, 2006.

[45] F. Ding, P. X. Liu, and H. Yang, "Parameter identification and intersample output estimation for dual-rate systems," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 38, no. 4, pp. 966–975, 2008.

[46] J. Li and F. Ding, "Maximum likelihood stochastic gradient estimation for Hammerstein systems with colored noise based on the key term separation technique," *Computers and Mathematics with Applications*, vol. 62, no. 11, pp. 4170–4177, 2011.

[47] X. L. Li, L. C. Zhou, R. Ding, and J. Shing, "Recursive least-squares estimation for Hammerstein nonlinear systems with nonuniform sampling," *Mathematical Problems in Engineering*, vol. 2013, Article ID 240929, 8 pages, 2013.

[48] F. Ding, X. P. Liu, and G. Liu, "Identification methods for Hammerstein nonlinear systems," *Digital Signal Processing: A Review Journal*, vol. 21, no. 2, pp. 215–238, 2011.

[49] F. Ding, "Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling," *Applied Mathematical Modelling*, vol. 37, no. 4, pp. 1694–1704, 2013.

[50] J. H. Li, "Parameter estimation for Hammerstein CARARMA systems based on the Newton iteration," *Applied Mathematics Letters*, vol. 26, no. 1, pp. 91–96, 2013.

[51] D. Wang, Y. Chu, and F. Ding, "Auxiliary model-based RELS and MI-ELS algorithm for Hammerstein OEMA systems," *Computers and Mathematics with Applications*, vol. 59, no. 9, pp. 3092–3098, 2010.

[52] J. Vörös, "Modeling and parameter identification of systems with multisegment piecewise-linear characteristics," *IEEE Transactions on Automatic Control*, vol. 47, no. 1, pp. 184–188, 2002.

[53] J. Vörös, "Modeling and identification of systems with backlash," *Automatica*, vol. 46, no. 2, pp. 369–374, 2010.

[54] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Systems and Control Letters*, vol. 56, no. 5, pp. 373–380, 2007.

[55] F. Ding, P. X. Liu, and G. Liu, "Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises," *Signal Processing*, vol. 89, no. 10, pp. 1883–1890, 2009.

[56] F. Ding and Y. Gu, "Performance analysis of the auxiliary model based least squares identification algorithm for one-step state delay systems," *International Journal of Computer Mathematics*, vol. 89, no. 15, pp. 2019–2028, 2012.

[57] F. Ding and Y. Gu, "Performance analysis of the auxiliary model-based stochastic gradient parameter estimation algorithm for state space systems with one-step state delay," *Circuits, Systems and Signal Processing*, vol. 32, no. 2, pp. 585–599, 2013.

[58] F. Ding, Y. Liu, and B. Bao, "Gradient-based and least-squares-based iterative estimation algorithms for multi-input multi-output systems," *Proceedings of the Institution of Mechanical Engineers I: Journal of Systems and Control Engineering*, vol. 226, no. 1, pp. 43–55, 2012.

[59] F. Ding, P. X. Liu, and G. Liu, "Gradient based and least-squares based iterative identification methods for OE and

OEMA systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 664–677, 2010.

[60] F. Ding, "Decomposition based fast least squares algorithm for output error systems," *Signal Processing*, vol. 93, no. 5, pp. 1235–1242, 2013.

[61] Y. Liu, D. Wang, and F. Ding, "Least squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data," *Digital Signal Processing*, vol. 20, no. 5, pp. 1458–1467, 2010.

[62] D. Q. Wang, "Least squares-based recursive and iterative estimation for output error moving average systems using data filtering," *IET Control Theory and Applications*, vol. 5, no. 14, pp. 1648–1657, 2011.

[63] F. Ding and T. Chen, "Performance analysis of multi-innovation gradient type identification methods," *Automatica*, vol. 43, no. 1, pp. 1–14, 2007.

[64] F. Ding, "Several multi-innovation identification methods," *Digital Signal Processing*, vol. 20, no. 4, pp. 1027–1039, 2010.

[65] F. Ding, H. Chen, and M. Li, "Multi-innovation least squares identification methods based on the auxiliary model for MISO systems," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 658–668, 2007.

[66] L. Han and F. Ding, "Multi-innovation stochastic gradient algorithms for multi-input multi-output systems," *Digital Signal Processing*, vol. 19, no. 4, pp. 545–554, 2009.

[67] D. Wang and F. Ding, "Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 750–762, 2010.

[68] L. Xie, Y. J. Liu, H. Z. Yang, and F. Ding, "Modelling and identification for non-uniformly periodically sampled-data systems," *IET Control Theory and Applications*, vol. 4, no. 5, pp. 784–794, 2010.

[69] Y. Liu, L. Yu, and F. Ding, "Multi-innovation extended stochastic gradient algorithm and its performance analysis," *Circuits, Systems, and Signal Processing*, vol. 29, no. 4, pp. 649–667, 2010.

[70] L. Han and F. Ding, "Identification for multirate multi-input systems using the multi-innovation identification theory," *Computers and Mathematics with Applications*, vol. 57, no. 9, pp. 1438–1449, 2009.

[71] F. Ding and T. Chen, "Hierarchical gradient-based identification of multivariable discrete-time systems," *Automatica*, vol. 41, no. 2, pp. 315–325, 2005.

[72] F. Ding and T. Chen, "Hierarchical least squares identification methods for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 397–402, 2005.

[73] H. Han, L. Xie, F. Ding, and X. Liu, "Hierarchical least-squares based iterative identification for multivariable systems with moving average noises," *Mathematical and Computer Modelling*, vol. 51, no. 9-10, pp. 1213–1220, 2010.

[74] Z. Zhang, F. Ding, and X. Liu, "Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems," *Computers and Mathematics with Applications*, vol. 61, no. 3, pp. 672–682, 2011.

[75] D. Q. Wang, R. Ding, and X. Z. Dong, "Iterative parameter estimation for a class of multivariable systems based on the hierarchical identification principle and the gradient search," *Circuits, Systems and Signal Processing*, vol. 31, no. 6, pp. 2167–2177, 2012.

[76] F. Ding and T. Chen, "On iterative solutions of general coupled matrix equations," *SIAM Journal on Control and Optimization*, vol. 44, no. 6, pp. 2269–2284, 2006.

[77] F. Ding, P. X. Liu, and J. Ding, "Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle," *Applied Mathematics and Computation*, vol. 197, no. 1, pp. 41–50, 2008.

[78] F. Ding and T. Chen, "Gradient based iterative algorithms for solving a class of matrix equations," *IEEE Transactions on Automatic Control*, vol. 50, no. 8, pp. 1216–1221, 2005.

[79] F. Ding and T. Chen, "Iterative least-squares solutions of coupled Sylvester matrix equations," *Systems and Control Letters*, vol. 54, no. 2, pp. 95–107, 2005.

[80] F. Ding, "Transformations between some special matrices," *Computers and Mathematics with Applications*, vol. 59, no. 8, pp. 2676–2695, 2010.

[81] L. Xie, J. Ding, and F. Ding, "Gradient based iterative solutions for general linear matrix equations," *Computers and Mathematics with Applications*, vol. 58, no. 7, pp. 1441–1448, 2009.

[82] J. Ding, Y. Liu, and F. Ding, "Iterative solutions to matrix equations of the form $A_i X B_i = F_i$," *Computers and Mathematics with Applications*, vol. 59, no. 11, pp. 3500–3507, 2010.

[83] L. Xie, Y. Liu, and H. Yang, "Gradient based and least squares based iterative algorithms for matrix equations $AXB + CX^T D = F$," *Applied Mathematics and Computation*, vol. 217, no. 5, pp. 2191–2199, 2010.

[84] F. Ding, "Two-stage least squares based iterative estimation algorithm for CARARMA system modeling," *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 4798–4808, 2013.

[85] F. Ding and H. H. Duan, "Two-stage parameter estimation algorithms for Box-Jenkins systems," *IET Signal Processing*, vol. 7, no. 8, pp. 646–654, 2013.

[86] J. Ding and F. Ding, "Bias compensation-based parameter estimation for output error moving average systems," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 12, pp. 1100–1111, 2011.

[87] Y. Zhang, "Unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods," *Mathematical and Computer Modelling*, vol. 53, no. 9-10, pp. 1810–1819, 2011.

[88] Y. Zhang and G. Cui, "Bias compensation methods for stochastic systems with colored noise," *Applied Mathematical Modelling*, vol. 35, no. 4, pp. 1709–1716, 2011.

[89] F. Ding, "Combined state and least squares parameter estimation algorithms for dynamic systems," *Applied Mathematical Modelling*, vol. 38, no. 1, pp. 403–412, 2014.

[90] F. Ding, X. M. Liu, H. B. Chen, and G. Y. Yao, "Hierarchical gradient based and hierarchical least squares based iterative parameter identification for CARARMA systems," *Signal Processing*, vol. 97, pp. 31–39, 2014.

[91] P. P. Hu and F. Ding, "Multistage least squares based iterative estimation for feedback nonlinear systems with moving average noises using the hierarchical identification principle," *Nonlinear Dynamics*, vol. 73, no. 1-2, pp. 583–592, 2013.

[92] J. Li, F. Ding, and G. Yang, "Maximum likelihood least squares identification method for input nonlinear finite impulse response moving average systems," *Mathematical and Computer Modelling*, vol. 55, no. 3-4, pp. 442–450, 2012.

[93] F. Ding and T. Chen, "Identification of Hammerstein nonlinear ARMAX systems," *Automatica*, vol. 41, no. 9, pp. 1479–1489, 2005.

[94] F. Ding, Y. Shi, and T. Chen, "Gradient-based identification methods for hammerstein nonlinear ARMAX models," *Nonlinear Dynamics*, vol. 45, no. 1-2, pp. 31–43, 2006.

[95] J. H. Li, F. Ding, and L. Hua, "Maximum likelihood Newton recursive and the Newton iterative estimation algorithms for Hammerstein CARAR systems," *Nonlinear Dynamics*, vol. 75, no. 1-2, pp. 234–245, 2014.

[96] X. Luan, P. Shi, and F. Liu, "Stabilization of networked control systems with random delays," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4323–4330, 2011.

[97] X. L. Luan, S. Y. Zhao, and F. Liu, "H-infinity control for discrete-time markov jump systems with uncertain transition probabilities," *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1566–1572, 2013.

[98] J. Chen, L. X. Lu, and R. Ding, "Parameter identification of systems with preload nonlinearities based on the finite impulse response model and negative gradient search," *Applied Mathematics and Computation*, vol. 219, no. 5, pp. 2498–2505, 2012.

[99] J. Chen, L. Lv, and R. F. Ding, "Multi-innovation stochastic gradient algorithms for dual-rate sampled systems with preload nonlinearity," *Applied Mathematics Letters*, vol. 26, no. 1, pp. 124–129, 2013.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 154679, 8 pages
http://dx.doi.org/10.1155/2014/154679

Research Article

Input-To-State Stability for a Class of Switched Stochastic Nonlinear Systems by an Improved Average Dwell Time Method

Rongwei Guo,[1,2] Ping Zhao,[3] and Chenghui Zhang[1]

[1] *School of Control Science and Engineering, Shandong University, Jinan 250061, China*
[2] *School of Science, Qilu University of Technology, Jinan 250353, China*
[3] *School of Electrical Engineering, University of Jinan, Jinan 250022, China*

Correspondence should be addressed to Chenghui Zhang; zchui@sdu.edu.cn

Received 21 November 2013; Accepted 11 March 2014; Published 6 April 2014

Academic Editor: Ming Gao

This paper investigates the input-to-state stable in the mean (ISSiM) property of the switched stochastic nonlinear (SSN) systems with an improved average dwell time (ADT) method in two cases: (i) all of the constituent subsystems are ISSiM and (ii) parts of the constituent subsystems are ISSiM. First, an improved ADT method for stability of SSN systems is introduced. Then, based on that not only a new ISSiM result for SSN systems whose subsystems are ISSiM is presented, but also a new ISSiM result for such systems in which parts of subsystems are ISSiM is established. In comparison with the existing ones, the main results obtained in this paper have some advantages. Finally, an illustrative example with numerical simulation is verified the correctness and validity of the proposed results.

1. Introduction

Switched systems, which provide a unified framework for mathematical modeling of many physical or man-made systems, display switching features such as communication networks, manufacturing, computer synchronization, auto pilot control design, automotive engine control, traffic control, and chemical processes. The systems consist of a collection of indexed differential or difference equations and a switching signal governing the switching among them. In the past two decades, increasing attention has been paid to the analysis and synthesis of switched systems because of their significance in both theory and applications, and many significant results have been established for the stability analysis and control design of such switched systems; see [1–12] and references therein. Regarding the stability analysis problem, there are two famous analysis methods, that is, common Lyapunov function (CLF) method [4, 5], and multiple Lyapunov functions (MLF) method [9]. Although the CLF method is very useful in stability analysis and control design, it is difficult to be applied in practice because of the following reason: for a given switched nonlinear system, there is no general method to determine whether all subsystems share a CLF or not, even for the switched linear systems. About the MLF method, it has been proved in [9] that the switched linear systems with stable subsystems are globally asymptotically stable (GAS) if the dwell time (DT) τ_d is sufficiently large. Therefore, a DT method [9] is established to analyze the stability analysis and control design of the switched systems; that is, given a constant $\tau_d > 0$, let $S_d[\tau_d]$ denote the set of all switching signals with interval between consecutive switchings being no smaller than τ_d, τ_d is called the “dwell time”. Recently, Ni et al. think it is necessary to find a minimum dwell time (MDT) τ_d^*, which ensures that the switched system stays on each mode for period greater than or equal to τ_d^*; the system is GAS and have obtained a new method called MDT method [13]. However, the above MDT method is only for the switched linear systems, and it is impossible to be extended to switched nonlinear systems concluding from the proofs of the results in [13]. It is well known that the ADT scheme characterizes a large class of stable switching signals than dwell time scheme. Thus, the ADT method is very important not only in practice but also in theory. Considerable attention has been paid, and many efforts have been done to take advantage of the ADT method to investigate the stability analysis and control design both in switched linear and nonlinear systems. In [14], we obtain an improved ADT method for the switched nonlinear systems, which have two advantages over the existing ADT methods [15–18]: one is that the conditions of the improved ADT method are less than those; the other one is that the obtained lower bound of ADT (i.e., τ_a^*) is also smaller than those obtained by the above ADT methods.

When a control system is affected by an external input, it is important to analyze how the external input affects the system's behavior. Input-to-state stable (ISS) property [19] characterizes the continuity of state trajectories on the initial states and the external inputs, and integral input-to-state stable (iISS) property is a weaker concept introduced in [20], and the iISS property has been shown to be a natural generalization of ISS. Both ISS and iISS have been proven to be useful in the stability analysis and control design of nonlinear systems; see [21–25] and the references therein. Various extensions of the ISS property have been made for different types of dynamical systems, such as discrete time systems [21], time-delay systems [22], impulsive system [23], and switched systems [24–26]. Many works about the ISSiM of SSN systems have been done, but this problem has not been solved completely so far. Thus, investigating the ISSiM of the SSN systems is not only very important in theory but is also reasonable in practice.

In this paper, we present several new sufficient conditions under which a SSN system with an improved ADT switching signal is ISSiM, also examine the case where parts of the constituent subsystems are not ISSiM. First, we introduce an improved ADT method, and by which we present a new sufficient condition for the SSN system whose subsystems are ISSiM. Then, we obtain some new ISSiM results for such switched systems that parts of the constituent subsystems are ISSiM. Finally, an illustrative example with numerical simulation is studied using the above obtained results. The study of example shows that our analysis methods work very well in analyzing the ISSiM of SSN systems.

The rest of the paper is organized as follows. Section 2 introduces some notations and preliminary results which are used in this paper. Section 3 presents the main results of this paper. In Section 4, an illustrative example with numerical simulation is given to support our new results, which is followed by the conclusion in Section 5.

2. Notations and Preliminary Results

Throughout this paper, $\mathbb{R}_+$ denotes the set of all nonnegative real numbers; $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote n-dimensional real space and $n \times m$ dimensional real matrix space, respectively. For vector $x \in \mathbb{R}^n$, $|x|$ denotes the Euclidean norm; that is, $|x| = (\sum_{i=1}^n x_i^2)^{1/2}$. All the vectors are column vectors unless otherwise specified; the transpose of vectors and matrices are denoted by superscript T; $\mathscr{C}^i$ denotes all the ith continuous differential functions. A function $\varphi(u)$ is said to belong to the class $\mathscr{K}$ if $\varphi \in \mathscr{C}(\mathbb{R}_+, \mathbb{R}_+)$, $\varphi(0) = 0$ and $\varphi(u)$ is strictly increasing in u. $\mathscr{K}_\infty$ is the subset of $\mathscr{K}$ functions that are unbounded.

Consider the following SSN systems:

$$dx = f_\sigma(x,u)\,dt + g_\sigma(x,u)\,dw, \quad t \geq 0, \tag{1}$$

where $x \in \mathbb{R}^n$ and $u \in \mathscr{L}_\infty^m$ are the system state and input, respectively; $\mathscr{L}_\infty^m$ denotes the set of all the measurable and locally essentially bounded input $u \in \mathbb{R}^m$ on $[0, \infty)$ with norm

$$\|u(t)\| = \inf_{\mathscr{A}\subset\Omega, P(\mathscr{A})=0} \sup\{|u(t,\omega)| : \omega \in \Omega \setminus \mathscr{A}\}. \tag{2}$$

w is an r-dimensional independent standard Wiener process (or Brownian motion); $\sigma(\cdot) : [0,\infty) \to \mathscr{I}$ ($\mathscr{I}$ is the index set, maybe infinite) is the switching path (or law, signal) and is right-hand continuous and piecewise constant on t. More specifically, we impose restrictions on the set of admissible switching signals by defining the set

$$D_T = \{\sigma(t) : t_{k+1} - t_k \geq T\}, \tag{3}$$

where t_k are the commutation instants and $T \geq 0$. For every $i \in \mathscr{I}$, $f_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $g_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times r}$ is continuous, uniformly locally Lipschitz, and satisfies $f_i(0,0) = g_i(0,0) = 0$; initial data $x_0 \in \mathbb{R}^n$. For an arbitrary matrix D, we define $|D| = [\lambda_M(D^T D)]^{1/2}$, where λ_M denotes the largest eigenvalue of $D^T D$.

For any given $V(x) \in C^2(\mathbb{R}^n; \mathbb{R}_+)$, associated with the SSN system (1), we define the differential operator $\mathscr{L}$ to every $i \in \mathscr{I}$ as follows:

$$\mathscr{L}V = \frac{\partial V}{\partial x} f_i(x,u) + \frac{1}{2}\operatorname{Tr}\left\{g_i^T(x,u)\frac{\partial^2 V}{\partial x^2} g_i(x,u)\right\}. \tag{4}$$

With the development of this paper, we first present some definitions.

Definition 1 (see [15]). For any switching signal $\sigma(t)$ and any $t \geq \tau$, let $N_\sigma(\tau,t)$ denote the number of switching of $\sigma(t)$ over the interval $[\tau, t)$ satisfying

$$N_\sigma(\tau,t) \leq N_0 + \frac{t-\tau}{\tau_a}, \tag{5}$$

where τ_a is called average dwell time and N_0 denotes the chatter bound.

Definition 2 (see [27]). The SSN system (1) is ISSiM if there exist $\beta \in \mathscr{KL}$ and $\alpha, \gamma \in \mathscr{K}_\infty$, such that for any $u \in \mathbb{R}^m$, $x_0 \in \mathbb{R}^n$, we have

$$E[\alpha(|x(t)|)] \leq \beta(|x_0|, t) + \gamma(\|u\|_{[0,t)}), \quad \forall t \geq 0. \tag{6}$$

The SSN system (1) is $e^{\lambda t}$-weighted ISSiM for some $\lambda > 0$; if there exist $\alpha_1, \alpha_2, \gamma \in \mathscr{K}_\infty$ such that for any $u \in \mathbb{R}^m$, $x_0 \in \mathbb{R}^n$, we have

$$\begin{aligned} &e^{\lambda t} E[\alpha_1(|x(t)|)] \\ &\quad \leq \alpha_2(|x_0|) + \sup_{s\in[0,t)}\{e^{\lambda s}\gamma(\|u(s)\|)\}, \quad \forall t \geq 0. \end{aligned} \tag{7}$$

The SSN system (1) is $e^{\lambda t}$-weighted integral ISSiM for some $\lambda > 0$; if there exist $\alpha_1, \alpha_2, \gamma \in \mathscr{K}_\infty$ such that for any $u \in \mathbb{R}^m$, $x_0 \in \mathbb{R}^n$, we have

$$\begin{aligned} & e^{\lambda t} E\left[\alpha_1(|x(t)|)\right] \\ & \leq \alpha_2\left(\left|x_0\right|\right)+\int_0^t e^{\lambda \tau} \gamma(\|u(\tau)\|) d \tau, \quad \forall t \geq 0. \end{aligned} \tag{8}$$

In [28], we have obtained an improved ADT method to investigate the stability of the SSN system (1) with $u \equiv 0$ in two cases: one is that all constituent subsystems are globally exponentially stable in the mean (GASiM) and the other is that some constituent subsystems are GASiM, while some of them are not GASiM. We introduce those results in the following.

Lemma 3 (see [28]). *For the SSN system* (1) *with $u \equiv 0$, if there exist $\mathscr{C}^1$ functions $V_i : \mathbb{R}^n \rightarrow [0, \infty), i \in \mathscr{I}$, and functions $\alpha, \beta \in \mathscr{K}_\infty$ such that*

$$\alpha(|x|) \leq V_i(x) \leq \beta(|x|), \quad \forall i \in \mathscr{I}, \tag{9}$$

$$\mathscr{L} V_i(x)\big|_{(i)} \leq -\lambda_i V_i(x), \tag{10}$$

where $\lambda_i > 0, i \in \mathscr{I}$, then the SSN system (1) *with $u \equiv 0$ is GASiM under any switching signal with ADT:*

$$\tau_a > \tau_a^* = \frac{a}{\lambda_{\min}}, \tag{11}$$

where

$$a = \ln \mu, \qquad \mu = \sup_{x \neq 0} \frac{\beta(|x(t)|)}{\alpha(|x(t)|)}, \qquad \lambda_{\min} = \min_{i \in \mathscr{I}} \lambda_i. \tag{12}$$

Remark 4. If $\mu = 1$, which implies that $V_i(x) \equiv V(x)$, $i \in \mathscr{I}$, $V(x)$ is a CLF for the SSN system (1) with $u \equiv 0$, and thus this system is GASiM under arbitrary switching. It is also noted that the ADT method proposed in [15–18] needs the conditions (9)-(10) and an additional condition as "$V_i(x) \leq \mu V_j(x)$, $\mu \geq 1, i \neq j, i, j \in \mathscr{I}$". Comparing Lemma 3 with the corresponding results in [15, 16], Lemma 3 needs fewer conditions and thus can be applied to a wider range of systems.

Moreover, it is noted that the above α and β in (9) should have the same order, and which can ensure that μ exists. Furthermore, if $V_i(x) = x^T P_i x, P_i > 0$, then inequality (9) becomes

$$\alpha_i |x|^2 \leq V_i(x) \leq \beta_i |x|^2 \tag{13}$$

and μ is given as

$$\mu = \max_{i \in \mathscr{I}} \frac{\beta_i}{\alpha_i}. \tag{14}$$

For this case, if we use the ADT method in [15–18], we can get

$$\mu' = \max_{i, j \in \mathscr{I}} \frac{\beta_i}{\alpha_j}. \tag{15}$$

Obviously, $\mu \leq \mu'$.

In particular, for the switched linear systems, the lower bound ADT τ_a^* obtained by Lemma 3 is smaller than the lower bound ADT τ_a' obtained in [15]; that is,

$$\tau_a^* = \max_{i \in \mathscr{I}} \left\{ \frac{\lambda_{\max}(P_i)}{\lambda_{\min}(P_i)\lambda_{\min}} \right\} \leq \max_{i, j \in \mathscr{I}} \left\{ \frac{\lambda_{\max}(P_j)}{\lambda_{\min}(P_i)\lambda_{\min}} \right\} = \tau_a', \tag{16}$$

where $V_i(x) = x^T P_i x$ with $P_i > 0$ is the Lyapunov function for the ith subsystem, $i \in \mathscr{I}$.

This improved ADT method can also be extended to analyze the stability of the SSN system (1) with $u \equiv 0$ in which both stable and unstable subsystems coexist. For the switching signal $\sigma(t)$ and any $t > \tau$, we let $T^u(\tau, t)$ (resp., $T^s(\tau, t)$) denote the total activation time of the unstable subsystems (resp., the stable subsystems) on interval $[\tau, t)$. Then, we let $\mathscr{I} = \mathscr{I}_s \cup \mathscr{I}_u$ such that $\mathscr{I}_s \cap \mathscr{I}_u = \emptyset$ and introduce a switching law form [16].

(S1) Determine the $\sigma(t)$ satisfying $T^s(t_0, t)/T^u(t_0, t) \geq (\lambda_u + \lambda^*)/(\lambda_s - \lambda^*)$ holds for any $t > t_0$, where $\lambda^* \in (0, \lambda_s)$; λ_u and λ_s are given as (19).

Next, we introduce the result in the following.

Lemma 5 (see [28]). *Consider the SSN system* (1) *with $u \equiv 0$; if there exist $\mathscr{C}^1$ functions $V_i(x)$: $\mathbb{R}^n \rightarrow [0, \infty)$, $i \in \mathscr{I}$, and functions α, $\beta \in \mathscr{K}_\infty$ such that* (9), *and*

$$\begin{aligned} & \mathscr{L} V_i(x)\big|_{(i)} \leq \lambda_i V_i(x), \quad i \in \mathscr{I}_u, \\ & \mathscr{L} V_i(x)\big|_{(i)} \leq -\lambda_i V_i(x), \quad i \in \mathscr{I}_s, \end{aligned} \tag{17}$$

where $\lambda_i > 0$, $i \in \mathscr{I}$, then under the switching law S1, the switched system (1) *with $u \equiv 0$ is GASiM for any switching signal with ADT:*

$$\tau_a > \tau_a^* = \frac{a}{\lambda^*}, \tag{18}$$

where a is given as (12), *and $\lambda^* \in (0, \lambda_s)$ is an arbitrarily chosen number,*

$$\lambda_s = \max_{i \in \mathscr{I}_s} \lambda_i, \qquad \lambda_u = \max_{i \in \mathscr{I}_u} \lambda_i. \tag{19}$$

Remark 6. Similar to Remark 4, comparing Lemma 5 with the corresponding existing results in [15, 16], Lemma 5 needs fewer conditions, and thus Lemma 5 is really an improvement of the existing results.

3. Main Results

3.1. All Subsystems Are ISSiM. In this section, we first investigate the ISSiM stability of the SSN system (1) in which all constituent subsystems are ISSiM. According to Lemma 3, we obtain the following result.

Theorem 7. *Considering the SSN system* (1), *if there exist* $\mathscr{C}^1$ *functions* $V_i : \mathbb{R}^n \to [0,\infty)$, $i \in \mathscr{I}$, *functions* $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\gamma} \in \mathscr{K}_\infty$ *and number* $\lambda_0 > 0$ *such that for all* $\xi \in \mathbb{R}^n$, $\eta \in \mathbb{R}^{\mathscr{L}}$,

$$\bar{\alpha}_1(|\xi|) \le V_i(\xi) \le \bar{\alpha}_2(|\xi|), \quad (20)$$

$$\mathscr{L}V_i(\xi)\big|_{(i)} \le -\lambda_0 V_i(\xi) + \bar{\gamma}(|\eta|), \quad i \in \mathscr{I}, \quad (21)$$

then

(I) *if* $\tau_a > \tau_a^* = a/\lambda_0$, *then the SSN system* (1) *is ISSiM;*

(II) *if* $\tau_a > \tau_a^* = a/(\lambda_0 - \lambda - \delta)$, *for some* $\lambda \in (0, \lambda_0 - \delta)$, *where* $\delta > 0$, *then the SSN system* (1) *is* $e^{\lambda t}$*-weighted ISSiM;*

(III) *if* $\tau_a > \tau_a^* = a/(\lambda_0 - \lambda)$, *for some* $\lambda \in (0, \lambda_0)$, *then the SSN system* (1) *is* $e^{\lambda t}$*-weighted integral ISSiM,*

where a is given as (12).

Proof. For notational brevity, define $G_a^b(\lambda) = \int_a^b e^{\lambda s}\bar{\gamma}(\|u(s)\|)ds$. Let $T > 0$ be an arbitrary time. Denote by $t_1, \ldots, t_{N_\sigma(0,T)}$ the switching times on the interval $(0,T)$ (by convention, $t_0 := 0$, $t_{N_\sigma(0,T)} := T$). Consider the piecewise continuously differentiable function

$$W(s) := e^{\lambda_0 s} V_{\sigma(s)}(x(s)). \quad (22)$$

On each interval $[t_i, t_{i+1})$, the switching signal is constant. Consider

$$\begin{aligned} W(t) = W(t_i) &+ \int_{t_i}^t e^{\lambda_0 s}\frac{\partial V_{\sigma(t_i)}}{\partial x} g_{\sigma(s)}(s, x(s))\,dw(s) \\ &+ \int_{t_i}^t e^{\lambda_0 s}\left(\mathscr{L}V_{\sigma(s)}(x(s)) + \lambda_0 V_{\sigma(s)}(x(s))\right)ds. \end{aligned} \quad (23)$$

If t is replaced by $t_r = \min\{t, \tau_r\}$ in the above, where $\tau_r = \inf\{s \ge 0 : |x(s)| \ge r\}$, then the stochastic integral (first integral) in (23) defines a martingale (with r fixed and t varying), not just a local martingale. Thus, on taking expectations in (23) with t_r in place of t and then using (21) on the right, we get

$$EW(t_r) \le EW(t_i) + E\left[\int_{t_i}^{t_r} e^{\lambda_0 s}\bar{\gamma}(\|u(s)\|)\,ds\right]. \quad (24)$$

On letting $r \to \infty$ and using Fatou's Lemma on the left and monotone convergence on the right, we conclude

$$EW(t) \le EW(t_i) + \int_{t_i}^t e^{\lambda_0 s}\bar{\gamma}(\|u(s)\|)\,ds. \quad (25)$$

According to inequality (20), we obtain

$$\begin{aligned} &E\bar{\alpha}_1(|x_T|) \\ &\le EV_{\sigma(t_k)}(x_T) \le e^{-\lambda_0(T-t_k)}EV_{\sigma(t_k)}(x_k) \\ &\quad + e^{-\lambda_0 T}G_{t_k}^T(\lambda_0) \\ &\le e^{-\lambda_0(T-t_k)}E\bar{\alpha}_2(|x_k|) + e^{-\lambda_0 T}G_{t_k}^T(\lambda_0) \\ &= e^{-\lambda_0(T-t_k)}\frac{E\bar{\alpha}_2(|x_k|)}{E\bar{\alpha}_1(|x_k|)}E\bar{\alpha}_1(|x_k|) + e^{-\lambda_0 T}G_{t_k}^T(\lambda_0) \\ &\le \mu e^{-\lambda_0(T-t_k)}E\bar{\alpha}_1(|x_k|) + e^{-\lambda_0 T}G_{t_k}^T(\lambda_0) \\ &\le \mu^2 e^{-\lambda_0(T-t_{k-1})}E\bar{\alpha}_1(|x_{k-1}|) + \mu e^{-\lambda_0 T}G_{t_{k-1}}^{t_k}(\lambda_0) \\ &\quad + e^{-\lambda_0 T}G_{t_k}^T(\lambda_0) \\ &\vdots \\ &\le \mu^{N_{\sigma(t)}(t_0,T)+1}e^{-\lambda_0(T-t_0)}E\bar{\alpha}_1(|x_0|) \\ &\quad + e^{-\lambda_0 T}\sum_{j=0}^{N_{\sigma(t)}(t_0,T)}\mu^{N_{\sigma(t)}(t_0,T)-j}G_{t_j}^{t_{j+1}}(\lambda_0) \\ &= \mu e^{aN_{\sigma(t)}(t_0,T)-\lambda_0(T-t_0)}E\bar{\alpha}_1(|x_0|) \\ &\quad + \sum_{j=0}^{N_{\sigma(t)}(t_0,T)}\mu^{N_{\sigma(t)}(t_0,T)-j}G_{t_j}^{t_{j+1}}(\lambda_0) \\ &\le \mu^{1+N_0}e^{(a/\tau_a-\lambda_0)(T-t_0)}E\bar{\alpha}_1(|x_0|) \\ &\quad + \sum_{j=0}^{N_{\sigma(t)}(t_0,T)}\mu^{N_{\sigma(t)}(t_0,T)-j}G_{t_j}^{t_{j+1}}(\lambda_0). \end{aligned} \quad (26)$$

For every $\delta \in [0, \lambda_0 - \lambda - a/\tau_a^*]$, that is, $\tau_a^* \ge a/(\lambda_0 - \lambda - \delta)$, where $\lambda > 0$,

$$\begin{aligned} G_{t_j}^{t_{j+1}}(\lambda_0) &\le e^{(\lambda_0-\lambda-\delta)t_{k+1}}G_{t_j}^{t_{j+1}}(\lambda+\delta), \\ \mu^{N_{\sigma(t)}(t_0,T)-j} &\le \mu^{1+N_0}e^{(\lambda_0-\lambda-\delta)(T-t_{j+1})}. \end{aligned} \quad (27)$$

Therefore,

$$\begin{aligned} &\sum_{j=0}^{N_{\sigma(t)}(t_0,T)}\mu^{N_{\sigma(t)}(t_0,T)-j}G_{t_j}^{t_{j+1}}(\lambda_0) \\ &\qquad \le \mu^{1+N_0}e^{(\lambda_0-\lambda-\delta)T}G_{t_0}^T(\lambda+\delta). \end{aligned} \quad (28)$$

Substituting inequality (28) to inequality (26), we get

$$\begin{aligned} E\bar{\alpha}_1(|x_T|) \le \mu^{1+N_0}&\left[e^{(a/\tau_a-\lambda_0)(T-t_0)}\bar{\alpha}_1(|x_0|)\right. \\ &\left.+ e^{-(\lambda+\delta)T}G_{t_0}^T(\lambda+\delta)\right], \end{aligned} \quad (29)$$

that is,

$$e^{(\lambda+\delta)T}E\,\bar{\alpha}_1\left(\left|x_T\right|\right) \le \mu^{1+N_0}\left[e^{\lambda+\delta t_0}e^{((a/\tau_a)-\lambda_0+\lambda+\delta)(T-t_0)} \times \bar{\alpha}_1\left(\left|x_0\right|\right) + G_{t_0}^{T}(\lambda+\delta)\right] = C\left[e^{(\lambda+\delta)t_0}e^{(a/\tau_a-\lambda_0+\lambda+\delta)(T-t_0)} \times \bar{\alpha}_1\left(\left|x_0\right|\right) + G_{t_0}^{T}(\lambda+\delta)\right], \tag{30}$$

where $C = \mu^{1+N_0}$.

For inequality (30), if $\delta = 0$, then

$$\begin{aligned} e^{\lambda T}E\,\bar{\alpha}_1\left(\left|x_T\right|\right) &\le Ce^{\lambda t_0}e^{(a/\tau_a-\lambda_0+\lambda)(T-t_0)}\bar{\alpha}_1\left(\left|x_0\right|\right) + CG_{t_0}^{T}(\lambda) \\ &= Ce^{\lambda t_0}e^{(a/\tau_a-\lambda_0+\lambda)(T-t_0)}\bar{\alpha}_1\left(\left|x_0\right|\right) + C\int_{t_0}^{T} e^{\lambda s}\bar{\gamma}\left(\left|u(s)\right|\right)ds. \end{aligned} \tag{31}$$

For inequality (31), if $\tau_a > a/(\lambda_0 - \lambda)$, we have property (8) with

$$\alpha_1\left(\left|x_T\right|\right) := \bar{\alpha}_1\left(\left|x_T\right|\right), \quad \alpha_2\left(\left|x_0\right|\right) := Ce^{\lambda t_0}e^{(a/\tau_a-\lambda_0+\lambda)(T-t_0)}\bar{\alpha}_1\left(\left|x_0\right|\right). \tag{32}$$

Note that

$$G_{t_0}^{T}(\lambda+\delta) \le \frac{1}{\lambda+\delta-\bar{\lambda}}e^{(\lambda+\delta-\bar{\lambda})T}\sup_{\tau\in[0,T)}\left\{e^{\bar{\lambda}\tau}\bar{\gamma}\left(\left|u(\tau)\right|\right)\right\}. \tag{33}$$

Substituting inequality (33) to inequality (30), we obtain

$$\begin{aligned} e^{(\lambda+\delta)T}E\,\bar{\alpha}_1\left(\left|x_T\right|\right) &\le Ce^{(\lambda+\delta)t_0}e^{(a/\tau_a-\lambda_0+\lambda+\delta)(T-t_0)}\bar{\alpha}_1\left(\left|x_0\right|\right) \\ &\quad + C\frac{1}{\lambda+\delta-\bar{\lambda}}e^{(\lambda+\delta-\bar{\lambda})T}\sup_{\tau\in[0,T)}\left\{e^{\bar{\lambda}\tau}\bar{\gamma}\left(\left|u(\tau)\right|\right)\right\}. \end{aligned} \tag{34}$$

For inequality (34), if $\delta \ne 0$, $\bar{\lambda} = \lambda$, we get

$$\begin{aligned} e^{\lambda T}E\,\bar{\alpha}_1\left(\left|x_T\right|\right) &\le Ce^{\lambda t_0}e^{(a/\tau_a-\lambda_0+\lambda+\delta)(T-t_0)} \\ &\quad \times E\,\bar{\alpha}_1\left(\left|x_T\right|\right) + C\frac{1}{\delta}e^{\delta T}\sup_{\tau\in[0,T)}\left\{e^{\bar{\lambda}\tau}\bar{\gamma}\left(\left|u(\tau)\right|\right)\right\}. \end{aligned} \tag{35}$$

For inequality (35), if $\tau_a > a/(\lambda_0-\lambda-\delta)$, we have property (7) with

$$\alpha_1\left(\left|x_T\right|\right) := e^{-\delta T}\frac{\delta}{C}\bar{\alpha}_1\left(\left|x_T\right|\right), \quad \alpha_2\left(\left|x_0\right|\right) := e^{-\delta T}\delta e^{\lambda t_0}e^{(a/\tau_a-\lambda_0+\lambda+\delta)(T-t_0)}\bar{\alpha}_1\left(\left|x_0\right|\right). \tag{36}$$

For inequality (34), if $\bar{\lambda} = \delta = 0$, we obtain that

$$\bar{\alpha}_1\left(\left|x_T\right|\right) \le Ce^{\lambda t_0}e^{(a/\tau_a-\lambda_0)(T-t_0)}\bar{\alpha}_1\left(\left|x_T\right|\right) + C\frac{1}{\lambda}\sup_{\tau\in[0,T)}\left\{\bar{\gamma}\left(\left|u(\tau)\right|\right)\right\}. \tag{37}$$

For inequality (37), if $\tau_a > a/(\lambda_0 - \lambda)$, we have property (6) with

$$\alpha_1\left(\left|x_T\right|\right) := Ce^{\lambda t_0}e^{(a/\tau_a-\lambda_0)(T-t_0)}\bar{\alpha}_1\left(\left|x_T\right|\right), \qquad \gamma := C\frac{1}{\lambda}\bar{\gamma}. \tag{38}$$

□

Remark 8. It should be pointed out that the result proposed in [25] needs conditions (20)-(21) and an additional condition as "$V_i(x) \le \mu V_j(x)$, $\mu \ge 1$, $i \ne j$, i, $j \in \mathcal{I}$". Comparing Theorem 7 with the existing result in [25], Theorem 7 needs fewer conditions and thus is an improvement of the existing result.

3.2. Some Subsystems Are Not ISSiM. In the next, we consider the SSN system (1) in which both ISSiM and not ISSiM subsystems coexist. Similarly, for the switching signal $\sigma(t)$ and any $t > \tau$, we let $T^u(\tau,t)$ (resp., $T^s(\tau,t)$) denote the total activation time of the not ISSiM subsystems (resp., the ISSiM subsystems) on interval $[\tau, t)$.

According to Lemma 5, we give the following result.

Theorem 9. *Considering the SSN system* (1), *if there exist* $\mathscr{C}^1$ *functions* $V_i : \mathbb{R}^n \to [0,\infty)$, $i \in \mathcal{I}$, *and functions* α_1, α_2, $\varphi_1 \in \mathscr{K}_\infty$, *constants* λ_s, $\lambda_u > 0$ *such that* (20) *for all* $x \in \mathbb{R}^n$, *and furthermore, the following inequalities hold:*

$$|x| \ge \varphi_1(u) \Longrightarrow \begin{cases} \mathscr{L}V_i(x)\big|_{(i)} \le \lambda_u V_i(x), & i \in \mathcal{I}_u, \\ \mathscr{L}V_i(x)\big|_{(i)} \le -\lambda_s V_i(x), & i \in \mathcal{I}_s. \end{cases} \tag{39}$$

Then, under the switching law S1, the SSN system (1) *is ISSiM for any switching signal with ADT:*

$$\tau_a > \tau_a^* = \frac{a}{\lambda^*}, \tag{40}$$

where a is given as (12), *and* $\lambda^* \in (0,\lambda_s)$ *is an arbitrarily chosen number;* λ_s *and* λ_u *are given as* (19).

Proof. Let $t_0 \ge 0$ be arbitrary. For $t \ge t_0$, define $\nu(t) := \varphi_1(\|u\|_{[t_0,t]})$ and $\xi(t) := \alpha_1^{-1}(\mu^{N_0}\alpha_1(\nu(t)))$, where N_0 comes from (5). Furthermore, define the balls around the origin $B_\nu(t) := \{x : |x| \le \nu(t)\}$, $B_\xi(t) := \{x : |x| \le \xi(t)\}$. Note that ν and thus also ξ are nondecreasing functions of time, and thus the balls B_ν and B_ξ have nondecreasing volume.

If $|x(t)| \ge \nu(t) \ge \varphi_1(\|u(t)\|)$ during some time interval $t \in [t', t'']$, then $x(t)$ can be bounded above by

$$E\left|x(t)\right| \le E\alpha_1^{-1}\left(\mu^{N_0}e^{-\lambda^*(t-t')}\alpha_1\left(\left|x\left(t'\right)\right|\right)\right) := \beta\left(\left|x\left(t'\right)\right|, t-t'\right) \tag{41}$$

for some $\lambda^* \in (0,\lambda_s)$.

In fact, on any interval $[\tau_i, \tau_{i+1}) \cap [t', t'']$, according to (39), we arrive at

$$E\alpha_1(|x(t)|) \le e^{aN_\sigma(t',t)+\lambda_u T^u(t',t)-\lambda_s T^s(t',t)} E\alpha_1\left(\left|x\left(t'\right)\right|\right). \tag{42}$$

Then, according to (5) and the switching law S1, we conclude from (42) that

$$\begin{aligned} E\alpha_1(|x(t)|) &\le e^{aN_0} e^{(a/\tau_a - \lambda^*)(t-t')} E\alpha_1\left(\left|x\left(t'\right)\right|\right) \\ &= \mu^{N_0} e^{(a/\tau_a-\lambda^*)(t-t')} E\alpha_1\left(\left|x\left(t'\right)\right|\right). \end{aligned} \tag{43}$$

Thus, if $\tau_a > \tau_a^* = a/\lambda^*$, we can get (41).

Denote the first time when $x(t) \in B_\nu(t)$ by $\check{t}_1$; that is, $\check{t}_1 := \inf\{t \ge t_0 : |x(t)| \le \nu(t)\}$. For $t_0 \le t \le \check{t}_1$, according to (41), we obtain

$$E|x(t)| \le \beta\left(\left|x_0\right|, t - t_0\right). \tag{44}$$

If $\check{t}_1 = \infty$, which only can happen if $\nu(t) \equiv 0$, that is, the input $u \equiv 0$ for all times, then the SSN system (1) is ISSiM. Hence in the following we assume that $\check{t}_1 < \infty$.

For $t > \check{t}_1$, $x(t)$ can be bounded above in terms of $\nu(t)$. Namely, let $\hat{t}_1 := \inf\{t > \check{t}_1 : |x(t)| > \nu(t)\}$. If this is an empty set, let $\hat{t}_1 := \infty$. Clearly, for all $t \in [\check{t}_1, \check{t}_2)$, it holds that $|x(t)| \le \nu(t) \le \xi(t)$. For the case that $\hat{t}_1 < \infty$, due to continuity of $x(\cdot)$ and monotonicity for $\nu(t)$, it holds that $|x(\hat{t}_1)| = \nu(\hat{t}_1)$. Furthermore, for all $\tau > \hat{t}_1$, if $|x(\tau)| > \nu(\tau)$, define

$$\hat{t} := \sup\{t < \tau : |x(t)| \le \nu(t)\} \tag{45}$$

which can be interpreted as the previous exit time of the trajectory $x(t)$ from the ball B_ν. Again, due to the same argument as above, one obtains that $E|x(\hat{t}_1)| = \nu(\hat{t}_1)$. But then, according to (41), it holds that

$$\begin{aligned} E|x(\tau)| &\le \beta\left(\nu(\hat{t}), \tau - \hat{t}\right) = E\alpha_1^{-1}\left(\mu^{N_0} e^{-\lambda^*(\tau-\hat{t})}\alpha_1\left(\nu\left(\hat{t}\right)\right)\right) \\ &\le E\alpha_1^{-1}\mu^{N_0} E\alpha_1\left(\nu\left(\hat{t}\right)\right) = \xi\left(\hat{t}\right) \le \xi(\tau). \end{aligned} \tag{46}$$

Summarizing the above, for all $t \ge \check{t}_1$, it holds that

$$\begin{aligned} E|x(t)| &\le \xi(t) = E\alpha_1^{-1}\left(\mu^{N_0}\alpha_1\left(\varphi_1\left(\|u\|_{[t_0,t]}\right)\right)\right) \\ &\le E\alpha_1^{-1}\mu^{N_0}\alpha_1\left(2\varphi_1\left(\|u\|_{[t_0,t]}\right)\right) := \gamma_1\left(\|u\|_{[t_0,t]}\right). \end{aligned} \tag{47}$$

Combining (44) and (47), we obtain that

$$E|x(t)| \le \beta\left(\left|x_0\right|, t - t_0\right) + \gamma_1\left(\|u\|_{[t_0,t]}\right) \tag{48}$$

for all $t \ge t_0$, which means the SSN system (1) is ISSiM and also completes the proof. □

4. An Illustrative Example

In this section, we give an illustrative example to show how to use the obtained results to analyze the ISSiM stability of SSN systems.

Example 1. Consider the following SSN system:

$$dx = f_i(x, u)\,dt + g_i(x)\,dw(t), \tag{49}$$

where $i \in \mathcal{I} = \{1, 2\}$, w is an r-dimensional standard Brownian motion, and

$$\begin{gathered} f_1(x,u) = \begin{pmatrix} -x_1 - x_1x_2^2 - x_1\sin^2 t \\ x_1^2x_2 - 3x_2 - x_2\cos^2 t \end{pmatrix}, \\ f_2(x,u) = \begin{pmatrix} 2x_1 + 2x_2 \\ x_1 + 3x_2 \end{pmatrix}, \\ g_1(x) = \begin{pmatrix} \dfrac{1}{2}x_1 - \dfrac{1}{2}x_2 \\ -\dfrac{1}{2}x_1 + \dfrac{1}{2}x_2 \end{pmatrix}, \\ g_2(x) = \begin{pmatrix} -\dfrac{1}{2}x_1 \\ \dfrac{\sqrt{3}}{2}x_2 \end{pmatrix}. \end{gathered} \tag{50}$$

It is easy to know that $V(x) = x^T x$ is a CLF for the system (49), and

$$\begin{aligned} \mathscr{L}V_1(x)\big|_{(1)} &= \frac{\partial V}{\partial x} f_i(x) + \frac{1}{2}\operatorname{Tr}\left\{g_i^T(x)\frac{\partial^2 V}{\partial x^2} g_i(x)\right\} \\ &= -2x_1^2 - 6x_2^2 \le -2V_1(x), \\ \mathscr{L}V_2(x)\big|_{(2)} &= \frac{\partial V}{\partial x} f_i(x) + \frac{1}{2}\operatorname{Tr}\left\{g_i^T(x)\frac{\partial^2 V}{\partial x^2} g_i(x)\right\} \\ &= 2x_1^2 + 4x_1x_2 + 3x_2^2 \le 5V_2(x). \end{aligned} \tag{51}$$

According to the above results, we obtain that $\lambda_u = 5$, $\lambda_s = 2$ and $a = 0$. Therefore, the lower bound ADT $\tau_a^* = 0$; that is, the ADT can be arbitrary. Next, we choose $\lambda^* = 0.1$; then the switching law S1 will require

$$\frac{T^s(t_0,t)}{T^u(t_0,t)} \ge \frac{\lambda_u + \lambda^*}{\lambda_s - \lambda^*} = \frac{5.1}{1.9} \approx 2.68. \tag{52}$$

According to Theorem 9, the switched system (49) is ISSiM under the above switching law S1.

To illustrate the correctness of the above conclusion, we carry out some simulation results with the following choices. Initial condition $[x_1(0), x_2(0)] = [-2.5, 3]$, and switching path

$$\sigma(t) = \begin{cases} 2, & t \in \left[t_{2m}, t_{2m+1}\right),\ t_{2m+1} - t_{2m} = 0.2 * \text{rand}, \\ 1, & t \in \left[t_{2m+1}, t_{2m+2}\right),\ t_{2m+2} - t_{2m+1} = 0.6 + 0.1 * \text{rand}, \end{cases} \tag{53}$$

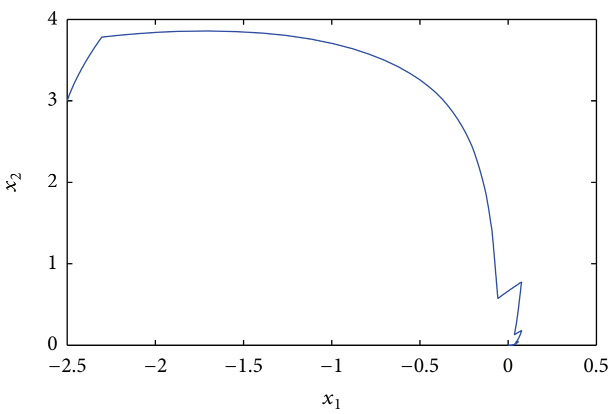


FIGURE 1: The state's response.

where $m = 0, 1, 2, \ldots$, rand $\in (0, 1)$ is a stochastic number. The simulation result is given in Figure 1, which is the response of the state under the above path $\sigma(t)$.

It can be observed from Figure 1 that the trajectory $x(t)$ converges to origin quickly. The simulation shows that Theorem 9 is very effective in analyzing the stability for the SSN systems with both unstable and ISSiM subsystems.

5. Conclusions

In this paper, we have investigated the ISSiM property of a class of SSN systems under ADT switching signals in two cases: (i) all of the constituent subsystems are ISSiM and (ii) parts of constituent subsystems are ISSiM and then proposed several new results about ISSiM of such systems. Firstly, a new ISSiM result for the SSN systems whose constituent subsystems are ISSiM has been obtained by applying an improved ADT method. Secondly, a new ISSiM result for the SSN system in which parts of subsystems are ISSiM has been given. In comparison with the existing results, the main results obtained in this paper have some advantages in some cases. Finally, an illustrative example with simulation has verified the validity and correctness of our results.

Conflict of Interests

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence their work; there is no professional or other personal interest of any nature or kind in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, this paper.

Acknowledgments

The paper was supported by the National Nature Science Foundation of China [61034007, 11226154, 61374074, 51107069, 61304133, and 61305130], China Postdoctoral Science Foundation funded Project [2013M541915 and 2013M541912], the Scientific Research Foundation of Shandong province Outstanding Young Scientist Award [BS2013SF023], and the Natural Science Foundation of Shandong province [ZR2012AQ01 and, ZR2013FQ003].

References

[1] J. Daafouz, P. Riedinger, and C. Iung, "Stability analysis and control synthesis for switched systems: a switched Lyapunov function approach," *IEEE Transactions on Automatic Control*, vol. 47, no. 11, pp. 1883–1887, 2002.

[2] M. Goncalves, A. Megretski, and A. Dahleh, "Global analysis of piecewise linear systems using impact maps and surface Lyapunov functions," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 2089–2106, 2003.

[3] P. Dayawansa and F. Martin, "A converse Lyapunov theorem for a class of dynamical systems which undergo switching," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 751–760, 1999.

[4] D. Liberzon and R. Tempo, "Common Lyapunov functions and gradient algorithms," *IEEE Transactions on Automatic Control*, vol. 49, no. 6, pp. 990–994, 2004.

[5] H. Lin and J. Antsaklis, "Stability and stabilizability of switched linear systems: a survey of recent results," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 1–15, 2008.

[6] X. Zhao and X. Liu, "Improved results on stability of continuous-time switched positive linea systems," *Automatica*, vol. 50, no. 2, pp. 614–621, 2014.

[7] X. Zhao, H. Liu, and J. Zhang, "Multiple-mode observer design for a class of switched linear systems linear systems," *IEEE Transactions on Automation Science and Engineering*, no. 99, pp. 1–9, 2013.

[8] X. Zhao, P. Shi, and L. Zhang, "Asynchronously switched control of a class of slowly switched linear systems," *Systems & Control Letters*, vol. 61, no. 12, pp. 1151–1156, 2012.

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

[10] R. Guo and Y. Wang, "Stability analysis for a class of switched linear systems," *Asian Journal of Control*, vol. 14, no. 3, pp. 817–826, 2012.

[11] P. Zhao, W. Feng, and Y. Kang, "Stochastic input-to-state stability of switched stochastic nonlinear systems," *Automatica*, vol. 48, no. 10, pp. 2569–2576, 2012.

[12] W. Feng, J. Tian, and P. Zhao, "Stability analysis of switched stochastic systems," *Automatica*, vol. 47, no. 1, pp. 148–157, 2011.

[13] W. Ni, D. Cheng, and X. Hu, "Minimum dwell time for stability and stabilization of switched linear systems," in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA '08)*, pp. 4109–4115, 2008.

[14] R. Guo, P. Zhao, and K. Li, "Input-to-state stability for a class of nonlinear switched systems by a new average dwell time method," in *Proceeding of the 32th Chinese Control Conference*, pp. 860–865, 2013.

[15] J. Hespanha and A. Morse, "Stability of switched systems with average dwell-time," in *Proceedings of the 38th IEEE Conference on Decision & Control (CDC '99)*, vol. 3, pp. 2655–2660, 1999.

[16] G. Zhai, B. Hu, K. Yasuda, and A. Michel, "Stability analysis of switched systems with stable and unstable subsystems: an average dwell time approach," in *Proceedings of the American Control Conference*, pp. 200–204, 2000.

[17] X. Zhao, L. Zhang, P. Shi, and M. Liu, "Stability of switched positive linear systems with average dwell time switching," *Automatica*, vol. 48, no. 6, pp. 1132–1137, 2012.

[18] M. A. Müller and F. Allgöwer, "Improving performance in model predictive control: switching cost functionals under average dwell-time," *Automatica*, vol. 48, no. 2, pp. 402–409, 2012.

[19] E. Sontag and Y. Wang, "On characterizations of the input-to-state stability property," *Systems and Control Letters*, vol. 24, no. 5, pp. 351–359, 1995.

[20] E. Sontag, "Comments on integral variants of ISS," *Systems & Control Letters*, vol. 34, no. 1-2, pp. 93–100, 1998.

[21] Z. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.

[22] N. Yeganefar, P. Pepe, and M. Dambrine, "Input-to-state stability of time-delay systems: a link with exponential stability," *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1526–1531, 2008.

[23] W. Chen and W. Zheng, "Input-to-state stability and integral input-to-state stability of nonlinear impulsive systems with delays," *Automatica*, vol. 45, no. 6, pp. 1481–1488, 2009.

[24] W. Xie, C. Wen, and Z. Li, "Input-to-state stabilization of switched nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1111–1116, 2001.

[25] L. Vu, D. Chatterjee, and D. Liberzon, "Input-to-state stability of switched systems and switching adaptive control," *Automatica*, vol. 43, no. 4, pp. 639–646, 2007.

[26] P. Zhao, W. Feng, and Y. Kang, "Stochastic input-to-state stability of switched stochastic nonlinear systems," *Automatica*, vol. 48, no. 10, pp. 2569–2576, 2012.

[27] P. Zhao and Z. Wang, "Input-to-state stability of switched stochastic nonlinear systems," in *Proceedings of the 30th Chinese Control Conference (CCC '11)*, pp. 1404–1406, 2011.

[28] R. Guo, P. Zhao, and C. Zhang, "Stability for a class of switched stochastic continuous-time nonlinear systems by animproved average dwell time method," . In press.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 907982, 15 pages
http://dx.doi.org/10.1155/2014/907982

Research Article

Exponential Stability of Neutral Stochastic Functional Differential Equations with Two-Time-Scale Markovian Switching

Junhao Hu and Zhiying Xu

College of Mathematics and Statistics, South-Central University for Nationalities, Wuhan 430074, China

Correspondence should be addressed to Junhao Hu; junhaohu74@gmail.com

Received 22 December 2013; Accepted 16 January 2014; Published 16 March 2014

Academic Editor: Weihai Zhang

We develop exponential stability of neutral stochastic functional differential equations with two-time-scale Markovian switching modeled by a continuous-time Markov chain which has a large state space. To overcome the computational effort and the complexity, we split the large-scale system into several classes and lump the states in each class into one class by the different states of changes of the subsystems; then, we give a limit system to effectively "replace" the large-scale system. Under suitable conditions, using the stability of the limit system as a bridge, the desired asymptotic properties of the large-scale system with Brownian motion and Poisson jump are obtained by utilizing perturbed Lyapunov function methods and Razumikhin-type criteria. Two examples are provided to demonstrate our results.

1. Introduction

In many practical dynamical systems such as neural networks, computer aided design, population ecology, chemical process simulation, and automatic control, stochastic differential equations represent the class of important dynamics (see [1–4]). During the recent several years, the asymptotic properties of neutral stochastic functional differential equations have been investigated by many authors (see [5–14]). Mao [10, 11] gave the exponential stability of neutral stochastic functional differential equations by using the Razumikhin-type theorems. Zhou and Hu [14] used the same argument to discuss the exponential stability in pth moment of neutral stochastic functional differential equations and neutral stochastic functional differential equations with Markovian switching. Wu et al. [13] examined the almost sure robust stability of nonlinear neutral stochastic functional differential equations with infinite delay, including the exponential stability and the polynomial stability. Song and Shen [12] investigated the asymptotic behavior of neutral stochastic functional differential equations under the more general conditions than the classical linear growth condition. Chen et al. [5] considered the exponential stability in mean square moment of mild solution for impulsive neutral stochastic partial functional differential equations by employing the inequality technique. The attraction and quasi-invariant sets of neutral stochastic partial functional differential equations were also studied in the recent paper [9].

In this paper, we will consider neutral stochastic functional differential equations with two-time-scale Markovian switching modeled by a continuous-time Markov chain which has a large state space. The computational effort and the complexity become a real concern. To overcome the difficulties, we have devoted much effort to the modeling and analysis of such systems, in which one of the main ideas is to split a large-scale system into several classes and lump the states in each class into one state (see [3, 15–21]). Khasminskii et al. for the first time established the asymptotic properties of the Markov chain $r^{\varepsilon}(\cdot)$ by introducing a small parameter $\varepsilon > 0$ (see [22]). Yin and Zhang developed the method in their book [4] that a complicated system can be replaced by the corresponding limit system that has a much simpler structure. Motivated by the papers [16, 21], under suitable conditions, using the stability of the limit system as a bridge, we will study the exponential stability of neutral stochastic functional differential equations with Brownian motion and Poisson jump by utilizing perturbed Lyapunov function methods and Razumikhin-type criteria.

The remainder of this paper is organized as follows. In Section 2, we introduce some notations and notions needed in our investigation. In Section 3, we state our main results, that is, exponential stability of neutral stochastic functional differential equations with two-time-scale Markovian switching. The exponential stability for neutral stochastic functional differential equations driven by pure jumps is also discussed in Section 4. Finally, two examples are presented to justify and illustrate applications of the theory in Section 5.

2. Preliminaries

Throughout this paper, unless otherwise specified, let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e., it is increasing and right continuous and $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets). Let $W(t) = (W_1(t), \ldots, W_m(t))^T$ be an m-dimensional Brownian motion defined on the probability space. For $\tau > 0$, let $C([-\tau, 0]; \mathbb{R}^n)$ denote the family of continuous functions φ from $[-\tau, 0]$ to $\mathbb{R}^n$ with norm $\|\varphi\| = \sup_{-\tau\leq\theta\leq 0}|\varphi(\theta)|$, where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$. If A is a vector or matrix, its transpose is denoted by A^T, while its trace norm is denoted by $|A| = \sqrt{\operatorname{trace}(A^T A)}$. Denote by $C^b_{\mathscr{F}_0}([-\tau, 0]; \mathbb{R}^n)$ the family of all $\mathscr{F}_0$ measurable and bounded $C([-\tau, 0]; \mathbb{R}^n)$-valued random variables. For $p > 0$ and $t \geq 0$, denote by $L^p_{\mathscr{F}_t}([-\tau, 0]; \mathbb{R}^n)$ the family of all $\mathscr{F}_t$-measurable $C([-\tau, 0]; \mathbb{R}^n)$-valued random variables $\phi = \{\phi(\theta) : -\tau \leq \theta \leq 0\}$ such that $\sup_{-\tau\leq\theta\leq 0} E|\phi(\theta)|^p < \infty$. We will denote the indicator function of a set G by I_G.

Consider an n-dimensional neutral stochastic functional differential equation with Markovian switching as follows:

$$\begin{aligned} d\left[x(t) - D(x_t, r(t))\right] \\ = f(x_t, t, r(t))\, dt + g(x_t, t, r(t))\, dw(t), \end{aligned} \tag{1}$$

on $t \geq 0$ with initial data $x_0 = \xi \in C([-\tau, 0]; \mathbb{R}^n)$ and $x_t = x(t+\theta) : -\tau \leq \theta \leq 0$, which is regarded as a $C([-\tau, 0]; \mathbb{R}^n)$-valued stochastic process. Moreover, $f : C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \mathbb{S} \rightarrow \mathbb{R}^n$, $g : C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \mathbb{S} \rightarrow \mathbb{R}^{n\times m}$, $D : C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{S} \rightarrow \mathbb{R}^n$.

Let $r(t)$ $(t \geq 0)$ be a right-continuous Markov chain on the probability space taking values in a finite state space $\mathbb{S} = \{1, 2, \ldots, M\}$ with generator $\Gamma = (\gamma_{ij})_{M\times M}$ given by

$$\mathbb{P}\left\{r(t+\Delta) = j \mid r(t) = i\right\} = \begin{cases} \gamma_{ij}\Delta + \circ(\Delta), & \text{if } i \neq j, \\ 1 + \gamma_{ii}\Delta + \circ(\Delta), & \text{if } i = j, \end{cases} \tag{2}$$

where $\Delta > 0$. Here, $\gamma_{ij} \geq 0$ is the transition rate from i to j if $i \neq j$, while $\gamma_{ii} = -\sum_{i\neq j}\gamma_{ij}$.

We assume the Markov $r(\cdot)$ is independent of the Brownian motion $W(\cdot)$. It is well known that almost every sample path $r(\cdot)$ is a right-continuous step function with finite number of simple jumps in any finite subinterval of $\mathbb{R}_+ := [0, \infty)$. As a standing hypothesis, we assume that the Markov chain is irreducible. This is equivalent to the condition that, for any $i, j \in \mathbb{S}$, we can find $i_1, i_2, \ldots, i_k \in \mathbb{S}$, such that

$$\gamma_{i,i_1}\gamma_{i_1,i_2}\cdots\gamma_{i_k,j} > 0. \tag{3}$$

Then, Γ always has an eigenvalue 0. The algebraic interpretation of irreducibility is rank $(\Gamma) = M-1$. Under this condition, the Markov chain has a unique stationary distribution $\pi\Gamma = 0$, subject to $\sum_{j=1}^{M}\pi_j = 1$ and $\pi_j > 0$ for all $j \in \mathbb{S}$. For a real-valued function $\sigma(\cdot)$ defined on $\mathbb{S}$, we define

$$\Gamma\sigma(\cdot)(i) := \sum_{j\in\mathbb{S}}\gamma_{ij}\sigma(j) = \sum_{j\neq i}\gamma_{ij}\left(\sigma(j) - \sigma(i)\right), \tag{4}$$

for each $i \in \mathbb{S}$.

Let $C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}; \mathbb{R}_+)$ denote the family of all nonnegative functions $V(x, t, i)$ on $\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}$, which are continuously twice differentiable in x and once differentiable in t. If $V(x, t, i) \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}; \mathbb{R}_+)$, define an operator $\mathscr{L}V$ from $C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \mathbb{S}$ to $\mathbb{R}$ by

$$\begin{aligned} \mathscr{L}V(\varphi, t, i) = {} & V_t(\varphi(0) - D(\varphi, i), t, i) \\ & + V_x(\varphi(0) - D(\varphi, i), t, i) f(\varphi, t, i) \\ & + \frac{1}{2}\operatorname{trace}\left[g^T(\varphi, t, i)\right. \\ & \qquad \times V_{xx}(\varphi(0) - D(\varphi, i), t, i) \\ & \qquad \left.\times\, g(\varphi, t, i)\right] \\ & + \sum_{j=1}^{l}\gamma_{ij}V(\varphi(0) - D(\varphi, j), t, j), \end{aligned} \tag{5}$$

where

$$\begin{gathered} \varphi \in C([-\tau, 0]; \mathbb{R}^n), \quad V_t = \frac{\partial V(x, t, i)}{\partial t}, \\ V_x = \left(\frac{\partial V(x, t, i)}{\partial x_1}, \frac{\partial V(x, t, i)}{\partial x_2}, \ldots, \frac{\partial V(x, t, i)}{\partial x_n}\right), \\ V_{xx} = \left(\frac{\partial^2 V(x, t, i)}{\partial x_i \partial x_j}\right)_{n\times n}. \end{gathered} \tag{6}$$

For a parameter $\varepsilon > 0$, we rewrite the Markov chain $r(t)$ as $r^{\varepsilon}(t)$ and the generator Γ as Γ^{ε}. Γ^{ε} is given by

$$\Gamma^{\varepsilon} = \frac{1}{\varepsilon}\overline{\Gamma} + \widehat{\Gamma}, \tag{7}$$

where $\overline{\Gamma}/\varepsilon$ represents the fast varying motions and $\widehat{\Gamma}$ represents the slowly changing dynamics. Set $\Gamma^{\varepsilon} = (\gamma^{\varepsilon}_{ij})_{M\times M}$, $\overline{\Gamma} = (\overline{\gamma}_{ij})_{M\times M}$, and $\widehat{\Gamma} = (\widehat{\gamma}_{ij})_{M\times M}$. For the sake of simplicity, suppose that

$$\mathbb{S} = \mathbb{S}^1 \cup \mathbb{S}^2 \cup \cdots \cup \mathbb{S}^l, \tag{8}$$

with $\mathbb{S}^k = \{s_{k1}, \ldots, s_{kM_k}\}$, $M = M_1 + M_2 + \cdots + M_l$, and

$$\overline{\Gamma} = \operatorname{diag}\left(\overline{\Gamma}^1, \ldots, \overline{\Gamma}^l\right), \tag{9}$$

where $\overline{\Gamma}^k$ is a generator of a Markov chain taking values in $\mathbb{S}^k$, for every $k \in \{1, \ldots, l\}$.

We give the first assumption as follows.

Assumption 1. For each $k \in \{1, \ldots, l\}$, $\overline{\Gamma}^k$ is irreducible.

In order to emphasize the effect of the fast switching, (1) can be given by

$$d\left[x^{\varepsilon} - D\left(x_t^{\varepsilon}, r^{\varepsilon}(t)\right)\right] = f\left(x_t^{\varepsilon}, t, r^{\varepsilon}(t)\right) dt + g\left(x_t^{\varepsilon}, t, r^{\varepsilon}(t)\right) dw(t), \quad x_0^{\varepsilon} = \xi \in C\left([-\tau, 0]; \mathbb{R}^n\right), \quad r^{\varepsilon} = r_0. \tag{10}$$

To assure the existence and uniqueness of the solution, we give the following standard assumptions.

Assumption 2 (local Lipschitz condition). For each integer $\alpha \geq 1$, there exists a constant $L_\alpha > 0$ such that

$$\left|f(\varphi, t, i) - f(\phi, t, i)\right| \vee \left|g(\varphi, t, i) - g(\phi, t, i)\right| \leq L_\alpha \|\varphi - \phi\|^2, \tag{11}$$

for all $i \in \mathbb{S}$, $t \geq 0$ and those $\varphi, \phi \in C([-\tau, 0]; \mathbb{R}^n)$ with $\|\varphi\| \vee \|\phi\| \leq \alpha$, and $f(0, t, i) \equiv 0$, $g(0, t, i) \equiv 0$.

Assumption 3 (linear growth condition). There is an $L > 0$, for any $\varphi \in C([-\tau, 0]; \mathbb{R}^n)$, $t \geq 0$, $i \in \mathbb{S}$ such that

$$\left|f(\varphi, t, i)\right|^2 \vee \left|g(\varphi, t, i)\right|^2 \leq L\left(1 + \|\varphi\|^2\right). \tag{12}$$

Assumption 4. For all $i \in \mathbb{S}$ and those $\varphi, \phi \in C([-\tau, 0]; \mathbb{R}^n)$, there is a constant $0 < \kappa < 1$ such that

$$\left|D(\varphi, i) - D(\phi, i)\right| \leq \kappa \|\varphi - \phi\|^2, \quad D(0, i) \equiv 0. \tag{13}$$

Under Assumptions 2, 3, and 4, (10) has a unique solution denoted by $x^{\varepsilon,\xi,i}(t)$ on $t \geq 0$, where $x^{\varepsilon,\xi,i}$ is dependent on the initial value (ξ, i) (see [8]). Moreover, for every $p > 0$ and any compact subset G of $C([-\tau, 0]; \mathbb{R}^n)$, there is a positive constant H which is independent of ε such that

$$\sup_{(\xi,i) \in G \times \mathbb{S}} E\left[\sup_{-\tau \leq s \leq t} \left|x^{\varepsilon,\xi,i}(s)\right|^p\right] \leq H, \quad t \geq 0. \tag{14}$$

Since the state space of the Markov chain is large, it is too complicated to deal with directly. We need to analyse the limit equation of (10). To continue, make all the states in each $\mathbb{S}^k$ into a single state and define an aggregated process $\widetilde{r}^{\varepsilon}(\cdot)$ as

$$\widetilde{r}^{\varepsilon}(t) = k, \quad \text{if } r^{\varepsilon}(t) \in \mathbb{S}^k. \tag{15}$$

Denote the state space of $\widetilde{r}^{\varepsilon}(t)$ by $\widetilde{\mathbb{S}} = \{1, \ldots, l\}$, the stationary distribution $\overline{\Gamma}^k$ by $\mu^k = (\mu_1^k, \ldots, \mu_{M_k}^k) \in \mathbb{R}^{1 \times M_k}$, and $\overline{\mu} = \operatorname{diag}(\mu^1, \ldots, \mu^l) \in \mathbb{R}^{l \times M}$. Define

$$\widetilde{\Gamma} = \left(\widetilde{\gamma}_{ij}\right)_{l \times l} = \overline{\mu} \widehat{\Gamma} \mathbf{1} \tag{16}$$

with $\mathbf{1} = \operatorname{diag}(\mathbf{1}_{M_1}, \ldots, \mathbf{1}_{M_l})$ and $\mathbf{1}_{M_k} = (1, \ldots, 1)^T \in \mathbb{R}^{M_k \times 1}$, $k = 1, \ldots, l$. It has been known that $\widetilde{r}^{\varepsilon}(\cdot)$ converges weakly to $\widetilde{r}(\cdot)$ as $\varepsilon \to 0$, where $\widetilde{r}(\cdot)$ is a continuous-time Markov chain with generator $\widetilde{\Gamma}$ and state space $\widetilde{\mathbb{S}}$ (see [4]). Define

$$\widetilde{D}(\varphi, k) = \sum_{j=1}^{M_k} \mu_j^k D\left(\varphi, s_{kj}\right),$$
$$\widetilde{f}(\varphi, t, k) = \sum_{j=1}^{M_k} \mu_j^k f\left(\varphi, t, s_{kj}\right), \tag{17}$$
$$\widetilde{g}(\varphi, t, k)\, \widetilde{g}^T(\varphi, t, k) = \sum_{j=1}^{M_k} \mu_j^k g\left(\varphi, t, s_{kj}\right) g^T\left(\varphi, t, s_{kj}\right),$$

for each $s_{kj} \in \mathbb{S}^k$ with $k \in \{1, \ldots, l\}$ and $j \in \{1, \ldots, M_k\}$. It is easy to know that $\widetilde{D}(\varphi, k)$, $\widetilde{f}(\varphi, t, k)$, and $\widetilde{g}(\varphi, t, k)$ are the limits with respect to the stationary distribution of the Markov chain. Consider that, for any $\varphi \neq 0$, $g(\varphi, t, s_{kj}) g^T(\varphi, t, s_{kj})$ are nonnegative definite matrices, so we denote its "square root" of $g(\varphi, t, s_{kj}) g^T(\varphi, t, s_{kj})$ by $\widetilde{g}(\varphi, t, k)$. For degenerate diffusions, we can see the argument in [23].

The limit equation of (10) is defined as follows:

$$d\left[\widetilde{\varphi}(0) - \widetilde{D}\left(\widetilde{\varphi}, \widetilde{r}(t)\right)\right] = \widetilde{f}\left(\widetilde{\varphi}, t, \widetilde{r}(t)\right) dt + \widetilde{g}\left(\widetilde{\varphi}, t, \widetilde{r}(t)\right) dw(t), \quad \widetilde{x}_0 = \xi, \quad \widetilde{r} = \widetilde{r}_0. \tag{18}$$

3. Exponential Stability of NSFDE with Two-Time-Scale Markovian Switching

In this section, we establish the Razumikhin-type theorem on the exponential stability for (10). Denote by $C^p(\mathbb{R}^n \times \mathbb{R}_+ \times \widetilde{\mathbb{S}}; \mathbb{R}_+)$ the family of nonnegative real-valued functions defined on $\mathbb{R}^n \times \mathbb{R}_+ \times \widetilde{\mathbb{S}}$ that are p-times continuously differentiable with respect to x. At the same time, we need another assumption and a lemma with respect to $V(x, t, i) \in C^p(\mathbb{R}^n \times \mathbb{R}_+ \times \widetilde{\mathbb{S}}; \mathbb{R}_+)$ for some $p \geq 4$.

Assumption 5. For each $k \in \widetilde{\mathbb{S}}$, $V(x, t, i) \to \infty$ as $|x| \to \infty$. Moreover, $\partial^p V(x, t, i) = O(1)$, $\partial^\iota V(x, t, i)(|x|^\iota + |y|^\iota) \leq K(|x|^p + |y|^p + 1)$ for $1 \leq \iota \leq p-1$, where $\partial^\iota V(x, t, i)$ denotes the ιth derivative of $V(x, t, i)$ with respect to x and $O(y)$ denotes the function of y satisfying $\sup_y |O(y)|/y < \infty$.

Lemma 6. *Suppose that $p \geq 1$; there is a positive constant $\kappa \in (0, 1)$ such that*

$$\mathbb{E}\left|D(\varphi, k)\right|^p \leq \kappa^p \sup_{-\tau \leq \theta \leq 0} e^{\gamma\theta} \|\varphi\|^p, \quad (\varphi, i) \in L_{\mathcal{F}_t}^p\left([-\tau, 0]; \mathbb{R}^n\right) \times \mathbb{S}. \tag{19}$$

Then, for any $\xi \in L^p_{\mathscr{F}_0}([-\tau,0];\mathbb{R}^n)$, *the solution for* (10) *satisfies*

$$\sup_{-\infty<s\le t} e^{\gamma s}\mathbb{E}|x(s)|^p \le \frac{\|\xi\|^p}{1-\kappa} \vee \frac{\sup_{0\le s\le t} e^{\gamma s}\mathbb{E}|x(s)-D(x_s,r(s))|^p}{(1-\kappa)^p}, \quad t\ge 0. \tag{20}$$

Proof. Note the following elementary inequality:

$$(x+y)^p = (1-\kappa_1)^{1-p}\left(x^p+\kappa_1{}^{1-p}y^p\right), \quad \forall x,y\ge 0, \quad \kappa_1>0. \tag{21}$$

We have from condition (20) that, for any $t\ge 0$,

$$\begin{aligned} e^{\gamma t}\mathbb{E}|x(t)|^p &\le e^{\gamma t}\left[(1-\kappa)^{1-p}\mathbb{E}|x(t)-D(x_t,r(t))|^p + \kappa^{1-p}\mathbb{E}|D(x_t,r(t))|^p\right] \\ &\le (1-\kappa)^{1-p}e^{\gamma t}\mathbb{E}|x(t)-D(x_t,r(t))|^p + \kappa e^{\gamma t}\sup_{-\tau\le\theta\le 0} e^{\gamma\theta}\mathbb{E}|x(t+\theta)|^p \\ &\le (1-\kappa)^{1-p}\sup_{0\le s\le t} e^{\gamma s}\mathbb{E}|x(s)-D(x_s,r(s))|^p + \kappa\sup_{-\tau\le\theta\le 0} e^{\gamma(s+\theta)}\mathbb{E}|x(s+\theta)|^p \\ &\le (1-\kappa)^{1-p}\sup_{0\le s\le t} e^{\gamma s}\mathbb{E}|x(s)-D(x_s,r(s))|^p + \kappa\sup_{-\infty<s\le t} e^{\gamma s}\mathbb{E}|x(s)|^p. \end{aligned} \tag{22}$$

Then,

$$\begin{aligned} \sup_{-\infty<s\le t} e^{\gamma s}\mathbb{E}|x(s)|^p &\le \left[\sup_{-\tau\le\theta\le 0}\mathbb{E}|x(\theta)|^2\right] \\ &\vee \left[(1-\kappa)^{1-p}\sup_{0\le s\le t} e^{\gamma s}\mathbb{E}|x(s)-D(x_s,r(s))|^p + \kappa\sup_{-\infty<s\le t} e^{\gamma s}\mathbb{E}|x(s)|^p\right]. \end{aligned} \tag{23}$$

Therefore, the desired result holds. □

Theorem 7. *Let Assumptions 1–4 hold and let* c_1, c_2, λ, p *be all positive numbers and* $q>1$. *Assume that there exists a function* $V(x,t,k)\in C^p(\mathbb{R}^n\times\mathbb{R}_+\times\widetilde{\mathbb{S}};\mathbb{R}_+)$ *satisfying Assumption 5, such that*

$$c_1|x|^p \le V(x,t,k)\le c_2|x|^p, \tag{24}$$

for all $(x,t,k)\in\mathbb{R}^n\times\mathbb{R}_+\times\widetilde{\mathbb{S}}$, $t\ge 0$, $k\in\widetilde{\mathbb{S}}$. *Consider the following:*

$$\mathbb{E}|D(\varphi,k)|^p \le \kappa^p\sup_{-\tau\le\theta\le 0} e^{\nu\theta}\|\varphi\|^p, \quad \kappa=\max\{\kappa_1,\dots,\kappa_k\}, \quad \varphi\in L^p_{\mathscr{F}_t}, \tag{25}$$

for all $t\ge 0$, $0<\kappa_\sigma<1$, $\sigma=\{1,\dots,k\}$, *and*

$$\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}\mathscr{L}V(\varphi,t,k)\right] \le -\lambda\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}V(\varphi(0)-D(\varphi,k),t,k)\right], \tag{26}$$

provided $\varphi=\{\varphi(\theta):-\tau\le\theta\le 0\}\in L^p_{\mathscr{F}_t}([-\tau,0];\mathbb{R}^n)$, *satisfying*

$$\mathbb{E}\left[\min_{k\in\widetilde{\mathbb{S}}}V(\varphi(\theta),t+\theta,k)\right] < q\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}V(\varphi(0)-D(\varphi,k),t,k)\right], \tag{27}$$

for all $-\tau\le\theta\le 0$. *Then, for all* $\xi\in C^b_{\mathscr{F}_0}([-\tau,0];\mathbb{R}^n)$, $t\ge 0$,

$$\limsup_{\varepsilon\to 0}\mathbb{E}|x^\varepsilon(t)|^p \le \frac{c_2(1+\kappa)^p}{c_1(1-\kappa)^p}\|\xi\|^p e^{-\nu t}, \tag{28}$$

where

$$\nu=\min\left\{\overline{\gamma}, \frac{1}{\tau}\log\frac{q}{(c_2/c_1)(1-\kappa)^p}\right\}, \tag{29}$$

$\overline{\gamma}$ *being the root of the following equation:*

$$\frac{c_2}{c_1}(1-\kappa)^p e^{\overline{\gamma}\overline{\tau}}=\lambda. \tag{30}$$

In other words, the trivial solution of (10) is pth moment exponentially stable and the pth moment Lyapunov exponent is not greater than $-\nu$.

Proof. Let

$$\widetilde{V}(\varphi,t,j)=\sum_{k=1}^{l}V(\varphi,t,k)I_{\{j\in\mathbb{S}^k\}}=V(\varphi,t,k), \quad \text{if } j\in\mathbb{S}^k. \tag{31}$$

By the definition of $\widetilde{V}$, we know that

$$\widetilde{V}(\varphi^\varepsilon,t,r^\varepsilon(t))=V(\varphi^\varepsilon,t,\widetilde{r}^\varepsilon(t)), \quad \sum_{i=1}^{M}\widetilde{\gamma}_{li}\widetilde{V}(\varphi,t,i)=\sum_{i=1}^{M}\widetilde{\gamma}_{li}\sum_{k=1}^{l}V(\varphi,t,k)I_{\{i\in\mathbb{S}^k\}}=0. \tag{32}$$

Extend $r(t)$ to $[-\tau,0]$ by setting $r(t)=r(0)$. Recalling the facts that $x(t)$ is continuous for all $-\tau\le\theta\le 0$ and $r(t)$ is right continuous, it is easy to see that $\mathbb{E}V(x(t),t,r(t))$ is right continuous on $t \geq -\tau$. Let $\gamma \in (0,\nu)$ be arbitrary and define

$$\begin{aligned} U(t) &:= \sup_{-\tau\leq\theta\leq 0}\left[e^{\gamma(t+\theta)}\mathbb{E}V\left(x^{\varepsilon}(t+\theta)-D\left(x^{\varepsilon}_{t+\theta},\widetilde{r}^{\varepsilon}(t+\theta)\right),\right.\right. \\ &\qquad \left.\left. t+\theta,\widetilde{r}^{\varepsilon}(t+\theta)\right)\right] \\ &= \sup_{-\tau\leq\theta\leq 0}\left[e^{\gamma(t+\theta)}\mathbb{E}\widetilde{V}\left(x^{\varepsilon}(t+\theta)-D\left(x^{\varepsilon}_{t+\theta},r^{\varepsilon}(t+\theta)\right),\right.\right. \\ &\qquad \left.\left. t+\theta,r^{\varepsilon}(t+\theta)\right)\right], \end{aligned} \tag{33}$$

for all $t \geq 0$. We claim that

$$D^{+}U(t)=\limsup_{h\to 0+}\frac{U(t+h)-U(t)}{h}\leq 0,\quad \forall t\geq 0. \tag{34}$$

Note that, for each $t \geq 0$, either $U(t) > e^{\gamma t}\mathbb{E}V(x^{\varepsilon}(t)-D(x^{\varepsilon}_t,\widetilde{r}^{\varepsilon}(t)),t,\widetilde{r}^{\varepsilon}(t))$ or $U(t) = e^{\gamma t}\mathbb{E}V(x^{\varepsilon}(t)-D(x^{\varepsilon}_t,\widetilde{r}^{\varepsilon}(t)),t,\widetilde{r}^{\varepsilon}(t))$.

If $U(t) > e^{\gamma t}\mathbb{E}V(x^{\varepsilon}(t)-D(x^{\varepsilon}_t,\widetilde{r}^{\varepsilon}(t)),t,\widetilde{r}^{\varepsilon}(t))$, because $\mathbb{E}V(x(t),t,r(t))$ is right continuous on $t \geq -\tau$, it is easy to obtain that, for all $h > 0$ sufficiently small, $U(t) > e^{\gamma(t+h)}\mathbb{E}V(x^{\varepsilon}(t+h)-D(x^{\varepsilon}_{t+h},\widetilde{r}^{\varepsilon}(t+h)),t+h,\widetilde{r}^{\varepsilon}(t+h))$; hence, $U(t+h)\leq U(t)$ and $D^{+}U(t)\leq 0$.

If $U(t) = e^{\gamma t}\mathbb{E}V(x^{\varepsilon}(t)-D(x^{\varepsilon}_t,\widetilde{r}^{\varepsilon}(t)),t,\widetilde{r}^{\varepsilon}(t))$, we have

$$\begin{aligned} &e^{\gamma(t+\theta)}\mathbb{E}V\left(x^{\varepsilon}(t+\theta)-D\left(x_{t+\theta},\widetilde{r}^{\varepsilon}(t+\theta)\right),t+\theta,\widetilde{r}^{\varepsilon}(t+\theta)\right) \\ &\quad\leq e^{\gamma t}\mathbb{E}V\left(x^{\varepsilon}(t)-D\left(x_t,\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right), \end{aligned} \tag{35}$$

for all $-\tau\leq\theta\leq 0$.

Then,

$$\begin{aligned} &\mathbb{E}V\left(x^{\varepsilon}(t+\theta)-D\left(x_{t+\theta},\widetilde{r}^{\varepsilon}(t+\theta)\right),t+\theta,\widetilde{r}^{\varepsilon}(t+\theta)\right) \\ &\quad\leq e^{-\gamma\theta}\mathbb{E}V\left(x^{\varepsilon}(t)-D\left(x_t,\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right) \\ &\quad\leq e^{\gamma\tau}\mathbb{E}V\left(x^{\varepsilon}(t)-D\left(x_t,\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right), \end{aligned} \tag{36}$$

for all $-\tau\leq\theta\leq 0$.

On the other hand, by Lemma 6, we derive

$$\begin{aligned} &e^{\gamma(t+\theta)}\mathbb{E}V\left(x^{\varepsilon}(t+\theta),t+\theta,\widetilde{r}^{\varepsilon}(t+\theta)\right) \\ &\quad\leq c_2e^{\gamma(t+\theta)}\mathbb{E}\left|x^{\varepsilon}(t+\theta)\right|^{p} \\ &\quad\leq c_2(1-\kappa)^{p}\sup_{0\leq s\leq t}e^{\gamma s}\mathbb{E}\left|x^{\varepsilon}(s)-D\left(x^{\varepsilon}_s,\widetilde{r}(s)\right)\right|^{p} \\ &\quad\leq \frac{c_2}{c_1}(1-\kappa)^{p}\sup_{0\leq s\leq t}e^{\gamma s}\mathbb{E}V\left(x^{\varepsilon}(s)-D\left(x^{\varepsilon}_s,\widetilde{r}(s)\right),s,\widetilde{r}(s)\right) \\ &\quad\leq \frac{c_2}{c_1}(1-\kappa)^{p}e^{\gamma t}\mathbb{E}V\left(x^{\varepsilon}(t)-D\left(x^{\varepsilon}_t,\widetilde{r}(t)\right),t,\widetilde{r}(t)\right). \end{aligned} \tag{37}$$

Then,

$$\begin{aligned} &\mathbb{E}V\left(x^{\varepsilon}(t+\theta),t+\theta,\widetilde{r}^{\varepsilon}(t+\theta)\right) \\ &\quad< q\mathbb{E}V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right), \end{aligned} \tag{38}$$

where $q > (c_2/c_1)(1-\kappa)^{p}e^{\gamma\tau}$; that is, $\gamma < (1/\tau)(\log(q/(c_2/c_1)(1-\kappa)^{p}))$.

Consequently, there exists a sufficiently small $\varepsilon_0 > 0$, such that, for any $\varepsilon\in(0,\varepsilon_0)$,

$$\begin{aligned} &\mathbb{E}\left[\min_{k\in\widetilde{\mathbb{S}}}V\left(\varphi^{\varepsilon}(\theta),t+\theta,k\right)\right] \\ &\quad< q\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},k\right),t,k\right)\right], \end{aligned} \tag{39}$$

for all $-\tau\leq\theta\leq 0$. Thus,

$$\begin{aligned} &\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}\mathscr{L}V\left(\varphi^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t)\right)\right] \\ &\quad\leq -\lambda\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right)\right], \end{aligned} \tag{40}$$

which implies that

$$\begin{aligned} &\mathbb{E}\left[\mathscr{L}V\left(\varphi^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t)\right)\right] \\ &\quad\leq -\lambda\mathbb{E}\left[V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right)\right]. \end{aligned} \tag{41}$$

By the condition of $\gamma < \nu \leq \lambda$, we get

$$\begin{aligned} &\mathbb{E}\left[\mathscr{L}V\left(\varphi^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t)\right)\right] \\ &\quad\leq -\gamma\mathbb{E}\left[V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)\right),t,\widetilde{r}^{\varepsilon}(t)\right)\right]. \end{aligned} \tag{42}$$

Next, we consider

$$\begin{aligned} &U(t+h)-U(t) \\ &\quad= \limsup_{\varepsilon\to 0}\left[e^{\gamma(t+\theta+h)}\mathbb{E}\left[V\left(x^{\varepsilon}(t+\theta+h)\right.\right.\right. \\ &\qquad\qquad -D\left(x^{\varepsilon}_{t+\theta+h},\widetilde{r}^{\varepsilon}(t+\theta+h)\right), \\ &\qquad\qquad \left.\left. t+\theta+h,\widetilde{r}^{\varepsilon}(t+\theta+h)\right)\right] \\ &\qquad - e^{\gamma(t+\theta)}\mathbb{E}\left[V\left(x^{\varepsilon}(t+\theta)-D\left(x^{\varepsilon}_{t+\theta},\widetilde{r}^{\varepsilon}(t+\theta)\right),\right.\right. \\ &\qquad\qquad \left.\left.\left. t+\theta,\widetilde{r}^{\varepsilon}(t+\theta)\right)\right]\right] \\ &\quad= \limsup_{\varepsilon\to 0}\mathbb{E}\int_{t+\theta}^{t+\theta+h}e^{\gamma s}\left[\mathscr{L}V\left(\varphi^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s)\right)\right. \\ &\qquad\qquad +\gamma V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)\right),\right. \\ &\qquad\qquad \left.\left. s,\widetilde{r}^{\varepsilon}(s)\right)\right]ds \\ &\quad= \limsup_{\varepsilon\to 0}\mathbb{E}\int_{t+\theta}^{t+\theta+h}e^{\gamma s}\left[\mathscr{L}\widetilde{V}\left(\varphi^{\varepsilon},s,r^{\varepsilon}(s)\right)\right. \\ &\qquad\qquad +\gamma V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)\right),\right. \\ &\qquad\qquad \left.\left. s,\widetilde{r}^{\varepsilon}(s)\right)\right]ds. \end{aligned} \tag{43}$$

By the definition of operator $\mathscr{L}$, we have

$$
\begin{aligned}
&\mathscr{L}\widetilde{V}\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\\
&\quad=\widetilde{V}_t\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right),t,r^{\varepsilon}(t)\right)\\
&\qquad+\widetilde{V}_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right),t,r^{\varepsilon}(t)\right)\\
&\qquad\times f\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\\
&\qquad+\frac{1}{2}\operatorname{trace}\left[g^T\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\right.\\
&\qquad\qquad\times\widetilde{V}_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right),t,r^{\varepsilon}(t)\right)\\
&\qquad\qquad\left.\times g\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\right]\\
&\qquad+\sum_{i=1}^{M}\gamma^{\varepsilon}_{r^{\varepsilon}(t)i}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,i\right),t,i\right)\\
&\quad=\widetilde{V}_t\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right),t,r^{\varepsilon}(t)\right)\\
&\qquad+\widetilde{V}_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right),t,r^{\varepsilon}(t)\right)\\
&\qquad\times f\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\\
&\qquad+\frac{1}{2}\operatorname{trace}\left[g^T\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\right.\\
&\qquad\qquad\times\widetilde{V}_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right),t,r^{\varepsilon}(t)\right)\\
&\qquad\qquad\left.\times g\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\right]\\
&\qquad+\sum_{i=1}^{M}\widehat{\gamma}_{r^{\varepsilon}(t)i}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,i\right),t,i\right)\\
&\quad=V_t\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad+V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\times\overline{f}\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad+\frac{1}{2}\operatorname{trace}\left[\overline{g}^T\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\right.\\
&\qquad\qquad\times V_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\left.\times\overline{g}\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\right]\\
&\qquad+\sum_{k=1}^{l}\overline{\gamma}^{\varepsilon}_{\overline{r}^{\varepsilon}(t)k}V\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,k\right),t,k\right)\\
&\qquad+V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\times\left[f\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)-\overline{f}\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\right]\\
&\qquad+\frac{1}{2}\operatorname{trace}\left[g^T\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\right.\\
&\qquad\qquad\times V_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\times g\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)-\overline{g}^T\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\times V_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\left.\times\overline{g}\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\right]\\
&\qquad+\sum_{i=1}^{M}\widehat{\gamma}_{r^{\varepsilon}(t)i}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,i\right),t,i\right)\\
&\qquad-\sum_{k=1}^{l}\overline{\gamma}_{\overline{r}^{\varepsilon}(t)k}V\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,k\right),t,k\right)\\
&\quad=\mathscr{L}V\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad+V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\times\left[f\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)-\overline{f}\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\right]\\
&\qquad+\frac{1}{2}\operatorname{trace}\left[g^T\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)\right.\\
&\qquad\qquad\times V_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\times g\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)-\overline{g}^T\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\times V_{xx}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(t)\right),t,\overline{r}^{\varepsilon}(t)\right)\\
&\qquad\qquad\left.\times\overline{g}\left(\varphi^{\varepsilon},t,\overline{r}^{\varepsilon}(t)\right)\right]\\
&\qquad+\sum_{i=1}^{M}\widehat{\gamma}_{r^{\varepsilon}(t)i}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,i\right),t,i\right)\\
&\qquad-\sum_{k=1}^{l}\overline{\gamma}_{\overline{r}^{\varepsilon}(t)k}V\left(\varphi^{\varepsilon}(0)-D\left(x^{\varepsilon}_t,k\right),t,k\right).
\end{aligned}\tag{44}
$$

Therefore,

$$
\begin{aligned}
&U(t+h)-U(t)\\
&\quad=\limsup_{\varepsilon\to0}\mathbb{E}\\
&\qquad\times\int_{t+\theta}^{t+\theta+h}e^{\gamma s}\\
&\qquad\qquad\times\left[\mathscr{L}V\left(\varphi^{\varepsilon},s,\overline{r}^{\varepsilon}(s)\right)\right.\\
&\qquad\qquad\quad\left.+\gamma V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(s)\right),s,\overline{r}^{\varepsilon}(s)\right)\right]ds\\
&\qquad+\limsup_{\varepsilon\to0}\mathbb{E}\\
&\qquad\times\int_{t+\theta}^{t+\theta+h}e^{\gamma t}V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\overline{r}^{\varepsilon}(s)\right),s,\overline{r}^{\varepsilon}(s)\right)\\
&\qquad\qquad\times\left[f\left(\varphi^{\varepsilon},s,r^{\varepsilon}(s)\right)-\overline{f}\left(\varphi^{\varepsilon},s,\overline{r}^{\varepsilon}(s)\right)\right]ds\\
&\qquad+\frac{1}{2}\limsup_{\varepsilon\to0}\mathbb{E}\\
&\qquad\times\int_{t+\theta}^{t+\theta+h}e^{\gamma s}
\end{aligned}
$$

$$\times \operatorname{trace}\left[g^T\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right)\right.$$
$$\times V_{xx}\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)$$
$$\times g\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right)-\tilde{g}^T\left(\varphi^\varepsilon, s, \tilde{r}^\varepsilon(s)\right)$$
$$\times V_{xx}\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)$$
$$\left.\times\, \tilde{g}\left(\varphi^\varepsilon, s, \tilde{r}^\varepsilon(s)\right)\right] ds$$
$$+\limsup_{\varepsilon\to 0}\mathbb{E}$$
$$\times \int_{t+\theta}^{t+\theta+h} e^{\gamma s}$$
$$\times\left(\sum_{i=1}^{M}\widehat{\gamma}_{r^\varepsilon(s)i}\widetilde{V}\left(\varphi^\varepsilon(0)-D\left(x_s^\varepsilon, i\right), s, i\right)\right.$$
$$\left.-\sum_{k=1}^{l}\widetilde{\gamma}_{\tilde{r}^\varepsilon(s)k}V\left(\varphi^\varepsilon(0)-D\left(x_s^\varepsilon, k\right), s, k\right)\right) ds$$
$$=: I_1+I_2+I_3+I_4. \quad (45)$$

By the definition of $\tilde{f}$,

$$f\left(\varphi^\varepsilon, t, r^\varepsilon(t)\right)-\tilde{f}\left(\varphi^\varepsilon, t, \tilde{r}^\varepsilon(t)\right) = \sum_{k=1}^{l}\sum_{j=1}^{M_k} f\left(\varphi^\varepsilon, t, s_{kj}\right)\times\left[I_{\{r^\varepsilon(t)=s_{kj}\}}-\mu_j^k I_{\{\tilde{r}^\varepsilon(t)=k\}}\right]. \quad (46)$$

This implies that

$$\lim_{\varepsilon\to 0}\mathbb{E}\int_{t+\theta}^{t+\theta+h} e^{\gamma s}V_x\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)$$
$$\times\left[f\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right)-\tilde{f}\left(\varphi^\varepsilon, s, \tilde{r}^\varepsilon(s)\right)\right] ds$$
$$\leq \lim_{\varepsilon\to 0}\left[\mathbb{E}\left|\int_{t+\theta}^{t+\theta+h} e^{\gamma s}V_x\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)\right.\right.$$
$$\left.\left.\times\left[f\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right)-\tilde{f}\left(\varphi^\varepsilon, s, \tilde{r}^\varepsilon(s)\right)\right] ds\right|^2\right]^{1/2}$$
$$=\lim_{\varepsilon\to 0}\left[\mathbb{E}\left|\int_{t+\theta}^{t+\theta+h} e^{\gamma s}V_x\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)\right.\right.$$
$$\times\sum_{k=1}^{l}\sum_{j=1}^{M_k} f\left(\varphi^\varepsilon, s, s_{kj}\right)$$
$$\left.\left.\times\left[I_{\{r^\varepsilon(s)=s_{kj}\}}-\mu_j^k I_{\{\tilde{r}^\varepsilon(s)=k\}}\right] ds\right|^2\right]^{1/2}$$
$$\leq \lim_{\varepsilon\to 0}\left[\mathbb{E}\left|\int_{t+\theta}^{t+\theta+h}\sum_{k=1}^{l}\sum_{j=1}^{M_k} e^{\gamma s}L\left(1+\|\varphi\|^p\right)\right.\right.$$
$$\left.\left.\times\left[I_{\{r^\varepsilon(s)=s_{kj}\}}-\mu_j^k I_{\{\tilde{r}^\varepsilon(s)=k\}}\right] ds\right|^2\right]^{1/2}. \quad (47)$$

By the argument of Lemma 7.14 in [4], the right side of the above inequality is equivalent to 0; that is, $I_2 = 0$. Similarly, we can show that

$$I_3 = \frac{1}{2}\limsup_{\varepsilon\to 0}\mathbb{E}$$
$$\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}$$
$$\times \operatorname{trace}\left[g^T\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right)\right.$$
$$\times V_{xx}\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)$$
$$\times g\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right)$$
$$-\tilde{g}^T\left(\varphi^\varepsilon, s, \tilde{r}^\varepsilon(s)\right)$$
$$\times V_{xx}\left(\varphi^\varepsilon(0)-D\left(\varphi^\varepsilon, \tilde{r}^\varepsilon(s)\right), s, \tilde{r}^\varepsilon(s)\right)$$
$$\left.\times\,\tilde{g}\left(\varphi^\varepsilon, s, \tilde{r}^\varepsilon(s)\right)\right] ds = 0. \quad (48)$$

By the definition of $\widehat{\Gamma}$ and $\widetilde{\Gamma}$, we have

$$\sum_{i=1}^{M}\widehat{\gamma}_{r^\varepsilon(t)i}\widetilde{V}\left(\varphi^\varepsilon(0)-D\left(x_t^\varepsilon, i\right), t, i\right) = \widehat{\Gamma}\widetilde{V}\left(\varphi^\varepsilon(0)-D\left(x_t^\varepsilon, i\right), t, \cdot\right)\left(r^\varepsilon(t)\right),$$
$$\sum_{k=1}^{l}\widetilde{\gamma}_{\tilde{r}^\varepsilon(t)k}V\left(\varphi(0)-D\left(x_t^\varepsilon, k\right), t, k\right) = \widetilde{\Gamma}V\left(\varphi(0)-D\left(x_t^\varepsilon, k\right), t, \cdot\right)\left(\tilde{r}^\varepsilon(t)\right). \quad (49)$$

Hence,

$$I_4 = \limsup_{\varepsilon\to 0}\mathbb{E}$$
$$\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}$$
$$\times\left(\sum_{i=1}^{M}\widehat{\gamma}_{r^\varepsilon(s)i}\widetilde{V}\left(\varphi^\varepsilon(0)-D\left(x_s^\varepsilon, i\right), s, i\right)\right.$$
$$\left.-\sum_{k=1}^{l}\widetilde{\gamma}_{\tilde{r}^\varepsilon(s)k}V\left(\varphi^\varepsilon(0)-D\left(x_s^\varepsilon, k\right), s, k\right)\right) ds$$

$$
\begin{aligned}
&= \limsup_{\varepsilon\to 0}\mathbb{E} \\
&\quad\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\
&\qquad\times\left(\widehat{\Gamma}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x_s^{\varepsilon},i\right),s,\cdot\right)\left(r^{\varepsilon}(s)\right)\right. \\
&\qquad\quad\left.-\widetilde{\Gamma}V\left(\varphi^{\varepsilon}(0)-D\left(x_s^{\varepsilon},k\right),s,\cdot\right)\left(s_{kj}\right)\right)ds \\
&= \limsup_{\varepsilon\to 0}\mathbb{E} \\
&\quad\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\sum_{k=1}^{l}\sum_{j=1}^{M_k}\widehat{\Gamma}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x_s^{\varepsilon},i\right),s,\cdot\right)\left(s_{kj}\right) \\
&\qquad\times\left[I_{\{r^{\varepsilon}(s)=s_{kj}\}}-\mu_j^k I_{\{\widetilde{r}^{\varepsilon}(s)=k\}}\right]ds \\
&\le \limsup_{\varepsilon\to 0}\left[\mathbb{E}\left|\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\right.\right. \\
&\qquad\times\sum_{k=1}^{l}\sum_{j=1}^{M_k}\widehat{\Gamma}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(x_s^{\varepsilon},i\right),s,\cdot\right)\left(s_{kj}\right) \\
&\qquad\left.\left.\times\left[I_{\{r^{\varepsilon}(s)=s_{kj}\}}-\mu_j^k I_{\{\widetilde{r}^{\varepsilon}(s)=k\}}\right]\right|^2\right]^{1/2}.
\end{aligned}
\tag{50}
$$

By the argument of Lemma 7.14 in [4], the right side of the above inequality is equivalent to 0; that is, $I_4 = 0$. Therefore,

$$
\begin{aligned}
&U(t+h)-U(t) \\
&\quad=\lim_{\varepsilon\to 0}\mathbb{E} \\
&\qquad\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\
&\qquad\quad\times\left[\mathscr{L}V\left(\varphi^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s)\right)\right. \\
&\qquad\qquad\left.+\gamma V\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)\right),s,\widetilde{r}^{\varepsilon}(s)\right)\right]ds \\
&\quad\le 0.
\end{aligned}
\tag{51}
$$

That is

$$U(t+h)\le U(t). \tag{52}$$

So, $U(t+h)=U(t)$ for all $h>0$ sufficiently small, and hence $D^{+}U(t)=0$. Inequality (34) holds.

It follows from (34) that $U(t)\le U(0)$ for all $t\ge 0$. By the definition of $U(t)$,

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0} e^{\gamma t}\mathbb{E}\left|x^{\varepsilon}(t)-D\left(x_t^{\varepsilon},\widetilde{r}^{\varepsilon}(t)\right)\right|^p \\
&\quad\le c_2\limsup_{\varepsilon\to 0}\sup_{\tau\le\theta\le 0} e^{\gamma\theta}\mathbb{E}\left|x^{\varepsilon}(\theta)-D\left(x_\theta^{\varepsilon},\widetilde{r}^{\varepsilon}(\theta)\right)\right|^p \\
&\quad\le c_2\limsup_{\varepsilon\to 0}\sup_{-\tau\le\theta\le 0}(1+\kappa)^{p-1} \\
&\qquad\times\left[\mathbb{E}\left|x^{\varepsilon}(\theta)\right|^p+\kappa^{1-p}\mathbb{E}\left|D\left(x_\theta^{\varepsilon},\widetilde{r}^{\varepsilon}(\theta)\right)\right|^p\right] \\
&\quad\le c_2(1+\kappa)^p\|\xi\|^p,\qquad t\ge 0.
\end{aligned}
\tag{53}
$$

By Lemma 6, we derive

$$\limsup_{\varepsilon\to 0} e^{\gamma t}\mathbb{E}\left|x^{\varepsilon}(t)\right|^p\le\frac{c_2(1+\kappa)^p}{c_1(1-\kappa)^p}\|\xi\|^p. \tag{54}$$

That is,

$$\limsup_{\varepsilon\to 0}\mathbb{E}\left|x^{\varepsilon}(t)\right|^p\le\frac{c_2(1+\kappa)^p}{c_1(1-\kappa)^p}\|\xi\|^p e^{-\gamma t},\quad\forall t\ge 0. \tag{55}$$

□

4. Neutral Stochastic Functional System with Pure Jump

In this section, we discuss the stability of the following neutral stochastic functional system with pure jump:

$$
\begin{aligned}
&d\left[\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},r^{\varepsilon}(t)\right)\right] \\
&\quad=f\left(\varphi^{\varepsilon},t,r^{\varepsilon}(t)\right)dt+\int_{\mathbb{R}^m} b\left(x_{t-}^{\varepsilon},t,r^{\varepsilon}(t),z\right)\widetilde{N}(dt,dz) \\
&\qquad x_0=\xi\in C\left([-\tau,0];\mathbb{R}^n\right),\quad r(0)\in\mathbb{S},
\end{aligned}
\tag{56}
$$

where $x_{t-}^{\varepsilon}=\lim_{s\uparrow t}x_s^{\varepsilon}$, $D: C([-\tau,0];\mathbb{R}^n)\times\mathbb{S}\to\mathbb{R}^n$, $b: C([-\tau,0];\mathbb{R}^n)\times\mathbb{R}_+\times\mathbb{S}\times\mathbb{R}^m\to\mathbb{R}^{n\times m}$. We assume that each column $b^{(\delta)}$ of the $n\times m$ matrix $b=[b_{ij}]$ depends on z only through the δth coordinate z_δ; that is,

$$
\begin{aligned}
&b^{(\delta)}(\varphi,t,i,z)=b^{(\delta)}(\varphi,t,i,z_\delta); \\
&z=(z_1,\ldots,z_m)\in\mathbb{R}^m,\quad i\in\mathbb{S}.
\end{aligned}
\tag{57}
$$

$N(t,z)$ is an m-dimensional Poisson process and the compensated Poisson process is defined by

$$
\begin{aligned}
\widetilde{N}(dt,dz)&=\left(\widetilde{N}_1(dt,dz_1),\ldots,\widetilde{N}_m(dt,dz_m)\right) \\
&=\left(N_1(dt,dz_1)-\lambda_1(dz_1)dt,\ldots,N_m(dt,dz_m)\right. \\
&\qquad\left.-\lambda_m(dz_m)dt\right),
\end{aligned}
\tag{58}
$$

where $\{N_\delta,\delta=1,\ldots,m\}$ are independent one-dimensional Poisson random measures with characteristic measure

$\{\lambda_\delta, \delta = 1, \ldots, m\}$ coming from m independent one-dimensional Poisson point processes. The limit system of (56) is defined as follows:

$$\begin{aligned} d\left[\varphi^{\varepsilon}(0) - D\left(\tilde{x}_t^{\varepsilon}, \tilde{r}^{\varepsilon}(t)\right)\right] &= \tilde{f}\left(\tilde{x}_t^{\varepsilon}, t, \tilde{r}^{\varepsilon}(t)\right) dt \\ &\quad + \int_{\mathbb{R}^m} \tilde{b}\left(\tilde{x}_{t-}^{\varepsilon}, t, \tilde{r}^{\varepsilon}(t), z\right) \widetilde{N}(dt, dz) \\ \tilde{x}_0 = \xi \in C\left([-\tau, 0]; \mathbb{R}^n\right), &\quad \tilde{r}(0) \in \mathbb{S}, \end{aligned} \tag{59}$$

where $\tilde{x}_{t-}^{\varepsilon} = \lim_{s\uparrow t} \tilde{x}_s^{\varepsilon}$ and $\tilde{b} : C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \widetilde{\mathbb{S}} \times \mathbb{R}^m \to \mathbb{R}^{n\times m}$. Similar to the definition of $\tilde{f}$, we define

$$\begin{aligned} \widetilde{D}(\varphi, k) &= \sum_{j=1}^{N_m} \mu_j^k D\left(\varphi, s_{kj}\right), \\ \tilde{b}(\varphi, t, k, z) &= \sum_{j=1}^{N_m} \mu_j^k b\left(\varphi, t, s_{kj}, z\right), \end{aligned} \tag{60}$$

for each $s_{kj} \in \mathbb{S}^k$ with $k \in \{1, \ldots, l\}$ and $j \in \{1, \ldots, N_m\}$.

To assure the existence and uniqueness of the solution of (59), we also give the following standard assumptions.

Assumption 8. For any integer ζ, there is a constant $L_\zeta > 0$, such that

$$\begin{aligned} &\left|f(\varphi, t, i) - f(\phi, t, i)\right| \\ &\quad \vee \sum_{\delta=1}^{m} \int_{\mathbb{R}} \left|b^{(\delta)}(\varphi, t, i, z_\delta) - b^{(\delta)}(\phi, t, i, z_\delta)\right| \lambda_\delta(dz_\delta) \\ &\quad \le L_\zeta \|\varphi - \phi\|^2, \end{aligned} \tag{61}$$

for all $i \in \mathbb{S}$ and those $\varphi, \phi \in C([-\tau, 0]; \mathbb{R}^n)$ with $\|\varphi\| \vee \|\phi\| \le \zeta$, and $f(0, t, i) \equiv 0$, $b(0, t, i, z) \equiv 0$.

Assumption 9. There is an $\bar{L} > 0$, such that, for any $\varphi, \phi \in C([-\tau, 0]; \mathbb{R}^n)$, $i \in \mathbb{S}$,

$$\begin{aligned} \left|f(\varphi, t, i)\right| \vee \sum_{\delta=1}^{m} \int_{\mathbb{R}} \left|b^{(\delta)}(\varphi, t, i, z_\delta)\right| \lambda_\delta(dz_\delta) \\ \le \bar{L}\left(1 + \|\varphi\|^2\right). \end{aligned} \tag{62}$$

Assumption 10. For all $i \in \mathbb{S}$ and those $\varphi, \phi \in C([-\tau, 0]; \mathbb{R}^n)$, there is a constant $0 < \kappa < 1$ such that

$$\begin{aligned} \left|D(\varphi, i) - D(\phi, i)\right| &\le \kappa \|\varphi - \phi\|^2, \\ D(0, i) &\equiv 0. \end{aligned} \tag{63}$$

Given that $V \in C^p(\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}; \mathbb{R}_+)$, define an operator $\mathbb{L}V$ by

$$\begin{aligned} &\mathbb{L}V(\varphi, t, i) \\ &\quad = V_t\left(\varphi(0) - D(\varphi, i), t, i\right) \\ &\qquad + V_x\left(\varphi(0) - D(\varphi, i), t, i\right) f(\varphi, t, i) \\ &\qquad + \sum_{j=1}^{N} \gamma_{ij} V\left(\varphi(0) - D(\varphi, i), t, j\right) \\ &\qquad + \int_{\mathbb{R}} \sum_{\delta=1}^{m} \left\{ V\left(\varphi(0) - D(\varphi, i) + b^{(\delta)}(\varphi, t, \iota, z_\delta), t, \iota\right) \right. \\ &\qquad\qquad - V\left(\varphi(0) - D(\varphi, i), t, i\right) \\ &\qquad\qquad - V_x\left(\varphi(0) - D(\varphi, i), t, i\right) \\ &\qquad\qquad \left. \times b^{(\delta)}(\varphi, t, \iota, z_\delta) \right\} \lambda_\delta(dz_\delta), \end{aligned} \tag{64}$$

where

$$\begin{aligned} &V_x\left(\varphi(0) - D(\varphi, i), t, i\right) \\ &\quad = \left(\frac{\partial V\left(\varphi(0) - D(\varphi, i), t, i\right)}{\partial x_1}, \ldots, \right. \\ &\qquad \left. \frac{\partial V\left(\varphi(0) - D(\varphi, i), t, i\right)}{\partial x_m} \right). \end{aligned} \tag{65}$$

Lemma 11 (see [20]). *Let Assumptions 1, 8, and 9 hold, as $\varepsilon \to 0$; then, $(x^{\varepsilon}(\cdot), \tilde{r}^{\varepsilon}(\cdot))$ converges weakly to $(\tilde{x}(\cdot), \tilde{r}(\cdot))$ in $D([0, \infty), \mathbb{R}^n \times \widetilde{\mathbb{S}})$, where $D([0, \infty), \mathbb{R}^n \times \widetilde{\mathbb{S}})$ is the space of functions defined on $[0, \infty)$ that are right continuous and have left limits taking values in $\mathbb{R}^n \times \widetilde{\mathbb{S}}$ and are endowed with the Skorohod topology.*

Theorem 12. *Let Assumptions 1 and 8–10 hold and let c_1, c_2, λ, p be all positive numbers and $q > 1$. Assume that there exists a function $V(x, t, k) \in C^p(\mathbb{R}^n \times \mathbb{R}_+ \times \widetilde{\mathbb{S}}; \mathbb{R}_+)$ satisfying Assumption 5, such that*

$$c_1 |x|^p \le V(x, t, k) \le c_2 |x|^p, \quad k \in \widetilde{\mathbb{S}}, \tag{66}$$

for all $(x, t, k) \in \mathbb{R}^n \times \mathbb{R}_+ \times \widetilde{\mathbb{S}}$ and $t \ge 0, k \in \widetilde{\mathbb{S}}$. Consider the following:

$$\begin{aligned} \mathbb{E}\left|D(\varphi, k)\right|^p &\le \kappa^p \sup_{-\tau \le \theta \le 0} e^{\bar{\nu}\theta} \|\varphi\|^p, \\ \kappa = \max\left\{\kappa_1, \ldots, \kappa_k\right\}, &\quad \varphi \in L^p_{\mathcal{F}_t}, \end{aligned} \tag{67}$$

for all $t \ge 0$, $0 < \kappa_\sigma < 1$, $\sigma = \{1, \ldots, k\}$, and

$$\mathbb{E}\left[\max_{k \in \widetilde{\mathbb{S}}} \mathbb{L}V(\varphi, t, k)\right] \le -\gamma \mathbb{E}\left[\max_{k \in \widetilde{\mathbb{S}}} V\left(\varphi(0) - D(\varphi, k)\right)\right], \tag{68}$$

provided $\varphi = \{\varphi(\theta) : -\tau \le \theta \le 0\} \in L^p_{\mathscr{F}_t}([-\tau, 0]; \mathbb{R}^n)$, *satisfying*

$$\mathbb{E}\left[\min_{k\in\widetilde{\mathbb{S}}} V(x(t+\theta), t+\theta, k)\right] < q\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}} V(\varphi(0) - D(\varphi, k), t, i)\right], \quad -\tau \le \theta \le 0. \tag{69}$$

Then, for all $\xi \in C([-\tau, 0]; \mathbb{R}^n)$, $t \ge 0$,

$$\limsup_{\varepsilon\to 0} \mathbb{E}|x^\varepsilon(t)|^p \le \frac{c_2(1+\kappa)^p}{c_1(1-\kappa)^p}\|\xi\|^p e^{-\overline{\nu}t}, \tag{70}$$

where

$$\overline{\nu} = \min\left\{\overline{\gamma}, \frac{1}{\tau}\log\frac{q}{(c_2/c_1)(1-\kappa)^p}\right\}, \tag{71}$$

$\overline{\gamma}$ *being the root of the following equation:*

$$\frac{c_2}{c_1}(1-\kappa)^p e^{\overline{\gamma}\tau} = \lambda. \tag{72}$$

Proof. Define

$$\widetilde{V}(\varphi, t, \rho) = \sum_{k=1}^{l} V(\varphi, t, k) I_{\{\rho\in\mathbb{S}^k\}} = V(\varphi, t, k), \quad \text{if } \rho \in \mathbb{S}^k. \tag{73}$$

Extend $r(t)$ to $[-\tau, 0]$ by setting $r(t) = r(0)$; then, $\mathbb{E}V(x(t), t, r(t))$ is right continuous on $t \ge -\tau$. Let $\gamma \in (0, \overline{\nu})$ be arbitrary and define

$$\begin{aligned} U(t) &:= \sup_{-\tau\le\theta\le 0}\left[e^{\gamma(t+\theta)}\mathbb{E}V\left(x^\varepsilon(t+\theta) - D\left(x^\varepsilon_{t+\theta}, \widetilde{r}^\varepsilon(t+\theta)\right), t+\theta, \widetilde{r}^\varepsilon(t+\theta)\right)\right] \\ &= \sup_{-\tau\le\theta\le 0}\left[e^{\gamma(t+\theta)}\mathbb{E}\widetilde{V}\left(x^\varepsilon(t+\theta) - D\left(x^\varepsilon_{t+\theta}, r^\varepsilon(t+\theta)\right), t+\theta, r^\varepsilon(t+\theta)\right)\right], \end{aligned} \tag{74}$$

for all $t \ge 0$. We claim that

$$D^+U(t) = \limsup_{h\to 0+}\frac{U(t+h) - U(t)}{h} \le 0, \quad \forall t \ge 0. \tag{75}$$

Similar to the proof of Theorem 7, we derive

$$\begin{aligned} &\mathbb{E}V\left(x^\varepsilon(t+\theta), t+\theta, \widetilde{r}^\varepsilon(t+\theta)\right) \\ &\quad < q\mathbb{E}V\left(\varphi(0) - D\left(\varphi^\varepsilon, \widetilde{r}^\varepsilon(t)\right), t, \widetilde{r}^\varepsilon(t)\right), \end{aligned} \tag{76}$$

for all $-\tau \le \theta \le 0$, where $q > (c_2/c_1)(1-\kappa)^p e^{\gamma\tau}$; that is, $\gamma < (1/\tau)(\log(q/(c_2/c_1)(1-\kappa)^p))$.

Thus,

$$\begin{aligned} &\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}}\mathbb{L}V\left(\varphi^\varepsilon, t, \widetilde{r}^\varepsilon(t)\right)\right] \\ &\quad \le -\lambda\mathbb{E}\left[\max_{k\in\widetilde{\mathbb{S}}} V\left(\varphi^\varepsilon(0) - D\left(\varphi^\varepsilon, \widetilde{r}^\varepsilon(t)\right), t, \widetilde{r}^\varepsilon(t)\right)\right], \end{aligned} \tag{77}$$

which implies that

$$\begin{aligned} &\mathbb{E}\left[\mathbb{L}V\left(\varphi^\varepsilon, t, \widetilde{r}^\varepsilon(t)\right)\right] \\ &\quad \le -\lambda\mathbb{E}\left[V\left(\varphi^\varepsilon(0) - D\left(\varphi^\varepsilon, \widetilde{r}^\varepsilon(t)\right), t, \widetilde{r}^\varepsilon(t)\right)\right]. \end{aligned} \tag{78}$$

By the condition of $\gamma < \overline{\nu} \le \lambda$, we get

$$\begin{aligned} &\mathbb{E}\left[\mathbb{L}V\left(\varphi^\varepsilon, t, \widetilde{r}^\varepsilon(t)\right)\right] \\ &\quad \le -\gamma\mathbb{E}\left[V\left(\varphi^\varepsilon(0) - D\left(\varphi^\varepsilon, \widetilde{r}^\varepsilon(t)\right), t, \widetilde{r}^\varepsilon(t)\right)\right]. \end{aligned} \tag{79}$$

We now consider

$$\begin{aligned} U(t+h) - U(t) &= \limsup_{\varepsilon\to 0}\Big[e^{\gamma(t+\theta+h)}\mathbb{E} \times\Big[V\Big(x^\varepsilon(t+\theta+h) - D\left(x^\varepsilon_{t+\theta+h}, \widetilde{r}^\varepsilon(t+\theta+h)\right), t+\theta+h, \widetilde{r}^\varepsilon(t+\theta+h)\Big)\Big] \\ &\qquad - e^{\gamma(t+\theta)}\mathbb{E} \times\Big[V\left(x^\varepsilon(t+\theta) - D\left(x^\varepsilon_{t+\theta}, \widetilde{r}^\varepsilon(t+\theta)\right), t+\theta, \widetilde{r}^\varepsilon(t+\theta)\right)\Big]\Big] \\ &= \limsup_{\varepsilon\to 0}\mathbb{E} \times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\left[\mathbb{L}V\left(\varphi^\varepsilon, s, \widetilde{r}^\varepsilon(s)\right) + \gamma V\left(\varphi^\varepsilon(0) - D\left(\varphi^\varepsilon, \widetilde{r}^\varepsilon(s)\right), s, \widetilde{r}^\varepsilon(t)\right)\right] ds \\ &= \limsup_{\varepsilon\to 0}\mathbb{E} \times\int_{t+\theta}^{t+\theta+h} e^{\gamma s} \times\Big[\mathbb{L}\widetilde{V}\left(\varphi^\varepsilon, s, r^\varepsilon(s)\right) + \gamma V\left(\varphi^\varepsilon(0) - D\left(\varphi^\varepsilon, \widetilde{r}^\varepsilon(t)\right), s, \widetilde{r}^\varepsilon(s)\right)\Big] dt. \end{aligned} \tag{80}$$

By the definition of the operator $\mathbb{L}$, we have

$$\begin{aligned}
&\mathbb{L}\widetilde{V}(\varphi^{\varepsilon},t,r^{\varepsilon}(t)) \\
&= \widetilde{V}_t(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},r^{\varepsilon}(t)),t,r^{\varepsilon}(t)) \\
&\quad + \widetilde{V}_x(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},r^{\varepsilon}(t)),t,r^{\varepsilon}(t)) \\
&\quad \times f(\varphi^{\varepsilon},t,r^{\varepsilon}(t)) \\
&\quad + \sum_{\delta=1}^{m}\int_{\mathbb{R}}\Big\{\widetilde{V}\Big(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},r^{\varepsilon}(t)) \\
&\qquad + b^{(\delta)}(x_{t-}^{\varepsilon},t,r^{\varepsilon}(t),z_{\delta}),t,r^{\varepsilon}(t)\Big) \\
&\qquad - \widetilde{V}(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},r^{\varepsilon}(t)),t,r^{\varepsilon}(t)) \\
&\qquad - \widetilde{V}_x(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},r^{\varepsilon}(t)),t,r^{\varepsilon}(t)) \\
&\qquad \times b^{(\delta)}(x_{t-}^{\varepsilon},t,r^{\varepsilon}(t),z_{\delta})\Big\}\lambda_{\delta}(dz_{\delta}) \\
&\quad + \sum_{j=1}^{N}\gamma_{r^{\varepsilon}(t)j}^{\varepsilon}\widetilde{V}(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},j),t,j) \\
&= \mathbb{L}V(\varphi^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t)) \\
&\quad + V_x(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)),t,\widetilde{r}^{\varepsilon}(t)) \\
&\quad \times \Big[f(\varphi^{\varepsilon},t,r^{\varepsilon}(t))-\widetilde{f}(\varphi^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t))\Big] \\
&\quad + \sum_{\delta=1}^{m}\int_{\mathbb{R}}\Big\{V\Big(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)) \\
&\qquad + b^{(\delta)}(x_{t-}^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t),z_{\delta}),t,\widetilde{r}^{\varepsilon}(t)\Big) \\
&\qquad - V\Big(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)) \\
&\qquad + \widetilde{b}^{(\delta)}(x_{t-}^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t),z_{\delta}),t,\widetilde{r}^{\varepsilon}(t)\Big)\Big\} \\
&\quad \times \lambda_{\delta}(dz_{\delta}) \\
&\quad - \sum_{\delta=1}^{m}\int_{\mathbb{R}}\Big\{V_x(\varphi^{\varepsilon}(0) \\
&\qquad - D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t)),t,\widetilde{r}^{\varepsilon}(t)) \\
&\qquad \times \Big(b^{(\delta)}(x_{t-}^{\varepsilon},t,r^{\varepsilon}(t),z_{\delta}) \\
&\qquad -\widetilde{b}^{(\delta)}(x_{t-}^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t),z_{\delta})\Big)\Big\}\lambda_{\delta}(dz_{\delta}) \\
&\quad + \sum_{j=1}^{N}\widehat{\gamma}_{r^{\varepsilon}(t)j}\widetilde{V}(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},j),t,j) \\
&\quad - \sum_{k=1}^{l}\widetilde{\gamma}_{\overline{r^{\varepsilon}}(t)k}(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},k),t,k).
\end{aligned} \tag{81}$$

This implies that

$$\begin{aligned}
&U(t+h)-U(t) \\
&= \limsup_{\varepsilon\to 0}\mathbb{E} \\
&\quad \times \int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\
&\qquad \times \Big[\mathbb{L}V(\varphi^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s)) \\
&\qquad + \gamma V(\varphi^{\varepsilon}(0) \\
&\qquad -D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)),s,\widetilde{r}^{\varepsilon}(s))\Big]ds \\
&\quad + \limsup_{\varepsilon\to 0}\mathbb{E} \\
&\quad \times \int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\
&\qquad \times V_x(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)),s,\widetilde{r}^{\varepsilon}(s)) \\
&\qquad \times \Big[f(\varphi^{\varepsilon},s,r^{\varepsilon}(s))-\widetilde{f}(\varphi^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s))\Big]ds \\
&\quad + \limsup_{\varepsilon\to 0}\mathbb{E} \\
&\quad \times \int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\
&\qquad \times \Bigg[\sum_{\delta=1}^{m}\int_{\mathbb{R}}\Big[V\Big(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)) \\
&\qquad +b^{(\delta)}(x_{s-}^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s),z_{\delta}),s,\widetilde{r}^{\varepsilon}(s)\Big) \\
&\qquad - V\Big(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)) \\
&\qquad + \widetilde{b}^{(\delta)}(x_{s-}^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s),z_{\delta}), \\
&\qquad s,\widetilde{r}^{\varepsilon}(s)\Big)\Big] \\
&\qquad \times\lambda_{\delta}(dz_{\delta})\Bigg]ds \\
&\quad - \limsup_{\varepsilon\to 0}\mathbb{E} \\
&\quad \times \int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\
&\qquad \times \Bigg\{\sum_{\delta=1}^{m}\int_{\mathbb{R}}\Big[V_x(\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)),s,\widetilde{r}^{\varepsilon}(s)) \\
&\qquad \times \Big(b^{(\delta)}(x_{s-}^{\varepsilon},s,r^{\varepsilon}(s),z_{\delta}) \\
&\qquad -\widetilde{b}^{(\delta)}(x_{s-}^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s),z_{\delta})\Big)\Big] \\
&\qquad \times\lambda_{\delta}(dz_{\delta})\Bigg\}ds
\end{aligned}$$

$$
\begin{aligned}
&+\limsup_{\varepsilon\to 0}\mathbb{E}\\
&\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\\
&\times\left(\sum_{j=1}^{N}\widehat{\gamma}_{r^{\varepsilon}(s)j}\widetilde{V}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},j\right),s,j\right)\right.\\
&\left.-\sum_{k=1}^{l}\widetilde{\gamma}_{\widetilde{r}^{\varepsilon}(s)k}\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},k\right),s,k\right)\right)ds\\
&=: J_1+J_2+J_3+J_4+J_5.
\end{aligned}
\tag{82}
$$

By the definition of $\widetilde{b}$,

$$
\begin{aligned}
&b^{(\delta)}\left(x_{t-}^{\varepsilon},t,r^{\varepsilon}(t),z_{\delta}\right)-\widetilde{b}^{(\delta)}\left(x_{t-}^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t),z_{\delta}\right)\\
&=\sum_{i=1}^{l}\sum_{j=1}^{N_k}b^{(\delta)}\left(x_{t-}^{\varepsilon},t,s_{kj},z_{\delta}\right)\\
&\times\left[I_{\{r^{\varepsilon}(t)=s_{kj}\}}-\mu_j^k I_{\{\widetilde{r}^{\varepsilon}(t)=k\}}\right].
\end{aligned}
\tag{83}
$$

By Assumption 8, we have

$$
\begin{aligned}
J_4&=\limsup_{\varepsilon\to 0}\sum_{\delta=1}^{m}\mathbb{E}\\
&\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\\
&\times V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)\right),s,\widetilde{r}^{\varepsilon}(s)\right)\\
&\times\int_{\mathbb{R}}\left[b^{(\delta)}\left(x_{s-}^{\varepsilon},s,r^{\varepsilon}(s),z_{\delta}\right)\right.\\
&\left.-\widetilde{b}^{(\delta)}\left(x_{s-}^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s),z_{\delta}\right)\right]\\
&\times\lambda_{\delta}\left(dz_{\delta}\right)ds\\
&=\limsup_{\varepsilon\to 0}\sum_{\delta=1}^{m}\sum_{k=1}^{l}\sum_{j=1}^{N_k}\mathbb{E}\\
&\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\\
&\times V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)\right),\right.\\
&\left.s,\widetilde{r}^{\varepsilon}(s)\right)\\
&\times\int_{\mathbb{R}}b^{(\delta)}\left(x_{s-}^{\varepsilon},s,s_{kj},z_{\delta}\right)\\
&\times\left[I_{\{r^{\varepsilon}(s)=s_{kj}\}}\right.\\
&\left.-\mu_j^k I_{\{\widetilde{r}^{\varepsilon}(s)=k\}}\right]\\
&\times\lambda_{\delta}\left(dz_{\delta}\right)ds\\
&\le\limsup_{\varepsilon\to 0}\sum_{\delta=1}^{m}\sum_{k=1}^{l}\sum_{j=1}^{N_k}\left[\mathbb{E}\left|\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\right.\right.\\
&\times V_x\left(\varphi^{\varepsilon}(0)-D\left(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(s)\right),\right.\\
&\left.s,\widetilde{r}^{\varepsilon}(s)\right)\\
&\times\int_{\mathbb{R}}b^{(\delta)}\left(x_{s-}^{\varepsilon},s,s_{kj},z_{\delta}\right)\\
&\times\left[I_{\{r^{\varepsilon}(s)=s_{kj}\}}\right.\\
&\left.-\mu_j^k I_{\{\widetilde{r}^{\varepsilon}(s)=k\}}\right]\\
&\left.\left.\times\lambda_{\delta}\left(dz_{\delta}\right)ds\right|^2\right]^{1/2}.
\end{aligned}
\tag{84}
$$

By the argument of Lemma 7.14 in [4], the right side of the above inequality is equivalent to 0; that is, $J_4=0$. Similarly, by mean-value theorem, we can show that there exists $\eta_t^{(\delta)}$ which is between $\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t))+b^{(\delta)}(x_{t-}^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t),z_{\delta})$ and $\varphi^{\varepsilon}(0)-D(\varphi^{\varepsilon},\widetilde{r}^{\varepsilon}(t))+\widetilde{b}^{(\delta)}(x_{t-}^{\varepsilon},t,\widetilde{r}^{\varepsilon}(t),z_{\delta})$ such that

$$
\begin{aligned}
J_3&=\lim_{\varepsilon\to 0}\sum_{\delta=1}^{m}\mathbb{E}\\
&\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}\left\{\int_{\mathbb{R}}V_x\left(\eta_s\right)\right.\\
&\times\left[b^{(\delta)}\left(x_{s-}^{\varepsilon},s,r^{\varepsilon}(s),z_{\delta}\right)\right.\\
&\left.-\widetilde{b}^{(\delta)}\left(x_{s-}^{\varepsilon},s,\widetilde{r}^{\varepsilon}(s),z_{\delta}\right)\right]\\
&\left.\times\lambda_{\delta}\left(dz_{\delta}\right)\right\}ds\\
&=\lim_{\varepsilon\to 0}\sum_{\delta=1}^{m}\sum_{k=1}^{l}\sum_{j=1}^{N_k}\mathbb{E}\\
&\times\int_{t+\theta}^{t+\theta+h} e^{\gamma s}V_x\left(\eta_s\right)\\
&\times\int_{\mathbb{R}}b^{(\delta)}\left(x_{s-}^{\varepsilon},s,s_{kj},z_{\delta}\right)\\
&\times\left[I_{\{r^{\varepsilon}(s)=s_{kj}\}}-\mu_j^k I_{\{\widetilde{r}^{\varepsilon}(s)=k\}}\right]\\
&\times\lambda_{\delta}\left(dz_{\delta}\right)ds\\
&\le\lim_{\varepsilon\to 0}\sum_{\delta=1}^{m}\sum_{k=1}^{l}\sum_{j=1}^{N_k}\left[\mathbb{E}\left|\int_{t+\theta}^{t+\theta+h} e^{\gamma s}V_x\left(\eta_t\right)\right.\right.\\
&\times\int_{\mathbb{R}}b^{(\delta)}\left(x_{s-}^{\varepsilon},s,s_{kj},z_{\delta}\right)
\end{aligned}
$$

$$\times \left[I_{\{r^\varepsilon(s)=s_{kj}\}} - \mu_j^k I_{\{\overline{r}^\varepsilon(s)=k\}} \right]$$
$$\left. \times \lambda_\delta(dz_\delta)ds \right|^2 \Bigg]^{1/2}. \quad (85)$$

By the argument of Lemma 7.14 in [4], we have $J_3 = 0$. Similar to the proof of Theorem 7, we derive $J_2 = 0$, $J_5 = 0$. Therefore, we arrive at

$$\begin{aligned} &U(t+h) - U(t) \\ &= \lim_{\varepsilon \to 0} \mathbb{E} \int_{t+\theta}^{t+\theta+h} e^{\gamma s} \\ &\quad \times [\mathbb{L}V(\varphi^\varepsilon, s, \overline{r}^\varepsilon(s)) \\ &\quad + \gamma V(\varphi^\varepsilon(0) - D(\varphi^\varepsilon, \overline{r}^\varepsilon(s)), s, \overline{r}^\varepsilon(s))]\, ds \\ &\le 0. \end{aligned} \quad (86)$$

Then,

$$U(t+h) \le U(t). \quad (87)$$

Similar to the proof of Theorem 7, we get

$$\mathbb{E}|x^\varepsilon(t)|^p \le \frac{c_2(1+\kappa)^p}{c_1(1-\kappa)^p} \|\xi\|^p e^{-\overline{\nu}t}. \quad (88)$$

The proof is therefore completed. □

5. Examples

We will give two examples to illustrate our theory.

Example 1. Let $r^\varepsilon(\cdot)$ be a Markov chain generated by Γ^ε given in (14) with

$$\overline{\Gamma} = \begin{pmatrix} -1 & 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 2 & 1 & -3 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \quad (89)$$

$$\widehat{\Gamma} = \begin{pmatrix} -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{pmatrix}. \quad (90)$$

The generator $\overline{\Gamma}$ is made up of two irreducible blocks; by

$$(\pi_1 \ \ \pi_2 \ \ \pi_3) \begin{pmatrix} -1 & 0 & 1 \\ 1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix} = 0 \quad (91)$$

and $\pi_1 + \pi_2 + \pi_3 = 1$, we get $\mu^1 = (5/8, 1/8, 1/4)$. In the same way, by

$$(\pi_4 \ \ \pi_5) \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = 0 \quad (92)$$

and $\pi_4 + \pi_5 = 1$, we have $\mu^2 = (1/2, 1/2)$. So,

$$\begin{aligned} \overline{\Gamma} = \overline{\mu}\widehat{\Gamma}\mathbf{1} &= \begin{pmatrix} \frac{5}{8} & \frac{1}{8} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &\quad \times \begin{pmatrix} -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{3}{8} & \frac{3}{8} \\ 1 & -1 \end{pmatrix}. \end{aligned} \quad (93)$$

Consider a one-dimensional neutral stochastic functional differential equation as follows:

$$\begin{aligned} &d[\varphi^\varepsilon(0) - D(\varphi^\varepsilon, r^\varepsilon(t))] \\ &\quad = f(\varphi^\varepsilon, r^\varepsilon(t))\, dt + g(\varphi^\varepsilon, r^\varepsilon(t))\, dw(t), \end{aligned} \quad (94)$$

with

$$D(\varphi, s_{11}) = -0.6 \int_{-1}^0 \varphi(\theta)\, d\theta,$$
$$D(\varphi, s_{12}) = -0.2 \int_{-1}^0 \varphi(\theta)\, d\theta,$$
$$D(\varphi, s_{13}) = -0.4 \int_{-1}^0 \varphi(\theta)\, d\theta,$$
$$f(\varphi, s_{11}) = -16\varphi(0) - 8\cos[\varphi(0)],$$
$$f(\varphi, s_{12}) = 8\varphi(0) + 4\cos[\varphi(0)],$$
$$f(\varphi, s_{13}) = 16\varphi(0),$$
$$g(\varphi, s_{11}) = \frac{\sqrt{10}}{10} \int_{-1}^0 \varphi(\theta)\, d\theta \cos\left[\int_{-1}^0 \varphi(\theta)\, d\theta\right],$$
$$g(\varphi, s_{12}) = -\frac{\sqrt{2}}{2} \int_{-1}^0 \varphi(\theta)\, d\theta \sin\left[\int_{-1}^0 \varphi(\theta)\, d\theta\right],$$
$$g(\varphi, s_{13}) = \frac{\sqrt{3}}{2} \int_{-1}^0 \varphi(\theta)\, d\theta,$$
$$D(\varphi, s_{21}) = 0.5 \int_{-1}^0 \varphi(\theta)\, d\theta,$$
$$D(\varphi, s_{22}) = 0.5 \int_{-1}^0 \varphi(\theta)\, d\theta,$$
$$f(\varphi, s_{21}) = -2\varphi(0), \qquad f(\varphi, s_{22}) = -2\varphi(0),$$

$$g(\varphi, s_{21}) = \frac{\int_{-1}^{0} \varphi(\theta)\, d\theta \sin\left[\int_{-1}^{0} \varphi(\theta)\, d\theta\right]}{4\sqrt{2}},$$
$$g(\varphi, s_{22}) = \frac{\int_{-1}^{0} \varphi(\theta)\, d\theta \cos\left[\int_{-1}^{0} \varphi(\theta)\, d\theta\right]}{4\sqrt{2}}. \tag{95}$$

For any $\varphi \in L^2_{\mathcal{F}_t}([-1,0];\mathbb{R})$ and $\kappa = \max\{0.6, 0.2, 0.4\} = 0.6$, applying the Hölder inequality yields

$$\mathbb{E}|D(\varphi, i)|^2 \le 0.6^2 \sup_{-1\le\theta\le 0} e^{\nu\theta}\mathbb{E}\left|\int_{-1}^{0} \varphi(\theta)\, d\theta\right|^2 \le 0.36\|\varphi\|^2, \tag{96}$$

which implies condition (24). Then, the limit equation is

$$d[\widetilde{\varphi}(0) - D(\widetilde{\varphi}, \widetilde{r}(t))] = f(\widetilde{\varphi}, \widetilde{r}(t))\, dt + g(\widetilde{\varphi}, \widetilde{r}(t))\, dw(t), \tag{97}$$

where $\widetilde{r}$ is the Markov chain generated by $\widetilde{\Gamma}$ and

$$\widetilde{D}(\varphi, 1) = -0.5 \int_{-1}^{0} \varphi(\theta)\, d\theta,$$
$$\widetilde{D}(\varphi, 2) = 0.5 \int_{-1}^{0} \varphi(\theta)\, d\theta,$$
$$\widetilde{f}(\varphi, 1) = -5\varphi(0), \qquad \widetilde{f}(\varphi, 2) = -2\varphi(0),$$
$$\widetilde{g}(\varphi, 1) = \frac{1}{2}\int_{-1}^{0} \varphi(\theta)\, d\theta, \qquad \widetilde{g}(\varphi, 2) = \frac{1}{4}\int_{-1}^{0} \varphi(\theta)\, d\theta. \tag{98}$$

We define $V(x,1) = 2x^2$, $V(x,2) = x^2$. And by simple calculation, we can get

$$\mathscr{L}V(\varphi, 1) \le -20\frac{3}{8}\varphi^2(0) + \frac{13}{32}\left|\int_{-1}^{0} \varphi(\theta)\, d\theta\right|^2,$$
$$\mathscr{L}V(\varphi, 2) \le -\frac{5}{2}\varphi^2(0) + \frac{13}{16}\left|\int_{-1}^{0} \varphi(\theta)\, d\theta\right|^2. \tag{99}$$

Consequently,

$$\begin{aligned}\max_{i=1,2}\mathscr{L}V(\varphi, i) &\le -\frac{5}{2}\varphi^2(0) + \frac{13}{16}\left|\int_{-1}^{0} \varphi(\theta)\, d\theta\right|^2 \\ &= -\frac{5}{4}\left[\max_{i=1,2} V(x,i)\right] + \frac{13}{16}\left[\min_{i=1,2} V(x,i)\right].\end{aligned} \tag{100}$$

It is easy to find a $q > 1$ such that $5/4 - 13q/16 > 0$. Therefore, for any $\phi \in L^2_{\mathcal{F}_t}([-1,0];\mathbb{R})$ satisfying $\mathbb{E}[\min_{i\in\mathbb{S}}\phi(\theta)] \le q\mathbb{E}[\max_{i\in\mathbb{S}}\phi(0)]$ on $-1 \le \theta \le 0$, (100) yields

$$\mathbb{E}\left[\max_{i\in\mathbb{S}} \mathscr{L}V(\varphi, i)\right] \le -\left(\frac{5}{4} - \frac{13q}{16}\right)\mathbb{E}\left[\max_{i=1,2} V(x,i)\right]. \tag{101}$$

Hence, by Theorem 7, the solution $x^\varepsilon(t)$ is mean square stable when ε is sufficiently small.

Example 2. Let $r^\varepsilon(\cdot)$ be a Markov chain generated by

$$\Gamma^\varepsilon = \frac{1}{\varepsilon}\widetilde{\Gamma} + \widehat{\Gamma} = \frac{1}{\varepsilon}\begin{pmatrix} -2 & 0 & 2 & 0 \\ 1 & -2 & 0 & 1 \\ 0 & 2 & -2 & 0 \\ 0 & 1 & 1 & -2 \end{pmatrix}. \tag{102}$$

Here, we set $\widehat{\Gamma} = 0$. By a similar way, we get the stationary distribution $\mu = (2/11, 4/11, 3/11, 2/11)$.

Consider the following one-dimensional equation:

$$\begin{aligned} &d[\varphi^\varepsilon(0) - D(\varphi^\varepsilon, r^\varepsilon(t))] \\ &\quad = f(\varphi^\varepsilon, r^\varepsilon(t))\, dt + \int_0^\infty \sigma(r^\varepsilon(t), z)\, x^\varepsilon_{t-}\widetilde{N}(dt, dz)\end{aligned} \tag{103}$$

with

$$D(\varphi, 1) = -0.9\int_{-1}^{0} \varphi(\theta)\, d\theta,$$
$$D(\varphi, 2) = -0.4\int_{-1}^{0} \varphi(\theta)\, d\theta,$$
$$D(\varphi, 3) = -0.5\int_{-1}^{0} \varphi(\theta)\, d\theta,$$
$$D(\varphi, 4) = -0.3\int_{-1}^{0} \varphi(\theta)\, d\theta, \tag{104}$$
$$f(\varphi, 1) = 2\sin[\varphi(0)], \qquad f(\varphi, 2) = -\frac{11}{2}\varphi(0),$$
$$f(\varphi, 3) = -\frac{11}{3}\varphi(0), \qquad f(\varphi, 4) = -2\sin[\varphi(0)].$$

Let

$$\alpha(z) = \frac{2}{11}\sigma(1,z) + \frac{4}{11}\sigma(2,z) + \frac{3}{11}\sigma(3,z) + \frac{2}{11}\sigma(4,z),$$
$$\int_0^\infty \alpha^2(z)\,\lambda(dz) < 2. \tag{105}$$

For any $\varphi \in L^2_{\mathcal{F}_t}([-1,0];\mathbb{R})$ and $\kappa = \max\{0.9, 0.4, 0.5, 0.3\} = 0.9$, applying the Hölder inequality yields

$$\mathbb{E}|D(\varphi, i)|^2 \le 0.9^2 \sup_{-1\le\theta\le 0} e^{\overline{\nu}\theta}\mathbb{E}\left|\int_{-1}^{0} \varphi(\theta)\, d\theta\right|^2 \le 0.81\|\varphi\|^2, \tag{106}$$

which implies condition (67). Then, the limit equation is

$$\begin{aligned} &d\left[\varphi(0) + 0.5\int_{-1}^{0} \varphi(\theta)\, d\theta\right] \\ &\quad = -3\varphi(0)\, dt + \int_0^\infty \alpha(z)\,\widetilde{x}_{t-}\widetilde{N}(dt, dz).\end{aligned} \tag{107}$$

Let $V(x) = x^2$; then

$$\mathbb{L}V(\varphi, i) \leq -6\varphi^2(0) + \int_0^\infty \alpha^2(z)\lambda(dz)\left|\int_{-1}^0 \varphi(\theta)\,d\theta\right|^2. \tag{108}$$

We can find a $q > 1$ such that $6-2q > 0$. Therefore, for any $\phi \in L^2_{\mathscr{F}_t}([-1,0];\mathbb{R})$ satisfying $\mathbb{E}[\min_{i\in\overline{\mathbb{S}}}\phi(\theta)] \leq q\mathbb{E}[\max_{i\in\overline{\mathbb{S}}}\phi(0)]$ on $-1 \leq \theta \leq 0$, (108) yields

$$\mathbb{E}\left[\max_{i\in\overline{\mathbb{S}}}\mathbb{L}V(\varphi, i)\right] \leq -(6-2q)\,\mathbb{E}\left[\max_{i=1,2} V(x,i)\right]. \tag{109}$$

Hence, by Theorem 12, the solution $x^\varepsilon(t)$ is mean square stable.

Conflict of Interests

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

Acknowledgment

This paper was supported by the National Science Foundation of China with Grant no. 61374085.

References

[1] Y. Shen and J. Wang, "Noise-induced stabilization of the recurrent neural networks with mixed time-varying delays and Markovian-switching parameters," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1857–1862, 2007.

[2] Y. Shen and J. Wang, "Almost sure exponential stability of recurrent neural networks with Markovian switching," *IEEE Transactions on Neural Networks*, vol. 20, no. 5, pp. 840–855, 2009.

[3] H. A. Simon and A. Ando, "Aggregation of variables in dynamic systems," *Econometrica*, vol. 29, pp. 111–138, 1961.

[4] G. G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Singular Perturbations Approach*, Springer, New York, NY, USA, 1998.

[5] H. Chen, C. Zhu, and Y. Zhang, "A note on exponential stability for impulsive neutral stochastic partial functional differential equations," *Applied Mathematics and Computation*, vol. 227, pp. 139–147, 2014.

[6] G. Hu and K. Wang, "Stability in distribution of neutral stochastic functional differential equations with Markovian switching," *Journal of Mathematical Analysis and Applications*, vol. 385, no. 2, pp. 757–769, 2012.

[7] S. Janković, M. Vasilova, and M. Krstić, "Some analytic approximations for neutral stochastic functional differential equations," *Applied Mathematics and Computation*, vol. 217, no. 8, pp. 3615–3623, 2010.

[8] V. Kolmanovskii, N. Koroleva, T. Maizenberg, X. Mao, and A. Matasov, "Neutral stochastic differential delay equations with Markovian switching," *Stochastic Analysis and Applications*, vol. 21, no. 4, pp. 819–847, 2003.

[9] D. Li and D. Xu, "Attracting and quasi-invariant sets of stochastic neutral partial functional differential equations," *Acta Mathematica Scientia B*, vol. 33, no. 2, pp. 578–588, 2013.

[10] X. Mao, "Exponential stability in mean square of neutral stochastic differential-functional equations," *Systems & Control Letters*, vol. 26, no. 4, pp. 245–251, 1995.

[11] X. Mao, "Razumikhin-type theorems on exponential stability of neutral stochastic functional-differential equations," *SIAM Journal on Mathematical Analysis*, vol. 28, no. 2, pp. 389–401, 1997.

[12] Y. Song and Y. Shen, "New criteria on asymptotic behavior of neutral stochastic functional differential equations," *Automatica*, vol. 49, no. 2, pp. 626–632, 2013.

[13] F. Wu, S. Hu, and C. Huang, "Robustness of general decay stability of nonlinear neutral stochastic functional differential equations with infinite delay," *Systems & Control Letters*, vol. 59, no. 3-4, pp. 195–202, 2010.

[14] S. Zhou and S. Hu, "Razumikhin-type theorems of neutral stochastic functional differential equations," *Acta Mathematica Scientia B*, vol. 29, no. 1, pp. 181–190, 2009.

[15] G. Badowski and G. G. Yin, "Stability of hybrid dynamic systems containing singularly perturbed random processes," *IEEE Transactions on Automatic Control*, vol. 47, no. 12, pp. 2021–2032, 2002.

[16] J. Hu, X. Mao, and C. Yuan, "Razumikhin-type theorems on exponential stability of SDDEs containing singularly perturbed random processes," *Abstract and Applied Analysis*, vol. 2013, Article ID 854743, 12 pages, 2013.

[17] A. A. Pervozvanskii and V. G. Gaitsgori, *Theory of Suboptimal Decisions: Decomposition and Aggregation*, Kluwer Academic, Dordrecht, The Netherlands, 1988.

[18] F. Wu, G. G. Yin, and L. Y. Wang, "Stability of a pure random delay system with two-time-scale Markovian switching," *Journal of Differential Equations*, vol. 253, no. 3, pp. 878–905, 2012.

[19] F. Wu, G. Yin, and L. Y. Wang, "Moment exponential stability of random delay systems with two-time-scale Markovian switching," *Nonlinear Analysis: Real World Applications*, vol. 13, no. 6, pp. 2476–2490, 2012.

[20] G. Yin and H. Yang, "Two-time-scale jump-diffusion models with Markovian switching regimes," *Stochastics and Stochastics Reports*, vol. 76, no. 2, pp. 77–99, 2004.

[21] C. Yuan and G. Yin, "Stability of hybrid stochastic delay systems whose discrete components have a large state space: a two-time-scale approach," *Journal of Mathematical Analysis and Applications*, vol. 368, no. 1, pp. 103–119, 2010.

[22] R. Z. Khasminskii, G. Yin, and Q. Zhang, "Asymptotic expansions of singularly perturbed systems involving rapidly fluctuating Markov chains," *SIAM Journal on Applied Mathematics*, vol. 56, no. 1, pp. 277–293, 1996.

[23] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, The MIT Press, Cambridge, Mass, USA, 1984.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 258093, 6 pages
http://dx.doi.org/10.1155/2014/258093

Research Article

Asymptotic Stabilization by State Feedback for a Class of Stochastic Nonlinear Systems with Time-Varying Coefficients

Hui Wang,[1] Wuquan Li,[1,2] and Xiuhong Wang[1]

[1] *School of Mathematics and Statistics Science, Ludong University, Yantai 264025, China*
[2] *Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*

Correspondence should be addressed to Wuquan Li; sea81@126.com

Received 16 December 2013; Accepted 31 January 2014; Published 11 March 2014

Academic Editor: Xuejun Xie

This paper investigates the problem of state-feedback stabilization for a class of upper-triangular stochastic nonlinear systems with time-varying control coefficients. By introducing effective coordinates, the original system is transformed into an equivalent one with tunable gain. After that, by using the low gain homogeneous domination technique and choosing the low gain parameter skillfully, the closed-loop system can be proved to be globally asymptotically stable in probability. The efficiency of the state-feedback controller is demonstrated by a simulation example.

1. Introduction

Consider a class of upper-triangular stochastic nonlinear systems with time-varying control coefficients described by

$$\begin{aligned}
dx_1 &= \left(d_1(t)\,x_2 + f_1(\tilde{x}_3)\right)dt + g_1^T(\tilde{x}_3)\,d\omega,\\
dx_2 &= \left(d_2(t)\,x_3 + f_2(\tilde{x}_4)\right)dt + g_2^T(\tilde{x}_4)\,d\omega,\\
&\ \ \vdots\\
dx_{n-2} &= \left(d_{n-2}(t)\,x_{n-1} + f_{n-2}(\tilde{x}_n)\right)dt + g_{n-2}^T(\tilde{x}_n)\,d\omega,\\
dx_{n-1} &= d_{n-1}(t)\,x_n dt,\\
dx_n &= d_n(t)\,u dt,
\end{aligned} \tag{1}$$

where $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $u \in \mathbb{R}$ are the measurable state and the input of system, respectively. $\tilde{x}_i = (x_i, \ldots, x_n)^T$. ω is an r-dimensional standard Wiener process defined on a probability space $(\Omega, \mathcal{F}, P)$, with Ω being a sample space, $\mathcal{F}$ being a filtration, and P being a probability measure. The functions $f_i : \mathbb{R}^{n-i-1} \to \mathbb{R}$ and $g_i : \mathbb{R}^{n-i-1} \to \mathbb{R}^r$, $i = 1, \ldots, n-2$, are assumed to be $\mathscr{C}^1$ with their arguments and $f_i(0) = 0$, $g_i(0) = 0$. $d_i : R_+ \to R$, $i = 1, \ldots, n$, are unknown time-varying control coefficients with known sign.

In recent years, the global controller design for stochastic nonlinear systems has been attracting more and more attention. According to the difference of selected Lyapunov functions, the existing literature on controller design can be mainly divided into two types. One type is to derive the backstepping controller design by using quadratic Lyapunov function and a risk-sensitive cost criterion [1–3]. Another essential improvement belongs to Krstić and Deng. By introducing the quartic Lyapunov function, [4–12] present asymptotical stabilization control under the assumption that the nonlinearities equal zero at the equilibrium point of the open-loop system. Subsequently, for several classes of stochastic high-order nonlinear systems, by combining Krstić and Deng's method with stochastic analysis, [13, 14] study the problem of state-feedback stabilization and the output-feedback stabilization problem is considered in [15, 16].

The study of stabilization control for upper-triangular nonlinear systems has long been recognized as difficult due to the inherent nonlinearity. In the existing literature, most results are established using the nested-saturation method [17, 18] and forwarding technique [19]. When no a priori information of the system nonlinearities is known, the work [20] proposes a universal stabilizer for feedforward nonlinear systems by employing a switching controller. Note that the listed results above do not consider the stochastic noise. However, from both practical and theoretical points of view, it is more important to study the control of upper-triangular stochastic nonlinear systems with time-varying control coefficients. Therefore, in this paper, under some appropriate assumptions, we consider the stabilization for system (1). To the best of the authors' knowledge, there are not any results about this topic.

In this paper, based on the low gain homogeneous domination technique, for system (1), we design a stabilization state-feedback controller, under which the closed-loop systems can be proved to be globally asymptotically stable in probability.

The contributions of this paper are highlighted as follows.

(i) This paper is the first result about state-feedback stabilization of upper-triangular stochastic nonlinear systems with time-varying control coefficients.

(ii) Due to the complex of upper-triangular system structure, how to deal with stochastic noise and time-varying control coefficients in the controller design is a nontrivial work.

The remainder of this paper is organized as follows. Section 2 offers some preliminary results. The state-feedback controller is designed and analyzed in Section 3. After that, in Section 4, a simulation example is presented to show the effectiveness of the state-feedback controller. Finally, the paper is concluded in Section 5.

2. Preliminary Results

The following notation will be used throughout the paper. $\mathbb{R}_+$ denotes the set of all nonnegative real numbers. For a given vector or matrix X, X^T denotes its transpose, $\mathrm{Tr}\{X\}$ denotes its trace when X is square, and $|X|$ is the Euclidean norm of a vector X. $\mathscr{C}^i$ denotes the set of all functions with continuous ith partial derivatives. $\mathscr{K}$ denotes the set of all functions: $\mathbb{R}_+ \to \mathbb{R}_+$, which are continuous, strictly increasing, and vanishing at zero; $\mathscr{K}_\infty$ denotes the set of all functions which are of class $\mathscr{K}$ and unbounded; $\mathscr{KL}$ denotes the set of all functions $\beta(s,t)$: $\mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, which are of $\mathscr{K}$ for each fixed t and decrease to zero as $t \to \infty$ for each fixed s.

Consider the following stochastic nonlinear system:

$$dx = f(x)\,dt + g^T(x)\,d\omega, \tag{2}$$

where $x \in \mathbb{R}^n$ is the state of the system and ω is an r-dimensional standard Wiener process defined on the probability space $(\Omega, \mathscr{F}, P)$. The Borel measurable functions $f: \mathbb{R}^n \to \mathbb{R}^n$ and $g^T: \mathbb{R}^n \to \mathbb{R}^{n\times r}$ are local Lipschitz in $x \in \mathbb{R}^n$.

The following definitions and lemma will be used throughout the paper.

Definition 1 (see [5]). For any given $V(x) \in \mathscr{C}^2$ associated with stochastic system (2), the differential operator $\mathscr{L}$ is defined as

$$\mathscr{L}V(x) \triangleq \frac{\partial V(x)}{\partial x} f(x) + \frac{1}{2}\mathrm{Tr}\left\{g(x)\frac{\partial^2 V(x)}{\partial x^2} g^T(x)\right\}. \tag{3}$$

Definition 2 (see [5]). For the stochastic system (2) with $f(0) = 0$, $g(0) = 0$, the equilibrium $x(t) = 0$ of (2) is globally asymptotically stable (GAS) in probability if, for any $\varepsilon > 0$, there exists a class $\mathscr{KL}$ function $\beta(\cdot,\cdot)$ such that $P\{|x(t)| < \beta(|x_0|, t)\} \ge 1-\varepsilon$ for any $t \ge 0$ and $x_0 \in \mathbb{R}^n \setminus \{0\}$.

Lemma 3 (see [5]). *Consider the stochastic system* (2); *if there exist a $\mathscr{C}^2$ function $V(x)$, class $\mathscr{K}_\infty$ functions α_1 and α_2, constants $c_1 > 0$ and $c_2 \ge 0$, and a nonnegative function $W(x)$ such that*

$$\alpha_1(|x|) \le V(x) \le \alpha_2(|x|), \qquad \mathscr{L}V \le -c_1 W(x) + c_2, \tag{4}$$

then

(a) *for* (2), *there exists an almost surely unique solution on* $[0,\infty)$;

(b) *when $c_2 = 0$, $f(0) = 0$, $g(0) = 0$, and $W(x) = \alpha_3(|x|)$, where $\alpha_3(\cdot)$ is a class $\mathscr{K}$ function, then the equilibrium $x = 0$ is GAS in probability and $P\{\lim_{t\to\infty}|x(t)| = 0\} = 1$.*

3. Controller Design and Stability Analysis

The following assumptions are made on system (1).

Assumption 1. For $i = 1,\ldots,n$, there exists a constant $b > 0$ such that

$$\begin{aligned}|f_i(\tilde{x}_{i+2})| &\le b(|x_{i+2}| + \cdots + |x_n|),\\ |g_i(\tilde{x}_{i+2})| &\le b(|x_{i+2}| + \cdots + |x_n|).\end{aligned} \tag{5}$$

Assumption 2. Without loss of generality, the sign of $d_i(t)$ is assumed to be positive, and there exist known positive constants λ_i and μ_i such that, for any $t \in \mathbb{R}^+$ and $i = 1,\ldots,n$,

$$0 < \lambda_i \le d_i(t) \le \mu_i. \tag{6}$$

Remark 4. From Assumption 1, the system investigated has an upper-triangular form. Due to the complex of upper-triangular system structure and the effect of stochastic noise, the stabilization of such systems is usually very difficult. In this paper, by using the low gain homogeneous domination approach, the state-feedback stabilization problem is investigated for the first time.

Remark 5. By Assumption 2, we know that $d_i(t)$s are time-varying control coefficients; how to effectively deal with them in the design process is nontrivial work.

Firstly, introduce the following coordinate transformation:

$$z_i = \frac{x_i}{\varepsilon^{i-1}}, \quad v = \frac{u}{\varepsilon^n}, \quad i = 1,\ldots,n, \tag{7}$$

where $0 < \varepsilon < 1$ is a parameter to be designed. System (1) can be rewritten as

$$\begin{aligned} dz_1 &= \left(\varepsilon d_1(t) z_2 + \overline{f}_1(\tilde{z}_3)\right) dt + \overline{g}_1^T(\tilde{z}_3) d\omega, \\ dz_2 &= \left(\varepsilon d_2(t) z_3 + \overline{f}_2(\tilde{z}_4)\right) dt + \overline{g}_2^T(\tilde{z}_4) d\omega, \\ &\vdots \\ dz_{n-2} &= \left(\varepsilon d_{n-2}(t) z_{n-1} + \overline{f}_{n-2}(\tilde{z}_n)\right) dt + \overline{g}_{n-2}^T(\tilde{z}_n) d\omega, \\ dz_{n-1} &= \varepsilon d_{n-1}(t) z_n dt, \\ dz_n &= \varepsilon d_n(t) v dt, \end{aligned} \tag{8}$$

where $\overline{f}_i(\tilde{z}_{i+2}) = f_i(\tilde{x}_{i+2})/\varepsilon^{i-1}$, $\overline{g}_i(\tilde{z}_{i+2}) = g_i(\tilde{x}_{i+2})/\varepsilon^{i-1}$.

The nominal system for (8) is

$$\begin{aligned} dz_1 &= d_1(t) z_2, \\ dz_2 &= d_2(t) z_3, \\ &\vdots \\ dz_{n-2} &= d_{n-2}(t) z_{n-1}, \\ dz_{n-1} &= d_{n-1}(t) z_n dt, \\ dz_n &= d_n(t) v dt. \end{aligned} \tag{9}$$

Theorem 6. *For nominal system* (9), *with Assumption 2, one can design a stabilizing state-feedback controller to guarantee that*

(1) *the closed-loop system has an almost surely unique solution on* $[0, \infty)$*;*

(2) *the equilibrium of the closed-loop system is GAS in probability.*

Proof. The controller design process proceeds step by step.

Step 1. Defining $\xi_1 = z_1$ and choosing $V_1 = (1/4) z_1^4$, from (9), it follows that

$$\mathscr{L}V_1 \le d_1(t) z_1^3 z_2. \tag{10}$$

Suppose that $z_2^* = -z_1\alpha_1 = -\xi_1\alpha_1$, where $\alpha_1 \ge 0$ is a constant to be chosen. Thus, by Assumption 2, we have

$$d_1(t) z_1^3 z_2^* \le \lambda_1 z_1^3 z_2^* \le 0. \tag{11}$$

By (10) and (11), one gets

$$\mathscr{L}V_1 \le \lambda_1 z_1^3 z_2^* + d_1(t) z_1^3 (z_2 - z_2^*). \tag{12}$$

Choosing the virtual smooth control z_2^* as

$$z_2^* = -\frac{n}{\lambda_1}\xi_1 \triangleq -\xi_1\alpha_1, \tag{13}$$

which substitutes into (12), yields

$$\mathscr{L}V_1 \le -n\xi_1^4 + d_1(t) z_1^3 (z_2 - z_2^*). \tag{14}$$

Deductive Step. Assume that, at step $k-1$, there are a $\mathscr{C}^2$, proper and positive definite Lyapunov function V_{k-1}, and the virtual controllers z_j^* defined by

$$\begin{aligned} z_1^* &= 0, & \xi_1 &= z_1 - z_1^*, \\ z_2^* &= -\xi_1\alpha_1, & \xi_2 &= z_2 - z_2^*, \\ &\vdots \\ z_k^* &= -\xi_{k-1}\alpha_{k-1}, & \xi_k &= z_k - z_k^*, \end{aligned} \tag{15}$$

where $\alpha_i \ge 0$, $1 \le i \le k-1$, are positive constants, such that

$$\mathscr{L}V_{k-1}(\overline{z}_{k-1}) \le -(n-k+2)\sum_{i=1}^{k-1}\xi_i^4 + d_{k-1}(t)\xi_{k-1}^3(z_k - z_k^*), \tag{16}$$

where $\overline{z}_{k-1} = (z_1, \ldots, z_{k-1})^T$. To complete the induction, at the kth step, one can choose the following Lyapunov function:

$$V_k(\overline{z}_k) = V_{k-1}(\overline{z}_{k-1}) + \frac{1}{4}\xi_k^4, \tag{17}$$

where $\overline{z}_k = (z_1, \ldots, z_k)^T$.

By (15)–(17), one has

$$\begin{aligned} \mathscr{L}V_k(\overline{z}_k) \le{}& -(n-k+2)\sum_{i=1}^{k-1}\xi_i^4 + d_{k-1}(t)\xi_{k-1}^3\xi_k \\ &+ \xi_k^3\left(d_k(t) z_{k+1} - \sum_{i=1}^{k-1}\frac{\partial z_k^*}{\partial z_i} d_i(t) z_{i+1}\right). \end{aligned} \tag{18}$$

By using Young's inequality and Assumption 2, one has

$$\begin{aligned} d_{k-1}(t)\xi_{k-1}^3\xi_k &\le \frac{1}{2}\xi_{k-1}^4 + c_k\xi_k^4, \\ -\xi_k^3\sum_{i=1}^{k-1}\frac{\partial z_k^*}{\partial z_i} d_i(t) z_{i+1} &\le c_{k1}|\xi_k|^3\sum_{i=1}^{k}|z_i| \le \frac{1}{2}\sum_{i=1}^{k-1}\xi_i^4 + \hat{c}_k\xi_k^4, \end{aligned} \tag{19}$$

where $c_k > 0$, $c_{k1} > 0$, and $\hat{c}_k > 0$ are constants. Suppose that

$$z_{k+1}^* = -\xi_k\alpha_k, \tag{20}$$

where $\alpha_k \ge 0$ is a constant to be chosen. Then, by Assumption 2, one has

$$d_k(t)\xi_k^3 z_{k+1}^* \le \lambda_k\xi_k^3 z_{k+1}^*. \tag{21}$$

Substituting (19) and (21) into (18) yields

$$\begin{aligned} \mathscr{L}V_k(\overline{z}_k) \le{}& -(n-k+1)\sum_{i=1}^{k-1}\xi_i^4 + d_k(t)\xi_k^3(z_{k+1} - z_{k+1}^*) \\ &+ \lambda_k\xi_k^3 z_{k+1}^* + (c_k + \hat{c}_k)\xi_k^4. \end{aligned} \tag{22}$$

Choosing the virtual smooth control

$$z_{k+1}^{*} = -\frac{1}{\lambda_k}\left(n-k+1+c_k+\widehat{c}_k\right)\xi_k \triangleq -\xi_k\alpha_k, \quad (23)$$

which substitutes into (22), yields

$$\mathscr{L}V_k\left(\overline{z}_k\right) \le -\left(n-k+1\right)\sum_{i=1}^{k}\xi_i^4 + d_k\left(t\right)\xi_k^3\left(z_{k+1}-z_{k+1}^{*}\right). \quad (24)$$

Step n. By choosing the actual control law

$$v = -\xi_n\alpha_n, \quad (25)$$

where $\alpha_n \ge 0$ is a constant and $\xi_n = x_n - x_n^{*}$, one gets

$$\mathscr{L}V_n\left(\overline{z}_n\right) \le -\sum_{i=1}^{n}\xi_i^4, \quad (26)$$

where

$$V_n\left(\overline{z}_n\right) = V_{n-1}\left(\overline{z}_{n-1}\right) + \frac{1}{4}\xi_n^4. \quad (27)$$

Finally, based on (26) and (27), by Lemma 3, one immediately gets the conclusion. □

Now, we are in a position to get the main results of this paper.

Theorem 7. *If Assumptions 1 and 2 hold for the upper-triangular stochastic nonlinear systems* (1)*, with the coordinate transformation* (7)*, by appropriately choosing the parameter* $0 < \varepsilon < 1$*, then, under the state-feedback controller* (25)*, one has the following:*

(1) *the closed-loop system has an almost surely unique solution on* $[0, \infty)$*;*

(2) *the equilibrium of the closed-loop system is GAS in probability.*

Proof. For system (8), with the state-feedback controller (25) and Lyapunov function (27), one has

$$\mathscr{L}V_n\left(\overline{z}_n\right) \le -\varepsilon\sum_{i=1}^{n}\xi_i^4 + \sum_{i=1}^{n}\frac{\partial V_n}{\partial z_i}\overline{f}_i\left(\widetilde{z}_{i+2}\right) + \frac{1}{2}\operatorname{Tr}\left\{G\frac{\partial^2 V_n}{\partial z^2}G^T\right\}, \quad (28)$$

where $z = (z_1, \ldots, z_n)^T$, $G = (\overline{g}_1, \ldots, \overline{g}_{n-2}, 0, 0)$. From (15) and (27), one has

$$V_n\left(\overline{z}_n\right) = \frac{1}{4}\sum_{i=1}^{n}\xi_i^4 = \frac{1}{4}\sum_{i=1}^{n}\left(z_i + c_{i,i-1}z_{i-1} + \cdots + c_{i,1}z_1\right)^4, \quad (29)$$

where $c_{i,j}$, $j = 1, \ldots, i-1$, are constants. By (7), (15) and Assumption 1, one can get

$$\begin{aligned}\left|\overline{f}_i\left(\widetilde{z}_{i+2}\right)\right| &= \left|\frac{f_i\left(\widetilde{x}_{i+2}\right)}{\varepsilon^{i-1}}\right| \le b\varepsilon^2\sum_{j=i+2}^{n}\left|z_j\right| \\ &\le b\varepsilon^2\sum_{j=i+2}^{n}\left(\left|\xi_j\right| + \alpha_{j-1}\left|\xi_{j-1}\right|\right).\end{aligned} \quad (30)$$

By Young's inequality, using (29) and (30), one has

$$\sum_{i=1}^{n}\frac{\partial V_n}{\partial z_i}\overline{f}_i\left(\widetilde{z}_{i+2}\right) \le b_f\varepsilon^2\sum_{i=1}^{n}\xi_i^4. \quad (31)$$

Similarly, one can prove that

$$\frac{1}{2}\operatorname{Tr}\left\{G\frac{\partial^2 V_n}{\partial z^2}G^T\right\} \le b_g\varepsilon^2\sum_{i=1}^{n}\xi_i^4. \quad (32)$$

Substituting (31) and (32) into (28), one has

$$\begin{aligned}\mathscr{L}V_n\left(\overline{z}_n\right)|_{(8)} &\le -\varepsilon\sum_{i=1}^{n}\xi_i^4 + \left(b_f+b_g\right)\varepsilon^2\sum_{i=1}^{n}\xi_i^4 \\ &= -\varepsilon\left(1-\left(b_f+b_g\right)\varepsilon\right)\sum_{i=1}^{n}\xi_i^4.\end{aligned} \quad (33)$$

By choosing $0 < \varepsilon < 1$ appropriately, (33) can be written as

$$\mathscr{L}V_n\left(\overline{z}_n\right)\big|_{(8)} \le -c_0\sum_{i=1}^{n}\xi_i^4, \quad (34)$$

where $c_0 > 0$ is a constant.

By (34) and the coordinate transformation (7), using Lemma 3, the conclusions hold. □

Remark 8. Theorems 6 and 7 provide us a new perspective to deal with the state-feedback control problem for upper-triangular stochastic nonlinear systems with time-varying coefficients. The main technical obstacle in the Lyapunov design for stochastic upper-triangular systems is that Itô stochastic differentiation involves not only the gradient but also the higher order Hessian term. The traditional design methods are invalid to deal with these terms. However, with the design methodology provided in Theorems 6 and 7, there is no need to estimate the bounds of drift and diffusion terms step by step. Based on this technique, a homogeneous nonlinear controller for the nominal nonlinear system is firstly constructed. Then we will design a scaled controller which can effectively dominate the drift and diffusion terms by taking advantage of the homogenous structure of the controller.

4. A Simulation Example

Consider the following system:

$$\begin{aligned}dx_1 &= \left(\left(2-\sin^2 t\right)x_2 + x_3\sin x_3\right)dt + x_3\cos x_3 d\omega, \\ dx_2 &= \left(2-\cos t\right)x_3 dt, \\ dx_3 &= \left(1+\sin^2 t\right)u dt.\end{aligned} \quad (35)$$

Obviously, Assumptions 1 and 2 hold.

Introduce the following coordinate transformation:

$$z_1 = x_1, \qquad z_2 = \frac{x_2}{\varepsilon}, \qquad z_3 = \frac{x_3}{\varepsilon^2}, \qquad v = \frac{u}{\varepsilon^3}, \quad (36)$$

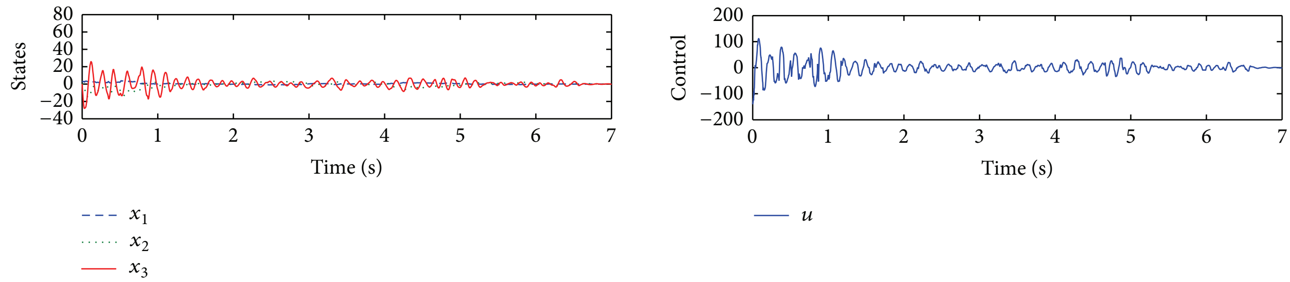


FIGURE 1: The response of closed-loop system (35)–(38).

where $0 < \varepsilon < 1$ is a design parameter. Then (35) can be written as

$$\begin{aligned} dz_1 &= \left(\varepsilon\left(2-\sin^2 t\right) z_2 + \varepsilon^2 z_3 \sin^2\left(\varepsilon^2 z_3\right)\right) dt \\ &\quad + \varepsilon^2 z_3 \cos\left(\varepsilon^2 z_3\right) d\omega, \\ dz_2 &= \varepsilon\left(2-\cos t\right) z_3 dt, \\ dz_3 &= \varepsilon\left(1+\sin^2 t\right) v dt. \end{aligned} \tag{37}$$

By following the design procedure in Section 3, one gets

$$v\left(z_1, z_2, z_3\right) = -1310\left(372 z_1 + 124 z_2 + z_3\right). \tag{38}$$

By choosing $\varepsilon = 0.001$, with the initial values $z_1(0) = 3$, $z_2(0) = 2$, and $z_3(0) = 5$, Figure 1 gives the system response of the closed-loop system consisting of (35)–(38), from which the efficiency of the tracking controller is demonstrated.

5. Concluding Remarks

For a class of upper-triangular stochastic nonlinear systems with time-varying control coefficients, this paper investigates the state-feedback stabilization problem. The designed controller can guarantee that the closed-loop system has a unique solution and the closed-loop system can be proved to be GAS in probability.

There are many related problems to be investigated, for example, how to generalize the result in this paper to more general stochastic upper-triangular nonlinear systems.

Conflict of Interests

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

Acknowledgments

This work is supported by National Natural Science Foundation (NNSF) of China under Grant nos. 61104128, 61174097; China Postdoctoral Science Foundation under Grant nos. 2012M520418, 2013T60185; Promotive Research Fund for Excellent Young and Middle-Aged Scientists of Shandong Province under Grant no. BS2013DX001; School Research Fund of Ludong University under Grant no. LY2012014.

References

[1] Z. Pan and T. Başar, "Adaptive controller design for tracking and disturbance attenuation in parametric strict-feedback nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 8, pp. 1066–1083, 1998.

[2] Z. Pan and T. Başar, "Backstepping controller design for nonlinear stochastic systems under a risk-sensitive cost criterion," *SIAM Journal on Control and Optimization*, vol. 37, no. 3, pp. 957–995, 1999.

[3] Z. Pan, K. Ezal, A. J. Krener, and P. V. Kokotović, "Backstepping design with local optimality matching," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1014–1027, 2001.

[4] H. Deng and M. Krstić, "Stochastic nonlinear stabilization. I. A backstepping design," *Systems & Control Letters*, vol. 32, no. 3, pp. 143–150, 1997.

[5] M. Krstić and H. Deng, *Stabilization of Uncertain Nonlinear Systems*, Springer, New York, NY, USA, 1998.

[6] H. Deng and M. Krstić, "Output-feedback stochastic nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 328–333, 1999.

[7] H. Deng and M. Krstić, "Output-feedback stabilization of stochastic nonlinear systems driven by noise of unknown covariance," *Systems & Control Letters*, vol. 39, no. 3, pp. 173–182, 2000.

[8] H. Deng, M. Krstić, and R. J. Williams, "Stabilization of stochastic nonlinear systems driven by noise of unknown covariance," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1237–1253, 2001.

[9] Z.-J. Wu, X.-J. Xie, and S.-Y. Zhang, "Adaptive backstepping controller design using stochastic small-gain theorem," *Automatica*, vol. 43, no. 4, pp. 608–620, 2007.

[10] Z.-J. Wu, X.-J. Xie, and S.-Y. Zhang, "Stochastic adaptive backstepping controller design by introducing dynamic signal and changing supply function," *International Journal of Control*, vol. 79, no. 12, pp. 1635–1646, 2006.

[11] X. Yu and X.-J. Xie, "Output feedback regulation of stochastic nonlinear systems with stochastic iISS inverse dynamics," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 304–320, 2010.

[12] X. Yu, X.-J. Xie, and N. Duan, "Small-gain control method for stochastic nonlinear systems with stochastic iISS inverse dynamics," *Automatica*, vol. 46, no. 11, pp. 1790–1798, 2010.

[13] W. Li, Y. Jing, and S. Zhang, "Adaptive state-feedback stabilization for a large class of high-order stochastic nonlinear systems," *Automatica*, vol. 47, no. 4, pp. 819–828, 2011.

[14] W. Li, X. Liu, and S. Zhang, "Further results on adaptive state-feedback stabilization for stochastic high-order nonlinear systems," *Automatica*, vol. 48, no. 8, pp. 1667–1675, 2012.

[15] X.-J. Xie and W.-Q. Li, "Output-feedback control of a class of high-order stochastic nonlinear systems," *International Journal of Control*, vol. 82, no. 9, pp. 1692–1705, 2009.

[16] W. Li, Y. Jing, and S. Zhang, "Output-feedback stabilization for stochastic nonlinear systems whose linearizations are not stabilizable," *Automatica*, vol. 46, no. 4, pp. 752–760, 2010.

[17] A. R. Teel, "Global stabilization and restricted tracking for multiple integrators with bounded controls," *Systems & Control Letters*, vol. 18, no. 3, pp. 165–171, 1992.

[18] F. Mazenc, "Stabilization of feedforward systems approximated by a non-linear chain of integrators," *Systems & Control Letters*, vol. 32, no. 4, pp. 223–229, 1997.

[19] F. Mazenc and L. Praly, "Adding integrations, saturated controls, and stabilization for feedforward systems," *IEEE Transactions on Automatic Control*, vol. 41, no. 11, pp. 1559–1578, 1996.

[20] X. Ye, "Universal stabilization of feedforward nonlinear systems," *Automatica*, vol. 39, no. 1, pp. 141–147, 2003.