TWO-TIME-SCALE SYSTEMS IN CONTINUOUS TIME WITH REGIME SWITCHING AND THEIR APPLICATIONS

by

YOUSEF TALAFHA

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2013

MAJOR: MATHEMATICS

Approved by:

Advisor Date

DEDICATION

To my mother

To soul my father

To soul my friend No'maan Alomari

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest sincere gratitude to my advisor, Professor Gang George Yin for giving me the chance to work on his interesting research. His energy and enthusiasm have been a constant source of motivation for me. It has a great honor to work with him and learn from him. I will be owe to him the most of the rest of my life.

I would also like to thank the other member of dissertation committee, Professors Pau-Liu Chow, Boris Mordukhovich, Tze-Chien Sun, and Wen Chen for serving on the committee.

Special acknowledgement go to all faculty members in the Department of Mathematics including Professors Daniel Frohardt, John Breckenridge, William Cohn, John Klein, Alexander Korostelev, Robert Bruner, Daniel Isakson, Choon-Jai Rhee, Ms. Patricia Bonesteel, Mary Klamo, and Tiana Fluker for helping me to grow as both a researcher and a teacher during my study at Wayne State University.

Special thanks are due to my colleagues and friends Zhixin Yang, Guangliang Zhao, Qi He, Quan Yuan, Yu Sun, Ren Zhao, Tamer Badawy, Jiuyi Zhu, and Li Fan for their conversations, the enjoyable discussions and Ph.D's politics. They made the graduate study life a great experience. I would like to thanks my best friend Dr. Musab Al-Hadrusi for helping me in the first year in both life and academic.

Last but not least, I like express my thanks to my mother and brothers for their infinite support and encouragement.

Yousef Talafha

July. 12th 2013, Detroit, Michigan

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

This dissertation is denoted to two-time-scale stochastic systems. The systems under consideration are modeled by stochastic differential equations with regime switching. Formulations using regime switching have appeared in a wide variety of situations including manufacturing systems, communication networks, financial engineering, ecology and biology modeling, multi-agent control systems etc. One of the main characteristics is the coexistence of continuous dynamics and discrete events. The continuous dynamics can be formulated by use of stochastic differential equations, whereas the discrete events have very different features. The focus in this work is placed on such systems with two-time scales. We consider the situation that the discrete events have a large state space with large number of elements. The large state space can however be partitioned into several subspaces. States within each subspace move rapidly, whereas the transitions from one subspace to another take place in a relatively slow pace. Using the idea of decomposition and aggregation, we lump all the states within each subspace into a "super" state. Then the total number of states in the newly aggregated state space is drastically reduced. The rationale is that using such a formulation, we can substantially reduce the computational complexity. In the early 1960s, the so-called nearly completely decomposable system models came into being. In the late 1990s, much work was done for two-time-scale Markov chains. Expanding on these ideas, this dissertation examines systems with switching diffusions with two-time scales.

The first part of the work focuses on studying Liénard equations with regime swashing. The well-known Liénard equations was named after the French physicist Alfred-Marie Liénard. Such equations have been studied extensively in the literature of dynamic systems and ordinary differential equations (ODEs). The standard Liénard equation has the form

$$\ddot{\eta}(t) + f(\eta(t))\dot{\eta}(t) + g(\eta(t)) = 0,$$

which may be written as a first-order system of equations. During the development of radio and vacuum tubes, the Liénard equations were used to model oscillating circuits. These equations have also been used to describe certain mechanical systems in physics and engineering. Parallel to the development of deterministic systems, there is a large amount of work on Liénard equations perturbed by white noise.

Recently, much interests are devoted to Liénard systems subject to both white noise perturbation and random environment influence in which an additional random switching process is added [35, 40]. Such models belong to a class of Markov processes involving both continuous states and discrete events. The discrete events are used to model random environment and other random factors that cannot be described by differential equations. In [34], we studied randomly-switching Liénard equations in which the switching process depends on the continuous states. A number of results including existence and uniqueness of solution of the underlying equations, regularity, and ergodicity were obtained. Moreover, in Chapter 3, we shall discuss the weak convergence limit of Liénard equations. The model of interest can defined as follows:

$$
\begin{aligned}
dX_1^\varepsilon(t) &= X_2^\varepsilon(t)dt, \\
dX_2^\varepsilon(t) &= -\big(X_2^\varepsilon(t)f(X_1^\varepsilon(t),\alpha^\varepsilon(t),t)+g(X_1^\varepsilon(t),\alpha^\varepsilon(t),t)\big)dt+h(X_1^\varepsilon(t),\alpha^\varepsilon(t),t)dw(t), \\
X^\varepsilon(0) &= x \ \text{ and } \ \alpha^\varepsilon(0)=\iota,
\end{aligned}
$$

where $w(t)$ is a standard one-dimensional Brownian motion. The random process $\alpha^\varepsilon(\cdot)$ is a continuous-time Markov chain with state space $\mathcal{M}$ and it is independent of the Brownian motion $w(\cdot)$.

The second part of the work is on near-optimal control of two-time-scale switching diffusions. The class of singularly perturbed Markov chain has been studied extensively in recent years. The notation of relaxed control is introduced for stochastic system in Fleming [9] and for deterministic optimal control problem in Warga [33]. The motivation for the first study is devoted to singulary perturbed controlled diffusion. The singularly perturbed problems of control has been considered in both stochastic and deterministic literature in [2, 15, 16, 18, 17, 22]. In [2] is devoted to studying the role of perturbations in problems of the optimal control of differential equations. Kokotovic and Khalil make a collection of reprint of 61 articles including with singular perturbation method to system analysis and control. The book by Kokotovic, Khalil and OReilly [17] is a good source for example make the book useful for students of undergraduate and graduate levels. In [22] weak convergence method and singularity perturbed were used to prove some important results in stochastic control.

We consider the following stochastic differential equation with regime switching

$$x^{\varepsilon}(t) = x_0 + \int_0^t b(x^{\varepsilon}(s), \alpha^{\varepsilon}(s), u^{\varepsilon}(s))ds + \int_0^t \sigma(\alpha^{\varepsilon}(s), x^{\varepsilon}(s))dw(s)$$

$$\alpha^{\varepsilon}(0) = i_0 \in \mathcal{M}$$

where $b(\cdot,\cdot,\cdot) : \mathbb{R}^k \times \mathcal{M} \times U \to \mathbb{R}^k$, $\sigma(\cdot) : \mathbb{R}^k \times \mathcal{M} \to \mathbb{R}^{2k}$ are given functions. Suppose that $w(\cdot)$ be a standard one-dimensional Brownian motion and $\alpha^{\varepsilon}(\cdot)$ is the continuous Markov chain with state space $\mathcal{M}$ such that the random process $\alpha^{\varepsilon}(\cdot)$ is independent of the Brownian motion $w(\cdot)$. The cost function has the form

$$J^{\varepsilon}(u^{\varepsilon}(\cdot)) = J^{\varepsilon}(x_0, i_0, u^{\varepsilon}(\cdot)) = E_{x_0,i_0} \int_0^T C(x^{\varepsilon}(s), \alpha^{\varepsilon}(s), u^{\varepsilon}(s))ds$$

where E_{x_0,i_0} denotes the expectation taken with $x^{\varepsilon}(0) = x_0$ and $\alpha^{\varepsilon}(0) = i_0$.

The next chapter is devoted to Van der Pol oscillator. The Van der Pol is one of an example of nonlinear oscillators. This model was investigated by Van der Pol (1889-1959) in 1927 [31] while he was an engineer at Philps company. This model has being extensively studied in [3, 4, 7, 25, 29, 30, 32]. The Van der Pol oscillator is used a variety of mechanical, electrical, physics and biological sciences. For example, Van der Pol and Van der Mark [30, 32] describes the hearts behavior. Fitzhagh [7] and Nagumo et. al. [25] describe the action potentials of neurons. Further application of Van der Pol oscillator we can found in [4, 29]. The standard form of the Van der Pol oscillator can defined as follows

$$\ddot{y} - \mu(1 - y^2)\dot{y} + y = 0$$

where x is the dynamical variable and $\mu > 0$ a parameter. In this chapter we investigate the behavior of systems of Van der Pol oscillator by introducing the noise. Simulation results are presented.

Finally, we conclude the dissertation by providing some further remarks. In addition, some preliminary results are recalled as well.

CHAPTER 2

Stochastic Liénard Equations with Random Switching and Two-time Scales

2.1 Introduction

This chapter is concerned with randomly switching Liénard equations. Our main concern is that the state space of the discrete events or the switching process is large. We focus on reducing the computational complexity. Treating large-scale systems, Simon and Ando noted in [27] that in a large-scale system, not all states change at the same rates. Some of them vary rapidly and others change slowly. In [5], the idea of decomposition and aggregation was brought in aiming at reducing the computational complexity. In this chapter, we are dealing with the models in which the state space of the discrete events is rather large. We use the idea of decomposition and aggregation to treat the problems under consideration. Denote the switching process by $\alpha(t)$. To highlight the different rates of actions or transition frequencies, we introduce a small parameter $\varepsilon > 0$ into the system, and write $\alpha(t) = \alpha^{\varepsilon}(t)$. So the system under consideration becomes one with two-time scales. Based on time-scale separation, our effort is then to obtain asymptotic properties of the underlying system. Specifically, we assume the state space of the discrete events can be decomposed into l weakly connected subspaces. Within each subspace, the transitions of the discrete events take place in a fast pace. Among different subspaces, the transition occurs relatively infrequently. We then aggregate the discrete states in each subspace into one super state and show that such aggregations lead to a limit system that can be represented as the average of the original system with respect to the quasi-invariance measure of the switching process. We analyze the limit properties by means of martingale problem formulation,

The idea of averaging has played an important role in numerous stochastic systems. For some of the developments, we mention the work [12, 13] for two-time-scale diffusions, [14] for singularly perturbed Markov chains, [1] for a class of problems involving switching processes with the use of diffusion approximation approach, [19, 20] for diffusion approximation in evolutionary systems using semi-Markov processes, and [26] for two-time-scale manufacturing systems and hierarchical decision making.

The rest of the chapter is arranged as follows. Section 2.2 presents the precise formulation of the problem. Section 2.3 concerns the limit properties of the system. Section 2.4 makes further remarks and concludes the chapter.

2.2 Formulation

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Consider the usual Lénard equation with an additional white noise perturbation. Moreover, assume the coefficients of the equation all depend on a switching process. Use z' to denote the transpose of $z \in \mathbb{R}^{l_1 \times l_2}$ for $l_1, l_2 \geq 1$. Setting $X(t) = (X_1(t), X_2(t))' = (\eta(t), \dot{\eta}(t))'$, to reflect the ε-dependence, we write $X(t)$ as

$X^\varepsilon(t)$ in what follows. We present the model of interest as follows:

$$
\begin{aligned}
dX_1^\varepsilon(t) &= X_2^\varepsilon(t)dt,\\
dX_2^\varepsilon(t) &= -\big(X_2^\varepsilon(t)f(X_1^\varepsilon(t),\alpha^\varepsilon(t),t)+g(X_1^\varepsilon(t),\alpha^\varepsilon(t),t)\big)dt+h(X_1^\varepsilon(t),\alpha^\varepsilon(t),t)dw(t),\\
X^\varepsilon(0) &= x \ \text{ and } \ \alpha^\varepsilon(0)=\iota,
\end{aligned}
\tag{2.1}
$$

where $w(t)$ is a standard one-dimensional Brownian motion. Often, one considers Liénard equations that are autonomous. Here, for more generality, we assume t depend in the dynamics. So the systems become time-varying. The random process $\alpha^\varepsilon(\cdot)$ is a continuous-time Markov chain with state space $\mathcal{M}$ and it is independent of the Brownian motion $w(\cdot)$. In this chapter, we consider the case that the state space $\mathcal{M}$ is large in that $|\mathcal{M}|$, the cardinality of $\mathcal{M}$ is a large number.

(A1) For each $\iota\in\mathcal{M}$ and each $t\in[0,T]$, the functions $f(x_1,\iota,t)$, $g(x_1,\iota,t)$, and $h(x_1,\iota,t)$ in (2.1) satisfy the local Lipschitz condition with respect to x_1. These functions are continuously differentiable with respect to t for each x_1 and $\iota\in\mathcal{M}$.

(A2) There exists a positive constant $K_0>0$ such that for each $\iota\in\mathcal{M}$,

$$
\begin{aligned}
&\inf\{f(x_1,\iota,t): x_1\in\mathbb{R}\}\geq -K_0 \ \text{ uniformly in } \ t,\\
&\int_0^{x_1} g(u,\iota,t)du\to\infty \quad \text{as} \quad |x_1|\to\infty \ \text{ uniformly in } \ t,
\end{aligned}
\tag{2.2}
$$

and that for each $\iota\in\mathcal{M}$, $h(x_1,\iota,t)$ is infinitely differentiable w.r.t. x_1 satisfying $0< h(x_1,\iota,t)\leq K_0$ for all $x_1\in\mathbb{R}$ uniformly in t.

Under the conditions above, we obtain the existence and uniqueness of the solution of (2.1). We state it as a lemma below and refer the reader to [34] for a detailed proof.

Lemma 2.1. *Under conditions* (A1) *and* (A2), (2.1) *has a unique strong solution for each initial condition* $(X^\varepsilon(0), \alpha^\varepsilon(0)) = (x, \iota)$.

Next, we assume that the generator of the switching process is a time-inhomogeneous Markov chain with the generator given by

$$Q^\varepsilon(t) = \frac{1}{\varepsilon}\widetilde{Q}(t) + \widehat{Q}(t), \tag{2.3}$$

such that

$$\widetilde{Q}(t) = \mathrm{diag}(\widetilde{Q}^1(t), \ldots, \widetilde{Q}^l(t)), \tag{2.4}$$

where $\mathrm{diag}(A^1, \ldots, A^k)$ denotes a block diagonal matrix with entries $A^1, \ldots, A^k$ of proper dimensions, that each $\widetilde{Q}^i(t)$ is a generator of suitable dimension and $\widehat{Q}(t)$ is another generator of a continuous-time Markov chain. The rationale is that we decompose the state space into weakly connected subspaces, $\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_l$ where $\mathcal{M}_i = \{s_{i1}, \ldots, s_{im_i}\}$. From now on, we often use a double index s_{ij} to denote a state in $\mathcal{M}$. To proceed, we need another condition.

(A3) For each $i \in \overline{\mathcal{M}} = \{1, \ldots, l\}$, $\widetilde{Q}^i(t) \in \mathbb{R}^{m_i \times m_i}$ is weakly irreducible in the sense that the system of equations

$$\nu^i(t)\widetilde{Q}^i(t) = 0, \quad \mathbb{1}_{m_i}\nu^i(t) = 1$$

has a unique solution, where $\mathbb{1}_{m_i} = (1, \ldots, 1)' \in \mathbb{R}^{m_i \times 1}$ and $\nu^i(t)$ is termed a quasi-stationary distribution satisfying $\nu^i(t) = (\nu^i_1(t), \ldots, \nu^i_{m_i}(t)) \in \mathbb{R}^{1 \times m_i}$ with $\nu^i_j(t) \geq 0$ for $j = 1, \ldots, m_i$. Both $\widetilde{Q}(\cdot)$ and $\widehat{Q}(\cdot)$ are bounded and measurable.

The motivation of the model is that although the state space $\mathcal{M}$ is not completely decomposable into l subspaces, the actions or transitions among different subspaces are weak. Define $\overline{\alpha}^{\varepsilon}(t) = i$ when $\alpha^{\varepsilon}(t) \in \mathcal{M}_i$. Note that $\overline{\alpha}^{\varepsilon}(\cdot)$ is not a Markov chain. Nevertheless, using (A3), we obtain the following weak limit of $\overline{\alpha}^{\varepsilon}(\cdot)$ by virtue of [37, Lemma 7.4].

Lemma 2.2. *The process $\overline{\alpha}^{\varepsilon}(\cdot)$ converges weakly to $\overline{\alpha}(\cdot)$ that is generated by*

$$\overline{Q}(t) = \mathrm{diag}(\nu^1(t), \dots, \nu^l(t))\widehat{Q}(t)\widetilde{\mathbb{1}}, \tag{2.5}$$

where $\widetilde{\mathbb{1}} = \mathrm{diag}(\mathbb{1}_{m_1}, \dots, \mathbb{1}_{m_l})$.

Note that in view of our partition, we may write the associated operator corresponding to the switching diffusion defined in (2.1) as: For each $i = 1, \dots, l$ and $j = 1, \dots, m_i$, $\ell = s_{ij}$, and suitable smooth function $V(\cdot, \ell, \cdot) \in C^{2,1}(\mathbb{R}^2 \times [0, T]; \mathbb{R})$, and $Q^{\varepsilon}(t) = (q^{\varepsilon}_{\ell\iota}(t))$,

$$\begin{aligned}\mathcal{L}^{\varepsilon}V(x, \ell, t) = \frac{\partial V(x, \ell, t)}{\partial t} + \frac{1}{2}h^2(x_1, \ell, t)\frac{\partial^2}{\partial x_2^2}V(x, \ell, t) + x_2\frac{\partial}{\partial x_1}V(x, \ell, t)\\ -\big[x_2 f(x_1, \ell, t) + g(x_1, \ell, t)\big]\frac{\partial}{\partial x_2}V(x, \ell, t) + \sum_{\iota\in\mathcal{M}} q^{\varepsilon}_{\ell\iota}(t)\big[V(x, \iota, t) - V(x, \ell, t)\big].\end{aligned} \tag{2.6}$$

2.3 Asymptotic Properties

Consider the pair of processes $Y^{\varepsilon}(\cdot) = (X^{\varepsilon}(\cdot), \overline{\alpha}^{\varepsilon}(\cdot))$. We aim to show that $Y^{\varepsilon}(\cdot)$ converges weakly to $Y(\cdot) = (X(\cdot), \overline{\alpha}(\cdot))$ such that $X(\cdot)$ is the solution of the switching Liénard equation

$$\begin{aligned}dX_1(t) &= X_2(t)dt\\ dX_2(t) &= -[X_2(t)\overline{f}(X_1(t), \overline{\alpha}(t), t) + \overline{g}(X_1(t), \overline{\alpha}(t), t)]dt + \overline{h}(X_1(t), \overline{\alpha}(t), t)dw(t),\end{aligned} \tag{2.7}$$

and that $\overline{\alpha}(\cdot)$ is a continuous-time Markov chain given by Lemma 2.2. where $\overline{h}(x_1,i,t)$ is the square root of $\overline{h}^2(x_1,i,t)$ with

$$\begin{aligned}
\overline{f}(x_1,i,t) &= \sum_{j=1}^{m_i} \nu_j^i(t) f(x_1,j,t),\\
\overline{g}(x_1,i,t) &= \sum_{j=1}^{m_i} \nu_j^i(t) g(x_1,j,t),\\
\overline{h}^2(x_1,i,t) &= \sum_{j=1}^{m_i} \nu_j^i(t) h^2(x_1,j,t).
\end{aligned} \tag{2.8}$$

Associated with the limit process, for each $i \in \overline{\mathcal{M}}$ and a suitably smooth function $V(\cdot,i,\cdot) \in C^{2,1}(\mathbb{R}^2 \times [0,T];\mathbb{R})$, define an operator $\overline{\mathcal{L}}$ as follows:

$$\begin{aligned}
\overline{\mathcal{L}}V(x,i,t) = \frac{\partial V(x,i,t)}{\partial t} &+ \frac{1}{2}\overline{h}^2(x_1,i,t)\frac{\partial^2}{\partial x_2^2}V(x,i,t) + x_2\frac{\partial}{\partial x_1}V(x,i,t)\\
&-\big[x_2\overline{f}(x_1,i,t) + \overline{g}(x_1,i,t)\big]\frac{\partial}{\partial x_2}V(x,i,t)\\
&+\sum_{j\in\overline{\mathcal{M}}} \overline{q}_{ij}(t)\big[V(x,j,t) - V(x,i,t)\big],
\end{aligned} \tag{2.9}$$

where $\overline{Q}(t)$ is given by (2.5).

Theorem 2.3. *Assume* (A1)–(A3)*, and suppose that* (2.7) *has a unique solution in the sense of in distribution for each initial condition. Then $Y^\varepsilon(\cdot)$ converges weakly to $Y(\cdot) = (X(\cdot),\overline{\alpha}(\cdot))$ such that $X(\cdot)$ is a solution of* (2.7) *and $\overline{\alpha}(\cdot)$ is a continuous-time Markov chain generated by $\overline{Q}(t)$ given in* (2.5).

Remark 2.4. An equivalent way of stating Theorem 2.3 is: $(X(\cdot),\overline{\alpha}(\cdot))$ is a solution of the martingale problem with operator $\overline{\mathcal{L}}$.

Proof. The proof of the theorem will be divided into several steps. These steps are realized by presenting a number of lemmas. We prove the weak convergence using a martingale problem formulation.

Step 1: Uniqueness of the martingale problem. The uniqueness in the sense of in distribution of the limit stochastic differential equation with switching (2.7) implies that the martingale problem with operator $\overline{\mathcal{L}}$ has a unique solution in the sense in distribution.

Step 2: A truncated process. To continue with the proof of weak convergence, we use a truncation methods; see [24, p. 284] for details. The idea is that for each $0 < N < \infty$, we work with $X^{\varepsilon,N}(\cdot)$ that is equal to $X^{\varepsilon}(\cdot)$ up until the first exit from $S_N = \{x \in \mathbb{R}^2 : |x| \leq N\}$, the ball with radius N. Such a process $X^{\varepsilon,N}(\cdot)$ is known as an N-truncation of $X^{\varepsilon}(\cdot)$. Define a truncation function as a smooth function such that

$$T_N(x) = \begin{cases} 1, \ x \in S_N, \\ \\ 0, \ x \in \mathbb{R}^2 - S_{N+1}. \end{cases}$$

Note that the truncation is such that it equals 1 when x is in the ball with radius N and is 0 outside the ball of radius $N+1$ and is smoothly connected. Consider

$$\begin{aligned} dX_1^N(t) &= X_2^N(t)dt \\ dX_2^N(t) &= -[X_2^N(t)\overline{f}^N(X_1^N(t), \overline{\alpha}(t), t) + \overline{g}^N(X_1^N(t), \overline{\alpha}(t), t)]dt \\ &\quad +\overline{h}^N(X_1^N(t), \overline{\alpha}(t), t)dw(t), \end{aligned} \tag{2.10}$$

where

$$\zeta^N(x,i,t) = \zeta(x,i,t)T_N(x) \text{ for } \zeta^N(\cdot) = f(\cdot), \text{ or } g(\cdot), \text{ or } h(\cdot).$$

Associated with the truncated process, we define a truncated operator as follows. For each $i \in \overline{\mathcal{M}}$ and a suitably smooth function $V(\cdot, i, \cdot) \in C^{2,1}(\mathbb{R}^2 \times [0,T]; \mathbb{R})$, define an operator $\overline{\mathcal{L}}^N$ as follows:

$$\begin{aligned}\overline{\mathcal{L}}^N V(x,i,t) = \frac{\partial V(x,i,t)}{\partial t} &+ \frac{1}{2}\overline{h}^{N,2}(x_1,i,t)\frac{\partial^2}{\partial x_2^2}V(x,i,t) + x_2\frac{\partial}{\partial x_1}V(x,i,t) \\ &-\big[x_2\overline{f}^N(x_1,i,t) + \overline{g}^N(x_1,i,t)\big]\frac{\partial}{\partial x_2}V(x,i,t) \\ &+\sum_{j\in\overline{\mathcal{M}}}\overline{q}_{ij}(t)\big[V(x,j,t) - V(x,i,t)\big],\end{aligned} \tag{2.11}$$

where we used the notation $h^{N,2}(x_1,\iota,t) = (h^N(x_1,\iota,t))^2$.

Proposition 2.5. *Under the conditions of Theorem* 2.3*,* $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ *converges weakly to* $(X^N(\cdot), \overline{\alpha}(\cdot))$ *such that* $X^N(\cdot)$ *is the solution of the truncated Liénard equation* (2.10) *and* $\overline{\alpha}(\cdot)$ *is the Markov chain generated by* $\overline{Q}(t)$ *given in* (2.5)*.*

The proof of Proposition 2.5 is again divided into several steps. We proceed to carry out the steps in what follows.

Step 2.1: Tightness of $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$. This is proved by Lemma 2.6. Denote by $D([0,T] : \mathbb{R}^2 \times \overline{\mathcal{M}})$ the space of functions defined in $[0,T]$ with values in $\mathbb{R}^2 \times \overline{\mathcal{M}}$ that are right continuous and have left limits endowed with the Skorohod topology (see [24, p. 228]).

Lemma 2.6. *Under the conditions of Proposition* 2.5*,* $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ *is tight in* $D([0,T] : \mathbb{R}^2 \times \overline{\mathcal{M}})$*.*

Proof of Lemma 2.6. For any $\delta > 0$, $t > 0$, and $s > 0$ satisfying $0 < s \leq \delta$, we have

$$\begin{aligned}
&E_t^\varepsilon \left| X^{\varepsilon,N}(t+s) - X^{\varepsilon,N}(t) \right|^2 \\
&\quad \leq K E_t^\varepsilon \left| \int_t^{t+s} X_2^{\varepsilon,N}(u) du \right|^2 \\
&\qquad + K E_t^\varepsilon \left| \int_t^{t+s} [-X_2^{\varepsilon,N}(u) f^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) + g^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u)] du \right|^2 \\
&\qquad + K E_t^\varepsilon \left| \int_t^{t+s} h^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) dw(u) \right|^2 .
\end{aligned} \tag{2.12}$$

By virtue of the familiar Hölder inequality and the boundedness $f^N(\cdot)$ and $g^N(\cdot)$ together with the choice of $\delta > 0$, it is readily seen that

$$\begin{aligned}
&E_t^\varepsilon \left| \int_t^{t+s} [-X_2^{\varepsilon,N}(u) f^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) + g^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u)] du \right|^2 \\
&\quad \leq K s^2 \leq K \delta^2 \leq K \delta.
\end{aligned} \tag{2.13}$$

Likewise, we obtain

$$E_t^\varepsilon \left| \int_t^{t+s} X_2^{\varepsilon,N}(u) du \right|^2 \leq K \delta. \tag{2.14}$$

where E_t^ε denotes the conditional expectation with the σ-algebra of $\sigma\{w(s), \alpha^\varepsilon(s); s \leq t\}$.

By using the properties of stochastic integrals, we obtain

$$
\begin{aligned}
&E_t^\varepsilon \left| \int_t^{t+s} h^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) dw(u) \right|^2 \\
&\le E \int_t^{t+s} |h^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u)|^2 du \\
&\le K\delta.
\end{aligned}
\tag{2.15}
$$

Combining (2.13)–(2.15) and applying these to (2.12), taking $\limsup_\varepsilon$ followed by $\lim_\delta$, we obtain

$$\lim_{\delta\to 0} \limsup_{\varepsilon\to 0} E_t^\varepsilon \{ \sup_{0\le s\le\delta} E_t^\varepsilon \left| X^{\varepsilon,N}(t+s) - X^{\varepsilon,N}(t) \right|^2 \} = 0,$$

where E_t^ε denotes the conditional expectation w.r.t. the σ-algebra generated by $\{w(u), \alpha^\varepsilon(u) : u \le t\}$. Thus the tightness of $\{X^{\varepsilon,N}(\cdot)\}$ follows. Furthermore, this tightness together with the tightness of $\{\overline{\alpha}^\varepsilon(\cdot)\}$ implies that of $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot)\}$. The lemma is proved.

Step 2.2: Characterization of the limit. Since $\{(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))\}$ is tight, we can extract convergent subsequences by Prophorov's theorem. Select such a sequence and for notational simplicity, still denote the sequence by $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$. Denote the limit by $(X^N(\cdot), \overline{\alpha}(\cdot))$. By Skorohod's representation, with a slight abuse of notation, we assume $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ converges to $(X^N(\cdot), \overline{\alpha}(\cdot))$ in the sense of w.p.1. The convergence is uniform in any bounded interval. We proceed to characterize the limit process. This is done by showing that the limit is a solution of the martingale problem with operator $\overline{\mathcal{L}}^N$.

First, we note that the following holds. By the weak convergence of $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ to $(X^N(\cdot), \overline{\alpha}(\cdot))$ and the Skorohod representation, for any bounded and continuous function $\rho(\cdot,\cdot) : \mathbb{R}^2 \times \mathcal{M} \mapsto \mathbb{R}$, any positive integer κ, any t and $s > 0$ with $t + s \le T$, and any $t_i \le t$

for any $i \le \kappa$, we have

$$\lim_{\varepsilon\to 0} E\rho(X^{\varepsilon,N}(t_i),\overline{\alpha}^\varepsilon(t_i) : i \le \kappa)[V(X^{\varepsilon,N}(t+s),\overline{\alpha}^\varepsilon(t+s),t+s) - V(X^{\varepsilon,N}(t),\overline{\alpha}^\varepsilon(t),t)]$$

$$= E\rho(X^N(t_i),\overline{\alpha}(t_i) : i \le \kappa)[V(X^N(t+s),\overline{\alpha}(t+s),t+s) - V(X^N(t),\overline{\alpha}(t),t)]. \tag{2.16}$$

To proceed, for each $\iota \in \mathcal{M}$, let $V(\cdot,\iota,\cdot) \in C_0^{2,1}(\mathbb{R}^2 \times [0,T] : \mathbb{R})$ (that is, $C^{2,1}$ functions with compact support). Define

$$\widehat{V}(x,\alpha,t) = \sum_{i=1}^{l} V(x,i,t) I_{\{\alpha\in\mathcal{M}_i\}}. \tag{2.17}$$

It is readily seen that $V(X^{\varepsilon,N}(t),\overline{\alpha}^\varepsilon(t),t) = \widehat{V}(X^{\varepsilon,N}(t),\alpha^\varepsilon(t),t)$. That is, to work with $\overline{\alpha}^\varepsilon(\cdot)$ is equivalent to work with $\alpha^\varepsilon(\cdot)$ using the structure of the function $\widehat{V}(\cdot)$. For any $(1/2) < \Delta < 1$, subdivide $[t,t+s]$ into subintervals of length $\varepsilon^{1-\Delta}$ by choosing $M_\varepsilon = \lfloor s/\varepsilon^{1-\Delta}\rfloor$ (the integer part of $s/\varepsilon^{1-\Delta}$) and $s_k = k\varepsilon^{1-\Delta}$ such that $t = s_0 \le s_1 \le \cdots \le s_{M_\varepsilon} = t+s$. Denote

$$L^\varepsilon(t,s) := E\rho(X^{\varepsilon,N}(t_i),\overline{\alpha}^\varepsilon(t_i) : i \le \kappa)$$

$$\times[V(X^{\varepsilon,N}(t+s),\overline{\alpha}^\varepsilon(t+s),t+s) - V(X^{\varepsilon,N}(t),\overline{\alpha}^\varepsilon(t),t)].$$

Then we have

$$L^\varepsilon(t,s) = E\rho(X^{\varepsilon,N}(t_i),\overline{\alpha}^\varepsilon(t_i) : i \le \kappa)$$

$$\times[\widehat{V}(X^{\varepsilon,N}(t+s),\alpha^\varepsilon(t+s),t+s) - \widehat{V}(X^{\varepsilon,N}(t),\alpha^\varepsilon(t),t)].$$

It then follows that

$$
\begin{aligned}
L^{\varepsilon}(t,s) = E\rho(X^{\varepsilon,N}(t_i),\overline{\alpha}^{\varepsilon}(t_i): i \le \kappa) \sum_{k=0}^{M_\varepsilon - 1} \Bigg\{ & \int_{s_k}^{s_{k+1}} \frac{\partial}{\partial u} \widehat{V}(X^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u)du \\
+ \int_{s_k}^{s_{k+1}} \Bigg[& \frac{1}{2} h^{N,2}(X_1^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u) \frac{\partial^2}{\partial x_2^2} \widehat{V}(X^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u) \\
& + X_2^{\varepsilon,N}(u) \frac{\partial}{\partial x_1} \widehat{V}(X^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u) \\
& - [X_2^{\varepsilon,N}(u) f^N(X_1^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u) \\
& \quad + g^N(X_1^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u)] \frac{\partial}{\partial x_2} \widehat{V}(X^{\varepsilon,N}(u),\alpha^{\varepsilon}(u),u) \\
& + Q^{\varepsilon}(u) \widehat{V}(X^{\varepsilon,N}(u),\cdot,u)(\alpha^{\varepsilon}(u)) \Bigg] \Bigg\} du.
\end{aligned}
\tag{2.18}
$$

Step 2.2.1: Piecewise constant approximation. Here, we estimate the difference of $X^{\varepsilon,N}(u)$ and $X^{\varepsilon,N}(s_k)$ for any $u \in [s_k, s_{k+1})$. Using (2.1) with the $X^{\varepsilon}(\cdot)$ replaced by $X^{\varepsilon,N}(\cdot)$, the Hölder inequality, the boundedness of $X^{\varepsilon,N}(\cdot)$, $f^N(\cdot)$, $g^N(\cdot)$, and $h^N(\cdot)$, and the well-known properties of stochastic integrals, we obtain that for any $u \in [s_k, s_{k+1})$,

$$\begin{aligned}
&E\left|X^{\varepsilon,N}(u) - X^{\varepsilon,N}(s_k)\right|^2 \\
&\quad \leq KE\left|\int_{s_k}^{u} \begin{pmatrix} X_2^{\varepsilon,N}(r) \\ \\ -(X_2^{\varepsilon,N}(r) f^N(X_1^{\varepsilon,N}(r), \alpha^\varepsilon(r), r) + g^N(X_1^{\varepsilon,N}(r), \alpha^\varepsilon(r), r) \end{pmatrix} dr\right|^2 \\
&\qquad +KE\left|\int_{s_k}^{u} \begin{pmatrix} 0 \\ \\ h(X_1^{\varepsilon,N}(r), \alpha^\varepsilon(t), t) dw(r) \end{pmatrix}\right|^2 \qquad (2.19) \\
&\quad \leq O((u - s_k)^2) + K\int_{s_k}^{u} E|h^N(X^{\varepsilon,N}(r), \alpha^\varepsilon(r), r)|^2 dr \\
&\quad \leq O(u - s_k) \leq K\varepsilon^{1-\Delta}.
\end{aligned}$$

Using the above obtained bounds, we further deduce

$$\lim_{\varepsilon\to 0} L^\varepsilon(t,s) = \lim_{\varepsilon\to 0} E\rho(X^{\varepsilon,N}(t_i),\overline{\alpha}^\varepsilon(t_i) : i \le \kappa) \sum_{k=0}^{M_\varepsilon - 1} \Bigg[\int_{s_k}^{s_{k+1}} \frac{\partial}{\partial u}\widehat{V}(X^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)du$$

$$+ \int_{s_k}^{s_{k+1}} \Big[\frac{1}{2} h^{N,2}(X_1^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)\frac{\partial^2}{\partial x_2^2}\widehat{V}(X^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)$$

$$+X_2^{\varepsilon,N}(u)\frac{\partial}{\partial x_1}\widehat{V}(X^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)$$

$$-\big[X_2^{\varepsilon,N}(s_k) f^N(X_1^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)$$

$$+g^N(X_1^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)\big]\frac{\partial}{\partial x_2}\widehat{V}(X^{\varepsilon,N}(s_k),\alpha^\varepsilon(u),u)$$

$$+Q^\varepsilon(u)\widehat{V}(X^{\varepsilon,N}(s_k),\cdot,u)(\alpha^\varepsilon(u))\Big]\Bigg]du. \tag{2.20}$$

Step 2.2.2: Further approximation. We begin with the last term in (2.18). The structures of the $\widehat{V}(\cdot)$ and $\widetilde{Q}(t)$ imply that for $u \in [s_k, s_{k+1})$,

$$
\begin{aligned}
& Q^\varepsilon(u)\widehat{V}(X^{\varepsilon,N}(s_k), \alpha^\varepsilon(u), u) \\
& \quad = \widehat{Q}(u)\widehat{V}(X^{\varepsilon,N}(s_k), \alpha^\varepsilon(u), u) \\
& \quad = \sum_{i=1}^{l}\sum_{j=1}^{m_i} \widehat{Q}(u)\widehat{V}(X^{\varepsilon,N}(s_k), s_{ij}, u) I_{\{\alpha^\varepsilon(u)=s_{ij}\}} \\
& \quad = \sum_{i=1}^{l}\sum_{j=1}^{m_i} \widehat{Q}(u)\widehat{V}(X^{\varepsilon,N}(s_k), s_{ij}, u)\nu^i_j(u) I_{\{\overline{\alpha}^\varepsilon(u)=i\}} \\
& \qquad + \sum_{i=1}^{l}\sum_{j=1}^{m_i} \widehat{Q}(u)\widehat{V}(X^{\varepsilon,N}(s_k), s_{ij}, u)[I_{\{\alpha^\varepsilon(u)=s_{ij}\}} - \nu^i_j(u) I_{\{\overline{\alpha}^\varepsilon(u)=i\}}].
\end{aligned}
\tag{2.21}
$$

By virtue of [37, Theorem 7.2],

$$
\begin{aligned}
& E\rho(X^{\varepsilon,N}(t_i), \overline{\alpha}^\varepsilon(t_i) : i \le \kappa) \sum_{k=0}^{M_\varepsilon - 1} \Bigg[\sum_{i=1}^{l}\sum_{j=1}^{m_i} \int_{s_k}^{s_{k+1}} \widehat{Q}(u)\widehat{V}(X^{\varepsilon,N}(s_k), s_{ij}, u)[I_{\{\alpha^\varepsilon(u)=s_{ij}\}} \\
& \qquad\qquad - \nu^i_j(u) I_{\{\overline{\alpha}^\varepsilon(u)=i\}}] \Bigg] \\
& \le K \sum_{k=0}^{M_\varepsilon - 1} E^{1/2} \left| \int_{s_k}^{s_{k+1}} [I_{\{\alpha^\varepsilon(u)=s_{ij}\}} - \nu^i_j(u) I_{\{\overline{\alpha}^\varepsilon(u)=i\}}] du \right|^2 \\
& \le K M_\varepsilon \varepsilon^{1/2} \le K\varepsilon^{\Delta-(1/2)} \to 0 \ \text{ as } \ \varepsilon \to 0.
\end{aligned}
$$

By the weak convergence of $(X^{\varepsilon,N}(\cdot), \overline{\alpha}^\varepsilon(\cdot))$, the Skorohod representation, and the definition of $\overline{Q}(t)$, we obtain that as $\varepsilon \to 0$,

$$E\rho(X^{\varepsilon,N}(t_i), \overline{\alpha}^\varepsilon(t_i) : i \le \kappa) \sum_{k=0}^{M_\varepsilon - 1} \left[\sum_{i=1}^{l} \sum_{j=1}^{m_i} \int_{s_k}^{s_{k+1}} \widehat{Q}(u) \widehat{V}(X^{\varepsilon,N}(s_k), s_{ij}, u) \nu_j^i(u) I_{\{\overline{\alpha}^\varepsilon(u)=i\}} \right]$$

$$\to E\rho(X^N(t_i), \overline{\alpha}(t_i) : i \le \kappa) \Big[\int_t^{t+s} \overline{Q}(u) V(X^N(u), \overline{\alpha}(u), u) du \Big].$$

This yields that as $\varepsilon \to 0$,

$$\begin{aligned} &E\rho(X^{\varepsilon,N}(t_i), \overline{\alpha}^\varepsilon(t_i) : i \le \kappa) \Big[Q^\varepsilon(u) \widehat{V}(X^{\varepsilon,N}(u), \cdot, u)(\alpha^\varepsilon(u)) du \Big] \\ &\quad \to E\rho(X^N(t_i), \overline{\alpha}(t_i) : i \le \kappa) \Big[\overline{Q}(u) V(X^N(u), \cdot, u)(\overline{\alpha}(u)) du \Big]. \end{aligned} \tag{2.22}$$

Using similar argument, we obtain that as $\varepsilon \to 0$,

$$\begin{aligned} &E\rho(X^{\varepsilon,N}(t_i), \overline{\alpha}^\varepsilon(t_i) : i \le \kappa) \Big[- \sum_{k=0}^{M_\varepsilon - 1} \int_{s_k}^{s_{k+1}} [X_2^{\varepsilon,N}(u) f^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) \\ &\qquad + g^N(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u)] \frac{\partial}{\partial x_2} \widehat{V}(X^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) du \Big] \\ &\to E\rho(X^N(t_i), \overline{\alpha}(t_i) : i \le \kappa) \Big[- \int_t^{t+s} [X_2^N(u) \overline{f}^N(X_1^N(u), \overline{\alpha}(u), u) \\ &\qquad + \overline{g}^N(X_1^N(u), \overline{\alpha}(u), u)] \frac{\partial}{\partial x_2} V(X^N(u), \overline{\alpha}(u), u) du \Big], \end{aligned} \tag{2.23}$$

$$\begin{aligned} &E\rho(X^{\varepsilon,N}(t_i), \overline{\alpha}^\varepsilon(t_i) : i \le \kappa) \Big[\sum_{k=0}^{M_\varepsilon - 1} \int_{s_k}^{s_{k+1}} \frac{1}{2} h^{N,2}(X_1^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) \frac{\partial^2}{\partial x_2^2} \widehat{V}(X^{\varepsilon,N}(u), \alpha^\varepsilon(u), u) \Big] \\ &\to E\rho(X^N(t_i), \overline{\alpha}(t_i) : i \le \kappa) \Big[\int_t^{t+s} \frac{1}{2} \overline{h}^{N,2}(X_1^N(u), \overline{\alpha}(u), u) \frac{\partial^2}{\partial x_2^2} V(X^N(u), \overline{\alpha}(u), u) \Big], \end{aligned} \tag{2.24}$$

and

$$E\rho(X^{\varepsilon,N}(t_i),\overline{\alpha}^\varepsilon(t_i):i\le\kappa)\Big[\sum_{k=0}^{M_\varepsilon-1}\int_{s_k}^{s_{k+1}}\frac{\partial}{\partial u}\widehat{V}(X^{\varepsilon,N}(u),\alpha^\varepsilon(u),u)du\Big]$$
$$\to E\rho(X^N(t_i),\overline{\alpha}(t_i):i\le\kappa)\Big[\int_t^{t+s}\frac{\partial}{\partial u}\overline{V}(X^N(u),\overline{\alpha}(u),u)du\Big]. \tag{2.25}$$

Combining the estimates obtained thus far, (2.16), (2.22), (2.23), (2.24), and (2.25) imply that

$$E\rho(X^N(t_i),\overline{\alpha}(t_i):i\le\kappa)\Big[V(X^N(t+s),\overline{\alpha}(t+s),t+s)-V(X^N(t),\overline{\alpha}(t),t)$$
$$-\int_t^{t+s}\overline{\mathcal{L}}V(X^B(u),\overline{\alpha}(u),u)du\Big]\ \text{ is a martingale.}$$

Therefore, $(X^N(\cdot),\overline{\alpha}(\cdot))$ is a solution of the martingale problem with operator $\overline{\mathcal{L}}^N$. Thus Proposition 2.5 is proved.

Proposition 2.7. *Under the conditions of Theorem* 2.3*,* $(X^\varepsilon(\cdot),\overline{\alpha}^\varepsilon(\cdot))$ *converges weakly to* $(X(\cdot),\overline{\alpha}(\cdot))$ *that is the solution of the Liénard equation* (2.7)*.*

Proof. The proof follows along the line of Step 4 of the proof of Theorem 5.1 [24, p. 285] (see also [21, p.46]). Let $P_{X(0)}(\cdot)$ and $P^N(\cdot)$ be the measures on the Borel subsets of $D([0,T],\mathbb{R}^2)$ induced by the solutions $X(\cdot)$ and $X^N(\cdot)$ of the corresponding Liénard equations, respectively. The measure $P^N(\cdot)$ is unique owing to the uniqueness of the solution to (2.10). Note that $P_{X(0)}(\cdot)$ and $P^N(\cdot)$ agree on all Borel subsets of the set of paths in $D([0,T],\mathbb{R}^2)$ whose values are in S_N. Note $P_{X(0)}(\sup_{t\le T}|X(t)|\le N)\to 1$ as $N\to\infty$. These together with the weak convergence of $X^{\varepsilon,N}(\cdot)$ to $X^N(\cdot)$ imply the weak convergence of $X^\varepsilon(\cdot)$ to $X(\cdot)$. Since the limit is unique owing to the argument in Step 1 of the proof of Theorem 2.3, the chosen subsequence is irrelevant. Thus $X^\varepsilon(\cdot)$ converges weakly to $X(\cdot)$.

By Proposition 2.7, we have established the desired convergence of $(X^\varepsilon(\cdot), \overline{\alpha}^\varepsilon(\cdot))$. Therefore, the proof of Theorem 2.3 is complete.

Remark 2.8. We have developed a weak convergence result Theorem 2.3. Here we consider a couple of specializations. First, suppose that there is no white noise perturbations involved. We consider instead

$$\begin{aligned} dX_1^\varepsilon(t) &= X_2^\varepsilon(t)dt, \\ dX_2^\varepsilon(t) &= -\big(X_2^\varepsilon(t) f(X_1^\varepsilon(t), \alpha^\varepsilon(t), t) + g(X_1^\varepsilon(t), \alpha^\varepsilon(t), t)\big)dt, \\ X^\varepsilon(0) &= x \quad \text{and} \quad \alpha^\varepsilon(0) = \iota. \end{aligned} \tag{2.26}$$

Thus we have fully degenerate case to deal with. Using the martingale problem formulation, we obtain that $(X^\varepsilon(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ converges weakly to $(X(\cdot), \overline{\alpha}(\cdot))$ such that $X(\cdot)$ is a solution of the averaged limit switching ODEs

$$\begin{aligned} dX_1(t) &= X_2(t)dt, \\ dX_2(t) &= -\big(X_2(t)\overline{f}(X_1(t), \overline{\alpha}(t), t) + \overline{g}(X_1(t), \overline{\alpha}(t), t)\big)dt, \end{aligned} \tag{2.27}$$

where $\overline{f}(\cdot)$ and $\overline{g}(\cdot)$ are as defined in (2.8). Such switched ODEs have become more important tools in modeling of networked systems in various situations.

The second specialization is: Suppose that the generator $\widetilde{Q}(t)$ itself is weakly irreducible. Denote its quasi-stationary distribution by $\nu(t) = (\nu_1(t), \ldots, \nu_m(t))$. That is, all of its states belong to the same weakly irreducible class. In this case, $\alpha^\varepsilon(\cdot)$ acts like a noise. The rationale is that as $\varepsilon \to 0$, the system changes more rapidly resulting in an average take place. Thus, we obtain

$$\begin{aligned} dX_1(t) &= X_2(t)dt, \\ dX_2(t) &= -(X_2(t)\overline{f}(X_1(t),t) + \overline{g}(X_1(t),t))dt, \end{aligned} \tag{2.28}$$

where

$$\begin{aligned} \overline{f}(x_1,t) &= \sum_{j=1}^{m} f(x_1,j,t)\nu_j(t), \\ \overline{g}(x_1,t) &= \sum_{j=1}^{m} g(x_1,j,t)\nu_j(t). \end{aligned} \tag{2.29}$$

If m is a rather large number, then substantial reduction of complexity is achieved. The original system is one with m discrete components, whereas the limit is a single Liénard equation.

2.4 Further Remarks

We have considered Liénard equations under white noise perturbation and regime switching. A limit system is obtained by means of martingale problem formulation. There are several extensions and generalizations. We mention them in what follows.

2.4.1 Inclusion of Transient States

The developments of the previous sections treat such Markov chains having only recurrent states. It is a natural generalization to examine stochastic Liénard equations with Markov switching in which the Markov chain also has transient states. To begin, let $\alpha^\varepsilon(t) \in \mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cdots \cup \mathcal{M}_l \cup M_*$, where $\mathcal{M}_i$ for $i = 1, \ldots, l$ are as before and $\mathcal{M}_* = \{s_{*1}, \ldots, s_{*m_*}\}$ represents the transient states. The corresponding generator is still of the form (2.3), but

$$\widetilde{Q}(t) = \begin{pmatrix} \widetilde{Q}^1(t) & & & \\ & \ddots & & \\ & & \widetilde{Q}^l(t) & \\ \widetilde{Q}^1_*(t) & \cdots & \widetilde{Q}^l_*(t) & \widetilde{Q}_*(t) \end{pmatrix}. \tag{2.30}$$

(A4) For all $t \in [0, T]$, and $i = 1, \ldots, l$, $\widetilde{Q}^i(t)$ are weakly irreducible, and all eigenvalues of $\widetilde{Q}_*(t)$ have negative real parts. Both $\widetilde{Q}(\cdot)$ and $\widehat{Q}(\cdot)$ are bounded and measurable.

Using the approach as in [37], partition

$$\widehat{Q}(t) = \begin{pmatrix} \widehat{Q}^{11}(t) & \widehat{Q}^{12}(t) \\ \widehat{Q}^{21}(t) & \widehat{Q}^{22}(t) \end{pmatrix},$$

where

$$\widehat{Q}^{11}(t) \in \mathbb{R}^{(m-m_*)\times(m-m_*)},\ \widehat{Q}^{12}(t) \in \mathbb{R}^{(m-m_*)\times m_*},$$

$$\widehat{Q}^{21}(t) \in \mathbb{R}^{m_*\times(m-m_*)},\ \text{ and }\ \widehat{Q}^{22}(t) \in \mathbb{R}^{m_*\times m_*},$$

and define

$$\overline{Q}_*(t) = \mathrm{diag}(\nu^1(t), \ldots, \nu^l(t))(\widehat{Q}^{11}(t)\widetilde{\mathbb{1}} + \widehat{Q}^{12}(t)(a_{m_1}(t), \ldots, a_{m_l}(t)))$$
$$\overline{Q}(t) = \mathrm{diag}(\overline{Q}_*(t), 0_{m_* \times m_*}), \tag{2.31}$$

where

$$\widetilde{\mathbb{1}} = \mathrm{diag}(\mathbb{1}_{m_1}, \ldots, \mathbb{1}_{m_l}),\ \mathbb{1}_{m_j} = (1, \ldots, 1)' \in \mathbb{R}^{m_j \times 1},$$

and

$$a_{m_i}(t) = -\widetilde{Q}_*^{-1}(t)\widetilde{Q}_*^i(t)\mathbb{1}_{m_i},\ \text{for } i = 1, \ldots, l. \tag{2.32}$$

Using essentially the same argument but with modification on the transient part with the help of using asymptotic results in [38], we obtain the following result.

Theorem 2.9. *Under the conditions of Theorem* 2.3 *with* (A3) *replaced by* (A4), $(X^\varepsilon(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ *converges weakly to* $(X(\cdot), \overline{\alpha}(\cdot))$ *such that* $X(\cdot)$ *is a solution of* (2.7) *and* $\overline{\alpha}(\cdot)$ *is a continuous-time Markov chain generated by* $\overline{Q}_*(t)$ *given in* (2.31).

2.4.2 Wide-band Noise Perturbations

In lieu of (2.1), we can consider

$$\begin{aligned} dX_1^\varepsilon(t) &= X_2^\varepsilon(t)dt, \\ dX_2^\varepsilon(t) &= -\big(X_2^\varepsilon(t)f(X_1^\varepsilon(t), \alpha^\varepsilon(t), t) + g(X_1^\varepsilon(t), \alpha^\varepsilon(t), t)\big)dt + \frac{1}{\varepsilon}h(X_1^\varepsilon(t), \alpha^\varepsilon(t), t)\xi^\varepsilon(t), \\ X^\varepsilon(0) &= x \ \text{ and } \ \alpha^\varepsilon(0) = \iota, \end{aligned} \tag{2.33}$$

where $\xi^\varepsilon(t) = \xi(t/\varepsilon^2)$ and $\xi(\cdot)$ is a stationary process with $E\xi(t) = 0$, $E\xi^2(t) = 1$, and $E\xi^{2+\delta}(t) < \infty$ for some $\delta > 0$ such that $\xi(\cdot)$ is independent of the Markov chain $\alpha^\varepsilon(\cdot)$. What we have here is to replace the white noise by a wide-band noise process so that it "approximates" the white noise. Recall that a wide-band noise process is one whose bandwidth is large and as $\varepsilon \to 0$, it approximates the white noise. This is a physical realization of the ideal white noise. It often appears in many applications. Under such a setup, we can still derive the desired limit result. We omit the details, but state the main result below.

Theorem 2.10. *Assume the conditions of Theorem* 2.3 *with the modification that the Brownian motion is replaced by $\xi^\varepsilon(t)/\varepsilon$. Then the conclusion of Theorem* 2.3 *continuous to hold. That is, $(X^\varepsilon(\cdot), \overline{\alpha}^\varepsilon(\cdot))$ converges weakly to $(X(\cdot), \overline{\alpha}(\cdot))$ such that $\overline{\alpha}(\cdot)$ is a continuous-time Markov chain with generator $\overline{Q}(t)$ and $X(\cdot)$ is a solution of* (2.7).

2.4.3 Future Study

For future study, we may consider the case that the generator $Q^\varepsilon(\cdot)$ is also x dependent. Specifically, we may consider that $Q^\varepsilon(x,t)$ has similar form as (2.3), but

$$Q^\varepsilon(x,t) = \frac{1}{\varepsilon}\widetilde{Q}(x,t) + \widehat{Q}(x,t), \tag{2.34}$$

where $\widetilde{Q}(x,t)$ have the same diagonal block form as (2.4) but are also x-dependent. We assume that for each $i = 1, \ldots, l$, $\widetilde{Q}^i(x,t)$ is weakly irreducible in the sense that the system of equation $\nu^i(x,t)\widetilde{Q}^i(x,t) = 0$ and $\nu^i(x,t)\mathbb{1} = 1$ has a unique solution. Assume that $\widetilde{Q}(x,t)$ and $\widehat{Q}(x,t)$ are bounded and continuous functions. In this case, we expect the averaged functions to have the forms

$$\begin{aligned}
\overline{f}(x,i,t) &= \sum_{j=1}^{m_i} \nu_j^i(x,t) f(x_1,j,t),\\
\overline{g}(x,i,t) &= \sum_{j=1}^{m_i} \nu_j^i(x,t) g(x_1,j,t), \qquad (2.35)\\
\overline{h}^2(x,i,t) &= \sum_{j=1}^{m_i} \nu_j^i(x,t) h^2(x_1,j,t).
\end{aligned}$$

Then it seems that we could proceed as in the previous case. Nevertheless, we note that $\alpha^\varepsilon(t)$ is no longer a Markov chain due to the x-dependence of its generator. Care needs to be exercised to handle such cases.

CHAPTER 3

Near-optimal Controls of Stochastic Differential Equations

3.1 Introduction

This chapter is devoted to the study of near-optimal controls of singularly perturbed control systems. This chapter is arranged as follows. We first give the problem formulation in Section 3.2. To proceed we use relaxed control formulation in Section 3.3. Finally, Section 3.4 proves the limit systems and shows that the optimal control for the original singular perturbed problem.

3.2 Problem Formulation

Consider a stochastic dynamical system with the states $x^\varepsilon(t) \in \mathbb{R}^k$ and a feedback control $u(\cdot)$ such that $u(t) \in \Gamma$, $t \geq 0$ where Γ is a compact subset of Euclidean space. Let $w(\cdot)$ be a standard one-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ with respect to the filtration $\mathcal{F}_t$. Suppose that $\alpha^\varepsilon(\cdot)$ is a continuous-time Markov Chain with ε being a small parameter with a finite-state space $\mathcal{M} = \{1, 2, ..., l\}$. Consider $b(\cdot,\cdot,\cdot) : \mathbb{R}^k \times \mathcal{M} \times U \to \mathbb{R}^k$, $\sigma(\cdot,\cdot) : \mathbb{R}^k \times \mathcal{M} \to \mathbb{R}^{2k}$ are given suitable functions. Assume that $\alpha^\varepsilon(\cdot)$ and $w(\cdot)$ are independent. Consider the SDE with regime switching as follows

$$x^{\varepsilon}(t) = x_0 + \int_0^t b(x^{\varepsilon}(s), \alpha^{\varepsilon}(s), u^{\varepsilon}(s))ds + \int_0^t \sigma(x^{\varepsilon}(s), \alpha^{\varepsilon}(s))dw(s) \tag{3.1}$$

$$\alpha^{\varepsilon}(0) = i_0 \in \mathcal{M}$$

Consider the cost function $J^{\varepsilon}(x_0, i_0, u^{\varepsilon}(\cdot))$ as

$$J^{\varepsilon}(u^{\varepsilon}(\cdot)) = J^{\varepsilon}(x_0, i_0, u^{\varepsilon}(\cdot)) = E_{x_0,i_0} \int_0^T C(x^{\varepsilon}(s), \alpha^{\varepsilon}(s), u^{\varepsilon}(s))ds \tag{3.2}$$

where E_{x_0,i_0} denotes the expectation taken with $x^{\varepsilon}(0) = x_0$ and $\alpha^{\varepsilon}(0) = i_0$. Our objective is to find the optimal control $u^{\varepsilon}(\cdot)$ that minimizes $J^{\varepsilon}(u^{\varepsilon}(\cdot))$.

Throughout the chapter we need the following assumption

(A1) Assume that the generator of the Markov Chain has the form

$$Q^{\varepsilon}(t) = \frac{1}{\varepsilon}\widetilde{Q}(t) + \widehat{Q}(t), \tag{3.3}$$

such that

$$\widetilde{Q}(t) = \text{diag}(\widetilde{Q}^1(t), \ldots, \widetilde{Q}^l(t)) \tag{3.4}$$

with all $\widetilde{Q}^i(t) \in \mathbb{R}^{m_i \times m_i}$ are irreducible generators for $k = 1, ..., l$. Both $\widetilde{Q}(t)$, $\widehat{Q}(t)$ are generators, bounded and measurable. To reduce the computational complexity we decompose the state space into connected subspace, the state space of $\alpha^{\varepsilon}(\cdot)$ is given by $\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_l = \{s_{11}, ..., s_{1m_1}, ..., s_{l1}, ..., s_{lm_i}\}$. We use a double index s_{ij} to denote a state.

(A2) There exist a positive constant K such that for each $\alpha \in \mathcal{M}$, each $c \in U$, and $x, y \in \mathbb{R}^k$,

$$|b(x, \alpha, c) - b(y, \alpha, c)| \leq K|x - y|,$$

$$|\sigma(x, \alpha) - \sigma(y, \alpha)| \leq K|x - y|.$$

(A3) There exist a positive constants K and positive integer κ such that for each $\alpha \in \mathcal{M}$ and any $c \in U$,

$$|C(x, \alpha, c)| \leq K(1 + |x|^\kappa).$$

In our setup, $\mathcal{M}$ is rather large in that the cardinality $|\mathcal{M}|$ of $\mathcal{M}$ is a large number. Treating reduction of dimensionality, in [37], it is illustrated that we can aggregate the elements corresponding to each $\mathcal{M}_i$ into a single state. Then the total number of states in the aggregated process will be l. Denote the reduced space by $\bar{\mathcal{M}} = \{1, \ldots, l\}$, and define $\bar{\alpha}^\varepsilon(t) = i$ if $\alpha^\varepsilon(t) \in \mathcal{M}_i$.

3.3 Relaxed Control Formulation

In this subsection we discuss the basic relaxed control and rewrite the control problem using relaxed control formulation. The relaxed stochastic control was inaugurated in [9] for stochastic system. The relaxed control was used in [33] for variational problem. The Near optimal state feedback controls for stochastic systems with wideband noise disturbance was discussed in [23]. Singularity perturbed stochastic control is devoted in [2]. Lately, much interests are devoted to perturbed stochastic control and filtering problems is added [22].

Recall that the control space U is a compact set in some Euclidean space. Assume that

$\mathcal{B}(S)$ is the σ-algebra of Borel subset of S. Suppose that

$$M^* = \{\tilde{m}(\cdot);\ \tilde{m}(\cdot) \text{ is a measure on } \mathcal{B}(U \times [0,\infty)) \text{ and}$$

$$\tilde{m}(U \times [0,t]) = t \text{ for all } t \geq 0\}.$$

A random M^*-valued measure $m(\cdot)$ is an admissible relaxed control if for each $B \in \mathcal{B}(U)$, the function defined by $\tilde{m}(B,t) \equiv \tilde{m}(B \times [0,t])$ is $\mathcal{F}_t$ measurable. An equivalent formulation reads that $\tilde{m}(\cdot)$ is a relaxed control if

$$\textstyle\int_0^t f(s,c)\tilde{m}(ds \times dc)$$

is progressively measurable with respect to $\mathcal{F}_t$ for each bounded and continuous function $f(\cdot)$.

It then follows that, if $\tilde{m}(\cdot)$ is an admissible relaxed control, then there is a measure-valued function $\tilde{m}_t(\cdot)$ (the "derivative") such that $\tilde{m}_t(dc)dt = \tilde{m}(dt \times dc)$ and for smooth function$f(\cdot)$,

$$\textstyle\int f(s,c)\tilde{m}(ds \times dc) = \int ds \int f(s,c)\tilde{m}_s(dc).$$

To proceed, we topologize M^* as follows. Let $\{f_{n_i}(\cdot); i < \infty\}$ be a countable dense (under the sup-norm) set of continuous functions on $U \times [0,n]$ for each n. Let

$$\langle \tilde{m}, f \rangle = \int f(s,c)\tilde{m}(ds \times dc), \text{ and}$$

$$d(\tilde{m}_1, \tilde{m}_2) = \sum_{n=1}^{\infty} 2^{-n} d_n(\tilde{m}_1, \tilde{m}_2),$$

where

$$d_n(\tilde{m}_1, \tilde{m}_2) = \sum_{i=1}^{\infty} 2^{-i} \frac{|(\tilde{m}_1 - \tilde{m}_2, f_{n_i})|}{1 + |(\tilde{m}_1 - \tilde{m}_2, f_{n_i})|}.$$

$\tilde{m}_n(\cdot) \Rightarrow \tilde{m}(\cdot)$ for a sequence of measures means the weak convergence in M^*.

By using the relaxed control we can rewrite the switching diffusion defined in (3.1) as follows

$$\begin{aligned} x^\varepsilon(t) &= x_0 + \textstyle\int_0^t \int_U b(x^\varepsilon(s), \alpha^\varepsilon(s), c))\tilde{m}_s^\varepsilon(dc)ds + \int_0^t \sigma(x^\varepsilon(s), \alpha^\varepsilon(s))dw(s) \\ \alpha^\varepsilon(0) &= i_0 \in \mathcal{M} \end{aligned} \tag{3.5}$$

and the cost function defined in (3.2) as

$$J^\varepsilon(\tilde{m}^\varepsilon(\cdot)) = J^\varepsilon(x_0, i_0, u^\varepsilon(\cdot)) = E_{x_0,i_0} \int_0^T \int_U C(x^\varepsilon(s), \alpha^\varepsilon(s), c)\tilde{m}_s^\varepsilon(dc)ds \tag{3.6}$$

where $\tilde{m}_s^\varepsilon$ is the relaxed control. Our objective is to find the optimal control $\tilde{m}^\varepsilon(\cdot)$ that minimizes $J^\varepsilon(\tilde{m}^\varepsilon(\cdot))$.

The associated operator for the switching diffusion defined in (3.5) is defined as follows; for each $i \in \mathcal{M}$, and for any twice continuously differentiable function $g(\cdot, i)$ define $\mathcal{L}$ by

$$\begin{aligned} \mathcal{L}g(x,\alpha) = &\frac{1}{2}\mathrm{tr}[\sigma(x,\alpha)\sigma'(x,\alpha)\nabla^2 g(x,\alpha)] + \int_U b'(x,\alpha,c)\nabla g(x,\alpha)\tilde{m}_t(dc) \\ &+Q^\varepsilon(t)g(x,\cdot)(\alpha),\ \ \alpha \in \mathcal{M} \end{aligned} \tag{3.7}$$

where

$$Q^\varepsilon(t)g(x,\alpha)(\cdot) = \sum_\beta q^\varepsilon_{\alpha\beta}(t)g(x,\beta),$$

$\nabla g(x,i)$ and $\nabla^2 g(x,i)$ denote the gradient and Hessian of $g(x,i)$, respectively. Note that the operator $\mathcal{L}$ in fact is control and ε dependent. So we could write it as $\mathcal{L}^{\varepsilon,\tilde{m}}$. However, for notational simplicity, we suppressed these dependence henceforth. To proceed, we state a couple of preliminary results.

Lemma 3.1. *The following assertions hold:*

(a) Let $\tilde{m}(\cdot)$ *be an admissible relaxed control for the limit problem* (3.1) *with* $x = x(0)$*. Then there is an* $\mathcal{F}_t = \sigma\{x(s), \alpha(s); s \leq t\}$ *adapted solution* $x(\cdot)$ *of the limit problem such that*

$$\sup_{0 \leq t \leq T} E|x(t)|^2 \leq K(1 + |x|^2). \tag{3.8}$$

(b) Let $\tilde{m}^n(\cdot)$ *converge weakly to* $m(\cdot)$*, where* $\tilde{m}^n(\cdot)$ *are admissible w.r.t. some Brownian motion* $w(\cdot)$*. Let* $x(\tilde{m}^n(\cdot), \cdot)$ *be the trajectory satisfying* (3.5) *with* $\tilde{m}^n(\cdot)$ *used. Then* $(x^n(\tilde{m}^n, \cdot), \tilde{m}^n(\cdot))$ *converges weakly to* $(x(\tilde{m}(\cdot), \cdot), \tilde{m}(\cdot))$ *satisfying* (3.5) *for some Brownian motion* $w(\cdot)$ *such that* $\tilde{m}(\cdot)$ *is admissible with respect to* $w(\cdot)$*.*

Lemma 3.2. *Assume that* (A2) *and equation* (3.10) *has a unique* (*in the weak sense*) *solution for each initial condition for the admissible triple* $(\tilde{m}(\cdot), w(\cdot))$*. Consider the cost function* (3.18)*. Given* $T > 0$ *and* $\Delta > 0$*, there are a finite set* $\{a_1^\Delta, \ldots, a_{k_\Delta}^\Delta\} = U^\Delta \subset U$*, a* $\delta > 0$*, and a* U^Δ*-valued ordinary admissible stochastic control* $u^\Delta(\cdot)$ *that is constant on each interval* $[i\delta, i\delta + \delta)$ *and is such that for all* m*,*

$$\begin{aligned} &P_{x,i_0}^{\tilde{m}}(\sup_{t \leq T} |x(t, u^\Delta) - x(t, \tilde{m})| > \Delta) \leq \Delta, \\ &|J(x, i_0, \tilde{m}) - J(x, i_0, u^\Delta)| \leq \Delta. \end{aligned} \tag{3.9}$$

If the solution to (3.10) *is unique in the weak sense for each admissible* $\tilde{m}(\cdot)$*, then* (3.9) *holds for all* $\tilde{m}(\cdot)$ *simultaneously.*

The proof of Lemma 3.1 (a) can be found in [39], and the proofs of Lemma 3.1 (b) as well as Lemma 3.2 can be proved using the same techniques as in [8, 22, 23]. We thus omit the details. Note that in Lemma 3.1, we can prove a strong result

$$E \sup_{0\le t\le T} |x^{\varepsilon}(t)|^2 < \infty.$$

However, for our current problem, Lemma 3.1 is sufficient.

3.4 Limit Results and Near-Optimal Control

In this section, we show that the weak limits defined in (3.1). We will use the ideal of martingale problem. Finally, we present the existence of the optimal control.

Theorem 3.3. *Assume* (A1)-(A3)*. Let* $\Delta_\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and let* $\tilde{m}^{\varepsilon}(\cdot)$ *be a* δ*-optimal admissible relaxed control for* (3.1) *with the cost defined in* (3.2)*. Then the following assertions hold:*

(a) $\{x^{\varepsilon}(\tilde{m}^{\varepsilon},(\cdot)), \bar{\alpha}^{\varepsilon}(\cdot), \tilde{m}^{\varepsilon}(\cdot)\}$ *is tight in* $D([0,T]:\mathbb{R}^k\times\mathcal{M}\times M^*)$*.*

(b) $(x^{\varepsilon}(\tilde{m}^{\varepsilon}(\cdot),(\cdot)),\cdot), \bar{\alpha}^{\varepsilon}(\cdot), \tilde{m}^{\varepsilon}(\cdot))$ *converge weakly to* $(x(\tilde{m}(\cdot), \bar{\alpha}(\cdot), \tilde{m}(\cdot))$ *such that* $x(\cdot)$ *is solution of the switching diffusion*

$$x(t) = x_0 + \int_0^t \int_U \bar{b}(x(s),\bar{\alpha}(s),c)m_s(dc) + \int_0^t \bar{\sigma}(x(s),\bar{\alpha}(s))dw(s) \tag{3.10}$$

where

$$\begin{aligned} \bar{b}(x,i,u) &= \sum_{j=1}^{m_i} \nu_j^i(t) b(x,s_{ij},u) \\ \bar{\sigma}(x,i)\bar{\sigma}'(x,i) &= \sum_{j=1}^{m_i} \nu_j^i(t)\sigma(x,s_{ij})\sigma'(x,s_{ij}). \end{aligned} \tag{3.11}$$

Proof. We divided the proof of the theorem into several steps.

Step 1. Tightness. First note that Since the space of relaxed control M^* is compact, $\{\tilde{m}^\varepsilon(\cdot)\}$ is tight in M^*. By virtue of [37, Theorem 7.4], $\{\bar{\alpha}^\varepsilon(\cdot)\}$ is tight.

Next, we work with $x^{\varepsilon,N}(\cdot)$. We use a truncation device. For any $0 < N < \infty$, let $x^{\varepsilon,N}(\cdot)$ be the N-truncation of $x^\varepsilon(\cdot)$, which is $x^{\varepsilon,N}(t) = x^\varepsilon(t)$ up until the first exit from the sphere $S_N = \{x : |x| \le N\}$. Define a sufficiently smooth truncation $\rho_N(x)$ as follows: $\rho(x) = 1$ if $x \in S_N$ and $\rho(x) = 0$ when $x \in \mathbb{R}^k - S_{N+1}$. Define

$$\phi_N(x,\alpha,c) = \phi(x,\alpha,c)\rho_N(x) \quad \text{for} \quad \phi = b \text{ or } \sigma.$$

Then we can rewrite the differential equation as

$$x^{\varepsilon,N}(t) = x_0 + \int_0^t \int_U b_N(x^{\varepsilon,N}(s),\alpha^\varepsilon(s),c))\tilde{m}_s^\varepsilon(dc)ds + \int_0^t \sigma_N(x^{\varepsilon,N}(s),\alpha^\varepsilon(s))dw(s). \quad (3.12)$$

By using the N-truncation and hence the boundedness of $x^{\varepsilon,N}(\cdot)$, the linear growth of the b and σ, the Hölder inequality, the basic properties of stochastic integration, and Lemma 3.1, we have that for any $\delta > 0$ and $\varepsilon > 0$, and for any $t, s \ge 0$ with $s \le \Delta$, there is a random variable $\gamma^\varepsilon(\delta) > 0$ such that

$$\begin{aligned}
&E_t^\varepsilon|x^{\varepsilon,N}(t+s) - x^{\varepsilon,N}(t)|^2 \\
&\leq KE_t^\varepsilon\left|\int_t^{t+s}\int_U b_N(x^{\varepsilon,N}(r),\alpha^\varepsilon(r),c))\tilde{m}_r^\varepsilon(dc)dr\right|^2 \\
&\quad +KE_t^\varepsilon\left|\int_t^{t+s}\sigma_N(x^{\varepsilon,N}(r),\alpha^\varepsilon(r))dw(r)\right|^2 \\
&\leq KsE_t^\varepsilon\int_t^{t+s}\left|\int_U b_N(x^{\varepsilon,N}(r),\alpha^\varepsilon(r),c))\tilde{m}_r^\varepsilon(dc)\right|^2 dr \\
&\quad +KE_t^\varepsilon\int_t^{t+s}|\sigma_N(x^{\varepsilon,N}(r),\alpha^\varepsilon(r))|^2 dr \\
&\leq E_t^\varepsilon\gamma^\varepsilon(\delta).
\end{aligned}$$

Taking $\limsup$ as $\varepsilon \to 0$ followed by $\lim_{\delta\to 0}$, we obtain

$$\lim_{\delta\to 0}\limsup_{\delta\to 0} E\gamma^\varepsilon(\delta) = 0.$$

Thus, the tightness criterion [21, Theorem 3, p. 47], $\{x^{\varepsilon,N}(\cdot)\}$ is tight. We thus obtain that $(x^{\varepsilon,N}(\cdot),\bar{\alpha}^\varepsilon(\cdot),\tilde{m}^\varepsilon(\cdot))$ is tight.

Step 2. Characterization of the limit. To begin, define a truncated operator by

$$\mathcal{L}_N g(x,\alpha) = \frac{1}{2}\mathrm{tr}[\sigma_N(x,\alpha)\sigma_N'(x,\alpha)\nabla^2 g(x,\alpha)] + \int_U b_N'(x,\alpha,c)\nabla g(x,\alpha)\tilde{m}_t^\varepsilon(dc),\ \ \alpha\in\mathcal{M}. \tag{3.13}$$

Again, we suppressed the ε and control dependence in $\mathcal{L}_N$. Since $(x^{\varepsilon,N}(\cdot),\bar{\alpha}^\varepsilon(\cdot),\tilde{m}^\varepsilon(\cdot))$ is tight, by Prohorov's theorem, it has a convergent subsequence. For notational simplicity, we still denote the sequence by $\{x^{\varepsilon,N}(\cdot),\bar{\alpha}^\varepsilon(\cdot),\tilde{m}^\varepsilon(\cdot)\}$ and denote the limit by $(x^N(\cdot),\bar{\alpha}(\cdot),\tilde{m}(\cdot))$. By

Skorohod representation, with a slight abuse of notation, we may assume that the convergence is in the sense of w.p.1. The convergence is uniform on any finite time interval. We proceed to characterize the limit process.

For each $\iota \in \bar{\mathcal{M}}$, each $f(\cdot, \iota) \in \mathcal{C}_0^2$ (C^2 functions with compact support), any bounded and contiguous function $h(\cdot)$, any $0 \le t, s < T$, any positive integers p, q, and any $t_i \le t$, we aim to show that

$$\begin{aligned}&Eh(x^N(t_i), \bar{\alpha}(t_i), \langle \varphi_j, \tilde{m} \rangle_{t_i}, i \le q, j \le p) \\ &\quad \times \Big[(f(x^N(t+s), \bar{\alpha}(t+s)) - f(x^N(t), \bar{\alpha}(t))) - \int_t^{t+s} \mathcal{L} f(x^N(\tau), \bar{\alpha}(\tau)) d\tau \Big] = 0.\end{aligned} \tag{3.14}$$

We start with the process index by ε. By the weak convergence and the Skorohod representation,

$$\begin{aligned}&\lim_{\varepsilon \to 0} Eh(x^{\varepsilon,N}(t_i), \bar{\alpha}^{\varepsilon}(t_i), \langle \varphi_j, \tilde{m}^{\varepsilon} \rangle_{t_i}, i \le q, j \le p)(f(x^{\varepsilon,N}(t+s), \bar{\alpha}^{\varepsilon}(t+s)) - f(x^{\varepsilon,N}(t), \bar{\alpha}^{\varepsilon}(t)) \\ &\qquad \to Eh(x^N(t_i), \bar{\alpha}(t_i), \langle \varphi_j, \tilde{m} \rangle_{t_i}, i \le q, j \le p)(f(x^N(t+s), \bar{\alpha}(t+s)) - f(x^N(t), \bar{\alpha}(t))).\end{aligned} \tag{3.15}$$

To proceed, define

$$\hat{f}(x, \alpha) = \sum_{i \in \bar{\mathcal{M}}} f(x, i) I_{\{\alpha \in \mathcal{M}_i\}}.$$

Note that $f(x^{\varepsilon}(t),\bar{\alpha}^{\varepsilon}(t))=\hat{f}(x^{\varepsilon}(t),\alpha^{\varepsilon}(t))$. It follows that

$$Eh(x^{\varepsilon,N}(t_i),\bar{\alpha}^{\varepsilon}(t_i),\langle\varphi_j,\tilde{m}^{\varepsilon}\rangle_{t_i},i\le q,j\le p)$$

$$\times\Big[(f(x^{\varepsilon,N}(t+s),\bar{\alpha}^{\varepsilon}(t+s))-f(x^{\varepsilon,N}(t),\bar{\alpha}^{\varepsilon}(t))\Big]$$

$$=Eh(x^{\varepsilon,N}(t_i),\bar{\alpha}^{\varepsilon}(t_i),\langle\varphi_j,\tilde{m}^{\varepsilon}\rangle_{t_i},i\le q,j\le p)$$

$$\times\Big[(\hat{f}(x^{\varepsilon,N}(t+s),\alpha^{\varepsilon}(t+s))-\hat{f}(x^{\varepsilon,N}(t),\alpha^{\varepsilon}(t))\Big]$$

$$=Eh(x^{\varepsilon,N}(t_i),\bar{\alpha}^{\varepsilon}(t_i),\langle\varphi_j,\tilde{m}^{\varepsilon}\rangle_{t_i},i\le q,j\le p)$$

$$\times\Big[\int_t^{t+s}\mathcal{L}_N\hat{f}(x^{\varepsilon}(\tau),\alpha^{\varepsilon}(\tau))d\tau\Big].$$

Direct computation reveals that

$$\int_t^{t+s}\mathcal{L}_N\hat{f}(x^{\varepsilon,N}(\tau),\alpha^{\varepsilon}(\tau))d\tau$$

$$=\int_t^{t+s}\int_U\nabla\hat{f}'(x^{\varepsilon,N}(r),\alpha^{\varepsilon}(r))b_N(x^{\varepsilon,N}(r),\alpha^{\varepsilon}(r),c)\tilde{m}_r(dc)dr$$

$$+\frac{1}{2}\int_t^{t+s}\mathrm{tr}[\sigma_N(x^{\varepsilon,N}(r),\alpha(r))\sigma_N'(x^{\varepsilon,N}(r),\alpha^{\varepsilon}(r))\nabla^2\hat{f}(x^{\varepsilon,N}(r),\alpha^{\varepsilon}(r))]dr$$

$$+\int_t^{t+s}Q^{\varepsilon}(r)\hat{f}(x^{\varepsilon,N}(r),\alpha^{\varepsilon}(r))dr.$$

Note that

$$\int_t^{t+s}\int_U \nabla \hat{f}'(x^{\varepsilon,N}(r),\alpha^\varepsilon(r))b_N(x^{\varepsilon,N}(r),\alpha^\varepsilon(r),c)\tilde{m}_r(dc)dr$$

$$= \sum_{\iota=1}^{l}\sum_{\ell=1}^{m_\iota}\int_t^{t+s}\int_U \nabla \hat{f}'(x^{\varepsilon,N}(r),s_{\iota\ell})b_N(x^{\varepsilon,N}(r),s_{\iota\ell},c)\tilde{m}_r(dc)I_{\{\alpha^\varepsilon(r)=s_{\iota\ell}\}}dr$$

$$= \sum_{\iota=1}^{l}\sum_{\ell=1}^{m_\iota}\int_t^{t+s}\int_U \nabla \hat{f}'(x^{\varepsilon,N}(r),s_{\iota\ell})b_N(x^{\varepsilon,N}(r),s_{\iota\ell},c)\tilde{m}_r(dc)\nu_\ell^\iota(r)I_{\{\bar{\alpha}^\varepsilon(r)=\iota\}}dr$$

$$+ \sum_{\iota=1}^{l}\sum_{\ell=1}^{m_\iota}\int_t^{t+s}\int_U \nabla \hat{f}'(x^{\varepsilon,N}(r),s_{\iota\ell})b_N(x^{\varepsilon,N}(r),s_{\iota\ell},c)\tilde{m}_r(dc)\nu_\ell^\iota(r)$$

$$\times[I_{\{\alpha^\varepsilon(r)=s_{\iota\ell}\}} - \nu_\ell^\iota(r)I_{\{\bar{\alpha}^\varepsilon(r)=\iota\}}]dr.$$

Using integration by parts and [37, Lemma 7.14] (see also [36, Theorem 3.6]), it can be shown that

$$Eh(x^{\varepsilon,N}(t_i),\bar{\alpha}^\varepsilon(t_i),\langle\varphi_j,\tilde{m}^\varepsilon\rangle_{t_i}, i\le q, j\le p)$$

$$\times\Big[\sum_{\iota=1}^{l}\sum_{\ell=1}^{m_\iota}\int_t^{t+s}\int_U \nabla \hat{f}'(x^{\varepsilon,N}(r),s_{\iota\ell})b_N(x^{\varepsilon,N}(r),s_{\iota\ell},c)\tilde{m}_r(dc)\nu_\ell^\iota(r)$$

$$\times[I_{\{\alpha^\varepsilon(r)=s_{\iota\ell}\}} - \nu_\ell^\iota(r)I_{\{\bar{\alpha}^\varepsilon(r)=\iota\}}]dr\Big]$$

$$\to 0 \quad \text{as} \quad \varepsilon\to 0.$$

Moreover, we can show that

$$Eh(x^{\varepsilon,N}(t_i), \bar{\alpha}^{\varepsilon}(t_i), \langle \varphi_j, \tilde{m}^{\varepsilon}\rangle_{t_i}, i \le q, j \le p)$$

$$\times\Big[\sum_{\iota=1}^{l}\sum_{\ell=1}^{m_\iota}\int_t^{t+s}\int_U \nabla \hat{f}'(x^{\varepsilon,N}(r), s_{\iota\ell}) b_N(x^{\varepsilon,N}(r), s_{\iota\ell}, c)\tilde{m}_r(dc)\nu_\ell^\iota(r) I_{\{\bar{\alpha}^{\varepsilon}(r)=\iota\}} dr\Big]$$

$$\to Eh(x^N(t_i), \bar{\alpha}(t_i), \langle \varphi_j, \tilde{m}\rangle_{t_i}, i \le q, j \le p)$$

$$\times\Big[\sum_{\iota=1}^{l}\sum_{\ell=1}^{m_\iota}\int_t^{t+s}\int_U \nabla \hat{f}'(x^N(r), s_{\iota\ell}) b_N(x^N(r), s_{\iota\ell}, c)\tilde{m}_r(dc)\nu_\ell^\iota(r) I_{\{\bar{\alpha}(r)=\iota\}} dr\Big]$$

$$= Eh(x^N(t_i), \bar{\alpha}(t_i), \langle \varphi_j, \tilde{m}\rangle_{t_i}, i \le q, j \le p)$$

$$\times\int_t^{t+s}\int_U \nabla \hat{f}'(x^N(r), \bar{\alpha}(r)) \bar{b}_N(x^N(r), \bar{\alpha}(r), c)\tilde{m}_r(dc).$$

Using similar techniques, we can prove that

$$Eh(x^{\varepsilon,N}(t_i), \bar{\alpha}^{\varepsilon}(t_i), \langle \varphi_j, \tilde{m}^{\varepsilon}\rangle_{t_i}, i \le q, j \le p)$$

$$\times\frac{1}{2}\int_t^{t+s} \mathrm{tr}[\sigma_N(x^{\varepsilon,N}(r), \alpha^{\varepsilon}(r))\sigma_N'(x^{\varepsilon,N}(r), \alpha^{\varepsilon}(r))\nabla^2 \hat{f}(x^{\varepsilon,N}(r), \alpha^{\varepsilon}(r))]dr$$

$$\to Eh(x^N(t_i), \bar{\alpha}(t_i), \langle \varphi_j, \tilde{m}\rangle_{t_i}, i \le q, j \le p)$$

$$\times\Big[\frac{1}{2}\int_t^{t+s} \mathrm{tr}[\sigma_N(x^N(r), \bar{\alpha}(r))\sigma_N'(x^N(r), \bar{\alpha}(r))\nabla^2 f(x^N(r), \bar{\alpha}(r))]dr\Big].$$

For the switching part, we have as $\varepsilon \to 0$,

$$
\begin{aligned}
&Eh(x^{\varepsilon,N}(t_i), \bar{\alpha}^\varepsilon(t_i), \langle \varphi_j, \tilde{m}^\varepsilon \rangle_{t_i}, i \le q, j \le p) \\
&\quad \times \Big[\int_t^{t+s} Q^\varepsilon(r) \hat{f}(x^{\varepsilon,N}(r), \alpha^\varepsilon(r)) dr \Big] \\
&\to Eh(x^N(t_i), \bar{\alpha}(t_i), \langle \varphi_j, \tilde{m} \rangle_{t_i}, i \le q, j \le p) \\
&\quad \times \Big[\int_t^{t+s} \bar{Q}(r) f(x^N(r), \bar{\alpha}(r) dr \Big],
\end{aligned}
$$

where

$$
\bar{Q}(t) = \mathrm{diag}(\nu^1, \dots, \nu^l(t)) \widehat{Q}(t) \mathrm{diag}(\mathbb{1}_{m_1}, \dots, \mathbb{1}_{m_l}).
$$

It can be thought of as an average of $\hat{Q}(t)$ with respect to the stationary measures $\nu^1(t), \dots, \nu^l(t))$.

Combing the estimates obtained so far, we have the weak convergence of $(x^{\varepsilon,N}(\cdot), \bar{\alpha}^\varepsilon(\cdot), \tilde{m}^\varepsilon(\cdot))$ to $(x^N(\cdot), \bar{\alpha}, \tilde{m}(\cdot))$. Finally, using the uniqueness of the limit problem and the techniques in [21, p. 46], we conclude that the untruncated process $(x^\varepsilon(\cdot), \bar{\alpha}^\varepsilon(\cdot), \tilde{m}^\varepsilon(\cdot))$ also converges. Effectively, we have shown that $(x(\cdot), \bar{\alpha}(\cdot), \tilde{m}(\cdot))$ is a solution of controlled martingale problem with operator

$$
\begin{aligned}
\bar{\mathcal{L}} g(x,i) = \frac{1}{2} \mathrm{tr}[\bar{\sigma}(x,i)\bar{\sigma}'(x,i) \nabla^2 g(x,i)] + \int_U \bar{b}'(x,i,c) \nabla g(x,i) \tilde{m}_t(dc) \\
+ \bar{Q}(t) g(x,\cdot)(i), \; i \in \bar{\mathcal{M}}.
\end{aligned}
\tag{3.16}
$$

Thus the desired result follows.

Theorem 3.4. *Assume the conditions of Theorem* 3.3. *Let* $\tilde{m}^\varepsilon(\cdot)$ *be admissible relaxed control* *and $\tilde{m}(\cdot)$ is admissible with respect to $w(\cdot)$, then*

$$J^{\varepsilon}(\tilde{m}^{\varepsilon}) \to J(\tilde{m}) \text{ as } \varepsilon \to 0 \tag{3.17}$$

where

$$J(\tilde{m}) = J(x, i_0, \tilde{m}(\cdot)) = E^{m}_{x_0,i_0} \int_0^T \int_U \overline{C}(x(t), \overline{\alpha}(t), c)\tilde{m}_t(dc)dt, \tag{3.18}$$

and

$$\overline{C}(x, i, c) = \sum_{j=1}^{m_i} \nu^i_j(t) C(x, s_{ij}, c).$$

Proof. By the weak convergence of $x^{\varepsilon}(\cdot)$ to $x(\cdot)$ together with the Skorohod representation, it can be seen that

$$\begin{aligned}
&\lim_{\varepsilon\to 0} J^{\varepsilon}(x_0, i_0, m^{\varepsilon}) \\
&= E^{m^{\varepsilon}}_{x,i_0} \int_0^T \int_U C(x^{\varepsilon}(t), \alpha^{\varepsilon}(t), c) m^{\varepsilon}_t(dc)dt \\
&= \lim_{\varepsilon\to 0} \sum_{i=1}^{l} \sum_{j=1}^{m_i} E^{m^{\varepsilon}}_{x_0,i_0} \int_0^T \int_U C(x^{\varepsilon}(t), s_{ij}, c) I_{\{\alpha^{\varepsilon}(t)=s_{ij}\}} m^{\varepsilon}_t(dc)dt \\
&= \lim_{\varepsilon\to 0} \sum_{i=1}^{l} \sum_{j=1}^{m_i} E^{m^{\varepsilon}}_{x_0,i_0} \int_0^T \int_U C(x^{\varepsilon}(t), s_{ij}, c)[I_{\{\alpha^{\varepsilon}(t)=s_{ij}\}} - \nu^i_j(t) I_{\{\overline{\alpha}^{\varepsilon}(t)=i\}}] m^{\varepsilon}_t(dc)dt \\
&\quad + \lim_{\varepsilon\to 0} \sum_{i=1}^{l} \sum_{j=1}^{m_i} E^{m^{\varepsilon}}_{x_0,i_0} \int_0^T \int_U C(x^{\varepsilon}(t), s_{ij}, c) \nu^i_j(t) I_{\{\overline{\alpha}^{\varepsilon}(t)=i\}} m^{\varepsilon}_t(dc)dt
\end{aligned} \tag{3.19}$$

By using Hölder's inequality and a [15, Theorem 7.29], we have

$$
\begin{aligned}
&E\left|\int_0^T \int_U C(x^\varepsilon(t), s_{ij}, c)[I_{\{\alpha^\varepsilon(t)=s_{ij}\}} - \nu_j^i(t) I_{\{\overline{\alpha}^\varepsilon(t)=i\}}] m_t^\varepsilon(dc)dt\right|^2 \\
&\le \Big[\int_0^T [1 + E|x^\varepsilon(t)|^{2n_0}]dt\Big] E\Big[\int_0^T [I_{\{\alpha^\varepsilon(t)=s_{ij}\}} - \nu_j^i(t) I_{\{\overline{\alpha}^\varepsilon(t)=i\}}]dt\Big]^2 \qquad (3.20) \\
&\le O(\varepsilon) \to 0 \ \text{ as } \ \varepsilon \to 0.
\end{aligned}
$$

Combining (3.19) and (3.20), we conclude that

$$
\begin{aligned}
&\lim_{\varepsilon\to 0} J^\varepsilon(x, i_0, m^\varepsilon) \\
&\quad = \sum_{i=1}^{l}\sum_{j=1}^{m_i} E_{x_0,i_0}^m \int_0^T \int_U C(x(t), s_{ij}, c)\nu_j^i(t) I_{\{\overline{\alpha}(t)=i\}} m_t(dc)dt \\
&\quad = \sum_{i=1}^{l}\sum_{j=1}^{m_i} E_{x_0,i_0}^m \int_0^T \int_U \overline{C}(x(t), i, c) I_{\{\overline{\alpha}(t)=i\}} m_t(dc)dt \\
&\quad = E_{x_0,i_0}^m \int_0^T \int_U \overline{C}(x(t), \overline{\alpha}(t), c) m_t(dc)dt
\end{aligned}
$$

The proof is complete.

Let $\mathcal{R}^\varepsilon$ be the set of all admissible controls, i.e.,

$$\mathcal{R}^\varepsilon = \{\tilde{m}^\varepsilon(\cdot) \in M^*; \tilde{m}^\varepsilon(\cdot) \text{ is } \mathcal{F} \text{ adapted }\}$$

and use $\mathcal{R}$ to denote the set of admissible controls for the limit problem,

$$\mathcal{R} = \{m(\cdot) \in M^*; m(\cdot) \text{ is } \mathcal{F}_t \text{ adapted }\}$$

where $\mathcal{F}_t = \sigma\{x(s), \alpha(s); s \le t\}$.

Theorem 3.5. *Assume that* (A1)-(A3). *Then there is δ-optimal control $u^\delta(\cdot)$ of the limit control* (3.10) *such that the cost function*

$$\limsup_{\varepsilon \to 0}[J^\varepsilon(u^\delta) - \inf_{\mathcal{R}^\varepsilon} J^\varepsilon(m)] \leq \delta. \tag{3.21}$$

Proof. By the weak convergence of Theorem 3.3 it follows that

$$x^\varepsilon(u^\delta, \cdot) \to x(u^\delta, \cdot) \text{ and } J^\varepsilon(u^\delta) \to J(u^\delta) \text{ as } \varepsilon \to 0 \tag{3.22}$$

Since u^δ is a δ-optimal control

$$\inf_{m \in \mathcal{R}} J(m) + \delta \geq J(u^\delta) \tag{3.23}$$

By virtue of Theorem 3.3 there exist $m^\varepsilon \in \mathcal{R}$ such that

$$\inf_{m \in \mathcal{R}^\varepsilon} J^\varepsilon(m) + \Delta_\varepsilon \geq J^\varepsilon(m^\varepsilon) \tag{3.24}$$

and

$$J^\varepsilon(u^\delta) = J(u^\delta) + \Delta_1(\varepsilon) \tag{3.25}$$

for some $\Delta_1(\varepsilon) \to 0$ as $\varepsilon \to 0$. Combining (3.23)-(3.25), we have

$$J^\varepsilon(u^\delta) - \inf_{m \in \mathcal{R}^\varepsilon} J^\varepsilon(m) \leq \Delta_\varepsilon - \Delta_1(\varepsilon) + \delta. \tag{3.26}$$

Letting $\varepsilon \to 0$ in (3.26). The proof is complete.

CHAPTER 4

Numerical Experiments on Van Der Pol Oscillator

4.1 Introduction

This chapter is concerned with the Van der Pol oscillator. This chapter is organized as follows. Section 4.2 contains the problem formulation. Section 4.3 presents some numerical experimental results. Finally, Section 4.4 makes further remarks.

4.2 Problem Formulation

The traditional Van der Pol oscillator is a non-conservative oscillator with nonlinear damping. The differential equation is

$$\ddot{y} - \mu(1 - y^2)\dot{y} + y = 0 \tag{4.1}$$

where y is real valued and μ is a positive constant. This model was originally proposed by the Dutch electrical engineer and physicist Balthasar Van der Pol in 1920. The Van der Pol oscillator has been used in both physical and biological sciences. Some noticeable properties include that it has decreasing oscillating solutions for $y^2 < 1$ and exponentially growing solution for $y^2 > 1$.

For the problem that we are interested in, we assume that it is also subject to random perturbations. We write the problem as a first-order system. Then the Van der Pol oscillator can be written in its two dimensional form. Set $x = (x_1, x_2) = (y, \dot{y})$, therefore (4.1) becomes

$$\dot{x}(t) = b(x(t)) \text{ where } b(x) = \begin{pmatrix} x_2 \\ \\ -\mu\left(x_1^2 - 1\right)x_2 - x_1 \end{pmatrix}. \tag{4.2}$$

We consider the Van der Pol equation driven by a white noise of constant intensity σ^2 of the following Itô equation

$$\begin{aligned} dx(t) &= b(x(t))\,dt + \lambda\sigma(x(t))\,dw(t) \\ x(0) &= x_0 \end{aligned} \tag{4.3}$$

$\lambda > 0$ is a parameter. We are interested in the asymptotic behavior as $\lambda \downarrow 0$.

4.3 MATLAB Simulation for Van Der Pol Oscillator

In this section we present the numerical experiments of Van der Pol equations. First, for small value $\mu = 0.1$. We test this model for a variety of λ to observes its behaviors. We obtain the following simulation results for the same initial conditions with $x(0) = 0, y(0) = 0.5$. The values of λ simulate are $0, 0.0001, 0.1$ and 1.

We note that when x is plotted against time t, its observed that the shape of the signal becomes less sinusoidal as μ increased and more fuzzy as λ increased. Next, for $\mu = 1$ we test the effect of different values of λ on the shape of the limit circle for the same initial condition with $x(0) = 0, y(0) = 0.5$.

We can see that when y is plotted against x and μ is small, the limit cycle is close to a sinusoidal oscillation with small λ. Moreover, we can see that the limit cycles become fuzzy as λ is increased from $0.0001, 0.1$ and 1. Fig. 1 - Fig. 10 provide some simulation results.

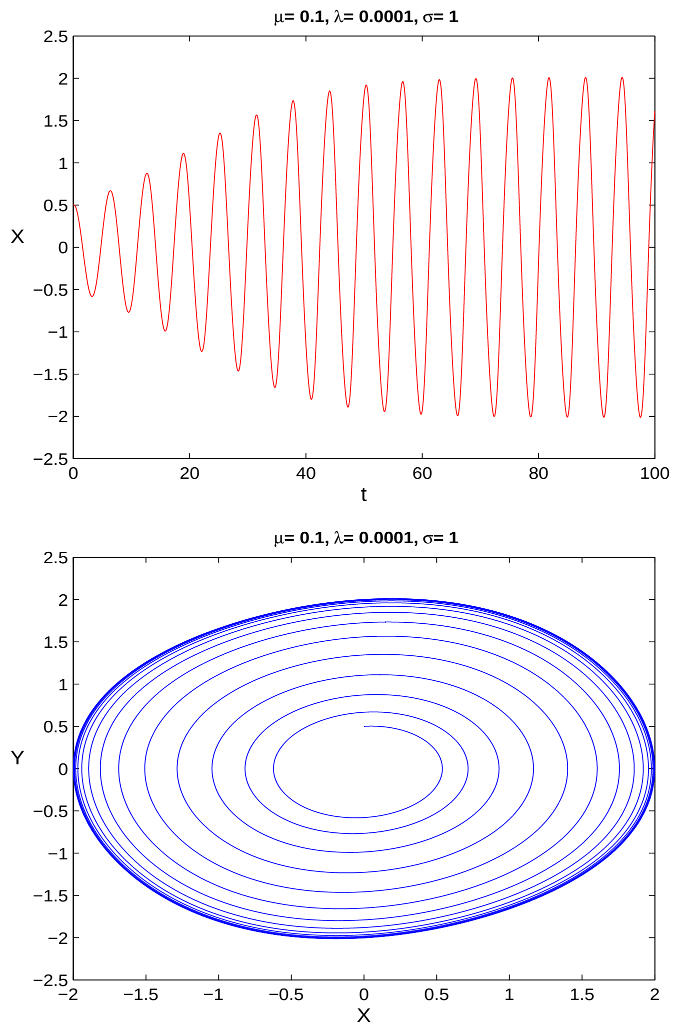

Figure 1: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 0.1, \lambda = 0.0001$

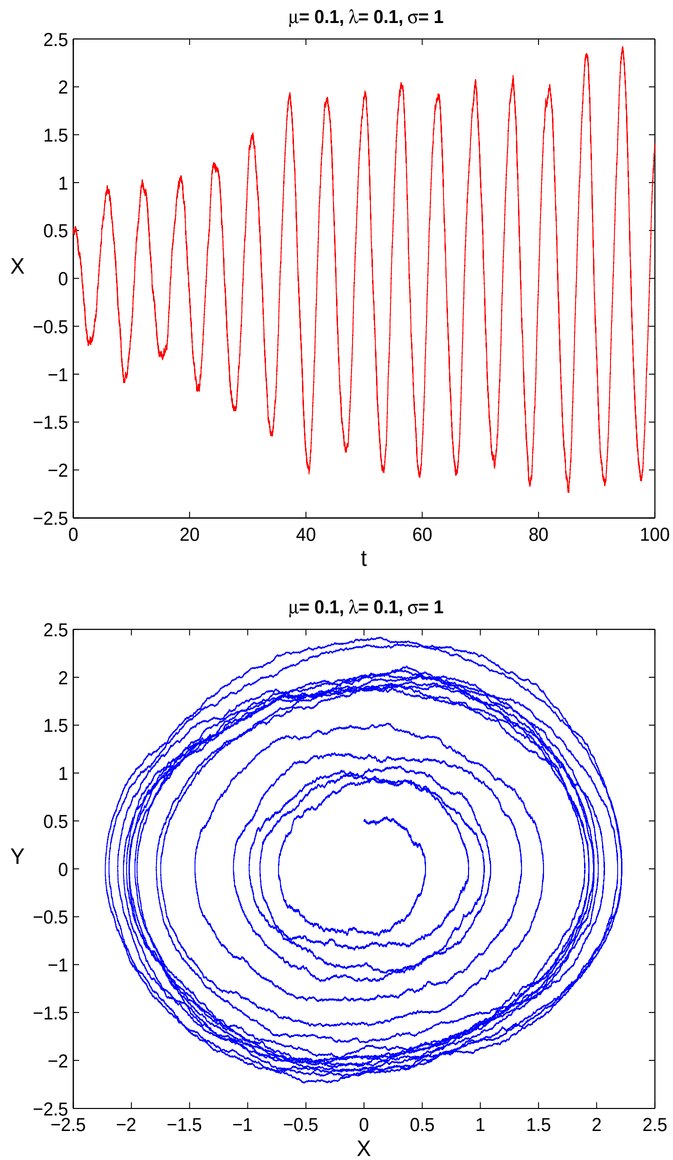

Figure 2: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 0.1, \lambda = 0.1$

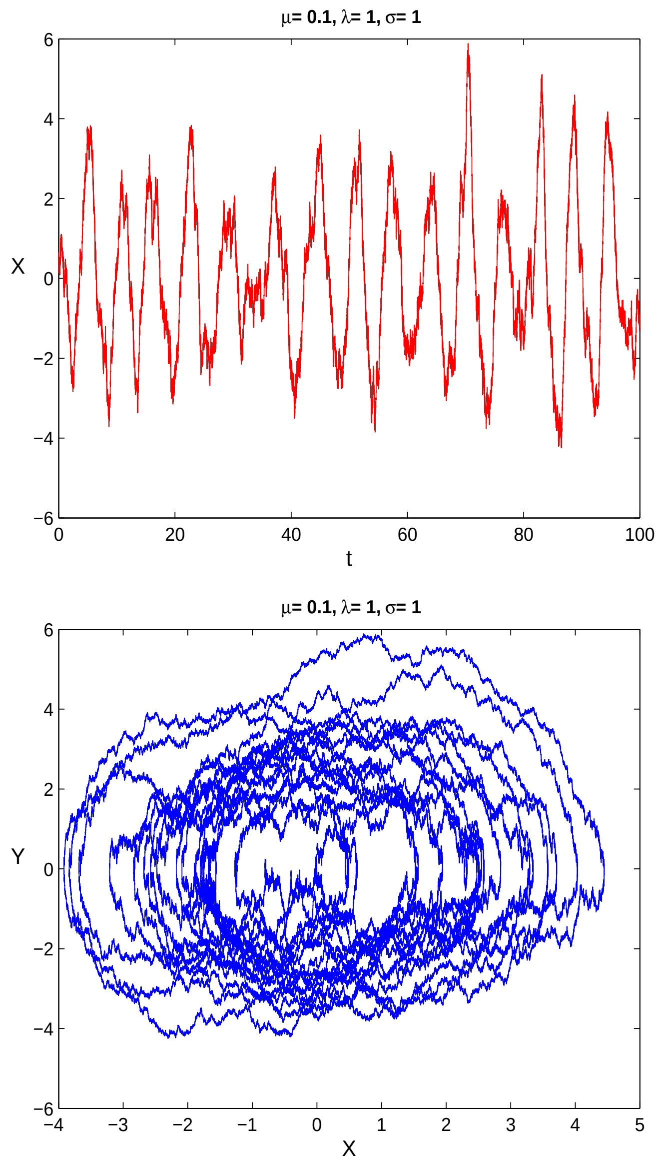

Figure 3: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 0.1, \lambda = 1$

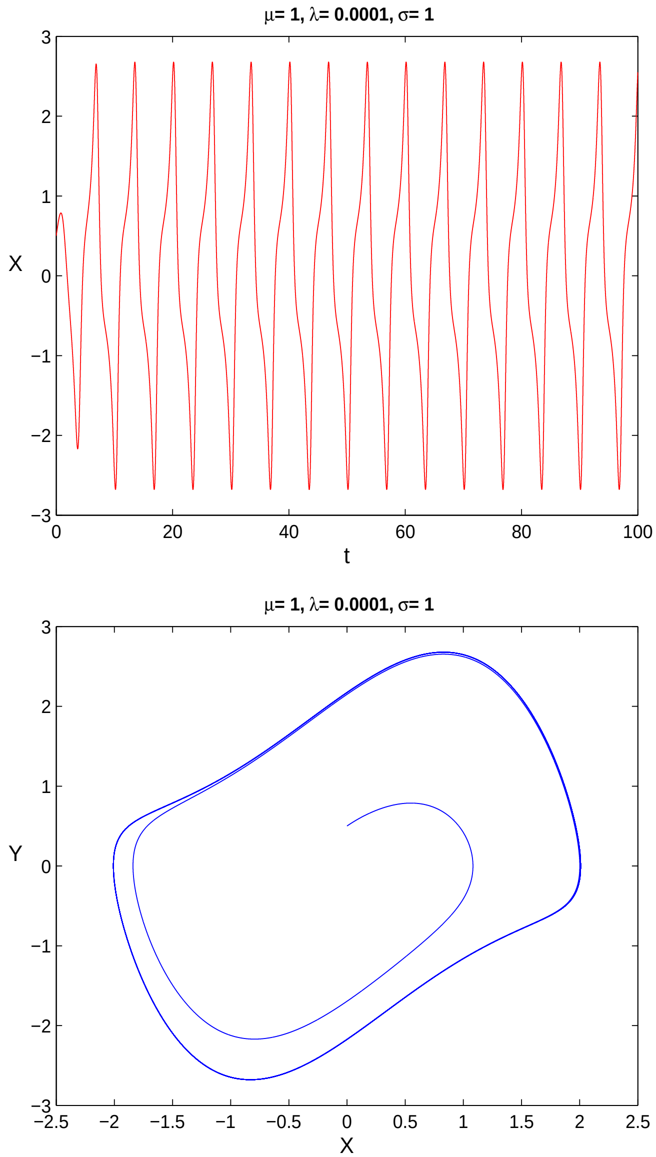

Figure 4: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 1, \lambda = 0.0001$

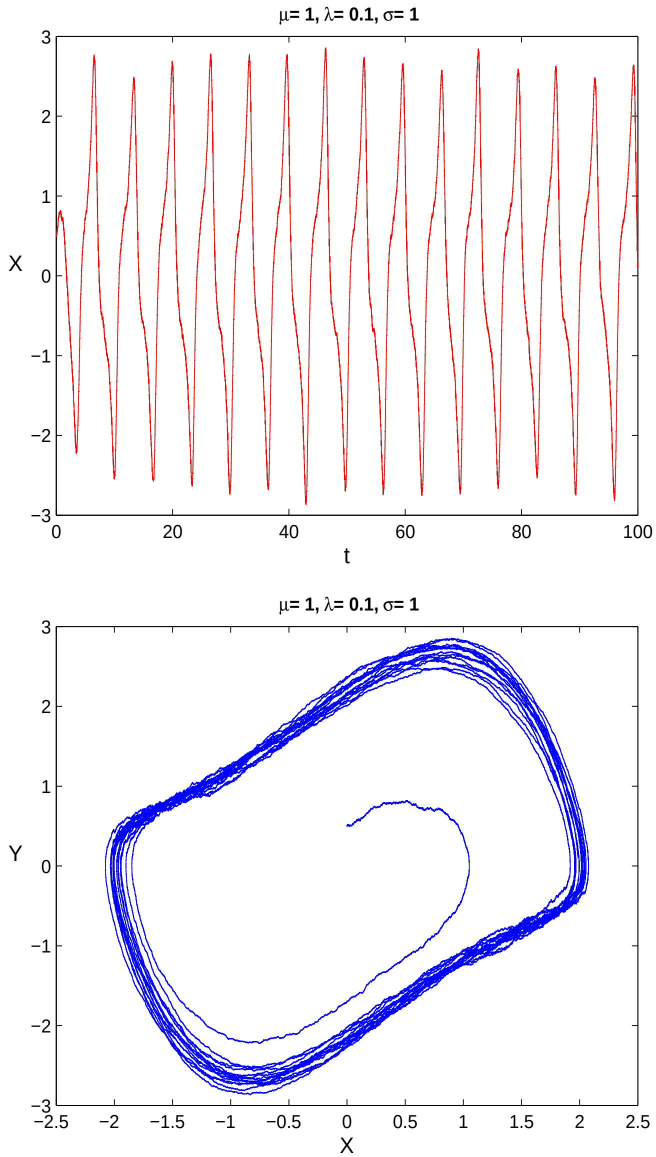

Figure 5: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 1, \lambda = 0.1$

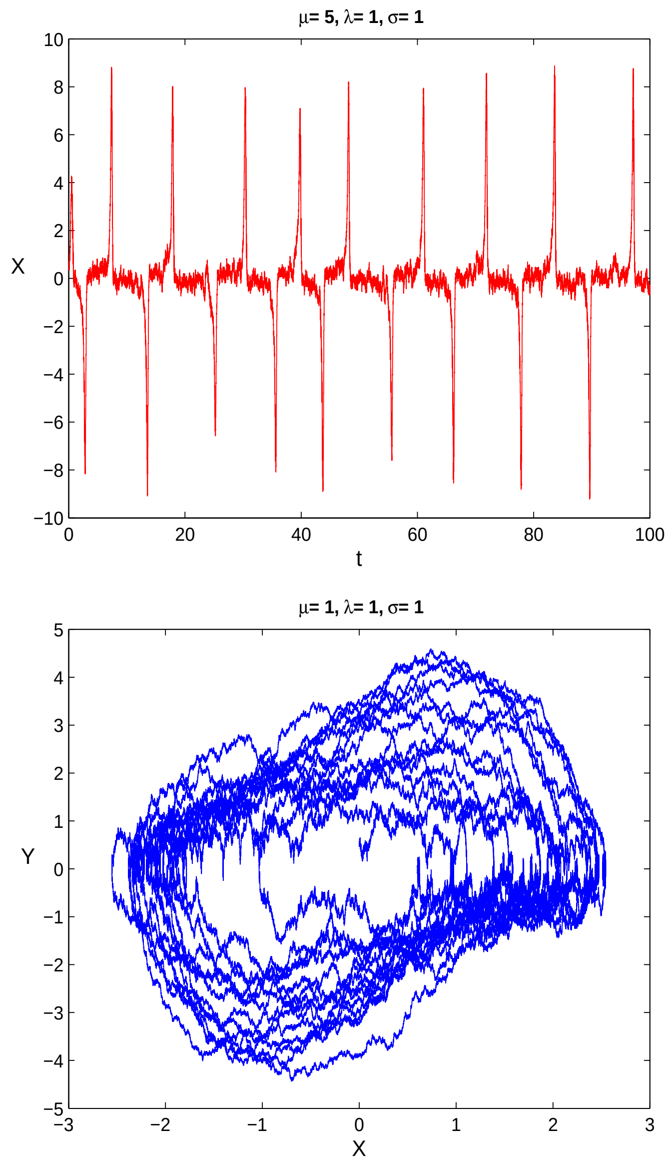

Figure 6: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 1, \lambda = 1$

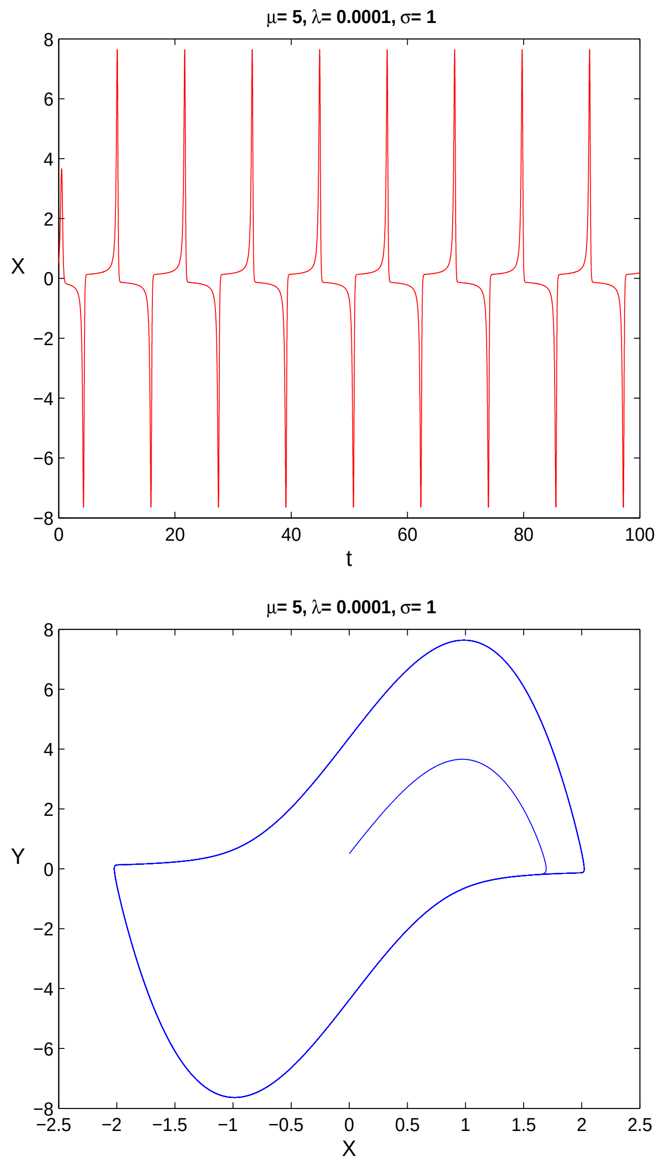

Figure 7: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 5, \lambda = 0.0001$

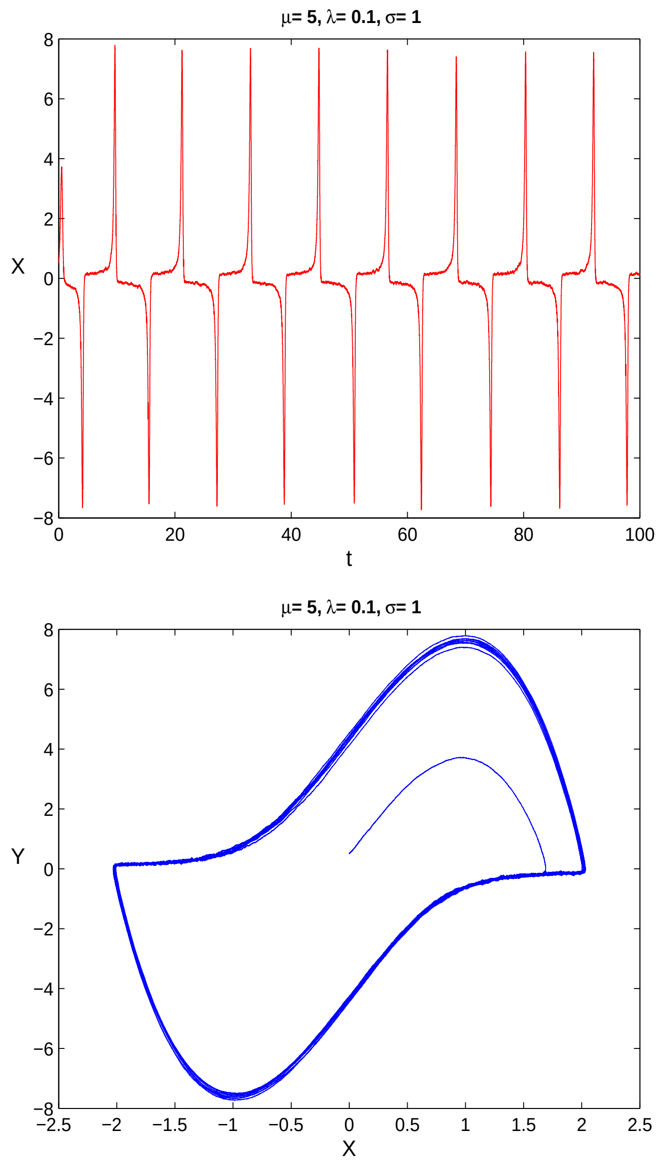

Figure 8: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 5, \lambda = 0.1$

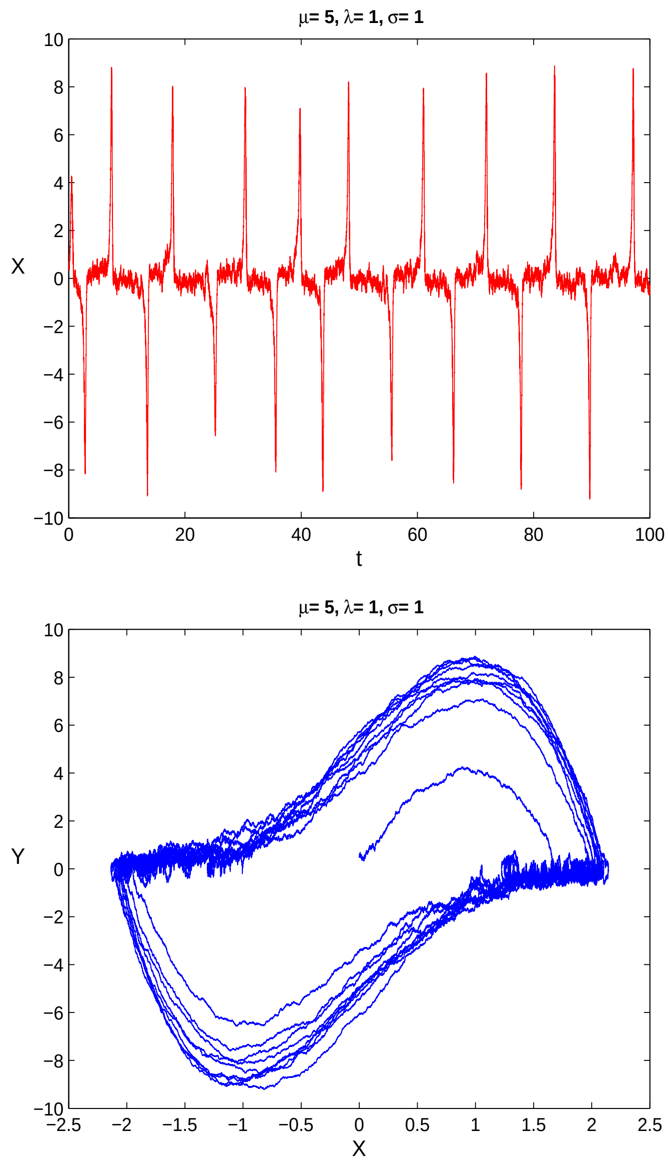

Figure 9: MATLAB Simulation for Van Der Pol Oscillator: $\mu = 5, \lambda = 1$

4.4 Remarks

The numerical examples considered in the last section can be thought of as random perturbations of deterministic dynamic systems. The solution of the stochastic differential equation may be written as $x^\lambda(t)$. Using the weak convergence methods, it can be shown that as $\lambda \to 0$, $x^\lambda(\cdot)$ converges weakly to $x(\cdot)$ such that $x(\cdot)$ is the solution of the deterministic differential equation

$$\dot{x}(t) = b(x(t)), \quad x(0) = x_0.$$

That is, averaging principle holds. It can further be demonstrated that

$$\lim_{\lambda\to 0} P(\sup_{0\le t\le T} |x^\lambda(t)) - x(t)| > \delta) = 0$$

for any $T > 0$ and $\delta > 0$. It can also be shown that as $\lambda \to 0$, $(x^\lambda(t) - x(t))/\sqrt{\lambda}$ converges weakly to a diffusion process. Furthermore, one may use the methods of Freidlin and Wentzel [10] to show that for any $T > 0$, $P(|x^\lambda(t) - x(t)| > \delta)$ is exponentially small.

Recently, there have been much effort in studying regime-switching dynamic systems.The basic premise is that the underlying system has both continuous dynamics and discrete event in which the discrete events cannot be modeled by the usual notion of differential equation. The switching process, for example, is a continuous-time Markov chain with a time-varying generator and state space $\mathcal{M} = \{1, \dots, m\}$. Associated with (4.2), we may consider a model

$$dx^\lambda(t) = b(x^\lambda(t), \alpha^\Delta(t), t)dt + \lambda\sigma(x^\lambda(t), \alpha^\Delta(t), t)dw(t),\ x^\lambda(0) = x_0, \tag{4.4}$$

where $\Delta \to 0$ as $\lambda \to 0$, $\alpha^\Delta(t)$ is a continuous-time Markov chain with generator $Q(t)/\Delta$, and $Q(t)$ is an irreducible generator. Note that in the above, the b and σ are allowed to depend on t as well. In [11] it was shown that $x^\lambda(\cdot)$ converges weakly to $x(\cdot)$ as $\lambda \to 0$ such that the limit is given by

$$\dot{x}(t) = \bar{b}(x(t), t),\ x(0) = x_0$$

where

$$\bar{b}(x,t) = \sum_{i=1}^{m} \nu_i(t) b(x,i,t)$$

and $\nu_i(t)$ is the quasi-stationary distribution associated with $Q(t)$. Moreover, three different cases

$$\lim_{\lambda\to 0} \frac{\Delta}{\lambda} = \begin{cases} \text{constant } \in (0,\infty), & \text{case 1} \\ \infty, & \text{case 2} \\ 0, & \text{case 3} \end{cases}$$

were treated. We describe one case below, namely, $\Delta = \lambda$.

Suppose that for each $i \in \mathcal{M}$, $b(\cdot, i, t)$ grows at most linearly in x and is Lipschitzian in x,

$$|b(x,i,t)| \leq K(1+|x|) \; \forall x \in \mathbb{R}^2, i \in \mathcal{M}$$

and that $\sigma(\cdot, i, t)$ is bounded and Lipschitz continuous. Suppose also $Q(t)$ is irreducible. Then it was proved in [11] that there exists a function $H(\cdot,\cdot,\cdot) : [0,\infty] \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ satisfying

$$\lim_{\lambda\to 0} \lambda \log E \exp\left(\frac{1}{\lambda} \int_0^T \langle b(x, \alpha^\lambda(s), s) ds + \frac{1}{2} \int_0^T a(x, \alpha^\lambda(s), s)\beta, \beta \rangle ds \right)$$

$$= \int_0^T H(x, \beta, s) ds.$$

Suppose that $\beta(t)$ is a step function on $[0,T]$. Then

$$\lim_{\lambda\to 0}\lambda\log E\exp\left\{\lambda^{-1}\int_0^T\big\langle b(x,\alpha^\lambda(s),s)ds+\frac{1}{2}\int_0^T a(x,\alpha^\lambda(s),s)\beta(s),\beta(s)\big\rangle ds\right\}$$

$$=\int_0^T H(x,\beta(s),s)ds,$$

where $a(x,i,t)=\sigma(x,i,t)\sigma'(x,i,t)$. Moreover, as $\lambda\to 0$, for any $s,\delta,h>0$, $\phi\in C([0,T],\mathbb{R}^2)$, and $\phi(0)=x$,

$$P\{\rho_{0T}(x^\lambda,\phi)<\delta\}\geq\exp\{-\frac{1}{\lambda}(S_T(\phi)+h)\},$$

$$P\{\rho_{0T}(x^\lambda,\Phi_x(s))>\delta\}\leq\exp\{-\frac{1}{\lambda}(s-h)\},$$

where $C([0,T],\mathbb{R}^2)$ denotes the space of continuous functions defined on $[0,T]$ taking values in $\mathbb{R}^2$, $\rho_{0T}(\cdot)$ is the distance function

$$\rho_{0T}(x,y)=\sup_{0\leq t\leq T}|x(t)-y(t)|, \tag{4.5}$$

$$\Phi_x(s)=\{\phi\in C([0,T],\mathbb{R}^2):\phi(0)=x,S_T(\phi)\leq s\},$$

and

$$S_T(\phi)=\begin{cases}\displaystyle\int_0^T L(\phi(s),\dot{\phi}(s),s)ds, & \text{if }\ \phi\in C([0,T],\mathbb{R}^2)\text{ is absolutely continuous,}\\[2ex] \infty\ \ \text{otherwise,}\end{cases}$$

$$L(x,\gamma,s)=\sup_{\beta\in\mathbb{R}^2}[\langle\gamma,\beta\rangle-H(x,\beta,s)].$$

In fact, more general case with $x(t)\in\mathbb{R}^k$ was considered in [11].

We next present the simulation result for this case.

Example 4.1. *Consider the stochastic Van der Pol equation switching diffusion* $(x(t), z(t))$ *with state* $\mathcal{M} = \{1, 2\}$ *and* $\lambda = \Delta = 0.0001$ *as follows*

$$\begin{aligned} dx_1(t) &= x_2(t)dt \\ dx_2(t) &= -\big(\alpha(z(t))x_2(t)(x_1^2(t) - 1) + \beta(z(t))x_2(t)\big)\, dt + \lambda\sigma dw \end{aligned}$$

where w *is a one-dimensional standard Brownian motion,* $\alpha(1) = 1$*,* $\alpha(2) = 2$*,* $\beta(1) = 1$ *and* $\beta(2) = 2$ *with*

$$Q(t) = \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix}.$$

Fig. 10 and Fig. 11 provide some simulation results.

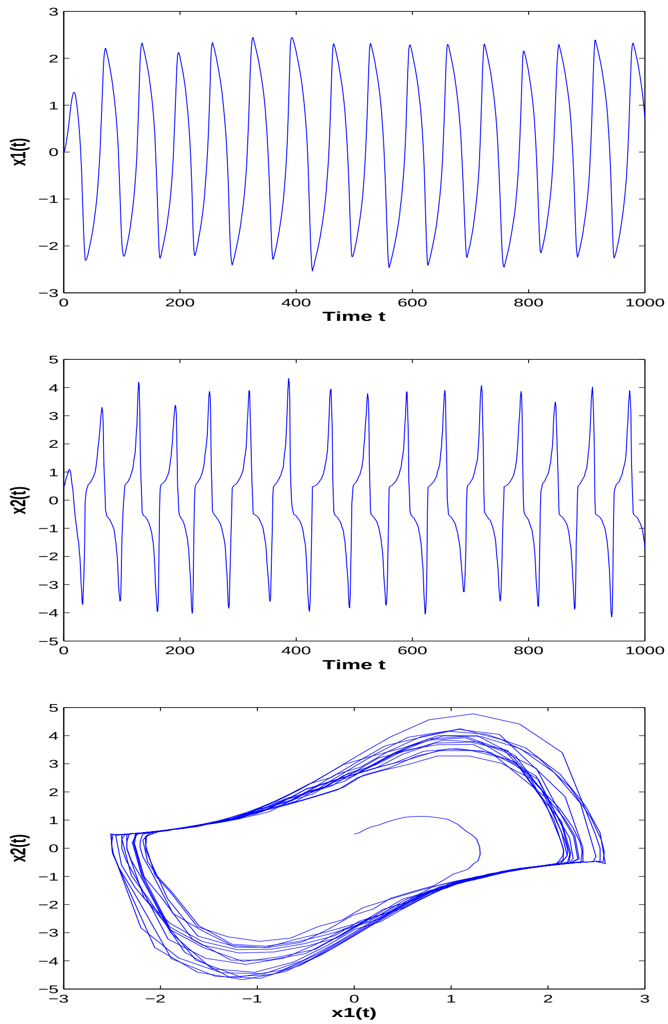

Figure 10: MATLAB Simulation for Van Der Pol Oscillator

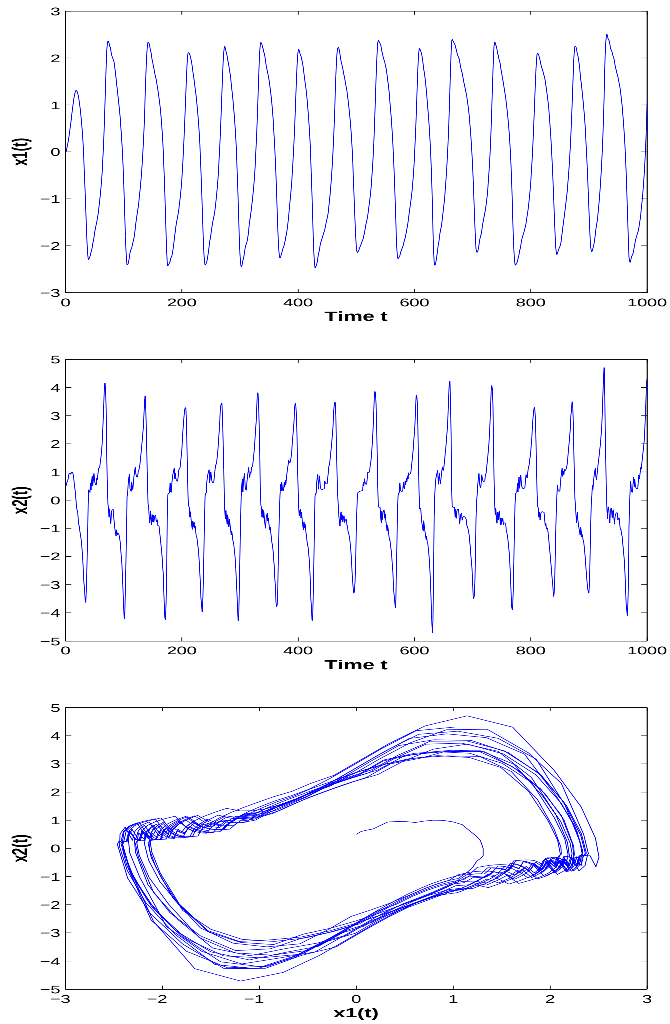

Figure 11: MATLAB Simulation for Van Der Pol Oscillator

APPENDIX A

Weak Convergence

In this section, we present a number of definitions of weak convergence including tightness, Prohorov's theorem, martingale problem and Skorohod representation.

Definition 1.1. (Weak Convergence). *Let P and P_k, $k = 1; 2; ....$, be probability measures defined on a metric space $\mathbb{S}$. The sequence P_k converges weakly to P if*

$$\int f dP_k \to \int f dP$$

for every bounded and continuous function $f(\cdot)$ on $\mathbb{S}$. Suppose that X_k and X are random variables associated with P_k and P, respectively. The sequence X_k converges to X weakly if for any bounded and continuous function $f(\cdot)$ on $\mathbb{S}$, $Ef(X_k) \to Ef(X)$ as $k \to 1$.

Definition 1.2. *Let $D([0,\infty), \mathbb{R}^r)$ be the set of all right continuous functions with left hand limits on $[0, \infty)$,i.e*

$$x : [0, \infty) \to \mathbb{R}^r, \ \lim_{s \downarrow t} x(s) = x(t) \ \& \ \lim_{s \uparrow t} x(s) = x(t-) \ \textit{exist} \ , \ \textit{for all } t \geqslant 0$$

Definition 1.3. *Let $\mathbb{L}'$ be the collection of strictly increasing functions $\lambda : [0, \infty) \to [0, \infty)$ such that the map is onto with $\lambda(0) = 0$, $\lim_{t \to \infty} \lambda(t) = \infty$ and λ is continuous. Let $\mathbb{L}$ be the set of Lipschitz continuous functions $\lambda \in \mathbb{L}'$ such that*

$$\gamma(\lambda) = \sup_{0 \leq s < t} \left| \log \left(\frac{\lambda(t) - \lambda(s)}{t - s} \right) \right| < 0$$

Definition 1.4. (Skorohod Topology). *For ζ, $\eta \in D([0; \infty); \mathbb{R}^r)$, the Skorohod topology $d(\cdot; \cdot)$ on $D([0; \infty); \mathbb{R}^r)$ is defined as*

$$d(\zeta,\eta)=\inf_{\lambda\in\mathbb{L}}\Big\{\gamma(\lambda)\vee\int_0^\infty e^{-\rho}\sup_{t\geq 0}\Big(1\wedge|\zeta(t\wedge\rho)-\eta(\lambda(t)\wedge\rho)|\Big)d\rho\Big\}$$

Definition 1.5. (Tightness). *A sequence of probability measures $\mathbb{P}_n$ on metric space $\mathbb{S}$ is tight if for every $\delta>0$ there exist a compact $K\subseteq\mathbb{S}$ and n_0 such that*

$$\mathbb{P}_n(K)>1-\delta \textit{ for all } n>n_0$$

Theorem 1.6. (Prohorov's Theorem) *Suppose that $\mathbb{P}_n$ is tight. Then it contains a weakly convergent subsequence $P_{nk}\Rightarrow P$.*

Theorem 1.7. (The Skorohod representation (Ethier and Kurtz [6])). *Let X_k and X be random elements belonging to $D([0,\infty);\mathbb{R}^r)$ such that X_k converges weakly to X. Then there exists a probability space $(\tilde{\Omega},\tilde{\mathbb{F}},\tilde{P})$ on which are defined random elements $\tilde{X}_k$, $k=1,2,...,$ and $\tilde{X}$ such that for any Borel set B and all $k<1$, $\tilde{P}(\tilde{X}_k\in B)=P_n(B)$, and $\tilde{P}(\tilde{X}\in B)=P(B)$ such that*

$$\lim_{n\to\infty}\tilde{X}_k=\tilde{X} \textit{ w.p.1.}$$

Definition 1.8. *Let $\mathbb{S}$ be a metric space and $\mathcal{A}$ be a linear operator on $B(\mathbb{S})$ (the set of all Borel measurable functions defined on $\mathbb{S}$). Let $X(\cdot)=\{X(t):t\geq 0\}$ be a right-continuous process with values in $\mathbb{S}$ such that for each $f(\cdot)$ in the domain of $\mathcal{A}$:*

$$f(X(t))-\textstyle\int_0^t\mathcal{A}f(X(s))ds$$

is a martingale with respect to the filtration $\sigma\{X((s):s\leq t\}$. Then $X(\cdot)$ is called a solution of the martingale problem with operator $\mathcal{A}$.

Theorem 1.9. *(Ethier and Kurtz [6, p. 174]). A right-continuous process $X(t), t \geq 0$, is a solution of the martingale problem for the operator $\mathcal{A}$ if and only if*

$$E\big[\prod_{j=1}^{i} h_j(X(t_j))\big[f(X(t_{i+1})) - f(X(t_i)) - \int_{t_i}^{t_{i+1}} \mathcal{A}f(X(s))ds\big]\big] = 0$$

whenever $0 \leq t_1 < t_2 < ... < t_{i+1}$, $f(\cdot)$ in the domain of $\mathcal{A}$, and $h_1, ..., h_i \in \mathcal{B}(\mathbb{S})$, the Borel field of $\mathbb{S}$.

Theorem 1.10. (Uniqueness of Martingale Problems, Ethier and Kurtz [6, p. 184]). *Let $X(\cdot)$ and $Y(\cdot)$ be two stochastic processes whose paths are in $D([0;T];\mathbb{R}^r)$. Denote an infinitesimal generator by $\mathcal{A}$. If for any function $f \in D(\mathcal{A})$ (the domain of $\mathcal{A}$),*

$$f(X(t)) - f(X(0)) - \textstyle\int_0^t \mathcal{A}f(X(s))ds, \ t \geq 0,$$

and

$$f(Y(t)) - f(Y(0)) - \textstyle\int_0^t \mathcal{A}f(Y(s))ds, \ t \geq 0,$$

are martingales and $X(t)$ and $Y(t)$ have the same distribution for each $t \geq 0$, $X(\cdot)$ and $Y(\cdot)$ have the same distribution on $D([0;\infty);\mathbb{R}^r)$.

REFERENCES

[1] V.V. Anisimov, *Switching processes: Averaging principle, diffusion approximation and applications*, *Acta Appl. Math.*, **40** (1995), 95–141.

[2] A. Bensoussan, *Perturbation Method in Optimal Control*, *J. Wiley, Chichester*, 1988.

[3] M.L. Cartwright M.L., *Balthazar Van der Pol*, *J. London Math. Soc.*, **35**, 367-376, (1960).

[4] J. Cartwright, V. Eguiluz, E. Hernandez-Garcia, and O. Piro, *Dynamics of elastic excitable media*, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, **9**, pp. 21972202, (1999)

[5] P.J. Courtois, *Decomposability: Queueing and Computer System Applications*, Academic Press, New York, NY, 1977.

[6] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, J. Wiley, New York, NY, 1986.

[7] R. Fitzhugh, *Impulses and Physiological in Theoretical Models of Nerve Membranes*, *Biophysics J*, **1**, 445-466, (1961).

[8] W.H. Fleming, *Generalized solution in optimal stochastic control* in Proc. *URI Conf. on Control*, 147-165, Kingston, Ri, 1982.

[9] W.H. Fleming and M. Nisio, *On stochastic relaxed controls for partially observable diffusions*, *Nagoya Math. J.* **93** (1984), 71–108.

[10] M.I. Freidlin and A.D. Wentzel, *Random Perturbations of Dynamical Systems*, 3nd Ed., Springer, Heidelberg, 2012.

[11] Q. He and G. Yin, Large deviations for multi-scale Markovian switching systems with a small diffusion, to appear in *Asymptotic Analysis*.

[12] R.Z. Khasminskii and G. Yin, *Asymptotic Series for Singularly Perturbed Kolmogorov–Fokker–Planck Equations*, *SIAM J. Appl. Math.*, **56** (1996), 1766–1793.

[13] R.Z. Khasminskii and G. Yin, *On averaging principles: An asymptotic expansion approach*, *SIAM J. Math. Anal.*, **35** (2004), 1534–1560.

[14] R.Z. Khashinskii, G. Yin, and Q. Zhang, *Asymptotic Expansions of Singularly Perturbed Systems Involving Rapidly Fluctuating Markov Chains*, *SIAM J. Appl. Math.*, **56** (1996), 277–293.

[15] P. Kokotovic, *Application of Singular Perturbation Techniques to Control Problems*, *SIAM Rev.***26** (1984), 501–550.

[16] P. Kokotovic, A. Bensoussan and G. Blankenship, EDs., *Singular Perturbations and Asymptotic Analysis in Control Systems*, Vol. 90, Lect. *Notes in Control and Information Science*, Springer, Berlin, 1987.

[17] P. Kokotovic, H. Khalil and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*, *Academic Press*, New York, 1096.

[18] P. V. Kokotovic and H. K. Khalil, *Singular Perturbations in Systems and Control*, *IEEE Press Selected Reprint Series. IEEE Press*, New York, 1986.

[19] V.S. Korolyuk and N. Limnios, *Diffusion approximation with equilibrium of evolutionary systems switched by semi-Markov processes*,translation in *Ukrainian Math. J.* **57** (2005), 1466–1476.

[20] V.S. Korolyuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Sci., Hackensack, NJ, 2005.

[21] H.J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.

[22] H.J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, MA, 1990.

[23] H.J. Kushner and W. Runggaldier, *Nearly optimal state feedback controls for stochastic systems with wideband noise disturbances*, *SIAM J. Control. Optim.* **25** (1987), 289–315.

[24] H.J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed., Springer-Verlag, New York, NY, 2003.

[25] J. Nagumo, S. Arimoto and S. Yoshizawa, *An Active Pulse Transmission Line Simulating Nerve Axon*, *Proc. IRE*, **50**, 2061-2070, (1962).

[26] S.P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, MA, 1994.

[27] H.A. Simon and A. Ando, *Aggregation of variables in dynamic systems*, Econometrica, **29** (1961), 111–138.

[28] A.V. Skorohod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Amer. Math. Soc., Providence, RI, 1989.

[29] T. J. Slight, B. Romeira, L. Wang, J.M.L Figueiredo, J.M.L.,E. Wasige, and C.N. Ironside, *A Linard Oscillator Resonant Tunnelling Diode-Laser Diode Hybrid Integrated Circuit: Model and Experiment*, *IEEE J. Quantum Electron.*, Vol. 44, No. **12**,pp. 1158-1163, December 2008.

[30] B. Van Der Pol, *On relaxation oscillations.*, *Philos Mag*, 1926;2:978.

[31] B. Van der Pol and J. Van der Mark, *Frequency demultiplication*, *Nature*, **120**, 363-364, (1927).

[32] B. Van Der Pol B, J. Van Der Mark, *The Heartbeat Considered as a Relaxation Oscillator and an Electrical Model of the Heart.*,*Philos Mag* 1928;**6**(Suppl.):763.

[33] J. Warga, *Relaxed variational problems*,*J. Math. Anal. Appl.* **4** (1962) 111–128.

[34] F.B. Xi and G. Yin, *Stochastic Liénard equations with state-dependent switching*, to appear in *Acta Math. Appl. Sinica*.

[35] F.B. Xi and L.Q. Zhao, *Stationary solution for a stochastic Liénard equation with Markovian switching*,*Acta Math. Sci. Ser. B*, **25** (2005), 95–104.

[36] G. Yin, H.Q. Zhang, and Q. Zhang, *Applications of Two-time-scale Markovian Systems*, Science Press, Beijing, China, 2013.

[37] G. Yin and Q. Zhang, *Continuous-time Markov Chains and Applications: A Singular Perturbations Approach*, Springer-Verlag, New York, NY, 1998.

[38] G. Yin, Q. Zhang, and G. Badowski, *Asymptotic properties of a singularly perturbed Markov chain with inclusion of transient states*, Ann. Appl. Probab., **10** (2000), 549–572.

[39] G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer, New York, 2010.

[40] L.Q. Zhao and F.B. Xi, *Explicit conditions for asymptotic stability of stochastic Liénard-type equations with Markovian switching*, J. Math. Anal. Appl., **348** (2008), 267–273.

ABSTRACT

TWO-TIME-SCALE SYSTEMS IN CONTINUOUS TIME WITH REGIME SWITCHING AND THEIR APPLICATIONS

by

YOUSEF TALAFHA

December 2013

Advisor: Dr. Gang George Yin

Major: Mathematics

Degree: Doctor of Philosophy

This dissertation focuses on two-time-scale stochastic systems represented by switching diffusions. In the model, a continuous-time Markov chain serves as a modulating force that enables the system to switch among a finite number of diffusion processes. A two-time-scale formulation is used to reduce the computational complexity. Liénard equations are examined. Then near-optimal controls of switching diffusions are treated. Then near-optimal controls for stochastic differential equation with regime switching. In addition, numerical experiments are performed for a class of Van der Pol equations.

The motivation of our study stems from modeling of complex systems in which both continuous dynamics and discrete events are present. In Chapter 2, the continuous component is a solution of a stochastic Liénard equation and the discrete component is a Markov chain, whereas in Chapter 3, the continuous component is a controlled diffusion and the discrete component is a Markov chain. In both cases, the Markov chains have a large state space. A distinct feature is that the processes under consideration are time inhomogeneous. Based on the idea of nearly decomposability and aggregation, the state space of the switching

process can be viewed as "nearly decomposable" into l subspaces that are connected with weak interactions among the subspaces. Using the idea of aggregation, we lump the states in each subspace into a single state. Considering the pair of process (continuous state, discrete state), under suitable conditions, we derive a weak convergence result by means of martingale problem formulation. The significance of the limit process is that it is substantially simpler than that of the original system. Thus, it can be used in the approximation and computation work to reduce the computational complexity. Finally, we investigate the system behavior of Van der Pol oscillator by introducing the noise. The system have been performed numerically and results are shown using Matlab. Simulations show that the proposed model gives limit cycles are more accurate as the noise decreased which the limit cycle is close to a sinusoidal oscillation and the shape of the signal becomes less sinusoidal as the noise increased.

AUTOBIOGRAPHICAL STATEMENT

YOUSEF TALAFHA

Education

- 2013, Ph.D. in Mathematics (*expected*)
  Wayne State University, Detroit, Michigan, USA.
- 2013, Master in Mathematical Statistics
  Wayne State University, Detroit, Michigan, USA.
- 2007, Master in Mathematics
  Al Al-Bayt University, Mafraq, Jordan.
- 2005, B.S. in Mathematics
  Al Al-Bayt University, Mafraq, Jordan.

Awards and Honors

- 2013 The Alfred L. Nelson Award for outstanding achievement in the Master Programme, Department of Mathematics, Wayne State University, April 2013.
- 2011-2012 Graduate Research Assistant, Department of Mathematics, Wayne State University.
- 2012 Graduate Student Professional Travel Award, Department of Mathematics, Wayne State University.
- 2009-2012 Scholarship from Tafila Technical University to get a Ph.D. Degree in Applied Mathematics from Wayne State University.
- 2009 Scholarship from New Mexico State University to get a Ph.D. Degree in Mathematics. (I declined the offer).
- 2003-2007 An Honor Certificate from Al Al-Bayt University, Jordan in B.Sc and M.Sc.

Papers and Conference Presentations

1. G. Yin, Xianping Guo, Yousef Talafha and Nicholas A. Baran, *Near-Optimal Controls of Differential Systems with Switching and Random Jumps Subject to Fast Switching and Wideband Noise Perturbation*. Submitted.
2. G. Yin, Yousef Talafha, and Fubao Xi, *Stochastic Liénard Equations with Random Switching and Two-time Scales*, to appear in Communications in Statistics - Theory and Methods.
3. Yin, G.; Talafha, Y.; *Mean-variance portfolio selection under regime-switching diffusion asset models: A two-time-scale limit, in Stochastic Processes, Finance and Control*, S.N. Cohen, D. Madan, T.K. Siu, and H. Yang Eds., World Scientific, 2012, New Jersey, 375–390.
4. Poster Presentation, Midwest Graduate Research Symposium, The University of Toledo, Ohio, April, 2013.
5. Talk in the Joint Mathematics Meetings AMS Special Session at the 2013 in SAN DIEGO, CA JANUARY 9-12, 2013.
6. Poster Presentation,Conference on Probability, Control and Finance in Honor of Ioannis Karatzas (Columbia University, June 4-8, 2012)