Asymptotic Properties and Finite Time Convergence of Classical and Modified Methods for Stochastic Differential Equations

Wei Liu

Department of Mathematics and Statistics

University of Strathclyde

Glasgow, UK

October 2013

This thesis is submitted to the University of Strathclyde for the degree of Doctor of Philosophy in the Faculty of Science.

Acknowledgements

Finally, it is the time to say goodbye to my PhD life. At this moment, I have a lot of words crowding in my throat and plenty of wonderful moments flashing in my mind. I wish I could give my appreciation to every single of you, and I could even write a book on it. However, as an acknowledge of a PhD thesis, I guess that readers must love to focus on mathematics and only give a as quick as possible glance of this page. So I project the unbounded set of thanks to the following a few words and hope that readers will enjoy rest of the thesis.

First of all, I would like thank my supervisor Professor Xuerong Mao, who introduces the topic to me, inspires me during my PhD study and encourages me to challenge myself.

My thanks also go to Dr Andrew Wade, Dr Alison Gray, Dr Vladislav Vyshemirsky and Dr Jiafeng Pan. The academic and nonacademic chats with them make my PhD life more colourful. I could not forget Martin, Alan, Steven, David, Eusebio and all the office chats with them.

For Hongrui, Wenwen and Chang, I need to say, without you I would have already died from loneliness.

I thank the university of Strathclyde and the Scottish government for their financial support, and Journal of Computational and Applied Mathematics and Applied Mathematics and Computation for accepting my papers.

Last but not the least, thank you mum and dad for the uncountable many reasons.

Abstract

As few stochastic differential equations have explicit solutions, the numerical schemes are studied to approximate the underlying solution. The fast development in computer science in recent years has made large scale simulations available, then the numerical analysis for stochastic differential equations has been blooming in past decades. However, the study on numerical solutions is still far behind the study on the underlying solutions. This thesis is devoted to mathematically rigorous investigation on the numerical solutions.

Among all those attractive mysteries in the numerical analysis of stochastic differential equations, one of the popular problems is that if the numerical solutions can reproduce different properties of the underlying solutions. In thesis, we present some interesting results on this topic, which includes the asymptotic moment boundedness, the stationary distribution and the almost sure stability. The methods considered in this part are two classical methods, the explicit Euler-Maruyama method and the backward Euler-Maruyama method, and one modified method, the Euler-Maruyama method with random variable step size, which is first introduced in this thesis. Another main focus of numerical analysis is the finite time convergence. Our work on this topic is to modify the explicit Euler-Maruyama method and investigate the strong convergence (in the L^2 sense) of it.

Our investigation first goes to reproduce the asymptotic boundedness in small moment of the underlying solutions. The explicit Euler-Maruyama method is shown to be able to achieve this goal if both the drift coefficient and the diffusion

coefficient are global Lipschitz. But with the global Lipschitz condition on the drift coefficient violated, a counter example indicates the failure of the explicit Euler-Maruyama method. A natural replacement, the backward Euler-Maruyama method, then is considered and successfully reproduce the asymptotic boundedness. In the case of small moment, we are only able to reproduce the boundedness property qualitatively so far. To answer another close related question that if we could reproduce the upper bound quantitatively, we strengthen the conditions and show that for the case of second moment the upper bound of the underlying solution can be reproduced as well.

As the moment boundedness is key to the existence and uniqueness of the stationary distribution, we next study this property for the numerical solution. Since the backward Euler-Maruyama method has better performance than the explicit Euler-Maruyama method, in this part we only discuss the backward Euler-Maruyama method. The coefficient related sufficient conditions are given for the existence and uniqueness of the stationary distribution of the backward Euler-Maruyama method. Then the numerical stationary distribution is proved to converge to the stationary distribution of the underlying solution as step size vanishes. These results largely extend the existing works to cover wider range of stochastic differential equations.

The almost sure stability is one of the hottest topics and many papers have studied the reproduction of this property by different kinds of classical methods. Therefore, we seek to study this property by one modified method, the Euler-Maruyama method with random variable step size. To our best knowledge, this is the first work to apply the random variable step size to the analysis of the almost sure stability of the explicit Euler-Maruyama method. One of our key contributions is that we show that the time variable is a stopping time, which were ignored by many researchers, and only under this circumstance the rest results hold. Compare with those fixed step size or nonrandom variable step size methods, the Euler-Maruyama method with random variable step size is shown to be able to reproduce the almost sure stability with much weaker conditions.

As the strong convergence of the classical methods has already been widely studied and the recent works have shown the good performance of the modified classical methods, we present our findings in this area by introducing the stopped Euler method and show the strong convergence of it to the underlying solution with the rate a half. Briefly, the stopped Euler method is the classical Euler-Maruyama method equipped with the stopping time technique. The stopping time is originally employed to preserve the non-negativity of the numerical solution, and it turns out that the non-negativity in return enables the strong convergence of the method with the rate arbitrarily close to a half. Compare with the explicit Euler-Maruyama method, the stopped Euler method can cover some highly non-linear stochastic differential equations.

Notation

a.s.	:	almost surely, or $\mathbb{P}$-almost surely, or with probability 1.
$A := B$	:	A is defined by B or A is denoted by B.
$A(x) \equiv B(x)$	:	$A(x)$ and $B(x)$ are identically equal, i.e. $A(x) = B(x)$ for all x.
$\emptyset$	:	the empty set.
$\mathbf{1}_A$	:	the indicator function of a set A, i.e. $\mathbf{1}_A(x) = 1$ if $x \in A$ or otherwise 0.
$\sigma(C)$	:	the σ-algebra generated by C.
$a \vee b$	:	the maximum of a and b.
$a \wedge b$	:	the minimum of a and b.
$\|x\|$	:	the Euclidean norm of a vector x.
M^T	:	the transpose of a vector or matrix M.
$\|M\|$	:	$= \sqrt{\text{trace}(M^T M)}$, i.e. the trace norm of a matrix M.
$\lambda_{\min}(M)$	:	the smallest eigenvalue of a matrix M.
$\lambda_{\max}(M)$	:	the largest eigenvalue of a matrix M.
$\lceil x \rceil$	:	the smallest integer larger than a real number x.
$\mathbb{R}_+$	:	the set of all nonnegative real numbers.
$\mathbb{N}$	:	the set of all nonnegative integers.
$\mathbb{Z}$	:	the set of all integers.
$\mathbb{Q}$	:	the set of all rational numbers.
$L^p(\Omega; \mathbb{R}^n)$	:	the family of $\mathbb{R}^n$-valued random variables θ with $\mathbb{E}\|\theta\|^p < \infty$.

$\mathcal{L}^p([a,b];\mathbb{R}^n)$: the family of $\mathbb{R}^n$-valued $\mathcal{F}_t$-adapted processes $\{f(t)\}_{a\le t\le b}$ such that $\int_a^b |f(t)|^p dt < \infty$.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

All our stories could date back to the year of 1827 when Scottish botanist Robert Brown observed the irregular motion of pollen grains in water through a microscope, but he was not able to determine the mechanisms that caused it. Decades after decades, no good explanation was given to this mystery. As time passed a new century, the paper of Albert Einstein published in 1905 (Einstein, 1905) finally came up with an explanation that the irregular motion of pollen grains was a result of being affected by individual water molecules. Since the number of water molecules is huge and the motion of them is so complicated, Einstein pointed out that one best way to describe the effects of water molecules on the pollen grain is statistical mechanics. In 1923, a mathematically rigorous description of Brownian motion was given by Norbert Wiener (Wiener, 1923). Since then, the study on stochastic differential modeling started to bloom. However, due to the unboundedness of the variation of Brownian motion, the Brownian sample path is nowhere differentiable and the Lebesgue-Stieltjes integral cannot be defied with respect to Brownian motion. But this is not the end of the world, as the quadratic variation of Brownian motion is finite. Motivated by this, the Japanese mathematician Kiyosi Itô firstly defined the stochastic integral (Itô, 1944; Itô, 1946) known as Itô integral. The corresponding differential equation is called stochastic differential equations of Itô type. In this thesis, all our studies focus on Itô stochastic differential equations (SDEs).

1.1 Motivation

After Itô's fundamental works, SDEs have become more and more important modeling tools in many disciplines, such as biology (Allen, 2007; Frank & Beek, 2001), physics (Schuss, 1980; Sobczyk, 2001), finance (Platen & Bruti-Liberati, 2010; Shreve, 2004) and chemistry (Gardiner, 2004; van Kampen, 1981), we just mention a few of them here.

Meanwhile, increasing attention has been paid on different properties of the solutions of SDEs. This is because, on one hand, One would like to know if the the dynamic of the SDE model is in line with the phenomenon it is trying to describe, for example if a population SDE model always has a nonnegative solution. One the other hand, the studies of SDEs could be standalone interest of mathematics.

Due to that the stochastic term plays an important role in the behaviour of the solution, the dynamic of the SDE could be so different from that of the counterpart ODE. For example, given an ODE whose solution is explosive as time tends to infinity, we can add a stochastic part of certain type on it and make the new SDE system stable to the trivial solution as time advances. This is called stochastic stabilization (Mao, 2008).

Many works have been done on the solutions of SDEs, such as existence and uniqueness, positivity, stability, boundedness, stationary distribution and hitting time, we just mention some of them here (Arnold, 1974; Khasminskii, 2012; Mao & Yuan, 2006; Mao, 2008; Oksendal, 2003) and the references therein.

Although many theories have been built on the existence and uniqueness of the solutions of SDEs, only a few SDEs can be solved explicitly. Therefore, in most practical problems we need to demonstrate the behaviour of the solutions numerically. Thanks to the fast development in the scientific computing technologies in recent years, the large scale computing simulations have become available.

Then one obviously interesting question is if the numerical solution approximates the underlying solution correctly. There already exist plenty of excellent works on numerical analysis for SDEs, which mainly focus on finite time (strong and weak) convergence, positivity, stability, numerical stationary distribution and boundedness. We mention some of the works here (Higham *et al.*, 2002; Higham *et al.*, 2007; Hu, 1996; Kloeden & Platen, 1992; Milstein, 1995; Milstein & Tretyakov, 2004) and the references therein. However, the study of numerical solutions is still far behind the study of underlying solutions. Due to the discontinuous nature of the numerical solutions, some essential techniques that works well on the underlying solutions can not be adapted to the numerical solution. Thus, alternatives need to be found and there are many gaps needing to be filled up in the study of numerical methods for SDEs.

This thesis contributes to finding the proper numerical methods to reproduce three asymptotic properties of the underlying solutions: asymptotic moment boundedness, stationary distribution and almost sure stability. We also do some works on the finite time strong convergence, in which the classical Euler-Maruyama method is modified by embedding with a stopping time and the modified method can cover larger range of SDEs.

1.2 Structure of This Thesis

In general, the next six chapters can be divided into three parts: Chapter 2 is devoted to the brief introductions to stochastic process, stochastic differential equations and related mathematical preliminaries which will be used in following chapters; Chapters 3, 4, 5 and 6 contain the main results of this thesis; Chapter 7 concludes this thesis and states some potential future research.

Chapter 3 studies one important asymptotic property of numerical solutions, the asymptotic boundedness. We first discuss a group of SDEs with the linear growth condition and reproduce the asymptotic boundedness in small moment using the Euler-Maruyama (EM) method. As the linear growth condition is violated, the EM method does not function well any more. Then we move to the backward Euler-Maruyama method and successfully reproduce the asymptotic boundedness in small moment under the one-sided Lipschitz condition. At last, by strengthening some conditions, we can reproduce not only the asymptotic boundedness of the underlying solution but also the upper bound accurately for the second moment.

In Chapter 4, we first investigate the existence of stationary distributions of numerical solutions. Then we study the convergence of the numerical stationary distribution to the stationary distribution of the underlying solution. The purpose of it is to avoid solving the nontrivial Kolmogorov-Fokker-Planck equations to find the stationary distribution of the underlying SDEs. In addition, as the Kolmogorov-Fokker-Planck equations can be regarded as connections between solutions of stochastic differential equations and deterministic differential equations, the numerical stationary distributions could also be used as numerical solutions to certain type of deterministic differential equations. It should be mentioned that the moment boundedness property studied and those techniques used in Chapter 3 are essential to the proofs of this chapter.

The step size of all the numerical methods discussed in Chapters 3 and 4 is constant. In Chapter 5, we introduce the random variable step size and embed it with the EM method to study the almost sure stability. Most previous works on the almost sure stability of numerical methods focused on the constant step size, and only a few of them looked at the nonrandom variable step size. To our best knowledge, this is the first work to apply the random variable step size to the analysis of the almost sure stability of the EM method. The key contribution of this chapter is that we prove the time variable is a stopping time and this observation is crucial to other proofs. As the payoff of using the EM method with random step size, the conditions for almost sure stability is largely released.

In Chapter 6, we continue to modify the classical EM method and study the finite time strong convergence of the modified method. As the classical EM has been proved to fail to work on some highly nonlinear SDEs, the classical EM method equipped with a stopping time is introduced in this chapter. The stopping time technique is employed to prevent the numerical solution from becoming negative. In return, this nonnegative numerical solution converges strongly (namely in the L^2 sense) to the underlying solution. In addition, we are able to prove the strong convergence rate is arbitrarily close to a half. Some Numerical simulations are conducted and the observations are in line with the theoretical results.

Each of those four main chapters has its only introduction section, in which more detailed literature review is given. The preliminary sections in those chapters specify the related notations, numerical methods, lemmas and theorems that are used in corresponding chapters.

Chapter 2

Mathematical Preliminaries

2.1 Random Variable, Stochastic Process and Martingale

For an experiment with uncertain outcomes, let Ω be the set of all the possible outcomes. The element of Ω is denoted by ω. Those subsets of Ω, that are of interest, are grouped together to form a family, $\mathcal{F}$, of subsets of Ω. A family $\mathcal{F}$ possessing the following three properties is called a σ-algebra:

- $\emptyset \in \mathcal{F}$,
- $S \in \mathcal{F} \Rightarrow S^C \in \mathcal{F}$,
- $\{S_i\}_{i\geq 1} \subset \mathcal{F} \Rightarrow \cup_{i=1}^{\infty} S_i \in \mathcal{F}$,

where $\emptyset$ denotes the empty set and S^C denotes the complement of S in Ω. Then the pair $(\Omega, \mathcal{F})$ is called a measurable space, and the element of $\mathcal{F}$ is called a $\mathcal{F}$-measurable set.

A probability measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a function: $\mathcal{F} \to [0,1]$ such that

- $\mathbb{P}(\Omega) = 1$,

- for any disjoint sequence $\{S_i\}_{i\geq 1} \subset \mathcal{F}$, $\mathbb{P}(\cup_{i=1}^{\infty} S_i) = \sum_{i=1}^{\infty} \mathbb{P}(S_i)$.

Then the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Define x to be a real-valued function on Ω, and it is said to be $\mathcal{F}$-measurable if

$$\{\omega : x(\omega) \leq c\} \in \mathcal{F} \quad \text{for all } c \in \mathbb{R}.$$

This function x is called a $\mathcal{F}$-measurable real-valued random variable as well. An $\mathbb{R}^n$-valued function, $x(\omega) = (x_1(\omega), x_2(\omega), ..., x_n(\omega))$, is said to be a $\mathcal{F}$-measurable $\mathbb{R}^n$-valued random variable if all the entries, $\{x_i(\omega)\}_{i=1,2,...,n}$, are $\mathcal{F}$-measurable real-valued random variables.

A filtration is a collection, $\{\mathcal{F}_t\}_{t\geq 0}$, of increasing sub-$\sigma$-algebra of $\mathcal{F}$, i.e. $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0 \leq s < t < \infty$. The filtration is said to be right continuous, if $\mathcal{F}_s = \cap_{t>s}\mathcal{F}_t$ for all $s \geq 0$.

Throughout this thesis, unless specified otherwise, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$, which is increasing and right continuous, with $\mathcal{F}_0$ containing all $\mathbb{P}$-null sets. We also define $\mathcal{F}_\infty$ as the σ-algebra generated by $\cup_{t\geq 0}\mathcal{F}_t$.

A collection of $\mathbb{R}^n$-valued random variables, $\{x_t\}_{t\geq 0}$, is called a stochastic process. The index t is considered to be time on $[0, \infty)$ for a convenient interpretation in this thesis. For each fixed $t \in [0, \infty)$, we have a random variable $x_t(\omega) \in \mathbb{R}^n$. On the other hand, for each fixed $\omega \in \Omega$, we have a function of t, $x_t(\omega) \in \mathbb{R}^n$, which is called a sample path of the process.

An $\mathbb{R}^n$-valued stochastic process $\{x_t\}_{t\geq 0}$ is said to be cadlag if it is right continuous and for almost all $\omega \in \Omega$ the left limit $\lim_{s\to t} x_s(\omega)$ exists and is finite for all $t > 0$. It is said to be adapted if for every t, x_t is $\mathcal{F}_t$-measurable. It is said to be progressively measurable if for every $T \geq 0$, $\{x_t\}_{0\leq t\leq T}$ regarded as a function of (t, ω) from $[0, T] \times \Omega$ to $\mathbb{R}^n$ is $\mathcal{B}([0, T]) \times \mathcal{F}_T$-measurable, where $\mathcal{B}([0, T])$ is the family of all Borel sub-sets of $[0, T]$.

In this thesis, we discuss both that the distribution of the numerical solution approximates the underlying distribution of the random variable $x_t(\omega)$ as $t \to \infty$, and that the numerical solution approximates the underlying sample path.

A random variable $\tau : \Omega \to [0, \infty]$ is called an $\{\mathcal{F}_t\}$-stopping time, if $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for any $t \geq 0$. Stopping time is essential for this thesis and we quote the following useful theorems. We refer the readers to Section 1.3 of (Mao, 2008) for more details.

Theorem 2.1.1 *If $\{x_t\}_{t\geq 0}$ is a progressively measurable process and τ is a stopping time, then $x_\tau \mathbf{1}_{\tau<\infty}$ is $\mathcal{F}_\tau$-measurable. In particular, if τ is finite, then x_τ is $\mathcal{F}_\tau$-measurable.*

Theorem 2.1.2 *Let $\{x_t\}_{t\geq 0}$ be an $\mathbb{R}^n$-valued cadlag $\{\mathcal{F}_t\}$-adapted process, and D be an open subset of $\mathbb{R}^n$. Define*

$$\tau = \inf\{t \geq 0 : x_t \notin D\},$$

where $\inf \emptyset = \infty$ is used for the convention. Then τ is an $\{\mathcal{F}_t\}$-stopping time, and is called the first exit time from D.

Conditional expectation plays an important role in this thesis, therefore we quote the following general concept of conditional expectation. Let $x \in L^1(\Omega; \mathbb{R})$, and $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra of $\mathcal{F}$. So $(\Omega, \mathcal{G})$ is a measurable space. In general, x is not $\mathcal{G}$-measurable. We now seek an integrable $\mathcal{G}$-measurable random variable y such that it has the same value as x on the average in the sense that

$$\mathbb{E}(\mathbf{1}_G y) = \mathbb{E}(\mathbf{1}_G x) \quad \text{i.e.} \quad \int_G y(\omega) d\mathbb{P}(\omega) = \int_G x(\omega) d\mathbb{P}(\omega) \quad \text{for all } G \in \mathcal{G}.$$

By the Radon-Nikodym theorem, there exists a unique y a.s. It is called the conditional expectation of x under the condition $\mathcal{G}$, and we denote it by

$$y = \mathbb{E}(x|\mathcal{G}). \tag{2.1}$$

If $\mathcal{G}$ is the σ-algebra generated by random variable y, define the σ-algebra generated by y by $\sigma(y)$, we write

$$\mathbb{E}(x|\mathcal{G}) = \mathbb{E}(x|\sigma(y)) = \mathbb{E}_y(x).$$

Some important properties of the conditional expectation we need in this thesis are listed as follows

- $\mathbb{E}(\mathbb{E}(x|\mathcal{G})) = \mathbb{E}(x)$;
- $x \geq 0$ a.s. $\Rightarrow \mathbb{E}(x|\mathcal{G}) \geq 0$;
- $\mathbb{E}(x|\mathcal{G}) = x$, if x is $\mathcal{G}$-measurable;
- x is $\mathcal{G}$-measurable $\Rightarrow \mathbb{E}(xy|\mathcal{G}) = x\mathbb{E}(y|\mathcal{G})$, particularly, x, y are independent $\Rightarrow \mathbb{E}(\mathbb{E}(x|\mathcal{G})y|\mathcal{G}) = \mathbb{E}(x|\mathcal{G})\mathbb{E}(y|\mathcal{G})$;
- $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F} \Rightarrow \mathbb{E}(\mathbb{E}(x|\mathcal{G}_2)|\mathcal{G}_1) = \mathbb{E}(x|\mathcal{G}_1)$.

The concept of martingale states that the best estimate of the expectation of future status given all the history information is the current status. Using the concept of the conditional expectation, formally an $\mathbb{R}^n$-valued $\{\mathcal{F}_t\}$-adapted integrable process $\{M_t\}_{t\geq 0}$ is called a martingale with respect to $\{\mathcal{F}_t\}$ if

$$\mathbb{E}(M_t|\mathcal{F}_s) = M_s \quad \text{a.s. for all } 0 \leq s < t < \infty. \tag{2.2}$$

If $x = \{x_t\}_{t\geq 0}$ is a progressively measurable process and τ is a stopping time, then $x^\tau = \{x_{\tau\wedge t}\}_{t\geq 0}$ is called a stopped process of x. The following is the well-known Doob martingale stopping theorem.

Theorem 2.1.3 *Let $\{M_t\}_{t\geq 0}$ be an $\mathbb{R}^n$-valued martingale with respect to $\{\mathcal{F}_t\}$, and τ, ρ be two finite stopping times. Then*

$$\mathbb{E}(M_\tau|\mathcal{F}_\rho) = M_{\tau\wedge\rho} \quad a.s.$$

Particularly, if τ is a stopping time, then

$$\mathbb{E}(M_{\tau\wedge t}|\mathcal{F}_s) = M_{\tau\wedge s} \quad a.s.$$

holds for all $0 \leq s < t < \infty$. That is, the stopped process $M^\tau = \{M_{\tau\wedge t}\}$ is still a martingale with respect to the same filtration $\{\mathcal{F}_t\}$.

If the equality in (2.2) is replaced by inequality, we have the two new concepts.

A real-valued $\{\mathcal{F}_t\}$-adapted integrable process $\{M_t\}_{t\geq 0}$ is called a supermartingale, if

$$\mathbb{E}(M_t|\mathcal{F}_s) \leq M_s \quad \text{a.s. for all } 0 \leq s < t < \infty.$$

And it is called a submartingale, if

$$\mathbb{E}(M_t|\mathcal{F}_s) \geq M_s \quad \text{a.s. for all } 0 \leq s < t < \infty.$$

In addition, a right continuous adapted process $M = \{M_t\}_{t\geq 0}$ is called a local martingale, if there exists a nondecreasing sequence $\{\tau_k\}_{k\geq 1}$ of stopping times with $\tau_k \to \infty$ a.s. (when $k \to \infty$) such that every $\{M_{\tau_k \wedge t} - M_0\}_{t\geq 0}$ is a martingale. By Theorem 2.1.3 it can be seen that every martingale is a local martingale, but the converse statement is not always true.

The following semi-martingale convergence theorem plays a key role in the stability analysis of this thesis.

Theorem 2.1.4 *Let $\{A_t\}_{t\geq 0}$ and $\{B_t\}_{t\geq 0}$ be two continuous adapted increasing precesses with $A_0 = B_0 = 0$ a.s. Let $\{M_t\}_{t\geq 0}$ be a real-valued continuous local martingale with $M_0 = 0$ a.s. Le ξ be a nonnegative $\mathcal{F}_0$-measurable random variable. Define*

$$X_t = \xi + A_t - B_t + M_t \quad \text{for } t \geq 0.$$

If X_t is nonnegative, then

$$\left\{\lim_{t\to\infty} A_t < \infty\right\} \subset \left\{\lim_{t\to\infty} X_t \text{ exists and is finite}\right\} \cap \left\{\lim_{t\to\infty} B_t < \infty\right\} \quad a.s.$$

where $D \subset G$ a.s. means $\mathbb{P}(D \cap G^c) = 0$. Particularly, if $\lim_{t\to\infty} A_t < \infty$ a.s., then for almost all $\omega \in \Omega$

$$\lim_{t\to\infty} X_t(\omega) \text{ exists and is finite, and } \lim_{t\to\infty} B_t(\omega) < \infty.$$

The next two theorems are often used in the proofs in the this thesis without explicitly referring.

Theorem 2.1.5 Monotonic convergence theorem *If $\{x_k\}$ is an increasing sequence of nonnegative random variables, then*

$$\lim_{k\to\infty} \mathbb{E}x_k = \mathbb{E}\left(\lim_{k\to\infty} x_k\right).$$

Theorem 2.1.6 Dominated convergence theorem *Let $p \le 1$, $\{x_k\} \subset L^p(\Omega; \mathbb{R}^n)$ and $y \in L^p(\Omega; \mathbb{R})$. Assume that $|x_k| \le y$ a.s. and $\{x_k\}$ converges to x in probability. Then $x \in L^p(\Omega; \mathbb{R})$, $\{x_k\}$ converges to x in L^p, and*

$$\lim_{k\to\infty} \mathbb{E}x_k = \mathbb{E}x.$$

We end up this section by quoting the well-known Borel-Cantelli lemma.

Lemma 2.1.7 Borel-Cantelli's lemma

- *If $\{A_k\} \subset \mathcal{F}$ and $\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty$, then*

$$\mathbb{P}\left(\limsup_{k\to\infty} A_k\right) = 0.$$

  That is, there exists a set $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ and an integer-valued random variable k_0 such that for every $\omega \in \Omega_0$ we have $\omega \notin A_k$ whenever $k \ge k_0(\omega)$.

- *If the sequence $\{A_k\} \subset \mathcal{F}$ is independent and $\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty$, then*

$$\mathbb{P}\left(\limsup_{k\to\infty} A_k\right) = 1.$$

  That is, there exists a set $\Omega_\theta \in \mathcal{F}$ with $\mathbb{P}(\Omega_\theta) = 1$ such that for every $\omega \in \Omega_\theta$, there exists a sub-sequence $\{A_{k_i}\}$ such that the ω belongs to every A_{k_i}.

2.2 Brownian Motion and Stochastic Integrals

Brownian motion depicts the random movement of pollen grains suspended in water, which was initially observed by the Scottish botanist Robert Brown through a microscope. And the rigorous mathematical explanation is due to Norbert Wiener.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$. A one-dimensional Brownian motion is a real-valued continuous $\{\mathcal{F}_t\}$-adapted process $\{B_t\}_{t\geq 0}$ possessing the following three properties:

- $B_0 = 0$ a.s.;
- for $0 \leq s < t < \infty$, the increment $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$;
- for $0 \leq s < t < \infty$, the increment $B_t - B_s$ is independent of $\mathcal{F}_s$.

A n-dimensional process $\{B_t = (B_t^1, ..., B_t^n)\}_{t\geq 0}$ is called a n-dimensional Brownian motion, if, for every i, $\{B_t^i\}$ is a one-dimensional Brownian motion and $\{B_t^1\}$, ..., $\{B_t^n\}$ are independent. The one-dimensional Brownian motion has many properties (similar properties hold for n-dimensional Brownian motion), we summarise some of them that are useful in this thesis and refer the readers to (Revuz & Yor, 1999) for more details:

- $\{B_t\}$ is a continuous square-integrable martingale and its quadratic variation is $\langle B, B\rangle_t = t$ for all $t \geq 0$;
- The strong law of large numbers indicates that $\lim_{t\to\infty} B_t/t = 0$ a.s.;
- For almost every $\omega \in \Omega$, the Brownian sample path $B_t(\omega)$ is nowhere differentiable.

Due to the last property, the integral with respect to the Brownian motion can not be defined in the ordinary way. Kiyosi Itô firstly defined the stochastic integral (Itô, 1944; Itô, 1946)

$$\int_a^b f(t)dB_t$$

with respect to a Brownian motion. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. Let $\{B_t\}_{t\geq 0}$ be a one-dimensional Brownian motion defined on the probability space adapted to the filtration.

Definition 2.2.1 *Let $0 \leq a < b < \infty$. Denote by $\mathcal{M}^2([a,b];\mathbb{R})$ the space of all real-valued measurable $\{\mathcal{F}_t\}$-adapted processes $f = \{f(t)\}_{a \leq t \leq b}$ such that*

$$\|f\|_{a,b}^2 = \mathbb{E}\int_a^b |f(t)|^2 dt < \infty.$$

Briefly speaking, to define the Itô integral $\int_a^b f(t)dB_t$, firstly we define the integral $\int_a^b g(t)dB_t$ for a class of simple processes $g = \{g(t)\}_{a \leq t \leq b}$. Then we show that each $f \in \mathcal{M}^2([a,b];\mathbb{R})$ can be approximated by such simple processes g. Finally, we define the limit of $\int_a^b g(t)dB_t$ as the integral $\int_a^b f(t)dB_t$.

There are other ways to define the stochastic integral, for example the Stratonovich integral (Stratonovich, 1966). One of the important difference between the Itô integral and the Stratonovich integral is that the former one is a martingale but the later one is not. In this thesis, we focus on the Itô integral.

The Itô stochastic integral has many sound properties, we list some of them as follows. Let $f, g \in \mathcal{M}^2([0,T];\mathbb{R})$, a, b be two real numbers with $0 \leq a \leq b \leq T$, and τ, ρ be two stopping times such that $0 \leq \tau \leq \rho \leq T$ a.s.

- $\int_a^b f(t)dB_t$ is $\mathcal{F}_b$-measurable;
- $\mathbb{E}\left(\int_a^b f(t)dB_t|\mathcal{F}_a\right) = 0$;
- $\mathbb{E}\left(|\int_a^b f(t)dB_t|^2|\mathcal{F}_a\right) = \mathbb{E}\left(\int_a^b |f(t)|^2dt|\mathcal{F}_a\right) = \int_a^b \mathbb{E}(|f(t)|^2|\mathcal{F}_a)dt$;
- $\mathbb{E}\left(\int_\tau^\rho f(t)dB_t|\mathcal{F}_\tau\right) = 0$;
- $\mathbb{E}\left(|\int_\tau^\rho f(t)dB_t|^2|\mathcal{F}_\tau\right) = \mathbb{E}\left(\int_\tau^\rho |f(t)|^2dt|\mathcal{F}_\tau\right)$.

To define the multi-dimensional Itô stochastic integral, let $f \in \mathcal{M}([0,T];\mathbb{R}^{n\times m})$ and $\{B_t = (B_t^1, ..., B_t^m)\}_{t\geq 0}$ be an m-dimensional Brownian motion, the n-dimensional stochastic integral $\int_0^t f(s)dB_s$ is a n-column-vector-valued process whose ith component is $\sum_{j=1}^m \int_0^t f_{ij}(s)dB_s^j$. Similar properties like those listed above hold for the multi-dimensional Itô stochastic integral as well.

2.3 Stochastic Differential Equations

This section is devoted to the basic concept of stochastic differential equations (SDEs) and the corresponding stochastic version of the chain rule for Itô, which is known as Itô formula.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. Let $B(t) = (B_1(t), ..., B_m(t))$ be an m-dimensional Brownian motion defined on the space. Let x_0 be an $\mathcal{F}_0$-measurable $\mathbb{R}^n$-valued random variable such that $\mathbb{E}|x_0|^2 < \infty$. For any $T > 0$, let $f : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times [0, T] \to \mathbb{R}^{n\times m}$ be both Borel measurable. Without loss of generality, we set the starting time $t_0 = 0$. The n-dimensional stochastic differential equation of Itô type is defined as

$$dx(t) = f(x(t))dt + g(x(t))dB_i(t), \quad t \in [0, T], \tag{2.3}$$

with initial value $x(0) = x_0$. In this thesis, we call $f(x(t))$ the drift coefficient and $g(x(t))$ the diffusion coefficient. This SDE is equivalent to the following stochastic integral equation:

$$x(t) = x_0 + \int_0^t f(x(s))ds + \int_0^t g(x(s))dB(s) \quad \text{for } t \in [0, T]. \tag{2.4}$$

The SDE (2.3) is the main equation that we focus on in this thesis, there may be slight changes in the notations in each of the following chapters. We first give the definition of the solution.

Definition 2.3.1 *An $\mathbb{R}^n$-valued stochastic process $\{x(t)\}_{0\leq t\leq T}$ is called a solution of SDE (2.3), if it possesses the following properties:*

- *$\{x(t)\}$ is continuous and $\mathcal{F}_t$-adapted;*
- *$\{f(x(t))\} \in \mathcal{L}^1([0, T]; \mathbb{R}^n)$ and $\{g(x(t))\} \in \mathcal{L}^2([0, T]; \mathbb{R}^{n\times m})$;*
- *the stochastic integral equation (2.4) holds for every $t \in [0, T]$ with probability 1.*

A solution $\{x(t)\}_{0\le t\le T}$ to (2.3) is said to be unique if any other solution $\{\bar{x}(t)\}_{0\le t\le T}$ is indistinguishable from $\{x(t)\}_{0\le t\le T}$, i.e.

$$\mathbb{P}(x(t) = \bar{x}(t) \text{ for all } 0 \le t \le T) = 1.$$

Next, we quote two theorems on the existence and uniqueness of the solutions. The conditions in the first one are coefficients related.

Theorem 2.3.2 *Assume that there exist two positive constants $\bar{K} = \bar{K}(R)$ and K such that*

Local Lipschitz condition *for all $x, y \in \mathbb{R}^n$ with $|x| \vee |y| \le R$*

$$|f(x) - f(y)|^2 \vee |g(x) - g(y)|^2 \le \bar{K}(R)|x - y|^2;$$

Linear growth condition *for all $x \in \mathbb{R}^n$*

$$|f(x)|^2 \vee |g(x)|^2 \le K(1 + |x|^2).$$

Then there exists a unique solution $x(t)$ to SDE (2.3) and the solution belongs to $\mathcal{M}^2([0,T];\mathbb{R}^n)$.

Before presenting a more general theorem, we state **the multi-dimensional Itô formula**.

Theorem 2.3.3 *Let $x(t)$ be a n-dimensional Itô process on $t \ge 0$ with the stochastic differential equation (2.3), here $f \in \mathcal{L}^1(\mathbb{R}_+;\mathbb{R}^n)$ and $g \in \mathcal{L}^2(\mathbb{R}_+;\mathbb{R}^{n\times m})$. Let $V(x(t),t) \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+;\mathbb{R})$, which is twice continuously differentiable with respect to x and continuously differentiable with respect to t. Then $V(x(t),t)$ is also an Itô process with the stochastic differential equation given by*

$$dV(x(t),t) = \Bigg[\frac{\partial V(x(t),t)}{\partial t} + \frac{\partial V(x(t),t)}{\partial x} f(x(t))$$
$$+ \quad \frac{1}{2}\text{trace}(g^T(x(t))\frac{\partial^2 V(x(t),t)}{\partial x^2} g(x(t)))\Bigg] dt + \frac{\partial V(x(t),t)}{\partial x} g(x(t)) dB(t).$$

Let $S_h = \{x \in \mathbb{R}^n : |x| < h\}$. For $0 < h \leq \infty$, let $V(x,t) \in C^{2,1}(S_h \times \mathbb{R}_+; \mathbb{R}_+)$. Define the differential operator L associate with SDE (2.3) by

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{n} f_i(x(t)) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} [g(x(t))g^T(x(t))]_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}.$$

If L acts on a function $V \in C^{2,1}(S_h \times \mathbb{R}_+; \mathbb{R}_+)$, by Itô formula we see that

$$dV(x(t),t) = LV(x(t),t)dt + \frac{\partial V(x(t),t)}{\partial x} g(x(t))dB(t).$$

The second theorem on the existence and uniqueness of the solution of (2.3) is in a certain sense "the best possible".

Theorem 2.3.4 *Suppose that the local Lipschitz condition holds, moreover, that there exists a nonnegative function $V \in C^{2,1}(S_h \times \mathbb{R}_+; \mathbb{R}_+)$ such that for some constant $c > 0$*

$$LV \leq cV,$$

$$V_R = \inf_{|x|>R} V(x,t) \to \infty \quad \text{as } R \to \infty.$$

Then there exists a unique solution $x(t)$ to SDE (2.3) and the solution belongs to $\mathcal{M}^2([0,T]; \mathbb{R}^n)$.

In general, the Markov property means that given a Markov process, the past and future are independent when the present is known. Mathematically, denote the σ-algebra generated by $\sigma\{x(t) : 0 \leq r \leq s\}$ by $\mathcal{F}_s$. A n-dimensional $\mathcal{F}_t$-adapted process $\{x(t)\}_{t \geq 0}$ is called a Markov process, if the following Markov property is satisfied: for all $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}^n$,

$$\mathbb{P}(x(t) \in A | \mathcal{F}_s) = \mathbb{P}(x(t) \in A | x(s)).$$

There are several equivalent formulations of the Markov property, for example for any bounded Borel measurable function $\phi : \mathbb{R}^n \to \mathbb{R}$ and $0 \leq s \leq t < \infty$,

$$\mathbb{E}(\phi(x(t)) | \mathcal{F}_s) = \mathbb{E}(\phi(x(t)) | x(s)).$$

We can also rewrite it as

$$\mathbb{E}(\phi(x(t))|\mathcal{F}_s) = \mathbb{E}_{x(s),s}(\phi(x(t))).$$

The transition probability of the Markov process is a function $\mathbb{P}(x(s), s; A, t)$, defined on $0 \leq s \leq t < \infty$, $x(s) \in \mathbb{R}^n$ and $A \in \mathcal{B}^n$, with the following properties

- For every $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}^n$,

$$\mathbb{P}(x(s), s; A, t) = \mathbb{P}(x(t) \in A|x(s));$$

- $\mathbb{P}(x(s), s; \cdot, t)$ is a probability measure on $\mathcal{B}^n$ for every $0 \leq s \leq t < \infty$ and $x(s) \in \mathbb{R}^n$;

- $\mathbb{P}(\cdot, s; A, t)$ is Borel measurable for every $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}^n$;

- The Chapman-Kolmogorov equation

$$\mathbb{P}(x(s), s, ; A, t) = \int_{\mathbb{R}^n} \mathbb{P}(y, r; A, t)\mathbb{P}(x(s), s; dy, r)$$

  holds for any $0 \leq s \leq r \leq t < \infty$, $x(s) \in \mathbb{R}^n$ and $A \in \mathbb{R}^n$.

If the constant time is replaced by a stopping time, we have the concept of strong Markov property. A n-dimensional process $\{x(t)\}_{t \geq 0}$ is called a strong Markov process, if the following strong Markov property is satisfied: for any bounded Borel measurable function $\phi : \mathbb{R}^n \to \mathbb{R}$, any finite $\mathcal{F}_t$-stopping time τ and $t \geq 0$,

$$\mathbb{E}(\phi(x(\tau + t))|\mathcal{F}_\tau) = \mathbb{E}(\phi(x(\tau + t))|x(\tau)).$$

It can also be written as

$$\mathbb{E}(\phi(x(\tau + t))|\mathcal{F}_\tau) = \mathbb{E}_{x(\tau),\tau}(\phi(x(\tau + t))).$$

In general, a Markov process is not a strong one. Then conditions that guarantee that a Markov process possesses the strong Markov property are right continuity of the sample paths and the Feller property. If, for any bounded continuous function $\phi : \mathbb{R}^n \to \mathbb{R}$, the mapping

$$(x(s), s) \to \int_{\mathbb{R}^n} \phi(y)\mathbb{P}(x(x), s; dy, s + \lambda)$$

is continuous, for any fixed $\lambda > 0$, we say that the transition probability (or the corresponding Markov process) satisfies the Feller property.

Next, we quote a theorem about the Markov property of the solutions of SDEs.

Theorem 2.3.5 *Let $x(t)$ be a solution of the SDE (2.3), whose coefficients satisfy the conditions of the existence and uniqueness theorem. Then $x(t)$ is a Markov process whose transition probability is defined by*

$$\mathbb{P}(x_s, s; A, t) = \mathbb{P}(x_{x_s,s}(t) \in A),$$

where $x_{x_s,s}(t)$ is the solution of the equation

$$x_{x_s,s}(t) = x_s + \int_s^t f(x_{x_s,s}(r))dr + \int_s^t g(x_{x_s,s}(r))dB(r) \quad \text{on } t \geq s.$$

For the strong Markov property of the solution, we need to strengthen the conditions. We quote one of the classical results as follows.

Theorem 2.3.6 *Let $x(t)$ be a solution of the Itô SDE (2.3). Assume the coefficients satisfy the* **global Lipschitz condition** *that there exists a positive constant $\bar{K}$ such that*

$$|f(x) - f(y)|^2 \vee |g(x) - g(y)|^2 \leq \bar{K}|x - y|^2$$

for all $x, y \in \mathbb{R}^n$. Then $x(t)$ is strong Markov process.

It is clear that the global Lipschitz condition can indicate the linear growth condition.

2.4 Useful Inequalities

Inequalities play a key role in many proofs in this thesis. Therefore, this section serves as an arsenal of inequalities that will be used in this thesis.

We start from the simplest one that for any $x, y \in \mathbb{R}$

$$2ab \leq a^2 + b^2,$$

which indicates a more flexible one that for any $a, b \in \mathbb{R}$ and any $\epsilon > 0$

$$2ab \leq \epsilon a^2 + \frac{1}{\epsilon} b^2.$$

Young's inequality states that for any $a, b \in \mathbb{R}$ and any $\epsilon \in [0, 1]$

$$|a|^{\epsilon}|b|^{(1-\epsilon)} \leq \epsilon|a| + (1-\epsilon)|b|.$$

Hölder's inequality that if $p > 1$, $1/p + 1/q = 1$, x and y are $\mathbb{R}^n$-valued random variables with $\mathbb{E}|x|^p < \infty$ and $\mathbb{E}|y|^q < \infty$ then

$$|\mathbb{E}(x^T y)| \leq (\mathbb{E}|x|^p)^{1/p}(\mathbb{E}|y|^q)^{1/q},$$

is frequently used in the study on finite time strong convergence. The next inequality will be used in the estimate of the difference of polynomials

$$|a^p - b^p| \leq p|a-b|(a^{p-1} + b^{p-1})$$

for any $a, b \geq 0$ and $p \geq 1$.

The Gronwall-type inequality has been widely applied in the theory of stochastic differential equations to prove the results on existence, uniqueness, boundedness and stability.

Theorem 2.4.1 Gronwall's inequality *Let* $T > 0$ *and* $c \geq 0$. *Let* $u(\cdot)$ *be a Borel measurable bounded non-negative function on* $[0, T]$, *and let* $v(\cdot)$ *be a non-negative integral function on* $[0, T]$. *If*

$$u(t) \leq c + \int_0^t v(s)u(s)ds \quad \text{for all } 0 \leq t \leq T,$$

then

$$u(t) \leq c \exp\left(\int_0^t v(s)ds\right) \quad \text{for all } 0 \leq t \leq T.$$

The discrete inequality of Gronwall type has been broadly used in the numerical analysis for SDEs.

Theorem 2.4.2 Discrete Gronwall's inequality *Let M be a positive integer. Let u_k and v_k be non-negative numbers for $k = 0, 1, ..., M$. If*

$$u_k \leq u_0 + \sum_{j=0}^{k-1} v_j u_j, \quad \forall k = 1, 2, ..., M,$$

then

$$u_k \leq u_0 \exp\left(\sum_{j=0}^{k-1} v_j\right), \quad \forall k = 1, 2, ..., M.$$

Chebyshev's inequality that if $c > 0$, $p > 0$ and $\mathbb{E}|x|^p < \infty$ then

$$\mathbb{P}(|x| \geq c) \leq c^{-p}\mathbb{E}|x|^p,$$

is often used in this thesis to relate properties in probability with properties in moment.

Some moment inequalities are presented below. Let $B(t) = (B_1(t), ..., B_m(t))^T$, $t \geq 0$ be an m-dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.

Theorem 2.4.3 *Let $p \geq 2$. Let $g \in \mathcal{M}^2([0,T]; \mathbb{R}^{n\times m})$ such that*

$$\mathbb{E}\int_0^T |g(s)|^p ds < \infty.$$

Then

$$\mathbb{E}\left|\int_0^T g(s)dB(s)\right|^p \leq \left(\frac{p(p-1)}{2}\right)^{\frac{p}{2}} T^{\frac{p-2}{2}}\mathbb{E}\int_0^T |g(s)|^p ds.$$

In particular, for $p = 2$, the equality holds.

Theorem 2.4.4 *Under the same assumptions as previous theorem,*

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left|\int_0^t g(s)dB(s)\right|^p\right) \leq \left(\frac{p^3}{2(p-1)}\right)^{\frac{p}{2}} T^{\frac{p-2}{2}}\mathbb{E}\int_0^T |g(s)|^p ds.$$

Theorem 2.4.5 Burkholder-Davis-Gundy inequality *Let* $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{n\times m})$*. Define, for* $t \geq 0$*,*

$$x(t) = \int_0^t g(s)dB(s) \quad \text{and} \quad A(t) = \int_0^t |g(s)|^2 ds.$$

Then for every $p > 0$*, there exist universal positive constants* c_p,C_p *dependent only on* p*, such that*

$$c_p \mathbb{E}|A(t)|^{\frac{p}{2}} \leq \mathbb{E}\left(\sup_{0\leq s\leq t} |x(s)|^p\right) \leq C_p \mathbb{E}|A(t)|^{\frac{p}{2}}$$

for all $t \geq 0$*. Particularly, one may take*

$$c_p = (p/2)^p, \qquad C_p = (32/p)^{p/2} \qquad \text{if } 0 < p < 2;$$
$$c_p = 1, \qquad C_p = 4 \qquad \text{if } p = 2;$$
$$c_p = (2p)^{-p/2}, \qquad C_p = (p^{p+1}/2(p-1)^{p-1})^{p/2} \qquad \text{if } p > 2.$$

Theorem 2.4.6 Exponential martingale inequality *Let* $g = (g_1, ..., g_m) \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{1\times m})$*, and let* T, α, β *be any positive numbers. Then*

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\left[\int_0^t g(s)dB(s) - \frac{\alpha}{2}\int_0^t |g(s)|^2 ds\right] > \beta\right) \leq e^{-\alpha\beta}.$$

2.5 Some Definitions

Definitions of different properties of both SDEs and numerical solutions are stated in this section.

Definition 2.5.1 *The trivial solution of SDE (2.3) is said to be almost surely stable if*

$$\lim_{t\to\infty} |x(t)| = 0 \quad a.s.$$

for all $x_0 \in \mathbb{R}^n$*. Particularly, if the rate of the stability is exponential, then*

$$\limsup_{t\to\infty} \frac{1}{t} \log |x(t)| < 0 \quad a.s.$$

for all $x_0 \in \mathbb{R}^n$*.*

Definition 2.5.2 *The trivial solution of SDE (2.3) is said to be pth moment stable, if*

$$\lim_{t\to\infty} \mathbb{E}|x(t)|^p = 0$$

for all $x_0 \in \mathbb{R}^n$. Particularly, if the rate is exponential, then there is a pair of positive constants λ and C such that

$$\mathbb{E}|x(t)|^p \leq C|x_0|^p e^{-\lambda t} \quad \text{on } t \geq 0,$$

for all $x_0 \in \mathbb{R}^n$.

In the case that $p = 2$, we call it mean square stability. When p is positive but far less than 1, we call it stability in small moment.

Definition 2.5.3 *Let $p > 0$. The solution of SDE (2.3) is said to be asymptotically bounded in pth moment if there is a positive constant C such that*

$$\limsup_{t\to\infty} \mathbb{E}|x(t)|^p \leq C$$

for all $x_0 \in \mathbb{R}^n$.

When $p = 2$, we say it the asymptotic boundedness in mean square. And if p is far less than 1, we call it the asymptotic boundedness in small moment.

Now we present the corresponding definitions for numerical solutions. Denote the numerical solution to SDE (2.3) by $\{X_k\}_{k=0,1,\ldots}$ (Here we just use it as a symbol, and will specify in corresponding chapters that which method the solution is derived from). Also we denote the step size by Δt.

Definition 2.5.4 *For all $X_0 = x_0 \in \mathbb{R}^n$, the numerical solution $\{X_k\}$ is said to be stable almost surely with respect to the trivial solution if*

$$\lim_{k\to\infty} |X_k| = 0 \quad a.s.$$

In addition, if the rate of the stability is exponential, then

$$\limsup_{k\to\infty} \frac{1}{k\Delta t} \log |X_k| < 0 \quad a.s.$$

for all $X_0 = x_0 \in \mathbb{R}^n$.

Definition 2.5.5 *For all $X_0 = x_0 \in \mathbb{R}^n$, the numerical solution $\{X_k\}$ is said to be pth moment stable with respect to the trivial solution if*

$$\lim_{k\to\infty} \mathbb{E}|X_k|^p = 0.$$

In addition, if the rate of the stability is exponential, then there is a pair of positive constants λ and C such that

$$\mathbb{E}|X_k|^p \le C|X_0|^p e^{-\lambda k \Delta t} \quad \text{on } k = 1, 2, ...$$

for all $X_0 = x_0 \in \mathbb{R}^n$.

Definition 2.5.6 *Let $p > 0$. The numerical solution $\{X_k\}$ is said to be asymptotically bounded in pth moment if there exists a positive constants C such that*

$$\limsup_{k\to\infty} \mathbb{E}|X_k|^p \le C$$

for all $X_0 = x_0 \in \mathbb{R}^n$.

Chapter 3

Asymptotic Moment Boundedness of Numerical Solutions

3.1 Introduction

Asymptotic properties of the solutions of SDEs have been widely studied in the past decades, particularly the stability theory has been attracting lots of attention (see for example, (Mao, 2008) and the references therein).

Due to the difficulty to find the explicit solutions to SDEs, different types of numerical methods have been introduced to approximate the underlying solutions (see, for example, (Hutzenthaler & Jentzen, 2012), (Kloeden & Platen, 1992), (Milstein & Tretyakov, 2004)). Thus the study of the stability of the numerical methods has naturally bloomed in recent years. We mention (Higham, 2000), (Saito & Mitsui, 1996) and (Schurz, 1997) here, as they are among those papers with original ideas. More recent works investigate the stability for different types of SDEs and different sorts of numerical methods, such as (Buckwar & Kelly, 2010; Burrage & Tian, 2000; De la Cruz Cancino *et al.*, 2010; Higham *et al.*, 2007;

Komori, 2008; Mitsui & Saito, 2007; Rodkina & Schurz, 2005; Schurz, 2005; Wu *et al.*, 2010) and the references therein. We also mention some works on stochastic difference equations (Appleby & Rodkina, 2009; Appleby *et al.*, 2009) as they are naturally related to discrete numerical solutions.

Another important asymptotic property of the SDE solutions, the asymptotic boundedness, has its own right. Unlike the stability property that requires the solutions be attracted by an equilibrium state, the boundedness property only requires the solutions stay within certain regime as time tends to infinity (Mao & Yuan, 2006). Works on the boundedness of the underlying SDE solutions can be found, such as (Luo *et al.*, 2011; Mao & Yuan, 2006; Schurz, 2007; Xing & Peng, 2012) and their references therein. But there are few papers investigating the asymptotic boundedness of the numerical solutions.
The main purpose of this chapter is to investigate the asymptotic moment boundedness of two classical numerical methods. We focus on two types of moment, small moment (i.e. pth moment with p much smaller than 1) and second moment. For the case of small moment, they do have some applications. For instance the stochastic permanence studied in stochastic population model, see for example (Li *et al.*, 2011), in which the probability of the solution larger than some constant can be estimated by the small moment together with Markov's inequality. The case of second moment is widely studied for many different asymptotic properties. In this chapter, we find that compare with the case of small moment stronger conditions are required in second moment but better results could be obtained (see Section 3.5 for details). In addition, thanks to Hölder's inequality the asymptotic pth moment boundedness for $1 < p < 2$ could be implied by the second moment boundedness.

Our key aim in this chapter is to answer the question: given that the solution of the underlying Itô type SDE is asymptotically bounded in moment, is there any numerical method that could preserve the boundedness property?

Due to the techniques used to deal with the small moment are much more complicated than those for the second moment, the majority of the chapter is devoted to the case of small moment. This chapter is constructed as follows. We briefly introduce the two classical numerical methods in Section 3.2. The main results of the small moment are developed in Sections 3.3 and 3.4. In each of these two sections we first present the results for the underlying true solution, the relative results for the numerical solution then follow. Section 3.3 is devoted to the asymptotic boundedness of the EM method under the linear growth condition, and Section 3.4 discusses the backward EM method applied to a set of SDEs on which the EM method fails to work. Section 3.5 discusses the results for the case of second moment. The last section summarizes the chapter and discusses some possible future research.

3.2 Mathematical Preliminaries

Throughout this chapter, we let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ which is increasing and right continuous, with $\mathcal{F}_0$ containing all $\mathbb{P}$-null sets. Let $B(t) = (B_1(t), ..., B_m(t))$ be an m-dimensional Brownian motion defined on the probability space. The inner product of x, y in $\mathbb{R}^n$ is denoted by $\langle x, y\rangle$. In this chapter, we consider the n-dimensional Itô SDE

$$dx(t) = f(x(t))dt + \sum_{i=1}^{m} g_i(x(t))dB_i(t), \quad t \geq 0, \quad x(0) \in \mathbb{R}^n. \tag{3.1}$$

We assume that $f, g_1, ..., g_m$: $\mathbb{R}^n \to \mathbb{R}^n$ are smooth enough for the SDE (3.1) to have a unique global solution on $[0, \infty)$ (see, for example, (Mao, 2008)).

Let us recall the two numerical methods we will use below. The reader is referred to (Kloeden & Platen, 1992) and (Milstein & Tretyakov, 2004) for more details on the numerical methods. The Euler–Maruyama (EM) method applied to (3.1) is defined by

$$Y_{k+1} = Y_k + f(Y_k)\Delta t + \sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}, \quad Y_0 = x(0), \tag{3.2}$$

for $k = 0, 1, ...,$ where Δt is the timestep and $\Delta B_{i,k} = B_i((k+1)\Delta t) - B_i(k\Delta t)$ is the Brownian increment.

The backward EM method (or the drift implicit EM method) is defined by

$$Y_{k+1} = Y_k + f(Y_{k+1})\Delta t + \sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}, \quad Y_0 = x(0), \tag{3.3}$$

for $k = 0, 1, ...$

3.3 Euler–Maruyama in Small Moment

We begin by imposing the linear growth condition on both drift and diffusion coefficients of the SDE (3.1):

$$|f(x)|^2 \vee |g_i(x)|^2 \leq K|x|^2 + \alpha \qquad \forall x \in \mathbb{R}^n \text{ and } 1 \leq i \leq m, \tag{3.4}$$

where K and α are positive constants. In this section, we will be concerned with the asymptotic boundedness in small moment of the solution $x(t)$ of (3.1) and the preservation of this property using the EM method.

3.3.1 Asymptotic Boundedness

We first give a sufficient condition for the asymptotic small moment boundedness of the SDE solution. It should be emphasized that more general sufficient condition exists (see, for example, Theorem 5.2, p157 in (Mao & Yuan, 2006)). The condition we employ in Theorem 3.3.1 is in line with the one for the boundedness of numerical solution in Theorem 3.3.2, and it is still an open question that whether there exists a numerical method could recover the asymptotic boundedness of the underlying SDE solution under the more general condition (for example, the condition given in Theorem 5.2 of (Mao & Yuan, 2006)).

Theorem 3.3.1 *Let (3.4) hold. If there exists a positive constant D such that for any $x \in \mathbb{R}^n$*

$$\frac{\langle x, f(x)\rangle + \frac{1}{2}\sum_{i=1}^m |g_i(x)|^2}{D + |x|^2} - \frac{\sum_{i=1}^m \langle x, g_i(x)\rangle^2}{(D + |x|^2)^2} \leq -\lambda + \frac{P_1(|x|)}{D + |x|^2} + \frac{P_3(|x|)}{(D + |x|^2)^2}, \tag{3.5}$$

where λ is a positive constant and $P_i(|x|)$ is a polynomial of $|x|$ with degree i, then there exists a $p^* \in (0,1)$ such that for all $0 < p < p^*$ the solution of (3.1) obeys

$$\limsup_{t\to\infty} \mathbb{E}(|x(t)|^p) \le C, \quad \forall x(0) \in \mathbb{R}^n, \tag{3.6}$$

where C is a positive constant dependent on K, α, p, D, but independent of $x(0)$.

Following the same technique used in Theorem 5.2 in (Mao & Yuan, 2006), by choosing the Lyapunov function $V(x) = (D + |x|^2)^{p/2}$, it is straightforward to prove this theorem. So we omit it here. Now we give the result for the EM solution.

Theorem 3.3.2 *Let (3.4) and (3.5) hold. Then for any $\varepsilon \in (0,\lambda)$, there exists a pair of constants $p^* \in (0,1)$ and $\Delta t^* \in (0,1)$ such that for $\forall p \in (0,p^*)$ and $\forall \Delta t \in (0,\Delta t^*)$, the EM solution (3.2) satisfies*

$$\limsup_{k\to\infty} \mathbb{E}|Y_k|^p \le \frac{C_2'}{p(\lambda-\varepsilon)}, \quad \forall Y_0 \in \mathbb{R}^n, \tag{3.7}$$

where C_2' is a constant dependent on K, α, p and D, but independent of Y_0 and Δt.

Proof. For the constant D in (3.5), we compute

$$\begin{aligned} & D + |Y_{k+1}|^2 \\ = \; & D + |Y_k|^2 + 2\langle Y_k, f(Y_k)\Delta t + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}\rangle + |f(Y_k)\Delta t + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}|^2. \end{aligned}$$

Let

$$\xi_k = \frac{1}{D+|Y_k|^2}(2\langle Y_k, f(Y_k)\Delta t + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}\rangle + |f(Y_k)\Delta t + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}|^2),$$

for any $p \in (0,1)$ we have

$$|D + |Y_{k+1}|^2|^{p/2} = |D + |Y_k|^2|^{p/2}(1+\xi_k)^{p/2}.$$

Clearly $\xi_k > -1$, recalling the fundamental inequality

$$(1+u)^{p/2} \le 1 + \frac{p}{2}u + \frac{p(p-2)}{8}u^2 + \frac{p(p-2)(p-4)}{2^3 \times 3!}u^3, \qquad u > -1, \tag{3.8}$$

we have

$$|D+|Y_{k+1}|^2|^{p/2} \le |D+|Y_k|^2|^{p/2}\left(1+\frac{p}{2}\xi_k+\frac{p(p-2)}{8}\xi_k^2+\frac{p(p-2)(p-4)}{2^3\times 3!}\xi_k^3\right).$$

Hence the conditional expectation

$$\begin{aligned} & \mathbb{E}(|D+|Y_{k+1}|^2|^{p/2}\big|\mathcal{F}_{k\Delta t}) \\ \le\ & |D+|Y_k|^2|^{p/2}\mathbb{E}\left(1+\frac{p}{2}\xi_k+\frac{p(p-2)}{8}\xi_k^2+\frac{p(p-2)(p-4)}{2^3\times 3!}\xi_k^3|\mathcal{F}_{k\Delta t}\right). \end{aligned} \tag{3.9}$$

Since $\Delta B_{i,k}$, $i=1,...,m$, is independent from each other and is independent of $\mathcal{F}_{k\Delta t}$, we have $\mathbb{E}(\Delta B_{i,k}\big|\mathcal{F}_{k\Delta t})=\mathbb{E}(\Delta B_{i,k})=0$, $\mathbb{E}((\Delta B_{i,k})^2\big|\mathcal{F}_{k\Delta t})=\mathbb{E}((\Delta B_{i,k})^2)=\Delta t$ and $\mathbb{E}(\Delta B_{i,k}\Delta B_{j,k}\big|\mathcal{F}_{k\Delta t})=\mathbb{E}(\Delta B_{i,k}\Delta B_{j,k})=\mathbb{E}(\Delta B_{i,k})\mathbb{E}(\Delta B_{j,k})=0$, for $i\neq j$. By (3.4) we can get

$$\begin{aligned} & \mathbb{E}(\xi_k\big|\mathcal{F}_{k\Delta t}) \\ =\ & \mathbb{E}(\frac{1}{D+|Y_k|^2}(2\langle Y_k, f(Y_k)\Delta t+\sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}\rangle \\ & +|f(Y_k)\Delta t+\sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}|^2)\big|\mathcal{F}_{k\Delta t}) \\ =\ & \frac{1}{D+|Y_k|^2}(2\langle Y_k, f(Y_k)\rangle+\sum_{i=1}^m|g_i(Y_k)|^2)\Delta t+\frac{1}{D+|Y_k|^2}|f(Y_k)|^2\Delta t^2 \\ \le\ & \frac{1}{D+|Y_k|^2}(2\langle Y_k, f(Y_k)\rangle+\sum_{i=1}^m|g_i(Y_k)|^2)\Delta t \\ & +K\Delta t^2+\frac{C_2}{D+|Y_k|^2}\Delta t^2. \end{aligned} \tag{3.10}$$

Similarly, we can show that

$$\mathbb{E}(\xi_k^2\big|\mathcal{F}_{k\Delta t}) \ge \frac{4}{(D+|Y_k|^2)^2}\sum_{i=1}^m\langle Y_k, g_i(Y_k)\rangle^2\Delta t-C_1\Delta t^2-\frac{C_2}{(D+|Y_k|^2)^2}\Delta t^2, \tag{3.11}$$

and

$$\mathbb{E}(\xi_k^3\big|\mathcal{F}_{k\Delta t}) \le C_1\Delta t^2+\frac{C_2}{(D+|Y_k|^2)^3}\Delta t^2, \tag{3.12}$$

where C_1 is a positive constant dependent on K, and C_2 is a positive constant dependent on α. C_1 and C_2 may change from line to line. Now consider the following two fractions,

$$\frac{(D+|Y_k|^2)^{p/2}P_1(|Y_k|)}{D+|Y_k|^2} \quad \text{and} \quad \frac{(D+|Y_k|^2)^{p/2}P_3(|Y_k|)}{(D+|Y_k|^2)^2}. \tag{3.13}$$

For $0 < p < 1$ the highest degrees of $|Y_k|$ in the numerators are $p+1$ and $p+3$ respectively, which are smaller than the corresponding highest degrees of $|Y_k|$ in the denominators. Thus for any $|Y_k| \in \mathbb{R}$ there exists an upper bound for both of the fractions. Also it is obvious that $C_2/(D+|Y_{k+1}|^2)^{i-p/2},\ i = 1, 2, 3$ are bounded by some constant that depends on α and D. Substituting (3.10), (3.11) and (3.12) into (3.9), then using (3.4), (3.5) and the arguement for (3.13) we have that

$$\begin{aligned}
&\mathbb{E}((D+|Y_{k+1}|^2)^{p/2}\big|\mathcal{F}_{k\Delta t}) \\
\leq\ & (D+|Y_k|^2)^{p/2}\Bigg(1+\frac{p}{2(D+|Y_{k+1}|^2)}(2\langle Y_k, f(Y_k)\rangle + \sum_{i=1}^m |g_i(Y_k)|^2)\Delta t \\
&+\frac{p(p-2)}{2(D+|Y_k|^2)^2}\sum_{i=1}^m \langle Y_k, g_i(Y_k)\rangle^2 \Delta t + C_1'\Delta t^2\Bigg) + C_2'\Delta t \\
=\ & (D+|Y_k|^2)^{p/2}\Bigg[1+p\Delta t\Bigg(\frac{\langle Y_k, f(Y_k)\rangle + \frac{1}{2}\sum_{i=1}^m |g_i(Y_k)|^2}{D+|Y_k|^2} \\
&-\frac{\sum_{i=1}^m \langle Y_k, g_i(Y_k)\rangle^2}{(D+|Y_k|^2)^2}\Bigg) + \frac{p^2\Delta t\sum_{i=1}^m \langle Y_k, g_i(Y_k)\rangle^2}{2(D+|Y_k|^2)^2} + C_1'\Delta t^2\Bigg] + C_2'\Delta t \\
\leq\ & (D+|Y_k|^2)^{p/2}\left(1-p\lambda\Delta t + \frac{mp^2\Delta t K}{2} + C_1'\Delta t^2\right) + C_2'\Delta t,
\end{aligned}$$

where C_1' is a positive constant dependent on K and p, C_2' is a positive constant dependent on K, α, p and D, and both of them may change from line to line. For any given $\varepsilon \in (0, \lambda)$, choose $p^* \in (0,1)$ sufficiently small for $mp^*K < \varepsilon$, then choose $\Delta t^* \in (0,1)$ sufficiently small for $p^*\lambda\Delta t^* \leq 1$ and $C_1'\Delta t^* \leq \frac{1}{2}p^*\varepsilon$. For any $p \in (0, p^*)$ and any $\Delta t \in (0, \Delta t^*)$ we have

$$\mathbb{E}((D+|Y_{k+1}|^2)^{p/2}\big|\mathcal{F}_{k\Delta t}) \leq (D+|Y_k|^2)^{p/2}(1-p(\lambda-\varepsilon)\Delta t) + C_2'\Delta t.$$

Taking expectations on both sides yields

$$\mathbb{E}((D+|Y_{k+1}|^2)^{p/2}) \leq \mathbb{E}((D+|Y_k|^2)^{p/2})(1-p(\lambda-\varepsilon)\Delta t) + C_2'\Delta t. \tag{3.14}$$

By iteration we have

$$\mathbb{E}((D+|Y_k|^2)^{p/2}) \le \mathbb{E}((D+|Y_0|^2)^{p/2})(1-p(\lambda-\varepsilon)\Delta t)^k + \frac{1-(1-p(\lambda-\varepsilon)\Delta t)^{k-1}}{p(\lambda-\varepsilon)} C_2'. \tag{3.15}$$

Since $\mathbb{E}(|Y_k|^p) \le \mathbb{E}((D+|Y_k|^2)^{p/2})$, we have

$$\mathbb{E}(|Y_k|^p) \le \mathbb{E}((D+|Y_0|^2)^{p/2})(1-p(\lambda-\varepsilon)\Delta t)^k + \frac{1-(1-p(\lambda-\varepsilon)\Delta t)^{k-1}}{p(\lambda-\varepsilon)} C_2'. \tag{3.16}$$

Letting $k \to \infty$, then (3.7) follows.

3.3.2 A Linear Scalar SDE Example

Let us consider a linear scalar SDE,

$$dx(t) = (\alpha_1 + \alpha_2 x(t))dt + (\sigma_1 + \sigma_2 x(t))dB(t), \qquad x(0) \in \mathbb{R}, \tag{3.17}$$

where $\alpha_1, \alpha_2, \sigma_1, \sigma_2$ are real numbers. We impose the condition, $\alpha_2 - \sigma_2^2/2 < 0$. By using this example, we will illustrate

- the existence of the constant, D, in condition (3.5) and how to choose it.

Obviously both drift and diffusion coefficients of (3.17) satisfy the linear growth condition (3.4). Now we consider the condition (3.5),

$$\frac{\langle Y_k, f(Y_k)\rangle + \frac{1}{2}|g(Y_k)|^2}{D+|Y_k|^2} = \frac{(\alpha_2 + \frac{1}{2}\sigma_2^2)Y_k^2}{D+|Y_k|^2} + \frac{(\alpha_1 + \sigma_1\sigma_2)Y_k + \sigma_1^2}{D+|Y_k|^2}, \tag{3.18}$$

and

$$\begin{aligned}
\langle Y_k, g(Y_k)\rangle^2 &= (\sigma_1 Y_k + \sigma_2 Y_k^2)^2 \\
&= \sigma_2^2 Y_k^4 + 2\sigma_1\sigma_2 Y_k^3 + \sigma_1^2 Y_k^2 \\
&= \sigma_2^2 \left(Y_k^2 + \frac{\sigma_1^2}{2\sigma_2^2}\right)^2 - \frac{\sigma_1^4}{4\sigma_2^2} + 2\sigma_1\sigma_2 Y_k^3.
\end{aligned} \tag{3.19}$$

Choose $D = (\sigma_1^2)/(2\sigma_2^2)$ we have

$$\begin{aligned}
&\frac{\langle Y_k, f(Y_k)\rangle + \frac{1}{2}|g(Y_k)|^2}{D+|Y_k|^2} - \frac{\langle Y_k, g(Y_k)\rangle^2}{(D+|Y_k|^2)^2} \\
\le\ & (\alpha_2 - \frac{1}{2}\sigma_2^2) + \frac{(\alpha_1 + \sigma_1\sigma_2)Y_k + \sigma_1^2}{D+|Y_k|^2} \\
& + \frac{1}{(D+|Y_k|^2)^2}\left(\frac{\sigma_1^4}{4\sigma_2^2} - 2\sigma_1\sigma_2 Y_k^3\right).
\end{aligned} \tag{3.20}$$

Thus $-\lambda = \alpha_2 - \sigma_2^2/2$, $P_1(Y_k) = (\alpha_1 + \sigma_1\sigma_2)Y_k + \sigma_1^2$ and $P_3(Y_k) = \sigma_1^4/(4\sigma_2^2) - 2\sigma_1\sigma_2 Y_k^3$. Then the similar process to the proof of Theorem 3.3.2 leads to the property (3.7) for the linear scalar SDE (3.17).

3.4 Backward Euler–Maruyama in Small Moment

So far, we have established some positive results on the asymptotic boundedness in small moment of the EM method under the linear growth condition (3.4). Now we consider to relax the constraint of the drift coefficient by imposing the one-sided Lipschitz condition,

$$\langle x - y, f(x) - f(y) \rangle \leq \bar{\mu}|x - y|^2 + \bar{\alpha} \quad \forall x, y \in \mathbb{R}^n,$$

where $\bar{\mu} \in \mathbb{R}$ and $\bar{\alpha} \in \mathbb{R}^+$. Without losing generality, we further assume for $\forall x \in \mathbb{R}^n$

$$\langle x, f(x) \rangle \leq \mu|x|^2 + \alpha, \tag{3.21}$$

where $\mu \in \mathbb{R}$ and $\alpha \in \mathbb{R}^+$. The diffusion coefficient still obeys the linear growth condition,

$$|g_i(x)|^2 \leq K|x|^2 + \alpha, \qquad 1 \leq i \leq m. \tag{3.22}$$

In this section, we start with a counter example to show that the EM solution will blow up under (3.21) and (3.22). Then we will show that the backward EM method can still preserve the boundedness property of the SDE solution under these conditions.

3.4.1 A Counter Example

Consider the following scalar SDE,

$$dx(t) = (-0.5x(t) - x^3(t) + 1)dt + (x(t) + 1)dB(t), \tag{3.23}$$

to which the EM method is applied.

Lemma 3.4.1 *Suppose $\Delta t \in (0,1)$ and $p \in (0,1)$, then for any $Y_0 \in \mathbb{R}$,*

$$\lim_{k\to\infty} \mathbb{E}|Y_k|^p = \infty. \tag{3.24}$$

Proof. By the property of conditional expectations, we have

$$\mathbb{E}|Y_{k+1}|^p = \mathbb{E}[\mathbb{E}(|Y_{k+1}|^p \big| Y_1)] \geq \mathbb{E}[\mathbf{1}_{\{|Y_1|^p \geq 2^3/\Delta t^{p/2}\}} \mathbb{E}(|Y_{k+1}|^p \big| Y_1)]. \tag{3.25}$$

Since there is a positive probability that the first Brownian motion increment will make $|Y_1|^p \geq 2^3/\Delta t^{p/2}$, we only need to show that for $|Y_1|^p \geq 2^3/\Delta t^{p/2}$, $\mathbb{E}(|Y_{k+1}|^p \big| Y_1) \geq 2^{k+3}/\Delta t^{p/2}$ for all $k \geq 0$. We show this by induction. Clearly, $\mathbb{E}(|Y_1|^p \big| Y_1) = |Y_1|^p \geq 2^3/\Delta t^{p/2}$. Suppose $\mathbb{E}(|Y_k|^p \big| Y_1) \geq 2^{k+2}/\Delta t^{p/2}$ for some $k \geq 1$, we will show that for any $\Delta t \in (0,1)$, $\mathbb{E}(|Y_{k+1}|^p \big| Y_1) \geq 2^{k+3}/\Delta t^{p/2}$. Applying the EM method to the SDE (3.23),

$$|Y_{k+1}| = |Y_k - 0.5\Delta t Y_k - \Delta t Y_k^3 + Y_k \Delta B_k + \Delta t + \Delta B_k|.$$

Then by the fundamental inequality, $|a+b|^p > |a|^p - |b|^p$, we have

$$\begin{aligned} |Y_{k+1}|^p &\geq |\Delta t Y_k^3 + (0.5\Delta t - 1) Y_k + Y_k \Delta B_k|^p - |\Delta t|^p - |\Delta B_k|^p \\ &\geq \Delta t^p |Y_k|^{3p} - (1 - 0.5\Delta t)^p |Y_k|^p - |Y_k \Delta B_k|^p - |\Delta t|^p - |\Delta B_k|^p. \end{aligned}$$

By the Hölder inequality, we have $\mathbb{E}(|Y_k|^{3p} \big| Y_1) \geq (\mathbb{E}(|Y_k|^p \big| Y_1))^3$. Since ΔB_k is independent of Y_1 for all $k > 0$, $\mathbb{E}(|\Delta B_k|^p \big| Y_1) = \mathbb{E}(|\Delta B_k|^p) < 2$. Then taking conditional expectation on both sides we have

$$\begin{aligned} &\mathbb{E}(|Y_{k+1}|^p \big| Y_1) \\ \geq\ &\mathbb{E}(|Y_k|^p \big| Y_1)(\Delta t^p (\mathbb{E}(|Y_k|^p \big| Y_1))^2 - (1 - 0.5\Delta t)^p - \mathbb{E}|\Delta B_k|^p) - |\Delta t|^p - \mathbb{E}|\Delta B_k|^p \\ \geq\ &\mathbb{E}(|Y_k|^p \big| Y_1)(\Delta t^p (\mathbb{E}(|Y_k|^p \big| Y_1))^2 - 1 - 2) - 1 - 2 \\ \geq\ &\frac{2^{k+2}}{\Delta t^{p/2}}(2^{2k+4} - 3) - 3 \\ \geq\ &\frac{2^{k+3}}{\Delta t^{p/2}}. \end{aligned}$$

Then substituting it back to (3.25) we obtain

$$\mathbb{E}|Y_{k+1}|^p \geq \frac{2^{k+3}}{\Delta t^{p/2}} \mathbb{P}(|Y_1|^p \geq \frac{2^3}{\Delta t^{p/2}}).$$

Hence the assertion holds.

This lemma states that for any initial value, the pth moment, $0 < p < 1$, of the EM solution will blow up. This contrasts to the initial-data-independent asymptotic boundedness of the underlying SDE solution, shown by Theorem 3.4.2. Hence the EM method is no longer a good candidate.

3.4.2 Asymptotic Boundedness

Let us present another theorem on the asymptotic boundedness of the solution of the SDE (3.1). The condition used in Theorem 3.4.2 will be employed in Theorem 3.4.3 as well.

Theorem 3.4.2 *Let (3.21) and (3.22) hold. If there exists a constant D such that*

$$\frac{\sum_{i=1}^{m} |g_i(x)|^2}{D + |x|^2} - \frac{\sum_{i=1}^{m} \langle x, g_i(x) \rangle^2}{(D + |x|^2)^2} \leq \rho + \frac{P_1(|x|)}{D + |x|^2} + \frac{P_3(|x|)}{(D + |x|^2)^2}, \tag{3.26}$$

where ρ is a constant with $\mu + \rho/2 < 0$, then there exists a $p^* \in (0,1)$ such that for all $0 < p < p^*$ the solution of SDE (3.1) obeys

$$\limsup_{t \to \infty} \mathbb{E}(|x(t)|^p) \leq C, \quad \forall x(0) \in \mathbb{R}^n, \tag{3.27}$$

where C is a constant dependent on μ, α, K, p and D, but independent of $x(0)$.

It is straightforward to adapt the proof of Theorem 3.3.1 to show Theorem 3.4.2.

Let us now begin to discuss the asymptotic boundedness in small moment of the backward EM solution (3.3) under conditions (3.21), (3.22) and (3.26).

Theorem 3.4.3 *Let (3.21), (3.22) and (3.26) hold. Then there exists a pair of constants $p^* \in (0,1)$ and $\Delta t^* \in (0, 1/(2|\mu|))$ such that for $\forall p \in (0, p^*)$ and $\forall \Delta t \in (0, \Delta t^*)$, the backward EM solution (3.3) satisfies*

$$\limsup_{k \to \infty} \mathbb{E}|Y_k|^p \leq \frac{C_2'}{p(\lambda - \varepsilon)}, \quad \forall Y_0 \in \mathbb{R}^n, \tag{3.28}$$

where $-\lambda = \mu + \rho/2 < 0$, $\varepsilon \in (0, |\mu + \rho/2|)$ and C_2' is a constant dependent on K, α, p and D, but independent of Y_0 and Δt.

Proof. From (3.3), we have

$$|Y_{k+1}|^2 = \langle Y_{k+1}, Y_k + \sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}\rangle + \langle Y_{k+1}, f(Y_{k+1})\Delta t\rangle.$$

By (3.21), we obtain

$$|Y_{k+1}|^2 \le \frac{1}{2}|Y_{k+1}|^2 + \frac{1}{2}|Y_k + \sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}|^2 + \mu\Delta t|Y_{k+1}|^2 + \alpha\Delta t.$$

Hence

$$\begin{aligned}\frac{D}{1-2\mu\Delta t} + |Y_{k+1}|^2 &\le \frac{D}{1-2\mu\Delta t} + \frac{1}{1-2\mu\Delta t}(|Y_k|^2 \\ &\quad +2\langle Y_k, \sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}\rangle + |\sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}|^2 + 2\alpha\Delta t) \\ &\le \frac{D+|Y_k|^2}{1-2\mu\Delta t}(1+\zeta_k),\end{aligned}$$

where

$$\zeta_k = \frac{1}{D+|Y_k|^2}(2\langle Y_k, \sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}\rangle + |\sum_{i=1}^{m} g_i(Y_k)\Delta B_{i,k}|^2 + 2\alpha\Delta t).$$

It is clear that $\zeta_k > -1$ for all $k \ge 0$. For any $p \in (0,1)$, by inequality (3.8) we have

$$\begin{aligned}&\mathbb{E}((D+|Y_{k+1}|^2)^{p/2}\big|\mathcal{F}_{k\Delta t}) \\ \le\ &\left(\frac{D+|Y_k|^2}{1-2\mu\Delta t}\right)^{p/2}\mathbb{E}\bigg(1+\frac{p}{2}\zeta_k + \frac{p(p-2)}{8}\zeta_k^2 \\ &+\frac{p(p-2)(p-4)}{2^3\times 3!}\zeta_k^3|\mathcal{F}_{k\Delta t}\bigg).\end{aligned} \tag{3.29}$$

Then following the same way as Theorem 3.3.2, by (3.22) we can show

$$\mathbb{E}(\zeta_k\big|\mathcal{F}_{k\Delta t}) = \frac{1}{D+|Y_k|^2}(\sum_{i=1}^{m}|g_i(Y_k)|^2\Delta t + 2\alpha\Delta t), \tag{3.30}$$

$$\mathbb{E}(\zeta_k^2\big|\mathcal{F}_{k\Delta t}) \ge \frac{4\sum_{i=1}^{m}\langle Y_k, g_i(Y_k)\rangle^2}{(D+|Y_k|^2)^2}\Delta t - \frac{P_2(|Y_k|)\Delta t^2}{(D+|Y_k|^2)^2}, \tag{3.31}$$

and

$$\mathbb{E}(\zeta_k^3\big|\mathcal{F}_{k\Delta t}) \le C_1\Delta t^2 + \frac{P_4(|Y_k|)\Delta t^2}{(D+|Y_k|^2)^3}, \tag{3.32}$$

where C_1 is a constant dependent on K. Substituting (3.30), (3.31) and (3.32) into (3.29), then using (3.22), (3.26) and the similar argument in (3.13) we obtain

$$\begin{aligned}
& \mathbb{E}((D+|Y_{k+1}|^2)^{p/2}\big|\mathcal{F}_{k\Delta t}) \\
\leq\ & \left(\frac{D+|Y_k|^2}{1-2\mu\Delta t}\right)^{p/2}\Bigg(1+\frac{p}{2}\frac{\sum_{i=1}^m|g_i(Y_k)|^2}{D+|Y_k|^2}\Delta t \\
& +\frac{p(p-2)}{8}\frac{4\sum_{i=1}^m\langle Y_k,g_i(Y_k)\rangle^2}{(D+|Y_k|^2)^2}\Delta t+\frac{p(p-2)(p-4)}{2^3\times 3!}C_1\Delta t^2\Bigg)+C_2'\Delta t \\
\leq\ & \frac{(D+|Y_k|^2)^{p/2}}{(1-2\mu\Delta t)^{p/2}}(1+\frac{1}{2}p\rho\Delta t+\frac{1}{2}p^2mK\Delta t+C_1'\Delta t^2)+C_2'\Delta t,
\end{aligned}$$

where C_1' is a positive constant dependent on K and p, C_2' is a positive constant dependent on K, α, μ, p and D, and both of them may change from line to line. Taking expectations on both sides, we obtain

$$\mathbb{E}((D+|Y_{k+1}|^2)^{p/2})\leq\frac{1+\frac{1}{2}p\rho\Delta t+\frac{1}{2}p^2mK\Delta t+C_1'\Delta t^2}{(1-2\mu\Delta t)^{p/2}}\mathbb{E}((D+|Y_k|^2)^{p/2})+C_2'\Delta t. \tag{3.33}$$

For any $\varepsilon\in(0,|\mu+\rho/2|)$, by choosing p^* sufficiently small such that $p^*mK\leq\varepsilon/4$ and sufficiently small Δt^*, for $p<p^*$ and $\Delta t<\Delta t^*$ we have

$$(1-2\mu\Delta t)^{p/2}\geq 1-p\mu\Delta t-C_3\Delta t^2>0, \tag{3.34}$$

where $C_3>0$ is a constant dependent on μ and p. Then further reducing Δt^* gives that for $\Delta t<\Delta t^*$

$$C_1'\Delta t<\frac{1}{8}p\varepsilon,\quad C_3\Delta t<\frac{1}{4}\varepsilon,\quad |p(\mu+\frac{1}{4}\varepsilon)\Delta t|<\frac{1}{2}.$$

Using these three inequalities together with (3.34), we have from (3.33) that

$$\mathbb{E}((D+|Y_{k+1}|^2)^{p/2})\leq\frac{1+\frac{1}{2}p(\rho+\frac{1}{2}\varepsilon)\Delta t}{1-p(\mu+\frac{1}{4}\varepsilon)\Delta t}\mathbb{E}((D+|Y_k|^2)^{p/2})+C_2'\Delta t. \tag{3.35}$$

Since that for any $h\in[-0.5,0.5]$

$$(1-h)^{-1}=1+h+h^2\sum_{i=0}^{\infty}h^i\leq 1+h+h^2\sum_{i=0}^{\infty}0.5^i=1+h+2h^2,$$

then by further reducing Δt^* such that for any $\Delta t<\Delta t^*$ we obtain

$$4p(\mu+\frac{1}{4}\varepsilon)^2\Delta t+(\rho+\frac{1}{2}\varepsilon)(p(\mu+\frac{1}{4}\varepsilon)\Delta t+2(p(\mu+\frac{1}{4}\varepsilon)\Delta t)^2)<\varepsilon.$$

Together with (3.35), we arrive at

$$\begin{aligned}
&\mathbb{E}((D+|Y_{k+1}|^2)^{p/2}) \\
\leq\ & (1+\frac{1}{2}p(\rho+\frac{1}{2}\varepsilon)\Delta t)(1+p(\mu+\frac{1}{4}\varepsilon)\Delta t \\
&+2(p(\mu+\frac{1}{4}\varepsilon)\Delta t)^2)\mathbb{E}((D+|Y_k|^2)^{p/2})+C_2'\Delta t \\
\leq\ & [1+p(\mu+\frac{1}{2}\rho+\varepsilon)\Delta t]\mathbb{E}((D+|Y_k|^2)^{p/2})+C_2'\Delta t. \qquad (3.36)
\end{aligned}$$

Then by iteration and let $k\to\infty$, we have

$$\limsup_{k\to\infty}\mathbb{E}(|Y_{k+1}|^p)\leq\limsup_{k\to\infty}\mathbb{E}((D+|Y_{k+1}|^2)^{p/2})\leq\frac{C_2'}{-p(\mu+\frac{1}{2}\rho+\varepsilon)}.$$

The proof is complete.

3.5 The Second Moment

In this section, we discuss the asymptotic boundedness in second moment for both the EM method and the backward EM method. Follow the same structure as previous sections, we first give the results for the underlying SDEs, then the results for numerical solutions are proved under the same conditions.

3.5.1 The EM Method

For the asymptotic second moment boundedness of the underlying solution, we still require condition (3.4) but replace condition (3.5) by the following condition that

$$\langle x,f(x)\rangle+\frac{1}{2}\sum_{i=1}^{m}|g_i(x)|^2\leq-\beta|x|^2+a_1, \qquad \forall x\in\mathbb{R}^n, \qquad (3.37)$$

where β and a_1 are positive constants.

Theorem 3.5.1 *Let (3.4) and (3.37) hold, then the equation (3.1) is asymptotically bounded in second moment*

$$\limsup_{t\to\infty}\mathbb{E}(|x(t)|^2)\leq\frac{a_1}{\beta}, \quad \forall x(0)\in\mathbb{R}^n. \qquad (3.38)$$

We refer to Chapter 5 of (Mao & Yuan, 2006) for the proof.

Now we consider to reproduce this boundedness property by the EM method.

Theorem 3.5.2 *Let (3.4) and (3.37) hold, then for any* $\Delta t < 2\beta/K$ *the EM solution (3.2) satisfies*

$$\limsup_{k\to\infty} \mathbb{E}|Y_k|^2 \le \frac{2a_1 + \alpha\Delta t}{2\beta - K\Delta t}, \quad \forall Y_0 \in \mathbb{R}^n.$$

Moreover, let the stepsize tend to zero, then

$$\lim_{\Delta t\to 0} \limsup_{k\to\infty} \mathbb{E}|Y_k|^2 \le \frac{a_1}{\beta}, \quad \forall Y_0 \in \mathbb{R}^n. \tag{3.39}$$

Proof. Since $\Delta B_{i,k}$, $i = 1, ..., m$, is independent from each other, we have $\mathbb{E}(\Delta B_{i,k}) = 0$, $\mathbb{E}((\Delta B_{i,k})^2) = \Delta t$ and $\mathbb{E}(\Delta B_{i,k}\Delta B_{j,k}) = \mathbb{E}(\Delta B_{i,k})\mathbb{E}(\Delta B_{j,k}) = 0$, for $i \neq j$. Taking square and expectation on both sides of the EM solution (3.2), we have

$$\begin{aligned}
\mathbb{E}|Y_{k+1}|^2 &\le \mathbb{E}|Y_k|^2 + \Delta t^2\mathbb{E}(|f(Y_k)|^2) + \Delta t\mathbb{E}(2\langle Y_k, f(Y_k)\rangle + \sum_{i=1}^{m} |g_i(Y_k)|^2) \\
&\le \mathbb{E}|Y_k|^2 + \Delta t^2(K\mathbb{E}|Y_k|^2 + \alpha) + \Delta t(-2\beta\mathbb{E}|Y_k|^2 + 2a_1) \\
&\le (1 - 2\beta\Delta t + K\Delta t^2)\mathbb{E}|Y_k|^2 + (\alpha\Delta t^2 + 2a_1\Delta t).
\end{aligned}$$

By iteration, we see

$$\mathbb{E}|Y_{k+1}|^2 \le (1-2\beta\Delta t+K\Delta t^2)^{k+1}\mathbb{E}|Y_0|^2+(\alpha\Delta t^2+2a_1\Delta t)\frac{1-(1-2\beta\Delta t+K\Delta t^2)^{k+1}}{1-(1-2\beta\Delta t+K\Delta t^2)}.$$

Choosing $\Delta t < 2\beta/K$, we have $1 - 2\beta\Delta t + K\Delta t^2 < 1$. Let k tend to infinity and Δt tend to 0, the assertion holds.

It is interesting to see that for the case of second moment, the EM method can reproduce not only the boundedness property but also the upper bound accurately, that is the upper bounds in (3.38) and (3.39) are identical. From this point of view, the result for the second moment is better than that for the small moment. However, it should be noticed that condition (3.37) is stronger than condition (3.5).

3.5.2 The Backward EM Method

To relax the constraint on the drift coefficient, we replace the linear growth condition by the one-sided Lipschitz condition

$$\langle x, f(x)\rangle \le -\eta |x|^2 + a_2, \qquad \forall x \in \mathbb{R}^n, \tag{3.40}$$

where η and a_2 are positive constant. We still need the linear growth condition (3.22) on the diffusion coefficient. For the asymptotic boundedness of the second moment of the underlying solution we state another theorem as follows and refer to Chapter 5 of (Mao & Yuan, 2006) for the proof.

Theorem 3.5.3 *Let (3.22) and (3.40) hold. If* $2\eta > mK$*, the equation (3.1) is asymptotically bounded in second moment*

$$\limsup_{t\to\infty} \mathbb{E}(|x(t)|^2) \le \frac{2a_2 + m\alpha}{2\eta - mK}, \quad \forall x(0) \in \mathbb{R}^n. \tag{3.41}$$

However, in the same spirit of Lemma 3.4.1, we see the second moment of the EM solution may blow up under condition (3.40). So we turn to the backward EM method.

Theorem 3.5.4 *Let (3.22) and (3.40) hold. If* $2\eta > mK$*, then for any* $\Delta t > 0$ *the BE solution (3.3) satisfies*

$$\limsup_{k\to\infty} \mathbb{E}|Y_k|^2 \le \frac{2a_2 + m\alpha}{2\eta - mK}, \quad \forall Y_0 \in \mathbb{R}^n. \tag{3.42}$$

Proof. Taking square on both sides of the backward EM solution, by (3.40) we obtain

$$\begin{aligned} |Y_{k+1}|^2 &= \langle Y_{k+1}, Y_k + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}\rangle + \langle Y_{k+1}, f(Y_{k+1})\Delta t\rangle \\ &\le \frac{1}{2}|Y_{k+1}|^2 + \frac{1}{2}|Y_k + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}|^2 - \eta\Delta t|Y_{k+1}|^2 + a_2\Delta t \\ &\le \frac{1}{1+2\eta\Delta t}|Y_k + \sum_{i=1}^m g_i(Y_k)\Delta B_{i,k}|^2 + \frac{2a_2\Delta t}{1+2\eta\Delta t}. \end{aligned}$$

Then taking expectation on both sides, by (3.22) we see

$$\begin{aligned}\mathbb{E}|Y_{k+1}|^2 &\leq \frac{1}{1+2\eta\Delta t}(\mathbb{E}|Y_k|^2 + mK\Delta t\mathbb{E}|Y_k|^2 + m\alpha\Delta t) + \frac{2a_2\Delta t}{1+2\eta\Delta t} \\ &\leq \frac{1+mK\Delta t}{1+2\eta\Delta t}\mathbb{E}|Y_k|^2 + \frac{(2a_2+m\alpha)\Delta t}{1+2\eta\Delta t}.\end{aligned}$$

By iteration, we have

$$\mathbb{E}|Y_{k+1}|^2 \leq \left(\frac{1+mK\Delta t}{1+2\eta\Delta t}\right)^{k+1}\mathbb{E}|Y_0|^2 + \frac{(2a_2+m\alpha)\Delta t}{1+2\eta\Delta t} \times \frac{1-((1+mK\Delta t)/(1+2\eta\Delta t))^{k+1}}{1-(1+mK\Delta t)/(1+2\eta\Delta t)}.$$

Due to $2\eta > mK$, let $k \to \infty$ the assertion holds.

We have three comments on Theorem 3.5.4.

- Compare the upper bounds in (3.41) and (3.42), we observe the backward EM method can reproduce the asymptotic upper bound of the underlying solution accurately as well.
- There is no constraint on the stepsize for the backward EM method.
- The conditions we imposed in the case of second moment are stronger than those used in the small moment in Section 3.4.

3.6 Conclusions and Future Research

In this chapter we have presented results on numerical asymptotic boundedness in both small moment and second moment. In both cases, the numerical methods are showed to be able to reproduce the asymptotic boundedness property of the underlying solution under certain conditions. It should be noted that the conditions for the small moment are weaker than those for the second moment, but better results are obtained for the second moment, that is the upper bound could be reproduced accurately and the requirement of the stepsize could be stated explicitly.

The asymptotic moment boundedness is an essential condition for the existence of stationary distribution of numerical solutions. Thus, in next chapter we are going to study the numerical stationary distribution.

Another obvious open question is in the case of small moment whether we could recover the upper bound of the true solution of the SDE accurately by using the numerical solution with carefully chosen D and Δt. Besides, although the asymptotic boundedness property for pth moment with $1 < p < 2$ could be implied by the second moment, it is still worth to investigate if there exists different (possibly weaker) condition for $p \in (1, 2)$. Also, the existence of sufficient conditions for the case of $p > 2$ is interesting for future research.

The work contained in this chapter has been published, and we refer the readers to (Liu & Mao, 2013a) for the published version.

Chapter 4

Stationary Distribution of Numerical Solutions

4.1 Introduction

As we mentioned in Chapter 1, stochastic differential equations (SDEs) have been widely used in modelling uncertain phenomena in many areas. The difficulty to find general explicit solutions to non-linear SDEs has been continuously stimulating the studies on the numerical approximations. As stated in the last section of Chapter 3, this chapter sees our further study on the asymptotic properties of the approximation solutions. Among all those different types of asymptotic properties, the asymptotic stability particularly interests most researchers, for example (Berkolaiko *et al.*, 2012; Buckwar & Sickenberger, 2012; Higham, 2000; Huang, 2012; Higham *et al.*, 2007; Mao *et al.*, 2011; Schurz, 1997; Wu *et al.*, 2010) and their references therein.

However, those stabilities mentioned above sometimes are too strong and in this case it is interesting to see if the numerical solution converges in distribution. Mao and Yuan's series papers (Mao *et al.*, 2005; Yuan & Mao, 2004; Yuan & Mao, 2005) are devoted to numerical stationary distribution of stochastic differential equations. The motivation is to avoid solving the nontrivial Kolmogorov-Fokker-

Planck equations to find the stationary distribution of underlying SDEs, in those series papers the authors proposed to use the numerical stationary distribution as an approximation to the underlying stationary distribution. Due to the simple structure and moderate computational cost (Higham, 2011), the explicit Euler–Maruyama (EM) method was used in those papers. However, the explicit EM method has its own restriction, as mentioned in (Hutzenthaler *et al.*, 2011), it may not deal well with the super-linear coefficient SDEs. Therefore, both the drift coefficient and the diffusion coefficient were required to be global Lipschitz in the series papers. Those restrictions, however, exclude many highly non-linear models, for example (Allen, 2007; Gray *et al.*, 2011; Higham, 2008) and references therein.

In this chapter, we propose the Backward Euler-Maruyama (BEM) method as the approximation solution. The BEM method, which is drift implicit, has been broadly investigated and shown better at dealing with the highly non-linear SDEs in both finite time convergence problems and asymptotic problems. We mention some of works (Higham & Kloeden, 2007; Higham *et al.*, 2002; Higham *et al.*, 2007; Hu, 1996; Mao *et al.*, 2011; Schurz, 1997; Szpruch *et al.*, 2011) and references therein. In this chapter, we are going to investigate the numerical stationary distribution of the BEM method and the convergence of the numerical stationary distribution to the underlying stationary distribution. One of our key contributions is that we release the global Lipschitz condition on the drift coefficient by assuming the one-sided Lipschitz condition instead, but we still require the global Lipschitz condition on the diffusion coefficient. We also observe that, due to the Kolmogorov-Fokker-Planck equations, the numerical stationary distributions of stochastic differential equations could be regarded as alternative numerical solutions to certain type of deterministic differential equations.

This chapter is constructed as follows. We first brief the method, definitions, conditions on the SDEs as well as other mathematical preliminaries in Section 4.2. Then, we proposed the coefficients related sufficient conditions for the existence and uniqueness of the numerical stationary distribution in Section 4.3.1. Under the same conditions, the stationary distribution of the underlying solution is discussed in Section 4.3.2. The convergence of the numerical stationary distribution is proved in Section 4.3.3. In Section 4.4, we demonstrate the theoretical results by some numerical simulations. We conclude this chapter and brief some future research in Section 4.5.

4.2 Mathematical Preliminaries

Throughout this chapter, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (that is, it is right continuous and increasing while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets).

Let $f, g : \mathbb{R}^d \to \mathbb{R}^d$. To keep symbols simple, let $B(t)$ be a scalar Brownian motion. The results in this chapter can be extended to the case of multi-dimensional Brownian motions. We consider the d-dimensional stochastic differential equation of Itô type

$$dx(t) = f(x(t))dt + g(x(t))dB(t) \tag{4.1}$$

with initial value $x(0) = x_0$.

We first assume that the drift coefficient satisfies the local Lipschitz condition and the diffusion coefficient satisfies the global Lipschitz condition.

Condition 4.2.1 *For any $h > 0$, there exists a constant $C_h > 0$ such that*

$$|f(x) - f(y)|^2 \leq C_h |x - y|^2,$$

for any $x, y \in \mathbb{R}^d$ with $\max(|x|, |y|) \leq h$.

Condition 4.2.2 *There exists a constant $\bar{K}_2 > 0$ such that*

$$|g(x) - g(y)|^2 \leq \bar{K}_2 |x - y|^2,$$

for any $x, y \in \mathbb{R}^d$.

We further impose the following condition on the the drift coefficient.

Condition 4.2.3 *Assume there exist a symmetric positive-definite matrix* $Q \in \mathbb{R}^{d\times d}$ *and a constant* $\bar{K}_1 \in \mathbb{R}$ *such that*

$$(x-y)^T Q(f(x) - f(y)) \le \bar{K}_1 (x-y)^T Q(x-y),$$

for any $x, y \in \mathbb{R}^d$.

From Condition 4.2.2 and 4.2.3, it is easy to see that for any $x \in \mathbb{R}^d$

$$x^T Q f(x) \le K_1 x^T Q x + \alpha_1, \tag{4.2}$$

and

$$|g(x)|^2 \le K_2 |x|^2 + \alpha_2, \tag{4.3}$$

with $K_2, \alpha_1, \alpha_2 > 0$ and $K_1 \in \mathbb{R}$.

4.2.1 The Backward Euler-Maruyama Method

The backward Euler-Maruyama method (BEM), also called the semi-implicit Euler method, to SDE (4.1) is defined by

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + g(X_k)\Delta B_k, \quad X(0) = x_0, \tag{4.4}$$

where $\Delta B_k = B(t_{k+1}) - B(t_k)$ is a Brownian motion increment and $t_k = k\Delta t$. We refer to (Kloeden & Platen, 1992; Milstein, 1995) for more details in numerical methods for SDEs.

Lemma 4.2.4 *Let Conditions 4.2.1, 4.2.2, 4.2.3 hold and* $\Delta t < 0.5|\bar{K}_1|^{-1}$*, the BEM solution (4.4) is well defined.*

Many papers have discussed the existence and uniqueness of the BEM solution (4.4), we therefore refer to (Mao *et al.*, 2011; Mao & Szpruch, 2013b) for the proof of the lemma above. From now on, we always assume $\Delta t < 0.5|\bar{K}_1|^{-1}$.

It is useful to write (4.4) as

$$X_{k+1} - f(X_{k+1})\Delta t = X_k + g(X_k)\Delta B_k.$$

Define a function $G : \mathbb{R}^d \to \mathbb{R}^d$ by $G(x) = x - f(x)\Delta t$. Then G has its inverse function $G^{-1} : \mathbb{R}^d \to \mathbb{R}^d$. Moreover, the BEM (4.4) can be represented as

$$X_{k+1} = G^{-1}(X_k + g(X_k)\Delta B_k). \tag{4.5}$$

Lemma 4.2.5 *Let Conditions 4.2.1, 4.2.2 and 4.2.3 hold, then*

$$\mathbb{P}(X_{k+1} \in B \big| X_k = x) = \mathbb{P}(X_1 \in B \big| X_0 = x)$$

for any Borel set $B \subset \mathbb{R}^d$.

Proof. If $X_k = x$ and $X_0 = x$, by (4.4) we see

$$X_{k+1} - f(X_{k+1})\Delta t = x + g(x)\Delta B_k,$$

and

$$X_1 - f(X_1)\Delta t = x + g(x)\Delta B_0.$$

Because ΔB_k and ΔB_0 are identical in probability law, comparing the two equations above, we know that $X_{k+1} - f(X_{k+1})$ and $X_1 - f(X_1)\Delta t$ have the identical probability law. Then, due to Lemma 4.2.4, we have that X_{k+1} and X_1 are identical in probability law under $X_k = x$ and $X_0 = x$. Therefore, the assertion holds.
To prove Theorem 4.2.7, we cite the following classical result (see, for example, Lemma 9.2 on p87 of (Mao, 2008)).

Lemma 4.2.6 *Let $h(x, \omega)$ be a scalar bounded measurable random function of x, independent of $\mathcal{F}_s$. Let ζ be an $\mathcal{F}_s$-measurable random variable. Then*

$$\mathbb{E}(h(\zeta, \omega)\big|\mathcal{F}_s) = H(\zeta),$$

where $H(x) = \mathbb{E}h(x, \omega)$.

For any $x \in \mathbb{R}^d$ and any Borel set $B \subset \mathbb{R}^d$, define

$$\mathbb{P}(x,B) := \mathbb{P}(X_1 \in B \big| X_0 = x) \text{ and } \mathbb{P}_k(x,B) := \mathbb{P}(X_k \in B \big| X_0 = x).$$

Theorem 4.2.7 *The BEM solution (4.4) is a homogeneous Markov process with transition probability kernel $\mathbb{P}(x,B)$.*

Proof. The homogeneous property follows Lemma 4.2.5, so we only need to show the Markov property. Define

$$Y^x_{k+1} = G^{-1}(x + g(x)\Delta B_k),$$

for $x \in \mathbb{R}^d$ and $k \geq 0$. By (4.5) we know that $X_{k+1} = Y^{X_k}_{k+1}$. Let $\Gamma_{t_{k+1}} = \sigma\{B(t_{k+1}) - B(t_k)\}$. Clearly, $\Gamma_{t_{k+1}}$ is independent of $\mathcal{F}_{t_k}$. Moreover, Y^x_{k+1} depends completely on the increment $B(t_{k+1}) - B(t_k)$, so is $\Gamma_{t_{k+1}}$-measurable. Hence, Y^x_{k+1} is independent of $\mathcal{F}_{t_k}$. Applying Lemma 4.2.6 with $h(x,\omega) = I_B(Y^x_{k+1})$, we compute that

$$\begin{aligned}\mathbb{P}(X_{k+1} \in B \big| \mathcal{F}_{t_k}) &= \mathbb{E}(I_B(X_{k+1}) \big| \mathcal{F}_{t_k}) = \mathbb{E}\left(I_B(Y^{X_k}_{k+1}) \big| \mathcal{F}_{t_k}\right) = \mathbb{E}\left(I_B(Y^x_{k+1})\right)\big|_{x=X_k} \\ &= \mathbb{P}(x,B)\big|_{x=X_k} = \mathbb{P}(X_k,B) = \mathbb{P}(X_{k+1} \in B \big| X_k).\end{aligned}$$

The proof is complete.

Therefore, we see that $\mathbb{P}(\cdot,\cdot)$ is the one-step transition probability and $\mathbb{P}_k(\cdot,\cdot)$ is the k-step transition probability, both of which are induced by the BEM solution.

4.2.2 Stationary Distributions

Denote the family of all probability measures on $\mathbb{R}^d$ by $\mathcal{P}(\mathbb{R}^d)$. Define by $\mathbb{L}$ the family of mappings $F : \mathbb{R}^d \to \mathbb{R}$ satisfying

$$|F(x) - F(y)| \leq |x - y| \quad \text{and} \quad |F(x)| \leq 1,$$

for any $x, y \in \mathbb{R}^d$. For $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}(\mathbb{R}^d)$, define metric $d_{\mathbb{L}}$ by

$$d_{\mathbb{L}}(\mathbb{P}_1,\mathbb{P}_2) = \sup_{F \in \mathbb{L}} \left| \int_{\mathbb{R}^d} F(x)\mathbb{P}_1(dx) - \int_{\mathbb{R}^d} F(x)\mathbb{P}_2(dx) \right|.$$

The weak convergence of probability measures can be illustrated in term of metric $d_{\mathbb{L}}$ (Ikeda & Watanabe, 1981). That is to say, a sequence of probability measures $\{\mathbb{P}_k\}_{k\geq 1}$ in $\mathcal{P}(\mathbb{R}^d)$ converges weakly to a probability measure $\mathbb{P} \in \mathcal{P}(\mathbb{R}^d)$ if and only if

$$\lim_{k\to\infty} d_{\mathbb{L}}(\mathbb{P}_k, \mathbb{P}) = 0.$$

Then we define the stationary distribution for $\{X_k\}_{k\geq 0}$ by using the concept of weak convergence.

Definition 4.2.8 *For any initial value $x \in \mathbb{R}^d$ and a given step size $\Delta t > 0$, $\{X_k\}_{k\geq 0}$ is said to have a stationary distribution $\Pi_{\Delta t} \in \mathcal{P}(\mathbb{R}^d)$ if the k-step transition probability measure $\mathbb{P}_k(x,\cdot)$ converges weakly to $\Pi_{\Delta t}(\cdot)$ as $k \to \infty$ for every $x \in \mathbb{R}^d$, that is*

$$\lim_{k\to\infty}\left(\sup_{F\in\mathbb{L}} |\mathbb{E}(F(X_k)) - E_{\Pi_{\Delta t}}(F)|\right) = 0,$$

where

$$E_{\Pi_{\Delta t}}(F) = \int_{\mathbb{R}^d} F(y)\Pi_{\Delta t}(dy).$$

In (Yuan & Mao, 2005), the authors provided the following three assumptions and proved that under those assumptions the Euler–Maruyama solution of stochastic differential equation has a unique stationary distribution. We observe that the three assumptions are very general and actually can cover many other types of one-step numerical methods including the BEM method. This is because that, in their proofs (Theorem 3.1 in (Yuan & Mao, 2005)), only the three assumptions were required but not the structure of the numerical method. Therefore, for any one-step numerical solution that is a homogeneous Markov process with a proper transition probability kernel and satisfies the three assumptions, Theorem 3.1 in (Yuan & Mao, 2005) always holds. To keep this chapter self-contained, we state the assumptions and the theorem as follows.

Assumption 4.2.9 *For any $\varepsilon > 0$ and $x_0 \in \mathbb{R}^d$, there exists a constant $R = R(\varepsilon, x_0) > 0$ such that*

$$\mathbb{P}(|X_k^{x_0}| \geq R) < \varepsilon, \quad \textit{for any } k \geq 0.$$

Assumption 4.2.10 *For any* $\varepsilon > 0$ *and any compact subset* K *of* $\mathbb{R}^d$*, there exists a positive integer* $k^* = k^*(\varepsilon, K)$ *such that*

$$\mathbb{P}(|X_k^{x_0} - X_k^{y_0}| < \varepsilon) \geq 1 - \varepsilon, \quad \textit{for any } k \geq k^* \textit{ and any } (x_0, y_0) \in K \times K.$$

Assumption 4.2.11 *For any* $\varepsilon > 0$*,* $n \geq 1$ *and any compact subset* K *of* $\mathbb{R}^d$*, there exists a* $R = R(\varepsilon, n, K) > 0$ *such that*

$$\mathbb{P}\left(\sup_{0 \leq k \leq n} |X_k^{x_0}| \leq R\right) > 1 - \varepsilon, \quad \textit{for any } x_0 \in K.$$

Theorem 4.2.12 *Under Assumptions 4.2.9, 4.2.10 and 4.2.11, the BEM solution* $\{X_k\}_{k \geq 0}$ *has a unique stationary distribution* $\Pi_{\Delta t}$*.*

We refer the readers to Theorem 3.1 in (Yuan & Mao, 2005) for the proof. However, those three assumptions are not easy to check as they are not directly related to the drift and diffusion coefficients of the underlying SDEs. In next section, we will provide some coefficients-related sufficient conditions for those assumptions. It should be noted that those sufficient conditions are method related, which makes them more constrained.

4.3 Main Results

This section is divided into three parts. In the first subsection, we propose three lemmas that are sufficient conditions for Assumption 4.2.9, 4.2.10 and 4.2.11. Then by Theorem 4.2.12, we see that the BEM solution has a unique stationary distribution. In the second subsection, we prove that given the same conditions in the three lemmas the underlying solution has a unique stationary distribution as well. The last subsection is devoted to the convergence of the numerical stationary distribution to the underlying stationary distribution.

4.3.1 Sufficient Conditions for the Numerical Stationary Distribution

Many works have discussed the second moment boundedness of the BEM solution in finite time, we only mention a few of them here (Kloeden & Platen, 1992; Mao & Szpruch, 2013b) and references therein. It should be emphasized that, comparing with techniques employed in Lemma 4.3.1, weaker conditions and more complicated techniques have already been developed in existing literature. But those weaker conditions may not be sufficient for other lemmas in this chapter. To keep the conditions consistent in this chapter and to make the chapter self-contained, we brief the following lemma. Without confusion, in some of the proofs we omit the superscript and simply denote $X_k^{x_0}$ by X_k.

Lemma 4.3.1 *Given Conditions 4.2.1, 4.2.2 and 4.2.3, the second moment of the BEM solution (4.4) obeys*

$$\mathbb{E}\left(\sup_{0\le k\le n}|X_k|^2\right) \le q\left(|x_0|^2 + C_1(n+1)\left(2\alpha_1\Delta t + \alpha_2\Delta t + 2\sqrt{2\alpha_2\Delta t/\pi}\right)\right)$$
$$\times \exp\left(q(n+1)C_1\left(1 + K_2\Delta t + 2(\sqrt{K_2}+\sqrt{\alpha_2})\sqrt{2\Delta t/\pi}\right)\right)$$

for any integer $n \ge 1$, where $C_1 = (1-2|K_1|\Delta t)^{-1}$ and $q = \lambda_{\max}(Q)/\lambda_{\min}(Q)$.

Proof. Fix any initial value $X(0) = x_0 \in \mathbb{R}^d$, from (4.4) we see that

$$X_{k+1}^T Q X_{k+1} = X_{k+1}^T Q(X_k + g(X_k)\Delta B_k) + X_{k+1}^T Q f(X_{k+1})\Delta t.$$

Since Q is a symmetric positive-definite matrix, by Cholesky decomposition there exists an unique lower triangular matrix L such that $Q = LL^T$. Then by applying the elementary inequality, Cauchy-Schwarz inequality and (4.2) we have

$$\begin{aligned} X_{k+1}^T Q X_{k+1} &\le \frac{1}{2}|X_{k+1}^T L|^2 + \frac{1}{2}|L^T(X_k + g(X_k)\Delta B_k)|^2 + (K_1 X_{k+1}^T Q X_{k+1} + \alpha_1)\Delta t \\ &\le \frac{1}{2}X_{k+1}^T Q X_{k+1} + \frac{1}{2}[X_k^T Q X_k + g^T(X_k) Q g(X_k)|\Delta B_k|^2 + 2X_k^T Q g(X_k)\Delta B_k] \\ &\quad + (K_1 X_{k+1}^T Q X_{k+1} + \alpha_1)\Delta t. \end{aligned}$$

This implies

$$X_{k+1}^T Q X_{k+1} \le C_1(X_k^T Q X_k + g^T(X_k)Qg(X_k)|\Delta B_k|^2 + 2X_k^T Qg(X_k)\Delta B_k) + 2C_1\alpha_1\Delta t,$$

where $C_1 = (1-2|K_1|\Delta t)^{-1}$. Taking sum on both sides gives

$$\begin{aligned} X_{k+1}^T Q X_{k+1} \le X_0^T Q X_0 + (C_1 - 1)\sum_{i=0}^{k}(X_i^T Q X_i) + 2\alpha_1 C_1 (k+1)\Delta t \\ + C_1 \sum_{i=0}^{k}(2X_i^T Qg(X_i)\Delta B_i + g^T(X_i)Qg(X_i)|\Delta B_i|^2). \end{aligned} \tag{4.6}$$

It is not difficult to show that

$$\mathbb{E}\left(\sup_{0\le k\le n}\left(\sum_{i=0}^{k} g^T(X_i)Qg(X_i)|\Delta B_i|^2\right)\right) \le \Delta t \lambda_{\max}(Q)\sum_{i=0}^{n}\mathbb{E}(K_2|X_i|^2 + \alpha_2),$$

and

$$\begin{aligned} &\mathbb{E}\left(\sup_{0\le k\le n}\left(\sum_{i=0}^{k} X_i^T Qg(X_i)\Delta B_i\right)\right) \le \lambda_{\max}(Q)\mathbb{E}\left(\sum_{i=0}^{n}|X_i||g(X_i)||\Delta B_i|\right) \\ &\le \lambda_{\max}(Q)(\sqrt{K_2}+\sqrt{\alpha_2})\sqrt{2\Delta t/\pi}\sum_{i=0}^{n}\mathbb{E}(|X_i|^2) + \lambda_{\max}(Q)\sqrt{2\alpha_2\Delta t/\pi}(n+1), \end{aligned}$$

where $\mathbb{E}|\Delta B_i| = \sqrt{2\Delta t/\pi}$ is used. Therefore, taking supremum and expectation on both sides of (4.6) yields

$$\begin{aligned} \mathbb{E}\left(\sup_{0\le k\le n}|X_k|^2\right) \le \frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)}\Bigg(|x_0|^2 + C_1(n+1)\left(2\alpha_1\Delta t + \alpha_2\Delta t + 2\sqrt{2\alpha_2\Delta t/\pi}\right) \\ + C_1\left(1 + K_2\Delta t + 2(\sqrt{K_2}+\sqrt{\alpha_2})\sqrt{2\Delta t/\pi}\right)\sum_{i=0}^{n}\mathbb{E}\left(\sup_{0\le k\le i}|X_k|^2\right)\Bigg). \end{aligned}$$

Then, using the discrete-type Gronwall inequality (see, for example, (Mao, 1991)) we see the assertion holds.

From Lemma 4.3.1, by the Chebyshev inequality we can conclude that Assumption 4.2.11 holds under Conditions 4.2.1, 4.2.2 and 4.2.3.

Lemma 4.3.2 *Let (4.2) and (4.3) hold. If, for the same Q in (4.2), there exists a positive constant D such that for any $x \in \mathbb{R}^d$*

$$\frac{g^T(x)Qg(x)}{D + x^T Q x} - \frac{2|x^T Qg(x)|^2}{(D + x^T Q x)^2} \le K_3 + \frac{P_1(|x|)}{D + x^T Q x} + \frac{P_3(|x|)}{(D + x^T Q x)^2} \tag{4.7}$$

where K_3 is a constant with $K_1 + 0.5K_3 < 0$ and $P_i(|x|)$ is a polynomial of $|x|$ with degree i, then there exists a pair of constants $(p^*, \Delta t^*)$ with $p^* \in (0,1)$ and $\Delta t^* \in (0, 0.5|K_1|^{-1})$ such that for any $p \in (0, p^*)$ and any $\Delta t \in (0, \Delta t^*)$ the BEM solution (4.4) has the property that for any $k \geq 1$

$$\mathbb{E}|X_k|^p \leq q(D^{p/2} + |X_0|^p - 2C_3'(p(K_1 + 0.5K_3))^{-1})$$

where $q = \lambda_{\max}(Q)/\lambda_{\min}(Q)$, and C_3' depends on K_1, α_1, D, Q and p.

Proof. Set $C_1 = (1 - 2K_1\Delta t)^{-1}$, from the proof of Lemma 4.3.1 we have that

$$\begin{aligned}
&DC_1 + X_{k+1}^T Q X_{k+1} \\
&\leq DC_1 + C_1(X_k^T Q X_k + 2X_k^T Q g(X_k)\Delta B_k + g^T(X_k) Q g(X_k)|\Delta B_k|^2 + 2\alpha_1 \Delta t) \\
&\leq C_1(D + X_k^T Q X_k)(1 + \zeta_k),
\end{aligned}$$

where $\zeta_k = (D + X_k^T Q X_k)^{-1}(2X_k^T Q g(X_k)\Delta B_k + g^T(X_k) Q g(X_k)|\Delta B_k|^2 + 2\alpha_1\Delta t)$. Clearly $\zeta_k > -1$. For any $p \in (0,1)$, thanks to the fundamental inequality that

$$(1+u)^{p/2} \leq 1 + \frac{p}{2}u + \frac{p(p-2)}{8}u^2 + \frac{p(p-2)(p-4)}{2^3 \times 3!}u^3, \quad u > -1, \tag{4.8}$$

we see that

$$\begin{aligned}
&\mathbb{E}((D + X_{k+1}^T Q X_{k+1})^{p/2} \big| \mathcal{F}_{k\Delta t}) \\
&\leq C_1^{p/2}(D + X_k^T Q X_k)^{p/2} \mathbb{E}\left(1 + \frac{p}{2}\zeta_k + \frac{p(p-2)}{8}\zeta_k^2 + \frac{p(p-2)(p-4)}{2^3 \times 3!}\zeta_k^3 | \mathcal{F}_{k\Delta t}\right).
\end{aligned} \tag{4.9}$$

Since ΔB_k is independent of $\mathcal{F}_{k\Delta t}$, we have that $\mathbb{E}(\Delta B_k \big| \mathcal{F}_{k\Delta t}) = \mathbb{E}(\Delta B_k) = 0$ and $\mathbb{E}(|\Delta B_k|^2 \big| \mathcal{F}_{k\Delta t}) = \mathbb{E}(|\Delta B_k|^2) = \Delta t$. Then

$$\begin{aligned}
&\mathbb{E}(\zeta_k \big| \mathcal{F}_{k\Delta t}) \\
&= \mathbb{E}\left((D + X_k^T Q X_k)^{-1}(2X_k^T Q g(X_k)\Delta B_k + g^T(X_k) Q g(X_k)|\Delta B_k|^2 + 2\alpha_1\Delta t) \big| \mathcal{F}_{k\Delta t}\right) \\
&= (D + X_k^T Q X_k)^{-1}(2X_k^T Q g(X_k)\mathbb{E}(\Delta B_k \big| \mathcal{F}_{k\Delta t}) + g^T(X_k) Q g(X_k)\mathbb{E}(|\Delta B_k|^2 \big| \mathcal{F}_{k\Delta t}) \\
&\quad + 2\alpha_1\Delta t) \\
&= (D + X_k^T Q X_k)^{-1}(g^T(X_k) Q g(X_k)\Delta t + 2\alpha_1\Delta t).
\end{aligned} \tag{4.10}$$

Using the facts that $\mathbb{E}(|\Delta B_k|^{2i}) = (2i-1)!!\Delta t^i$ and $\mathbb{E}((\Delta B_k)^{2i+1}) = 0$, where $(2i-1)!!$ denotes the double factorial, i.e. $(2i-1)!! = (2i-1)(2i-3)\cdots 3\cdot 1$, similarly we get that

$$\begin{aligned}\mathbb{E}(\zeta_k^2|\mathcal{F}_{k\Delta t}) &= (D+X_k^TQX_k)^{-2}(4|X_k^TQg(X_k)|^2\Delta t + 3|g^T(X_k)Qg(X_k)|^2\Delta t^2 + 4\alpha_1^2\Delta t^2 \\ &\quad + 4\alpha_1 g^T(X_k)Qg(X_k)\Delta t^2) \\ &\geq (D+X_k^TQX_k)^{-2}(4|X_k^TQg(X_k)|^2\Delta t), \end{aligned} \tag{4.11}$$

and

$$\begin{aligned}\mathbb{E}(\zeta_k^3|\mathcal{F}_{k\Delta t}) &= (D+X_k^TQX_k)^{-3}(15|g^T(X_k)Qg(X_k)|^3\Delta t^3 + 12\alpha_1^2 g^T(X_k)Qg(X_k)\Delta t^3 \\ &\quad + 8\alpha_1^3\Delta t^3 + 24\alpha_1|X_k^TQg(X_k)|^2\Delta t^2 + 18\alpha_1|g^T(X_k)Qg(X_k)|^2\Delta t^3 \\ &\quad + 36|X_k^TQg(X_k)|^2 g^T(X_k)Qg(X_k)\Delta t^2) \\ &\leq C_2\Delta t^2, \end{aligned} \tag{4.12}$$

where C_2 is a constant dependent on K_2, α_1, α_2, $\lambda_{\max}(Q)$, $\lambda_{\min}(Q)$ and D.

Substituting (4.10), (4.11) and (4.12) back to (4.9) yields

$$\begin{aligned}&\mathbb{E}((D+X_{k+1}^TQX_{k+1})^{p/2}|\mathcal{F}_{k\Delta t}) \\ &\leq C_1^{p/2}(D+X_k^TQX_k)^{p/2}\mathbb{E}\Bigg(1+\frac{p}{2}\left(\frac{g^T(X_k)Qg(X_k)}{D+X_k^TQX_k} - \frac{2|X_k^TQg(X_k)|^2}{(D+X_k^TQX_k)^2}\right)\Delta t \\ &+\frac{p^2}{2}\frac{|X_k^TQg(X_k)|^2}{(D+X_k^TQX_k)^2}\Delta t + \frac{p(p-2)(p-4)}{2^3\times 3!}C_2\Delta t^2\Bigg) + C_3\Delta t\end{aligned}$$

where C_3 depends on K_1, α_1, D, $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$. Considering the following two fractions,

$$\frac{(D+X_k^TQX_k)^{p/2}P_1(|X_k|)}{D+X_k^TQX_k} \quad \text{and} \quad \frac{(D+X_k^TQX_k)^{p/2}P_3(|X_k|)}{(D+X_k^TQX_k)^2}.$$

For $0<p<1$ the highest degrees of $|X_k|$ in the numerators are $p+1$ and $p+3$ respectively, which are smaller than the corresponding highest degrees of $|X_k|$ in the denominators. Thus, for any $|X_k|\in\mathbb{R}$ there exists an upper bound for both of the fractions. By (4.7), we have

$$\begin{aligned}&\mathbb{E}((D+X_{k+1}^TQX_{k+1})^{p/2}|\mathcal{F}_{k\Delta t}) \\ &\leq C_1^{p/2}(D+X_k^TQX_k)^{p/2}(1+\frac{p}{2}K_3\Delta t + \frac{p^2}{2}K_2 q\Delta t + C_2'\Delta t^2) + C_3'\Delta t\end{aligned}$$

where C_2' depends C_2 and p, and C_3' depends C_3 and p. Taking expectation on both sides, we have

$$\begin{aligned}&\mathbb{E}((D+X_{k+1}^TQX_{k+1})^{p/2})\\ &\leq C_1^{p/2}(1+\frac{p}{2}K_3\Delta t+\frac{p^2}{2}K_2q\Delta t+C_2'\Delta t^2)\mathbb{E}((D+X_k^TQX_k)^{p/2})+C_3'\Delta t.\end{aligned}\tag{4.13}$$

Set $\varepsilon=0.5|K_1+0.5K_3|$, choose p^* sufficient small such that $p^*K_2q\leq 0.25\varepsilon$, then choose Δt^* sufficient small such that for $p\in(0,p^*)$ and $\Delta t\in(0,\Delta t^*)$ we have

$$C_1=(1-2K_1\Delta t)^{-1}\geq 1-pK_1\Delta t-C_4\Delta t^2>0,\tag{4.14}$$

where C_4 is a positive constant dependent on K_1 and p. By further reducing Δt^* such that for any $\Delta t\in(0,\Delta t^*)$

$$C_2'\Delta t<\frac{1}{8}p\varepsilon,\quad C_4\Delta t<\frac{1}{4}\varepsilon,\quad |p(K_1+\frac{1}{4}\varepsilon)\Delta t|<\frac{1}{2}.$$

Now using these three inequalities and (4.14), we derive from (4.13) that

$$\mathbb{E}((D+X_{k+1}^TQX_{k+1})^{p/2})\leq\frac{1+0.5p(K_3+0.5\varepsilon)\Delta t}{1-p(K_1+0.25\varepsilon)\Delta t}\mathbb{E}((D+X_k^TQX_k)^{p/2})+C_3'\Delta t.\tag{4.15}$$

Considering the estimate that for any $\kappa\in[-0.5,0.5]$

$$(1-\kappa)^{-1}=1+\kappa+\kappa^2\sum_{i=0}^{\infty}\kappa^i\leq 1+\kappa+\kappa^2\sum_{i=0}^{\infty}0.5^i=1+\kappa+2\kappa^2,$$

by further reducing Δt^* we see that for $\Delta t\in(0,\Delta t)$

$$4p(K_1+\frac{1}{4}\varepsilon)^2\Delta t+(K_3+\frac{1}{2}\varepsilon)(p(K_1+\frac{1}{4}\varepsilon)\Delta t+2(p(K_1+\frac{1}{4}\varepsilon)\Delta t)^2)<\varepsilon.$$

Then (4.15) indicates that

$$\begin{aligned}\mathbb{E}((D+X_{k+1}^TQX_{k+1})^{p/2})&\leq(1+0.5p(K_3+0.5\varepsilon)\Delta t)(1+p(K_1+0.25\varepsilon)\Delta t\\ &\quad+2(p(K_1+0.25\varepsilon)\Delta t)^2)\mathbb{E}((D+X_k^TQX_k)^{p/2})+C_3'\Delta t\\ &\leq(1+p(K_1+0.5K_3+\varepsilon)\Delta t)\mathbb{E}((D+X_k^TQX_k)^{p/2})+C_3'\Delta t.\end{aligned}$$

By iteration, we obtain that

$$\mathbb{E}((D + X_{k+1}^T Q X_{k+1})^{p/2}) \le (1 + p(K_1 + 0.5K_3 + \varepsilon)\Delta t)^{k+1}(D + X_0^T Q X_0)^{p/2} + \frac{1 - (1 + p(K_1 + 0.5K_3 + \varepsilon)\Delta t)^{k+1}}{1 - (1 + p(K_1 + 0.5K_3 + \varepsilon)\Delta t)} C_3' \Delta t.$$

Since $(1 + p(K_1 + 0.5K_3 + \varepsilon)\Delta t) \in (0, 1)$ for any $p \in (0, p^*)$ and $\Delta t \in (0, \Delta t^*)$, we see that

$$\mathbb{E}((D + X_{k+1}^T Q X_{k+1})^{p/2}) \le (D + X_0^T Q X_0)^{p/2} - 2(p(K_1 + 0.5K_3))^{-1} C_3'.$$

Because Q is a symmetric positive-definite matrix, the assertion holds.
From Lemma 4.3.2, we can conclude that Assumption 4.2.9 holds for sufficiently small Δt.

Now we are investigating the sufficient condition for Assumption 4.2.10. The techniques used in the proof of Lemma 4.3.3 are similar to those in Lemma 4.3.2.

Lemma 4.3.3 *Let Conditions 4.2.1, 4.2.2 and 4.2.3 hold. Assume that, for the same Q in (4.2.3) and any $x, y \in \mathbb{R}^d$ with $x \neq y$,*

$$\frac{(g(x) - g(y))^T Q (g(x) - g(y))}{(x - y)^T Q (x - y)} - \frac{2|(x - y)^T Q (g(x) - g(y))|^2}{|(x - y)^T Q (x - y)|^2} \le K_4, \qquad (4.16)$$

where K_4 is constant with $\bar{K}_1 + 0.5K_4 < 0$. Then for any two different initial values $x, y \in \mathbb{R}^d$, the BEM solution (4.4) has the property that for any $k \ge 1$ there are sufficiently small Δt^* and p^* such that for any pair of Δt and p with $\Delta t \in (0, \Delta t^*)$ and $p \in (0, p^*)$

$$\mathbb{E}(|X_k^x - X_k^y|^p) \le q(1 + 0.5p(\bar{K}_1 + 0.5K_4)\Delta t)^k \mathbb{E}(|x - y|^p),$$

where $q = \lambda_{\max}(Q)/\lambda_{\min}(Q)$. Therefore, Assumption 4.2.10 follows.

Proof. From (4.4) we have

$$X_{k+1}^x - X_{k+1}^y = X_k^x - X_k^y + (f(X_{k+1}^x) - f(X_{k+1}^y))\Delta t + (g(X_k^x) - g(X_k^y))\Delta B_k.$$

Then, in the similar manner as the proof of Lemma 4.3.1, we see that

$$\begin{aligned}&(X_{k+1}^x - X_{k+1}^y)^T Q (X_{k+1}^x - X_{k+1}^y)\\ &\leq (1-2\bar{K}_1\Delta t)^{-1}((X_k^x - X_k^y)^T Q (X_k^x - X_k^y) + 2(X_k^x - X_k^y)^T Q (g(X_k^x) - g(X_k^y))\Delta B_k\\ &+ (g(X_k^x) - g(X_k^y))^T Q (g(X_k^x) - g(X_k^y))|\Delta B_k|^2).\end{aligned}$$

Set

$$\eta_k = \frac{2(X_k^x - X_k^y)^T Q (g(X_k^x) - g(X_k^y))\Delta B_k + (g(X_k^x) - g(X_k^y))^T Q (g(X_k^x) - g(X_k^y))|\Delta B_k|^2}{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)}$$

we can have

$$(X_{k+1}^x - X_{k+1}^y)^T Q (X_{k+1}^x - X_{k+1}^y) \leq \frac{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)}{1-2\bar{K}_1\Delta t}(1+\eta_k).$$

Taking conditional expectation on both sides and using the fundamental inequality (4.8), for any $p \in (0,1)$ we have that

$$\begin{aligned}&\mathbb{E}(|(X_{k+1}^x - X_{k+1}^y)^T Q (X_{k+1}^x - X_{k+1}^y)|^{p/2}\big|\mathcal{F}_{k\Delta t})\\ &\leq \left|\frac{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)}{1-2\bar{K}_1\Delta t}\right|^{p/2} \mathbb{E}\left(1 + \frac{p}{2}\eta_k + \frac{p(p-2)}{8}\eta_k^2 + \frac{p(p-2)(p-4)}{2^3 \times 3!}\eta_k^3|\mathcal{F}_{k\Delta t}\right).\end{aligned} \tag{4.17}$$

It is not difficult to show that

$$\mathbb{E}(\eta_k\big|\mathcal{F}_{k\Delta t}) = \frac{(g(X_k^x) - g(X_k^y))^T Q (g(X_k^x) - g(X_k^y))}{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)}\Delta t,$$

$$\mathbb{E}(\eta_k^2\big|\mathcal{F}_{k\Delta t}) \geq \frac{4|(X_k^x - X_k^y)^T Q (g(X_k^x) - g(X_k^y))|^2}{|(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)|^2}\Delta t,$$

and

$$\mathbb{E}(\eta_k^3\big|\mathcal{F}_{k\Delta t}) \leq C_5 \Delta t^2,$$

where C_5 depends on K_2, $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$. Together with (4.16) we derive from (4.17) that

$$
\begin{aligned}
&\mathbb{E}(|(X_{k+1}^x - X_{k+1}^y)^T Q (X_{k+1}^x - X_{k+1}^y)|^{p/2} \big| \mathcal{F}_{k\Delta t}) \\
\leq& \left| \frac{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)}{1 - 2\bar{K}_1 \Delta t} \right|^{p/2} \left(1 + \frac{p}{2} \frac{(g(X_k^x) - g(X_k^y))^T Q (g(X_k^x) - g(X_k^y))}{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)} \Delta t \right. \\
&\left. + \frac{p(p-2)}{8} \frac{4|(X_k^x - X_k^y)^T Q (g(X_k^x) - g(X_k^y))|^2}{|(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)|^2} \Delta t + \frac{p(p-2)(p-4)}{2^3 \times 3!} C_5 \Delta t^2 \right) \\
\leq& \left| \frac{(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)}{1 - 2\bar{K}_1 \Delta t} \right|^{p/2} \left(1 + \frac{p}{2} K_4 \Delta t + \frac{p^2}{2} \bar{K}_2 q \Delta t + \frac{p(p-2)(p-4)}{2^3 \times 3!} C_5 \Delta t^2 \right).
\end{aligned}
$$

In the same way as in the proof of Lemma 4.3.2, we can choose sufficiently small Δt^* and p^* such that for any $p \in (0, p^*)$ and $\Delta t \in (0, \Delta t^*)$

$$
\begin{aligned}
&\mathbb{E}(|(X_{k+1}^x - X_{k+1}^y)^T Q (X_{k+1}^x - X_{k+1}^y)|^{p/2}) \\
&\leq (1 + 0.5p(\bar{K}_1 + 0.5K_4)\Delta t) \mathbb{E}(|(X_k^x - X_k^y)^T Q (X_k^x - X_k^y)|^{p/2}).
\end{aligned}
$$

Therefor, by iteration and the fact that Q is a symmetric positive-definite matrix we show the assertion.

Therefore, given the conditions in Lemma 4.3.1, 4.3.2 and 4.3.3, from Theorem 4.2.12 we conclude that there exists a unique stationary distribution for the BEM solution as time tends to infinity.

4.3.2 The Underlying Stationary Distribution

The existence and uniqueness of the stationary distribution for the underlying solution is discussed in this part under the same conditions as previous subsection.

Lemma 4.3.4 *Assume Conditions 4.2.1, 4.2.2 and 4.2.3 hold, the second moment of the solution of (4.1) satisfies*

$$
\mathbb{E}\left(\sup_{0 \leq t \leq T} |x(t)|^2 \right) \leq (1 + \mathbb{E}|x_0|^2) \exp(2T \times max(K_1 \lambda_{\max}(Q) + K_2, \alpha_1 + \alpha_2)),
$$

for any $T > 0$.

We refer the readers to Theorem 2.4.1 in (Mao, 2008) for the proof.

Lemma 4.3.5 *Assume the conditions in Lemma 4.3.2 hold, there exists a constant* $p^* \in (0,1)$ *such that for any* $p \in (0, p^*)$

$$\mathbb{E}|x(t)|^p \leq q(c_1 t + \mathbb{E}|x_0|^p + D^{p/2}) \exp\left(p\left[K_1 + \frac{1}{2}K_3 + \frac{p}{2}K_2 q\right]t\right) < \infty,$$

holds for any $t > 0$, *where* $q = \lambda_{\max}(Q)/\lambda_{\min}(Q)$ *and* c_1 *is a positive constant dependent on* p, K_1, K_2, α_1, α_2, D, $\lambda_{min}(Q)$ *and* $\lambda_{max}(Q)$.

Proof. For $p \in (0,1)$, from the Itô formula,

$$\begin{aligned}
& d|x^T(t)Qx(t) + D|^{p/2} \\
= & [p|x^T(t)Qx(t) + D|^{p/2-1}(x^T(t)Qf(x(t))) \\
& + p(\frac{p}{2} - 1)|x^T(t)Qx(t) + D|^{p/2-2}|x^T(t)Qg(x(t))|^2 \\
& + \frac{p}{2}|x^T(t)Qx(t) + D|^{p/2-1}(g^T(x(t))Qg(x(t)))]dt \\
& + p|x^T(t)Qx(t) + D|^{p/2-1}(x^T(t)Qg(x(t)))dB(t) \\
= & p|x^T(t)Qx(t) + D|^{p/2}\left[\frac{x^T(t)Qf(x(t))}{x^T(t)Qx(t) + D} + \frac{\frac{1}{2}g^T(x(t))Qg(x(t))}{x^T(t)Qx(t) + D} - \frac{|x^T(t)Qg(x(t))|^2}{|x^T(t)Qx(t) + D|^2}\right. \\
& \left. + \frac{p}{2}\frac{|x^T(t)Qg(x(t))|^2}{|x^T(t)Qx(t) + D|^2}\right]dt + p|x^T(t)Qx(t) + D|^{p/2-1}(x^T(t)Qg(x(t)))dB(t).
\end{aligned}$$

Under (4.2), (4.3) and (4.7) it implies

$$\begin{aligned}
d|x^T(t)Qx(t) + D|^{p/2} \leq & \, p|x^T(t)Qx(t) + D|^{p/2}\left[K_1 + \frac{1}{2}K_3 + \frac{p}{2}K_2 q\right]dt + c_1 dt \\
& + p|x^T(t)Qx(t) + D|^{p/2-1}(x^T(t)Qg(x(t)))dB(t),
\end{aligned}$$

where c_1 is a positive constant dependent on p, K_1, K_2, α_1, α_2, D, $\lambda_{min}(Q)$ and $\lambda_{max}(Q)$. Since $K_1 + 0.5K_3 < 0$, given $\varepsilon \in (0, |K_1 + 0.5K_3|)$ we may choose $p^* \in (0,1)$ so small that $0.5p^* K_2 q < \varepsilon$, then for any $p \in (0, p^*)$ we see that

$$\begin{aligned}
\mathbb{E}|x^T(t)Qx(t) + D|^{p/2} \leq & \, p\left[K_1 + \frac{1}{2}K_3 + \frac{p}{2}K_2 q\right]\int_0^t \mathbb{E}|x^T(s)Qx(s) + D|^{p/2}ds \\
& + c_1 t + \mathbb{E}|x_0^T Qx_0 + D|^{p/2}.
\end{aligned}$$

By Gronwall's inequality, we see that

$$\mathbb{E}|x^T(t)Qx(t) + D|^{p/2} \leq (c_1 t + \mathbb{E}|x_0^T Q x_0 + D|^{p/2}) \exp\left(p\left[K_1 + \frac{1}{2}K_3 + \frac{p}{2}K_2 q\right] t\right).$$

Although the time variable, t, appears in both the coefficient of the exponentiation term and the exponent, the choice of the p and the fact that $K_1 + 0.5K_3 < 0$ guarantee that exponentiation term decreases as t increases. Thus, the term on the right hand side of the inequality above has an upper.

Lemma 4.3.6 *Assume the conditions in Lemma 4.3.3 hold, for any two different initial values $x_0, y_0 \in \mathbb{R}^d$, there exists a constant $p^* \in (0,1)$ such that for any $p \in (0, p^*)$*

$$\mathbb{E}|x^{x_0}(t) - x^{y_0}(t))|^p \leq q\mathbb{E}|(x_0 - y_0)|^p \exp(p(\bar{K}_1 + 0.5K_4 + 0.5p\bar{K}_2 q)t),$$

where $q = \lambda_{\max}(Q)/\lambda_{\min}(Q)$.

Proof. For $p \in (0,1)$, from the Itô formula,

$$\begin{aligned}
&d|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2} \\
&= [p|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2-1}(x^{x_0}(t) - x^{y_0}(t))^T Q(f(x^{x_0}(t)) - f(x^{y_0}(t))) \\
&+ p(\frac{p}{2} - 1)|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2-2} \\
&\times |(x^{x_0}(t) - x^{y_0}(t))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))|^2 \\
&+ \frac{p}{2}|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2-1} \\
&\times (g(x^{x_0}(t)) - g(x^{y_0}(t)))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))]dt \\
&+ p|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2-1}(x^{x_0}(t) - x^{y_0}(t))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))dB(t) \\
&= p|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2}\Bigg(\frac{(x^{x_0}(t) - x^{y_0}(t))^T Q(f(x^{x_0}(t)) - f(x^{y_0}(t)))}{(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))} \\
&+ \frac{(g(x^{x_0}(t)) - g(x^{y_0}(t)))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))}{2(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))} \\
&- \frac{|(x^{x_0}(t) - x^{y_0}(t))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))|^2}{|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^2} \\
&+ \frac{p}{2}\frac{|(x^{x_0}(t) - x^{y_0}(t))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))|^2}{|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^2}\Bigg)dt \\
&+ p|(x^{x_0}(t) - x^{y_0}(t))^T Q(x^{x_0}(t) - x^{y_0}(t))|^{p/2-1}(x^{x_0}(t) - x^{y_0}(t))^T Q(g(x^{x_0}(t)) - g(x^{y_0}(t)))dB(t).
\end{aligned}$$

Under Condition 4.2.2, 4.2.3 and (4.16) this implies

$$\begin{aligned} & d|(x^{x_0}(t)-x^{y_0}(t))^TQ(x^{x_0}(t)-x^{y_0}(t))|^{p/2} \\ \le & \, p|(x^{x_0}(t)-x^{y_0}(t))^TQ(x^{x_0}(t)-x^{y_0}(t))|^{p/2}\left(\bar{K}_1+0.5K_4+0.5p\bar{K}_2q\right)dt \\ & + p|(x^{x_0}(t)-x^{y_0}(t))^TQ(x^{x_0}(t)-x^{y_0}(t))|^{p/2-1}(x^{x_0}(t)-x^{y_0}(t))^TQ(g(x^{x_0}(t))-g(x^{y_0}(t)))dB(t). \end{aligned}$$

Since $\bar{K}_1+0.5K_4<0$, given $\varepsilon\in(0,|\bar{K}_1+0.5K_4|)$ we may choose $p^*\in(0,1)$ so small that $0.5p\bar{K}_2q<\varepsilon$, then for any $p\in(0,p^*)$ we have that

$$\begin{aligned} & \mathbb{E}|(x^{x_0}(t)-x^{y_0}(t))^TQ(x^{x_0}(t)-x^{y_0}(t))|^{p/2} \\ \le & \, \mathbb{E}|(x_0-y_0)^TQ(x_0-y_0)|^{p/2} \\ & + p(\bar{K}_1+0.5K_4+0.5p\bar{K}_2q)\int_0^t \mathbb{E}|(x^{x_0}(s)-x^{y_0}(s))^TQ(x^{x_0}(s)-x^{y_0}(s))|^{p/2}ds. \end{aligned}$$

Then Gronwall's inequality indicates that

$$\begin{aligned} & \mathbb{E}|(x^{x_0}(t)-x^{y_0}(t))^TQ(x^{x_0}(t)-x^{y_0}(t))|^{p/2} \\ \le & \, \mathbb{E}|(x_0-y_0)^TQ(x_0-y_0)|^{p/2}\exp(p(\bar{K}_1+0.5K_4+0.5p\bar{K}_2q)t). \end{aligned}$$

As Q is a symmetric positive-definite matrix, the proof is complete.

We conclude this part by the following theorem.

Theorem 4.3.7 *Given the conditions in Lemma 4.3.1, 4.3.2 and 4.3.3, the solution of (4.1) has a unique stationary distribution denoted by* $\pi(\cdot)$.

Having Lemma 4.3.4, 4.3.5 and 4.3.6, the proof of this theorem follows from Theorem 3.1 in (Yuan & Mao, 2003).

4.3.3 The Convergence

Given Conditions 4.2.1, 4.2.2, 4.2.3 and those conditions assumed in Lemma 4.3.1, 4.3.2, 4.3.3, the convergence of the numerical stationary distribution to the underlying stationary distribution is discussed in this subsection.

Recall that the probability measure induced by the numerical solution, X_k, is denoted by $\mathbb{P}_k(\cdot,\cdot)$, similarly we denote the probability measure induced by the underlying solution,$x(t)$, by $\bar{\mathbb{P}}_t(\cdot,\cdot)$.

Lemma 4.3.8 *Let Conditions 4.2.1, 4.2.2, 4.2.3 hold and fix any initial value $x_0 \in \mathbb{R}^d$. Then, for any given $T > 0$ and $\varepsilon > 0$. there exists a sufficiently small $\Delta t^* > 0$ such that*

$$d_{\mathbb{L}}(\bar{\mathbb{P}}_{k\Delta t}(x_0, \cdot), \mathbb{P}_k(x_0, \cdot)) < \varepsilon$$

provided that $\Delta t < \Delta t^*$ and $k\Delta t \leq T$.

The result can be derived from the fact that the BEM solution converges strongly to the underlying solution in finite time (Higham *et al.*, 2002; Hu, 1996; Kloeden & Platen, 1992).

Now we are ready to show that the numerical stationary distribution converges to the underlying stationary distribution as time step diminishes.

Theorem 4.3.9 *Given Conditions 4.2.1, 4.2.2, 4.2.3, (4.7) and (4.16),*

$$\lim_{\Delta t \to 0} d_{\mathbb{L}}(\Pi_{\Delta t}(\cdot), \pi(\cdot)) = 0.$$

Proof. Fix any initial value $x_0 \in \mathbb{R}^d$ and set $\varepsilon > 0$ be arbitrary real number. According to Theorem 4.3.7, there exists a $\Theta^* > 0$ such that for any $t > \Theta^*$

$$d_{\mathbb{L}}(\bar{\mathbb{P}}_t(x_0, \cdot), \pi(\cdot)) < \varepsilon/3.$$

Similarly, by Theorem 4.2.12, there exists a pair of $\Delta t^{**} > 0$ and $\Theta^{**} > 0$ such that

$$d_{\mathbb{L}}(\mathbb{P}_k(x_0, \cdot), \Pi_{\Delta t}(\cdot)) < \varepsilon/3$$

for all $\Delta t < \Delta t^{**}$ and $k\Delta t > \Theta^{**}$. Let $\Theta = \max(\Theta^*, \Theta^{**})$, from Lemma 4.3.8 there exists a Δt^* such that for any $\Delta t < \Delta t^*$ and $k\Delta t < \Theta + 1$

$$d_{\mathbb{L}}(\bar{\mathbb{P}}_{k\Delta t}(x_0, \cdot), \mathbb{P}_k(x_0, \cdot)) < \varepsilon/3.$$

Therefore, for any $\Delta t < \min(\Delta t^*, \Delta t^{**})$, set $k = [\Theta/\Delta t] + 1/\Delta t$, we see the assertion holds by the triangle inequality.

4.4 Examples

In this section, we illustrate the theoretical results by three examples. First, we consider a two-dimensional SDE with scalar Brownian motion.

Example 4.4.1

$$dx(t) = (\mathrm{diag}(x_1(t), x_2(t))b + \mathrm{diag}(x_1(t), x_2(t))A\mathrm{diag}(x_1(t), x_2(t))x(t) + c_1)\, dt \\ + (\mathrm{diag}(x_1(t), x_2(t))\sigma + c_2)\, dB(t), \tag{4.18}$$

where $x(t) = (x_1(t), x_2(t))^T$, $\mathrm{diag}(x_1(t), x_2(t))$ *denotes a diagonal matrix with non-zero entries* $x_1(t)$ *and* $x_2(t)$ *on the diagonal,* $b = (1, 1)^T$, $A = (a_{ij})_{i,j=1,2}$ *with* $a_{1,1} = -1, a_{1,2} = -0.7, a_{2,1} = -1.2, a_{2,2} = -2$, $c_1 = (0.5, 0.7)^T$, $c_2 = (3.5, 4)^T$ *and* $\sigma = (3.5, 4)^T$.

Choosing Q to be identity matrix, it is clear that the drift and diffusion coefficients of (4.18) satisfy Conditions 4.2.1, 4.2.2, 4.2.3 and (4.2) with $\bar{K}_1 = 1$ and $K_1 = 1.7$, which indicates that Lemma 4.3.1 holds. To check conditions for Lemma 4.3.2, we see that

$$\frac{(3.5x_1 + 0.3)^2 + (4x_2 + 0.2)^2}{D + (x_1^2 + x_2^2)} - \frac{2|3.5x_1^2 + 0.3x_1 + 4x_2^2 + 0.2x_2|^2}{(D + (x_1^2 + x_2^2))^2}.$$

Set $D = 0.04/25$, we can derive that (4.7) is satisfied with $K_3 = -7$ and $K_1 + 0.5K_3 < 0$, then Lemma 4.3.2 holds. Finally, we have that (4.16) is satisfied with $K_4 = -7$ and $\bar{K}_1 + 0.5K_4 < 0$, that is Lemma 4.3.3 holds.

We simulate 1000 paths, each of which has 10000 iterations. In Figure 4.1, we plot one path of the BEM solution for $x_1(t)$ and $x_2(t)$. Intuitively, some stationary behaviour displays.

We further plot the empirical cumulative distribution function (ECDF) of the last iterations of the 1000 paths and the ECDF of last 1000 iterations of one path in Figure 4.2. It can be seen that the shapes and the intervals of the ECDFs are similar. To measure the similarity quantitatively, we use the Kolmogorov-Smirnov test (K-S test) (Massey J., 1951) to test the alternative hypothesis that

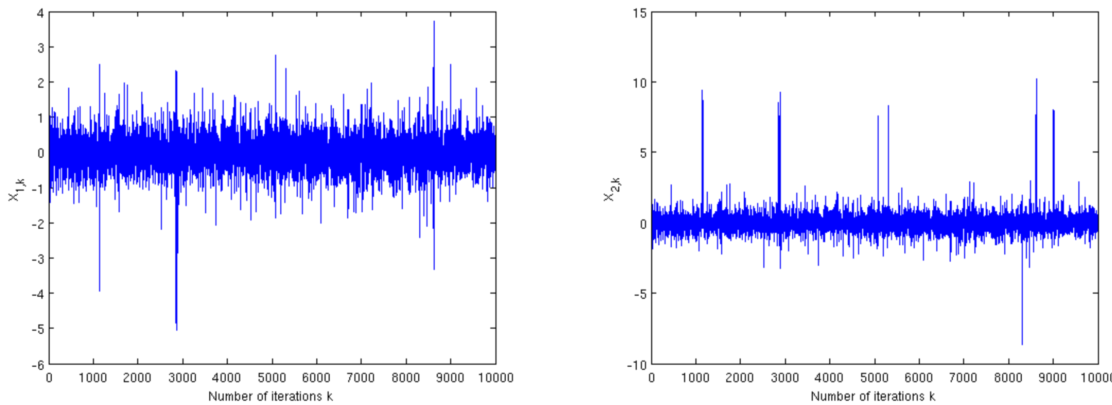

Figure 4.1: Left: the BEM solution to $x_1(t)$; Right: the BEM solution to $x_2(t)$.

the last iterations of the 1000 paths and last 1000 iterations of one path are from different distributions against the null hypothesis that they are from the same distribution for both $x_1(t)$ and $x_2(t)$. With 5% significance level, the K-S test indicates that we can not reject the null hypothesis. This example illustrates the existence of the stationary distribution as the time variable becomes large. Moreover, it may indicate that instead of simulating many paths to construct the stationary distribution, one could just use last few iterations of one path to approximate the stationary distribution.

To compare the numerical stationary distribution with the theoretical one, we next consider a nonlinear scalar SDE, whose stationary distribution can be explicitly derived from the Kolmogorov-Fokker-Planck equation.

Example 4.4.2

$$dx(t) = -0.5(x + x^3) + dB(t).$$

It is straightforward to see that $\bar{K}_1 = K_1 = -0.5$ and $K_3 = K_4 = 0$, hence all the conditions required in Section 4.2 and 4.3 are satisfied. The corresponding Kolmogorov-Fokker-Planck equation for the theoretical probability density function of the stationary distribution $p(x)$ is

$$0.5\frac{d^2p(x)}{dx^2} - \frac{d}{dx}(-0.5(x + x^3)p(x)) = 0.$$

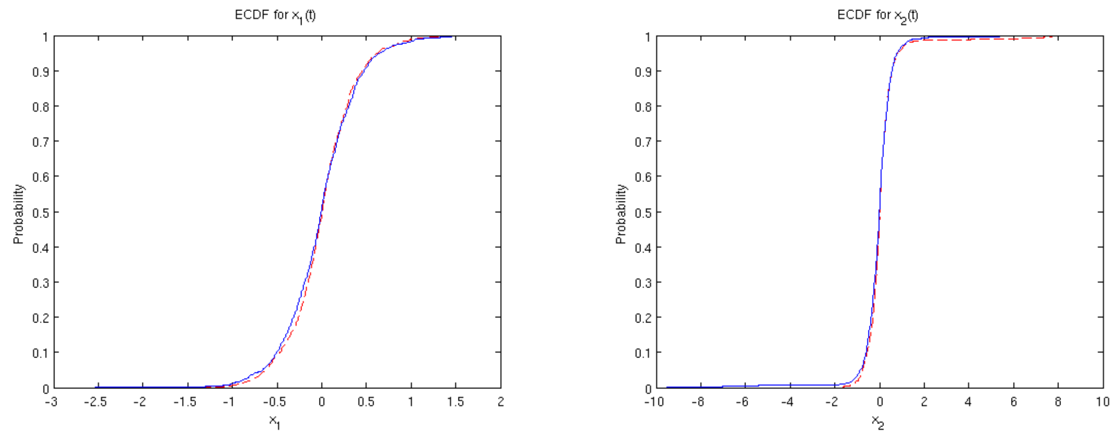

Figure 4.2: Left: ECDFs for x_1; Right: ECDFs for x_2. Red dashed line is last 1000 iterations of one path; Blue solid line is last iterations of the 1000 paths.

And the exact solution is known to be (Soong, 1973)

$$p(x) = \frac{1}{I_{\frac{1}{4}}(\frac{1}{8}) + I_{-\frac{1}{4}}(\frac{1}{8})} \exp(\frac{1}{8} - \frac{1}{2}x^2 - \frac{1}{4}x^4),$$

where $I_\nu(x)$ is modified Bessel function of the first kind. We simulate one path with 100000 iterations and plot the ECDF of last 20000 iterations in red dashed line on Figure 4.3. The theoretical cumulative distribution function is plotted on the same figure in blue solid line. The similarity of those two distribution is clear to see, which indicates the numerical stationary distribution is a good approximation to the theoretical one. The mean and variance of the numerical stationary distribution are 0 and 0.453, respectively, which are close to the theoretical counterparts 0 and 0.466.

This example also demonstrates that the numerical method for stochastic differential equations can serve as an alternative way to approximate deterministic differetial equations.

At last, we consider a linear scalar equation, the Langevin equation (Uhlenbeck & Ornstein, 1930). The comparison of the BEM method in this chapter with the EM method studied in (Yuan & Mao, 2004) demonstrates that the BEM has less constraint on the step size.

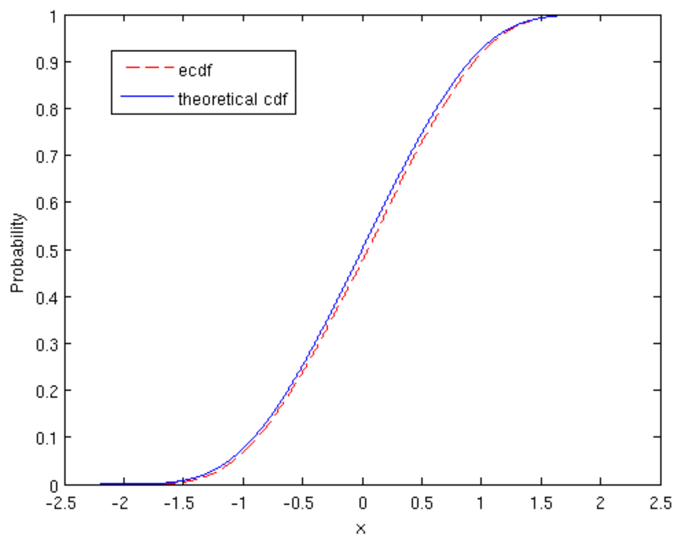

Figure 4.3: Comparison of the ECDF with the theoretical cumulative distribution function

Example 4.4.3 *We write the Itô type equation of the Langevin equation as*

$$dx(t) = -\alpha x(t)dt + \sigma dB(t) \quad \text{on } t \geq 0, \tag{4.19}$$

where $\alpha > 0$ *and* $\sigma \in \mathbb{R}$*.*

From (4.4), given the initial value $X_0 = x(0) \in \mathbb{R}$ we have

$$X_{k+1} = X_k + -\alpha X_{k+1}\Delta t + \sigma \Delta B_k.$$

This gives that X_{k+1} is normally distributed with mean

$$\mathbb{E}(X_{k+1}) = (1 + \alpha\Delta t)^{-(k+1)}x(0)$$

and variance

$$\begin{aligned} Var(X_{k+1}) &= (1 + \alpha\Delta t)^{-2}Var(X_k) + \sigma^2(1 + \alpha\Delta t)^{-2}\Delta t \\ &= \sigma^2\Delta t[(1 + \alpha\Delta t)^{-2} + (1 + \alpha\Delta t)^{-4} + ... + (1 + \alpha\Delta t)^{-2(k+1)}] \\ &= \sigma^2\Delta t\frac{1 - (1 + \alpha\Delta t)^{-(k+1)}}{(1 + \alpha\Delta t)^2 - 1} \\ &= \frac{\sigma^2}{2\alpha + \alpha^2\Delta t}. \end{aligned}$$

Therefore, the distribution of the BEM solution approaches the normal distribution $N(0, \sigma^2/(2\alpha + \alpha^2 \Delta t))$ as $k \to \infty$ for any $\Delta t > 0$. Recall, from Example 3.5.1 in (Mao, 2008), that the underlying solution of (4.19) approaches its stationary distribution $N(0, \sigma^2/(2\alpha))$ as $t \to \infty$, then it is interesting to observe that $N(0, \sigma^2/(2\alpha + \alpha^2 \Delta t))$ will further converge to stationary distribution of the true solution as $\Delta t \to 0$.

4.5 Conclusions and Future Research

This chapter largely extend the results in Mao and Yuan's series papers (Mao *et al.*, 2005; Yuan & Mao, 2004; Yuan & Mao, 2005). By using the Backward Euler-Maruyama method, the linear growth condition on the drift coefficient is released to the one-sided Lipschitz condition and the stationary distribution of many more SDEs can be approximated by the numerical stationary distribution. However, it should be mentioned that, compared to the three assumptions in Section 4.2, those sufficient conditions in Section 4.3.1 are still quite strong. And this is because that those assumptions are in probability, while those sufficient conditions are in term of moment. Therefore, it is interesting to construct some coefficients related sufficient conditions which are in probability.

Chapter 5

Almost Sure Stability with Random Variable Step Size

5.1 Introduction

To continue the study on the asymptotic properties of numerical solutions, we present some our findings on one of the most popular topics, almost sure stability. This chapter is devoted to the analyses of the almost sure stability of numerical methods for stochastic differential equations (SDEs) by using the well-known semimartingale convergence theory, see for example (Shiryaev, 1996) in terms of equalities and (Appleby *et al.*, 2006) in terms of inequalities. There are many papers that have adopted this approach to study the numerical almost sure stability, for example (Buckwar & Kelly, 2010; Mao & Szpruch, 2013a; Rodkina & Schurz, 2005; Rodkina *et al.*, 2008; Wu *et al.*, 2010; Wu *et al.*, 2011; Yu, 2011) and the references therein. However, in most of the papers the step size is either fixed or nonrandom variable.

Unlike the preceding two chapters, in which the constant step size methods are considered, the methods discussed in current chapter and the coming chapter are modified methods. In this chapter the random variable step size is introduced to embed into the Euler-Maruyama (EM) method. Our key contribution is that we prove the time variable is a stopping time. Moreover, the stopping time is essential for the application of the semimartingale convergence theory in our approach. Benefiting from the random variable step size, the sufficient conditions for the almost sure stability of the EM method obtained in this chapter are much weaker than those established in (Mao & Szpruch, 2013a) and (Wu *et al.*, 2010). To our best knowledge, this is the first work to apply the random variable step size (with clear proof of the stopping time) to the analysis of the almost sure stability of the EM method.

It should be noted that the technique of adjusting the size of each step has been broadly used in the multi-stage methods (see for example (Burrage & Burrage, 2002; Burrage *et al.*, 2004; Römisch & Winkler, 2006), and the references therein). Due to the application of the local error control technique, some steps could be rejected then smaller steps may be retreated. Since the step size in those methods is dependent on the state of the solution, it is indeed a random variable. However, the current step size may be decided after future information available and this indicates the time variable can not be a stopping time (Mauthner, 1998). In fact, not like the case in this chapter the stopping time is not necessary for those methods (Gaines & Lyons, 1997).

The Euler-type methods with the random variable step size, were also considered in different aspects, for instance in (Dávila *et al.*, 2005) to reproduce the finite time explosion of SDEs, in (Lamba *et al.*, 2007) to study convergence and ergodicity, and in (Müller-Gronbach, 2002) to optimise the error constant.

We also mention here that there are lots of other approaches to study the almost sure stability of the numerical methods for SDEs, for example by the local error control, by directly applying the the strong law of large numbers, and by the Chebyshev inequality and the Borel-Cantelli lemma the almost sure stability can be derived from the moment exponential stability. We refer to some of the works (Higham *et al.*, 2007; Lamba & Seaman, 2006; Mao *et al.*, 2011; Pang *et al.*, 2008; Schurz, 2005) and the references therein.

This chapter is constructed as follows. Section 5.2 is devoted to the mathematical notation and some preparation for the main result. In Section 5.3 we present our main result, Theorem 5.3.1, in which we demonstrate the strategy of choosing the step size, give the proof of the stopping time and conclude the almost sure stability of the EM method with random variable step size. Section 5.4 sees the computer simulations of the proposed method. In Section 5.5, alternative sufficient conditions for the numerical almost sure stability are proposed, which enable the EM method with random variable step size to cover wider range of SDEs. Proofs in the last section are only briefed as the same techniques to those in Theorem 5.3.1 are employed.

5.2 Mathematical Preliminaries

Throughout this chapter, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ which is increasing and right continuous, with $\mathcal{F}_0$ containing all $\mathbb{P}$-null sets. Let $B(t) = (B_1(t), ..., B_m(t))^T$ be an m-dimensional Brownian motion defined on the probability space. The inner product of x, y in $\mathbb{R}^n$ is denoted by $\langle x, y\rangle$.

In this chapter, we investigate the numerical methods for the n-dimensional SDE

$$dx(t) = f(x(t))dt + g(x(t))dB(t), \quad x(0) \in \mathbb{R}^n, \tag{5.1}$$

where f: $\mathbb{R}^n \to \mathbb{R}^n$ and g: $\mathbb{R}^n \to \mathbb{R}^{n\times m}$. The following two conditions are imposed on the drift and diffusion coefficients. For every integer $R \geq 1$, there exists a positive constant $C(R)$ such that, for all $x, y \in \mathbb{R}^n$ with $|x| \vee |y| \leq R$,

$$|f(x) - f(y)|^2 \vee |g(x) - g(y)|^2 \leq C(R)|x - y|^2. \tag{5.2}$$

And $\forall x \in \mathbb{R}^n$

$$-z(x) := 2\langle x, f(x)\rangle + |g(x)|^2 \leq 0. \tag{5.3}$$

From (5.3), we can see that the monotone condition holds automatically. Therefore under (5.2) and (5.3), there exists a unique solution to (5.1) for any given initial value $x(0) \in \mathbb{R}^n$ (see, for example Theorem 2.3.5 in (Mao, 2008)). The theorem for the almost sure asymptotic stability for the SDE (5.1) is presented as follows.

Theorem 5.2.1 *Let (5.2) and (5.3) hold. Assume $z(x) = 0$ if and only if $x = 0$, then for any initial value $x(0) \in \mathbb{R}^n$*

$$\lim_{t\to\infty} x(t) = 0 \quad a.s.$$

We refer to the stochastic version of the LaSalle theorem in (Shen *et al.*, 2006) for the proof of this theorem.

Lemma 5.2.2 *Assume $z(x)$, defined by (5.3), is zero if and only if $x = 0$. Then both $f(x) = 0$ and $g(x) = 0$ if $x = 0$, and $f(x) \neq 0$ if $x \neq 0$.*

Proof. We first prove $f(x) \neq 0$ if $x \neq 0$. Assume $f(\bar{x}) = 0$ for some $\bar{x} \neq 0$, then by (5.3) we have $-z(\bar{x}) = |g(\bar{x})|^2 \geq 0$. But this contradicts that $-z(x) < 0$ for $x \neq 0$.

We now prove $f(x) = 0$ if $x = 0$. Assume $f(0) \neq 0$, that is $f(0) = (f_1(0), ..., f_n(0))^T \neq 0$. Without loss of generality, we assume $f_1(0) < 0$. Due to the continuity of $f(x)$, for some sufficiently small $\varepsilon > 0$ we have $f_1(x) < 0$ for some vector x, where the first entry lies in $(-\varepsilon, \varepsilon)$ and all the rest are zeros. Then given $\bar{x} = (-\varepsilon/2, 0, ..., 0)^T$, we have $\langle \bar{x}, f(\bar{x})\rangle > 0$. But this contradicts to $-z(\bar{x}) < 0$.

Suppose $x = 0$, by (5.3) it is easy to see that $|g(0)|^2 = -z(0) = 0$, i.e. $g(0) = 0$.

The next lemma is the discrete version of the semimartingale convergence theorem (Shiryaev, 1996).

Lemma 5.2.3 *Let $\{A_i\}$ and $\{B_i\}$ be two nonnegative $\mathcal{F}_i$-measurable processes for $i = 0, 1, 2, ...$ with $A_0 = B_0 = 0$ a.s. and $\{M_i\}$ be $\mathcal{F}_i$-measurable local martingale for $i = 0, 1, 2, ...$ with $M_0 = 0$. If a nonnegative stochastic process $\{Z_i\}_{i=0,1,...}$ can be decomposed as $Z_i = Z_0 + A_i - B_i + M_i$, then*

$$\left\{\lim_{i\to\infty} A_i < \infty\right\} \subseteq \left\{\lim_{i\to\infty} B_i < \infty\right\} \cap \left\{\lim_{i\to\infty} Z_i \ \ \textit{exists and is finite}\right\} \quad a.s.$$

5.3 The EM Method with Random Variable Step Size

In this section, we present our main results about the variable step size EM method. To keep the proof simple and clear we specify the choice of the step size in the proof, but readers should notice that there are other choices. We emphasise here that there are two important properties of the variable step size that the sum of the steps is a stopping time and divergent. The feature of stopping time is essential to the proof of the local martingale term in Theorem 5.3.1, and the divergence guarantees the time is able to tend to infinity.

The first main result is that the variable step size method can reproduce the stability of the SDE shown in Theorem 5.2.1.

Theorem 5.3.1 *Let (5.2) and (5.3) hold. Assume $z(x) = 0$ if and only if $x = 0$, and*

$$\liminf_{|x|\to 0} \frac{z(x)}{|f(x)|^2} > 0. \tag{5.4}$$

Define the EM method with variable step size as

$$Y_{i+1} = Y_i + f(Y_i)\Delta t_i + g(Y_i)\Delta B_i, \quad Y_0 = x(0), \quad i \geq 0, \tag{5.5}$$

where $\Delta B_i = B(t_i) - B(t_{i-1})$ with $t_i = \sum_{k=0}^{i} \Delta t_k$ for $i = 0, 1, 2...$ and $t_{-1} = 0$, Δt_i is chosen to be 2^{-n_i} with $n_i = \lceil 1 - \log_2(z(Y_i)/|f(Y_i)|^2) \rceil$ for $|Y_i| \neq 0$ and 2^{-2} for $|Y_i| = 0$. Then t_i is an $\{\mathcal{F}_t\}$-stopping time for each $i = 0, 1, 2...$, and the sequence of time steps obeys $\sum_{i=0}^{\infty} \Delta t_i = \infty$ a.s. Moreover, for any initial value $Y_0 \in \mathbb{R}^n$

$$\lim_{i\to\infty} Y_i = 0 \quad a.s.$$

Proof. Taking square on both sides of (5.5), we have

$$\begin{aligned} |Y_{i+1}|^2 &= |Y_i|^2 + 2\langle Y_i, f(Y_i)\Delta t_i + g(Y_i)\Delta B_i\rangle + |f(Y_i)\Delta t_i + g(Y_i)\Delta B_i|^2 \\ &= |Y_i|^2 + \Delta t_i(2\langle Y_i, f(Y_i)\rangle + |g(Y_i)|^2 + |f(Y_i)|^2\Delta t_i) + \Delta m_i, \end{aligned} \tag{5.6}$$

where $\Delta m_i = 2\langle Y_i, g(Y_i)\Delta B_i\rangle + 2\langle f(Y_i)\Delta t_i, g(Y_i)\Delta B_i\rangle + |g(Y_i)|^2(|\Delta B_i|^2 - \Delta t_i)$.

The proof is divided into three parts. Firstly, we demonstrate the strategy of choosing the step size Δt_i in each time step and show that t_i is an $\{\mathcal{F}_t\}$-stopping time for every $i = 0, 1, ...$ Then we prove that $m_i = \sum_{k=0}^{i} \Delta m_k$ is a local martingale for $i = 0, 1, ...$ At last, we give the proof of the divergence of the sequence of the timesteps and conclude the almost sure stability.

Step 1

Since (5.3), in each step we can choose sufficiently small and rational step size Δt_i such that

$$-U(Y_i, \Delta t_i) := -z(Y_i) + |f(Y_i)|^2 \Delta t_i \leq 0. \tag{5.7}$$

For example, when $Y_i \neq 0$ (by Lemma 5.2.2 we know $f(Y_i) \neq 0$) we could choose $\Delta t_i = 2^{-n_i}$ with $n_i = \lceil 1 - \log_2(z(Y_i)/|f(Y_i)|^2) \rceil$. Then it is obvious that $\Delta t_i \leq z(Y_i)/(2|f(Y_i)|^2)$, thus the inequality (5.7) holds. When $Y_i = 0$ (i.e. $z(Y_i) = 0$ and $f(Y_i) = 0$), any choice of Δt_i will satisfy (5.7) and we simply choose, for example $\Delta t_i = 2^{-2}$. From the iteration (5.5), we know that if at some time point the solution becomes zero, the solution afterwards will stay at zero. Hence in this case the step size is fixed and the almost sure stability follows naturally. In the following, we focus on the case when $\Delta t_i = 2^{-n_i}$ with $n_i = \lceil 1 - \log_2(z(Y_i)/|f(Y_i)|^2) \rceil$. We emphasise here that the requirement that each Δt_i is a rational number is key to the following proof that $t_i = \sum_{k=0}^{i} \Delta t_k = \sum_{k=0}^{i} 2^{-n_k}$ is an $\{\mathcal{F}_t\}$-stopping time for every $i = 0, 1, ...$

Assume t_i is an $\{\mathcal{F}_t\}$-stopping time for some $i \geq 0$, i.e. $\{t_i \leq t\} \in \mathcal{F}_t$ for any $t \geq 0$. Note that Y_{i+1} is $\mathcal{F}_{t_i}$-measurable. Because the choice of Δt_{i+1} is dependent on Y_{i+1} we have that Δt_{i+1} is $\mathcal{F}_{t_i}$-measurable. Then we need to show $t_{i+1} = t_i + \Delta t_{i+1}$ is an $\{\mathcal{F}_t\}$-stopping time, that is to show $\{t_i + \Delta t_{i+1} \leq t\} \in \mathcal{F}_t$ for any $t \geq 0$. For any $s \in \mathbb{Z}$ and any $j \in \mathbb{N}$ with $j2^s \in [0, t]$, we have $\{t_i \leq j2^s\} \in \mathcal{F}_{j2^s} \subseteq \mathcal{F}_t$, and $\{\Delta t_{i+1} \leq t - j2^s\} \in \mathcal{F}_{t_i} \subset \mathcal{F}$. Thus we have $\{t_i \leq j2^s\} \cap \{\Delta t_{i+1} \leq t - j2^s\} \in \mathcal{F}_t$ (see for example (Mao, 2008)). As both $\mathbb{Z}$ and $\mathbb{N}$ are countable sets, we have that for any $t \geq 0$ (Gihman & Skorohod, 1974)

$$\{t_i + \Delta t_{i+1} \leq t\} = \bigcup_{\{0 \leq j2^s \leq t, s \in \mathbb{Z}, j \in \mathbb{N}\}} (\{t_i \leq j2^s\} \cap \{\Delta t_{i+1} \leq t - j2^s\}) \in \mathcal{F}_t.$$

Thus we have proved that t_{i+1} is an $\{\mathcal{F}_t\}$-stopping time. Since Δt_0 is dependent on the given initial value Y_0, we have Δt_0 and Y_0 are $\mathcal{F}_{t_{-1}}$-measurable (recalling $t_{-1} = 0$). By induction we conclude that t_i is an $\{\mathcal{F}_t\}$-stopping time for each $i = 0, 1, \ldots$. Substituting (5.7) into (5.6), we obtain

$$|Y_{i+1}|^2 = |Y_i|^2 - U(Y_i, \Delta t_i)\Delta t_i + \Delta m_i.$$

Then taking sum on i we have

$$|Y_{i+1}|^2 = |Y_0|^2 - \sum_{k=0}^{i} U(Y_k, \Delta t_k)\Delta t_k + m_i, \tag{5.8}$$

where $m_i = \sum_{k=0}^{i} \Delta m_k$.

Step 2

Due to (5.5) and the definition of t_i, it is clear that Y_i is $\mathcal{F}_{t_{i-1}}$-measurable for $i = 0, 1, \ldots$. We define another filtration $\{\mathcal{G}_i\}_{i=-1,0,1,\ldots}$ by $\mathcal{G}_i = \mathcal{F}_{t_i}$ for $i = -1, 0, 1, \ldots$. So Y_i is $\mathcal{G}_{i-1}$-measurable and m_i is $\mathcal{G}_i$-measurable. We are going to prove that $\{m_i\}_{i \geq 0}$ is a $\{\mathcal{G}_i\}$-local martingale. Choosing R s.t. $|x(0)| < R$, we define a stopping time

$$\rho_R = \inf\{i \geq 0, |Y_i| > R\}.$$

Clearly, $\rho_R \to \infty$ a.s. when $R \to \infty$. It is easy to see that ρ_R is a $\{\mathcal{G}_{i-1}\}$-stopping time i.e. $\{\rho_R \leq i\} \in \mathcal{G}_{i-1}$. This indicates $\{\rho_R - 1 \leq i\} \in \mathcal{G}_i$. Denoting $\tau_R = \rho_R - 1$, we have τ_R is a $\{\mathcal{G}_i\}$-stopping time. By the definition of ρ_R, we have that $|Y_{i \wedge (\rho_R - 1)}| \leq R$ a.s. so $|Y_{i \wedge \tau_R}| \leq R$ a.s. for all $i \geq 0$.

We claim that $t_{i \wedge \tau_R}$ and $t_{(i-1) \wedge \tau_R}$ are $\{\mathcal{F}_t\}$-stopping times. For $t_{i \wedge \tau_R}$ we have for any $t \geq 0$

$$\{t_{i \wedge \tau_R} \leq t\} = \{\{t_i \leq t\} \cap \{\tau_R \geq i\}\} \cup \{\{t_{\tau_R} \leq t\} \cap \{\tau_R < i\}\}.$$

Denote the complement set of A by A^c. It can be seen that $\{\tau_R \le i-1\} \in \mathcal{G}_{i-1}$, so the complement $\{\tau_R > i-1\} = \{\tau_R \le i-1\}^c \in \mathcal{G}_{i-1}$. Because $\{\tau_R > i-1\}$ is equivalent to $\{\tau_R \ge i\}$, we have $\{\tau_R \ge i\} \in \mathcal{G}_{i-1}$.

Since $\{\tau_R \ge i\} \in \mathcal{G}_{i-1} \subset \mathcal{G}_i = \mathcal{F}_{t_i}$, we have $\{t_i \le t\} \cap \{\tau_R \ge i\} \in \mathcal{F}_t$. And

$$\{t_{\tau_R} \le t\} \cap \{\tau_R < i\} = \bigcup_{j=0}^{i-1} (\{\{t_j \le t\} \cap \{\tau_R = j\}),$$

because $\{\tau_R = j\} \in \mathcal{F}_{t_i}$ for $j = 0, 1, ...i-1$ we have $\{t_{\tau_R} \le t\} \cap \{\tau_R < i\} \in \mathcal{F}_t$. Hence $\{t_{i\wedge\tau_R} \le t\} \in \mathcal{F}_t$. Similarly for $t_{(i-1)\wedge\tau_R}$, we have

$$\{t_{(i-1)\wedge\tau_R} \le t\} = \{\{t_{i-1} \le t\} \cap \{\tau_R \ge i-1\}\} \cup \{\{t_{\tau_R} \le t\} \cap \{\tau_R < i-1\}\}.$$

Similarly, we have that $\{\tau_R \le i-2\} \in \mathcal{G}_{i-2}$ indicates $\{\tau_R > i-2\} = \{\tau_R \le i-2\}^c \in \mathcal{G}_{i-2}$. As $\{\tau_R > i-2\}$ is equivalent to $\{\tau_R \ge i-1\}$, we see $\{\tau_R \ge i-1\} \in \mathcal{G}_{i-2}$.

Since $\{\tau_R \ge i-1\} \in \mathcal{G}_{i-2} \subset \mathcal{F}_{t_i}$, we have $\{t_{i-1} \le t\} \cap \{\tau_R \ge i-1\} \in \mathcal{F}_t$. And

$$\{t_{\tau_R} \le t\} \cap \{\tau_R < i-1\} = \bigcup_{j=0}^{i-2} (\{\{t_j \le t\} \cap \{\tau_R = j\}),$$

we have $\{\{t_{\tau_R} \le t\} \cap \{\tau_R < i-1\}\} \in \mathcal{F}_t$ due to $\{\tau_R = j\} \in \mathcal{G}_j \subset \mathcal{F}_{t_i}$ for $j = 0, 1, ...i-2$. Thus $\{t_{(i-1)\wedge\tau_R} \le t\} \in \mathcal{F}_t$.

Due to the iteration (5.5) and the fact that $|Y_{k\wedge\tau_R}| = |Y_{\tau_R}|$ for any $k \ge \tau_R$, we define the Brownian motion increment with the stopping time by $\Delta B_{i\wedge\tau_R} = B(t_{i\wedge\tau_R}) - B(t_{(i-1)\wedge\tau_R})$ and the time step with the stopping time by $\Delta t_{i\wedge\tau_R} = t_{i\wedge\tau_R} - t_{(i-1)\wedge\tau_R}$. Since $\tau_R \to \infty$ a.s. when $R \to \infty$, those two definitions can reproduce the original ones we used in the statement of the theorem. Thus they are valid. In addition, we have

$$m_{i\wedge\tau_R} = \sum_{k=0}^{i\wedge\tau_R} \Delta m_k = \sum_{k=0}^{i} \Delta m_{k\wedge\tau_R} \quad \text{and} \quad m_{i\wedge\tau_R} = m_{(i-1)\wedge\tau_R} + \Delta m_{i\wedge\tau_R}.$$

From condition (5.2) and Lemma 5.2.2, for $|x| \le R$ there exists a constant $c(R)$ dependent on R such that $|f(x)| \vee |g(x)| \le c(R)$. By the elementary inequality, we have

$$\begin{aligned}
|m_{i\wedge\tau_R}| &= \left|\sum_{k=0}^{i\wedge\tau_R}\Delta m_k\right| \\
&\le \sum_{k=0}^{i}|\Delta m_{k\wedge\tau_R}| \\
&\le \sum_{k=0}^{i}(2|Y_{k\wedge\tau_R}||g(Y_{k\wedge\tau_R})||\Delta B_{k\wedge\tau_R}| + 2|f(Y_{k\wedge\tau_R})||g(Y_{k\wedge\tau_R})|\Delta t_{k\wedge\tau_R}|\Delta B_{k\wedge\tau_R}| \\
&\quad +|g(Y_{k\wedge\tau_R})|^2||\Delta B_{k\wedge\tau_R}|^2 - \Delta t_{k\wedge\tau_R}|) \\
&\le \sum_{k=0}^{i}(c_1(R)|\Delta B_{k\wedge\tau_R}| + c_2(R)|\Delta B_{k\wedge\tau_R}|^2), \qquad (5.9)
\end{aligned}$$

where $c_1(R)$ and $c_2(R)$ are constants dependent on R only. Hence we have

$$\mathbb{E}|m_{i\wedge\tau_R}| \le \sum_{k=0}^{i}(c_1(R)\mathbb{E}|\Delta B_{k\wedge\tau_R}| + c_2(R)\mathbb{E}|\Delta B_{k\wedge\tau_R}|^2) < \infty.$$

Also we have

$$\mathbb{E}(m_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) = \mathbb{E}(m_{(i-1)\wedge\tau_R} + \Delta m_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) = m_{(i-1)\wedge\tau_R} + \mathbb{E}(\Delta m_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}). \qquad (5.10)$$

Because $\{\tau_R > i-1\} \in \mathcal{G}_{i-1}$ and ΔB_i is independent of $\mathcal{G}_{i-1}$, we have

$$\begin{aligned}
&\mathbb{E}(\Delta B_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) \\
={}& \mathbb{E}[(B(t_i) - B(t_{i-1}))\mathbf{1}_{\{\tau_R>i-1\}}\big|\mathcal{G}_{i-1}] + \mathbb{E}[(B(t_{\tau_R}) - B(t_{\tau_R}))\mathbf{1}_{\{\tau_R\le i-1\}}\big|\mathcal{G}_{i-1}] \\
={}& \mathbf{1}_{\{\tau_R>i-1\}}\mathbb{E}[B(t_i) - B(t_{i-1})] \\
={}& 0,
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}(|\Delta B_{i\wedge\tau_R}|^2\big|\mathcal{G}_{i-1}) \\
={}& \mathbb{E}[|B(t_i) - B(t_{i-1})|^2\mathbf{1}_{\{\tau_R>i-1\}}\big|\mathcal{G}_{i-1}] + \mathbb{E}[|B(t_{\tau_R}) - B(t_{\tau_R})|^2\mathbf{1}_{\{\tau_R\le i-1\}}\big|\mathcal{G}_{i-1}] \\
={}& \mathbf{1}_{\{\tau_R>i-1\}}\mathbb{E}[|B(t_i) - B(t_{i-1})|^2] \\
={}& \mathbf{1}_{\{\tau_R>i-1\}}(t_i - t_{i-1}),
\end{aligned}$$

and

$$\begin{aligned} & \mathbb{E}(\Delta t_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) \\ = \;& \Delta t_{i\wedge\tau_R} \\ = \;& \mathbf{1}_{\{\tau_R>i-1\}}(t_i - t_{i-1}) + \mathbf{1}_{\{\tau_R\le i-1\}}(t_{\tau_R} - t_{\tau_R}) \\ = \;& \mathbf{1}_{\{\tau_R>i-1\}}(t_i - t_{i-1}). \end{aligned}$$

Hence

$$\begin{aligned} & \mathbb{E}(\Delta m_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) \\ = \;& \mathbb{E}(2\langle Y_{i\wedge\tau_R}, g(Y_{i\wedge\tau_R})\Delta B_{i\wedge\tau_R}\rangle + 2\langle f(Y_{i\wedge\tau_R})\Delta t_{i\wedge\tau_R}, g(Y_{i\wedge\tau_R})\Delta B_{i\wedge\tau_R}\rangle \\ & +|g(Y_{i\wedge\tau_R})|^2(|\Delta B_{i\wedge\tau_R}|^2 - \Delta t_{i\wedge\tau_R})\big|\mathcal{G}_{i-1}) \\ = \;& 2\langle Y_{i\wedge\tau_R}, g(Y_{i\wedge\tau_R})\rangle\mathbb{E}(\Delta B_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) + 2\langle f(Y_{i\wedge\tau_R}), g(Y_{i\wedge\tau_R})\rangle\Delta t_{i\wedge\tau_R}\mathbb{E}(\Delta B_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) \\ & +|g(Y_{i\wedge\tau_R})|^2(\mathbb{E}(|\Delta B_{i\wedge\tau_R}|^2\big|\mathcal{G}_{i-1}) - \mathbb{E}(\Delta t_{i\wedge\tau_R}\big|\mathcal{G}_{i-1})) \\ = \;& 0. \end{aligned} \tag{5.11}$$

Combining (5.10) and (5.11), we achieve the required

$$\mathbb{E}(m_{i\wedge\tau_R}\big|\mathcal{G}_{i-1}) = m_{(i-1)\wedge\tau_R}.$$

This means that $\{m_{i\wedge\tau_R}\}_{i\ge0}$ is a $\{\mathcal{G}_i\}$-martingale. Recalling that $\tau_R\to\infty$ a.s. when $R\to\infty$, we see that $\{m_i\}_{i\ge0}$ is a $\{\mathcal{G}_i\}$-local martingale.

Step 3

Therefore from (5.8) and Lemma 5.2.3, we have

$$\lim_{i\to\infty}|Y_i|^2 < \infty \quad \text{a.s.} \tag{5.12}$$

and

$$\sum_{k=0}^{\infty} U(Y_k, \Delta t_i)\Delta t_k < \infty \quad \text{a.s.} \tag{5.13}$$

From (5.13), we have $\lim_{i\to\infty} U(Y_i, \Delta t_i)\Delta t_i = 0$ a.s. We next show the time step Δt_i will never tend to zero as i goes to infinity, that is $\liminf_{i\to\infty}\Delta t_i > 0$ a.s.

According to (5.12) for almost all $\omega \in \Omega$, there exists $C(\omega) \in \mathbb{R}_+$ such that $\lim_{i\to\infty} |Y_i(\omega)| = C(\omega)$. Fix any such ω, write $C(\omega) = C$ and $Y_i(\omega) = Y_i$. Consider two cases:

(i) For the case when $C \neq 0$, there exists a sufficiently large integer i_1^* such that for all $i > i_1^*$, $0.5C < |Y_i| < 1.5C$. This indicates either $0.5C < Y_i < 1.5C$ or $-1.5C < Y_i < -0.5C$. Because that $z(x) = 0$ and $f(x) = 0$ if and only if $x = 0$, in both of the two intervals we have $z(Y_i) \neq 0$ and $f(Y_i) \neq 0$. Furthermore, due to the continuity of $z(x)$ and $f(x)$, we have

$$\min_{0.5C \leq |x| \leq 1.5C} \frac{z(x)}{|f(x)|^2} = \eta > 0.$$

So for any $i > i_1^*$, we have

$$\frac{z(Y_i)}{|f(Y_i)|^2} \geq \eta > 0$$

then

$$1 - \log_2(z(Y_i)/|f(Y_i)|^2) \leq 1 - \log_2(\eta).$$

Recalling the choice of the step size, we see

$$n_i = \lceil 1 - \log_2(z(Y_i)/|f(Y_i)|^2) \rceil \leq \lceil 1 - \log_2(\eta) \rceil$$

then

$$\Delta t_i = 2^{-n_i} \geq 2^{-\lceil 1-\log_2(\eta) \rceil} > 0.$$

(ii) For the case when $C = 0$, suppose the limit of (5.4) be $D > 0$. There exists a constant $\delta = \delta(D) > 0$ such that $|z(x)/|f(x)|^2 - D| < 0.5D$ for all $|x| \in (0, \delta)$. Also, there exists an integer i_2^* such that for all $i > i_2^*$, $|Y_i| \in (0, \delta)$, which indicates $|z(Y_i)/|f(Y_i)|^2 - D| < 0.5D$. So for any $i > i_2^*$, we have

$$1 - \log_2(1.5D) < 1 - \log_2(z(Y_i)/|f(Y_i)|^2) < 1 - \log_2(0.5D).$$

Recalling the choice of the step size, we see

$$\Delta t_i = 2^{-n_i} > 2^{-\lceil 1-\log_2(0.5D) \rceil} > 0.$$

Thus Δt_i will never tend to 0 as i tends to infinity. Hence we have $\sum_{i=0}^{\infty} \Delta t_i = \infty$ a.s.

Now we have $\lim_{i\to\infty} U(Y_i, \Delta t_i) = 0$ a.s. Due to (5.7) and the choice of Δt_i that $\Delta t_i \leq z(Y_i)/(2|f(Y_i)|^2)$, we have

$$U(Y_i, \Delta t_i) = z(Y_i) - |f(Y_i)|^2 \Delta t_i \geq 0.5 z(Y_i) \geq 0.$$

Therefore $\lim_{i\to\infty} z(Y_i) = 0$ a.s. Given the condition "$z(x) = 0 \Leftrightarrow x = 0$", we obtain that $\lim_{i\to\infty} Y_i = 0$ a.s. Hence the proof is complete.

We have two comments on the proof.

- The condition in Theorem 5.3.1 for the EM method with variable step size is weaker than the condition for the EM method with fixed step size (i.e. when $\theta = 0$) stated in Theorem 5.3 of (Mao & Szpruch, 2013a). For example, a scalar SDE $dx(t) = (-x^3(t) - x(t))dt + x^2(t)dB(t)$ satisfies the conditions in Theorem 5.3.1, but not in Theorem 5.3 of (Mao & Szpruch, 2013a).

- When conducting computer simulation, the step size is naturally rational number as computers can only deal with finite number of decimals. Thus we may simply set each step size to be $\alpha z(Y_i)/(|f(Y_i)|^2)$ for any rational number $\alpha \in (0, 1)$. We generalise Theorem 5.3.1 to the next theorem.

Theorem 5.3.2 *Let (5.2) and (5.3) hold. Assume $z(x) = 0$ if and only if $x = 0$, and (5.4). For the EM method with variable step size (5.5), Δt_i is chosen to be rational number satisfying $\Delta t_i = \alpha z(Y_i)/(|f(Y_i)|^2)$ with $\alpha \in (0, 1)$ for $|Y_i| \neq 0$, and any nonzero rational number for $|Y_i| = 0$. Then t_i is an $\{\mathcal{F}_t\}$-stopping time for each $i = 0, 1, 2...$, and the sequence of time steps obeys $\sum_{i=0}^{\infty} \Delta t_i = \infty$ a.s. Moreover, for any initial value $Y_0 \in \mathbb{R}^n$*

$$\lim_{i\to\infty} Y_i = 0 \quad a.s.$$

Most part of the proof of Theorem 5.3.2 is similar to the proof of Theorem 5.3.1, and the only different part is the proof of the stopping time as follows.

Assume t_i is an $\{\mathcal{F}_t\}$-stopping time for some $i \geq 0$, i.e. $\{t_i \leq t\} \in \mathcal{F}_t$ for any $t \geq 0$. Note that Y_{i+1} is $\mathcal{F}_{t_i}$-measurable, because the choice of Δt_{i+1} is dependent on Y_{i+1} we have that Δt_{i+1} is $\mathcal{F}_{t_i}$-measurable. Then we need to show $t_{i+1} = t_i + \Delta t_{i+1}$ is an $\{\mathcal{F}_t\}$-stopping time, that is to show $\{t_i + \Delta t_{i+1} \leq t\} \in \mathcal{F}_t$ for any $t \geq 0$. For any rational number $s \in [0, t]$, we have $\{t_i \leq s\} \in \mathcal{F}_s \subseteq \mathcal{F}_t$, and $\{\Delta t_{i+1} \leq t - s\} \in \mathcal{F}_{t_i} \subseteq \mathcal{F}$. Thus we have $\{t_i \leq s\} \cap \{\Delta t_{i+1} \leq t - s\} \in \mathcal{F}_t$ (see for example (Mao, 2008)). As the set of all rational number $s \in [0, t]$ is a countable set, we have that for any $t \geq 0$ (Gihman & Skorohod, 1974)

$$\{t_i + \Delta t_{i+1} \leq t\} = \bigcup_{\{0 \leq s \leq t, s \in \mathbb{Q}\}} (\{t_i \leq s\} \cap \{\Delta t_{i+1} \leq t - s\}) \in \mathcal{F}_t.$$

Thus we have proved that t_{i+1} is an $\{\mathcal{F}_t\}$-stopping time. Since Δt_0 is dependent on the given initial value Y_0, we have Δt_0 and Y_0 are $\mathcal{F}_{t_{-1}}$-measurable (recalling $t_{-1} = 0$). By induction we conclude that t_i is an $\{\mathcal{F}_t\}$-stopping time for each $i = 0, 1, ...$

5.4 Examples

We first consider a scalar SDE

$$dx(t) = (-x^3(t) - x(t))dt + x^2(t)dB(t) \tag{5.14}$$

with a given initial value $x(0) = 1$. It is easy to verify that for any $x \in \mathbb{R}$ and $x \neq 0$

$$-z(x) := 2\langle x, f(x)\rangle + g^2(x) = -2x^2 - x^4 < 0.$$

It is clear that $z(x) = 0 \Leftrightarrow x = 0$, by Theorem 5.2.1 we have the solution of the underlying SDE is asymptotically almost surely stable. Moreover,

$$\liminf_{|x| \to 0} \frac{z(x)}{|f(x)|^2} = \liminf_{|x| \to 0} \frac{2x^2 + x^4}{x^2 + 2x^4 + x^6} = 2 > 0.$$

Choose the step size, for example $\Delta t_i = 0.98 z(Y_i)/|f(Y_i)|^2$ in each step, from Theorem 5.3.2 we obtain the variable step size EM solution is asymptotically almost surely stable as well. Set $Y_0 = 1$, we simulated 1000 time steps of one path of the variable step size EM solution. The left plot on Figure 5.1 is the solution path, from which we can see that the oscillation decays and the solution tends zero as time increases. This is in line with the theoretical result. The plot on the right of Figure 5.1 is the size of each time step. It is clear that with the solution approaching the origin the step size tends to 1.96 and this is due to the limit 2 and the choice of factor 0.98. In addition, the plot also shows that the step size does not need to tend to zero, thus we have $\sum_{i=0}^{\infty} \Delta t_i = \infty$ a.s.

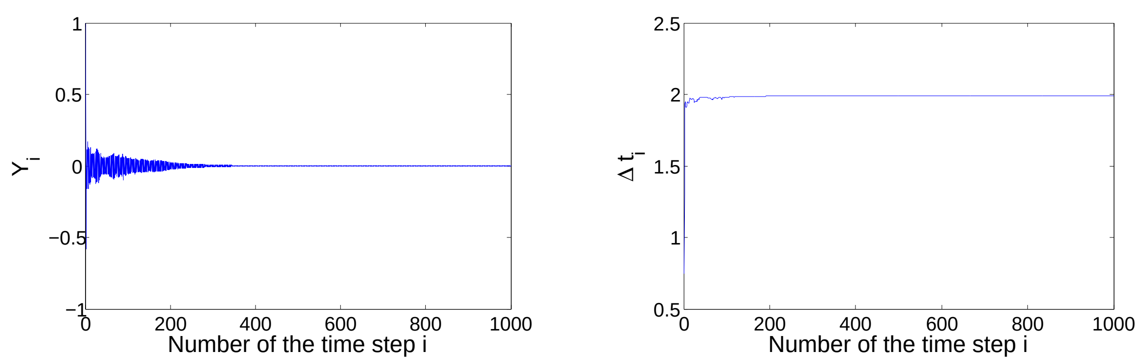

Figure 5.1: Left: One simulation path, Right: The step size of each time step

Now we consider a two-dimensional case

$$dx(t) = \text{diag}(x_1(t), x_2(t)) \left((b + A\text{diag}(x_1(t), x_2(t))x(t))\, dt + \sigma dB(t) \right), \qquad (5.15)$$

where $\text{diag}(x_1(t), x_2(t))$ denotes diagonal matrix with nonzero entries $x_1(t)$ and

$x_2(t)$ on the diagonal, $x(t) = (x_1(t), x_2(t))^T$, $b = (b_1, b_2)^T$, $A = (a_{ij})_{i,j\in\{1,2\}}$, $\sigma = (\sigma_{ij})_{i,j\in\{1,2\}}$ and $B(t) = (B_1(t), B_2(t))^T$.

We set $b = (-1,-2)^T, a_{11} = a_{22} = -1, a_{12} = -2, a_{21} = 1, \sigma_{11} = \sigma_{12} = 0.5, \sigma_{21} = 1, \sigma_{22} = -1$. It is easy to verify that for any $x \in \mathbb{R}^2$ and $x \neq 0$

$$\begin{aligned} & 2\langle x, f(x)\rangle + g^2(x) \\ = \; & (2b_1 + \sigma_{11}^2 + \sigma_{12}^2)x_1^2 + (2b_2 + \sigma_{21}^2 + \sigma_{22}^2)x_2^2 + (a_{12} + a_{21})x_1^2 x_2^2 + a_{11}x_1^4 + a_{22}x_2^4 < 0. \end{aligned}$$

From Theorem 5.2.1, we know the SDE solution is almost surely stable. In addition, by the elementary inequality $ab \leq a^2 + b^2$ we have

$$\begin{aligned} & \liminf_{|x|\to 0} \frac{z(x)}{|f(x)|^2} \\ = \; & \liminf_{|x|\to 0} \frac{1.5x_1^2 + 2x_2^2 + x_1^2x_2^2 + x_1^4 + x_2^4}{x_1^2 + 2x_1^4 + x_1^6 + 2x_1^2x_2^4 + 5x_1^4x_2^2 + 4x_2^2 + 4x_2^4 + x_2^6} \\ \geq \; & \liminf_{|x|\to 0} \frac{|x|^2}{4|x|^2 + 6.5|x|^4 + |x|^6 + 2.5|x|^8} \\ = \; & \frac{1}{4} > 0. \end{aligned}$$

By choosing the step size, for example $\Delta t_i = 0.1z(Y_i)/|f(Y_i)|^2$ in each step, we have from Theorem 5.3.1 that the variable step size EM solution is almost surely stable as well.

We simulated 10000 time steps and plotted the two solution paths on the left of Figure 5.2. It can be seen that as time increases both the solutions tend to zero. And from the plot on the right of Figure 5.2 the size of the time step approaches to 0.025 as the solutions go to zeros, which shows the step size will not tend to zero. Hence both the simulations of the one-dimensional and the multi-dimensional cases are in line with the theoretical result.

5.5 Other Sufficient Conditions

In this section, we propose some other sufficient conditions which can cover some SDEs that are not included in Section 5.3.

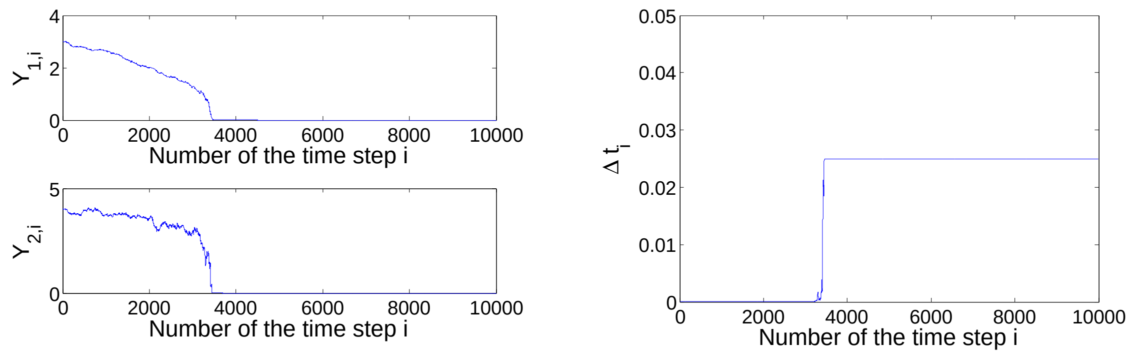

Figure 5.2: Left: One simulation path of $Y_{1,\cdot}$ and $Y_{2,\cdot}$, Right: The step size of each time step.

A slightly better condition than (5.3) is to assume there exists a symmetric positive-definite $n \times n$ matrix Q such that for $\forall x \in \mathbb{R}^n$

$$-\bar{z}(x) := 2x^T Q f(x) + trace(g^T(x) Q g(x)) \leq 0. \tag{5.16}$$

Thanks to the stochastic version of the LaSalle theorem in (Shen *et al.*, 2006), we have that the underlying solution of (5.1) is almost surely asymptotically stable if (5.2) and (5.16) hold, and $\bar{z}(x) = 0$ if and only if $x = 0$. In addition, it is obvious that given the condition that $\bar{z}(x) = 0$ if and only if $x = 0$ the results in Lemma 5.2.2 still hold for $f(x)$ and $g(x)$. Denote the smallest and largest eigenvalue of Q by $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$ respectively. Now we are ready to present the following theorem.

Theorem 5.5.1 *Let (5.2) and (5.16) hold. Assume* $\bar{z}(x) = 0$ *if and only if* $x = 0$, *and*

$$\liminf_{|x|\to 0} \frac{\bar{z}(x)}{|f(x)|^2} > 0.$$

For the EM method with variable step size (5.5), Δt_i *is chosen to be rational number satisfying* $\Delta t_i = \alpha\bar{z}(Y_i)/(\lambda_{\max}(Q)|f(Y_i)|^2)$ *with* $\alpha \in (0,1)$ *for* $|Y_i| \neq 0$, *and any nonzero rational number for* $|Y_i| = 0$. *Then* t_i *is an* $\{\mathcal{F}_t\}$*-stopping time for each* $i = 0, 1, 2...$, *and the sequence of time steps obeys* $\sum_{i=0}^{\infty} \Delta t_i = \infty$ *a.s. Moreover, for any initial value* $Y_0 \in \mathbb{R}^n$

$$\lim_{i\to\infty} Y_i = 0 \quad a.s.$$

Proof. Since Q is a symmetric positive-definite $n \times n$ matrix, it is clear that for any $i \geq 0$

$$\lambda_{\min}(Q)|Y_i|^2 \leq Y_i^T Q Y_i \leq \lambda_{\max}(Q)|Y_i|^2$$

and

$$\lambda_{\min}(Q)|f(Y_i)|^2 \leq f^T(Y_i)Qf(Y_i) \leq \lambda_{\max}(Q)|f(Y_i)|^2.$$

From (5.5) we have

$$Y_{i+1}^T Q Y_{i+1} = Y_i^T Q Y_i + \Delta t_i[2Y_i^T Q f(Y_i) + trace(g^T(Y_i)Qg(Y_i)) + f^T(Y_i)Qf(Y_i)\Delta t_i] + \Delta m_i,$$

where

$$\begin{aligned}\Delta m_i &= 2Y_i^T Q g(Y_i)\Delta B_i + 2f^T(Y_i)Qg(Y_i)\Delta B_i \\ &\quad + (g(Y_i)\Delta B_i)^T Q(g(Y_i)\Delta B_i) - trace(g^T(Y_i)Qg(Y_i))\Delta t_i.\end{aligned}$$

Then the proof can be completed by adapting the same procedure used in Theorem 5.3.1.

We see condition (5.16) as a generalisation of (5.3) as we can recover (5.3) by choosing Q to be identity matrix in (5.16).

To keep the notations simple in the next theorem, we investigate the SDEs with the scalar Brownian motion

$$dx(t) = f(x(t))dt + g(x(t))dB(t), \quad x(0) \in \mathbb{R}^n,$$

where f: $\mathbb{R}^n \to \mathbb{R}^n$, g: $\mathbb{R}^n \to \mathbb{R}^n$ and $B(t)$ is a scalar Brownian motion. We still assume condition (5.2), but replace condition (5.3) by the following condition: there exists a constant $p \in (0,2)$ such that

$$-v := \sup_{x\in\mathbb{R}^n, x\neq 0} \left(\frac{2\langle x, f(x)\rangle + |g(x)|^2}{|x|^2} + (p-2)\frac{\langle x, g(x)\rangle^2}{|x|^4} \right) < 0. \tag{5.17}$$

Also we assume $f(0) = 0$ and $g(0) = 0$.

Under (5.2) and (5.17), the true solution of SDE (5.1) is almost surely asymptotically stable (Shen *et al.*, 2006). Now we study the numerical solution.

Theorem 5.5.2 *Let (5.2) and (5.17) hold. Assume*

$$\limsup_{|x|\to 0} \frac{|f(x)|}{|x|} < \infty, \tag{5.18}$$

and

$$\limsup_{|x|\to 0} \frac{|g(x)|}{|x|} < \infty. \tag{5.19}$$

Define the EM method with variable step size as

$$Y_{i+1} = Y_i + f(Y_i)\Delta t_i + g(Y_i)\Delta B_i, \quad Y_0 = x(0), \quad i \geq 0, \tag{5.20}$$

where $\Delta B_i = B(t_i) - B(t_{i-1})$ *with* $t_i = \sum_{k=0}^{i} \Delta t_k$ *for* $i = 0, 1, 2...$ *and* $t_{-1} = 0$.

For $Y_i \neq 0$, Δt_i *is chosen to be rational number satisfying* $\Delta t_i \leq (p/12)\times \min_{\{j=1,2,3,4,5\}}\{(v/A_j(Y_i))^{(1/j)}\}$, *where* $\{A_j\}_{j=1,2,3,4,5}$ *are defined in the proof. For* $Y_i = 0$, Δt_i *is chosen to be any nonzero rational number. Then* t_i *is an* $\{\mathcal{F}_t\}$-*stopping time for each* $i = 0, 1, 2...$, *and the sequence of time steps obeys* $\sum_{i=0}^{\infty} \Delta t_i = \infty$ *a.s. Moreover, for any initial value* $Y_0 \in \mathbb{R}^n$

$$\lim_{i\to\infty} Y_i = 0 \quad a.s.$$

Proof. From the first line of (5.6), we have that for the p given in (5.17) and $Y_i \neq 0$

$$|Y_{i+1}|^p = |Y_i|^p \left(1 + \frac{2\langle Y_i, f(Y_i)\Delta t_i + g(Y_i)\Delta B_i\rangle + |f(Y_i)\Delta t_i + g(Y_i)\Delta B_i|^2}{|Y_i|^2} \right)^{p/2}.$$

When $Y_i = 0$ (i.e. $f(Y_i) = 0$ and $g(Y_i) = 0$) for some $i > 0$, due to the iteration (5.20) the solution will stay at zero afterwards. In this case Δt_i could be set to be any nonzero rational number. In the following we focus on the case that $Y_i \neq 0$ for all $i \geq 0$. Let

$$\zeta = \frac{2\langle Y_i, f(Y_i)\Delta t_i + g(Y_i)\Delta B_i\rangle + |f(Y_i)\Delta t_i + g(Y_i)\Delta B_i|^2}{|Y_i|^2},$$

and by the fundamental inequality that for any $\zeta \geq -1$

$$(1+\zeta)^{p/2} \leq 1 + \frac{p}{2}\zeta + \frac{p(p-2)}{8}\zeta^2 + \frac{p(p-2)(p-4)}{2^3 \times 3!}\zeta^3,$$

we have

$$|Y_{i+1}|^p \leq |Y_i|^p \left(1 + \frac{p}{2}\zeta + \frac{p(p-2)}{8}\zeta^2 + \frac{p(p-2)(p-4)}{2^3 \times 3!}\zeta^3\right). \tag{5.21}$$

We compute

$$\begin{aligned}\zeta \;=\;& \frac{1}{|Y_i|^2}(\Delta t_i(2\langle Y_i, f(Y_i)\rangle + |g(Y_i)|^2) + \Delta t_i^2|f(Y_i)|^2 \\ &+2\langle Y_i, g(Y_i)\rangle\Delta B_i + 2f^T(Y_i)g(Y_i)\Delta t_i\Delta B_i + |g(Y_i)|^2(\Delta B_i^2 - \Delta t_i)),\end{aligned}$$

$$\begin{aligned}&\zeta^2 = \\ &\frac{1}{|Y_i|^4}(\Delta t_i(4\langle Y_i, g(Y_i)\rangle^2) \\ &+\Delta t_i^2(4\langle Y_i, f(Y_i)\rangle^2 + |g(Y_i)|^4 + 4\langle Y_i, f(Y_i)\rangle|g(Y_i)|^2 + 8\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)) \\ &+\Delta t_i^3(6|f(Y_i)|^2|g(Y_i)|^2 + 4\langle Y_i, f(Y_i)\rangle|f(Y_i)|^2) + \Delta t_i^4|f(Y_i)|^4 \\ &+4\langle Y_i, g(Y_i)\rangle^2(\Delta B_i^2 - \Delta t_i) + |g(Y_i)|^4(\Delta B_i^4 - \Delta t_i^2) + 4\langle Y_i, f(Y_i)\rangle|g(Y_i)|^2\Delta t_i(\Delta B_i^2 - \Delta t_i) \\ &+8\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)\Delta t_i(\Delta B_i^2 - \Delta t_i) + 6|f(Y_i)|^2|g(Y_i)|^2\Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\ &+8\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle\Delta t_i\Delta B_i + 4|f(Y_i)|^2 f^T(Y_i)g(Y_i)\Delta t_i^3\Delta B_i \\ &+4f^T(Y_i)g(Y_i)|g(Y_i)|^2\Delta t_i\Delta B_i^3 \\ &+8\langle Y_i, f(Y_i)\rangle f^T(Y_i)g(Y_i)\Delta t_i^2\Delta B_i + 4\langle Y_i, g(Y_i)\rangle|f(Y_i)|^2\Delta t_i^2\Delta B_i + 4\langle Y_i, g(Y_i)\rangle|g(Y_i)|^2\Delta B_i^3),\end{aligned}$$

and

$$
\begin{aligned}
&\zeta^3 = \\
&\frac{1}{|Y_i|^6}(\Delta t_i^2(24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle^2 + 12\langle Y_i, g(Y_i)\rangle^2|g(Y_i)|^2) \\
&+\Delta t_i^3(8\langle Y_i, f(Y_i)\rangle^3 + 12\langle Y_i, f(Y_i)\rangle^2|g(Y_i)|^2 + 48\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i) \\
&+12\langle Y_i, g(Y_i)\rangle^2|f(Y_i)|^2 + 6\langle Y_i, f(Y_i)\rangle|g(Y_i)|^4 + |g(Y_i)|^6 + 24\langle Y_i, g(Y_i)\rangle|f(Y_i)||g(Y_i)|^3) \\
&+\Delta t_i^4(12\langle Y_i, f(Y_i)\rangle^2|f(Y_i)|^2 + 36\langle Y_i, f(Y_i)\rangle|f(Y_i)|^2|g(Y_i)|^2 + 15|f(Y_i)|^2|g(Y_i)|^4 \\
&+24\langle Y_i, g(Y_i)\rangle|f(Y_i)|^3|g(Y_i)|) \\
&+\Delta t_i^5(6\langle Y_i, f(Y_i)\rangle|f(Y_i)|^4 + 15|f(Y_i)|^4|g(Y_i)|^2) \\
&+\Delta t_i^6(|f(Y_i)|^6) \\
&+24\langle Y_i, f(Y_i)\rangle^2\langle Y_i, g(Y_i)\rangle\Delta t_i^2\Delta B_i + 24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle^2\Delta t_i(\Delta B_i^2 - \Delta t_i) \\
&+8\langle Y_i, g(Y_i)\rangle^3\Delta B_i^3 + 24\langle Y_i, f(Y_i)\rangle^2 f^T(Y_i)g(Y_i)\Delta t_i^3\Delta B_i \\
&+12\langle Y_i, f(Y_i)\rangle^2|g(Y_i)|^2\Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle|f(Y_i)|^2\Delta t_i^3\Delta B_i \\
&+48\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)\Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle|g(Y_i)|^2\Delta t_i\Delta B_i^3 + 12\langle Y_i, g(Y_i)\rangle^2|f(Y_i)|^2\Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, g(Y_i)\rangle^2 f^T(Y_i)g(Y_i)\Delta t_i\Delta B_i^3 + 12\langle Y_i, g(Y_i)\rangle^2|g(Y_i)|^2(\Delta B_i^4 - \Delta t_i^2) \\
&+24\langle Y_i, f(Y_i)\rangle|f(Y_i)|^2 f^T(Y_i)g(Y_i)\Delta t_i^4\Delta B_i + 36\langle Y_i, f(Y_i)\rangle|f(Y_i)|^2|g(Y_i)|^2\Delta t_i^3(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, f(Y_i)\rangle f^T(Y_i)g(Y_i)|g(Y_i)|^2\Delta t_i^2\Delta B_i^3 + 6\langle Y_i, f(Y_i)\rangle|g(Y_i)|^4\Delta t_i(\Delta B_i^4 - \Delta t_i^2) \\
&+6|f(Y_i)|^4 f^T(Y_i)g(Y_i)\Delta t_i^5\Delta B_i + 15|f(Y_i)|^4|g(Y_i)|^2\Delta t_i^4(\Delta B_i^2 - \Delta t_i) \\
&+20|f(Y_i)|^2 f^T(Y_i)g(Y_i)|g(Y_i)|^2\Delta t_i^3\Delta B_i^3 \\
&+15|f(Y_i)|^2|g(Y_i)|^4\Delta t_i^2(\Delta B_i^4 - \Delta t_i^2) + 6f^T(Y_i)g(Y_i)|g(Y_i)|^4\Delta t_i\Delta B_i^5 + |g(Y_i)|^6(\Delta B_i^6 - \Delta t_i^3) \\
&+6\langle Y_i, g(Y_i)\rangle|f(Y_i)|^4\Delta t_i^4\Delta B_i + 24\langle Y_i, g(Y_i)\rangle|f(Y_i)|^3|g(Y_i)|\Delta t_i^3(\Delta B_i^2 - \Delta t_i) \\
&+36\langle Y_i, g(Y_i)\rangle|f(Y_i)|^2|g(Y_i)|^2\Delta t_i^2\Delta B_i^3 + 24\langle Y_i, g(Y_i)\rangle|f(Y_i)||g(Y_i)|^3\Delta t_i(\Delta B_i^4 - \Delta t_i^2) \\
&+6\langle Y_i, g(Y_i)\rangle|g(Y_i)|^4\Delta B_i^5).
\end{aligned}
$$

Then we can rearrange (5.21) into

$$|Y_{i+1}|^p \leq |Y_i|^p - |Y_i|^p \Delta t_i U_1(\Delta t_i, Y_i) + \Delta m_i, \tag{5.22}$$

where

$$\begin{aligned}
&-U_1(\Delta t_i, Y_i) := \\
&\frac{p}{2}\Bigg(\frac{2\langle Y_i, f(Y_i)\rangle + |g(Y_i)|^2}{|Y_i|^2} + \frac{p-2}{4}\frac{4\langle Y_i, g(Y_i)\rangle^2}{|Y_i|^4}\Bigg) \\
+\quad & \Delta t_i \Bigg(\frac{p}{2}\frac{|f(Y_i)|^2}{|Y_i|^2} \\
&+\frac{p(p-2)}{8}\frac{4\langle Y_i, f(Y_i)\rangle^2 + |g(Y_i)|^4 + 4\langle Y_i, f(Y_i)\rangle|g(Y_i)|^2 + 8\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)}{|Y_i|^4} \\
&+\frac{p(p-2)(p-4)}{2^3\times 3!}\frac{24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle^2 + 12\langle Y_i, g(Y_i)\rangle^2|g(Y_i)|^2}{|Y_i|^6}\Bigg) \\
+\quad & \Delta t_i^2 \Bigg(\frac{p(p-2)}{8}\frac{6|f(Y_i)|^2|g(Y_i)|^2 + 4\langle Y_i, f(Y_i)\rangle|f(Y_i)|^2}{|Y_i|^4} + \frac{p(p-2)(p-4)}{2^3\times 3!}\times \\
&\Bigg(\frac{8\langle Y_i, f(Y_i)\rangle^3 + 12\langle Y_i, f(Y_i)\rangle^2|g(Y_i)|^2 + 48\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)}{|Y_i|^6} \\
&+\frac{12\langle Y_i, g(Y_i)\rangle^2|f(Y_i)|^2 + 6\langle Y_i, f(Y_i)\rangle|g(Y_i)|^4 + |g(Y_i)|^6 + 24\langle Y_i, g(Y_i)\rangle|f(Y_i)||g(Y_i)|^3}{|Y_i|^6}\Bigg)\Bigg) \\
+\quad & \Delta t_i^3 \Bigg(\frac{p(p-2)}{8}\frac{|f(Y_i)|^4}{|Y_i|^4} + \frac{p(p-2)(p-4)}{2^3\times 3!}\times\Bigg(\frac{12\langle Y_i, f(Y_i)\rangle^2|f(Y_i)|^2}{|Y_i|^6} \\
+\quad & \frac{36\langle Y_i, f(Y_i)\rangle|f(Y_i)|^2|g(Y_i)|^2 + 15|f(Y_i)|^2|g(Y_i)|^4 + 24\langle Y_i, g(Y_i)\rangle|f(Y_i)|^3|g(Y_i)|}{|Y_i|^6}\Bigg)\Bigg) \\
+\quad & \Delta t_i^4 \Bigg(\frac{p(p-2)(p-4)}{2^3\times 3!}\frac{6\langle Y_i, f(Y_i)\rangle|f(Y_i)|^4 + 15|f(Y_i)|^4|g(Y_i)|^2}{|Y_i|^6}\Bigg) \\
+\quad & \Delta t_i^5 \Bigg(\frac{p(p-2)(p-4)}{2^3\times 3!}\frac{|f(Y_i)|^6}{|Y_i|^6}\Bigg),
\end{aligned}$$

and

$$
\begin{aligned}
&\Delta m_i = \\
&|Y_i|^p(\frac{1}{|Y_i|^2}(2\langle Y_i, g(Y_i)\rangle \Delta B_i + 2f^T(Y_i)g(Y_i)\Delta t_i \Delta B_i + |g(Y_i)|^2(\Delta B_i^2 - \Delta t_i)) \\
&+\frac{1}{|Y_i|^4}(4\langle Y_i, g(Y_i)\rangle^2(\Delta B_i^2 - \Delta t_i) \\
&+|g(Y_i)|^4(\Delta B_i^4 - \Delta t_i^2) + 4\langle Y_i, f(Y_i)\rangle |g(Y_i)|^2 \Delta t_i(\Delta B_i^2 - \Delta t_i) \\
&+8\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)\Delta t_i(\Delta B_i^2 - \Delta t_i) + 6|f(Y_i)|^2|g(Y_i)|^2\Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+8\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle \Delta t_i \Delta B_i + 4|f(Y_i)|^2 f^T(Y_i)g(Y_i)\Delta t_i^3 \Delta B_i \\
&+4f^T(Y_i)g(Y_i)|g(Y_i)|^2\Delta t_i \Delta B_i^3 \\
&+8\langle Y_i, f(Y_i)\rangle f^T(Y_i)g(Y_i)\Delta t_i^2 \Delta B_i + 4\langle Y_i, g(Y_i)\rangle |f(Y_i)|^2 \Delta t_i^2 \Delta B_i + 4\langle Y_i, g(Y_i)\rangle |g(Y_i)|^2 \Delta B_i^3) \\
&+\frac{1}{|Y_i|^6}(24\langle Y_i, f(Y_i)\rangle^2 \langle Y_i, g(Y_i)\rangle \Delta t_i^2 \Delta B_i + 24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle^2 \Delta t_i(\Delta B_i^2 - \Delta t_i) \\
&+8\langle Y_i, g(Y_i)\rangle^3 \Delta B_i^3 + 24\langle Y_i, f(Y_i)\rangle^2 f^T(Y_i)g(Y_i)\Delta t_i^3 \Delta B_i \\
&+12\langle Y_i, f(Y_i)\rangle^2 |g(Y_i)|^2 \Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle |f(Y_i)|^2 \Delta t_i^3 \Delta B_i \\
&+48\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle f^T(Y_i)g(Y_i)\Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, f(Y_i)\rangle\langle Y_i, g(Y_i)\rangle |g(Y_i)|^2 \Delta t_i \Delta B_i^3 + 12\langle Y_i, g(Y_i)\rangle^2 |f(Y_i)|^2 \Delta t_i^2(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, g(Y_i)\rangle^2 f^T(Y_i)g(Y_i)\Delta t_i \Delta B_i^3 + 12\langle Y_i, g(Y_i)\rangle^2 |g(Y_i)|^2(\Delta B_i^4 - \Delta t_i^2) \\
&+24\langle Y_i, f(Y_i)\rangle |f(Y_i)|^2 f^T(Y_i)g(Y_i)\Delta t_i^4 \Delta B_i + 36\langle Y_i, f(Y_i)\rangle |f(Y_i)|^2 |g(Y_i)|^2 \Delta t_i^3(\Delta B_i^2 - \Delta t_i) \\
&+24\langle Y_i, f(Y_i)\rangle f^T(Y_i)g(Y_i)|g(Y_i)|^2 \Delta t_i^2 \Delta B_i^3 + 6\langle Y_i, f(Y_i)\rangle |g(Y_i)|^4 \Delta t_i(\Delta B_i^4 - \Delta t_i^2) \\
&+6|f(Y_i)|^4 f^T(Y_i)g(Y_i)\Delta t_i^5 \Delta B_i + 15|f(Y_i)|^4 |g(Y_i)|^2 \Delta t_i^4(\Delta B_i^2 - \Delta t_i) \\
&+20|f(Y_i)|^2 f^T(Y_i)g(Y_i)|g(Y_i)|^2 \Delta t_i^3 \Delta B_i^3 \\
&+15|f(Y_i)|^2|g(Y_i)|^4 \Delta t_i^2(\Delta B_i^4 - \Delta t_i^2) + 6f^T(Y_i)g(Y_i)|g(Y_i)|^4 \Delta t_i \Delta B_i^5 + |g(Y_i)|^6(\Delta B_i^6 - \Delta t_i^3) \\
&+6\langle Y_i, g(Y_i)\rangle |f(Y_i)|^4 \Delta t_i^4 \Delta B_i + 24\langle Y_i, g(Y_i)\rangle |f(Y_i)|^3 |g(Y_i)| \Delta t_i^3(\Delta B_i^2 - \Delta t_i) \\
&+36\langle Y_i, g(Y_i)\rangle |f(Y_i)|^2 |g(Y_i)|^2 \Delta t_i^2 \Delta B_i^3 + 24\langle Y_i, g(Y_i)\rangle |f(Y_i)||g(Y_i)|^3 \Delta t_i(\Delta B_i^4 - \Delta t_i^2) \\
&+6\langle Y_i, g(Y_i)\rangle |g(Y_i)|^4 \Delta B_i^5)).
\end{aligned}
$$

In each step, we need to choose Δt_i such that $U_1(\Delta t_i, Y_i) < 0$. To do this, we could choose Δt_i such that

$$-U_2(\Delta t_i, Y_i) := -\frac{p}{2}v + A_1(Y_i)\Delta t_i + A_2(Y_i)\Delta t_i^2 + A_3(Y_i)\Delta t_i^3 + A_4(Y_i)\Delta t_i^4 + A_5(Y_i)\Delta t_i^5 < 0,$$

where

$$\begin{aligned} A_1(Y_i) &= \frac{p}{2}\frac{|f(Y_i)|^2}{|Y_i|^2} + \frac{p(2-p)}{8}\frac{4|Y_i||g(Y_i)|^3 + 8|Y_i||f(Y_i)|^2|g(Y_i)|}{|Y_i|^4} \\ &\quad + \frac{p(p-2)(p-4)}{2^3 \times 3!}\frac{24|Y_i|^3|f(Y_i)||g(Y_i)|^2 + 12|Y_i|^2|g(Y_i)|^4}{|Y_i|^6}, \end{aligned}$$

$$\begin{aligned} A_2(Y_i) &= \frac{p(2-p)}{8}\frac{4|Y_i||f(Y_i)|^3}{|Y_i|^4} + \frac{p(p-2)(p-4)}{2^3 \times 3!} \times \\ &\quad \Bigg(\frac{8|Y_i|^3|f(Y_i)|^3 + 12|Y_i|^2|f(Y_i)|^2|g(Y_i)|^2 + 48|Y_i|^2|f(Y_i)|^2|g(Y_i)|^2}{|Y_i|^6} \\ &\quad + \frac{12|Y_i|^2|f(Y_i)|^2|g(Y_i)|^2 + 6|Y_i||f(Y_i)||g(Y_i)|^4 + |g(Y_i)|^6 + 24|Y_i||f(Y_i)||g(Y_i)|^4}{|Y_i|^6}\Bigg), \end{aligned}$$

$$A_3(Y_i) = \frac{p(p-2)(p-4)}{2^3 \times 3!} \times \frac{12|Y_i|^2|f(Y_i)|^4 + 36|Y_i||f(Y_i)|^3|g(Y_i)|^2 + 15|f(Y_i)|^2|g(Y_i)|^4 + 24|Y_i|f(Y_i)|^3|g(Y_i)|^2}{|Y_i|^6},$$

$$A_4(Y_i) = \frac{p(p-2)(p-4)}{2^3 \times 3!}\frac{6|Y_i||f(Y_i)|^5 + 15|f(Y_i)|^4|g(Y_i)|^2}{|Y_i|^6},$$

and

$$A_5(Y_i) = \frac{p(p-2)(p-4)}{2^3 \times 3!}\frac{|f(Y_i)|^6}{|Y_i|^6}.$$

By the elementary inequality $\langle a, b\rangle \le |a||b|$, it is clear that $-U_1(\Delta t_i, Y_i) < -U_2(\Delta t_i, Y_i)$ a.s. We choose rational number Δt_i such that

$$\Delta t_i \le \frac{p}{12} \min_{\{A_j(Y_i)\neq 0, j=1,2,3,4,5\}} \{(v/A_j(Y_i))^{(1/j)}\}.$$

Apply the same techniques used in Theorem 5.3.1, we can prove that t_i is an $\{\mathcal{F}_t\}$-stopping time for each $i = 0, 1, ...$ and $\{m_i = \sum_{k=0}^{i} \Delta m_k\}_{i\ge 0}$ is a $\mathcal{G}_i$-local martingale. Now from (5.22), we have

$$|Y_{i+1}|^p \le |Y_0|^p - \sum_{k=0}^{i} \Delta t_k |Y_k|^p U_1(\Delta t_k, Y_k) + m_i.$$

By Lemma 5.2.3, we conclude

$$\lim_{i\to\infty} |Y_i|^p < \infty \quad \text{a.s.} \quad \text{and} \quad \sum_{k=0}^{i} \Delta t_k |Y_k|^p U_1(\Delta t_k, Y_k) < \infty \quad \text{a.s.}$$

Hence we have $\lim_{i\to\infty} \Delta t_i |Y_i|^p U_1(\Delta t_i, Y_i) = 0$ a.s. For almost all $\omega \in \Omega$, there exists $C(\omega) \in \mathbb{R}_+$ such that $\lim_{i\to\infty} |Y_i(\omega)| = C(\omega)$. Fix any such ω, write $C(\omega) = C$ and $Y_i(\omega) = Y_i$. Due to the choice of Δt_i, we have $U_1 > pv/12 > 0$. Since (5.18) and (5.19), applying the same techniques employed in Theorem 5.3.1 we have $\liminf_{i\to\infty} v/A_j(Y_i) > 0$ for each $j = 1, 2, 3, 4, 5$. That is to say there is no requirement that Δt_i vanishes as i increases, thus $\sum_{i=0}^{\infty} \Delta t_i = \infty$ a.s. Hence we can only have $\lim_{i\to\infty} |Y_i|^p = 0$. The proof is complete.

Because of the extra negative term in the condition (5.17), $2\langle x, f(x)\rangle + |g(x)|^2$ is not necessarily less than 0 for all nonzero x. Therefore Theorem 5.5.2 does cover some SDEs that can not be covered by Theorem 5.3.1. But it should be noted that Theorem 5.3.1 is not fully included in Theorem 5.5.2. For example a scalar SDE with $f(x) = -0.5x^3 - x^5$ and $g(x) = x^2$. We check the conditions (5.3) and (5.4) that for any $x \in \mathbb{R}^n$ with $x \neq 0$

$$2\langle x, f(x)\rangle + g^2(x) = -2x^6 < 0 \quad \text{and} \quad \liminf_{|x|\to 0} \frac{z(x)}{|f(x)|^2} = \frac{2}{0.25} > 0,$$

i.e. all the conditions in Theorem 5.3.1 hold. To check the condition (5.17) in Theorem 5.5.2, we have

$$\frac{2\langle x, f(x)\rangle + |g(x)|^2}{|x|^2} + (p-2)\frac{\langle x, g(x)\rangle^2}{|x|^4} = -x^4 + (p-2)x^2.$$

But for any $p \in (0, 2)$, we can not find a $v > 0$ to satisfy (5.17).

5.6 Conclusions and Future Research

In this chapter, we investigate the Euler–Maruyama method with random variable step size and successfully reproduce the almost sure stability of the true solution using this method with the semimartingale convergence theory. Conditions we impose on the drift and diffusion coefficients for the random variable step size method are much weaker than those for the fixed or nonrandom variable step size methods. Our key contribution also goes to the proof that the time variable is a stopping time, and only when this is true the rest of our proof is proper.

Considering that the random variable step size method works well for the stability, it is interesting to investigate other asymptotic properties of this method. Other numerical methods with random variable step size, such as the stochastic θ-method, are also worth to investigate.

Chapter 6

Strong Convergence of the Stopped Euler Method

6.1 Introduction

We have enjoyed the benefit brought by modifying the classical EM method in last chapter and this chapter sees an alternative modification. The current chapter is devoted to another important aspect of numerical analysis for SDEs, finite time strong convergence.

The convergence of numerical methods for stochastic differential equations (SDEs) has been broadly studied. Since the classical explicit Euler-Maruyama (EM) method has its simple algebraic structure, cheap computational cost and acceptable convergence rate under the global Lipschitz condition, its has been attracting lots of attention (Higham, 2011). Under the global Lipschitz condition, the convergence of the classical EM method is well established (see (Kloeden & Platen, 1992; Milstein & Tretyakov, 2004)). However, most SDE models in real life do not obey the global Lipschitz condition. In (Higham *et al.*, 2002), the authors studied the SDEs without global Lipschitz condition, and proved the strong convergence of the classical EM method under the assumption of pth moment boundedness of both true solution and numerical solution to the underlying SDE.

However, the authors in (Higham *et al.*, 2002) pointed out that in general, it is not clear when such moment bounds can be expected to hold for the EM method even when both drift coefficient and the diffusion coefficient are C^1 (unbounded derivatives of course). More recently, the authors in (Hutzenthaler *et al.*, 2011) answered the question negatively by proving that the explicit EM methods will diverge in finite time for those SDEs with either the drift coefficient or the diffusion coefficient being superlinear. The same group of authors then developed an explicit method to approximate SDEs with one-sided Lipschitz drift coefficient and the linear growth diffusion coefficient in (Hutzenthaler *et al.*, 2012). The method is called the tamed EM method, and to the best our knowledge it is the first explicit method to converge strongly to that type of SDEs. In (Wang & Gan, 2012), the authors used similar idea and constructed a higher order method, the tamed Milstein method.

Implicit methods are widely discussed as well. We refer here to the papers (Burrage & Tian, 2002; Higham *et al.*, 2002; Hu, 1996; Klauder & Petersen, 1985; Milstein *et al.*, 1998; Saito & Mitsui, 1993) and the book (Kloeden & Platen, 1992). Compare with the explicit methods, the implicit methods are better at tackling the nonlinear SDEs but are computationally expensive. Methods with variable step size also attract a lot of attention (Burrage & Burrage, 2002; Müller-Gronbach, 2002; Valinejad & Hosseini, 2010; Römisch & Winkler, 2006). Other weak forms of convergence, say weak convergence, convergence in probability and pathwise convergence, are discussed in (Anderson & Mattingly, 2011; Gyöngy, 1998; Jentzen *et al.*, 2009; Kloeden & Platen, 1992; Mao, 2011; Marion *et al.*, 2002; Milstein & Tretyakov, 2005; Wu *et al.*, 2008), just to mention a few.

We propose a new Euler-type method in this chapter and study the strong convergence of that method for one-dimensional SDEs in the form of

$$dx(t) = f(x(t))dt + g(x(t))dB(t), \quad x(0) = x_0 > 0, \tag{6.1}$$

where $B(t)$ be a scalar Brownian motion. We assume that $f(x(t))$ can be decomposed into two parts denoted by $f(x(t)) = h_1(x(t)) - h_2(x(t))$. Denote $\max(a, b)$ and $\min(a.b)$ by $a \vee b$ and $a \wedge b$ respectively. We impose several conditions on the drift and diffusion coefficients in the following.

For $h_1(x)$ and $g(x)$, we require them to satisfy the global Lipschitz condition, i.e. for $\forall x, y \in \mathbb{R}$ there exists a constant $K_1 \in \mathbb{R}^+$ such that

$$|h_1(x) - h_1(y)|^2 \vee |g(x) - g(y)|^2 \leq K_1|x - y|^2. \tag{6.2}$$

For $h_2(x)$, we need $h_2(x) \geq 0$ for all $x > 0$, and the polynomial growth condition, i.e. for $\forall x, y \in \mathbb{R}$, there exist constants $K_2 \in \mathbb{R}^+$ and $a \in \mathbb{Z}^+$ such that

$$|h_2(x) - h_2(y)|^2 \leq K_2(1 + |x|^a + |y|^a)|x - y|^2. \tag{6.3}$$

Also for $f(x) = h_1(x) - h_2(x)$, we assume for all $x, y \geq 0$ there exists a constant K_3 such that

$$(x - y)[f(x) - f(y)] \leq K_3|x - y|^2. \tag{6.4}$$

With further assuming $h_1(0) = h_2(0) = g(0) = 0$, we see that for $\forall x \in \mathbb{R}$

$$|h_1(x)|^2 \vee |g(x)|^2 \leq K_1|x|^2, \tag{6.5}$$

and for some constant $a > 0$

$$|h_2(x)|^2 \leq K_2(1 + |x|^{a+2}). \tag{6.6}$$

There are many models satisfying those conditions. For example by choosing $h_1(x(t)) = bx(t)$, $h_2(x(t)) = ax^2(t)$ and $g(x(t)) = \sigma x(t)$, we recover the stochastic Lotka-Volterra model (see for example, Chapter 11 of (Mao, 2008)). Also when $h_1(x(t)) = 0.5\sigma^2 x(t)$, $h_2(x(t)) = x^3(t)$ and $g(x(t)) = \sigma x(t)$ we get the stochastic Ginzburg-Landau equation (Ginzurg & Landau, 1950). Due to requirement of the nonnegativity of the solution for biological models and financial models, numerical solutions to these models need to be nonnegative as well.

For a given fixed time step Δt and the initial value $X_0 = x(0)$, the classic EM method for (6.1) is defined by

$$X_{i+1} = X_i + (h_1(X_i) - h_2(X_i))\Delta t + g(X_i)\Delta B_i, \quad i = 0, 1, 2, \cdots, \tag{6.7}$$

where $\Delta B_i = B(t_{i+1}) - B(t_i)$ is a Brownian motion increment and $t_i = i\Delta t$. In our analysis it will be more natural to work with the continuous version

$$X(t) = X_0 + \int_0^t [h_1(\bar{X}(s)) - h_2(\bar{X}(s))]ds + \int_0^t g(\bar{X}(s))dB(s), \tag{6.8}$$

where $\bar{X}(t)$ is defined by

$$\bar{X}(t) = X_i \quad \text{for } t \in [t_i, t_{i+1}).$$

The existing known results have only so far shown that the classic EM solutions converge to the true solution in probability (Mao, 2011; Marion *et al.*, 2002). Since the drift coefficient of (6.1) does not satisfy the linear growth condition, the theory established in (Higham *et al.*, 2002) is not applicable. Besides, as $f(x)$ satisfies the one-sided Lipschitz condition only for nonnegative x, the result of the tamed Euler method in (Hutzenthaler *et al.*, 2012) may not be applicable either. Moreover, the drift coefficient is in fact superlinear. Thus, according to the recent theory in (Hutzenthaler *et al.*, 2011), the classical EM method (6.7) will diverge in L^2 sense in finite time from the true solution of (6.1). All of these known results strongly indicate that the classical EM method is not good enough for the underlying SDE (6.1). However, the classical EM method has its great advantage due to its simple algebraic structure and cheap computational cost. The question is:

- Can we modify the classical EM method so that not only the modified method will preserve the simple algebraic structure and cheap computational cost of the classical EM method but also the approximate solutions based on the modified method will converge to the true solution in the strong sense (namely in L^2)?

To tackle this problem, we introduce a stopped EM method in this chapter. Firstly, the classical EM method (6.8) equipped with a stopping time is considered. Define the stopping time

$$\rho = \inf \{t \geq 0 : X(t) \leq 0\},$$

where throughout this chapter we set $\inf \emptyset = \infty$ (as usual $\emptyset$ denotes the empty set). It should be noted that the stopping time, ρ, depends on the continuous version of the approximate solution. Then the continuous approximate solution of the stopped EM method is defined by

$$X_\Delta(t) = X(t \wedge \rho), \quad t \geq 0, \tag{6.9}$$

while the corresponding discrete one is defined by

$$\bar{X}_\Delta(t) = \bar{X}(t \wedge \rho), \quad t \geq 0. \tag{6.10}$$

There are two benefits from the technique of stopping time. Firstly, the stopping time can guarantee the non-negativity of the numerical solution, i.e. $X_\Delta(t) \geq 0$ for all $t \geq 0$ almost surely. It is easy to see from (6.7) that the classical EM numerical solution could become negative due to the random effect of the Brownian motion. However, from Theorem 6.2.1 in Section 6.2 we can see that the underlying SDE solution is always positive. Hence, numerical methods that can preserve non-negativity are more desired. Next, and more importantly, both continuous and approximate solutions of the stopped EM method converge strongly to the true solution of the underlying SDE (6.1). We should mention that the preservation of non-negativity has been discussed in e.g. (Appleby *et al.*, 2010; Berkaoui *et al.*, 2008; Deelstra & Delbaen, 1998), and the technique of stopping time has widely been used to control diffusion processes (see e.g. (Gobet & Menozzi, 2004; Gobet & Menozzi, 2010)). We also mention the classical projection scheme that prevents the numerical solution from escaping from a closed domain (Pettersson, 2000) and its application on simulating reflected SDEs (Dangerfield *et al.*, 2012).

It should emphasized that the stopping time ρ functions on the continuous version of the method (6.8) which requires the whole Brownian path. However, in practice when we implement the stopped EM method in computer simulation we actually use a finite number of points of a Brownian path, thus a different stopping time which functions on the discrete version of the method $\bar{X}(t)$ is introduced and defined by

$$\bar{\rho} = \inf\left\{t \geq 0 : \bar{X}(t) \leq 0\right\}.$$

It is clear that $\rho \leq \bar{\rho}$ a.s. We denote the discrete version of the method with this stopping time by $\bar{X}^*_\Delta(t) = \bar{X}(t \wedge \bar{\rho})$. Our main result, Theorem 6.3.6, shows that $\bar{X}^*_\Delta(t)$ converges strongly to the true solution of the SDE (6.1) with a order arbitrarily close to a half. It should be emphasized that most of the existing results are about the strong convergence of the continuous version of the approximate solution, but in this chapter the main result is about the strong convergence of the discrete version of the approximate solution. It can be seen in Section 6.3 that, compared with the continuous version, to deal with the discrete version needs much more mathematical techniques for the stopped EM method.

This chapter is organized in the following way. In section 6.2, the positivity and the *pth* moment boundedness of the true solution to the underlying SDE will be discussed. The properties of our stopped EM method will be studied in detail in section 6.3, where several useful lemmas as well as our main results will be established. Section 6.4 presents the numerical simulations to demonstrate the convergence of our stopped EM method.

6.2 Properties of the Underlying SDE

Throughout this chapter, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (that is, it is right continuous and increasing while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets), and $\mathbb{E}$ denotes the expectation corresponding to $\mathbb{P}$. Let $B(t)$ be a scalar Brownian motion defined on the space. Define the σ-algebra generated by $y(t)$ by $\sigma(y(t))$ and we denote the conditional expectation by $\mathbb{E}_{t,y(t)}(\cdot) = \mathbb{E}(\cdot|\sigma(y(t)))$.

We state some known results about the underlying SDE solution that will be used in the proofs in next section.

Theorem 6.2.1 *Assume $h_1(0) = h_2(0) = g(0) = 0$ and, (6.2) and (6.4) hold, for any given initial value $x(0) > 0$, there exists a unique positive global solution $x(t)$ to the SDE (6.1) on $t \geq 0$.*

This is because the Feller's boundary classification assures the boundary 0 not accessible and the drift coefficient satisfies the one-sided Lipschitz condition (see for example, Theorem 2.3.6 and Lemma 4.3.2 in (Mao, 2008)).

Lemma 6.2.2 *Given $x_0 > 0$, for each $p > 2$ and $T > 0$ there is $A_1 = A_1(p, T, K_1) > 0$, such that*

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |x(t)|^p\right) \leq A_1 \mathbb{E}|x_0|^p.$$

Due to (6.2), (6.4), (6.5), (6.6) and Theorem 6.2.1, it is straightforward to prove this lemma (see for example, Theorem 2.4.1 in (Mao, 2008)).

6.3 Strong Convergence

Let us first present a number of useful lemmas before we prove our main results.

Lemma 6.3.1 *Given any initial value $X(0) > 0$, for each $p > 2$ and $T > 0$ there is A_2 dependent on p, T and K_1, but independent of Δt, such that*

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_\Delta(t)|^p\right) \leq A_2 \mathbb{E}|X(0)|^p.$$

Proof. By definition of $X_\Delta(t)$ and continuity of $X(t)$, we know $X_\Delta(t) = X(\rho) = 0$ for all $t > \rho$. For every integer $n \geq 1$, define the stopping time

$$\kappa_n = T \wedge \inf\{t \in [0, T] : |X(t)| \geq n\}.$$

Clearly, $\kappa_n \uparrow T$ a.s. By Itô's formula, using condition (6.5) and $h_2(x) > 0$ a.s. for all $x > 0$ a.s. we have, for $0 \leq t \leq \rho \wedge \kappa_n$

$$\begin{aligned} d|X(t)|^2 &= \left[2X(t)[h_1(\bar{X}(t)) - h_2(\bar{X}(t))] + |g(\bar{X}(t))|^2\right] dt + 2X(t)g(\bar{X}(t))dB(t) \\ &\leq \left[2\sqrt{K_1}X(t)\bar{X}(t) + K_1\bar{X}^2(t)\right] dt + 2X(t)g(\bar{X}(t))dB(t), \end{aligned}$$

Taking integration on both sides yields

$$\begin{aligned}|X(t\wedge\rho\wedge\kappa_n)|^2 &\le |X(0)|^2+\int_0^{t\wedge\rho\wedge\kappa_n}\left[2\sqrt{K_1}X(s)\bar{X}(s)+K_1\bar{X}^2(s)\right]ds\\ &\quad +\int_0^{t\wedge\rho\wedge\kappa_n}2X(s)g(\bar{X}(s))dB(s).\end{aligned}$$

For $t_1\in[0,T]$ and $p>2$,

$$\begin{aligned}\sup_{0\le t\le t_1\wedge\kappa_n}|X(t\wedge\rho)|^p &\le 4^{\frac{p}{2}-1}\Bigg[|X(0)|^p+(2\sqrt{K_1})^{\frac{p}{2}}\left(\int_0^{t_1\wedge\rho\wedge\kappa_n}X(s)\bar{X}(s)ds\right)^{\frac{p}{2}}\\ &\quad +K_1^{\frac{p}{2}}\left(\int_0^{t_1\wedge\rho\wedge\kappa_n}\bar{X}^2(s)ds\right)^{\frac{p}{2}}\\ &\quad +(2\sqrt{K_1})^{\frac{p}{2}}\sup_{0\le t\le t_1}\left|\int_0^{t\wedge\rho\wedge\kappa_n}X(s)\bar{X}(s)dB(s)\right|^{\frac{p}{2}}\Bigg].\end{aligned}$$

Taking expectation on both sides and by the Hölder inequality, we have

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\le t\le t_1}|X(t\wedge\rho\wedge\kappa_n)|^p\right)\\ \le\ & 4^{\frac{p}{2}-1}\Bigg[\mathbb{E}|X(0)|^p+(2\sqrt{K_1})^{\frac{p}{2}}t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}(X(s\wedge\rho\wedge\kappa_n)\bar{X}(s\wedge\rho\wedge\kappa_n))^{\frac{p}{2}}ds\\ &+K_1^{\frac{p}{2}}t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}(|\bar{X}(s\wedge\rho\wedge\kappa_n)|^p)ds\\ &+(2\sqrt{K_1})^{\frac{p}{2}}\mathbb{E}\left(\sup_{0\le t\le t_1}\left|\int_0^t X(s\wedge\rho\wedge\kappa_n)\bar{X}(s\wedge\rho\wedge\kappa_n)dB(s)\right|^{\frac{p}{2}}\right)\Bigg].\qquad(6.11)\end{aligned}$$

It is clear that

$$\begin{aligned}\mathbb{E}(X(s\wedge\rho\wedge\kappa_n)\bar{X}(s\wedge\rho\wedge\kappa_n))^{\frac{p}{2}} &\le \mathbb{E}(|X(s\wedge\rho\wedge\kappa_n)|^{\frac{p}{2}}|\bar{X}(s\wedge\rho\wedge\kappa_n)|^{\frac{p}{2}})\\ &\le \mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^{\frac{p}{2}}\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^{\frac{p}{2}}\right)\\ &\le \mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right),\end{aligned}$$

and

$$\mathbb{E}\left(|\bar{X}(s\wedge\rho\wedge\kappa_n)|^p\right)\le\mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right).$$

By the Burkholder-Davis-Gundy inequality

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\le t\le t_1}\left|\int_0^t X(s\wedge\rho\wedge\kappa_n)\bar{X}(s\wedge\rho\wedge\kappa_n)dB(s)\right|^{\frac{p}{2}}\right)\\ \le\ & C_p\mathbb{E}\left(\int_0^{t_1}|X(s\wedge\rho\wedge\kappa_n)|^2\bar{X}^2(s\wedge\rho\wedge\kappa_n)ds\right)^{\frac{p}{4}},\end{aligned}$$

where C_p depends on p only. By the elementary inequality and the Hölder inequality,

$$
\begin{aligned}
&\mathbb{E}\left(\int_0^{t_1} |X(s\wedge\rho\wedge\kappa_n)|^2\bar{X}^2(s\wedge\rho\wedge\kappa_n)ds\right)^{\frac{p}{4}} \\
\le\;& \mathbb{E}\left[\sup_{0\le s\le t_1}|X(s\wedge\rho\wedge\kappa_n)|^{\frac{p}{2}}\left(\int_0^{t_1}\bar{X}^2(s\wedge\rho\wedge\kappa_n)ds\right)^{\frac{p}{4}}\right] \\
\le\;& \frac{1}{2K}\mathbb{E}\left(\sup_{0\le s\le t_1}|X(s\wedge\rho\wedge\kappa_n)|^p\right)+\frac{K}{2}\mathbb{E}\left(\int_0^{t_1}\bar{X}^2(s\wedge\rho\wedge\kappa_n)ds\right)^{\frac{p}{2}} \\
\le\;& \frac{1}{2K}\mathbb{E}\left(\sup_{0\le s\le t_1}|X(s\wedge\rho\wedge\kappa_n)|^p\right)+\frac{K}{2}t_1^{\frac{p-2}{2}}\mathbb{E}\int_0^{t_1}|\bar{X}(s\wedge\rho\wedge\kappa_n)|^p ds \\
\le\;& \frac{1}{2K}\mathbb{E}\left(\sup_{0\le s\le t_1}|X(s\wedge\rho\wedge\kappa_n)|^p\right)+\frac{K}{2}t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right)ds.
\end{aligned}
$$

Hence, substituting these into (6.11) and choosing $K = 4^{p/2}(2\sqrt{K_1})^{p/2}C_p/6$ yields

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le t\le t_1}|X(t\wedge\rho\wedge\kappa_n)|^p\right) \\
\le\;& 4^{\frac{p}{2}-1}\Bigg[\mathbb{E}|X(0)|^p+\left((2\sqrt{K_1})^{\frac{p}{2}}+K_1^{\frac{p}{2}}\right)t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right)ds \\
&+(2\sqrt{K_1})^{\frac{p}{2}}C_p\frac{K}{2}t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right)ds\Bigg] \\
&+\frac{3}{4}\mathbb{E}\left(\sup_{0\le s\le t_1}|X(s\wedge\rho\wedge\kappa_n)|^p\right).
\end{aligned}
$$

Namely,

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le t\le t_1}|X(t\wedge\rho\wedge\kappa_n)|^p\right) \\
\le\;& 4^{\frac{p}{2}}\Bigg[\mathbb{E}|X(0)|^p+\left((2\sqrt{K_1})^{\frac{p}{2}}+K_1^{\frac{p}{2}}\right)t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right)ds \\
&+(2\sqrt{K_1})^{\frac{p}{2}}C_p\frac{K}{2}t_1^{\frac{p-2}{2}}\int_0^{t_1}\mathbb{E}\left(\sup_{0\le r\le s}|X(r\wedge\rho\wedge\kappa_n)|^p\right)ds\Bigg].
\end{aligned}
$$

Applying the Gronwall inequality and let $n \to \infty$, we obtain the required assertion with

$$A_2 = 4^{p/2}\exp[(4T)^{p/2}(K_1^p + (2\sqrt{K_1})^{p/2} + 0.5KC_p(2\sqrt{K_1})^{p/2})].$$

The proof is complete.

Since $\mathbb{E}(\sup_{0\le t\le T}|\bar{X}_\Delta(t)|^p) \le \mathbb{E}(\sup_{0\le t\le T}|X_\Delta(t)|^p)$, the moment bound for $\bar{X}_\Delta(t)$ is easy to obtain from Lemma 6.3.1. However, it should be emphasized that $\bar{X}^*_\Delta(t)$ is different from $\bar{X}_\Delta(t)$, as $\rho \le \bar{\rho}$ a.s. Thus we can not obtain a moment bound result for $\bar{X}^*_\Delta(t)$ directly, but need some extra effort. Before presenting the result about the moment bound of $\bar{X}^*_\Delta(t)$, we first prove the strong convergence of $X_\Delta(t)$ to $\bar{X}_\Delta(t)$ with the rate arbitrarily close to one half.

It is easy to see from (6.5), (6.6), Lemma 6.2.2 and Lemma 6.3.1 that there exists a constant A' dependent on q, A_1, A_2 and $x(0)$ such that for $\forall q \ge 1$ and $F = h_1, h_2, g$,

$$\mathbb{E}\left(\sup_{0\le t\le T}|F(x(t))|^q\right) \le A', \quad \mathbb{E}\left(\sup_{0\le t\le T}|F(X_\Delta(t))|^q\right) \le A',$$

and

$$\mathbb{E}\left(\sup_{0\le t\le T}|F(\bar{X}_\Delta(t))|^q\right) \le A'.$$

The technique used to deal with the expectation of Brownian motion increment with stopping time in the next lemma is essential and will be employed several times in the rest of the chapter. We emphasize that given any real-valued stopping time α, we do not have $\mathbb{E}|\Delta B(\alpha)|^2 = \mathbb{E}|B(\alpha+\Delta t) - B(\alpha)|^2 = \Delta t$, but need the technique of raising power to handle it. A similar approach was used in (Mao, 2011).

Lemma 6.3.2 *For any $T > 0$ and any integer $r \ge 2$, let $p > 1$ be any integer sufficiently large for*

$$\left(\frac{2rp}{2rp-1}\right)^r (T+1)^{\frac{1}{2p}} < 2. \tag{6.12}$$

Then

$$\mathbb{E}\left(\sup_{0\le t\le T}|X_\Delta(t) - \bar{X}_\Delta(t)|^r\right) \le G\Delta t^{\frac{r}{2}-\frac{1}{2p}}, \tag{6.13}$$

where $G = (2^{2r-1} + 2^{r+1}rp)(A'+1)$.

Proof. For $t \in [0, T\wedge\rho]$, let $i = i(t)$ be the integer part of $t/\Delta t$. So $t \in [t_i, t_{i+1})$ and

$$X_\Delta(t) - \bar{X}_\Delta(t) = (h_1(X_i) - h_2(X_i))(t - t_i) + g(X_i)(B(t) - B(t_i)).$$

Raising the power to r on both sides then taking the supremum and expectation, we compute, by the elementary inequality and the Hölder inequality, that

$$\begin{aligned}
& \mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|X_\Delta(t)-\bar{X}_\Delta(t)|^r\right) \\
= \ & \mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|(h_1(X_i)-h_2(X_i))(t-t_i)+g(X_i)(B(t)-B(t_i))|^r\right) \\
\le \ & \mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}2^{r-1}\left(|h_1(X_i)-h_2(X_i)|^r(t-t_i)^r+|g(X_i)|^r|B(t)-B(t_i)|^r\right)\right) \\
\le \ & \mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}2^{r-1}(2^{r-1}(|h_1(X_i)|^r+|h_2(X_i)|^r)(t-t_i)^r+|g(X_i)|^r|B(t)-B(t_i)|^r)\right) \\
\le \ & 2^{2r-2}\mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|h_1(X_i)|^r\right)\Delta t^r+2^{2r-2}\mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|h_2(X_i)|^r\right)\Delta t^r \\
& +2^{r-1}\mathbb{E}\left[\left(\sup_{0\le t\le T\wedge\rho}|g(X_i)|^r\right)\left(\sup_{0\le t\le T\wedge\rho}|B(t)-B(t_i)|^r\right)\right] \\
\le \ & 2^{2r-1}A'\Delta t^r+2^{r-1}\left[\mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|g(X_i)|^{2r}\right)\right]^{\frac{1}{2}}\left[\mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|B(t)-B(t_i)|^{2r}\right)\right]^{\frac{1}{2}} \\
\le \ & 2^{2r-1}A'\Delta t^r+2^{r-1}\sqrt{A'}\left[\mathbb{E}\left(\sup_{u=0,1,\dots N}\ \sup_{t_u\le t\le t_{u+1}\wedge T}|B(t)-B(t_u)|^{2rp}\right)\right]^{\frac{1}{2p}}, \qquad (6.14)
\end{aligned}$$

where N is the integer part of $T/\Delta t$. By the Doob martingale inequality,

$$\begin{aligned}
& \mathbb{E}\left(\sup_{u=0,1,\dots N}\ \sup_{t_u\le t\le t_{u+1}\wedge T}|B(t)-B(t_u)|^{2rp}\right) \\
\le \ & \sum_{u=0}^{N}\mathbb{E}\left(\sup_{t_u\le t\le t_{u+1}\wedge T}|B(t)-B(t_u)|^{2rp}\right) \\
\le \ & \left(\frac{2rp}{2rp-1}\right)^{2rp}\sum_{u=0}^{N}\mathbb{E}|B(t_{u+1}\wedge T)-B(t_u)|^{2rp} \\
\le \ & \left(\frac{2rp}{2rp-1}\right)^{2rp}\sum_{u=0}^{N}(2rp-1)!!\Delta t^{rp} \\
\le \ & \left(\frac{2rp}{2rp-1}\right)^{2rp}(2rp-1)!!(T+1)\Delta t^{rp-1},
\end{aligned}$$

where $(2rp-1)!!$ denotes the double factorial, i.e. $(2rp-1)!!=(2rp-1)(2rp-3)\cdots 3\cdot 1$. Substituting this into (6.14), and recalling (6.12) and $[(2rp-1)!!]^{1/p}\le$

$2rp$, we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le t\le T\wedge\rho}|X_\Delta(t)-\bar{X}_\Delta(t)|^r\right)\\
\le\ & 2^{2r-1}A'\Delta t^r+2^{r-1}\sqrt{A'}\left[\left(\frac{2rp}{2rp-1}\right)^{2rp}(2rp-1)!!(T+1)\Delta t^{rp-1}\right]^{\frac{1}{2p}}\\
\le\ & 2^{2r-1}A'\Delta t^r+2^{r-1}\sqrt{A'}\left(\frac{2rp}{2rp-1}\right)^{r}(T+1)^{\frac{1}{2p}}[(2rp-1)!!]^{\frac{1}{2p}}\Delta t^{\frac{rp-1}{2p}}\\
\le\ & 2^{2r-1}A'\Delta t^r+2^{r+1}\sqrt{A'}rp\Delta t^{\frac{r}{2}-\frac{1}{2p}}\\
\le\ & (2^{2r-1}+2^{r+1}rp)(A'+1)\Delta t^{\frac{r}{2}-\frac{1}{2p}}.
\end{aligned}$$

In the case where $\rho<T$, we have that for $\rho<t\le T$

$$\mathbb{E}\left(\sup_{\rho<t\le T}|X_\Delta(t)-\bar{X}_\Delta(t)|^r\right)=\mathbb{E}\left(|X_\Delta(\rho)-\bar{X}_\Delta(\rho)|^r\right)\le(2^{2r-1}+2^{r+1}rp)A'\Delta t^{\frac{r}{2}-\frac{1}{2p}}.$$

Hence the proof is complete.

Now we are ready to prove the pth moment boundedness of $\bar{X}^*_\Delta(t)$.

Lemma 6.3.3 *Given any initial value $X(0)>0$, for each integer $p\ge 2$ and $T>0$, there exists a constant A_3 dependent on T, p and K_1, but independent of Δt such that*

$$\mathbb{E}\left(\sup_{0\le t\le T}|\bar{X}^*_\Delta(t)|^p\right)<A_3\mathbb{E}|X_0|^p,$$

and for any $\varepsilon>0$

$$\mathbb{E}(\mathbf{1}_{\{\bar{\rho}\le T\}}|\bar{X}(\bar{\rho})|^2)=O(\Delta t^{1-\varepsilon}).$$

Proof. For any nonnegative real number α, we denote the integer part of it by $[\alpha]$. For every integer $n\ge 1$, define the stopping time

$$\bar{\kappa}_n=T\wedge\inf\{t\in[0,T]:|\bar{X}(t)|\ge n\}.$$

Clearly, $n\uparrow T$ a.s. For $0\le t<\bar{\rho}\wedge\bar{\kappa}_n$, define $\tau=[t/\Delta t]\Delta t$. Because (6.8) and $h_2(x)\ge 0$ for all $x>0$, we have that

$$\begin{aligned}
0<\bar{X}^*_\Delta(t)=X(\tau)\ &=\ X_0+\int_0^\tau[h_1(\bar{X}(s))-h_2(\bar{X}(s))]ds+\int_0^\tau g(\bar{X}(s))dB(s)\\
&\le\ X_0+\int_0^\tau h_1(\bar{X}(s))ds+\int_0^\tau g(\bar{X}(s))dB(s).
\end{aligned}$$

Then for any integer $p \geq 2$, raising the power of both sides to p we have

$$|\bar{X}^*_\Delta(t)|^p \leq 3^{p-1}\left(|X_0|^p + \left|\int_0^\tau h_1(\bar{X}(s))ds\right|^p + \left|\int_0^\tau g(\bar{X}(s))dB(s)\right|^p\right).$$

Taking supremum and expectation, for $t_1 \in [0,T]$ we have

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge t_1\wedge\bar{\kappa}_n}|\bar{X}^*_\Delta(t)|^p\right) &\leq 3^{p-1}\Bigg(\mathbb{E}|X_0|^p + \mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge t_1\wedge\bar{\kappa}_n}\left|\int_0^\tau h_1(\bar{X}(s))ds\right|^p\right)\\ &\quad+\mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge t_1\wedge\bar{\kappa}_n}\left|\int_0^\tau g(\bar{X}(s))dB(s)\right|^p\right)\Bigg)\\ &\leq 3^{p-1}\Bigg(\mathbb{E}|X_0|^p + \mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge t_1\wedge\bar{\kappa}_n}\left|\int_0^t h_1(\bar{X}(s))ds\right|^p\right)\\ &\quad+\mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge t_1\wedge\bar{\kappa}_n}\left|\int_0^t g(\bar{X}(s))dB(s)\right|^p\right)\Bigg).\end{aligned}$$

By (6.5), the Hölder inequality and the Burkholder-Davis-Gundy inequality, we obtain

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge\bar{\kappa}_n}|\bar{X}^*_\Delta(t)|^p\right)\\ \leq\;& 3^{p-1}\mathbb{E}|X_0|^p + 3^{p-1}K_1^{\frac{p}{2}}T^{\frac{p}{2}-1}\left(T^{\frac{p}{2}} + \left(\frac{p^3}{2(p-1)}\right)^{\frac{p}{2}}\right)\int_0^T \mathbb{E}\left(\sup_{0\leq r\leq s\wedge\bar{\rho}\wedge\bar{\kappa}_n}|\bar{X}(r)|^p\right)ds.\end{aligned}$$

Then applying the Gronwall inequality and letting $n \to \infty$ give

$$\mathbb{E}\left(\sup_{0\leq t\leq\bar{\rho}\wedge T}|\bar{X}^*_\Delta(t)|^p\right) < A_3\mathbb{E}|X_0|^p, \tag{6.15}$$

where $A_3 = 3^{p-1}\exp\{3^{p-1}K_1^{p/2}T^{p/2-1}(T^{p/2} + (p^3/(2(p-1)))^{p/2}T\}$.

By the definition of $\bar{\rho}$, we know $\bar{\rho}$ is a multiple of Δt, and denote $\bar{\rho}/\Delta t$ by $n_{\bar{\rho}}$. Since $X_{n_{\bar{\rho}}-1} > 0$ and $X_{n_{\bar{\rho}}} \leq 0$ when $\omega \in \{\bar{\rho} \leq T\}$, we have

$$\begin{aligned}X_{n_{\bar{\rho}}} &= X_{n_{\bar{\rho}}-1} + (h_1(X_{n_{\bar{\rho}}-1}) - h_2(X_{n_{\bar{\rho}}-1}))\Delta t + g(X_{n_{\bar{\rho}}-1})\Delta B_{n_{\bar{\rho}}-1}\\ &> (h_1(X_{n_{\bar{\rho}}-1}) - h_2(X_{n_{\bar{\rho}}-1}))\Delta t + g(X_{n_{\bar{\rho}}-1})\Delta B_{n_{\bar{\rho}}-1}.\end{aligned}$$

Taking square and expectation on both sides, we have

$$\begin{aligned}\mathbb{E}(\mathbf{1}_{\{\bar{\rho}\leq T\}}X^2_{n_{\bar{\rho}}}) &\leq 4\Delta t^2(\mathbb{E}(\mathbf{1}_{\{\bar{\rho}\leq T\}}|h_1(X_{n_{\bar{\rho}}-1})|^2) + \mathbb{E}(\mathbf{1}_{\{\bar{\rho}\leq T\}}|h_2(X_{n_{\bar{\rho}}-1})|^2))\\ &\quad+2\mathbb{E}(\mathbf{1}_{\{\bar{\rho}\leq T\}}|g(X_{n_{\bar{\rho}}-1})|^2|\Delta B_{n_{\bar{\rho}}-1}|^2)\\ &\leq 4\Delta t^2(\mathbb{E}|h_1(X_{n_{\bar{\rho}}-1})|^2 + \mathbb{E}|h_2(X_{n_{\bar{\rho}}-1})|^2)\\ &\quad+2\mathbb{E}(|g(X_{n_{\bar{\rho}}-1})|^2)\mathbb{E}(|\Delta B_{n_{\bar{\rho}}-1}|^2),\end{aligned}$$

where $\mathbf{1}_A$ is the indicator function of A. The Brownian motion increment in the last term involves the stopping time $\bar{\rho}$, thus we employ the technique of raising power used in Lemma 6.3.2 here and by a similar approach we can show that $\mathbb{E}(|\Delta B_{n_{\bar{\rho}}-1}|^2) = O(\Delta t^{1-\varepsilon})$ for any $\varepsilon > 0$. Let $q > 1$ be any integer sufficiently large for

$$\left(\frac{2q}{2q-1}\right)^2 (T+1)^{\frac{1}{q}} < 2,$$

by the Hölder inequality, the Doob martingale inequality and $[(2q-1)!!]^{1/q} \leq 2q$ we have

$$\begin{aligned}
\mathbb{E}(|\Delta B_{n_{\bar{\rho}}-1}|^2) &= \mathbb{E}(|B(\bar{\rho}) - B(\bar{\rho} - \Delta t)|^2) \\
&\leq \mathbb{E}\left(\sup_{0\leq t\leq \bar{\rho}\wedge T} |B(t) - B(t-\Delta t)|^2\right) \\
&\leq \left[\mathbb{E}\left(\sup_{u=1,\ldots,N} \sup_{t_u \leq t \leq t_{u+1}\wedge T} |B(t) - B(t_u)|^{2q}\right)\right]^{\frac{1}{q}} \\
&\leq \left[\sum_{u=1}^{N} \mathbb{E}\left(\sup_{t_u \leq t \leq t_{u+1}\wedge T} |B(t) - B(t_u)|^{2q}\right)\right]^{\frac{1}{q}} \\
&\leq \left[\left(\frac{2q}{2q-1}\right)^{2q} \sum_{u=1}^{N} \mathbb{E}|B(t_{u+1}\wedge T) - B(t_u)|^{2q}\right]^{\frac{1}{q}} \\
&\leq \left[\left(\frac{2q}{2q-1}\right)^{2q} \sum_{u=1}^{N} (2q-1)!!\Delta t^q\right]^{\frac{1}{q}} \\
&\leq \left[\left(\frac{2q}{2q-1}\right)^{2q} (2q-1)!!(T+1)\Delta t^{q-1}\right]^{\frac{1}{q}} \\
&\leq 4q\Delta t^{1-\frac{1}{q}}.
\end{aligned}$$

The required rate follows. By (6.5), (6.6) and (6.15), we know all the three terms, $\mathbb{E}|h_1(X_{n_{\bar{\rho}}-1})|^2$, $\mathbb{E}|h_2(X_{n_{\bar{\rho}}-1})|^2$ and $\mathbb{E}|g(X_{n_{\bar{\rho}}-1})|^2$, are bounded by some finite number independent of Δt, so

$$\mathbb{E}(\mathbf{1}_{\{\bar{\rho}\leq T\}}|\bar{X}(\bar{\rho})|^2) = \mathbb{E}(\mathbf{1}_{\{\bar{\rho}\leq T\}}|X_{n_{\bar{\rho}}}|^2) = O(\Delta t^{1-\varepsilon}). \tag{6.16}$$

Hence the proof is complete.

It is straightforward to adapt the proof above to show that for a fixed $T > 0$ and any integer k_0 such that $k_0\Delta t \in (0, T)$ we have

$$\mathbb{E}_{k_0\Delta t, \bar{X}(k_0\Delta t)} \left(\sup_{k_0\Delta t \le t \le T} |\bar{X}^*_\Delta(t)|^p \right) < A_3 |\bar{X}(k_0\Delta t)|^p. \tag{6.17}$$

Lemma 6.3.4 *For $\forall \epsilon > 0$, the continuous approximate solution (6.9) of the stopped EM method satisfies*

$$\mathbb{E} \left(\sup_{0 \le t \le T} |x(t) - X_\Delta(t)|^2 \right) = O(\Delta t^{1-\epsilon}). \tag{6.18}$$

Proof. Set $e(t) = x(t) - X_\Delta(t)$. By Itô formula and inequalities (6.2), (6.3) and (6.4), for $t \in [0, \rho \wedge T]$ we compute

$$\begin{aligned}
|e(t)|^2 &= \int_0^t 2(f(x(s)) - f(\bar{X}_\Delta(s)))e(s)ds + \int_0^t |g(x(s)) - g(\bar{X}_\Delta(s))|^2 ds \\
&\quad + M(t) \\
&\le \int_0^t \left[2(f(x(s)) - f(X_\Delta(s)))e(s) + K_1|x(s) - \bar{X}_\Delta(s)|^2 \right] ds \\
&\quad + 2\int_0^t (f(X_\Delta(s)) - f(\bar{X}_\Delta(s)))e(s)ds + M(t) \\
&\le 2\int_0^t (K_3|e(s)|^2 + K_1|e(s)|^2 + K_1|X_\Delta(s) - \bar{X}_\Delta(s)|^2)ds \\
&\quad + \int_0^t (|f(X_\Delta(s)) - f(\bar{X}_\Delta(s))|^2 + |e(s)|^2)ds + M(t) \\
&\le (1 + 2(K_1 + K_3)) \int_0^t |e(s)|^2 ds \\
&\quad + \int_0^t D'(1 + |X_\Delta(s)|^a + |\bar{X}_\Delta(s)|^a)|X_\Delta(s) - \bar{X}_\Delta(s)|^2 ds + M(t),
\end{aligned}$$

where $M(t) = \int_0^t 2e(s)(g(x(s)) - g(\bar{X}_\Delta(s)))dB(s)$ and D' is a positive constant dependent only on K_1 and K_2, which may change from line to line in the following proof. Using the Cauchy-Schwarz and Lemma 6.3.2 with $r = 4$, we further compute

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}|e(s)|^2\right)\\
\le\ & (1+2(K_1+K_3))\int_0^t \mathbb{E}|e(s)|^2ds\\
&+\int_0^t D'\Big[\mathbb{E}(1+|X_\Delta(s)|^a+|\bar{X}_\Delta(s)|^a)^2\mathbb{E}|X_\Delta(s)-\bar{X}_\Delta(s)|^4\Big]^{\frac{1}{2}}ds+m(t)\\
\le\ & (1+2(K_1+K_3))\int_0^t \mathbb{E}|e(s)|^2ds\\
&+GD'\Delta t^{1-\frac{1}{4p}}\int_0^t \mathbb{E}(1+|X_\Delta(s)|^{2a}+|\bar{X}_\Delta(s)|^{2a})ds+m(t)\\
\le\ & (1+2(K_1+K_3))\int_0^t \mathbb{E}|e(s)|^2ds+GD'T(1+2A')\Delta t^{1-\frac{1}{4p}}+m(t),
\end{aligned}$$

where $m(t) = \mathbb{E}(\sup_{0\le s\le t}|M(s)|)$. But, by the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned}
m(t) \le\ & 16\mathbb{E}\left[\int_0^t |e(s)|^2|g(x(s))-g(\bar{X}_\Delta(s))|^2ds\right]^{\frac{1}{2}}\\
\le\ & 16\mathbb{E}\left[\sup_{0\le s\le t}|e(s)|^2\int_0^t K_1|x(s)-\bar{X}_\Delta(s)|^2ds\right]^{\frac{1}{2}}\\
\le\ & \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}|e(s)|^2\right)+128K_1\mathbb{E}\int_0^t |x(s)-\bar{X}_\Delta(s)|^2ds\\
\le\ & \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}|e(s)|^2\right)+256K_1\int_0^t (\mathbb{E}|e(s)|^2+\mathbb{E}|X_\Delta(s)-\bar{X}_\Delta(s)|^2)ds\\
\le\ & \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}|e(s)|^2\right)+256K_1\int_0^t \mathbb{E}|e(s)|^2ds+GT\Delta t^{1-\frac{1}{2p}}.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}|e(s)|^2\right)\\
\le\ & (1+258K_1+2K_3)\int_0^t \mathbb{E}\left(\sup_{0\le r\le s}|e(r)|^2\right)ds+(GD'(1+2A')+G)T\Delta t^{1-\frac{1}{2p}}.
\end{aligned}$$

Since p can be arbitrarily large, by the Gronwall inequality, we see

$$\mathbb{E}\left(\sup_{0\le t\le \rho\wedge T}|e(t)|^2\right) = O(\Delta t^{1-\epsilon}). \tag{6.19}$$

In the case where $\rho < T$, by the strong Markov property and (6.19) we have that

$$\begin{aligned}
\mathbb{E}(\sup_{\rho<t\leq T}|x(t)-X(t\wedge\rho)|^2) &= \mathbb{E}(\sup_{\rho<t\leq T}|x(t)-X(\rho)|^2) \\
&= \mathbb{E}(\sup_{\rho<t\leq T}|x(t)|^2) \\
&\leq \mathbb{E}\left[\mathbb{E}\left(\sup_{\rho<t\leq T}|x(t)|^2\Big|\sigma(x(\rho))\right)\right] \\
&\leq \mathbb{E}\left[\mathbb{E}_{\rho,x(\rho)}\left(\sup_{\rho<t\leq T}|x(t)|^2\right)\right] \\
&\leq \mathbb{E}\left(A_1|x(\rho)|^2\right) \\
&= A_1\mathbb{E}\left(A_1|x(\rho)-X(\rho)|^2\right) \\
&= O(\Delta t^{1-\epsilon}).
\end{aligned}$$

The proof is therefore complete.

The next theorem follows from Lemma 6.3.2, Lemma 6.3.4 and the triangle inequality directly.

Theorem 6.3.5 *For $\forall\epsilon > 0$, the discrete approximate solution (6.10) of the stopped EM method satisfies*

$$\mathbb{E}\left(\sup_{0\leq t\leq T}|x(t)-\bar{X}_\Delta(t)|^2\right) = O(\Delta t^{1-\epsilon}).$$

However due to the definition of the stopping time ρ, neither $X_\Delta(t)$ nor $\bar{X}_\Delta(t)$ can be implemented in practice. So it is actually the approximate solution $\bar{X}^*_\Delta(t)$ that we use for simulation. The following theorem is therefore more important.

Theorem 6.3.6 *Assume (6.2), (6.3), (6.4), (6.5) and (6.6) hold, then for $\forall\epsilon > 0$ the approximate solution $\bar{X}^*_\Delta(t)$ converges strongly to the true solution $x(t)$ with order $1-\epsilon$,*

$$\mathbb{E}\left(\sup_{0\leq t\leq T}|x(t)-\bar{X}^*_\Delta(t))|^2\right) = O(\Delta t^{1-\varepsilon}).$$

Proof. To prove the assertion, by triangle inequality we need to show

$$\mathbb{E}\left(\sup_{0\leq t\leq T}|x(t)-X(t\wedge\rho)|^2\right) = O(\Delta t^{1-\varepsilon}), \tag{6.20}$$

and

$$\mathbb{E}\left(\sup_{0\le t\le T}|X(t\wedge\rho)-\bar{X}(t\wedge\bar{\rho})|^2\right)=O(\Delta t^{1-\varepsilon}). \tag{6.21}$$

We already proved (6.20) in Lemma 6.3.4, now we try to prove (6.21). By the definition of ρ, we know that $X(t\wedge\rho)=X(\rho)=0$ for $\rho\le t\le T$. It is clear that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le t\le T}|X(t\wedge\rho)-\bar{X}(t\wedge\bar{\rho})|^2\right)\\
\le\ &\mathbb{E}\left(\sup_{0\le t\le\rho\wedge T}|X(t)-\bar{X}(t)|^2\right)+\mathbb{E}\left(\sup_{\rho\wedge T<t\le\bar{\rho}\wedge T}|X(\rho)-\bar{X}(t)|^2\right)\\
&+\mathbb{E}\left(\sup_{\bar{\rho}\wedge T<t\le T}|X(\rho)-\bar{X}(\bar{\rho})|^2\right).
\end{aligned}$$

According to Lemma 6.3.2, the first term on the right hand side of the inequality equals $O(\Delta t^{1-\varepsilon})$.

Now we consider the second term. For $\omega\in\{\rho<T\}$, denote the next time point larger than ρ by $\tau_1=([\rho/\Delta t]+1)\Delta t$, by (6.8) for $\rho\wedge T<t\le\tau_1\wedge T$, we have

$$X(t)=X(\rho)+\int_{\rho\wedge T}^{t}[h_1(\bar{X}(s))-h_2(\bar{X}(s))]ds+\int_{\rho\wedge T}^{t}g(\bar{X}(s))dB(s).$$

Due to Lemma 6.3.3, (6.5) and (6.6), we know h_1, h_2 and g are all bounded, together with $\tau_1-\rho\le\Delta t$ a.s. it is easy to obtain

$$\mathbb{E}\left(\sup_{\rho\wedge T<t\le\tau_1\wedge T}|\bar{X}(t)|^2\right)=O(\Delta t^{1-\varepsilon}).$$

The degradation of the rate is again due to the stopping time in the Brownian motion increment and the application of the technique of raising power. Then by

(6.17) and the strong Markov property we obtain

$$\begin{aligned}
& \mathbb{E}\left(\sup_{\rho\wedge T<t\leq\bar\rho\wedge T}|X(\rho)-\bar X(t)|^2\right) \\
\leq\ & \mathbb{E}\left(\sup_{\tau_1\wedge T<t\leq\bar\rho\wedge T}|\bar X(t)|^2\right)+\mathbb{E}\left(\sup_{\rho\wedge T<t\leq\tau_1\wedge T}|\bar X(t)|^2\right) \\
\leq\ & \mathbb{E}\left[\mathbb{E}\left(\sup_{\tau_1\wedge T<t\leq\bar\rho\wedge T}|\bar X(t)|^2\Big|\sigma(\bar X(\tau_1\wedge T))\right)\right]+O(\Delta t^{1-\varepsilon}) \\
\leq\ & \mathbb{E}\left[\mathbb{E}_{\tau_1\wedge T,\bar X(\tau_1\wedge T)}\left(\sup_{\tau_1\wedge T<t\leq\bar\rho\wedge T}|\bar X(t)|^2\right)\right]+O(\Delta t^{1-\varepsilon}) \\
\leq\ & \mathbb{E}\left(A_3|\bar X(\tau_1\wedge T)|^2\right)+O(\Delta t^{1-\varepsilon}) \\
=\ & O(\Delta t^{1-\varepsilon}).
\end{aligned}$$

Also by Lemma 6.3.3 we have

$$\mathbb{E}\left(\sup_{\bar\rho\wedge T<t\leq T}|X(\rho)-\bar X(\bar\rho)|^2\right)=\mathbb{E}\left(\mathbf{1}_{\{\bar\rho\leq T\}}|\bar X(\bar\rho)|^2\right)=O(\Delta t^{1-\varepsilon}).$$

Hence we can obtain (6.21). The proof is therefore complete.

6.4 Numerical Simulation

In this section we present two SDEs and their numerical simulations to illustrate the strong convergence as well as the convergence rate of the stopped EM method. We choose the SDEs with their explicit solutions in order to test the efficiency of our stopped EM method.

Firstly, we consider the stochastic Lotka-Volterra model (see for example, Chapter 11 of (Mao, 2008)), namely the SDE (6.1) with $f(x(t)) = bx(t) - ax^2(t)$ and $g(x(t)) = \sigma x(t)$, where $a, b, \sigma > 0$. Obviously, the assumptions of f, h_1, h_2 and g are satisfied. In (Oksendal, 2003), the explicit solution is expressed by

$$x(t) = \frac{x_0\exp((b-\frac{1}{2}\sigma^2)t+\sigma B(t))}{1+ax_0\int_0^t\exp((b-\frac{1}{2}\sigma^2)s+\sigma B(s))ds}. \tag{6.22}$$

Although the integral in the denominator cannot be computed analytically, we will use a very small step size, say 10^{-6}, to approximate it. The approximate solution obtained in this way is regarded as the true solution in the following numerical tests.

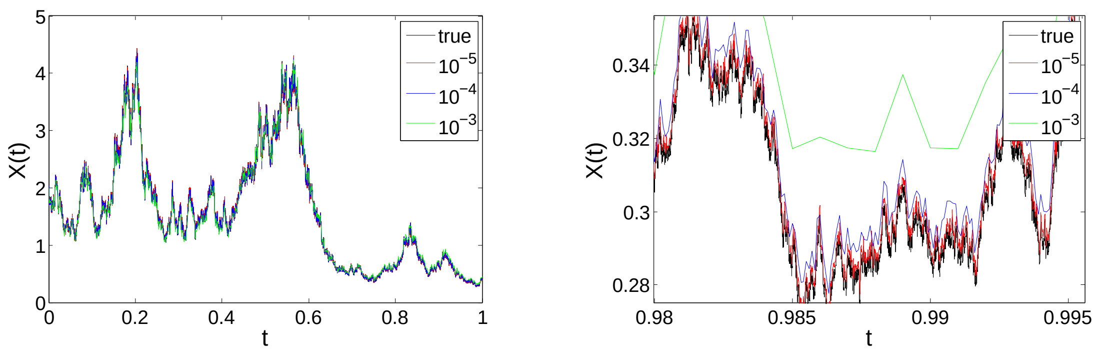

Figure 6.1: Simulation of the EM solutions with step size 10^{-5}, 10^{-4} and 10^{-3} respectively. The coefficients, $a = 1$ $b = 1$ $\sigma = 2$. Left: the true solution and the EM solution on $[0, 1]$. Right: zoomed in plot of the left one.

We set $a = 1$, $b = 1$ and $\sigma = 2$ for all the experiments in this section. The stopped EM solutions with step sizes of 10^{-5}, 10^{-4} and 10^{-3} respectively are plotted in Figure 6.1. The initial value is $X(0) = 2$. It is difficult to distinguish the true solution from the three simulated solutions on the left plot, which indicates that all the three step sizes can result in a good approximation. To find which step size is more precise, a detailed graph is plotted. It can be seen that with the step size decreasing the stopped EM solution makes better approximation to the true solution. In the simulation, if the numerical solution touches 0 or becomes negative at some time point, we set the numerical solution at this time point and all the time points afterwards be zero.

We calculate the strong error by

$$e = \mathbb{E}\left(\max_{i=1,2,\ldots,n} |x(i\Delta t) - \bar{X}^*_\Delta(i\Delta t)|\right).$$

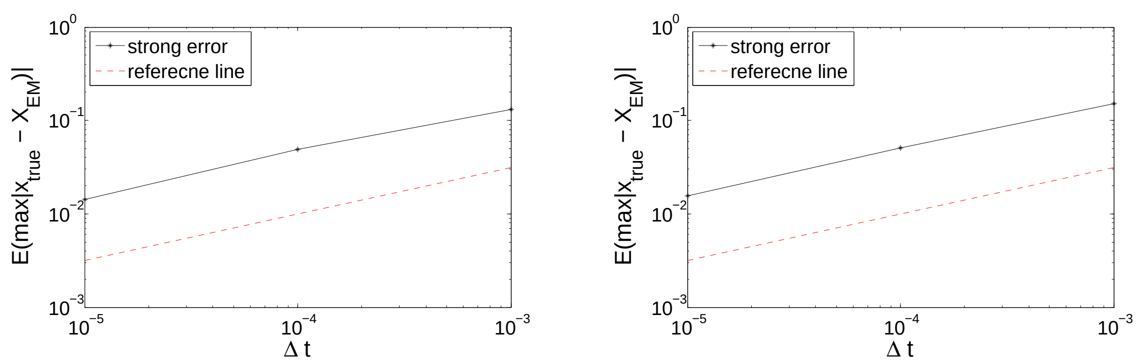

Figure 6.2: Left: The strong error plot for the stochastic Lotka-Volterra equation. Right: The strong error plot for the stochastic Ginzburg-Landau equation. The dashed line of slope 0.5 is the reference slope.

We carry 1000 simulations for the true solution and each of the three stopped EM solution. The strong error is calculated based on those simulations. In Figure 6.2, the strong error plot on the left indicates better approximation with smaller step size, which is in line with the right plot in Figure 6.1. In addition, compared with the dashed line of slope 0.5, we obverse the order of about 0.5 for the strong convergence.

Next we consider the stochastic Ginzburg-Landau equation (Ginzurg & Landau, 1950)

$$dx(t) = \left(\frac{1}{2}\sigma^2 x(t) - x^3(t)\right) dt + \sigma x(t) dB(t), \tag{6.23}$$

to which the true solution is known (Kloeden & Platen, 1992)

$$x(t) = \frac{x_0 \exp(\sigma B(t))}{\sqrt{1 + 2x_0^2 \int_0^t \exp(2\sigma B(s)) ds}}. \tag{6.24}$$

The same simulation as the stochastic Lotka-Volterra equation is carried with $\sigma = 7$. The strong error is on the right plot of Figure 6.2. It can be seen that the order of the strong error is around 0.5 as well, which is in line with Theorem 6.3.6. Equation (6.23) was also considered in (Hutzenthaler *et al.*, 2011) to illustrate the divergence of the classical EM method. In (Hutzenthaler *et al.*, 2011), the authors found that, for the classical EM method with fixed step size, the second moment of the classical EM solution at the terminal time point approaches the second moment of the true solution very poorly for $\sigma = 5, 6$ and even is explosive for $\sigma = 7$. Since the finite time moment boundedness is key to the strong convergence, in Table 6.1 we display the simulation results for the stopped EM method at the terminal point $T = 3$. The initial value is $X(0) = 1$ and the step size is $\Delta t = 1 \times 10^{-5}$.

It can be seen that the stopped EM performs well in approaching the second moment of the true solution for all the σ and no explosion occurs in such a large number of Monte Carlo runs.

Hence from both the theoretical results and the simulations, we can conclude that the stopped EM method outperforms the classical EM method greatly for the SDE (6.1) and the use of the stopping time contributes mostly to that outperformance. One of our future work is to extend our theory to multi-dimensional nonlinear SDEs.

The work contained in this chapter has been published, and we refer the readers to (Liu & Mao, 2013b) for the published version.

Table 6.1: Simulation of $\mathbb{E}(x(3))^2$ and $\mathbb{E}(\bar{X}^*_\Delta(3))^2$ for the stochastic Ginzburg-Landau equation where the step size is $\Delta t = 1 \times 10^{-5}$. The number of Monte Carlo runs is 10^5 for both the stopped EM solution and the exact solution.

	$\sigma = 2$	$\sigma = 3$	$\sigma = 4$	$\sigma = 5$	$\sigma = 6$	$\sigma = 7$
exact solution $E[x^2(3)]$	0.4740	0.6917	0.9233	1.1478	1.3960	1.5878
simulation 1	0.4640	0.6847	0.9166	1.1138	1.3286	1.5854
simulation 2	0.4721	0.6919	0.8963	1.1302	1.3364	1.6377
simulation 3	0.4784	0.7101	0.9329	1.1373	1.3521	1.6360
simulation 4	0.4700	0.6974	0.9088	1.1386	1.3871	1.6280
simulation 5	0.4706	0.6895	0.9125	1.1127	1.3565	1.5657
simulation 6	0.4666	0.6927	0.9187	1.1452	1.3724	1.5977
simulation 7	0.4638	0.6947	0.9255	1.1596	1.3946	1.6391
simulation 8	0.4757	0.7081	0.9442	1.1824	1.4218	1.6706
simulation 9	0.4677	0.6947	0.9240	1.1568	1.3909	1.6026
simulation 10	0.4695	0.6823	0.9050	1.1287	1.3521	1.5926

Chapter 7

Conclusions and Future Research

7.1 Conclusions

Two important aspects of numerical analysis for stochastic differential equations, the asymptotic properties and the finite time convergence, have been investigated in this thesis. The classical explicit Euler-Maruyama method and the (implicit) backward Euler-Maruyama method are the fundamental schemes for this thesis. Besides, we developped two new schemes by modifying the classical method: the Euler-Maruyama method with random variable step size was introduced to study the almost sure stability and the stopped Euler method was developped to cope with the finite time strong convergence.

In Chapter 3, we presented our observations on the asymptotic boundedness for numerical solutions. Theorem 3.3.2 and Theorem 3.4.3 state our findings on the asymptotic boundedness in small moment for the EM method and the backward EM method, respectively. The EM method works well when the linear growth conditions hold and the backward EM method is a good replacement for it when the linear growth condition on the drift coefficient is violated. However, in the case of small moment, we have only been able to reproduce the asymptotic boundedness qualitatively, but not the exact value of the bound. Then in Theorem 3.5.2 and Theorem 3.5.4, by strengthening some conditions we successfully reproduced the asymptotic boundedness both qualitatively and quantitatively for the second moment.

Chapter 4 could be treated as a continuous study of Chapter 3, as the moment boundedness obtained in Chapter 3 is essential for the study of the numerical stationary distribution. The coefficient-related conditions for the existence and uniqueness of the stationary distribution of the backward Euler-Maruyama method were given in Lemmas 4.3.1, 4.3.2 and 4.3.3. Then we further studied the convergence of the numerical stationary distribution to the stationary distribution of the underlying solution and stated the main result in Theorem 4.3.9. The numerical simulations in Section 4.4 are in line with the theoretical results. In addition, we observed that the numerical stationary distributions could be used as numerical solutions to certain type of deterministic differential equations.

The EM method with random variable step size was introduced in Chapter 5. The first almost surely stability theorem of it was given in Theorem 5.3.1. To our best knowledge, Chapter 5 is the first work to apply the random variable step size (with clear proof of the stopping time) to the analysis of the almost sure stability of the EM method. Compare with the existing results, by employing the random variable step size the new scheme is able to reproduce the almost sure stability of much larger range of SDEs. Other sufficient conditions were provided in Theorems 5.3.2, 5.5.1 and 5.5.2, which make the new scheme more applicable.

Having tasted the sweet of modification of the classical method in Chapter 5, we considered to use the modified EM method to approximate the underlying solution in finite time in Chapter 6. The idea of embedding a stopping time into the classical EM method was initially adopted to preserve the non-negativity of the numerical solution, and it turned out that the non-negativity in return guarantees the scheme to converge to the underlying solution strongly with the rate a half. Compare with the classical method, the stopped EM method can cover highly nonlinear SDEs with just a little computational cost added. The main results were stated in Theorems 6.3.5 and 6.3.6.

7.2 Future Research

As we mentioned at the beginning chapter of this thesis, the study on the numerical solutions of SDEs is far behind its underlying counterpart. Those results presented in this thesis is just a tip of the iceberg, and there is still a boundless ocean to explore.

This thesis focuses on stochastic differential equations, and it is worth to investigate if those results obtained still hold for, such as stochastic functional differential equations and stochastic differential equations with jumps.

From Chapters 5 and 6, the non-fixed step size methods look promising. As we mentioned at the start of Chapter 5, there already exist some works on the adapted step size methods showing better performance than the constant step size methods. However, the methods with constant step size still dominate the literatures. Therefore, the study on the adapted method is an interesting direction.

References

Allen, E. 2007. *Modeling with Itô stochastic differential equations.* Mathematical Modelling: Theory and Applications, vol. 22. Dordrecht: Springer.

Anderson, D. F., & Mattingly, J. C. 2011. A weak trapezoidal method for a class of stochastic differential equations. *Commun. Math. Sci.*, **9**(1), 301–318.

Appleby, J. A. D., & Rodkina, A. 2009. Stability of nonlinear stochastic Volterra difference equations with respect to a fading perturbation. *Int. J. Difference Equ.*, **4**(2), 165–184.

Appleby, J. A. D., Mao, X., & Rodkina, A. 2006. On stochastic stabilization of difference equations. *Discrete Contin. Dyn. Syst.*, **15**(3), 843–857.

Appleby, J. A. D., Berkolaiko, G., & Rodkina, A. 2009. Non-exponential stability and decay rates in nonlinear stochastic difference equations with unbounded noise. *Stochastics*, **81**(2), 99–127.

Appleby, J. A. D., Guzowska, M., Kelly, C., & Rodkina, A. 2010. Preserving positivity in solutions of discretised stochastic differential equations. *Appl. Math. Comput.*, **217**(2), 763–774.

Arnold, L. 1974. *Stochastic differential equations: theory and applications.* New York: Wiley-Interscience [John Wiley & Sons]. Translated from the German.

Berkaoui, A., Bossy, M., & Diop, A. 2008. Euler scheme for SDEs with non-

Lipschitz diffusion coefficient: strong convergence. *ESAIM Probab. Stat.*, **12**, 1–11 (electronic).

Berkolaiko, G., Buckwar, E., Kelly, C., & Rodkina, A. 2012. Almost sure asymptotic stability analysis of the θ-Maruyama method applied to a test system with stabilising and destabilising stochastic perturbations. *LMS J. Comput. Math.*, **15**, 71–83.

Buckwar, E., & Kelly, C. 2010. Towards a systematic linear stability analysis of numerical methods for systems of stochastic differential equations. *SIAM J. Numer. Anal.*, **48**(1), 298–321.

Buckwar, E., & Sickenberger, T. 2012. A structural analysis of asymptotic mean-square stability for multi-dimensional linear stochastic differential systems. *Appl. Numer. Math.*, **62**(7), 842–859.

Burrage, K., & Tian, T. 2000. A note on the stability properties of the Euler methods for solving stochastic differential equations. *New Zealand J. Math.*, **29**(2), 115–127. Dedicated to John Butcher.

Burrage, K., & Tian, T. 2002. Predictor-corrector methods of Runge-Kutta type for stochastic differential equations. *SIAM J. Numer. Anal.*, **40**(4), 1516–1537 (electronic).

Burrage, P. M., & Burrage, K. 2002. A variable stepsize implementation for stochastic differential equations. *SIAM J. Sci. Comput.*, **24**(3), 848–864.

Burrage, P. M., Herdiana, R., & Burrage, K. 2004. Adaptive stepsize based on control theory for stochastic differential equations. *J. Comput. Appl. Math.*, **170**(2), 317–336.

Dangerfield, C. E., Kay, D., & Burrage, K. 2012. Modeling ion channel dynamics through reflected stochastic differential equations. *Physical Review E*, **85**(5), 051907.

Dávila, J., Bonder, J. F., Rossi, J. D., Groisman, P., & Sued, M. 2005. Numerical analysis of stochastic differential equations with explosions. *Stoch. Anal. Appl.*, **23**(4), 809–825.

De la Cruz Cancino, H., Biscay, R. J., Jimenez, J. C., Carbonell, F., & Ozaki, T. 2010. High order local linearization methods: an approach for constructing A-stable explicit schemes for stochastic differential equations with additive noise. *BIT*, **50**(3), 509–539.

Deelstra, G., & Delbaen, F. 1998. Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied stochastic models and data analysis*, **14**(1), 77–84.

Einstein, A. 1905. On the motion of small particles suspended in a stationary liquid, as required by the molecular kinetic theory of heat. *Annalen der Physik*, **17**(6), 132–148.

Frank, T. D., & Beek, P. J. 2001. Stationary solutions of linear stochastic delay differential equations: Applications to biological systems. *Physical Review E*, **64**(2), 021917.

Gaines, J. G., & Lyons, T. J. 1997. Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.*, **57**(5), 1455–1484.

Gardiner, C. W. 2004. *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Third edn. Springer Series in Synergetics, vol. 13. Berlin: Springer-Verlag.

Gihman, I. I., & Skorohod, A. V. 1974. *The theory of stochastic processes. I*. New York: Springer-Verlag. Translated from the Russian by S. Kotz, Die Grundlehren der mathematischen Wissenschaften, Band 210.

Ginzurg, V.L., & Landau, L.D. 1950. Concerning the theory of superconductivity. *J. Exptl. Theoret. Phys.*, **20**, 1064–1082.

Gobet, E., & Menozzi, S. 2004. Exact approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme. *Stochastic Process. Appl.*, **112**(2), 201–223.

Gobet, E., & Menozzi, S. 2010. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Process. Appl.*, **120**(2), 130–162.

Gray, A., Greenhalgh, D., Hu, L., Mao, X., & Pan, J. 2011. A stochastic differential equation SIS epidemic model. *SIAM J. Appl. Math.*, **71**(3), 876–902.

Gyöngy, I. 1998. A note on Euler's approximations. *Potential Anal.*, **8**(3), 205–216.

Higham, D. J. 2000. Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, **38**(3), 753–769 (electronic).

Higham, D. J. 2008. Modeling and simulating chemical reactions. *SIAM Rev.*, **50**(2), 347–368.

Higham, D. J. 2011. Stochastic ordinary differential equations in applied and computational mathematics. *IMA J. Appl. Math.*, **76**(3), 449–474.

Higham, D. J., & Kloeden, P. E. 2007. Strong convergence rates for backward Euler on a class of nonlinear jump-diffusion problems. *J. Comput. Appl. Math.*, **205**(2), 949–956.

Higham, D. J., Mao, X., & Stuart, A. M. 2002. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, **40**(3), 1041–1063 (electronic).

Higham, D. J., Mao, X., & Yuan, C. 2007. Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations. *SIAM J. Numer. Anal.*, **45**(2), 592–609 (electronic).

Hu, Y. 1996. Semi-implicit Euler-Maruyama scheme for stiff stochastic equations. *Pages 183–202 of: Stochastic analysis and related topics, V (Silivri, 1994)*. Progr. Probab., vol. 38. Boston, MA: Birkhäuser Boston.

Huang, C. 2012. Exponential mean square stability of numerical methods for systems of stochastic differential equations. *J. Comput. Appl. Math.*, **236**(16), 4016–4026.

Hutzenthaler, M., & Jentzen, A. 2012. Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. *arXiv preprint arXiv:1203.5809*.

Hutzenthaler, M., Jentzen, A., & Kloeden, P. E. 2011. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **467**(2130), 1563–1576.

Hutzenthaler, M., Jentzen, A., & Kloeden, P. E. 2012. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.*, **22**(4), 1611–1641.

Ikeda, N., & Watanabe, S. 1981. *Stochastic differential equations and diffusion processes*. North-Holland Mathematical Library, vol. 24. Amsterdam: North-Holland Publishing Co.

Itô, K. 1944. Stochastic integral. *Proc. Imp. Acad. Tokyo*, **20**, 519–524.

Itô, K. 1946. On a stochastic integral equation. *Proc. Japan Acad.*, **22**(nos. 1-4), 32–35.

Jentzen, A., Kloeden, P. E., & Neuenkirch, A. 2009. Pathwise approximation of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients. *Numer. Math.*, **112**(1), 41–64.

Khasminskii, R. 2012. *Stochastic stability of differential equations*. second edn. Stochastic Modelling and Applied Probability, vol. 66. Heidelberg: Springer. With contributions by G. N. Milstein and M. B. Nevelson.

Klauder, J. R., & Petersen, W. P. 1985. Numerical integration of multiplicative-noise stochastic differential equations. *SIAM J. Numer. Anal.*, **22**(6), 1153–1166.

Kloeden, P. E., & Platen, E. 1992. *Numerical solution of stochastic differential equations*. Applications of Mathematics (New York), vol. 23. Berlin: Springer-Verlag.

Komori, Y. 2008. Weak first- or second-order implicit Runge-Kutta methods for stochastic differential equations with a scalar Wiener process. *J. Comput. Appl. Math.*, **217**(1), 166–179.

Lamba, H., & Seaman, T. 2006. Mean-square stability properties of an adaptive time-stepping SDE solver. *J. Comput. Appl. Math.*, **194**(2), 245–254.

Lamba, H., Mattingly, J. C., & Stuart, A. M. 2007. An adaptive Euler-Maruyama scheme for SDEs: convergence and stability. *IMA J. Numer. Anal.*, **27**(3), 479–506.

Li, X., Gray, A., Jiang, D., & Mao, X. 2011. Sufficient and necessary conditions of stochastic permanence and extinction for stochastic logistic populations under regime switching. *J. Math. Anal. Appl.*, **376**(1), 11–28.

Liu, W., & Mao, X. 2013a. Asymptotic moment boundedness of the numerical solutions of stochastic differential equations. *J. Comput. Appl. Math.*, **251**, 22–32.

Liu, W., & Mao, X. 2013b. Strong convergence of the stopped Euler-Maruyama method for nonlinear stochastic differential equations. *Appl. Math. Comput.*, **223**, 389–400.

Luo, Q., Mao, X., & Shen, Y. 2011. Generalised theory on asymptotic stability and boundedness of stochastic functional differential equations. *Automatica J. IFAC*, **47**(9), 2075–2081.

Mao, X. 1991. *Stability of stochastic differential equations with respect to semimartingales*. Pitman Research Notes in Mathematics Series, vol. 251. Harlow: Longman Scientific & Technical.

Mao, X. 2008. *Stochastic differential equations and applications*. Second edn. Horwood Publishing Limited, Chichester.

Mao, X. 2011. Numerical solutions of stochastic differential delay equations under the generalized Khasminskii-type conditions. *Appl. Math. Comput.*, **217**(12), 5512–5524.

Mao, X., & Szpruch, L. 2013a. Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients. *J. Comput. Appl. Math.*, **238**, 14–28.

Mao, X., & Szpruch, L. 2013b. Strong convergence rates for backward Euler-Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. *Stochastics*, **85**(1), 144–171.

Mao, X., & Yuan, C. 2006. *Stochastic differential equations with Markovian switching*. London: Imperial College Press.

Mao, X., Yuan, C., & Yin, G. 2005. Numerical method for stationary distribution of stochastic differential equations with Markovian switching. *J. Comput. Appl. Math.*, **174**(1), 1–27.

Mao, X., Shen, Y., & Gray, A. 2011. Almost sure exponential stability of backward Euler-Maruyama discretizations for hybrid stochastic differential equations. *J. Comput. Appl. Math.*, **235**(5), 1213–1226.

Marion, G., Mao, X., & Renshaw, E. 2002. Convergence of the Euler scheme for a class of stochastic differential equation. *Int. Math. J.*, **1**(1), 9–22.

Massey J., Frank J. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, **46**(253), 68–78.

Mauthner, S. 1998. Step size control in the numerical solution of stochastic differential equations. *J. Comput. Appl. Math.*, **100**(1), 93–109.

Milstein, G. N. 1995. *Numerical integration of stochastic differential equations.* Mathematics and its Applications, vol. 313. Dordrecht: Kluwer Academic Publishers Group. Translated and revised from the 1988 Russian original.

Milstein, G. N., & Tretyakov, M. V. 2004. *Stochastic numerics for mathematical physics.* Scientific Computation. Berlin: Springer-Verlag.

Milstein, G. N., & Tretyakov, M. V. 2005. Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients. *SIAM J. Numer. Anal.*, **43**(3), 1139–1154 (electronic).

Milstein, G. N., Platen, E., & Schurz, H. 1998. Balanced implicit methods for stiff stochastic systems. *SIAM J. Numer. Anal.*, **35**(3), 1010–1019 (electronic).

Mitsui, T., & Saito, Y. 2007. MS-stability analysis for numerical solutions of stochastic differential equations—beyond single-step single dim. *Pages 181–194 of: Some topics in industrial and applied mathematics.* Ser. Contemp. Appl. Math. CAM, vol. 8. Beijing: Higher Ed. Press.

Müller-Gronbach, T. 2002. The optimal uniform approximation of systems of stochastic differential equations. *Ann. Appl. Probab.*, **12**(2), 664–690.

Oksendal, B. 2003. *Stochastic differential equations.* Sixth edn. Universitext. Berlin: Springer-Verlag. An introduction with applications.

Pang, S., Deng, F., & Mao, X. 2008. Almost sure and moment exponential stability of Euler-Maruyama discretizations for hybrid stochastic differential equations. *J. Comput. Appl. Math.*, **213**(1), 127–141.

Pettersson, R. 2000. Projection scheme for stochastic differential equations with convex constraints. *Stochastic Process. Appl.*, **88**(1), 125–134.

Platen, E., & Bruti-Liberati, N. 2010. *Numerical solution of stochastic differential equations with jumps in finance*. Stochastic Modelling and Applied Probability, vol. 64. Berlin: Springer-Verlag.

Revuz, D., & Yor, M. 1999. *Continuous martingales and Brownian motion*. Third edn. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293. Berlin: Springer-Verlag.

Rodkina, A., & Schurz, H. 2005. Almost sure asymptotic stability of drift-implicit θ-methods for bilinear ordinary stochastic differential equations in R^1. *J. Comput. Appl. Math.*, **180**(1), 13–31.

Rodkina, A., Schurz, H., & Shaikhet, L. 2008. Almost sure stability of some stochastic dynamical systems with memory. *Discrete Contin. Dyn. Syst.*, **21**(2), 571–593.

Römisch, W., & Winkler, R. 2006. Stepsize control for mean-square numerical methods for stochastic differential equations with small noise. *SIAM J. Sci. Comput.*, **28**(2), 604–625 (electronic).

Saito, Y., & Mitsui, T. 1993. T-stability of numerical scheme for stochastic differential equations. *Pages 333–344 of: Contributions in numerical mathematics*. World Sci. Ser. Appl. Anal., vol. 2. World Sci. Publ., River Edge, NJ.

Saito, Y., & Mitsui, T. 1996. Stability analysis of numerical schemes for stochastic differential equations. *SIAM J. Numer. Anal.*, **33**(6), 2254–2267.

Schurz, H. 1997. *Stability, stationarity, and boundedness of some implicit numerical methods for stochastic differential equations and applications.* Logos Verlag Berlin, Berlin.

Schurz, H. 2005. Stability of numerical methods for ordinary stochastic differential equations along Lyapunov-type and other functions with variable step sizes. *Electron. Trans. Numer. Anal.*, **20**, 27–49 (electronic).

Schurz, H. 2007. Modeling, analysis and discretization of stochastic logistic equations. *Int. J. Numer. Anal. Model.*, **4**(2), 178–197.

Schuss, Z. 1980. Singular perturbation methods in stochastic differential equations of mathematical physics. *Siam Review*, **22**(2), 119–155.

Shen, Y., Luo, Q., & Mao, X. 2006. The improved LaSalle-type theorems for stochastic functional differential equations. *J. Math. Anal. Appl.*, **318**(1), 134–154.

Shiryaev, A. N. 1996. *Probability.* Second edn. Graduate Texts in Mathematics, vol. 95. New York: Springer-Verlag. Translated from the first (1980) Russian edition by R. P. Boas.

Shreve, S. E. 2004. *Stochastic calculus for finance. II.* Springer Finance. New York: Springer-Verlag. Continuous-time models.

Sobczyk, K. 2001. *Stochastic differential equations: with applications to physics and engineering.* Vol. 40. Springer.

Soong, T. T. 1973. *Random differential equations in science and engineering.* New York: Academic Press [Harcourt Brace Jovanovich Publishers]. Mathematics in Science and Engineering, Vol. 103.

Stratonovich, R. L. 1966. A new representation for stochastic integrals and equations. *SIAM J. Control*, **4**, 362–371.

Szpruch, L., Mao, X., Higham, D. J., & Pan, J. 2011. Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. *BIT*, **51**(2), 405–425.

Uhlenbeck, G. E., & Ornstein, L. S. 1930. On the theory of the Brownian motion. *Physical review*, **36**(5), 823.

Valinejad, A., & Hosseini, S. M. 2010. A variable step-size control algorithm for the weak approximation of stochastic differential equations. *Numer. Algorithms*, **55**(4), 429–446.

van Kampen, N. G. 1981. *Stochastic processes in physics and chemistry.* Lecture Notes in Mathematics, vol. 888. Amsterdam: North-Holland Publishing Co.

Wang, X., & Gan, S. 2012. The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients. *Journal of Difference Equations and Applications*, 1–25.

Wiener, N. 1923. Differential space. *J. Math. and Phys.*, **2**, 131–174.

Wu, F., Mao, X., & Chen, K. 2008. A highly sensitive mean-reverting process in finance and the Euler-Maruyama approximations. *J. Math. Anal. Appl.*, **348**(1), 540–554.

Wu, F., Mao, X., & Szpruch, L. 2010. Almost sure exponential stability of numerical solutions for stochastic delay differential equations. *Numer. Math.*, **115**(4), 681–697.

Wu, F., Mao, X., & Kloeden, P. E. 2011. Almost sure exponential stability of the Euler–Maruyama approximations for stochastic functional differential equations. *Random Operators and Stochastic Equations*, **19**(2), 165–186.

Xing, Z., & Peng, J. 2012. Boundedness, persistence and extinction of a stochastic non-autonomous logistic system with time delays. *Appl. Math. Model.*, **36**(8), 3379–3386.

Yu, Z. 2011. Almost surely asymptotic stability of exact and numerical solutions for neutral stochastic pantograph equations. *Abstr. Appl. Anal.*, Art. ID 143079, 14.

Yuan, C., & Mao, X. 2003. Asymptotic stability in distribution of stochastic differential equations with Markovian switching. *Stochastic Process. Appl.*, **103**(2), 277–291.

Yuan, C., & Mao, X. 2004. Stability in distribution of numerical solutions for stochastic differential equations. *Stochastic Anal. Appl.*, **22**(5), 1133–1150.

Yuan, C., & Mao, X. 2005. Stationary distributions of Euler-Maruyama-type stochastic difference equations with Markovian switching and their convergence. *J. Difference Equ. Appl.*, **11**(1), 29–48.