# Stochastic Modelling and Monte Carlo Simulations in Finance

Emmanuel Coffie

Department of Mathematics and Statistics
University of Strathclyde
Glasgow, UK

September 2021

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

# Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.



...........................,
Emmanuel Coffie,
September 2021.

Eyi ne N⊃nye, Akorsiwa, si megali o.

# Acknowledgements

I am deeply grateful to my supervisor, Prof. Mao Xuerong, for the support and guidance he gave me to complete this thesis on time. I wish to extend my deepest gratitude to my family and friends for their unconditional support and encouragement throughout this journey. I am very thankful to the University of Strathclyde for the financial support. Finally, I wish to thank all the staff and colleagues of the Department of Mathematics and Statistics at the University of Strathclyde for the friendly and simulating environment they provided me to complete this task.

Many thanks to you all.

# Abstract

Stochastic modelling of interest rates is very important for calibrating and evaluating expected payoffs of interest-rate products. Many well-known univariate linear drift stochastic models have been proposed to explain interest rate dynamics. However, by testing parametric models by comparing their implied parametric density to the same density estimated non-parametrically, Ait-Sahalia revealed all the existing univariate linear drift stochastic models could not explain well the dynamics of Euro-dollar interest rates. As a result, he proposed a new class of highly non-linear stochastic interest rate models. The original Ait-Sahalia interest rate model has been found to have considerable use for modelling time series evolution of interest rates. However, this model does not possess certain specifications to provide adequate descriptions of interest rates against unexpected empirical phenomena such as volatility 'skews' and 'smiles', jump behaviour, market regulatory lapses, economic crisis, financial clashes, political instability, among others collectively. In this thesis, we propose a modified version of this model by incorporating additional features to help collectively describe these empirical phenomena adequately. However, the proposed model does not have explicit solution. Hence, we split it into three stochastic interest rate models and construct a new implementable truncated EM scheme to approximate them numerically. Further, we study finite time strong convergence of the truncated EM solutions to the exact solutions of the three models under the local Lipschitz condition plus the Khasminskii-type condition. Moreover, we perform numerical simulations to validate the strong convergence results and justify these results within Monte Carlo frameworks to evaluate expected payoffs of some financial products.

# Contents

## ◇ Table of contents

# ♢ List of figures

## ◇ List of tables

# General notations

positive : $> 0$.

nonpositive : $\leq 0$.

negative : $< 0$.

nonnegative : $\geq 0$.

a.s. : almost surely, or with probaility 1.

$\emptyset$ : empty set

$1_A$ : the indicator function of a set $A$ i.e.

$1_A(x) = 1$ if $x \in A$ or otherwise 0.

$\sigma(C)$ : the $\sigma$-algebra generated by $C$.

$a \wedge b$ : $\min\{a, b\}$.

$a \vee b$ : $\max\{a, b\}$.

$f : A \to B$ : the mapping $f$ from $A$ to $B$.

$\mathbb{R} = \mathbb{R}^1$ : the real line.

$\mathbb{R}_+$ : the set of all nonnegative real numbers, i.e. $\mathbb{R}_+ \in [0, \infty)$.

$\mathbb{R}^d$ : the d-dimensional Euclidean space.

$\mathbb{R}^d$ : $= \{x \in \mathbb{R}^d : x_i > 0, 1 \leq i \leq d\}$, i.e. the positive cone.

$\mathcal{B}$ : the Borel-$\sigma$-algebra on $\mathbb{R}^d$.

$|x|$ : the Euclidean norm of a vector $x$.

$C(D; \mathbb{R}^d)$ : the family of continuous $\mathbb{R}^d$-valued functions defined on $D$.

$C^m(D;\mathbb{R})$ : the family of continuous $m-$times differentiable $\mathbb{R}^d$-valued functions defined on $D$.

$C_0^m(D;\mathbb{R}^d)$ : the family of functions in $C^m(D;\mathbb{R})$ with compact support in $D$.

$C^{2,1}(D\times\mathbb{R}_+;\mathbb{R})$ : the family of all real-valued functions $H(x,t)$ defined on $D\times\mathbb{R}_+$ which are continuously twice differentiable in $x\in D$ and once differentiable in $t\in\mathbb{R}_+$.

$$H_x := (H_{x_1},\cdots,H_{x_d}) = \Big(\frac{\partial H}{\partial x_1},\cdots,\frac{\partial H}{\partial x_d}\Big).$$

$$H_{xx} := (H_{x_ix_j})_{d\times d} = \Big(\frac{\partial^2 H}{\partial x_i\partial x_j}\Big)_{d\times d}.$$

$||\xi||_{L^p}$ : $(\mathbb{E}|\xi|^p)^{1/p}$.

$L^p(\Omega;\mathbb{R}^d)$ : the family of $\mathbb{R}^d$-valued random variables $\xi$ with $\mathbb{E}|\xi|^p<\infty$.

$L^p_{\mathcal{F}_t}(\Omega;\mathbb{R}^d)$ : the family of $\mathbb{R}^d$-valued $\mathcal{F}_t$-measurable random variables $\xi$ with $\mathbb{E}|\xi|^p<\infty$.

$C([-\tau,0];\mathbb{R}^d)$ : the space of all continuous $\mathbb{R}^d$-valued functions $\varphi$ defined on $[-\tau,0]$ with a norm $||\varphi|| = \sup_{-\tau\le\theta\le 0}|\varphi(\theta)|$.

$L^p_{\mathcal{F}}([-\tau,0];\mathbb{R}^d)$ : the family of all $C([-\tau,0];\mathbb{R}^d)$-valued random variables $\phi$ such that $\mathbb{E}||\phi||^p<\infty$.

$L^p_{\mathcal{F}_t}([-\tau,0];\mathbb{R}^d)$ : the family of all $\mathcal{F}_t$-measurable $C([-\tau,0];\mathbb{R}^d)$-valued random variables $\phi$ such that $\mathbb{E}||\phi||^p<\infty$.

$C^b_{\mathcal{F}_t}([-\tau,0];\mathbb{R}^d)$ : the family of all $\mathcal{F}_t$-measurable bounded $C([-\tau,0];\mathbb{R}^d)$-valued random variables.

$L^p([a,b];\mathbb{R}^d)$ : the family of Borel measurable functions $h:[a,b]\to\mathbb{R}^d$

such that $\int_a^b |h(t)|^p dt < \infty$.

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

$\mathcal{M}^p([a,b];\mathbb{R}^d)$ : the family of processes $\{f(t)\}_{a\le t\le b} \in \mathcal{L}^p([a,b];\mathbb{R}^d)$ such that $\mathbb{E}\int_a^b |f(t)|^p dt < \infty$.

$\mathcal{L}^p(\mathbb{R}_+;\mathbb{R}^d)$ : the family of processes $\{f(t)\}_{t\ge 0}$ such that for every $T > 0$, $\{f(t)\}_{0\le t\le T} \in \mathcal{L}^p([0,T];\mathbb{R}^d)$.

$\mathcal{M}^p(\mathbb{R}_+;\mathbb{R}^d)$ : the family of processes $\{f(t)\}_{t\ge 0}$ such that for every $T > 0$, $\{f(t)\}_{0\le t\le T} \in \mathcal{M}^p([0,T];\mathbb{R}^d)$.

Other notations will be explained where they appear.

# Chapter 1

# Introduction

The probabilistic nature of interest rates is important since it affects every nature of interest-rate products. In several pricing applications, interest rate is regarded as a deterministic function of time. This is usually motivated under the assumption that the variability of interest rates contributes to the price of financial products such as equity or FX options by a smaller order of magnitude based on the underlying's movements. However, when we deal with interest-rate products, the main variability of importance is that of the interest rates themselves. This gave rise to a new approach for modelling random evolution of interest rates through time with stochastic models (see, e.g., [1, 2] and references cited therein).

Several well-known univariate linear drift stochastic models have been proposed over the years to model dynamics of stochastic interest rates. These models, for example, include Black and Scholes (1973) in [3], Merton (1973) in [4], Vasicek (1977) in [5], Dothan (1978) in [6], Brennan and Schwartz (1980) in [7] and, Cox, Ingersoll and Ross (CIR) (1980 and 1985) in [8] and [9] respectively. These models constitute the popular mean-reverting process with parameter $\theta \in [1/2, 1]$. However, it has been espoused in some empirical studies that the most successful continuous-time models for explaining the dynamics of interest rate are those that allow the volatility of interest changes to be highly sensitive to the level of the rate. By applying $\chi^2$ tests to US treasury bill data, those models with $\theta < 1$ were rejected in favour of those with $\theta \geq 1$. For instance, using the generalised moment method, Chan et al. (1992) in [10] revealed that $\theta = 1.449$. With the same data, Nowman (1997) in [11] used the Gaussian estimation method to also estimate $\theta = 1.361$. Hence it becomes more evident empirically that $\theta > 1$. Is it in this spirit that Lewis (2000) in [12] generalised all these models as a non-linear mean-reverting-theta process of the form

$$dx(t) = \alpha(\mu - x(t))dt + \sigma x(t)^{\theta} dB(t) \tag{1.1}$$

for any $t > 0$ with initial data $x(0) = x_0$, where $\alpha, \mu$ and $\sigma$ are constants, $\theta > 1$ and $B(t)$ is a scalar Brownian motion. This SDE model is widely used to explain time-series evolution of stochastic interest rate, asset price, volatility and other financial quantities. There have been several extensive literature concerning with SDE (1.1) with parametric restrictions. For example, Mao [17] studied analytical properties and strong convergence theory of the numerical solutions of SDE (1.1) for $\theta \in [1/2, 1]$. Higham and Mao [15] examined the strong convergence of Monte Carlo simulations of SDE (1.1) for $\theta = 1/2$. Wu et al. [16] established analytical properties of SDE (1.1) and convergence in probability of the EM approximate solutions for $\theta > 1$.

Ait-Sahalia discovered through empirical studies that all the existing univariate linear drift models could not explain well the dynamics of Euro-dollar interest rates. He then proposed in [18] a new class of stochastic interest rate model for capturing time-series evolution of term structure of interest rates. This model is governed by a highly non-linear SDE of the form

$$dx(t) = (\alpha_{-1}x(t)^{-1} - \alpha_0 + \alpha_1 x(t) - \alpha_2 x(t)^2)dt + \sigma x(t)^{\theta} dB(t), \tag{1.2}$$

on $t \geq 0$ with initial data $x(0) = x_0$, where $\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2 > 0$ and $B(t)$ is a scalar Brownian motion. Similarly, based on the empirical evidence, he retained $\theta > 1$. In his seminal paper, he used Feller test to show conditions under which almost surely the solution of SDE (1.2) will not explode to infinity in finite time. The SDE (1.2) has been studied extensively by many authors. For example, Cheng [19] studied analytical properties including positivity of solution and finite moments of SDE (1.2) and established convergence of EM approximate solutions to the exact solution in probability. Szpruch et al. in [20] generalised SDE (1.2) to

$$dx(t) = (\alpha_{-1}x(t)^{-1} - \alpha_0 + \alpha_1 x(t) - \alpha_2 x(t)^{\rho})dt + \sigma x(t)^{\theta} dB(t), \tag{1.3}$$

for any $t > 0$, $\rho > 1$, with initial data $x(0) = x_0$, and established strong convergence of the implicit EM method as well as preservation of positive approximate solutions of this method when a monotone condition is fulfilled. Dung [21] in 2016 derived explicit estimates for tail probabilities of the solutions to the generalised form of this model. Deng et al. [13] in 2019 studied analytical properties of the generalised form of this model with Poisson-jump and revealed weak convergence of the explicit EM method.

Despite of the wide applications of the aforementioned linear drift stochastic models, they may not be well-specified adequately to fully explain financial variables against certain types of phenomena which have been observed empirically from most financial markets. For example, phenomena such as volatility 'skews' and 'smiles', economic crisis, financial clashes and tail distribution or jump behaviour which have been observed empirically from various sources of financial data may not be adequately explained by these stochastic models (e.g., see [26, 35, 46, 51]).

Several research works have been devoted in recent times to adequately explain dynamical behaviours of financial variables against unexpected occurrences of these empirical phenomena. For instance, the shortcoming of the continuous-time model of Black-Scholes [3] in describing convex phenomena of implied volatility exhibited by most historical financial data led to the underlying assumption of constant volatility to be questioned. In particular, this assumption fails to describe volatility 'skews' and 'smiles' which are typically important for evaluating complex derivative instruments. However, many empirical results found that volatility can be regarded as an endogenous factor and any good financial model should possess important characteristic of reproducing volatility 'smiles' and 'skews' as evidenced in option markets (e.g., see [23,26]). There are several extensive literature where stochastic models with inherent features of past dependency have been used to describe volatility 'smiles' and 'skews' adequately. For example, Kind et al. justified in [25] that the instantaneous volatility is modelled in terms of the sample variance of the log-prices over a past interval of fixed length. Mao and Sabanis [26] also extended the geometric Brownian motion (GBM) to a delay geometric Brownian motion (DGBM) described by a stochastic delay differential equation (SDDE), where the volatility is modelled as a function of delay in asset price and justified it as a rich alternative for modelling financial quantities in a complete market setting.

Additionally, it has also been well known that asset prices admit jumps in response to lack of information or unexpected catastrophic news. This phenomenon typically generates price vibrations with larger quantiles than normal (see [46]). Apparently, this violates the efficient market hypothesis that all available information are reflected in current asset prices. There are several existing rich literature where the authors employed jump-diffusions models to describe jump behaviour of asset prices arising from lack of information or unexpected catastrophic news (e.g., see [33–36]).

Furthermore, hybrid models driven by finite-state Markovian chains have also been considerably used to model uncertainty in modern financial and economic systems. These models have a major characteristic of randomly switching between finite number of regimes in anticipation to unexpected structural changes of unobservable underlying economic or financial settings and mechanisms. The reader is referred, for instance, to [48], [49] and [51] for relevant coverage of applications of hybrid stochastic models in finance.

Interestingly, while the SDE (1.3) enjoys significant patronage from researchers, academic experts and, market practitioners and participants, it may also not possess inherent features to provide full descriptions of dynamical behaviours of interest rates against unexpected joint effects of extreme volatility, jumps, financial clashes, economic crisis, regulatory lapses, political instability, among others. To help describe joint effects of these empirical phenomena adequately, it may be appropriate to specify SDE (1.3) as a hybrid SDDE with Poisson-driven jump governed by

$$\begin{cases} dx(t) = (\alpha_{-1}(r(t))x(t^-)^{-1} - \alpha_0(r(t)) + \alpha_1(r(t))x(t^-) - \alpha_2(r(t))x(t^-)^{\rho})dt \\ \qquad\qquad +\varphi(x((t-\tau)^-), r(t))x(t^-)^{\theta}dB(t) + \alpha_3(r(t))x(t^-)dN(t), \quad t > 0, \\ x(t) = \xi(t), r(0) = r_0, \quad t \in [-\tau, 0], \end{cases} \tag{1.4}$$

where $\rho, \theta > 1$, $r(\cdot)$ is a Markov chain with finite space $\mathcal{S} = \{1, 2, \cdots, N\}$, $\alpha_{-1}$, $\alpha_0$, $\alpha_1$, $\alpha_2$ and $\alpha_3$ are functions of $r(\cdot)$, the volatility function $\varphi(\cdot, \cdot)$ depends on $r(\cdot)$ and $x(t-\tau)$, $\tau > 0$ and $x(t-\tau)$ denotes delay in $x(t)$. Moreover, $x(t^-) = \lim_{s \to t^-} x(s)$, $N(t)$ is a scalar Poisson process independent of a scalar Brownian motion $B(t)$, with the compensated Poisson process given by $\widetilde{N}(t) = N(t) - \lambda t$, where $\lambda$ is the jump intensity.

The SDDE (1.4) integrates three main unique specifications under a unified framework. For instance, the delayed volatility function may be useful in capturing 'smiles' and 'skews' of market implied volatility. On the other hand, the Poisson-driven jump may account for responses of interest rates to discontinuous random effects generated in connection with unexpected catastrophic news or lack of information. The Markovian switching term may address effects of unpredictable market shocks which may arise from abrupt changes such us regulatory lapses, financial clashes, economic crisis, political instability or unobservable states of the underlying market mechanisms or frameworks.

The solution to SDDE (1.4) obviously cannot be found by a closed-form formula. It is also obvious the drift and diffusion terms of SDDE (1.4) are of super-linear growth. This is further complicated by the stiff function $\alpha_{-1}(r(t))x(t)^{-1}$ in the drift which may explode to infinity in finite time around the origin, and the delayed volatility function $\varphi(x((t-\tau)^-), r(t))$ in the diffusion term. As a result, we cannot employ the classical global Lipschitz-based techniques for numerical analysis of SDDE (1.4). Unfortunately, to the best of our knowledge, SDDE (1.4) has never been theoretically and numerically analysed in any existing research literature in the strong sense. Hence, we recognise the need to fill these important gaps by theoretically and numerically investigating feasibility of this model from viewpoint of financial applications such as pricing and evaluating expected payoffs of path-dependent financial products. However, it is widely known that the Monte Carlo approach is one of the most powerful numerical methods used to price and evaluate expected payoffs of many path-dependent financial products quickly and the strong convergence of numerical approximations guarantees convergence in Monte Carlo simulations (e.g., see [15, 61]).

Meanwhile, the well-known existing strong convergence theory for explicit Euler scheme required the drift and diffusion coefficients of SDEs to be globally Lipschitzian (e.g., see [17, 60] for detailed coverage). Higham et al. (2002) in [31] established strong convergence theory for explicit EM scheme for SDEs with non-globally Lipschitz coefficients under the local Lipschitz condition plus the linear growth condition. Hutzenthaler et al. revealed later in [22] that the explicit EM scheme diverges in strong mean-square sense at finite point for SDEs with super-linear coefficients. This unlocked a new chapter for development of suitable class of efficient numerical schemes with cheap computational cost as modified versions of the explicit EM scheme for investigation of convergent approximations of SDEs with super-linear coefficients. For instance, the tamed EM method was developed in 2012 to approximate SDE models with one-sided Lipschitz drift coefficient and the linear growth diffusion coefficient. In 2013, the stopped EM method was also developed to approximate SDE models with non-globally Lipschitz continuous coefficients. Recently, Mao in [27] developed a new explicit numerical method called the truncated EM method, for SDE models with non-globally Lipschitz continuous coefficients. Moreover, the truncated EM method has been further applied in various literature. For example, the authors in [28] and [40] applied the truncated EM method to numerically study non-linear SDE models with constant delay and Poisson-driven jumps respectively. For further existing rich literature in connection with truncated EM methods, the reader is referred, for instance, to [41–44].

In this thesis, we split SDDE (1.4) into three sets of stochastic interest rate models. We examine analytical properties of the three proposed models. Then, we construct several new truncated EM techniques to approximate these three proposed models and explore $L^p(p \geq 2)$ finite time strong convergence of the truncated EM approximate solutions to the exact solutions under the local Lipschitz condition plus the Khasminskii-type condition, where $p$ is a parameter in connection with the Khasminskii-type condition. We perform some numerical examples to validate the strong convergence results established. These results are then applied within Monte Carlo settings to justify calibration and valuation of some financial products such as a debt and a path-dependent derivative instruments. However, we accomplish these tasks in [39], [57] and [58], and extract the results to form Chapters 3-6 of the thesis.

The rest of the thesis is organised as follows: We provide mathematical settings and frameworks in Chapter 2. Chapter 3 is extracted from the paper *The truncated EM numerical method for generalised Ait-Sahalia-type interest rate model with delay*, [39], which I co-authored with my supervisor, Prof. Mao Xuerong. In this chapter, we examine analytical properties of the Ait-Sahalia-type interest rate model with delay and prove that the truncated EM approximation of this model is convergent when the step size is sufficiently small. Chapter 4 is extracted from the paper *Delay stochastic interest rate model with jump and strong convergence in Monte Carlo simulations*, [57], which is solely authored. In this chapter, we investigate analytical properties of the delay Ait-Sahalia-type interest rate model with Poisson-jump and employ the truncated EM techniques to establish the strong convergence results. Chapter 5 is extracted from the paper *Numerical approximation of hybrid Poisson-jump Ait-Sahalia-type interest rate model with delay*, [58], which is also solely authored. In this chapter, we study analytical properties of the hybrid Poisson-jump Ait-Sahalia-type interest rate model with delay and show that the truncated EM approximate solutions converge strongly to the exact solution of this model. The financial applications in Chapter 6 are extracted from the three papers. We conclude the thesis with discussions involving drawbacks, limitations, further applications and future extensions of the theoretical and numerical findings in Chapter 7.

# Chapter 2

# Preliminaries

We need relevant mathematical apparatus to make this thesis self-sufficient. In this chapter, we discuss basic concepts of stochastic analysis. We begin by introducing some concepts from probability theory. We then exploit the fundamental concepts of Brownian motion, stochastic integrals, Itô calculus, stochastic differential equations, Poisson and Markov processes in connection with stochastic differential equations, and proceed to recall some useful well-known inequalities in the last section. We borrowed the contents of this chapter from [17], [60] and [55].

## 2.1 Basic probability concepts

Let us now present the key mathematical concepts of probability theory. If $\Omega$ is a given set, then the $\sigma$-*algebra* $\mathcal{F}$ on $\Omega$ is a family of subsets of $\Omega$ with the following properties:

i. $\emptyset \in \mathcal{F}$, where $\emptyset$ denotes the empty set;

ii. $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$, where $A^C = \Omega - A$ is the complement of $A$ in $\Omega$.;

iii. $\{A_i\}_{i\geq 1} \in \mathcal{F} \Rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*, and the elements of $\mathcal{F}$ is henceforth called $\mathcal{F}$-*measurable sets* instead of events. If $\mathcal{C}$ is a family of subsets of $\Omega$, then there exists a smallest $\sigma$-algebra $\sigma(\mathcal{C})$ on $\Omega$ which contains $\mathcal{C}$. This $\sigma(\mathcal{C})$ is called the *$\sigma$-algebra generated by* $\mathcal{C}$. If $\Omega = \mathbb{R}^d$ and $\mathcal{C}$ is the family of all open sets in $\mathbb{R}^d$, then $\mathcal{B}^d = \sigma(\mathcal{C})$ is called the *Borel $\sigma$-algebra* and the elements of $\mathcal{B}^d$ are called the *Borel sets.*

A real-valued function $X : \Omega \to \mathbb{R}$ is said to be *$\mathcal{F}$-measurable* if

$$\{\omega : X(\omega) \le a\} \in \mathcal{F} \quad \text{for all } a \in \mathbb{R}.$$

The function $X$ is also called a real-valued ($\mathcal{F}$-measurable) *random variable.* An $\mathbb{R}^d$-valued function $X(\omega) = (X_1(\omega), X_2(\omega), \cdots, X_d(\omega))^T$ is said to be *$\mathcal{F}$-measurable* if all the elements $X_i$ are $\mathcal{F}$-measurable. Similarly, a $d \times m$-matrix-valued function $X(\omega) = (X_{ij}(\omega))_{d\times m}$ is said to be *$\mathcal{F}$-measurable* if all the elements $X_{ij}$ are $\mathcal{F}$-measurable.

The *indicator function* $1_A$ of a set $A \subset \Omega$ is defined by

$$1_A(\omega) = \begin{cases} 1 & \text{for } \omega \in A \\ 0 & \text{for } \omega \notin A. \end{cases}$$

The indicator function $1_A$ is $\mathcal{F}$-measurable if and only if $A$ is an $\mathcal{F}$-measurable set, i.e. $A \in \mathcal{F}$. If the measurable space is $(\mathbb{R}^d, \mathcal{B}^d)$, a $\mathcal{B}^d$-measurable function is then called a *Borel measurable function.* More generally, let $(\Omega', \mathcal{F}')$ be another measurable space.

A mapping $X : \Omega \to \Omega'$ is said to be *$(\mathcal{F}, \mathcal{F}')$-measurable* if

$$\{\omega : X(\omega) \in A'\} \in \mathcal{F} \quad \text{for all } A' \in \mathcal{F}.$$

The mapping X is then called an $\Omega'$-valued $(\mathcal{F}, \mathcal{F}')$-measurable (or simply, $\mathcal{F}$-measurable) random variable.

Let $X : \Omega \to \mathbb{R}^d$ be any function. The *$\sigma$-algebra* $\sigma(X)$ *generated by* $X$ is the smallest $\sigma$-algebra on $\Omega$ containing all the sets $\{\omega : X(\omega) \in U\}$, $U \subset \mathbb{R}^d$ open. That is

$$\sigma(X) = \sigma(\{\omega : X(\omega) \in U\} : U \subset \mathbb{R}^d \text{ open}).$$

Clearly, $X$ will then be $\sigma(X)$-measurable and $\sigma(X)$ is the smallest $\sigma$-algebra with this property. If $X$ is $\mathcal{F}$-measurable, then $\sigma(X) \subset \mathcal{F}$, i.e. $X$ generates a sub-$\sigma$-algebra of $\mathcal{F}$. If $\{X_i : i \in I\}$ is a collection of $\mathbb{R}^d$-valued functions, define

$$\sigma(X_i : i \in I) = \sigma\Big(\bigcup_{i\in I} \sigma(X_i)\Big)$$

which is called the *$\sigma$-algebra generated by* $\{X_i : i \in I\}$. It is the smallest $\sigma$-algebra with respect to which every $X_i$ is measurable.

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

i. $\mathbb{P}(\emptyset) = 0$;

ii. $\mathbb{P}(\Omega) = 1$;

iii. For any disjoint sequence $\{A_i\}_{i\geq 1} \subset \mathcal{F}$ (*i.e.* $A_i \cap A_j = \emptyset$ if $i \neq j$)

$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

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

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, we set

$$\bar{\mathcal{F}} = \{A \in \Omega : \exists B, C \in \mathcal{F} \text{ such that } B \subset A \subset C,\ \mathbb{P}(B) = \mathbb{P}(C)\}.$$

Then $\bar{\mathcal{F}}$ is a $\sigma$-algebra and is called the *completion* of $\mathcal{F}$. If $\mathcal{F} = \bar{\mathcal{F}}$, the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be *complete*. In the sequel, we let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a given complete probability space.

A *random variable* $X$ is an $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^d$. Every random variable induces a probability measure $\mu_x$ on the Borel measurable space $(\mathbb{R}^d, \mathcal{B}^d)$, defined by

$$\mu_X(B) = \mathbb{P}\{\omega : X(\omega) \in B\} \quad \text{for } B \in \mathcal{B}^d,$$

and $\mu_X$ is called the *distribution* of $X$.

If $X$ is a real-valued random variable and is *integrable* with respect to the probability measure $\mathbb{P}$, then the number

$$\mathbb{E}X = \int_\Omega X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} x d\mu_X(x)$$

is called the *expectation* of $X$ with respect to $\mathbb{P}$. The number

$$\mathbb{V}(X) = \mathbb{E}(X - \mathbb{E}(X))^2$$

is called the *variance* of X.

More generally, if $f : \mathbb{R}^d \to \mathbb{R}^m$ is Borel measurable and $\int_\Omega |f(X(\omega))| d\mathbb{P}(\omega) < \infty$, then we have

$$\mathbb{E}f(X) = \int_\Omega f(X(\omega)) d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} f(x) d\mu_X(x).$$

The number $\mathbb{E}|X|^p$ for $p > 0$ is called the pth moment of $X$ i.e. $\mathbb{E}|X|^p = \int_\Omega |X(\omega)|^p d\mathbb{P}(\omega)$. For $p \in (0, \infty)$, let $L^p = L^p(\Omega; \mathbb{R}^d)$ be the family of $\mathbb{R}^d$-valued random variables $X$ with $\mathbb{E}|X|^p < \infty$. In $L^1$, we have $|\mathbb{E}X| \leq \mathbb{E}|X|$. Moreover, the following three inequalities hold true.

i **Hölder's inequality:** if $p > 1$, $1/p + 1/q = 1$, $X \in L^p$ and $Y \in L^q$, then

$$|\mathbb{E}(X^T Y)| \leq (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}|Y|^q)^{1/q};$$

ii **Minkovski's inequality:** if $p > 1$ and $X, Y \in L^p$, then

$$(\mathbb{E}|X + Y|^p)^{1/p} \leq (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}|Y|^p)^{1/p};$$

iii **Chebyshev's inequality:** if $c > 0$, $p > 0$ and $X \in L^p$, then

$$\mathbb{P}\{\omega : |X(\omega)| \geq c\} \leq \frac{1}{c^p}\mathbb{E}|X|^p.$$

A simple application of Hölder's inequality implies

$$(\mathbb{E}|X|^r)^{1/r} \leq (\mathbb{E}|X|^p)^{1/p}$$

if $0 < r < p < \infty$, $X \in L^p$.

Let $X$ and $X_k$, $k \geq 1$, be $\mathbb{R}^d$-valued random variables. The following four convergence concepts are very important:

a. If there exists a $\mathbb{P}$-null set $\Omega_0 \in \mathcal{F}$ such that for every $\omega \notin \Omega_0$, the sequence $\{X_k(\omega)\}$ converges to $X(\omega)$ in the usual sense in $\mathbb{R}^d$, then $\{X_k\}$ is said to converge to $X$ *almost surely or with a probability 1*, and we write

$$\lim_{k \to \infty} X_k = X \text{ a.s.}$$

b. If for every $\epsilon > 0$, $\mathbb{P}\{\omega : |X_k(\omega) - X(\omega)| > \epsilon\} \to 0$ as $k \to \infty$, then $\{X_k\}$ is said to converge to $X$ *stochastically or in probability.*

c. If $X_k$ and $X$ belong to $L^p$ and $\mathbb{E}|X_k - X|^p \to 0$, then $\{X_k\}$ is said to converge to $X$ *in pth moment or in* $L^p$.

d. If for every real-valued continuous bounded function $g$ defined on $\mathbb{R}^d$,

$$\lim_{k \to \infty} \mathbb{E}g(X_k) = \mathbb{E}g(X),$$

then $\{X_k\}$ is said to converge to $X$ *in distribution*.

These convergence concepts have the following relationship:

$$\begin{array}{ccc} & & \text{convergence in } L^p \\ & & \Downarrow \\ \text{a.s. convergence} & \Rightarrow & \text{convergence in probability} \\ & & \Downarrow \\ & & \text{convergence in distribution} \end{array}$$

Furthermore, a sequence converges in probability if and only if every subsequence of it contains an almost surely convergent subsequence. A sufficient condition for

$$\lim_{k\to\infty} X_k = X \text{ a.s.}$$

is the condition

$$\sum_{k=1}^{\infty} \mathbb{E}|X_k - X|^p < \infty \qquad \text{for some } p > 0.$$

Let now state two very important integration convergence theorems.

**Theorem 2.1.1.** (***Monotonic convergence theorem***) *If $\{X_k\}$ is an increasing sequence of non-negative random variables , then*

$$\lim_{k\to\infty} \mathbb{E}X_k = \mathbb{E}\Big(\lim_{k\to\infty} X_k\Big).$$

**Theorem 2.1.2.** (***Dominated convergence theorem***) *Let $p \geq 1$, $\{X_k\} \subset L^p(\Omega;\mathbb{R}^d)$ and $Y \in L^p(\Omega;\mathbb{R})$. Assume that $|X_k| \leq Y$ a.s. and $\{X_k\}$ converges to $X$ in probability. Then $X \in L^p(\Omega;\mathbb{R}^d)$, $\{X_k\}$ converges to $X$ in $L^p$, and*

$$\lim_{k\to\infty} \mathbb{E}X_k = \mathbb{E}X.$$

*When $Y$ is bounded, this is referred to as the bounded convergence theorem.*

Let $I$ be an index set. A collection of sets $\{A_i : i \in I\} \subset \mathcal{F}$ is said to be *independent* if

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k})$$

for all possible choices of indices $i_1, i_2, \cdots, i_k \in I$. Two sub-$\sigma$-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ of $\mathcal{F}$ are said to be *independent* if

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) \quad \text{for all } A_1 \in \mathcal{F}_1 \text{ and } A_2 \in \mathcal{F}_2.$$

A collection of sub-$\sigma$-algebras $\{\mathcal{F}_i : i \in I\}$ is said to be *independent* if for every possible choice of indices $i_1, i_2, \cdots, i_k \in I$,

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k})$$

holds for all $A_{i_1} \in \mathcal{F}_{i_1}$, $A_{i2} \in \mathcal{F}_{i_2}$,$\cdots$,$A_{i_k} \in \mathcal{F}_{ik}$. A family of random variables $\{X_i : i \in I\}$ (whose ranges may differ for different values of the index) is said to be *independent* if the $\sigma$-algebras $\sigma(X_i)$, $i \in I$ generated by them are independent. For example, two random variables $X : \Omega \longrightarrow \mathbb{R}^d$ and $Y : \Omega \longrightarrow \mathbb{R}^m$ are independent if and only if

$$\mathbb{P}\{\omega : X(\omega) \in A, Y(\omega) \in B\} = \mathbb{P}\{\omega : X(\omega) \in A\}\mathbb{P}\{\omega : Y(\omega) \in B\}$$

holds for all $A \in \mathcal{B}^d$, $B \in \mathcal{B}^m$. If $X$ and $Y$ are two independent real-valued integrable random variables, then $XY$ is also integrable and

$$\mathbb{E}(XY) = \mathbb{E}X\,\mathbb{E}Y.$$

If $X, Y \in L^2(\Omega; \mathbb{R})$ are uncorrelated then

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$

If the $X$ and $Y$ are independent, they are uncorrelated. If $(X, Y)$ has a normal distribution, then $X$ and $Y$ are independent if and only if they are uncorrelated.

Let $\{A_k\}$ be a sequence of sets in $\mathcal{F}$. Define the *upper limit of the sets* by

$$\lim_{k\to\infty} A_k = \{\omega : \omega \in A_k \text{ for infinitely many } k\} = \bigcap_{i=1}^{\infty}\bigcup_{k=i}^{\infty} A_k.$$

Clearly, it belongs to $\mathcal{F}$. With regard to its probability, we have the following well-known *Borel-Cantelli lemma*.

**Lemma 2.1.3.** (***Borel-Cantelli's lemma***)

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

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

*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* $k \geq k_0(\omega)$.

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

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

*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* $\omega$ *belongs to every* $A_{k_i}$.

Let $A, B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$. The *conditional probability of* $A$ *under condition* $B$ is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Let now present a more general concepts of *conditional expectation.* Let $X \in L^1(\Omega; \mathbb{R})$. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. So, $(\Omega, \mathcal{G})$ is a measurable space. In general, $X$ is not $\mathcal{G}$-measurable. Let seek an integrable $\mathcal{G}$-measurable random variable $Y$ such that

$$\mathbb{E}(I_G Y) = \mathbb{E}(I_G X),$$

that is, for all $G \in \mathcal{G}$

$$\int_G Y(\omega) d\mathbb{P}(\omega) = \int_G X(\omega) d\mathbb{P}(\omega).$$

By the Randon-Nikodym theorem, there exists one such $Y$, almost surely unique. This is called the *conditional expectation of $X$ under the condition $\mathcal{G}$*, and is given by

$$Y = \mathbb{E}(X|\mathcal{G}).$$

If $\mathcal{G}$ is the $\sigma$-algebra generated by a random variable $Y$, we write

$$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X|Y).$$

Now consider a collection of sets $\{A_k\} \in \mathcal{F}$ with

$$\bigcup_k A_k = \Omega, \quad \mathbb{P}(A_k) > 0, \quad A_k \cap A_i = \emptyset \text{ if } k \neq i.$$

Let $\mathcal{G} = \sigma(\{A_k\})$, i.e. $\mathcal{G}$ is generated by $\{A_k\}$. Then $\mathbb{E}(X|\mathcal{G})$ is a *step function* on $\Omega$ given by

$$\mathbb{E}(X|\mathcal{G}) = \sum_k \frac{1_{A_k} \mathbb{E}(1_{A_k} X)}{\mathbb{P}(A_k)}.$$

In other words, if $\omega \in A_k$,

$$\mathbb{E}(X|\mathcal{G})(\omega) = \frac{\mathbb{E}(1_{A_k} X)}{\mathbb{P}(A_k)}.$$

It then follows from the definition that

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$$

and

$$|\mathbb{E}(X|\mathcal{G})| \leq \mathbb{E}(|X||\mathcal{G}) \quad \text{a.s.}$$

## 2.2 Stochastic processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *filtration* is a family $\{\mathcal{F}\}_{t\geq 0}$ of increasing sub-$\sigma$-algebras of $\mathcal{F}$ (i.e. $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$ for all $0 \leq t < s < \infty$). The filtration is said to be *right continuous* if $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \geq 0$. When the probability space is complete, the filtration is said to satisfy the *usual conditions* if it is right continuous and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets.

*From now on, unless otherwise specified, we let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. We also define $\mathcal{F}_\infty = \sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$, i.e. the $\sigma$-algebra generated by $\bigcup_{t\geq 0} \mathcal{F}_t$.*

A family $\{X_t\}_{t\in I}$ of $\mathbb{R}^d$-valued random variables is called a *stochastic process* with *parameter set* or *index set* $I$ and *state space* $\mathbb{R}^d$. The parameter set $I$ is usually the halfline $\mathbb{R}_+ = [0, \infty)$, but it may also be an interval $[a, b]$, the non-negative integers or even subsets of $\mathbb{R}^d$. For each fixed $t \in I$, we have a random variable

$$\Omega \ni \omega \to X_t(\omega) \in \mathbb{R}^d.$$

Moreover, for each fixed $\omega \in \Omega$, we have a function

$$I \ni t \to X_t(\omega) \in \mathbb{R}^d$$

which is called a *sample path* of the process, and we shall write $X_\cdot(\omega)$ for the path. Mostly, it is convenient to write $X(t, \omega)$ instead of $X_t(\omega)$, and the stochastic process may be regarded as a function of two variables $(t, \omega)$ from $I \times \Omega$ to $\mathbb{R}^d$. We often write stochastic process $\{X_t\}_{t\geq 0}$ as $\{X_t\}$, $X_t$ or $X(t)$. In this work, we use the variable *$x(t)$ to denote a stochastic process.*

Let $\{X_t\}_{t\geq 0}$ be an $\mathbb{R}^d$-valued stochastic process. The stochastic process is said to be *continuous* (resp. *right continuous, left continuous*) if for almost all $\omega \in \Omega$, the function $X_t(\omega)$ is continuous (resp. right continuous, left continuous) on $t \geq 0$. It is said to be *càdlàg* (*right continuous and left limit*) if it is right continuous and for almost all $\omega \in \Omega$, the left limit $\lim_{s \uparrow t} X_s(\omega)$ exists and is finite for all $t > 0$. It is said to be *integrable* if for every $t \geq 0$, $X_t$ is an integrable random variable. It is said to be $\{\mathcal{F}_t\}$-*adapted* if for every $t \geq 0$, $X_t$ is $\mathcal{F}_t$-measurable. It is said to be *measurable* if the stochastic process regarded as a function of two random variables $(t, \omega)$ from $\mathbb{R}_+ \times \Omega$ to $\mathbb{R}^d$ is $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$-measurable, where $\mathcal{B}(\mathbb{R}_+)$ is the family of all Borel sub-sets of $\mathbb{R}_+$.

A random variable $\tau : \Omega \to [0, \infty]$ (it may take the value $\infty$) is called $\{\mathcal{F}_t\}$ *stopping time* if $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for any $t \geq 0$.

**Theorem 2.2.1.** *If $\{X_t\}_{t \geq 0}$ is a progressively measurable process and $\tau$ is a stopping time, then $X_\tau I_{\{\tau < \infty\}}$ is $\mathcal{F}_\tau$-measurable. In particular, if $\tau$ is finite, then $X_\tau$ is $\mathcal{F}_\tau$-measurable.*

**Theorem 2.2.2.** *Let $\{X_t\}_{t \geq 0}$ be an $\mathbb{R}^d$-valued càdlàg $\{\mathcal{F}_t\}$-adapted process, and $D$ an open subset of $\mathbb{R}^d$. Define*

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

*where we use the convention* $\inf \emptyset = \infty$. *Then $\tau$ is an $\{\mathcal{F}_t\}$-stopping time, and is called the first exit time from D. Moreover, if $\rho$ is a stopping time, then*

$$\theta = \inf\{t \geq \rho : X_t \notin D\}$$

*is also called $\{\mathcal{F}_t\}$-stopping time, and is called the first exit time from $D$ after $\rho$.*

An $\mathbb{R}^d$-valued $\{\mathcal{F}_t\}$-adapted integrable process $\{M_t\}_{t \geq 0}$ is called a *martingale with respect to* $\{\mathcal{F}_t\}$ (or simply, *martingale*) if

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

Note every martingale has a *càdlàg modification* since we always assume that the filtration $\{\mathcal{F}_t\}$ is right continuous.

If $X = \{X_t\}_{t\geq 0}$ is a progressively measurable process and $\tau$ 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 2.2.3.** *Let $\{M_t\}_{t\geq 0}$ be an $\mathbb{R}^d$-valued martingale with respect to $\{\mathcal{F}_t\}$, and let $\theta$, $\rho$ be two finite stopping times. Then*

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

*In particular, if $\tau$ is a stopping time, then*

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

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

A stochastic process $X = \{X_t\}_{t\geq 0}$ is called *square-integrable* if $\mathbb{E}|X_t|^2 < \infty$ for every $t \geq 0$. If $M = \{M_t\}_{t\geq 0}$ is a real-valued square-integrable continuous martingale, then there exists a unique continuous integrable adapted increasing process denoted by $\{\langle M, M\rangle_t\}$ such that $\{M_t^2 - \langle M, M\rangle_t\}$ is a continuous martingale vanishing at $t = 0$. The process $\{\langle M, M\rangle_t\}$ is called the *quadratic variation* of $M$. In particular, for any finite stopping time $\tau$,

$$\mathbb{E}M_\tau^2 = \mathbb{E}\langle M, M\rangle_\tau.$$

If $N = \{N_t\}_{t\geq 0}$ is another real-valued square-integrable continuous martingale, we define

$$\langle M, N\rangle_t = \frac{1}{2}\Big(\langle M+N, M+N\rangle_t - \langle M, M\rangle_t - \langle N, N\rangle_t\Big),$$

and call $\{\langle M, N\rangle_t\}$ the *joint quadratic variation* of $M$ and $N$. It is useful to know that $\{\langle M, N\rangle_t\}$ is the unique continuous integrable adapted process of finite variation such that $\{M_tN_t - \langle M, N\rangle_t\}$ is a continuous martingale vanishing at $t = 0$. In particular, for any finite stopping time $\tau$,

$$\mathbb{E}M_\tau N_\tau = \mathbb{E}\langle M, N\rangle_\tau.$$

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 $\tau_k \uparrow \infty$ a.s. such that every $\{M_{\tau_k \wedge t} - M_0\}_{t\geq 0}$ is a martingale. Every martingale is a local martingale (by Theorem 2.2.3) but the converse is not true. If $M = \{M_t\}_{t\geq 0}$ and $N = \{N_t\}_{t\geq 0}$ are two real-valued continuous local martingales, their *joint quadratic variation* $\{\langle M, N\rangle\}_{t\geq 0}$ is the unique continuous adapted process of finite variation such that $\{M_t N_t - \langle M, N\rangle_t\}_{t\geq 0}$ is a continuous local martingale vanishing at $t = 0$. When $M = N$, $\{\langle M, M\rangle\}_{t\geq 0}$ is called the *quadratic variation* of $M$.

The following result is the useful strong law of large numbers.

**Theorem 2.2.4.** (***Strong law of large numbers***) *Let $M = \{M_t\}_{t\geq 0}$ be a real-valued continuous local martingale varnishing at $t = 0$. Then*

$$\lim_{t\to\infty} \langle M, M\rangle_t = \infty \quad a.s. \quad \Rightarrow \quad \lim_{t\to\infty} \frac{M_t}{\langle M, M\rangle_t} = 0 \quad a.s.$$

*and also*

$$\limsup_{t\to\infty} \frac{\langle M, M\rangle_t}{t} < \infty \quad a.s. \quad \Rightarrow \quad \lim_{t\to\infty} \frac{M_t}{t} = 0 \quad a.s.$$

*More generally, if $A = \{A_t\}_{t\geq 0}$ is a continuous adapted increasing process such that*

$$\lim_{t\to\infty} A_t = \infty \quad and \quad \int_0^\infty \frac{d\langle M, M\rangle_t}{(1 + A_t)^2} < \infty \quad a.s.$$

*then*

$$\lim_{t\to\infty} \frac{M_t}{A_t} = 0 \quad a.s.$$

A real-valued $\{\mathcal{F}_t\}$-adapted integrable process $\{M_t\}_{t\geq 0}$ is called a *supermartingale* (*with respect to* $\{\mathcal{F}_t\}$) if

$$\mathbb{E}(M_t|\mathcal{F}_s) \leq M_s$$

and a *submartingale* (*with respect to* $\{\mathcal{F}_t\}$) if

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

Obviously, $\{M_t\}$ is submartingale if and only if $\{-M_t\}$ is a supermartingale. For a real-valued martingale $\{M_t\}$, $\{M_t^+ := \max(M_t, 0)\}$ and $\{M_t^- := \max(0, -M_t)\}$ are submartingales. For a supermartingale (resp. submartingale), $\mathbb{E}M_t$ is monotonically decreasing (resp. increasing). Moreover, if $p \geq 1$ and $\{M_t\}$ is an $\mathbb{R}^d$-valued martingale such that $M_t \in L^p(\Omega; \mathbb{R}^d)$, then $\{|M_t|^p\}$ is a non-negative submartingale. Note Doob's stopping Theorem 2.2.3 holds as well for supermartingales and submartingales.

## 2.3 Brownian motion

In 1828, the Scottish botanist Robert Brown observed that pollens suspended in liquid performed an irregular motion. The motion was later explained by a random collisions with the molecules of the liquid. To describe the motion mathematically, it is natural to use the concept of a stochastic process $B_t(\omega)$, interpreted as the position at time $t$ of the pollen grain $\omega$. The Brownian motion is the most fundamental continuous-time stochastic process. It has useful applications in several stochastic systems. The mathematical concepts of Brownian motion form the basis for stochastic analysis. Let us now give the mathematical definition of a Brownian motion.

**Definition 2.3.1.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space with a filtration* $\{\mathcal{F}_t\}_{t\geq 0}$. *A* (*standard*) *one-dimensional Brownian motion is a real-valued continuous* $\{\mathcal{F}_t\}$-*adapted process* $\{B_t\}_{t\geq 0}$ *with the following properties:*

*i.* $B_0 = 0$ *a.s.;*

*ii. for* $0 \leq s < t < \infty$, *the increment* $B_t - B_s$ *is normally distributed with mean zero and variance* $t - s$*;*

*iii. for $0 \le s < t < \infty$, the increment $B_t - B_s$ is independent of $\mathcal{F}_s$.*

*iv. Almost surely, the sample path $t \to B_t(\omega)$ is continuous.*

Let $\{B_t\}_{0 \le t \le T}$ on $[0,T]$ for some $T > 0$. If $\{B_t\}_{t \ge 0}$ is Brownian motion and $0 \le t_0 < t_1 < \cdots < t_k < \infty$, then the increments $B_{t_i} - B_{t_{i-1}}$, $1 \le i \le k$ are independent, and we say that the Brownian motion has *independent increments*. Moreover, the distribution of $B_{t_i} - B_{t_{i-1}}$ depends only on the difference $t_i - t_{i-1}$, and we say that the Brownian motion has *stationary increments*. The filtration $\{\mathcal{F}_t\}$ is a part of the definition of Brownian motion.

The following are important properties of Brownian motion.

i. $\{-B_t\}$ is a Brownian motion with respect to the same filtration $\{\mathcal{F}_t\}$.

ii. Let $c > 0$. Define
$$X_t = \frac{B_{ct}}{\sqrt{c}} \quad \text{for } t \ge 0.$$
The $\{X_t\}$ is a Brownian motion with respect to the filtration $\{\mathcal{F}_{ct}\}$.

iii. $\{B_t\}$ is a continuous square-integrable martingale and its quadratic variation $\langle B, B \rangle_t = t$ for all $t \ge 0$.

iv. The strong law of large numbers states that
$$\lim_{t \to \infty} \frac{B_t}{t} = 0 \quad \text{a.s.}$$

v. For almost every $\omega \in \Omega$, the Brownian sample path $B_\cdot(\omega)$ is nowhere differentiable.

vi. For almost every $\omega \in \Omega$, the Brownian sample path $B_\cdot(\omega)$ is locally Hölder continuous with exponent $\delta$ if $\delta \in (0, 1/2)$. However, for almost every $\omega \in \Omega$, the Brownian sample path $B_\cdot(\omega)$ is nowhere Hölder continuous with exponent $\delta > 1/2$.

## 2.4 Stochastic integrals

We present in this section the mathematical framework of stochastic integral. Let us now define the stochastic integral

$$\int_0^t f(s)dB_s$$

with respect to an $m$-dimensional Brownian motion $\{B_t\}$ for a class of $d\times m$-matrix-valued stochastic processes $\{f(t)\}$. Since for almost all $\omega\in\Omega$, the Brownian sample path $B_\cdot(\omega)$ is of infinite variation and nowhere differentiable, the integral cannot be defined in the ordinary way. This integral was first defined by K. Itô in 1949 and is now known as *Itô stochastic integral*.

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 $B=\{B_t\}_{t\geq 0}$ be a one-dimensional Brownian motion defined on the probability space adapted to the filtration.

**Definition 2.4.1.** *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*

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

*We identify $f$ and $\bar{f}$ in $\mathcal{M}^2([a,b];\mathbb{R})$ if $||f-\bar{f}||^2_{a,b}=0$. In this case we say that $f$ and $\bar{f}$ are equivalent and write $f=\bar{f}$.*

The stochastic processes $f\in\mathcal{M}^2([a,b];\mathbb{R})$ would help define the Itô stochastic integral. The idea is natural: First we define the integral $\int_a^b g(t)dB_t$ for a class of simple processes $g$. Then we show that each $f\in\mathcal{M}^2([a,b];\mathbb{R})$ can be approximated by such simple processes $g's$ and we define the limit of $\int_a^b g(t)dB_t$ as the integral of $\int_a^b f(t)dB_t$. Let us first introduce the concept of simple processes.

**Definition 2.4.2.** *A real-valued stochastic process $g=\{g(t)\}_{a\leq t\leq b}$ is called a simple (or step) process if there exists a partition $a=t_0<t_1<\cdots<t_k=b$ of $[a,b]$, and*

*bounded random variables $\xi_i$, $0 \le i \le k-1$ such that $\xi_i$ is $\mathcal{F}_{t_i}$-measurable and*

$$g(t) = \xi_0 1_{[t_0,t_1]}(t) + \sum_{i=1}^{k-1} \xi_i 1_{(t_i,t_{i+1}]}(t). \tag{2.1}$$

*Denote by $\mathcal{M}_0([a,b];\mathbb{R})$ the family of all such processes.*

Apparently, $\mathcal{M}_0([a,b];\mathbb{R}) \subset \mathcal{M}^2([a,b];\mathbb{R})$. Let us now provide the definition of Itô stochastic integral for such simple processes.

**Definition 2.4.3.** *For a simple process $g$ with the form of (2.1) in $\mathcal{M}_0([a,b];\mathbb{R})$, define*

$$\int_a^b g(t)dB_t = \sum_{i=0}^{k-1} \xi_i (B_{t_{i+1}} - B_{t_i}) \tag{2.2}$$

*and call it the stochastic integral of $g$ with respect to the Brownian motion $\{B_t\}$ or the Itô integral.*

Clearly, the stochastic integral $\int_a^b g(t)dB_t$ is $\mathcal{F}_b$-measurable. By extension of (2.2) into $\mathcal{M}^2([a,b];\mathbb{R})$ yields the following definition.

**Definition 2.4.4.** *Let $f \in \mathcal{M}^2([a,b];\mathbb{R})$. The Itô integral of $f$ with respect to $\{B_t\}$ is defined by*

$$\int_a^b f(t)dB_t = \lim_{n\to\infty} \int_a^b g_n(t)dB_t \quad \text{in } L^2(\Omega;\mathbb{R}),$$

*where $\{g_n\}$ is a sequence of simple processes such that*

$$\lim_{n\to\infty} \mathbb{E} \int_a^b |f(t) - g_n(t)|^2 dt = 0.$$

Let now present the following useful properties of Itô integral.

**Theorem 2.4.5.** *Let $f, g \in \mathcal{M}^2([a,b];\mathbb{R})$, and let $\alpha, \beta$ be two real numbers. Then*

*i.* $\int_a^b f(t)dB_t$ *is $\mathcal{F}_b$-measurable;*

*ii.* $\mathbb{E}\int_a^b f(t)dB_t = 0$;

*iii.* $\mathbb{E}|\int_a^b f(t)dB_t|^2 = \mathbb{E}\int_a^b |f(t)|^2 dt$;

*iv.* $\int_a^b [\alpha f(t) + \beta g(t)]dB_t = \alpha \int_a^b f(t)dB_t + \beta \int_a^b g(t)dB_t$.

The indefinite Itô integral is defined below.

**Definition 2.4.6.** *Let $f \in \mathcal{M}^2([a,b];\mathbb{R})$. Define*

$$I(t) = \int_0^t f(s)dB_s \quad \text{for } 0 \le t \le T,$$

*where, by definition, $I(0) = \int_0^0 f(s)dB_s = 0$. We call $I(t)$ the indefinite Itô integral of $f$.*

Clearly, $\{I(t)\}$ is $\{\mathcal{F}_t\}$-adapted. Let us now present the very important martingale property of the indefinite Itô integral.

**Theorem 2.4.7.** *Let $f \in \mathcal{M}^2([a,b];\mathbb{R})$, then the indefinite Itô integral $\{I(t)\}_{0\le t\le T}$ is a square-integrable martingale with respect to the filtration $\{\mathcal{F}_t\}$. In particular,*

$$\mathbb{E}\Big[\sup_{0\le t\le T}\Big|\int_0^t f(s)dB_s\Big|^2\Big] \le 4\mathbb{E}\int_0^T |f(s)|^2 ds.$$

**Theorem 2.4.8.** *If $f \in \mathcal{M}^2([a,b];\mathbb{R})$, then the indefinite Itô integral $\{I(t)\}_{0\le t\le T}$ has a continuous version.*

**Theorem 2.4.9.** *Let $f \in \mathcal{M}^2([a,b];\mathbb{R})$. Then the indefinite Itô integral $I = \{I(t)\}_{0\le t\le T}$ is a square-integrable continuous martingale and its quadratic variation is given by*

$$\langle I, I\rangle_t = \int_0^t |f(s)|^2 ds, \quad 0 \le t \le T.$$

## 2.5 Itô formula

We use Itô formula to evaluate Itô integral. That is, we use Itô formula to simplify stochastic integrals to Lebesgue integrals for easy evaluation. In this section, we shall first establish the one-dimensional Itô formula and then generalise it to the multi-dimensional case.

Let $B = \{B_t\}_{t\geq 0}$ be a one-dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Let $\mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^d)$ denote the family of all $\mathbb{R}^d$-valued measurable $\{\mathcal{F}_t\}$-adapted processes $f = \{f(t)\}_{t\geq 0}$ such that

$$\int_0^T |f(t)|dt < \infty \quad \text{a.s. for every } T > 0.$$

We require Itô process to define Itô formula. Let us now define the Itô process.

**Definition 2.5.1.** *A $d$-dimensional Itô process is an $\mathbb{R}^d$-valued continuous adapted process $x(t) = (x_1(t), \dots, x_d(t))^T$ on $t \geq 0$ of the form*

$$x(t) = x(0) + \int_0^t f(s)ds \int_0^t g(s)dB(s),$$

*where $f = (f_1, \cdots, f_d)^T \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^d)$ and $g = (g_{ij})_{d\times m} \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{d\times m})$. We shall say that $x(t)$ has stochastic differential $dx(t)$ on $t \geq 0$ given by*

$$dx(t) = f(t)dt + g(t)dB(t).$$

Let $C^{2,1}(\mathbb{R}^d \times \mathbb{R}_+; \mathbb{R})$ denote the family of all real-valued functions $H(x,t)$ defined on $\mathbb{R}^d \times \mathbb{R}_+$ such that they are continuously twice differentiable in $x$ and once in $t$. If $H \in C^{2,1}(\mathbb{R}^d \times \mathbb{R}_+; \mathbb{R})$, we set

$$H_t = \frac{\partial H}{\partial t}, \quad H_x = \Big(\frac{\partial H}{\partial x_1}, \cdots, \frac{\partial H}{\partial x_d}\Big)$$

and

$$H_{xx} = \Big(\frac{\partial^2 H}{\partial x_i \partial x_j}\Big)_{d\times d} = \begin{pmatrix} \frac{\partial^2 H}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 H}{\partial x_1 \partial x_d} \\ \vdots & & \vdots \\ \frac{\partial^2 H}{\partial x_d \partial x_1} & \cdots & \frac{\partial^2 H}{\partial x_d \partial x_d} \end{pmatrix}.$$

**Theorem 2.5.2.** (***The multi-dimensional Itô formula***) *Let $x(t)$ be a $d$-dimensional Itô process on $t \geq 0$ with the stochastic differential*

$$dx(t) = f(t)dt + g(t)dB(t),$$

*where $f \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^d)$ and $g =\in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{d\times m})$. Let $H \in C^{2,1}(\mathbb{R}^d \times \mathbb{R}_+; \mathbb{R})$. Then $H(x(t), t)$ is again an Itô process with the stochastic differential given by*

$$\begin{aligned} dH(x(t), t) = \Big[&H_t(x(t), t) + H_x(x(t), t)f(t) \\ &+ \frac{1}{2}trace(g^T(t)H_{xx}(x(t), t)g(t))\Big]dt + H_x(x(t), t)g(t)dB(t) \quad a.s. \end{aligned}$$

Let us now present formally a multiplication table:

$$\begin{aligned} dtdt = 0, \quad & dB_i dt = 0, \\ dB_i dB_i = dt, \quad & dB_i dB_j = 0 \quad \text{if } i \neq j. \end{aligned}$$

Then, for example,

$$dx_i(t)dx_j(t) = \sum_{k=1}^{m} g_{ik}(t)g_{jk}(t)dt.$$

Moreover, the Itô formula can be written as

$$\begin{aligned} dH(x(t), t) = H_t(x(t), t)dt + H_x(x(t), t)dx(t) \\ \frac{1}{2}dx^T(t)H_{xx}(x(t), t)dx(t). \end{aligned}$$

## 2.6 Stochastic differential equations (SDEs)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. Throughout this section, unless otherwise specified, we let $B(t) = (B_1(t), \cdots, B_m(t))^T$, $t \geq 0$ be an $m$-dimensional Brownian motion defined on the space. Let $0 \leq t_0 < T < \infty$. Let $x_0$ be an $\mathcal{F}_{t_0}$-measurable $\mathbb{R}^d$-valued random variable such that $\mathbb{E}|x_0|^2 < \infty$. Let $f : \mathbb{R}^d \times [t_0, T] \to \mathbb{R}^d$ and $g : \mathbb{R}^d \times [t_0, T] \to \mathbb{R}^{d\times m}$ be both Borel measurable. Consider the $d$-dimensional stochastic differential equation of Itô type

$$dx(t) = f(x(t), t)dt + g(x(t), t)dB(t) \quad \text{on } t_0 \le t \le T, \tag{2.3}$$

with initial value $x(t_0) = x_0$. By the definition of stochastic differential, this equation is equivalent to the following integral equation:

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s), s)ds + \int_{t_0}^{t} g(x(s), s)dB(s) \quad \text{on } t_0 \le t \le T. \tag{2.4}$$

Let us now provide the definition of the solution.

**Definition 2.6.1.** *An $\mathbb{R}^d$-valued stochastic process $\{x(t)\}_{t_0\le t\le T}$ is called a solution of equation* (2.3) *if it has the following properties:*

*i.* *$\{x(t)\}$ is continuous and $\mathcal{F}_t$-adapted;*

*ii.* *$\{f(x(t), t)\} \in \mathcal{L}^1([t_0, T]; \mathbb{R}^d)$ and $\{g(x(t), t)\} \in \mathcal{L}^2([t_0, T]; \mathbb{R}^{d\times m})$;*

*iii.* *equation* (2.4) *holds for every $t \in [t_0, T]$ with probability 1.*

*A solution $\{x(t)\}$ is said to be unique if any other solution $\{\bar{x}(t)\}$ is indistinguishable from $\{x(t)\}$, that is*

$$\mathbb{P}\{x(t) = \bar{x}(t) \text{ for all } t_0 \le t \le T\} = 1.$$

The following theorem provides conditions to guarantee existence and uniqueness of the solution to SDE (2.3)

**Theorem 2.6.2.** *Assume that there exist two positive constants $\bar{K}$ and $K$ such that*

*i.* *(Lipschitz condition) for all $x, \bar{x} \in \mathbb{R}^d$ and $t \in [t_0, T]$*

$$|f(x, t) - f(\bar{x}, t)|^2 \vee |g(x, t) - g(\bar{x}, t)|^2 \le \bar{K}|x - \bar{x}|^2; \tag{2.5}$$

*ii. (Linear growth condition) for all* $(x,t) \in \mathbb{R}^d \times [t_0, T]$

$$|f(x,t)|^2 \vee |g(x,t)|^2 \leq K(1 + |x|^2). \tag{2.6}$$

*Then there exists a unique solution* $x(t)$ *to equation* (2.3) *and the solution belongs* $\mathcal{M}^2([t_0, T]; \mathbb{R}^d)$.

The Lipschitz condition (2.5) means that the coefficients $f(x,t)$ and $g(x,t)$ do not change faster than a linear function of $x$ as change in $x$. This implies in particular the continuity of $f(x,t)$ and $g(x,t)$ in $x$ for all $t \in [t_0, T]$. Hence, functions that are discontinuous with respect to $x$ are excluded as the coefficients. This shows that the Lipschitz condition is too restrictive. The following theorem is the generalisation of Theorem 2.6.2. in which this (uniform) Lipschitz condition is replaced by the local Lipschitz condition.

**Theorem 2.6.3.** *Assume that the linear growth condition* (2.6) *holds, but the Lipschitz condition* (2.5) *is replaced with the following local Lipschitz condition: For every integer* $n \geq 1$, *there exists a positive constant* $K_n$ *such that, for all* $t \in [t_0, T]$ *and all* $x, \bar{x} \in \mathbb{R}^d$ *with* $|x| \vee |\bar{x}| \leq n$

$$|f(x,t) - f(\bar{x},t)|^2 \vee |g(x,t) - g(\bar{x},t)|^2 \leq K_n |x - \bar{x}|^2. \tag{2.7}$$

*Then there exists a unique solution* $x(t)$ *to equation* (2.3) *and the solution belongs* $\mathcal{M}^2([t_0, T]; \mathbb{R}^d)$.

The local Lipschitz condition allows us to include many functions. However, the linear growth condition still excludes some important functions. The following result improves the situation.

**Theorem 2.6.4.** *Assume that the local Lipschitz condition* (2.7) *holds but the linear growth condition* (2.6) *is replaced with the following monotone condition: There exists a positive constant* $K$ *such that for all* $(x,t) \in \mathbb{R}^d \times [t_0, T]$

$$x^T f(x,t) + \frac{1}{2}|g(x,t)|^2 \leq K(1 + |x|^2). \tag{2.8}$$

*Then there exists a unique solution $x(t)$ to equation* (2.3) *in* $\mathcal{M}^2([t_0, T]; \mathbb{R}^d)$.

## 2.7 Poisson processes

The Brownian motion and Poisson process are the two basic examples of stochastic processes. In this section, we present the key mathematical concept of a Poisson process. We let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions.

**Definition 2.7.1.** *A stochastic process $N = \{N(t)\}_{t\geq 0}$ taking values in $\{0, 1, 2, \cdots\}$ is said to be a Poisson process with intensity $\lambda > 0$ if the following conditions hold:*

*i.* $N(0) = 0$ *a.s.;*

*ii. For any $0 \leq t_0 < t_1 < \cdots < t_k < \infty$ and $1 \leq n \leq k-1$, the increments $N(t_{n+1}) - N(t_n)$ are independent Poisson random variables with means $\lambda(t_{n+1} - t_n)$;*

*iii. The sample paths $\{N(t, \omega)\}_{t\geq 0}$ of the process $N$ are right-continuous with left limits a.s.*

It follows from Definition 2.7.1 that

$$\mathbb{P}(N(t) = n) = \frac{e^{\lambda t}(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \cdots .$$

**Theorem 2.7.2.** *Let $N$ be a Poisson process with intensity $\lambda$. Then $N(t) - \lambda t$ is a martingale and is called a compensated Poisson process.*

## 2.8 SDEs with Poisson jumps

Let $B(t) = (B_1(t), \cdots, B_m(t))^T$, $t \geq 0$, be an $m$-dimensional Brownian motion defined on the above probability space. Let $N(t)$ be Poisson process independent of $B(t)$ with compensated Poisson process $\widetilde{N}(t) = N(t) - \lambda t$, where $\lambda$ is the jump intensity, also defined on the above probability space. The $d$-dimensional stochastic differential equations with Poisson jumps is given by:

$$dx(t) = f(x(t^-), t)dt + g(x(t^-), t)dB(t) + h(x(t^-), t)dN(t) \tag{2.9}$$

on $t \geq 0$ with initial value $x(0) = x_0 \in \mathbb{R}^d$. Here $f : \mathbb{R}^d \times [t_0, T] \to \mathbb{R}^d$, $g : \mathbb{R}^d \times [t_0, T] \to \mathbb{R}^{d\times m}$ and $h : \mathbb{R}^d \times [t_0, T] \to \mathbb{R}^d$ are Borel measurable and $x(t^-) = \lim_{s\to t^-} x(s)$. The following theorem reveals conditions for the existence and uniqueness of the solution to equation (2.9).

**Theorem 2.8.1.** *Assume that there exist two positive constants $\bar{K}_1$ and $K_1$ such that*

*i. (Lipschitz condition) for all $x, \bar{x} \in \mathbb{R}^d$ and $t \in [t_0, T]$*

$$|f(x,t) - f(\bar{x},t)|^2 \vee |g(x,t) - g(\bar{x},t)|^2 \vee |h(x,t) - h(\bar{x},t)|^2 \leq \bar{K}_1|x - \bar{x}|^2; \tag{2.10}$$

*ii. (Linear growth condition) for all $(x,t) \in \mathbb{R}^d \times [t_0, T]$*

$$|f(x,t)|^2 \vee |g(x,t)|^2 \vee |h(x,t)|^2 \leq K_1(1 + |x|^2). \tag{2.11}$$

*Then there exists a unique solution $x(t)$ to equation (2.9) and the solution belongs $\mathcal{M}^2([t_0, T]; \mathbb{R}^d)$.*

## 2.9 Markov processes

We will recall some fundamental concepts of a Markov process in this section. An $n$-dimensional $\mathcal{F}_t$-adapted process $X = \{X_t\}_{t\geq 0}$ is called a *Markov process* if the following *Markov property* is satisfied: for all $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}(\mathbb{R}^n)$,

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

This is equivalent to the following one: for any bounded Borel measurable function $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $0 \leq s \leq t < \infty$,

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

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

i. For every $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}(\mathbb{R}^n)$,

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

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

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

iv. The Kolmogorov-Chapman equation

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

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

A stochastic process $X = \{X(t)\}_{t \geq 0}$, defined on a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, with values in a countable set $\Xi$ (to be called the *state space* of the process), is called a *continuous-time Markov chain* if for any finite set $0 \leq t_1 < t_2 < \cdots < t_n < t_{n+1}$ of "times", and corresponding set $i_1, i_2, \cdots, i_{n-1}, i, j$ of states in $\Xi$ such that $\mathbb{P}\{X(t_n) = i, X(t_{n-1}) = i_{n-1}, \cdots, X(t_1) = i_1\} > 0$, we have

$$\begin{aligned}
&\mathbb{P}\{X(t_{n+1}) = j|X(t_n) = i, X(t_{n-1}) = i_{n-1}, \cdots, X(t_1) = i_1\} \\
&\quad = \mathbb{P}\{X(t_{n+1}) = j|X(t_n) = i\}.
\end{aligned}$$

If for all $s, t$ such that $0 \leq s \leq t < \infty$ and all $i, j \in \Xi$ the conditional probability $\mathbb{P}\{X(t) = j|X(s) = i\}$ depends only on $t - s$, we say that the process $X = \{X(t)\}_{t \geq 0}$

is *homogeneous*. In this case, then, $\mathbb{P}\{X(t) = j|X(s) = i\} = \mathbb{P}\{X(t-s) = j|X(0) = i\}$, and the function

$$P_{ij}(t) =: \mathbb{P}\{X(t) = j|X(s) = i\}, \quad i, j \in \Xi, t \geq 0,$$

is called the *transition function or transition probability* of the process. The function $P_{ij}(t)$ is called *standard* if $\lim_{t\to 0} P_{ii}(t) = 1$ for all $i \in \Xi$.

**Theorem 2.9.1.** *Let $P_{ij}(t)$ be a standard transition function, then*

$$\gamma_i := \frac{\lim_{t\to 0}[1 - P_{ii}(t)]}{t}$$

*exists (but may be $\infty$) for all $i \in \Xi$.*

A state $i \in \Xi$ is said to be *stable* if $\gamma_i < \infty$.

**Theorem 2.9.2.** *Let $P_{ij}(t)$ be a standard transition function, and let $j$ be a stable state. Then $\gamma_{ij} = P'_{ij}(0)$ exists and is finite for all $i \in \Xi$.*

Let $\gamma_{ii} = -\gamma_i$ and $\Gamma = (\gamma_{ij})_{i,j\in\Xi}$. $\Gamma$ is called the *generator* of the Markov chain. If the state space is *finite* which we can take to be $\mathcal{S} = \{1, 2, \cdots, N\}$, then the process is called a continuous-time *finite* Markov chain. We assume that all Markov chains are finite and all states are stable throughout this thesis. For such a Markov chain, almost every sample path is a right continuous step function.

**Theorem 2.9.3.** *Let $P(t) = (P_{ij}(t))_{N\times N}$ be the transition probability matrix and $\Gamma = (\gamma_{ij})_{N\times N}$ be the generator of a finite Markov chain. Then*

$$P(t) = e^{t\Gamma}.$$

It is useful to note that a continuous-time Markov chain $X(t)$ with generator $\Gamma = (\gamma_{ij})_{N\times N}$ can be represented as a stochastic integral with respect to a Poisson random measure. Indeed, let $\Delta_{ij}$ be consecutive, left closed, right open intervals of the real line each having length $\gamma_{ij}$ such that

$$\begin{aligned}
\Delta_{12} &= [0, \gamma_{12}), \\
\Delta_{13} &= [\gamma_{12}, \gamma_{12} + \gamma_{13}), \\
&\vdots \\
\Delta_{1N} &= \left[ \sum_{j=2}^{N-1} \gamma_{1j}, \sum_{j=2}^{N} \gamma_{1j} \right), \\
\Delta_{21} &= \left[ \sum_{j=2}^{N} \gamma_{1j}, \sum_{j=2}^{N} \gamma_{1j} + \gamma_{21} \right), \\
\Delta_{23} &= \left[ \sum_{j=2}^{N} \gamma_{1j} + \gamma_{21}, \sum_{j=2}^{N} \gamma_{1j} + \gamma_{21} + \gamma_{23} \right), \\
&\vdots \\
\Delta_{2N} &= \left[ \sum_{j=2}^{N} \gamma_{1j} + \sum_{j=1, j\neq 2}^{N-1} \gamma_{2j}, \sum_{j=2}^{N} \gamma_{1j} + \sum_{j=1, j\neq 2}^{N} \gamma_{2j} \right)
\end{aligned}$$

and so on. Define a function

$$h : \mathcal{S} \times \mathbb{R} \to \mathbb{R}$$

by

$$h(i, y) = \begin{cases} j - i & \text{if } y \in \Delta_{ij} \ , \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

Then

$$dX(t) = \int_{\mathbb{R}} h(X^-, y) \nu(dt, dy),$$

with initial condition $X(0) = i_0$, where $\nu(dt, dy)$ is a Poisson random measure with intensity $dt \times \mu(dy)$, in which $\mu$ is the Lebesgue measure on $\mathbb{R}$.

## 2.10 Generalised Itô's formula

Let $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\}$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e, it is increasing and right continuous while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets). Let $B(t) = (B_1(t), \cdots, B_m(t))^T$, $t \geq 0$, be an $m$-dimensional Brownian motion defined on the above probability space. Also let $r(t), t \geq 0$, be a right-continuous Markov chain defined on the above probability space taking values in a finite state space $\mathcal{S} = \{1, 2, \cdots, N\}$ with the generator $\Gamma = (\gamma_{ij})_{N\times N}$ given by

$$\mathbb{P}\{r(t+\delta) = j | r(t) = i\} = \begin{cases} \gamma_{ij}\delta + o(\delta) & \text{if } i \neq j, \\ 1 + \gamma_{ii}\delta + o(\delta) & \text{if } i = j, \end{cases}$$

where $\delta > 0$. Here $\gamma_{ij} \geq 0$ is the transition rate from $i$ to $j$ if $i \neq j$ while

$$\gamma_{ii} = -\sum_{i\neq j} \gamma_{ij}.$$

We assume that the Markov chain $r(\cdot)$ is $\mathcal{F}_t$-adapted but independent of the Brownian motion $B(\cdot)$.

Let $x(t)$ be an $n$-dimensional Itô process on $t \geq 0$ with the stochastic differential

$$dx(t) = f(t)dt + g(t)dB(t),$$

where $f \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^n)$ and $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{n\times m})$. The Itô formula established in Section 2.5 shows that a $C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+; \mathbb{R}_+)$-function $H$ maps the Itô process $x(t)$ into another Itô process $H(x(t), t)$. Here, we let a function $H : \mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}$ map a paired process $(x(t), r(t))$ into another process $H(x(t), r(t), t)$. For this purpose, let $C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}; \mathbb{R})$ denote the family of all real-valued functions $H(x, t, i)$ on $\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}$ which are continuously twice differentiable in $x$ and once in $t$. If

$H \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}; \mathbb{R})$, define $LH$ from $\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}$ to $\mathbb{R}$ by

$$
\begin{aligned}
LH(x,t,i) &= H_t(x,t,i) + H_x(x,t,i)f(t) \\
&\qquad + \frac{1}{2}\mathrm{trace}(g^T(t)H_{xx}(x,t,i)g(t)) + \sum_{j=1}^{N} \gamma_{ij} H(x,t,j),
\end{aligned}
$$

where

$$
H_t(x,t,i) = \frac{\partial H(x,t,i)}{\partial t}, \quad H_x(x,t,i) = \Big(\frac{\partial H(x,t,i)}{\partial x_1}, \cdots, \frac{\partial H(x,t,i)}{\partial x_n}\Big)
$$

$$
H_{xx}(x,t,i) = \Big(\frac{\partial^2 H(x,t,i)}{\partial x_i \partial x_j}\Big)_{n\times n}.
$$

The following theorem is known as the generalised Itô's formula.

**Theorem 2.10.1.** *If $H \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}; \mathbb{R})$, then for any $t \geq 0$*

$$
\begin{aligned}
&H(x(t), t, r(t)) \\
&= H(x(0), 0, r(0)) + \int_0^t LH(x(s), s, r(s))ds \\
&+ \int_0^t H_x(x(s), s, r(s))g(x(s), s, r(s))dB(s) \\
&+ \int_0^t \int_{\mathbb{R}} (H(x(s), s, i_0 + h(r(s), l)) - H(x(s), s, r(s)))\mu(ds, dl),
\end{aligned}
$$

*where the function $h$ is defined by* (2.12) *and $\mu(ds, dl) = \nu(ds, dl) - \mu(dl)ds$ is a martingale measure, while $\nu$ and $\mu$ have been defined in the end of Section 2.9.*

## 2.11 SDEs with Markovian switching

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 us assume that the Markov chain $r(\cdot)$ is $\mathcal{F}_t$-adapted but independent of the Brownian motion $B(\cdot)$, Consider an SDE with Markovian switching of the form

$$dx(t) = f(x(t), t, r(t))dt + g(x(t), t, r(t))dB(t), \quad t_0 \le t \le T \tag{2.13}$$

with initial data $x(t_0) = x_0 \in L^2_{\mathcal{F}_{t_0}}(\Omega; \mathbb{R}^n)$ and $r(t_0) = r_0$, where $r_0$ is an $\mathcal{S}$-valued $\mathcal{F}_{t_0}$-measurable random variable and

$$f : \mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}^n \quad \text{and} \quad g : \mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}^{n \times m}.$$

**Definition 2.11.1.** *An $\mathbb{R}^n$-valued stochastic process $\{x(t)\}_{t_0 \le t \le T}$ is called a solution of equation* (2.13) *if it has the following properties.*

*i.* $\{x(t)\}_{t_0 \le t \le T}$ *is continuous and $\mathcal{F}_t$-adapted;*

*ii.* $\{f(x(t), t, r(t))\}_{t_0 \le t \le T} \in \mathcal{L}^1([t_0, T]; \mathbb{R}^n)$ *while* $\{g(x(t), t, r(t))\}_{t_0 \le t \le T} \in \mathcal{L}^2([t_0, T]; \mathbb{R}^{n \times m})$*;*

*iii. for any $t \in [t_0, T]$, equation*

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(s), s, r(s))dt + \int_{t_0}^{t} g(x(s), s, r(s))dB(s),$$

*holds with probability 1.*

**Theorem 2.11.2.** *Assume that there exist two positive constants $\bar{K}_2$ and $K_2$ such that*

*i.* (*Lipschitz condition*) *for all $x, \bar{x} \in \mathbb{R}^n$, $t \in [t_0, T]$ and $i \in \mathcal{S}$*

$$|f(x, t, i) - f(\bar{x}, t, i)|^2 \vee |g(x, t, i) - g(\bar{x}, t, i)|^2 \le \bar{K}_2 |x - \bar{x}|^2; \tag{2.14}$$

*ii.* (*Linear growth condition*) *for all $(x, t, i) \in \mathbb{R}^n \times [t_0, T] \times \mathcal{S}$*

$$|f(x, t, i)|^2 \vee |g(x, t, i)|^2 \le K_2(1 + |x|^2). \tag{2.15}$$

*Then there exists a unique solution $x(t)$ to equation* (2.13) *and the solution belongs* $\mathcal{M}^2([t_0, T]; \mathbb{R}^n)$.

The following theorem shows the existence of unique maximal local solution under the Local Lipschitz condition without the linear growth condition.

**Theorem 2.11.3.** *Assume that* (*local Lipschitz condition*) *for every integer $k \geq 1$, there exists a positive constant $h_k$ such that, for all $t \in [t_0, T]$, $i \in \mathcal{S}$ and those $x, \bar{x} \in \mathbb{R}^n$ with $|x| \vee |\bar{x}| \leq k$,*

$$|f(x,t,i) - f(\bar{x},t,i)|^2 \vee |g(x,t,i) - g(\bar{x},t,i)|^2 \leq h_k|x - \bar{x}|^2. \tag{2.16}$$

*Then there exists a unique maximal local solution $x(t)$ to equation* (2.13).

The following theorem is an improved version of Theorem 2.11.3.

**Theorem 2.11.4.** *Assume that the local Lipschitz condition* (2.16) *holds but the linear growth condition* (2.15) *is replaced with the following monotone condition: There exists a positive constant $K$ such that for all $(x, t, i) \in \mathbb{R}^n \times [t_0, T] \times \mathcal{S}$*

$$x^T f(x,t,i) + \frac{1}{2}|g(x,t,i)|^2 \leq K(1 + |x|^2). \tag{2.17}$$

*Then there exists a unique solution $x(t)$ to equation* (2.13) *in $\mathcal{M}^2([t_0, T]; \mathbb{R}^n)$.*

## 2.12 Some useful inequalities

Let us also present some useful inequalities which are used frequently in this thesis. Let us start with the simplest inequality

$$2ab \leq a^2 + b^2, \quad \forall a, b \in \mathbb{R}.$$

From this follows

$$2ab \leq \epsilon a^2 + \frac{1}{\epsilon}b^2, \quad \forall a, b \in \mathbb{R} \text{ and } \forall \epsilon > 0.$$

Let us also proceed to the Young inequality

$$|a|^{\beta}|b|^{(1-\beta)} \leq \beta|a| + (1-\beta)|b|, \quad \forall a, b \in \mathbb{R} \text{ and } \forall \beta \in [0,1].$$

**Theorem 2.12.1.** (***Jensen's inequality***) *If $\varphi : \Omega \to \mathbb{R}$ is a convex function while $\xi : \mathbb{R} \to \mathbb{R}$ is a random varaible on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}|\xi| < \infty$, then*

$$\varphi(\mathbb{E}\xi) \leq \mathbb{E}(\varphi(\xi)).$$

**Theorem 2.12.2.** (***Doob's martingale inequalities***) *Let $\{M_t\}_{t\geq 0}$ be an $\mathbb{R}^d$-valued martingale. Let $[a,b]$ be a bounded interval in $\mathbb{R}_+$.*

*i. If $p \geq 1$ and $M_t \in L^p(\Omega; \mathbb{R}^d)$, then*

$$\mathbb{P}\Big\{\omega : \sup_{a\leq t\leq b} |M_t(\omega)| \geq c\Big\} \leq \frac{\mathbb{E}|M_b|^p}{c^p}$$

*holds for all $c > 0$.*

*ii. If $p > 1$ and $M_t \in L^p(\Omega; \mathbb{R}^d)$, then*

$$\mathbb{E}\Big(\sup_{a\leq t\leq b} |M_t|^p\Big) \leq \Big(\frac{p}{p-1}\Big)^p \mathbb{E}|M_b|^p.$$

**Theorem 2.12.3.** *Let $p \geq 2$. Let $g \in \mathcal{M}^2([0,T]; \mathbb{R}^{d\times m})$ such that*

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

*Then*

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

*In particular, for $p = 2$, there is equality.*

**Theorem 2.12.4.** *Let $p \geq 2$. Let $g \in \mathcal{M}^2([0,T]; \mathbb{R}^{d\times m})$ such that*

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

*Then*

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

**Theorem 2.12.5.** (***Burkholder-Davis-Gundy inequality***) *Let $g \in \mathcal{L}^2(\mathbb{R}_+;\mathbb{R}^{d\times m})$. Define, for $t \ge 0$,*

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

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

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

*for all $t \ge 0$. In particular, one may take*

$$\begin{aligned}
&c_p = (p/2)^p, && C_p = (32/p)^{\frac{p}{2}} && \textit{if } 0 < p < 2;\\
&c_p = 1, && C_p = 4 && \textit{if } p = 2;\\
&c_p = (2p)^{-\frac{p}{2}}, && C_p = [p^{p+1}/2(p-1)^{p-1}]^{\frac{p}{2}} && \textit{if } p > 2.
\end{aligned}$$

**Theorem 2.12.6.** (***Gronwall's inequality***) *Let $T > 0$ and $c \ge 0$. Let $u(\cdot)$ be a Borel measurable bounded non-negative function on $[0,T]$, and let $v(\cdot)$ be a non-negative integrable function on $[0,T]$. If*

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

*then*

$$u(t) \le c\,exp\Big(\int_0^t v(s)ds\Big) \quad \textit{for all } 0 \le t \le T.$$

# Chapter 3

# Truncated Euler-Maruyama method for Ait-Sahalia-type interest rate model with delay

## 3.1 Introduction

The original Ait-Sahalia model of the spot interest rate proposed by Ait-Sahalia assumes constant volatility. As supported by several empirical results, volatility is never constant in most financial markets. From application viewpoint, it is important we generalise the Ait-Sahalia model to incorporate volatility as a function of delay in the spot rate. In this chapter, we study analytical properties of the exact solution to this model. Apparently, the solution to this model cannot be found by a closed-form formula. Therefore, we construct a new implementable truncated EM method to study numerical properties of this model under the local Lipschitz condition plus the Khasminskii-type condition.

The rest of the chapter is organised as follows: We introduce the Ait-Sahalia-type interest rate model with delay in Section 3.2. In Section 3.3, we verify the existence and uniqueness of the solution to the proposed model and show that the solution will never become negative. We also study analytical properties such as boundedness of moments of the exact solution in Section 3.3. In Section 3.4, we construct a new implementable truncated EM scheme for the proposed model. We explore numerical properties to investigate a finite time strong convergence of this scheme in Section 3.5. In Section 3.6, we perform some numerical examples to support the established results and provide a brief summary of the results in Section 3.7.

## 3.2 The Ait-Sahalia-type model with delay

We 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 us now incorporate a delayed volatility function into SDE (1.3) to obtain a dynamics

$$\begin{cases} dx(t) = (\alpha_{-1}x(t)^{-1} - \alpha_0 + \alpha_1 x(t) - \alpha_2 x(t)^{\rho})dt + \varphi(x(t-\tau))x(t)^{\theta}dB(t), t > 0, \\ x(t) = \xi(t), \quad t \in [-\tau, 0], \end{cases} \tag{3.1}$$

for time-series evolution of interest rates. Here $\varphi(\cdot)$ is a volatility function which depends on $x(t-\tau)$, where $\tau > 0$ and $x(t-\tau)$ denotes delay in $x(t)$. The delayed volatility function is past-level-dependent in this case and hence, may describe dynamics of volatility 'smiles' and 'skews' adequately (e.g., see [25, 26]).

Consider the following scalar dynamics

$$dx(t) = f(x(t))dt + \varphi(x(t-\tau))g(x(t))dB(t), \tag{3.2}$$

as equation of SDDE (3.1) on $t \in [-\tau, \infty)$ with initial data $x(t) = \xi(t)$, where $f(x) = \alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^{\rho}$, $g(x) = x^{\theta}$, $\forall x \in \mathbb{R}_+$ and $\varphi(y) \in C(\mathbb{R}_+; \mathbb{R}_+)$. Let $C^{2,1}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ be the family of all real-valued functions $H(x,t)$ defined on $\mathbb{R} \times \mathbb{R}_+$ such that $H(x,t)$ is twice continuously differentiable in $x$ and once in $t$. Given $H \in C^{2,1}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$, we note the operator $LH : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is defined by

$$LH(x,y,t) = H_t(x,t) + H_x(x,t)f(x) + \frac{1}{2}H_{xx}(x,t)\varphi(y)^2 g(x)^2, \tag{3.3}$$

where $H_t(x,t)$ and $H_x(x,t)$ are first-order partial derivatives with respect to $t$ and $x$, and $H_{xx}(x,t)$ is a second-order partial derivative with respect to $x$. The Itô formula can now be written as

$$dH(x(t),t) = LH(x(t),x(t-\tau),t)dt + H_x(x(t),t)\varphi(x(t-\tau))g(x(t))dB(t) \quad \text{a.s.} \tag{3.4}$$

## 3.3 Analytical properties

We observe the $f$ and $g$ coefficient terms of SDDE (3.2) are non-globally Lipschitz continuous. Naturally, for SDDE (3.2) to have a pathwise unique global solution for any given initial data, both drift and diffusion terms are required to satisfy local Lipschitz condition plus super-linear growth condition (e.g., see [20] for more details). Clearly, this means we have to assume the volatility function $\varphi(\cdot)$ is locally Lipschitz continuous and bounded. The following theorem illustrates that the SDDE (3.2) admits a unique positive global solution. Moreover, since the SDDE (3.2) describes interest rate dynamics in the financial market, it is important the solution $x(t)$ should always be positive. The following conditions are however sufficient to establish a pathwise unique positive global solution $x(t)$ to SDDE (3.2).

**Assumption 3.3.1.** *The volatility function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ of* SDDE (3.2) *is Borel-measurable and bounded by a positive constant $\lambda$, that is*

$$\varphi(y) \le \lambda, \quad \forall y \in \mathbb{R}_+. \tag{3.5}$$

See, for instance, Mao and Sabanis [26] for detailed coverage of the above assumption. In addition to Assumption 3.3.1, we also require the following assumption on the parameter values to help control the potential growth likely to emerge from the diffusion term.

**Assumption 3.3.2.** *The parameters of the* SDDE (3.2) *satisfy*

$$1 + \rho > 2\theta, \quad \rho, \theta > 1.$$

### 3.3.1 Existence and uniqueness of solution

**Theorem 3.3.3.** *Let Assumptions 3.3.1 and 3.3.2 hold. Then for any given initial data*

$$\{x(t) : -\tau \leq t \leq 0\} = \xi(t) \in C([-\tau, 0] : \mathbb{R}_+), \tag{3.6}$$

*there exists a unique global solution $x(t)$ to* SDDE (3.2) *on $t \in [-\tau, \infty)$ and $x(t) > 0$ almost surely. This solution can be computed by the following step by step procedure: for $k = 0, 1, 2, \cdots$ and $t \in [k\tau, (k+1)\tau]$,*

$$x(t) = x(k\tau) + \int_{k\tau}^{t} f(x(s))ds + \int_{k\tau}^{t} \varphi(x(s-\tau))g(x(s))dB(s). \tag{3.7}$$

*Moreover, for any $T > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(\tau_n \leq T) = 0, \tag{3.8}$$

*where*

$$\tau_n = \inf\{t \geq 0 : x(t) \notin (1/n, n)\} \tag{3.9}$$

*for every sufficiently large integer $n$.*

We employ an inductive argument to establish this proof.

*Proof.* For $t \in [0, \tau]$, the SDDE (3.2) becomes the following SDE

$$dx(t) = f(x(t))dt + \varphi(\xi(t-\tau))g(x(t))dB(t),$$

with initial value $x(0) = \xi(0)$ and has a well-known unique positive global solution

$$x(t) = \xi(0) + \int_{0}^{t} f(x(s))ds + \int_{0}^{t} \varphi(\xi(s-\tau))g(x(s))dB(s). \tag{3.10}$$

The solution $x(t)$ to SDE (3.10) on $t \geq 0$ has been however established in various literature to satisfy (3.8) (see, e.g., [18,20] for more details). This implies (3.7) holds for $k = 0$. As $x(t)$ is now known on $t \in [0, \tau]$, we may repeat this procedure over the interval $t \in [\tau, 2\tau]$ to obtain the SDE

$$dx(t) = f(x(t))dt + \varphi(x(t-\tau))g(x(t))dB(t).$$

This SDE has a unique positive global solution

$$x(t) = \xi(0) + \int_0^t f(x(s))ds + \int_0^t \varphi(x(s-\tau))g(x(s))dB(s). \tag{3.11}$$

Clearly the solution $x(t)$ is a continuous stochastic process on $t \in [0, \tau]$ and so both integrals are well defined. Hence the (3.7) holds for $k = 1$. Given that the solution $x(t)$ to SDE (3.10) on $t \geq 0$ satisfies (3.8) implies it also satisfies (3.8) for SDE (3.11). Repeating this procedure for all $k \geq 0$, we obtain a unique positive global solution to SDDE (3.2) which satisfies (3.8). □

### 3.3.2 Moment bounds

Finiteness of moments is essential for evaluating and pricing financial quantities. The following lemmas give boundedness property of the exact solution to SDDE (3.2).

**Lemma 3.3.4.** *Let Assumptions 3.3.1 and 3.3.2 hold. Then for any $p \geq 2$, the solution $x(t)$ to* SDDE (3.2) *satisfies*

$$\sup_{0 \leq t < \infty} (\mathbb{E}|x(t)|^p) \leq C_1 \tag{3.12}$$

*and*

$$\sup_{0 \leq t < \infty} \left(\mathbb{E}|\frac{1}{x(t)}|^p\right) \leq C_2, \tag{3.13}$$

*where $C_1$ and $C_2$ are constants which depend on the initial value $\xi(t)$ and $p$.*

*Proof.* Let $n_0 > 0$ be sufficiently large such that

$$\frac{1}{n_0} < \min_{-\tau \leq t \leq 0} |\xi(t)| \leq \max_{-\tau \leq t \leq 0} |\xi(t)| < n_0.$$

For each integer $n \geq n_0$, define the stopping time by

$$\tau_n = \inf\{t \geq 0 : x(t) \notin (1/n, n)\}.$$

Applying the diffusion operator to $H(x,t) = e^t x^p$, we compute

$$\begin{aligned}
LH(x,y,t) &= e^t x^p + pe^t x^{p-1} f(x) + \frac{1}{2}p(p-1)e^t x^{p-2}(\varphi(y)g(x))^2 \\
&= e^t x^p + pe^t x^{p-1}\big(\alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^{\rho}\big) \\
&+ \frac{1}{2}p(p-1)e^t x^{p-2}\varphi(y)^2 x^{2\theta} \\
&\leq e^t\big[x^p + px^{p-2}(\alpha_{-1} - \alpha_0 x + \alpha_1 x^2 - \alpha_2 x^{\rho+1} + \frac{(p-1)}{2}\lambda^2 x^{2\theta})\big],
\end{aligned}$$

where Assumption 3.3.1 has been used. Moreover, by Assumption 3.3.2, there exists a constant $K$ such that

$$LH(x,y,t) \leq e^t K. \tag{3.14}$$

By the Itô formula, we obtain

$$\begin{aligned}
\mathbb{E}[e^{t\wedge\tau_n}|x(t\wedge\tau_n)|^p] &\leq |\xi(0)|^p + \mathbb{E}\int_0^{t\wedge\tau_n} Ke^s ds \\
&\leq |\xi(0)|^p + Ke^t.
\end{aligned}$$

Applying the Fatou lemma and letting $n \to \infty$ gives

$$\mathbb{E}|x(t)|^p \leq \frac{|\xi(0)|^p}{e^t} + K < \infty, \tag{3.15}$$

and hence

$$\sup_{0\leq t<\infty}(\mathbb{E}|x(t)|^p) \leq C_1. \tag{3.16}$$

Similarly, we can show (3.13) in the same way by using the Itô formula on $H(x,t) = e^t/x^p$, applying the Fatou lemma and letting $n \to \infty$. $\square$

**Lemma 3.3.5.** *Let Assumptions 3.3.1 and 3.3.2 hold. Then for any $p \geq 2$, the* *solution $x(t)$ to* SDDE (3.2) *satisfies*

$$\mathbb{E}\Big(\sup_{0\le t\le T}|x(t)|^p\Big) \le C_3, \tag{3.17}$$

*where $C_3$ is a constant.*

*Proof.* Define a function $H \in C^2(\mathbb{R}_+, \mathbb{R}_+)$ by

$$H(x) = x^p. \tag{3.18}$$

By the Itô formula, we compute

$$\begin{aligned} dH(x(t)) &= px^{p-1}dx(t) + \frac{1}{2}p(p-1)x^{p-2}(dx(t))^2 \\ &= px^{p-1}(\alpha_{-1}x(t)^{-1} - \alpha_0 + \alpha_1 x(t) - \alpha_2 x(t)^\rho \\ &+ \frac{1}{2}p(p-1)x(t)^{2(\theta-1)+p}\varphi(y)^2)dt + px(t)^{p+\theta-1}\varphi(y)dB(t) \\ &\le \big[px^{p-2}(\alpha_{-1} - \alpha_0 x(t) + \alpha_1 x(t)^2 - \alpha_2 x(t)^{\rho+1} \\ &+ \frac{(p-1)}{2}\lambda^2 x(t)^{2\theta})\big]dt + \lambda p x(t)^{p+\theta-1}dB(t), \end{aligned}$$

where Assumption 3.3.1 has been used. We now have

$$\begin{aligned} \mathbb{E}(\sup_{0\le t\le T}|x(t)|^p) \le |\xi(0)|^P + \mathbb{E}\int_0^T [px^{p-2}(\alpha_{-1} - \alpha_0 x(t) + \alpha_1 x(t)^2 - \alpha_2 x(t)^{\rho+1} \\ + \frac{(p-1)}{2}\lambda^2 x(t)^{2\theta})]dt + \mathbb{E}[\sup_{0\le t\le T}\int_0^t \lambda p x(s)^{p+\theta-1}dB(s)]. \end{aligned}$$

By Assumption 3.3.2, there exists a constant $\mathcal{K}$ such that

$$\mathbb{E}(\sup_{0\le t\le T}|x(t)|^p) \le |\xi(0)|^P + \mathcal{K}T + \mathbb{E}[\sup_{0\le t\le T}\int_0^t \lambda p x(s)^{p+\theta-1}dB(s)].$$

By the Hölder and Burkholder-Davis Gundy inequalities we then obtain,

$$\mathbb{E}(\sup_{0\le t\le T}|x(t)|^p) \le |\xi(0)|^P + \mathcal{K}T + \mathcal{C}\Big(\int_0^T \mathbb{E}x(s)^{2(p+\theta-1)}ds\Big)^{1/2},$$

where $\mathcal{C}$ is a constant which may vary from line to line. Hence

$$\mathbb{E}\Big(\sup_{0\le t\le T}|x(t)|^p\Big) \le C_3.$$

□

## 3.4 Numerical method

As we have already noted, the truncated EM method for SDEs under local Lipschitz condition plus Khasminskii-type condition was developed in [27]. This numerical method was further developed in [28] to study SDDEs under local Lipschitz condition plus generalised Khasminskii-type condition. Hence, in order to study SDDE (3.2) using the truncated EM techniques, we need to imposed the following conditions on the coefficient terms.

**Assumption 3.4.1.** *For any $R > 0$, there exists a positive constant $L_R$ such that the volatility function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ of* SDDE (3.2) *satisfies*

$$|\varphi(y) - \varphi(\bar{y})| \le L_R|y - \bar{y}| \tag{3.19}$$

*for all $y, \bar{y} \in [\frac{1}{R}, R]$.*

**Lemma 3.4.2.** *For any $R > 0$, there exists a positive constant $K_R > 0$ such that the $f$ and $g$ coefficient terms of* SDDE (3.2) *satisfy*

$$|f(x) - f(\bar{x})| \vee |g(x) - g(\bar{x})| \le K_R|x - \bar{x}| \tag{3.20}$$

*for all $x, \bar{x} \in [\frac{1}{R}, R]$.*

**Lemma 3.4.3.** *Let Assumptions 3.3.1 and 3.3.2 hold. For any $p \geq 2$, there exists $K_1 = K(p) > 0$ such that the coefficients of* SDDE (3.2) *satisfy*

$$xf(x) + \frac{p-1}{2}|\varphi(y)g(x)|^2 \leq K_1(1 + |x|^2) \tag{3.21}$$

*for all $x, y \in \mathbb{R}_+$.*

*Proof.* By Assumption 3.3.1, $\forall x, y > 0$, we have that

$$\begin{aligned} xf(x) + \frac{p-1}{2}|\varphi(y)g(x)|^2 &= x(\alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^\rho) + \frac{p-1}{2}|\varphi(y)x^\theta|^2 \\ &\leq \alpha_{-1} - \alpha_0 x + \alpha_1 x^2 - \alpha_2 x^{\rho+1} + \frac{p-1}{2}\lambda^2 x^{2\theta}. \end{aligned}$$

By Assumption 3.3.2,

$$\begin{aligned} xf(x) + \frac{p-1}{2}|\varphi(y)g(x)|^2 &\leq \alpha_{-1} - \alpha_0 x + \alpha_1 x^2 + K(p) \\ &\leq \alpha_{-1} + \alpha_1 x^2 + K(p) \\ &\leq K_1(1 + |x|^2), \end{aligned}$$

where $K(p) \geq -\alpha_2 x^{\rho+1} + \frac{p-1}{2}\lambda^2 x^{2\theta}$ and $K_1 = [(\alpha_{-1} + K(p)) \vee \alpha_1]$. $\square$

### 3.4.1 The truncated EM method

Before we proceed to construct the truncated EM scheme, let us extend the domain of the volatility function $\varphi(y)$ from $\mathbb{R}_+$ to $\mathbb{R}$ by setting the volatility function $\varphi(x) = \varphi(0)$ for $x < 0$. It is worth to note that the solution for the SDDE (3.2) is already known to always be positive, so this extension does not in any way influence the solution. The local Lipschitz condition in Assumption 3.4.1 and the boundedness condition on $\varphi(y)$ in (3.5) are also well preserved.

To define the truncated EM numerical solutions for the SDDE (3.2), we first choose a strictly increasing continuous function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\mu(r) \to \infty$ as $r \to \infty$

and

$$\sup_{1/r \le x \le r} (|f(x)| \vee g(x)) \le \mu(r), \quad \forall r > 1. \tag{3.22}$$

Denote by $\mu^{-1}$ the inverse function of $\mu$. We define a strictly decreasing function $h : (0,1) \to \mathbb{R}_+$ such that

$$\lim_{\Delta \to 0} h(\Delta) = \infty \text{ and } \Delta^{1/4} h(\Delta) \le 1, \quad \forall \Delta \in (0,1]. \tag{3.23}$$

Find $\Delta^* \in (0,1)$ such that $\mu^{-1}(h(\Delta^*)) > 1$ and $f(x) > 0$ for $0 < x < \Delta^*$. For a given step size $\Delta \in (0, \Delta^*)$, let us define the truncated functions

$$f_\Delta(x) = f\Big(1/\mu^{-1}(h(\Delta)) \vee (x \wedge \mu^{-1}(h(\Delta)))\Big), \quad \forall x \in \mathbb{R}$$

and

$$g_\Delta(x) = \begin{cases} g\Big(x \wedge \mu^{-1}(h(\Delta))\Big), & \text{if } x \ge 0 \\ 0, & \text{if } x < 0. \end{cases}$$

That is, for $x < 1/\mu^{-1}(h(\Delta))$, we have $f_\Delta(x) = f(1/\mu^{-1}(h(\Delta)))$ and if $x < 1 \wedge \mu^{-1}(h(\Delta))$, $g_\Delta(x) = g(x)$ or 0 if $x < 0$. For $x > \mu^{-1}(h(\Delta))$, we have $f_\Delta(x) = f(\mu^{-1}(h(\Delta)))$ and $g_\Delta(x) = g(\mu^{-1}(h(\Delta)))$. Moreover, for $x \in [1/\mu^{-1}(h(\Delta)), \mu^{-1}(h(\Delta))]$, we have $f_\Delta(x) = f(x)$ and $g_\Delta(x) = g(x)$ with

$$\begin{aligned} |f_\Delta(x)| = |f(x)| &\le \max_{1/\mu^{-1}(h(\Delta)) \le z \le \mu^{-1}(h(\Delta))} |f(z)| \\ &\le \mu(\mu^{-1}(h(\Delta))) \\ &= h(\Delta) \end{aligned}$$

and

$$g_\Delta(x) \le \mu(\mu^{-1}(h(\Delta))) = h(\Delta).$$

It is easy to see that

$$|f_\Delta(x)| \vee g_\Delta(x) \le h(\Delta), \quad \forall x \in \mathbb{R}. \tag{3.24}$$

Obviously, both truncated functions $f_\Delta$ and $g_\Delta$ are bounded although both $f$ and $g$ may not. The following lemma illustrates $f_\Delta$ and $g_\Delta$ preserve the Khasminskii-type condition in (3.21) very well.

**Lemma 3.4.4.** *Let Assumption 3.3.1 and 3.3.2 hold. Then, for all $\Delta \in (0, \Delta^*)$ and $p \ge 2$, the truncated functions satisfy*

$$xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \le \bar{K}(1 + |x|^2) \tag{3.25}$$

*$\forall x, y \in \mathbb{R}$, where $\bar{K}$ is a positive constant independent of $\Delta$.*

*Proof.* Fix any $\Delta \in (0, \Delta^*)$. For $x, y \in \mathbb{R}$ with $x \in [1/\mu^{-1}(h(\Delta)), \mu^{-1}(h(\Delta))]$, by (3.21), we have

$$xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 = xf(x) + \frac{p-1}{2}|\varphi(y)g(x)|^2 \le K_1(1 + |x|^2)$$

as the required assertion. For $x \in \mathbb{R}$ with $x \in (0, 1/\mu^{-1}(h(\Delta)))$, we have

$$0 < x\mu^{-1}(h(\Delta)) < 1$$

. So by (3.21), we get

$$\begin{aligned}
& xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \\
& = xf(1/\mu^{-1}(h(\Delta))) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \\
& = x\mu^{-1}(h(\Delta))\frac{1}{\mu^{-1}(h(\Delta))}f(1/\mu^{-1}(h(\Delta))) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \\
& \le K_1 x\mu^{-1}(h(\Delta))(1 + [1/\mu^{-1}(h(\Delta))^2]) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2.
\end{aligned}$$

This follows that

$$\begin{aligned} xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 &\le K_1(1+1) + K_1(1+|x|^2) \\ &= 2K_1 + K_1(1+|x|^2) \\ &\le K_2(1+|x|^2), \end{aligned}$$

where $K_2 = 3K_1$. But for $x, y \in \mathbb{R}$ with $x \le 0$, we have

$$f_\Delta(x) = f(1/\mu^{-1}(h(\Delta))) > 0 \text{ and } g_\Delta(x) = 0.$$

Therefore,

$$xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \le 0 \le K_1(1+|x|^2).$$

Finally, for $x, y \in \mathbb{R}$ with $x > \mu^{-1}(h(\Delta))$, we have

$$\begin{aligned} & xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \\ & \le xf(1/\mu^{-1}(h(\Delta)) \vee \mu^{-1}(h(\Delta))) + \frac{p-1}{2}|\varphi(y)g(\mu^{-1}(h(\Delta)))|^2 \\ & \le \mu^{-1}(h(\Delta))f(\mu^{-1}(h(\Delta))) + \frac{p-1}{2}|\varphi(y)g(\mu^{-1}(h(\Delta))|^2 \\ & + (\frac{x}{\mu^{-1}(h(\Delta))} - 1)\mu^{-1}(h(\Delta))f(\mu^{-1}(h(\Delta))) \\ & \le K_1(1+[\mu^{-1}(h(\Delta))]^2) + (\frac{x}{\mu^{-1}(h(\Delta))} - 1)\mu^{-1}(h(\Delta))f(\mu^{-1}(h(\Delta))), \end{aligned}$$

where (3.21) with $K_1$ independent of $\Delta$ has been used. But once again we see from (3.21) that $xf(x) \le K_1(1+|x|^2)$ for any $x \in \mathbb{R}_+$. We therefore have

$$\begin{aligned} xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 &\le K_1(1+[\mu^{-1}(h(\Delta))]^2) \\ &+ \Big(\frac{x}{\mu^{-1}(h(\Delta))} - 1\Big)K_1(1+[\mu^{-1}(h(\Delta))]^2) \\ &\le \frac{x}{\mu^{-1}(h(\Delta))}K_1(1+[\mu^{-1}(h(\Delta))]^2) \end{aligned}$$

$$\leq xK_1(1+\mu^{-1}(h(\Delta)))$$
$$\leq xK_1(1+x) \leq 2K_1(1+|x|^2).$$

It is worthwhile to note that $\bar{K} = (K_1 \vee K_2)$. □

From now on, we will let the step size $\Delta \in (0,1)$ be a fraction of $\tau$. That is, we will use $\Delta = \tau/N$ for sufficiently large integer $N$. Let form the discrete-time truncated approximation for SDDE (3.2). Define $t_k = k\Delta$ for $k = -N, -(N-1), \cdots, 0, 1, 2, \cdots$. Set $X_\Delta(t_k) = \xi(t_k)$ for $k = -N, -(N-1), \cdots, 0$ and form

$$X_\Delta(t_{k+1}) = X_\Delta(t_k) + f_\Delta(X_\Delta(t_k))\Delta + \varphi(X_\Delta(t_{k-N}))g_\Delta(X_\Delta(t_k))\Delta B_k \tag{3.26}$$

for $k = 0, 1, 2, \cdots$, where $\Delta B_k = B(t_{k+1}) - B(t_k)$. Let us now form two versions of the continuous-time truncated EM solutions. The first is defined by

$$\bar{x}_\Delta(t) = \sum_{k=-N}^{\infty} X_\Delta(t_k) 1_{[k\Delta,(k+1)\Delta)}(t). \tag{3.27}$$

This is the continuous-time step-process $\bar{x}_\Delta(t)$ on $t \in [-\tau, \infty]$, where $1_{[k\Delta,(k+1)\Delta]}$ is the indicator function on $[k\Delta, (k+1)\Delta]$. The other is the continuous-time continuous process $x_\Delta(t)$ on $t \in [-\tau, \infty]$ defined by setting $x_\Delta(t) = \xi(t)$ for $t \in [-\tau, 0]$ while for $t \geq 0$

$$x_\Delta(t) = \xi(0) + \int_0^t f_\Delta(\bar{x}_\Delta(s))ds + \int_0^t \varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))dB(s). \tag{3.28}$$

We see that $x_\Delta(t)$ is an Itô process on $t \geq 0$ with its Itô differential

$$dx_\Delta(t) = f_\Delta(\bar{x}_\Delta(t))dt + \varphi(\bar{x}_\Delta(t-\tau))g_\Delta(\bar{x}_\Delta(t))dB(t). \tag{3.29}$$

We can clearly observe that $x_\Delta(t_k) = \bar{x}_\Delta(t_k) = X_\Delta(t_k)$ for all $k = -N, -(N-1), \cdots$. That is $x_\Delta(t)$ and $\bar{x}_\Delta(t)$ coincide with the discrete truncated EM approximate solution at the gridpoints. We would like to point out that this numerical scheme is not positivity-preserving. This will however be tackled elsewhere.

## 3.5 Numerical properties

Under this section, we establish boundedness of moments and strong convergence theory for the truncated EM solutions.

### 3.5.1 Moment bounds

To upper bound the pth moment of the truncated EM solution, we need the following lemma which shows $x_\Delta(t)$ and $\bar{x}_\Delta(t)$ are close to each other in the strong sense.

**Lemma 3.5.1.** *Let Assumption 3.3.1 hold. For any fixed $\Delta \in (0, \Delta^*]$ and $p \geq 2$, we have that*

$$\mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p \leq C_p \Delta^{p/2}(h(\Delta))^p, \quad \forall t \geq 0, \tag{3.30}$$

*where $C_p$ stands for generic positive real constants dependent only on $p$ and may change between occurrences. Consequently,*

$$\lim_{\Delta \to 0} \mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p = 0, \quad \forall t \geq 0. \tag{3.31}$$

*Proof.* Fix any $\Delta \in (0, \Delta^*)$ and $t \geq 0$. There exists an integer $k \geq 0$ such that $t_k \leq t \leq t_{k+1}$. By elementary inequality, (3.24) and Assumption 3.3.1, we obtain from (3.28) that

$$\begin{aligned}
&\mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p \\
&\leq 2^{p-1}\Big(\mathbb{E}|\int_{t_k}^t f_\Delta(\bar{x}_\Delta(s))ds|^p + \mathbb{E}|\int_{t_k}^t \varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))B(s)|^p\Big) \\
&\leq 2^{p-1}\Big(\Delta^{p-1}\mathbb{E}\int_{t_k}^t |f_\Delta(\bar{x}_\Delta(s))|^p ds + \bar{c}_p\Delta^{(p-2)/2}\mathbb{E}\int_{t_k}^t |\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))|^p ds\Big) \\
&\leq 2^{p-1}\Big(\Delta^{p-1}\Delta(h(\Delta))^p + \bar{c}_p\Delta^{(p-2)/2}\Delta(\lambda h(\Delta))^p\Big) \\
&\leq 2^{p-1}(1 \vee \bar{c}_p\lambda^p)\Delta^{p/2}(h(\Delta))^p
\end{aligned}$$

$$\leq C_p \Delta^{p/2}(h(\Delta))^p,$$

where $\bar{c}_p$ depends on $p$ and $C_p = 2^{p-1}(1 \vee \bar{c}_p \lambda^p)$. Noting from (3.23) that

$$\Delta^{p/2}(h(\Delta))^p \leq \Delta^{p/4},$$

we get (3.31) from (3.30). □

The following lemma reveals the upper bound of the truncated EM solutions.

**Lemma 3.5.2.** *Let Assumptions 3.3.1 and 3.3.2 hold. Then for any $p \geq 2$,we have*

$$\sup_{0 \leq \Delta \leq \Delta^*} \sup_{0 \leq t \leq T} (\mathbb{E}|x_\Delta(t)|^p) \leq C_4, \quad \forall T > 0, \tag{3.32}$$

*where $C_4$ stands for generic positive real constants dependent on $T, p, \bar{K}, \xi$ but independent of $\Delta$ and may change between occurrences.*

*Proof.* Fix any $\Delta \in (0, \Delta^*)$ and $T \geq 0$. By the Itô formula, we derive from (3.28) that, for $0 \leq t \leq T$,

$$\begin{aligned}
\mathbb{E}|x_\Delta(t)|^p \leq\, & |\xi(0)|^p + \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}\Big(x_\Delta(s) f_\Delta(\bar{x}_\Delta(s)) \\
& + \frac{p-1}{2}|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))|^2\Big)ds \\
=\, & |\xi(0)|^p + \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}\Big(\bar{x}_\Delta(s) f_\Delta(\bar{x}_\Delta(s)) \\
& + \frac{p-1}{2}|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))|^2\Big)ds \\
& + \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}(x_\Delta(s) - \bar{x}_\Delta(s)) f_\Delta(\bar{x}_\Delta(s))ds.
\end{aligned}$$

By Lemma 3.4.4 and the Young inequality, we then have

$$\mathbb{E}|x_\Delta(t)|^p \leq |\xi(0)|^p + \mathbb{E}\int_0^t \bar{K}|x_\Delta(s)|^{p-2}(1 + |\bar{x}_\Delta(s)|^2)ds$$

$$
\begin{aligned}
&+ (p-2)\mathbb{E}\int_0^t |x_\Delta(s)|^p ds + 2\mathbb{E}\int_0^t |x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2}|f_\Delta(\bar{x}_\Delta(s)|^{p/2} ds \\
&\leq C_5 + C_6 \int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p) ds \\
&+ 2\mathbb{E}\int_0^t |x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2}|f_\Delta(\bar{x}_\Delta(s)|^{p/2} ds,
\end{aligned}
$$

where $C_5$ and $C_6$ are positive constants independent of $\Delta$. By Lemma 3.5.1 and inequalities (3.24) and (3.23), we have

$$
\begin{aligned}
\mathbb{E}\int_0^t |x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2}|f_\Delta(\bar{x}_\Delta(s)|^{p/2} ds &\leq (h(\Delta))^{p/2}\int_0^T \mathbb{E}(|x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2}) ds \\
&\leq (h(\Delta))^{p/2}\int_0^T (\mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^p)^{1/2} ds \\
&\leq (h(\Delta))^{p/2}\int_0^T (C_p\Delta^{p/2}(h(\Delta))^p)^{1/2} ds \\
&\leq C_p T (h(\Delta))^p \Delta^{p/4} \leq C_p T.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\mathbb{E}|x_\Delta(t)|^p &\leq C_5 + 2C_pT + C_6\int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p) ds \\
&\leq C_5 + 2C_pT + 2C_6\int_0^t \sup_{0\leq u\leq s}\Big(\mathbb{E}|x_\Delta(u)|^p\Big) ds.
\end{aligned}
$$

As this holds for any $t \in [0, T]$ while the right-hand side is non-decreasing in t, we then see

$$
\sup_{0\leq u\leq t}(\mathbb{E}|x_\Delta(u)|^p) \leq C_5 + 2C_pT + 2C_6\int_0^t \sup_{0\leq u\leq s}\Big(\mathbb{E}|x_\Delta(u)|^p\Big) ds.
$$

The well-known Gronwall inequality gives us

$$
\sup_{0\leq u\leq T}(\mathbb{E}|x_\Delta(u)|^p) \leq C_4.
$$

As this holds for any $\Delta \in (0, \Delta^*)$ while $C_4 = (C_5 + 2C_pT)e^{2C_6T}$ is independent of $\Delta$, we obtain the required assertion. $\square$

### 3.5.2 Finite time strong convergence

For the numerical solution to converge in finite time to the exact solution in the strong sense, we need the following condition on the initial data (see, e.g, [29]).

**Assumption 3.5.3.** *There is a pair of constant $K_4 > 0$ and $\gamma \in (0, 1]$ such that for all $-\tau \le s \le t \le 0$, the initial data $\xi$ satisfies*

$$|\xi(t) - \xi(s)| \le K_4|t - s|^\gamma. \tag{3.33}$$

In addition to the above condition, we also need the following lemma.

**Lemma 3.5.4.** *Let Assumptions 3.3.1, 3.3.2, 3.4.1 and 3.5.3 hold and $T > 0$ be fixed. Then for any $\varepsilon \in (0, 1)$, there exists a pair of positive constants $n = n(\varepsilon)$ and $\Delta_1 = \Delta_1(\varepsilon)$ such that*

$$\mathbb{P}(\rho_n \le T) \le \varepsilon \tag{3.34}$$

*for each $\Delta \in (0, \Delta_1]$, where*

$$\rho_n = \rho_n(\Delta) = \inf\{t \in [0, T] : x_\Delta(t) \notin (1/n, n)\}$$

*is the stopping time.*

*Proof.* Define a $C^2$-function, $H : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$H(x) = 1/x^2 + x^2. \tag{3.35}$$

Clearly, $H(x) \to \infty$ as $x \to \infty$ or $x \to 0$. For $s \in [0, t \wedge \rho_n]$, we can derive from the Itô formula that

$$\mathbb{E}(H(x_\Delta(t \wedge \rho_n))) = H(\xi(0)) + \mathbb{E}\int_0^{t\wedge\rho_n} \Big(H_x(x_\Delta(s))f_\Delta(\bar{x}_\Delta(s)) \tag{3.36}$$

$$+\frac{1}{2}H_{xx}(x_\Delta(s))\varphi(\bar{x}_\Delta(s-\tau))^2 g_\Delta(\bar{x}_\Delta(s))^2\Big)ds.$$

But we note

$$\begin{aligned}
&H_x(x_\Delta(s))f_\Delta(\bar{x}_\Delta(s)) + \frac{1}{2}H_{xx}(x_\Delta(s))\varphi(\bar{x}_\Delta(s-\tau))^2 g_\Delta(\bar{x}_\Delta(s))^2\\
&\leq LH(x_\Delta(s), x_\Delta(s-\tau)) + H_x(x_\Delta(s))\Big(f_\Delta(\bar{x}_\Delta(s)) - f_\Delta(x_\Delta(s))\Big)\\
&+\frac{1}{2}H_{xx}(x_\Delta(s))\Big(\varphi(\bar{x}_\Delta(s-\tau))^2 g_\Delta(\bar{x}_\Delta(s))^2 - \varphi(x_\Delta(s-\tau))^2 g_\Delta(x_\Delta(s))^2\Big),
\end{aligned}$$

where $LH$ is (3.3) with $H$ independent of $t$, defined here by

$$LH(x_\Delta(s), x_\Delta(s-\tau)) = H_x(x_\Delta(s))f_\Delta(x_\Delta(s)) + \frac{1}{2}H_{xx}(x_\Delta(s))\varphi(x_\Delta(s-\tau))^2 g_\Delta(x_\Delta(s))^2.$$

By Assumptions 3.3.1 and 3.3.2, there exists a constant $K_3$ such that

$$LH(x_\Delta(s), x_\Delta(s-\tau)) \leq K_3$$

and

$$\begin{aligned}
&H_x(x_\Delta(s))f_\Delta(\bar{x}_\Delta(s)) + \frac{1}{2}H_{xx}(x_\Delta(s))\varphi(\bar{x}_\Delta(s-\tau))^2 g_\Delta(\bar{x}_\Delta(s))^2\\
&\leq K_3 + H_x(x_\Delta(s))\Big(f_\Delta(\bar{x}_\Delta(s)) - f_\Delta(x_\Delta(s))\Big) + \frac{1}{2}H_{xx}(x_\Delta(s))\Big(\varphi(\bar{x}_\Delta(s-\tau))^2 g_\Delta(\bar{x}_\Delta(s))^2\\
&-\varphi(x_\Delta(s-\tau))^2 g_\Delta(x_\Delta(s))^2\Big).
\end{aligned}$$

We recall from the definition of the truncated functions $f_\Delta$ and $g_\Delta$ that

$$f_\Delta(\bar{x}_\Delta(s)) = f(\bar{x}_\Delta(s)) \text{ and } g_\Delta(\bar{x}_\Delta(s)) = g(\bar{x}_\Delta(s))$$

for $s \in [0, t\wedge\rho_n]$. So by Lemma 3.4.2, we have that for $s \in [0, t\wedge\rho_n]$

$$|f(\bar{x}_\Delta(s)) - f(x_\Delta(s))| \vee |g(\bar{x}_\Delta(s)) - g(x_\Delta(s))| \leq K_n|\bar{x}_\Delta(s) - x_\Delta(s)|.$$

We observe that for any $\bar{x}_\Delta(s), x_\Delta(s) \in [1/n, n]$, by (3.22), we have

$$|g(\bar{x}_\Delta(s))| \vee |g(x_\Delta(s))| \le \mu(n).$$

So by Lemma 3.4.2, we have that for $s \in [0, t \wedge \rho_n]$

$$\begin{aligned}|g(\bar{x}_\Delta(s))^2 - g(x_\Delta(s))^2| &= |g(\bar{x}_\Delta(s)) - g(x_\Delta(s))||g(\bar{x}_\Delta(s)) + g(x_\Delta(s))| \\ &\le 2\mu(n)K_n|\bar{x}_\Delta(s) - x_\Delta(s)|.\end{aligned}$$

Moreover, for $s \in [0, t \wedge \rho_n]$, we obtain from Assumptions 3.3.1 and 3.4.1 that

$$\begin{aligned}&|\varphi(\bar{x}_\Delta(s-\tau))^2 - \varphi(x_\Delta(s-\tau))^2| \\ &= |\varphi(\bar{x}_\Delta(s-\tau)) - \varphi(x_\Delta(s-\tau))||\varphi(\bar{x}_\Delta(s-\tau)) + \varphi(x_\Delta(s-\tau))| \\ &\le 2\lambda L_n|\bar{x}_\Delta(s-\tau) - x_\Delta(s-\tau)|.\end{aligned}$$

Consequently,

$$\begin{aligned}&\varphi(\bar{x}_\Delta(s-\tau))^2 g(\bar{x}_\Delta(s))^2 - \varphi(x_\Delta(s-\tau))^2 g(x_\Delta(s))^2 = \varphi(\bar{x}_\Delta(s-\tau))^2 g(\bar{x}_\Delta(s))^2 \\ &- \varphi(\bar{x}_\Delta(s-\tau))^2 g(x_\Delta(s))^2 + \varphi(\bar{x}_\Delta(s-\tau))^2 g(x_\Delta(s))^2 - \varphi(x_\Delta(s-\tau))^2 g(x_\Delta(s))^2 \\ &= g(x_\Delta(s))^2(\varphi(\bar{x}_\Delta(s-\tau))^2 - \varphi(x_\Delta(s-\tau))^2) + \varphi(\bar{x}_\Delta(s-\tau))^2(g(\bar{x}_\Delta(s))^2 - g(x_\Delta(s))^2) \\ &\le 2\lambda(\mu(n))^2 L_n|\bar{x}_\Delta(s-\tau) - x_\Delta(s-\tau)| + 2\lambda^2\mu(n)K_n|\bar{x}_\Delta(s) - x_\Delta(s)|.\end{aligned}$$

So we get

$$\begin{aligned}&H_x(x_\Delta(s))f_\Delta(\bar{x}_\Delta(s)) + \frac{1}{2}H_{xx}(x_\Delta(s))\varphi(\bar{x}_\Delta(s-\tau))^2 g_\Delta(\bar{x}_\Delta(s))^2 \\ &\le K_3 + \lambda(\mu(n))^2 L_n H_{xx}(x_\Delta(s))|\bar{x}_\Delta(s-\tau) - x_\Delta(s-\tau)| \\ &+ \Big(K_n H_x(x_\Delta(s)) + \lambda^2 K_n\mu(n)H_{xx}(x_\Delta(s))\Big)|\bar{x}_\Delta(s) - x_\Delta(s)| \\ &\le K_3 + \zeta_n|\bar{x}_\Delta(s-\tau) - x_\Delta(s-\tau)| + \zeta_n^*|\bar{x}_\Delta(s) - x_\Delta(s)|,\end{aligned}$$

where

$$\zeta_n = \max_{1/n \le x \le n} \Big[\lambda(\mu(n))^2 L_n H_{xx}(x)\Big]$$

and

$$\zeta_n^* = \max_{1/n \le x \le n} \Big[K_n H_x(x) + \lambda^2 K_n \mu(n) H_{xx}(x)\Big].$$

We now have

$$\begin{aligned}
\mathbb{E}(H(x_\Delta(t \wedge \rho_n))) &\le H(\xi(0)) + \mathbb{E}\int_0^{t\wedge\rho_n} (K_3 + \zeta_n|\bar{x}_\Delta(s-\tau) - x_\Delta(s-\tau)| \\
&\quad + \zeta_n^*|\bar{x}_\Delta(s) - x_\Delta(s)|)ds \\
&\le H(\xi(0)) + K_3T + \zeta_n\mathbb{E}\int_{-\tau}^0 |\xi([s/\Delta]\Delta) - \xi(s)|ds \\
&\quad + (\zeta_n + \zeta_n^*)\int_0^T \mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|ds \\
&\le H(\xi(0)) + K_3T + \zeta_n K_4 \Delta^\gamma \tau \\
&\quad + (\zeta_n + \zeta_n^*)\int_0^T (\mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^p)^{1/p} ds.
\end{aligned}$$

By Lemma 3.5.1 and (3.23), we obtain

$$\mathbb{E}(H(x_\Delta(t \wedge \rho_n))) \le H(\xi(0)) + K_3T + \zeta_n K_4 \Delta^\gamma \tau + (\zeta_n + \zeta_n^*)TC_p^{1/p}\Delta^{1/4}.$$

Therefore

$$\mathbb{P}(\rho_n \le T) \le \frac{H(\xi(0)) + K_3T + \zeta_n K_4 \Delta^\gamma \tau + (\zeta_n + \zeta_n^*)TC_p^{1/p}\Delta^{1/4}}{H(1/n) \wedge H(n)}. \tag{3.37}$$

For $\varepsilon \in (0,1)$, we may choose sufficiently large $n$ such that

$$\frac{H(\xi(0)) + K_3T}{H(1/n) \wedge H(n)} \le \frac{\varepsilon}{2} \tag{3.38}$$

and sufficiently small step size $\Delta \in (0, \Delta_1]$ such that

$$\frac{\zeta_n K_4 \Delta^\gamma \tau + (\zeta_n + \zeta_n^*) T C_p^{1/p} \Delta^{1/4}}{H(1/n) \wedge H(n)} \leq \frac{\varepsilon}{2}. \tag{3.39}$$

Combining (3.38) and (3.39), we get the required assertion. □

To establish the strong convergence of the truncated EM scheme, we first define the stopping time

$$\upsilon_n = \tau_n \wedge \rho_n, \tag{3.40}$$

where $\tau_n$ and $\rho_n$ are (3.9) and (3.34) respectively.

**Lemma 3.5.5.** *Let Assumptions 3.3.1, 3.4.1 and 3.5.3 hold. Then, for any $p \geq 2$, $T > 0$, $\Delta \in (0, \Delta^*]$ and sufficiently large $n$*

$$\mathbb{E}\Big(\sup_{0 \leq t \leq T} |x_\Delta(t \wedge \upsilon_n) - x(t \wedge \upsilon_n)|^p\Big) \leq \boldsymbol{C} \Delta^{p(1/4 \wedge \gamma)} \tag{3.41}$$

*and*

$$\lim_{\Delta \to 0} \mathbb{E}\Big(\sup_{0 \leq t \leq T} |x_\Delta(t \wedge \upsilon_n) - x(t \wedge \upsilon_n)|^p\Big) = 0 \tag{3.42}$$

*where $\boldsymbol{C}$ is a constant independent of $\Delta$.*

*Proof.* It follows from (3.2) and (3.29) that

$$[x_\Delta(t \wedge \upsilon_n) - x(t \wedge \upsilon_n)] = \int_0^{t \wedge \upsilon_n} [f_\Delta(\bar{x}_\Delta(s)) - f(x(s))] ds$$
$$+ \int_0^{t \wedge \upsilon_n} [\varphi(\bar{x}_\Delta(s-\tau)) g_\Delta(\bar{x}_\Delta(s)) - \varphi(x(s-\tau)) g(x(s))] dB(s).$$

We now apply elementary inequality to have

$$|x_\Delta(t \wedge \upsilon_n) - x(t \wedge \upsilon_n)|^p \leq 2^{p-1} \Big( \Big| \int_0^{t \wedge \upsilon_n} [f_\Delta(\bar{x}_\Delta(s)) - f(x(x))] ds \Big|^p$$
$$+ \Big| \int_0^{t \wedge \upsilon_n} [\varphi(\bar{x}_\Delta(s-\tau)) g_\Delta(\bar{x}_\Delta(s)) - \varphi(x(s-\tau)) g(x(s))] dB(s) \Big|^p \Big).$$

So for $t_1 \in [0,T]$, we obtain

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge \upsilon_n)-x(t\wedge \upsilon_n)|^p\Big)\le 2^{p-1}\Big(\mathbb{E}\Big|\int_0^{t_1\wedge \upsilon_n}[f_\Delta(\bar{x}_\Delta(s))-f(x(s))]ds\Big|^p$$
$$+\,\mathbb{E}(\sup_{0\le t\le t_1}\Big|\int_0^{t_1\wedge \upsilon_n}[\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))]dB(s)\Big|^p)\Big).$$

By the Hölder inequality,

$$\mathbb{E}\Big(|\int_0^{t_1\wedge \upsilon_n}[f_\Delta(\bar{x}_\Delta(s))-f(x(s))]ds|^p\Big)\le T^{p-1}\mathbb{E}\Big(\int_0^{t_1\wedge \upsilon_n}|f_\Delta(\bar{x}_\Delta(s))-f(x(s))|^p ds\Big).$$

Also by the Burkholder-Davis-Gundy inequality, we obtain

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}\Big|\int_0^{t_1\wedge \upsilon_n}(\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s)))dB(s)\Big|^p\Big)$$
$$\le T^{\frac{p-2}{2}}C(p)\mathbb{E}\Big(\int_0^{t_1\wedge \upsilon_n}|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p ds\Big),$$

where $C(p)$ is a constant. We now have

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge \upsilon_n)-x(t\wedge \upsilon_n)|^p\Big)\le 2^{p-1}\Big(T^{p-1}\mathbb{E}\int_0^{t_1\wedge \upsilon_n}|f_\Delta(\bar{x}_\Delta(s))-f(x(s))|^p ds$$
$$+\,T^{\frac{p-2}{2}}C(p)\mathbb{E}\int_0^{t_1\wedge \upsilon_n}|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p ds\Big).$$

Meanwhile

$$\mathbb{E}\int_0^{t_1\wedge \upsilon_n}(|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p)ds$$
$$=\mathbb{E}\int_0^{t_1\wedge \upsilon_n}(|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g_\Delta(\bar{x}_\Delta(s))+\varphi(x(s-\tau))g_\Delta(\bar{x}_\Delta(s))$$
$$-\,\varphi(x(s-\tau))g(x(s))|^p)ds.$$

By elementary inequality,

$$\begin{aligned}
&\mathbb{E}\int_0^{t_1\wedge \upsilon_n}(|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p)ds \\
&\le 2^{p-1}\mathbb{E}\int_0^{t_1\wedge \upsilon_n}(|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g_\Delta(\bar{x}_\Delta(s))|^p \\
&+|\varphi(x(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p)ds \\
&\le 2^{p-1}\mathbb{E}\int_0^{t_1\wedge \upsilon_n} g_\Delta(\bar{x}_\Delta(s))^p|\varphi(\bar{x}_\Delta(s-\tau))-\varphi(x(s-\tau))|^p \\
&+\varphi(x(s-\tau))^p|g_\Delta(\bar{x}_\Delta(s))-g(x(s))|^p)ds.
\end{aligned}$$

By Assumption 3.3.1, we get

$$\begin{aligned}
&\mathbb{E}\int_0^{t_1\wedge \upsilon_n}(|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p ds) \\
&\le 2^{p-1}\mathbb{E}\int_0^{t_1\wedge \upsilon_n} g_\Delta(\bar{x}_\Delta(s))^p|\varphi(\bar{x}_\Delta(s-\tau))-\varphi(x(s-\tau))|^p ds \\
&+2^{p-1}\lambda^p\mathbb{E}\int_0^{t_1\wedge \upsilon_n}|g_\Delta(\bar{x}_\Delta(s))-g(x(s))|^p ds.
\end{aligned}$$

Moreover, by (3.22), we note $|g_\Delta(\bar{x}_\Delta(s))|\le \mu(n)$ for any $\bar{x}_\Delta(s)\in[1/n,n]$. Hence,

$$\begin{aligned}
&\mathbb{E}\int_0^{t_1\wedge \upsilon_n}(|\varphi(\bar{x}_\Delta(s-\tau))g_\Delta(\bar{x}_\Delta(s))-\varphi(x(s-\tau))g(x(s))|^p ds) \\
&\le 2^{p-1}(\mu(n))^p\mathbb{E}\int_0^{t_1\wedge \upsilon_n}|\varphi(\bar{x}_\Delta(s-\tau))-\varphi(x(s-\tau))|^p ds \\
&+2^{p-1}\lambda^p\mathbb{E}\int_0^{t_1\wedge \upsilon_n}|g_\Delta(\bar{x}_\Delta(s))-g(x(s))|^p ds.
\end{aligned}$$

We note from Assumption 3.4.1 that

$$|\varphi(\bar{x}_\Delta(s-\tau))-\varphi(x(s-\tau))|^p\le L_n^p|\bar{x}_\Delta(s-\tau)-x(s-\tau)|^p$$

for $s \in [0, t_1 \wedge \upsilon_n]$. So by Assumption 3.5.3, we get

$$\begin{aligned}
&\mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\varphi(\bar{x}_\Delta(s-\tau)) - \varphi(x(s-\tau))|^p ds \\
&\le L_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s-\tau) - x(s-\tau)|^p ds \\
&\le L_n^p \mathbb{E}\int_{-\tau}^{0} |\xi([s/\Delta]\Delta) - \xi(s)|^p ds + L_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds \\
&\le L_n^p K_4^p \Delta^{p\gamma}\tau + L_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds.
\end{aligned}$$

We now have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1} |x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p\Big) \le 2^{p-1}T^{p-1}\mathbb{E}\int_0^{t_1\wedge\upsilon_n} |f_\Delta(\bar{x}_\Delta(s)) - f(x(s)|^p ds \\
&+ 4^{p-1}T^{\frac{p-2}{2}}(\mu(n))^p C(p) L_n^p K_4^p \Delta^{p\gamma}\tau \\
&+ 4^{p-1}T^{\frac{p-2}{2}}(\mu(n))^p C(p) L_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds \\
&+ 4^{p-1}T^{\frac{p-2}{2}}\lambda^p C(p)\mathbb{E}\int_0^{t_1\wedge\upsilon_n} |g_\Delta(\bar{x}_\Delta(s)) - g(x(s))|^p ds.
\end{aligned}$$

We note from the definition of the truncated functions $f_\Delta$ and $g_\Delta$ that

$$f_\Delta(\bar{x}_\Delta(s)) = f(\bar{x}_\Delta(s)) \text{ and } g_\Delta(\bar{x}_\Delta(s)) = g(\bar{x}_\Delta(s))$$

for $s \in [0, t_1 \wedge \upsilon_n]$. Hence by Lemma 3.4.2, we have

$$|f(\bar{x}_\Delta(s)) - f(x(s))|^p \vee |g(\bar{x}_\Delta(s)) - g(x(s))|^p \le K_n^p |\bar{x}_\Delta(s) - x(s)|^p$$

for $s \in [0, t_1 \wedge \upsilon_n]$. We now get

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1} |x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p\Big) \\
&\le 4^{p-1}T^{\frac{p-2}{2}} C(p) L_n^p K_4^p (\mu(n))^p \Delta^{p\gamma}\tau + 2^{p-1}T^{p-1}K_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds
\end{aligned}$$

$$+ 4^{p-1}T^{\frac{p-2}{2}}(\mu(n))^p C(p)L_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds$$
$$+ 4^{p-1}T^{\frac{p-2}{2}}\lambda^p C(p)K_n^p \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds.$$

This implies

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p\Big) \le 4^{p-1}T^{\frac{p-2}{2}}C(p)L_n^p K_4^p(\mu(n))^p\Delta^{p\gamma}\tau$$
$$+ 4^{p-1}\Big(T^{\frac{p-2}{2}}(\mu(n))^p C(p)L_n^p + T^{\frac{p-2}{2}}\lambda^p C(p)K_n^p + 2^{1-p}T^{p-1}K_n^p\Big)$$
$$\times \mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s)|^p ds.$$

By elementary inequality, we have

$$\mathbb{E}\int_0^{t_1\wedge\upsilon_n} |\bar{x}_\Delta(s) - x(s) + x_\Delta(s) - x_\Delta(s)|^p ds$$
$$\le 2^{p-1}\mathbb{E}\Big(\int_0^{t_1\wedge\upsilon_n}(|\bar{x}_\Delta(s) - x_\Delta(s)|^p + |x_\Delta(s) - x(s)|^p)ds\Big)$$
$$\le 2^{p-1}\Big(\int_0^T \mathbb{E}|\bar{x}_\Delta(s) - x_\Delta(s)|^p ds + \mathbb{E}\int_0^{t_1}\sup_{0\le t\le s}|x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p ds\Big).$$

This yields

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p\Big)$$
$$\le 4^{p-1}T^{\frac{p-2}{2}}C(p)L_n^p K_4^p\tau(\mu(n))^p\Delta^{p\gamma}$$
$$+ 8^{p-1}\Big(T^{\frac{p-2}{2}}(\mu(n))^p C(p)L_n^p + T^{\frac{p-2}{2}}\lambda^p C(p)K_n^p + 2^{1-p}T^{p-1}K_n^p\Big)$$
$$\times\Big(\int_0^T \mathbb{E}|\bar{x}_\Delta(s) - x_\Delta(s)|^p ds + \mathbb{E}\int_0^{t_1}\sup_{0\le t\le s}|x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p ds\Big).$$

By Lemma 3.5.1, we get

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\upsilon_n) - x(t\wedge\upsilon_n)|^p\Big)$$

$$
\begin{aligned}
&\leq 4^{p-1}T^{\frac{p-2}{2}}C(p)L_n^pK_4^p\tau(\mu(n))^p\Delta^{p\gamma}\\
&+8^{p-1}C_p\Big(T^{\frac{p}{2}}(\mu(n))^pC(p)L_n^p+T^{\frac{p}{2}}\lambda^pC(p)K_n^p+2^{1-p}T^pK_n^p\Big)\Delta^{p/4}\\
&+8^{p-1}\Big(T^{\frac{p-2}{2}}(\mu(n))^pC(p)L_n^p+T^{\frac{p-2}{2}}\lambda^pC(p)K_n^p+2^{1-p}T^{p-1}K_n^p\Big)\\
&\times\mathbb{E}\int_0^{t_1}\sup_{0\leq t\leq s}|x_\Delta(t\wedge\upsilon_n)-x(t\wedge\upsilon_n)|^pds.
\end{aligned}
$$

This also means

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\leq t\leq t_1}|x_\Delta(t\wedge\upsilon_n)-x(t\wedge\upsilon_n)|^p\Big)\leq(C_1(n,p,T)+C_2(n,p,T))\Delta^{p(1/4\wedge\gamma)}\\
&+C_3(n,p,T)\mathbb{E}\int_0^{t_1}\sup_{0\leq t\leq s}|x_\Delta(t\wedge\upsilon_n)-x(t\wedge\upsilon_n)|^pds,
\end{aligned}
$$

where

$$
\begin{aligned}
&C_1(n,p,T)=4^{p-1}T^{\frac{p-2}{2}}C(p)L_n^pK_4^p\tau(\mu(n))^p,\\
&C_2(n,p,T)=8^{p-1}C_p\Big(T^{\frac{p}{2}}(\mu(n))^pC(p)L_n^p+T^{\frac{p}{2}}\lambda^pC(p)K_n^p+2^{1-p}T^pK_n^p\Big)
\end{aligned}
$$

and

$$
C_3(n,p,T)=8^{p-1}\Big(T^{\frac{p-2}{2}}(\mu(n))^pC(p)L_n^p+T^{\frac{p-2}{2}}\lambda^pC(p)K_n^p+2^{1-p}T^{p-1}K_n^p\Big).
$$

By the Grownwall inequality, we arrive at

$$
\mathbb{E}\Big(\sup_{0\leq t\leq T}|x_\Delta(t\wedge\upsilon_n)-x(t\wedge\upsilon_n)|^p\Big)\leq\mathbf{C}\Delta^{p(1/4\wedge\gamma)}
$$

as the required assertion, where $\mathbf{C}=(C_1(n,p,T)+C_2(n,p,T))e^{C_3(n,p,T)}$. Moreover, we obtain (3.42) by letting $\Delta\to 0$. □

**Theorem 3.5.6.** *Let Assumptions 3.3.1, 3.3.2, 3.4.1 and 3.5.3 hold. Then, for any $p\geq 2$, we get*

$$
\lim_{\Delta\to 0}\mathbb{E}\Big(\sup_{0\leq t\leq T}|x_\Delta(t)-x(t)|^p\Big)=0. \tag{3.43}
$$

*and consequently*

$$\lim_{\Delta\to 0}\mathbb{E}\Big(\sup_{0\le t\le T}|\bar{x}_\Delta(t)-x(t)|^p\Big)=0. \tag{3.44}$$

*Proof.* Let $\upsilon_n$ be the same as before. Set

$$e_\Delta(t)=x_\Delta(t)-x(t).$$

Clearly

$$\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p\Big)=\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\tau_n>T\text{ and }\rho_n>T\}}\Big)$$
$$+\,\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\tau_n\le T\text{ or }\rho_n\le T\}}\Big). \tag{3.45}$$

For any arbitrary $\varrho>0$, the Young inequality gives us

$$\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p\Big)1_{\{\tau_n\le T\text{ or }\rho_n\le T\}}\le\frac{\varrho}{2}\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^{2p}\Big)+\frac{1}{2\varrho}\mathbb{P}(\tau_n\le T\text{ or }\rho_n\le T).$$

Consequently,

$$\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p\Big)\le\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\tau_n>T\text{ and }\rho_n>T\}}\Big)+\frac{\varrho}{2}\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^{2p}\Big)$$
$$+\frac{1}{2\varrho}\mathbb{P}(\tau_n\le T\text{ or }\rho_n\le T). \tag{3.46}$$

By elementary inequality, we can derive to obtain

$$\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^{2p}\Big)\le 2^p\mathbb{E}\Big(\sup_{0\le t\le T}(|x_\Delta(t)|^{2p}+|x(t)|^{2p})\Big)$$
$$\le 2^{2p}\mathbb{E}\Big(\sup_{0\le t\le T}(|x_\Delta(t)|^p)\vee\sup_{0\le t\le T}(|x(t)|^p)\Big)^2.$$

So by Lemmas 3.3.4 and 3.5.2,

$$\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^{2p}\Big)\le 2^{2p}(C_1\vee C_4)^2. \tag{3.47}$$

Moreover,

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\tau_n>T \text{ and } \rho_n>T\}}\Big)\\
&\le \mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\upsilon_n>T\}}\Big)\\
&\le \mathbb{E}\Big(\sup_{0\le t\le T}|x_\Delta(t\wedge\upsilon_n)-x(t\wedge\upsilon_n)|^p\Big).
\end{aligned}$$

So by Lemma 3.5.5,

$$\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^q 1_{\{\tau_n>T \text{ and } \rho_n>T\}}\Big) \le \mathbf{C}\Delta^{p(1/4\wedge\gamma)}. \tag{3.48}$$

Also,

$$\mathbb{P}(\tau_n\le T \text{ or } \rho_n\le T)\le \mathbb{P}(\tau_n\le T)+\mathbb{P}(\rho_n\le T). \tag{3.49}$$

Substituting the inequalities (3.47), (3.48) and (3.49) into (3.46), we obtain

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^q\Big) \le{}& \frac{2^p(C_1\vee C_4)^2\varrho}{2}+\mathbf{C}\Delta^{p(1/4\wedge\gamma)}\\
&+\frac{1}{2\varrho}\mathbb{P}(\tau_n\le T)+\frac{1}{2\varrho}\mathbb{P}(\rho_n\le T).
\end{aligned}$$

For any given $\varepsilon\in(0,1)$, we may choose $\varrho$ such that

$$\frac{2^p(C_1\vee C_4)^2\varrho}{2}\le\frac{\varepsilon}{4}. \tag{3.50}$$

By Theorem 3.3.3 and Lemma 3.5.4, for any given $\varepsilon\in(0,1)$ , there exists $n_o$ such that for $n\ge n_o$ we may choose $\varrho$ to have

$$\frac{1}{2\varrho}\mathbb{P}(\tau_n\le T)\le\frac{\varepsilon}{4} \tag{3.51}$$

and choose $n(\varepsilon) \le n_o$ such that for $\Delta \in (0, \Delta_1]$

$$\frac{1}{2\varrho}\mathbb{P}(\rho_n \le T) \le \frac{\varepsilon}{4}. \tag{3.52}$$

Lastly, we may select $\Delta \in (0, \Delta_1]$ sufficiently small for $\varepsilon \in (0, 1)$ such that

$$\mathbf{C}\Delta^{p(1/4\wedge\gamma)} \le \frac{\varepsilon}{4}. \tag{3.53}$$

Combining (3.50), (3.51), (3.52) and (3.53), we get

$$\mathbb{E}\Big(\sup_{0\le t\le T} |x_\Delta(t) - x(t)|^p\Big) \le \varepsilon$$

as the required assertion. Moreover, by Lemma 3.5.1, we obtain (3.44). □

## 3.6 Numerical experiments

To illustrate efficiency of the proposed truncated EM scheme for SDDE (3.1), we perform two numerical examples with different Ait-Sahalia-type models. In the first numerical example, we implement Ait-Sahalia-type model with $\alpha_{-1}x(t)^{-1}$ term in the drift and delayed volatility function. In the second numerical example, we implement and perform comparative assessment of the delayed Ait-Sahalia-type model without $\alpha_{-1}x(t)^{-1}$ term in the drift using both truncated EM (TEM) and backward EM (BEM) schemes. This becomes necessary because it is unknown if the backward EM scheme could cope with $\alpha_{-1}x(t)^{-1}$ term at the origin. We would like to point out that we consider this case and use BEM scheme in the numerical study because the only well-known available literature for one half strong convergent approximation of Ait-Sahalia-type model focuses on the BEM method (see e.g [31]). There is so far no relevant literature devoted to strong convergent approximation of Ait-Sahalia-type model with $\alpha_{-1}x(t)^{-1}$ term and delayed volatility function.

### 3.6.1 Numerical example I

In this numerical illustration, we consider the following delayed Ait-Sahalia-type model

$$dx(t) = (0.1x(t)^{-1} - 0.3 + x(t) - 0.5x(t)^3)dt + \varphi(x(t-1))x(t)^{3/2}dB(t), \qquad (3.54)$$

with initial data $\xi(t) = 0.2$ and $\varphi(y)$ is defined by

$$\varphi(y) = \begin{cases} \frac{1}{2}\frac{(1+(e^y - e^{-y}))}{(e^y + e^{-y})}, & \text{if } y \geq 0 \\ \frac{1}{4}, & \text{Otherwise.} \end{cases} \qquad (3.55)$$

Note (3.55) is a special type of sigmoid function. Naturally, sigmoid functions like (3.55) are bounded, real-valued functions and hence fulfil Assumption 3.3.1. Moreover, parameterising sigmoid-based functions in financial models on past data are observed to capture volatility skews and smiles (see, e.g., [32]). Do also note the drift and diffusion coefficient terms of (3.54) satisfy

$$\sup_{1/u \leq x \leq u} (|f(x)| \vee g(x)) \leq 1.9u^3, \quad u \geq 1.$$

This means we can have $\mu(u) = 1.9u^2$ with inverse $\mu^{-1}(u) = (u/1.9)^{1/3}$. If we define $\Delta = 10^{-2}$ and $h(\Delta) = \Delta^{-2/3}$, then $\mu^{-1}(h(\Delta)) = (\Delta^{-2/3}/1.9)^{1/3}$ and $1/\mu^{-1}(h(\Delta)) = (\Delta^{-2/3}/1.9)^{-1/2}$. Displayed in Figure 3.1 is a Monte Carlo simulated sample path of $x(t)$ with step size $10^{-2}$ using the TEM scheme.

### 3.6.2 Numerical example II

In this subsection, we assess the performance of TEM scheme with BEM scheme. We already noted there exists no relevant literature on strong convergent approximation of SDDE (3.56). Hence, we have to fall on the BEM method which has one half strong order approximation of Ait-Sahalia-type model without the delayed volatility function. Consider the following delayed Ait-Sahalia-type model

$$dx(t) = (0.2 + 0.3x(t) - 0.5x(t)^2)dt + \varphi(x(t-1))x(t)^{4/3}dB(t), \tag{3.56}$$

with initial data $\xi(t) = 0.2$ and the same volatility function $\varphi(\cdot)$ in (3.55). Clearly, we have $\mu(u) = u^2$ with inverse $\mu^{-1}(u) = u^{1/2}$. Using TEM and BEM schemes with step size $10^{-2}$, we obtain Monte Carlo simulated sample paths of $x(t)$ in Figure 3.2. We notice that both simulated sample paths are almost the same. Figure 3.3 depicts the log-log plot of the strong errors between TEM and BEM numerical solutions based on step sizes $10^{-3}$, $10^{-4}$, $10^{-5}$ and $10^{-6}$. For the purpose of comparison, we also plotted the reference line with slope 1.0. We can see the strong errors between TEM and BEM numerical solutions have order 1.0 although this has not been proved theoretically.

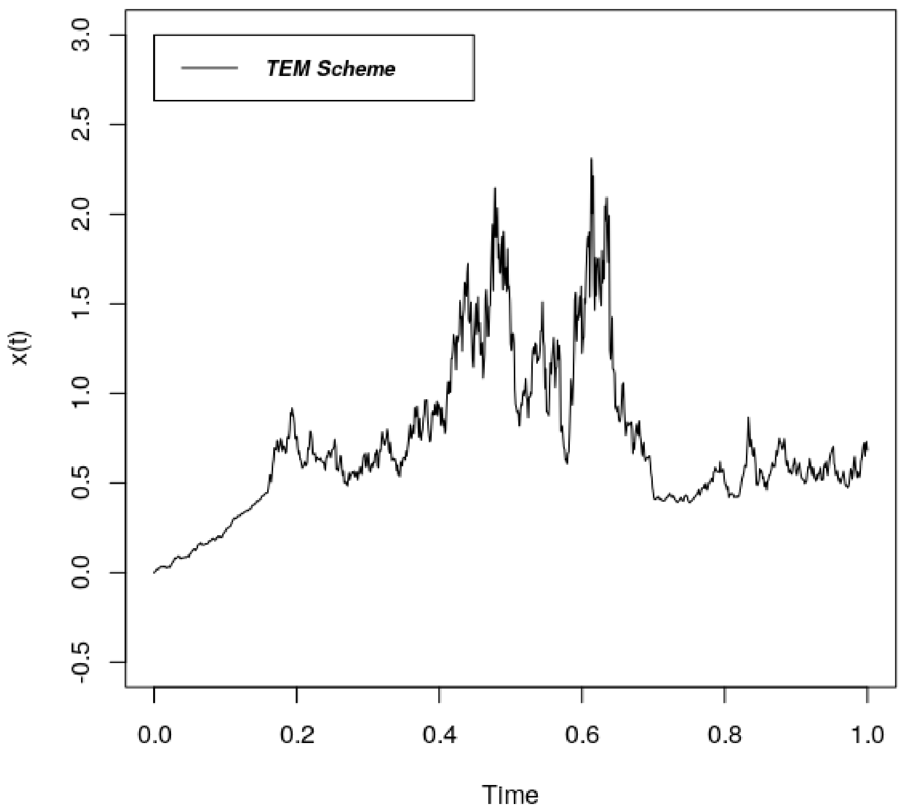


Figure 3.1: Simulated sample path of $x(t)$ when $\Delta = 0.01$

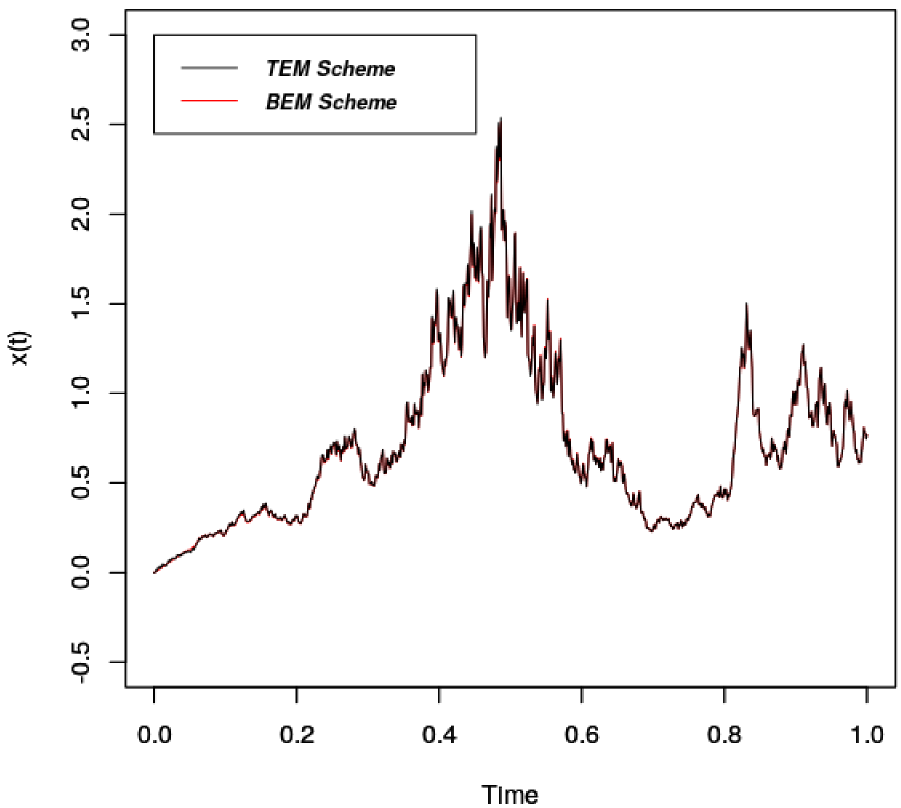


Figure 3.2: Convergence of TEM and BEM solutions when $\Delta = 0.01$

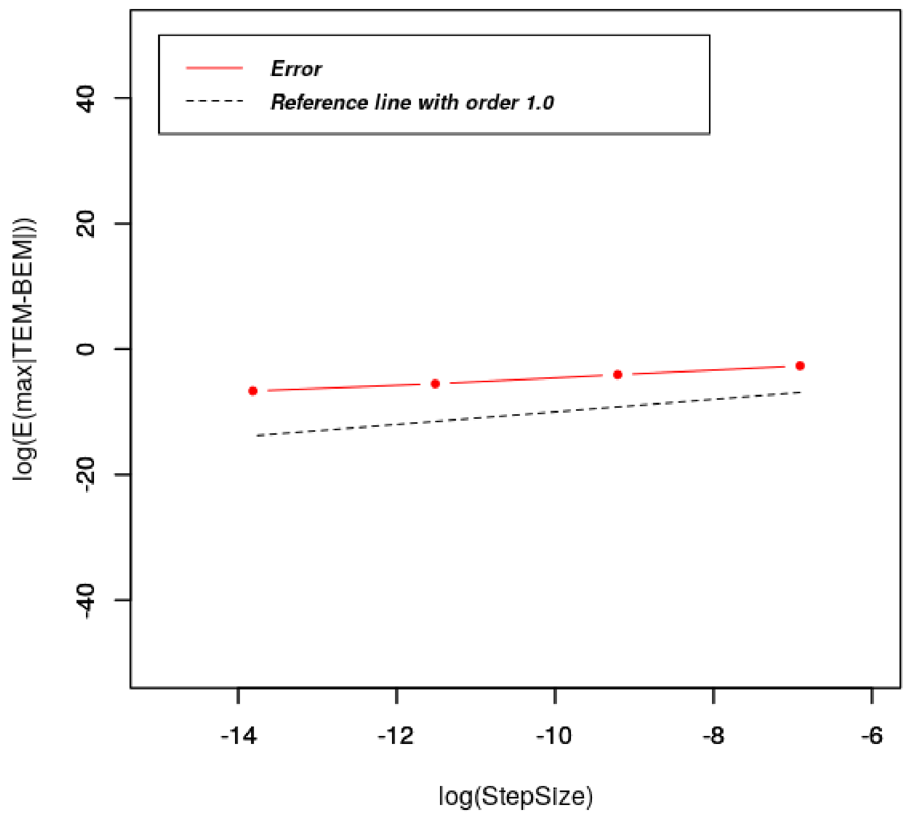


Figure 3.3: Strong errors between TEM and BEM schemes

## 3.7 Summary

As supported by empirical findings, stochastic volatility models with inherent features of past dependency are suitable models for describing convex phenomena of implied volatility against market anomalies. This motivated the need to replace the constant volatility of the Ait-Sahalia-type interest rate model with a delayed volatility function. Then we discussed analytical properties such as existence of pathwise unique positive global solution and boundedness of moments of the exact solution.

We moved on to construct a new implementable truncated EM scheme which could cope around the origin with the inverse function in the drift of the proposed model. We also proved numerical properties such as boundedness of moment in the strong sense and established finite time strong convergence of the truncated EM approximate solutions to the exact solution under the local Lipschitz condition plus the Khasminskii-type condition. The strong convergence result implies that in practice, the truncated EM approximate solutions can be used to compute some debt and path-dependent financial products. We obtained $\mathbf{C}\Delta^{p(1/4\wedge\gamma)}$ as the strong pathwise error. Finally we implemented some numerical examples to validate the theoretical results.

# Chapter 4

# Numerical approximation of Poisson-jump Ait-Sahalia-type interest rate model with delay

## 4.1 Introduction

We study analytical properties of the exact solution to the generalised Poisson-jump Ait-Sahalia-type interest rate model with delay in this chapter. Since this model does not have explicit solution, we employ several new truncated EM techniques to investigate finite time strong convergence theory of the numerical solutions under the local Lipschitz condition plus the Khasminskii-type condition.

The rest of the chapter is organised as follows: In Section 4.2, we present the Poisson-jump Ait-Sahalia-type interest rate model with delay. We study the existence of a unique global solution to the proposed model and show that the solution will always be positive in Section 4.3. We also establish boundedness of moments of the exact solution in this section. In Section 4.4, we introduce the truncated EM approximation scheme for the proposed model. Section 4.5 is entirely devoted to exploration of numerical properties of the truncated EM scheme. These include boundedness of moments and $L^p(p \geq 2)$ finite time strong convergence of the truncated EM approximate solutions to the exact solution. In Section 4.6, we perform some numerical illustrations to support the theoretical results. We briefly summarise the findings in the last section.

## 4.2 The Poisson-jump Ait-Sahalia-type interest rate model with delay

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 us extend SDE (1.3) to incorporate delayed volatility function and Poisson-driven jump described by

$$\begin{aligned} dx(t) &= (\alpha_{-1}x(t^-)^{-1} - \alpha_0 + \alpha_1 x(t^-) - \alpha_2 x(t^-)^{\rho})dt + \varphi(x((t-\tau)^-))x(t^-)^{\theta}dB(t) \\ &\quad + \alpha_3 x(t^-)dN(t) \end{aligned} \tag{4.1}$$

on $t \geq -\tau$ with initial data $x(t) = \xi(t)$ for $t \in [-\tau, 0]$. Here $x(t^-) = \lim_{s\to t^-} x(s)$, $x((t-\tau)^-)$ denotes delay in $x(t^-)$, $\varphi(\cdot)$ depends on $x((t-\tau)^-)$ with $\tau > 0$. The delayed volatility function and Poisson-driven jump may, for instance, explain joint effects of volatility 'skews' and 'smiles', and tail distribution of interest rates which pervade most financial markets. The reader is referred to [14] for relevant information about this extension.

Now let the following scalar dynamics

$$dx(t) = f(x(t^-))dt + \varphi(x((t-\tau)^-))g(x(t^-))dB(t) + h(x(t^-))dN(t), \tag{4.2}$$

$x(t) = \xi(t)$, on $t \in [-\tau, \infty)$, denote equation of SDDE (4.1) such that $f(x) = \alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^{\rho}$, $g(x) = x^{\theta}$ and $h(x) = \alpha_3 x$, $\forall x \in \mathbb{R}_+$, with $\varphi(y)$ defined in $C(\mathbb{R}_+; \mathbb{R}_+)$. Let $C^{2,1}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ be the family of all real-valued functions $Z(x,t)$ defined on $\mathbb{R} \times \mathbb{R}_+$ such that $Z(x,t)$ is twice continuously differentiable in $x$ and once in $t$. For each $Z \in C^{2,1}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$, define the jump-diffusion operator $LZ$ :

$\mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ by

$$LZ(x, y, t) = \ell(x, y, t) + \lambda(Z(x + h(x), t) - Z(x, t)), \tag{4.3}$$

for SDDE (4.2) associated with the $C^{2,1}$-function $Z$, where

$$\ell(x, y, t) = Z_t(x, t) + Z_x(x, t)f(x) + \frac{1}{2}Z_{xx}(x, t)\varphi(y)^2 g(x)^2, \tag{4.4}$$

$\ell Z : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$, is the diffusion operator. Here, $Z_t(x, t)$ and $Z_x(x, t)$ are first-order partial derivatives with respect to $t$ and $x$ respectively, and $Z_{xx}(x, t)$ is a second-order partial derivative with respect to $x$. With the jump-diffusion operator defined, the Itô formula then yields

$$\begin{aligned} dZ(x(t), t) = {} & LZ(x(t^-), x((t - \tau)^-), t)dt \\ & + \varphi(x((t - \tau)^-))Z_x(x(t^-), t)g(x(t^-))dB(t) \\ & + (Z(x(t^-) + h(x(t^-)), t) - Z(x(t^-), t))d\widetilde{N}(t) \end{aligned} \tag{4.5}$$

a.s. (e.g., see [47] for detailed coverage).

## 4.3 Analytical properties

In this section, we survey the analytical properties such as existence-and-uniqueness theorem and boundedness of moments of the exact solution to SDDE (4.2).

### 4.3.1 Existence and uniqueness of solution

Before we show existence of positive solution to SDDE (4.2), we are required to assume the volatility function $\varphi(\cdot)$ is locally Lipschitz continuous and bounded (see, e.g., [26] for detailed accounts of these conditions). The following conditions are thus sufficient to establish existence of a unique positive global or nonexplosive solution to SDDE (4.2).

**Assumption 4.3.1.** *The volatility function* $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ *of* SDDE (4.2) *is Borel-measurable and bounded by a positive constant* $\sigma$*, i.e.*

$$\varphi(y) \leq \sigma, \tag{4.6}$$

$\forall y \in \mathbb{R}_+$.

**Assumption 4.3.2.** *For any* $R > 0$*, there exists a constant* $L_R > 0$ *such that the volatility function* $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ *of* SDDE (4.2) *satisfies*

$$|\varphi(y) - \varphi(\bar{y})| \leq L_R|y - \bar{y}|, \tag{4.7}$$

$\forall y, \bar{y} \in [\frac{1}{R}, R]$.

**Assumption 4.3.3.** *The parameters of* SDDE (4.2) *satisfy*

$$1 + \rho > 2\theta, \quad \rho, \theta > 1. \tag{4.8}$$

The following theorem reveals the SDDE (4.2) admits a pathwise-unique positive global solution $x(t)$ on $t \in [-\tau, \infty)$. Since SDDE (4.2) describes interest rate dynamics, the solution will always remain nonnegative a.s.

**Theorem 4.3.4.** *Let Assumptions 4.3.1 and 4.3.3 hold. Then for any given initial data*

$$\{x(t) : -\tau \leq t \leq 0\} = \xi(t) \in C([-\tau, 0] : \mathbb{R}_+), \tag{4.9}$$

*there exists a unique global solution* $x(t)$ *to* SDDE (4.2) *on* $t \in [-\tau, \infty)$ *and* $x(t) > 0$ *a.s.*

*Proof.* Since the coefficient terms of SDDE (4.2) are locally Lipschitz continuous in $[-\tau, \infty)$, then there exists a unique positive maximal local solution $x(t) \in [-\tau, \tau_e)$ for any given initial data (4.9), where $\tau_e$ is the explosion time (e.g., see [13] and the classical methods in [17]). Let $n_0 > 0$ be sufficiently large such that

$$\frac{1}{n_0} < \min_{-\tau \leq t \leq 0} |\xi(t)| \leq \max_{-\tau \leq t \leq 0} |\xi(t)| < n_0.$$

For each integer $n \geq n_0$, define the stopping time

$$\tau_n = \inf\{t \in [0, \tau_e) : x(t) \notin (1/n, n)\}. \tag{4.10}$$

Obviously, $\tau_n$ is increasing as $n \to \infty$. Set $\tau_\infty = \lim_{n\to\infty} \tau_n$, whence $\tau_\infty \leq \tau_e$ a.s. In other words, we need to show that $\tau_\infty = \infty$ a.s. to complete the proof. For any $\beta \in (0, 1)$, define a $C^2$-function $Z : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$Z(x) = x^\beta - 1 - \beta\log(x). \tag{4.11}$$

Clearly $Z(x) \to \infty$ as $x \to \infty$ or $x \to 0$. By Assumption 4.3.1, we get from the operator in (4.3) that

$$\begin{aligned}
LZ(x, y) &\leq \ell Z(x, y) + \lambda\Big((x + \alpha_3 x)^\beta - 1 - \beta\log(x + \alpha_3 x) - (x^\beta - 1 - \beta\log(x))\Big) \\
&= \ell Z(x, y) + \lambda\Big(((x + \alpha_3 x)^\beta - x^\beta) - \beta\log(x(1 + \alpha_3)/x)\Big) \\
&= \ell Z(x, y) + \lambda((1 + \alpha_3)^\beta - 1)x^\beta - \lambda\beta\log(1 + \alpha_3),
\end{aligned}$$

where

$$\begin{aligned}
\ell Z(x, y) &= \beta(x^{\beta-1} - x^{-1})\big(\alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^\rho\big) + \frac{1}{2}(\beta(\beta-1)x^{\beta-2} + \beta x^{-2})\varphi(y)^2 x^{2\theta} \\
&\leq \alpha_{-1}\beta x^{\beta-2} - \alpha_0\beta x^{\beta-1} + \alpha_1\beta x^\beta - \alpha_2\beta x^{\rho+\beta-1} - \alpha_{-1}\beta x^{-2} + \alpha_0\beta x^{-1} \\
&\quad - \alpha_1\beta + \alpha_2\beta x^{\rho-1} + \frac{\sigma^2}{2}\beta(\beta-1)x^{\beta+2\theta-2} + \frac{\sigma^2}{2}\beta x^{2\theta-2}.
\end{aligned}$$

Since $\beta \in (0, 1)$ and by Assumption 4.3.3, we note $-\alpha_{-1}\beta x^{-2}$ leads and tends to $-\infty$ for small $x$ and for large $x$, $-\alpha_2\beta x^{\rho+\beta-1}$ leads and also tends to $-\infty$. Hence there exists a constant $K_0$ such that

$$LZ(x, y) \leq K_0. \tag{4.12}$$

So for $t_1 \in [0, \tau]$, we derive from the Itô formula

$$\mathbb{E}[Z(x(\tau_n \wedge t_1))] \leq Z(\xi(0)) + \int_0^{\tau_n \wedge t_1} K_0 dt,$$

$\forall n \geq n_0$. It then follows that

$$\mathbb{P}(\tau_n \leq \tau) \leq \frac{Z(\xi(0)) + K_0\tau}{Z(1/n) \wedge Z(n)}.$$

As $n \to \infty$, $\mathbb{P}(\tau_n \leq \tau) \to 0$. This implies $\tau_\infty > \tau$ a.s. Also for $t_1 \in [0, 2\tau]$, the Itô formula yields

$$\mathbb{E}[Z(x(\tau_n \wedge t_1))] \leq Z(\xi(0)) + \int_0^{\tau_n \wedge t_1} K_0 dt,$$

$\forall n \geq n_0$ and consequently,

$$\mathbb{P}(\tau_n \leq 2\tau) \leq \frac{Z(\xi(0)) + 2K_0\tau}{Z(1/n) \wedge Z(n)}.$$

As $n \to \infty$, we get $\tau_\infty > 2\tau$ a.s. Repeating this procedure for $t_1 \in [0, \infty)$, we obtain $\mathbb{P}(\tau_\infty \leq \infty) \to 0$ by letting $n \to \infty$. This means $\tau_\infty = \infty$ a.s. and hence $\tau_e = \infty$ a.s. The proof is now complete. □

### 4.3.2 Moment bounds

The following lemmas show the finite moments of the exact solution to SDDE (4.2).

**Lemma 4.3.5.** *Let Assumptions 4.3.1 and 4.3.3 hold. Then for any $p \geq 2$, there exists a constant $\rho_1$ such that the solution of* SDDE (4.2) *satisfies*

$$\sup_{0 \leq t < \infty} \left(\mathbb{E}|x(t)|^p\right) \leq \rho_1. \tag{4.13}$$

*Proof.* Define the stopping time for every sufficiently large integer $n$ by

$$\tau_n = \inf\{t \geq 0 : x(t) \notin (1/n, n)\}. \tag{4.14}$$

Define a function $Z \in C^{2,1}(\mathbb{R}_+ \times \mathbb{R}_+; \mathbb{R}_+)$ by $Z(x,t) = e^t x^p$ . By Assumption 4.3.1, the jump-diffusion operator in (4.3) gives us

$$\begin{aligned} LZ(x,y,t) &\leq \ell Z(x,y,t) + \lambda[e^t(x+\alpha_3 x)^p - e^t x^p] \\ &= \ell Z(x,y,t) + \lambda e^t x^p[(1+\alpha_3)^p - 1], \end{aligned}$$

where

$$\begin{aligned} \ell Z(x,y,t) &= e^t x^p + pe^t x^{p-1}\big(\alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^\rho\big) + \frac{1}{2}p(p-1)e^t x^{p-2}\varphi^2(y)x^{2\theta} \\ &\leq e^t\big[x^p + \alpha_{-1}px^{p-2} - \alpha_0 px^{p-1} + \alpha_1 px^p - \alpha_2 px^{\rho+p-1} + \frac{p(p-1)}{2}\sigma^2 x^{2\theta+p-2})\big]. \end{aligned}$$

By Assumption 4.3.3, $-p\alpha_2 x^{\rho+p-1}$ dominates and tends to $-\infty$ for large $x$. Hence we can find a constant $K_1$ such that

$$LZ(x,y,t) \leq K_1 e^t.$$

The Itô formula gives us

$$\mathbb{E}[e^{t\wedge\tau_n}|x(t\wedge\tau_n)|^p] \leq |\xi(0)|^p + K_1 e^t.$$

Applying the Fatou lemma and letting $n \to \infty$ yields

$$\mathbb{E}|x(t)|^p < e^{-t}|\xi(0)|^p + K_1$$

and consequently,

$$\sup_{0\leq t<\infty} (\mathbb{E}|x(t)|^p) \leq \rho_1.$$

as the required assertion in (4.13). $\square$

**Lemma 4.3.6.** *Let Assumptions 4.3.1 and 4.3.3 hold. For any $p > 2 \vee (\rho - 1)$, there exists a constant $\rho_2$ such that the solution of* SDDE (4.2) *satisfies*

$$\sup_{0 \le t < \infty} \left( \mathbb{E} |\frac{1}{x(t)}|^p \right) \le \rho_2. \tag{4.15}$$

*Proof.* Let $\tau_n$ be the same as in (4.14). By applying (4.3) to $Z(x,t) = e^t/x^p$, we compute

$$\begin{aligned} LZ(x,y,t) &\le \ell Z(x,y,t) + \lambda[e^t(x + \alpha_3 x)^{-p} - e^t x^{-p}] \\ &= \ell Z(x,y,t) + \lambda e^t x^{-p}[(1+\alpha_3)^{-p} - 1], \end{aligned}$$

where Assumption 4.3.1 has been used and here, we have

$$\begin{aligned} \ell Z(x,y,t) &= e^t x^{-p} - p e^t x^{-(p+1)}(\alpha_{-1} x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^\rho) \\ &+ \frac{1}{2} p(p+1) e^t x^{-(p+2)} \varphi(y)^2 x^{2\theta} \\ &\le e^t [x^{-p} - \alpha_{-1} p x^{-(p+2)} + \alpha_0 p x^{-(p+1)} - \alpha_1 p x^{-p} + \alpha_2 x^{\rho - p - 1} \\ &- \frac{p(p+1)}{2} \sigma^2 x^{2\theta - p - 2})]. \end{aligned}$$

By Assumption 4.3.3 and noting that $p > 2 \vee (\rho - 1)$, we observe $-\alpha_{-1} p x^{-(p+2)}$ leads and tends to $-\infty$ for small $x$ and for large $x$, $p\alpha_2 x^{\rho - p - 1}$ dominates and tends to 0. Hence there exists a constant $K_2$ such that

$$LZ(x,y,t,) \le K_2 e^t.$$

We can now use the Itô formula, apply Fatou lemma and let $n \to \infty$ to arrive at

$$\mathbb{E}|x(t)|^{-p} < e^{-t} |\xi(0)|^{-p} + K_2$$

and consequently the required assertion in (4.15). $\square$

## 4.4 The truncated EM method

In this section, we present the truncated EM scheme for numerical approximation of SDDE (4.2). Meanwhile, we need the following useful assumption on the initial data. This is needed to prove the results of this chapter.

**Assumption 4.4.1.** *There is a pair of constant $K_3 > 0$ and $\gamma \in (0, 1]$ such that for all $-\tau \le s \le t \le 0$, the initial data $\xi$ satisfies*

$$|\xi(t) - \xi(s)| \le K_3|t - s|^\gamma. \tag{4.16}$$

In the sequel, we also need the following lemmas.

**Lemma 4.4.2.** *For any $R > 0$, there exists a constant $K_R > 0$ such that the coefficient terms $f$, $g$ and $h$ of* SDDE (4.2) *satisfy*

$$|f(x) - f(\bar{x})| \vee |g(x) - g(\bar{x})| \vee |h(x) - h(\bar{x})| \le K_R|x - \bar{x}|, \tag{4.17}$$

$\forall x, \bar{x} \in [\frac{1}{R}, R]$.

**Lemma 4.4.3.** *Let Assumptions 4.3.1 and 4.3.3 hold. Then for any $p \ge 2$, there exists $K_4 > 0$ such that the drift and diffusion terms of* SDDE (4.2) *satisfy*

$$xf(x) + \frac{p-1}{2}|\varphi(y)g(x)|^2 \le K_4(1 + |x|^2), \tag{4.18}$$

$\forall x, y \in \mathbb{R}_+$, *where $K_4$ is a constant (see [39] for the proof).*

### 4.4.1 Numerical approximation

Before we proceed, let us extend the volatility function $\varphi(y)$ and the jump term $h(x)$ from $\mathbb{R}_+$ to $\mathbb{R}$ by setting $\varphi(y) = \varphi(0)$ and $h(x) = 0$ for $x < 0$. Apparently, Theorem 4.3.4 as well as conditions (4.6), (4.7), (4.17) and (4.18) are well maintained. Moreover, we do not need to truncate the jump term since it is of linear growth. To define the truncated EM scheme for SDDE (4.2), we first choose a strictly increasing continuous function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\mu(r) \to \infty$ as $r \to \infty$ and

$$\sup_{1/r \le x \le r} (|f(x)| \vee g(x)) \le \mu(r), \quad \forall r > 1. \tag{4.19}$$

Denote by $\mu^{-1}$ the inverse function of $\mu$. We define a strictly decreasing function $\pi : (0,1) \to \mathbb{R}_+$ such that

$$\lim_{\Delta \to 0} \pi(\Delta) = \infty \text{ and } \Delta^{1/4}\pi(\Delta) \le 1, \quad \forall \Delta \in (0,1]. \tag{4.20}$$

Find $\Delta^* \in (0,1)$ such that $\mu^{-1}(\pi(\Delta^*)) > 1$ and $f(x) > 0$ for $0 < x < \Delta^*$. For a given step size $\Delta \in (0, \Delta^*)$, let us define the truncated functions

$$f_\Delta(x) = f\Big(1/\mu^{-1}(\pi(\Delta)) \vee (x \wedge \mu^{-1}(\pi(\Delta)))\Big), \quad \forall x \in \mathbb{R}$$

and

$$g_\Delta(x) = \begin{cases} g\Big(x \wedge \mu^{-1}(\pi(\Delta))\Big), & \text{if } x \ge 0 \\ 0, & \text{if } x < 0. \end{cases}$$

So for $x \in [1/\mu^{-1}(\pi(\Delta)), \mu^{-1}(\pi(\Delta))]$, we have

$$\begin{aligned} |f_\Delta(x)| = |f(x)| &\le \max_{1/\mu^{-1}(\pi(\Delta)) \le w \le \mu^{-1}(\pi(\Delta))} |f(w)| \\ &\le \mu(\mu^{-1}(\pi(\Delta))) = \pi(\Delta) \end{aligned}$$

and

$$g_\Delta(x) \le \mu(\mu^{-1}(\pi(\Delta))) = \pi(\Delta).$$

We easily observe that

$$|f_\Delta(x)| \vee g_\Delta(x) \le \pi(\Delta), \quad \forall x \in \mathbb{R}. \tag{4.21}$$

That is, both truncated functions $f_\Delta$ and $g_\Delta$ are bounded although both $f$ and $g$ may not. The following lemma shows $f_\Delta$ and $g_\Delta$ maintain (4.18) nicely.

**Lemma 4.4.4.** *Let Assumptions 4.3.1 and 4.3.3 hold. Then, for all $\Delta \in (0, \Delta^*)$ and $p \geq 2$, the truncated functions satisfy*

$$xf_\Delta(x) + \frac{p-1}{2}|\varphi(y)g_\Delta(x)|^2 \leq K_5(1+|x|^2) \tag{4.22}$$

*$\forall x, y \in \mathbb{R}$, where $K_5$ is a constant independent of $\Delta$ (see [39] for the proof).*

From now on, let $T > 0$ be fixed arbitrarily and the step size $\Delta \in (0, \Delta^*]$ be a fraction of $\tau$. We define $\Delta = \tau/M$ for some positive integer $M$. Let now form the discrete-time truncated EM approximation of SDDE (4.2). Define $t_k = k\Delta$ for $k = -M, -(M-1), \cdots, 0, 1, 2, \cdots$. Set $X_\Delta(t_k) = \xi(t_k)$ for $k = -M, -(M-1), \cdots, 0$ and then compute

$$X_\Delta(t_{k+1}) = X_\Delta(t_k) + f_\Delta(X_\Delta(t_k))\Delta + \varphi(X_\Delta(t_{k-M}))g_\Delta(X_\Delta(t_k))\Delta B_k + h(X_\Delta(t_k))\Delta N_k \tag{4.23}$$

for $k = 0, 1, 2, \cdots$, where $\Delta B_k = B(t_{k+1}) - B(t_k)$ and $\Delta N_k = N(t_{k+1}) - N(t_k)$. Let now form two versions of the continuous-time truncated EM solutions. The first one is defined by

$$\bar{x}_\Delta(t) = \sum_{k=-M}^{\infty} X_\Delta(t_k) 1_{[t_k, t_{k+1})}(t). \tag{4.24}$$

This is the continuous-time step-process $\bar{x}_\Delta(t)$ on $t \in [-\tau, \infty]$, where $1_{[t_k,t_{k+1})}$ is the indicator function on $[t_k, t_{k+1})$. The other one is the continuous-time continuous process $x_\Delta(t)$ on $t \geq -\tau$ defined conveniently by setting $x_\Delta(t) = \xi(t)$ for $t \in [-\tau, 0]$ while for $t \geq 0$

$$\begin{aligned} x_\Delta(t) = \xi(0) &+ \int_0^t f_\Delta(\bar{x}_\Delta(s^-))ds + \int_0^t \varphi(\bar{x}_\Delta((s-\tau)^-))g_\Delta(\bar{x}_\Delta(s^-))dB(s) \\ &+ \int_0^t h(\bar{x}_\Delta(s^-))dN(s). \end{aligned} \tag{4.25}$$

Obviously $x_\Delta(t)$ is an Itô process on $t \ge 0$ satisfying Itô differential

$$dx_\Delta(t) = f_\Delta(\bar{x}_\Delta(t^-))dt + \varphi(\bar{x}_\Delta((t-\tau)^-))g_\Delta(\bar{x}_\Delta(t^-))dB(t) + h(\bar{x}_\Delta(t^-))dN(t). \quad (4.26)$$

For all $k = -M, -(M-1), \cdots$, it is useful to see that $x_\Delta(t_k) = \bar{x}_\Delta(t_k) = X_\Delta(t_k)$.

# 4.5 Numerical properties

In this section, we establish the finite moment and finite time strong convergence theory of the truncated EM solutions to SDDE (4.2).

## 4.5.1 Moment bounds

To upper bound the moment of the truncated EM solution, let first define

$$k(t) = [t/\Delta]\Delta,$$

for any $t \in [0, T]$, where $[t/\Delta]$ denotes the integer part of $t/\Delta$. The following lemma shows $x_\Delta(t)$ and $\bar{x}_\Delta(t)$ are close to each other in $L^p$.

**Lemma 4.5.1.** *Let Assumption 4.3.1 hold. Then for any fixed $\Delta \in (0, \Delta^*]$, we have*

$$\mathbb{E}\Big(|x_\Delta(t) - \bar{x}_\Delta(t)|^p \big| \mathcal{F}_{k(t)}\Big) \le \mathfrak{D}_1\Big(\Delta^{p/2}(\pi(\Delta))^p + \Delta\Big)|\bar{x}_\Delta(t)|^p, \quad p \in [2, \infty) \quad (4.27)$$

*and*

$$\mathbb{E}\Big(|x_\Delta(t) - \bar{x}_\Delta(t)|^p \big| \mathcal{F}_{k(t)}\Big) \le \mathfrak{D}_2\Big(\Delta^{p/2}(\pi(\Delta))^p\Big)|\bar{x}_\Delta(t)|^p, \quad p \in (0, 2), \quad (4.28)$$

*for all $t \ge 0$, where $\mathfrak{D}_1$ and $\mathfrak{D}_2$ denote positive generic constants which depend only on $p$ and may change between occurrences.*

*Proof.* Fix any $\Delta \in (0, \Delta^*)$ and $t \in [0, T]$. Then for $p \in [2, \infty)$, we derive

$$
\begin{aligned}
&\mathbb{E}\Big(|x_\Delta(t) - \bar{x}_\Delta(t)|^p \Big| \mathcal{F}_{k(t)}\Big) \\
&\le 3^{p-1}\Big(\mathbb{E}\big(|\int_{k(t)}^t f_\Delta(\bar{x}_\Delta(s))ds|^p \big| \mathcal{F}_{k(t)}\big) + \mathbb{E}\big(|\int_{k(t)}^t \varphi(\bar{x}_\Delta((s-\tau)))g_\Delta(\bar{x}_\Delta(s))dB(s)|^p \big| \mathcal{F}_{k(t)}\big) \\
&+ \mathbb{E}\big(|\int_{k(t)}^t h(\bar{x}_\Delta(s))dN(s)|^p \big| \mathcal{F}_{k(t)}\big)\Big) \\
&\le 3^{p-1}\Big(\Delta^{p-1}\mathbb{E}(\int_{k(t)}^t |f_\Delta(\bar{x}_\Delta(s))|^p ds \big| \mathcal{F}_{k(t)}) \\
&+ c(p)\Delta^{(p-2)/2}\mathbb{E}(\int_{k(t)}^t |\varphi(\bar{x}_\Delta((s-\tau)))g_\Delta(\bar{x}_\Delta(s))|^p ds \big| \mathcal{F}_{k(t)}) \\
&+ \mathbb{E}(|\int_{k(t)}^t h(\bar{x}_\Delta(s))dN(s)|^p \big| \mathcal{F}_{k(t)})\Big) \\
&\le 3^{p-1}\Big(\Delta^{p-1}\Delta(\pi(\Delta))^p + c(p)\Delta^{(p-2)/2}\Delta(\sigma\pi(\Delta))^p + \mathbb{E}(|\int_{k(t)}^t h(\bar{x}_\Delta(s))dN(s)|^p \big| \mathcal{F}_{k(t)})\Big),
\end{aligned}
$$

where Assumption 4.3.1 and (4.21) have been used and $c(p)$ depends on $p$. By the characteristic function's argument (see [45]), we have

$$\mathbb{E}|\Delta N_k|^p \le \bar{c}\Delta, \quad \forall \Delta \in (0, \Delta^*),$$

where $\bar{c}$ is a positive constant independent of $\Delta$. We now obtain

$$\mathbb{E}(|\int_{k(t)}^t h(\bar{x}_\Delta(s))dN(s)|^p \big| \mathcal{F}_{k(t)}) = |h(\bar{x}_\Delta(t))|^p \mathbb{E}|\Delta N_k|^p.$$

This implies

$$
\begin{aligned}
\mathbb{E}\Big(|x_\Delta(t) - \bar{x}_\Delta(t)|^p \Big| \mathcal{F}_{k(t)}\Big) \le 3^{p-1}\Big(&\Delta^{p-1}\Delta(\pi(\Delta))^p \\
&+ c(p)\Delta^{(p-2)/2}\Delta(\sigma\pi(\Delta))^p + |h(\bar{x}_\Delta(t))|^p \mathbb{E}|\Delta N_k|^p\Big),
\end{aligned}
$$

where $h(\bar{x}_\Delta(t))$ is independent of $N_k$. We now have

$$\begin{aligned}\mathbb{E}\Big(|x_\Delta(t)-\bar{x}_\Delta(t)|^p\Big|\mathcal{F}_{k(t)}\Big) &\le 3^{p-1}\Big((1\vee c(p)\sigma^p)\Delta^{p/2}(\pi(\Delta))^p+\bar{c}\alpha_3^p|\bar{x}_\Delta(t^-)|^p\Delta\Big)\\ &\le 3^{p-1}(1\vee c(p)\sigma^p\vee\bar{c}\alpha_3^p)\Big(\Delta^{p/2}(\pi(\Delta))^p+|\bar{x}_\Delta(t)|^p\Delta\Big)\\ &\le \mathfrak{D}_1\Big(\Delta^{p/2}(\pi(\Delta))^p+\Delta\Big)|\bar{x}_\Delta(t)|^p,\end{aligned}$$

which is (4.27), where $\mathfrak{D}_1=3^{p-1}[(1\vee c(p)\sigma^p)\vee\bar{c}\alpha_3^p]$. For $p\in(0,2)$, the Jensen inequality yields

$$\begin{aligned}\mathbb{E}\Big(|x_\Delta(t)-\bar{x}_\Delta(t)|^p\Big|\mathcal{F}_{k(t)}\Big) &\le \Big\{\mathbb{E}\Big(|x_\Delta(t)-\bar{x}_\Delta(t)|^2\Big|\mathcal{F}_{k(t)}\Big)\Big\}^{p/2}\\ &\le \Big\{\mathfrak{D}_1\Big(\Delta(\pi(\Delta))^2+\Delta\Big)|\bar{x}_\Delta(t)|^p\Big\}^{p/2}\\ &\le 2^{p/2-1}\mathfrak{D}_1^{p/2}\Big(\Delta^{p/2}(\pi(\Delta))^p+\Delta^{p/2}\Big)(|\bar{x}_\Delta(t)|^p)^{p/2}\\ &\le \mathfrak{D}_2\Big(\Delta^{p/2}(\pi(\Delta))^p\Big)|\bar{x}_\Delta(t)|^p,\end{aligned}$$

which is the required assertion in (4.28), where $\mathfrak{D}_2=2^{p/2}\mathfrak{D}_1^{p/2}$. The proof is thus complete. □

The finite moment of the truncated EM solution is revealed in the following lemma.

**Lemma 4.5.2.** *Let Assumptions 4.3.1 and 4.3.3 hold. Then for any $p\ge 3$*

$$\sup_{0\le\Delta\le\Delta^*}\sup_{0\le t\le T}(\mathbb{E}|x_\Delta(t)|^p)\le\rho_3,\quad \forall T>0, \tag{4.29}$$

*where $\rho_3:=\rho_3(T,p,K_5,\xi)$ and may change between occurrences.*

*Proof.* Fix any $\Delta\in(0,\Delta^*)$ and $T\ge 0$. For $t\in[0,T]$, we derive from (4.3), (4.19) and Lemma 4.4.4 that

$$\mathbb{E}|x_\Delta(t)|^p-|\xi(0)|^p$$

$$
\begin{aligned}
&\le \mathbb{E}\int_0^t p|x_\Delta(s^-)|^{p-2}\Big(\bar{x}_\Delta(s^-)f_\Delta(\bar{x}_\Delta(s^-)) + \frac{p-1}{2}|\varphi(\bar{x}_\Delta((s-\tau)^-))g_\Delta(\bar{x}_\Delta(s^-))|^2\Big)ds \\
&+ \mathbb{E}\int_0^t p|x_\Delta(s^-)|^{p-2}(x_\Delta(s^-) - \bar{x}_\Delta(s^-))f_\Delta(\bar{x}_\Delta(s^-))ds \\
&+ \lambda\mathbb{E}\Big(\int_0^t |x_\Delta(s^-) + h(\bar{x}_\Delta(s^-))|^p - |x_\Delta(s^-)|^p\Big)ds \\
&\le H_{11} + H_{12} + H_{13},
\end{aligned}
$$

where

$$
\begin{aligned}
H_{11} &= \mathbb{E}\int_0^t K_5 p|x_\Delta(s^-)|^{p-2}(1 + |\bar{x}_\Delta(s^-)|^2)ds, \\
H_{12} &= \mathbb{E}\int_0^t p|x_\Delta(s^-)|^{p-2}\Big(x_\Delta(s^-) - \bar{x}_\Delta(s^-)\Big)f_\Delta(\bar{x}_\Delta(s^-))ds
\end{aligned}
$$

and

$$
H_{13} = \lambda\mathbb{E}\Big(\int_0^t |x_\Delta(s^-) + h(\bar{x}_\Delta(s^-))|^p - |x_\Delta(s^-)|^p\Big)ds.
$$

Applying the Young inequality, we obtain

$$
\begin{aligned}
H_{11} &= K_5 p\mathbb{E}\int_0^t |x_\Delta(s^-)|^{p-2}(1 + |\bar{x}_\Delta(s^-)|^2)ds \\
&\le K_5 p\int_0^t \Big(\frac{(p-2)}{p}\mathbb{E}|x_\Delta(s^-)|^p + \frac{2}{p}\mathbb{E}(1 + |\bar{x}_\Delta(s^-)|)^p\Big)ds \\
&\le K_5\int_0^t \Big((p-2)\mathbb{E}|x_\Delta(s^-)|^p + 2^p(1 + \mathbb{E}|\bar{x}_\Delta(s^-)|^p)\Big)ds \\
&\le c_1\int_0^t (1 + \mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p)ds,
\end{aligned}
$$

where $c_1 = K_5[(p-2)\vee 2^p]$. For $s \in [0,t]$, we note from the triangle inequality

$$
|x_\Delta(s^-)| \le |x_\Delta(s^-) - \bar{x}_\Delta(s^-)| + |\bar{x}_\Delta(s^-)|.
$$

This implies for $p \geq 3$, we obtain

$$\begin{aligned}
H_{12} &\leq p\mathbb{E}\int_0^t \Big(|x_\Delta(s^-) - \bar{x}_\Delta(s^-)| + |\bar{x}_\Delta(s^-)|\Big)^{p-2} \\
&\times |x_\Delta(s^-) - \bar{x}_\Delta(s^-)||f_\Delta(\bar{x}_\Delta(s^-))|ds \\
&\leq 2^{(p-3)}p\mathbb{E}\int_0^t \Big(|x_\Delta(s^-) - \bar{x}_\Delta(s^-)|^{p-2} + |\bar{x}_\Delta(s^-)|^{p-2}\Big) \\
&\times |x_\Delta(s^-) - \bar{x}_\Delta(s^-)||f_\Delta(\bar{x}_\Delta(s^-))|ds \\
&= H_{121} + H_{122},
\end{aligned}$$

where

$$H_{121} = 2^{(p-3)}p\mathbb{E}\int_0^t |\bar{x}_\Delta(s^-)|^{p-2}|x_\Delta(s^-) - \bar{x}_\Delta(s)||f_\Delta(\bar{x}_\Delta(s^-))|ds$$

and

$$H_{122} = 2^{(p-3)}p\mathbb{E}\int_0^t |x_\Delta(s^-) - \bar{x}_\Delta(s^-)|^{p-1}|f_\Delta(\bar{x}_\Delta(s^-))|ds.$$

By Lemma 4.5.1 and (4.21), we now have

$$\begin{aligned}
H_{121} &\leq 2^{(p-3)}p\int_0^t \mathbb{E}\Big\{|\bar{x}_\Delta(s)|^{p-2}|f_\Delta(\bar{x}_\Delta(s))|\mathbb{E}\Big(|x_\Delta(s) - \bar{x}_\Delta(s)|\mathcal{F}_{k(s)})\Big)\Big\}ds \\
&\leq 2^{(p-3)}p\mathfrak{D}_2(\pi(\Delta))\Delta^{1/2}(\pi(\Delta))\int_0^t \mathbb{E}\Big\{|\bar{x}_\Delta(s)|(|\bar{x}_\Delta(s)|^{p-2})\Big\}ds \\
&\leq 2^{(p-3)}p\mathfrak{D}_2(\pi(\Delta))\Delta^{1/2}(\pi(\Delta))\int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1}ds \\
&\leq 2^{(p-3)}p\mathfrak{D}_2(\pi(\Delta))^2\Delta^{1/2}\int_0^t \Big(\frac{1}{p} + \frac{(p-1)}{p}\mathbb{E}|\bar{x}_\Delta(s)|^p\Big)ds \\
&\leq c_2 + c_3\int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^p ds, \qquad (4.30)
\end{aligned}$$

where $c_2 = 2^{(p-3)}\mathfrak{D}_2 T$ and $c_3 = 2^{(p-3)}\mathfrak{D}_2(p-1)$ , noting that $(\pi(\Delta))\Delta^{1/4} \leq 1$ and hence

$$[(\pi(\Delta))\Delta^{1/4}]^2 \leq 1.$$

Also by (4.21), we have

$$H_{122} \leq 2^{(p-3)} p\pi(\Delta) \int_0^t \mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^{p-1} ds. \tag{4.31}$$

Do note for $p \geq 3$ and $\bar{w} \in (0, 1/4]$, we have $p\bar{w} \leq (p-1)/2$ and then

$$\Delta^{(p-1)/2 - \bar{w}p} \leq 1. \tag{4.32}$$

So for $p \geq 3$ and $\bar{w} = 1/4$, we obtain from (4.31), Lemma 4.5.1, (4.32) and the Young inequality

$$\begin{aligned}
H_{122} &\leq 2^{(p-3)} p\mathfrak{D}_1 \Big( \Delta^{(p-1)/2} (\pi(\Delta))^{p-1} (\pi(\Delta)) + \Delta(\pi(\Delta)) \Big) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\leq 2^{(p-3)} p\mathfrak{D}_1 \Big( \Delta^{(p-1)/2} (\pi(\Delta))^{p} + \Delta(\pi(\Delta)) \Big) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\leq 2^{(p-3)} p\mathfrak{D}_1 \Big( \Delta^{(p-2)/4} + \Delta(\pi(\Delta)) \Big) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\leq 2^{(p-2)} p\mathfrak{D}_1 \int_0^t \Big( \frac{1}{p} + \frac{(p-1)}{p} \mathbb{E}|\bar{x}_\Delta(s)|^{p} \Big) ds \\
&\leq c_4 + c_5 \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^p ds,
\end{aligned}$$

where $c_4 = 2^{(p-2)}\mathfrak{D}_1 T$ and $c_5 = 2^{(p-2)}\mathfrak{D}_1(p-1)$. We now combine $H_{121}$ and $H_{122}$ to have

$$\begin{aligned}
H_{12} &\leq c_2 + c_4 + (c_3 + c_5) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^p ds \\
&\leq c_6 + c_7 \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^p ds,
\end{aligned}$$

where $c_6 = c_2 + c_4$ and $c_7 = c_3 + c_5$. Also we estimate $H_{13}$ as

$$\begin{aligned}
H_{13} &= \lambda\mathbb{E}\Big(\int_0^t |x_\Delta(s^-) + h(\bar{x}_\Delta(s^-))|^p - |x_\Delta(s^-)|^p\Big)ds \\
&\leq \lambda\mathbb{E}\Big(\int_0^t 2^{p-1}|x_\Delta(s^-)|^p + 2^{p-1}|h(\bar{x}_\Delta(s^-))|^p - |x_\Delta(s^-)|^p\Big)ds \\
&\leq \lambda\mathbb{E}\Big(\int_0^t (2^{p-1}-1)|x_\Delta(s^-)|^p + 2^{p-1}\alpha_3^p|\bar{x}_\Delta(s^-)|^p\Big)ds \\
&\leq c_8\int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p)ds,
\end{aligned}$$

where $c_8 = \lambda[(2^{p-1}-1)\vee 2^{p-1}\alpha_3^p]$. Combining $H_{11}$, $H_{12}$ and $H_{13}$, we have

$$\begin{aligned}
\mathbb{E}|x_\Delta(t)|^p &\leq |\xi(0)|^p + (c_1T + c_6) + \int_0^t \Big((c_1+c_8)\mathbb{E}|x_\Delta(s)|^p + (c_1+c_7+c_8)\mathbb{E}|\bar{x}_\Delta(s)|^p\Big)ds \\
&\leq c_9 + 2c_{10}\int_0^t \sup_{0\leq u\leq s}\Big(\mathbb{E}|x_\Delta(u)|^p\Big)ds,
\end{aligned}$$

where $c_9 = |\xi(0)|^p + c_1T + c_6$ and $c_{10} = (c_1+c_8)\vee(c_1+c_7+c_8)$. As this holds for any $t \in [0,T]$, we then have

$$\sup_{0\leq u\leq t}(\mathbb{E}|x_\Delta(u)|^p) \leq c_9 + 2c_{10}\int_0^t \sup_{0\leq u\leq s}\Big(\mathbb{E}|x_\Delta(u)|^p\Big)ds.$$

The Gronwall inequality yields

$$\sup_{0\leq u\leq T}(\mathbb{E}|x_\Delta(u)|^p) \leq \rho_3$$

as the required assertion, where $\rho_3 = c_9 e^{2c_{10}T}$ is independent of $\Delta$. □

### 4.5.2 Finite time strong convergence

We can now establish the finite time strong convergence of the truncated EM solutions to the exact solution of SDDE (4.2). Before that, let us first establish the following useful lemma.

**Lemma 4.5.3.** *Suppose Assumptions 4.3.1, 4.3.3 and 4.4.1 hold and fix $T > 0$. Then for any $\epsilon \in (0,1)$, there exists a pair $n = n(\epsilon) > 0$ and $\bar{\Delta} = \bar{\Delta}(\epsilon) > 0$ such that*

$$\mathbb{P}(\vartheta_{\Delta,n} \leq T) \leq \epsilon \tag{4.33}$$

*as long as $\Delta \in (0, \bar{\Delta}]$, where*

$$\vartheta_{\Delta,n} = \inf\{t \in [0,T] : x_\Delta(t) \notin (1/n, n)\} \tag{4.34}$$

*is a stopping time.*

*Proof.* Let $Z(\cdot)$ be the Lyapunov function in (4.11). Then for $t \in [0,T]$, the Itô formula gives us

$$\begin{aligned}
&\mathbb{E}(Z(x_\Delta(t \wedge \vartheta_{\Delta,n})) - Z(\xi(0))) \\
&= \mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} \Big[Z_x(x_\Delta(s^-))f_\Delta(\bar{x}_\Delta(s^-)) + \frac{1}{2}Z_{xx}(x_\Delta(s^-))\varphi(\bar{x}_\Delta((s-\tau)^-))^2 g_\Delta(\bar{x}_\Delta(s^-))^2 \\
&+ \lambda(Z(x_\Delta(s^-) + h(\bar{x}_\Delta(s^-))) - Z(x_\Delta(s^-)))\Big]ds.
\end{aligned}$$

By expansion, we obtain

$$\begin{aligned}
&\mathbb{E}(Z(x_\Delta(t \wedge \vartheta_{\Delta,n})) - Z(\xi(0))) \\
&\leq \mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} \Big[\Big(Z_x(x_\Delta(s^-))f_\Delta(x_\Delta(s^-)) \\
&+ \frac{1}{2}Z_{xx}(x_\Delta(s^-))\varphi(x_\Delta((s-\tau)^-))^2 g_\Delta(x_\Delta(s^-))^2 \\
&+ \lambda(Z(x_\Delta(s^-) + h(x_\Delta(s^-))) - Z(x_\Delta(s^-)))\Big) \\
&+ Z_x(x_\Delta(s^-))\Big(f_\Delta(\bar{x}_\Delta(s^-)) - f_\Delta(x_\Delta(s^-))\Big) \\
&+ \frac{1}{2}Z_{xx}(x_\Delta(s^-))\Big(\varphi(\bar{x}_\Delta((s-\tau)^-))^2 g_\Delta(\bar{x}_\Delta(s^-))^2 - \varphi(x_\Delta((s-\tau)^-))^2 g_\Delta(x_\Delta(s^-))^2\Big) \\
&+ \lambda\Big(Z(x_\Delta(s^-) + h(\bar{x}_\Delta(s^-))) - Z(x_\Delta(s^-) + h(x_\Delta(s^-)))\Big)\Big]ds \\
&\leq \mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} L(x_\Delta(s^-), x_\Delta((s-\tau)^-))ds + H_{21} + H_{22} + H_{23}
\end{aligned}$$

Here,

$$
\begin{aligned}
L(x_\Delta(s^-), x_\Delta((s-\tau)^-)) \le\ & \ell(x_\Delta(s^-), x_\Delta((s-\tau)^-)) \\
& + \lambda(Z(x_\Delta(s^-) + h(x_\Delta(s^-))) - Z(x_\Delta(s^-)))
\end{aligned}
$$

is the operator (4.3) which is independent of $t$ with

$$
\begin{aligned}
\ell(x_\Delta(s^-), x_\Delta((s-\tau)^-)) =\ & Z_x(x_\Delta(s^-)) f_\Delta(x_\Delta(s^-)) \\
& + \frac{1}{2} Z_{xx}(x_\Delta(s^-)) \varphi(x_\Delta((s-\tau)^-))^2 g_\Delta(x_\Delta(s^-))^2,
\end{aligned}
$$

and

$$
\begin{aligned}
H_{21} &= \mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_x(x_\Delta(s^-))\Big(f_\Delta(\bar{x}_\Delta(s^-)) - f_\Delta(x_\Delta(s^-))\Big)ds, \\
H_{22} &= \frac{1}{2}\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_{xx}(x_\Delta(s^-))\Big(\varphi(\bar{x}_\Delta((s-\tau)^-))^2 g_\Delta(\bar{x}_\Delta(s^-))^2 \\
&\quad - \varphi(x_\Delta((s-\tau)^-))^2 g_\Delta(x_\Delta(s^-))^2\Big)ds, \\
H_{23} &= \lambda\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} \Big(Z(x_\Delta(s^-) + h(\bar{x}_\Delta(s^-))) - Z(x_\Delta(s^-) + h(x_\Delta(s^-)))\Big)ds.
\end{aligned}
$$

By Assumption 4.3.3, there exists a constant $K_6 > 0$ such that for $s \in [0, t\wedge\vartheta_{\Delta,n}]$

$$
L(x_\Delta(s^-), x_\Delta((s-\tau)^-)) \le K_6.
$$

Also by Lemma 4.4.2, we have

$$
H_{21} \le K_n \mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_x(x_\Delta(s^-))|\bar{x}_\Delta(s^-) - x_\Delta(s^-)|ds.
$$

Meanwhile, for $x_\Delta(s^-), \bar{x}_\Delta(s^-) \in [1/n, n]$, we derive that

$$
H_{22} = \frac{1}{2}\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_{xx}\Big(g_\Delta(x_\Delta(s^-))^2|\varphi(\bar{x}_\Delta((s-\tau)^-))^2 - \varphi(x_\Delta((s-\tau)^-))^2|
$$

$$
\begin{aligned}
&+\varphi(\bar{x}_\Delta((s-\tau)^-))^2|g_\Delta(\bar{x}_\Delta(s^-))^2-g_\Delta(x_\Delta(s^-))^2|\Big)ds\\
&\leq \mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_{xx}\Big(x_\Delta(s^-))(\sigma^2\mu(n)K_n|\bar{x}_\Delta(s^-)-x_\Delta(s^-)|\\
&+\sigma(\mu(n))^2L_n|\bar{x}_\Delta((s-\tau)^-)-x_\Delta((s-\tau)^-)|\Big)ds,
\end{aligned}
$$

where (4.6), (4.16) and (4.17) have been used. Moreover, by the definition of (4.11), we have

$$
\begin{aligned}
H_{23}\leq \lambda\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} &\Big((x_\Delta(s^-)+h(\bar{x}_\Delta(s^-)))^\beta-1-\beta\log(x_\Delta(s^-)+h(\bar{x}_\Delta(s^-)))\\
&-(x_\Delta(s^-)+h(x_\Delta(s^-)))^\beta+1+\beta\log(x_\Delta(s^-)+h(x_\Delta(s^-)))\Big)ds\\
&\leq H_{231}+H_{232},
\end{aligned}
$$

where

$$
H_{231}=\lambda\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}}|(x_\Delta(s^-)+\alpha_3\bar{x}_\Delta(s^-))^\beta-(x_\Delta(s^-)+\alpha_3x_\Delta(s^-))^\beta|ds
$$

and

$$
H_{232}=\lambda\beta\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}}|\log(x_\Delta(s^-)+\alpha_3\bar{x}_\Delta(s^-))-\log(x_\Delta(s^-)+\alpha_3x_\Delta(s^-))|ds.
$$

Applying the mean value theorem, we obtain

$$
\begin{aligned}
H_{231}&\leq n\lambda\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}}|x_\Delta(s^-)+\alpha_3\bar{x}_\Delta(s^-)-\alpha_3x_\Delta(s^-)-x_\Delta(s^-)|ds\\
&=n\lambda\alpha_3\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}}|\bar{x}_\Delta(s^-)-x_\Delta(s^-)|ds.
\end{aligned}
$$

Similarly, we also have

$$
H_{232}\leq n\lambda\beta\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}}|x_\Delta(s^-)+\alpha_3\bar{x}_\Delta(s^-)-\alpha_3x_\Delta(s^-)-x_\Delta(s^-)|ds
$$

$$= n\lambda\alpha_3\beta\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x_\Delta(s^-)|ds.$$

Substituting $H_{231}$ and $H_{232}$ back into $H_{23}$, we have

$$H_{23} \leq n\lambda\alpha_3(1+\beta)\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x_\Delta(s^-)|ds.$$

We thus combine the $H_{21}$, $H_{22}$ and $H_{23}$ to have

$$\begin{aligned}
&\mathbb{E}(Z(x_\Delta(t\wedge\vartheta_{\Delta,n}))) \leq Z(\xi(0)) + K_6T \\
&+ \sigma(\mu(n))^2L_n\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_{xx}(x_\Delta(s^-))|\bar{x}_\Delta((s-\tau)^-) - x_\Delta((s-\tau)^-)|ds \\
&+ K_n\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_x(x_\Delta(s^-))|\bar{x}_\Delta(s^-) - x_\Delta(s^-)|ds \\
&+ \sigma^2\mu(n)K_n\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} Z_{xx}(x_\Delta(s^-))|\bar{x}_\Delta(s^-) - x_\Delta(s^-)|ds \\
&+ n\lambda\alpha_3(1+\beta)\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x_\Delta(s^-)|ds.
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
\mathbb{E}(Z(x_\Delta(t\wedge\vartheta_{\Delta,n}))) &\leq Z(\xi(0)) + K_6T + K_7\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s-\tau) - x_\Delta(s-\tau)|ds \\
&+ K_8\mathbb{E}\int_0^{t\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s) - x_\Delta(s)|ds \\
&\leq Z(\xi(0)) + K_6T + K_7\mathbb{E}\int_{-\tau}^0 |\xi([s/\Delta]\Delta) - \xi(s)|ds \\
&+ (K_7+K_8)\int_0^T \mathbb{E}\Big(\mathbb{E}|\bar{x}_\Delta(s) - x_\Delta(s)|^p\Big|\mathcal{F}_{k(s)}\Big)^{1/p}ds
\end{aligned}$$

where

$$K_7 = \max_{1/n\leq x\leq n}\{Z_{xx}(x)\sigma(\mu(n))^2L_n\}$$

and

$$K_8 = \max_{1/n \le x \le n} \{Z_x(x)K_n + Z_{xx}(x)\sigma^2\mu(n)K_n + n\lambda\alpha_3(1+\beta)\}.$$

By Lemma 4.5.1 and 4.5.2, we now have

$$\begin{aligned}
\mathbb{E}(Z(x_\Delta(t\wedge\vartheta_{\Delta,n}))) &\le Z(\xi(0)) + K_6T + K_3K_7T\Delta^\gamma + (K_7+K_8)\mathfrak{D}_1^{1/p} \\
&\times \Big(\Delta^{p/2}(\pi(\Delta))^p + \Delta\Big)^{1/p} \int_0^T (\mathbb{E}|\bar{x}_\Delta(s)|^p)^{1/p} ds \\
&\le Z(\xi(0)) + K_6T + K_3K_7T\Delta^\gamma + (K_7+K_8)\mathfrak{D}_1^{1/p} \\
&\times \Big(\Delta^{p/2}(\pi(\Delta))^p + \Delta\Big)^{1/p} \int_0^T (\sup_{0\le u\le s} (\mathbb{E}|\bar{x}_\Delta(u)|^p))^{1/p} ds \\
&\le Z(\xi(0)) + K_6T + \nu_1\Delta^\gamma + \nu_2(\Delta^{p/2}(\pi(\Delta))^p + \Delta)^{1/p}\rho_3^{1/p}T.
\end{aligned}$$

where $\nu_1 = K_3K_7T$ and $\nu_2 = (K_7+K_8)\mathfrak{D}_1^{1/p}$. Hence,

$$\mathbb{P}(\vartheta_{\Delta,n} \le T) \le \frac{Z(\xi(0)) + K_6T + \nu_1\Delta^\gamma + \nu_2(\Delta^{p/2}(\pi(\Delta))^p + \Delta)^{1/p}\rho_3^{1/p}T}{Z(1/n)\wedge Z(n)}. \tag{4.35}$$

For any $\epsilon \in (0,1)$, we may select sufficiently large $n$ such that

$$\frac{Z(\xi(0)) + K_6T}{Z(1/n)\wedge Z(n)} \le \frac{\epsilon}{2} \tag{4.36}$$

and sufficiently small of each step size $\Delta \in (0, \bar{\Delta}]$ such that

$$\frac{\nu_1\Delta^\gamma + \nu_2(\Delta^{p/2}(\pi(\Delta))^p + \Delta)^{1/p}\rho_3^{1/p}T}{Z(1/n)\wedge Z(n)} \le \frac{\epsilon}{2}. \tag{4.37}$$

Therefore, we obtain (4.33) by combining (4.36) and (4.37). □

The following lemma shows the finite time strong convergence theory of the truncated EM solutions.

**Lemma 4.5.4.** *Let Assumptions 4.3.1, 4.3.3, 4.4.1 and 4.3.2 hold. Set*

$$\varsigma_{\Delta,n} = \vartheta_{\Delta,n} \wedge \tau_n,$$

*where* $\vartheta_{\Delta,n}$ *and* $\tau_n$ *are* (4.10) *and* (4.34) *respectively. Then for any* $p \geq 2$, $T > 0$, *we have*

$$\mathbb{E}\Big(\sup_{0\leq t\leq T} |x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big) \leq \mathcal{K}\Delta^{p(1/4\wedge\gamma\wedge1/p)} \tag{4.38}$$

*for any sufficiently large* $n$ *and any* $\Delta \in (0, \Delta^*]$, *where* $\mathcal{K}$ *is a constant independent of* $\Delta$. *Consequently, we have*

$$\lim_{\Delta\to0} \mathbb{E}\Big(\sup_{0\leq t\leq T} |x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big) = 0. \tag{4.39}$$

*Proof.* By elementary inequality, it follows from (4.2) and (4.26) that for $t_1 \in [0,T]$

$$\mathbb{E}\Big(\sup_{0\leq t\leq t_1} |x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big) \leq H_{31} + H_{32} + H_{33}.$$

where

$$\begin{aligned} H_{31} &= 3^{p-1}\mathbb{E}\Big(|\int_0^{t_1\wedge\varsigma_{\Delta,n}} [f_\Delta(\bar{x}_\Delta(s^-)) - f(x(s^-))]ds|^p\Big), \\ H_{32} &= 3^{p-1}\mathbb{E}\Big(\sup_{0\leq t\leq t_1} |\int_0^{t\wedge\varsigma_{\Delta,n}} [\varphi(\bar{x}_\Delta((s-\tau)^-))g_\Delta(\bar{x}_\Delta(s^-)) \\ &\quad - \varphi(x((s-\tau)^-))g(x(s^-))]dB(s)|^p\Big) \end{aligned}$$

and

$$H_{33} = 3^{p-1}\mathbb{E}\Big(\sup_{0\leq t\leq t} |\int_0^{t_1\wedge\varsigma_{\Delta,n}} [h(\bar{x}_\Delta(s^-)) - h(x(s^-))]dN(s)|^p\Big).$$

By the Hölder inequality and Lemma 4.4.2, we have

$$H_{31} \leq 3^{p-1}T^{p-1}K_n^p\mathbb{E}\int_0^{t_1\wedge\varsigma_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds, \tag{4.40}$$

Furthermore, the Hölder and Burkholder-Davis Gundy inequalities yield

$$H_{32} \leq 3^{p-1}T^{\frac{p-2}{2}}c_p\mathbb{E}\int_0^{t_1\wedge\varsigma_{\Delta,n}}\Big(|\varphi(\bar{x}_\Delta((s-\tau)^-))g_\Delta(\bar{x}_\Delta(s^-))-\varphi(x((s-\tau)^-))g_\Delta(\bar{x}_\Delta(s^-))$$
$$+\varphi(x((s-\tau)^-))g_\Delta(\bar{x}_\Delta(s^-))-\varphi(x((s-\tau)^-))g(x(s^-))|^p\Big)ds$$
$$\leq 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_p\mathbb{E}\int_0^{t_1\wedge\varsigma_{\Delta,n}}\Big(g_\Delta(\bar{x}_\Delta(s^-))^p|\varphi(\bar{x}_\Delta((s-\tau)^-))-\varphi(x((s-\tau)^-))|^p$$
$$+\varphi(x((s-\tau)^-))^p|g_\Delta(\bar{x}_\Delta(s^-))-g(x(s^-))|^p\Big)ds,$$

where $c_p$ is a positive constant. For $s\in[0,t_1\wedge\varsigma_{\Delta,n}]$, we have $x_\Delta(s^-),\bar{x}_\Delta(s^-)\in[1/n,n]$. So by Assumption 4.4.1, Lemma 4.4.2 and (4.19), we now have

$$H_{32} \leq 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_pL_n^p(\mu(n))^p\mathbb{E}\int_{-\tau}^0|\xi([s/\Delta]\Delta)-\xi(s)|^pds \tag{4.41}$$
$$+2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_p(L_n^p(\mu(n))^p+K_n^p\sigma^p)\mathbb{E}\int_0^{t_1\wedge\varsigma_{\Delta,n}}|\bar{x}_\Delta(s^-)-x(s^-)|^pds.$$
$$\leq 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_pL_n^pK_3^p(\mu(n))^p\Delta^{p\gamma}\tau+2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_p\Big(L_n^p(\mu(n))^p+K_n^p\sigma^p\Big)$$
$$\times\mathbb{E}\int_0^{t_1\wedge\varsigma_{\Delta,n}}|\bar{x}_\Delta(s^-)-x(s^-)|^pds. \tag{4.42}$$

Moreover, we obtain from elementary inequality

$$H_{33}\leq 3^{p-1}\mathbb{E}\Big(\sup_{0\leq t\leq t_1}|\int_0^{t\wedge\varsigma_{\Delta,n}}[h(\bar{x}_\Delta(s^-))-h(x(s^-))]d\widetilde{N}(s)$$
$$+\lambda\int_0^{t\wedge\varsigma_{\Delta,n}}[h(\bar{x}_\Delta(s^-))-h(x(s^-))]ds|^p\Big)$$
$$\leq H_{331}+H_{332},$$

where

$$H_{331}=2^{p-1}3^{p-1}\mathbb{E}\Big(\sup_{0\leq t\leq t_1}|\int_0^{t\wedge\varsigma_{\Delta,n}}[h(\bar{x}_\Delta(s^-))-h(x(s^-))]d\widetilde{N}(s)|^p\Big)$$

and

$$H_{332} = 2^{p-1}3^{p-1}\lambda^p \mathbb{E}\Big(\sup_{0\le t\le t_1} |\int_0^{t\wedge \varsigma_{\Delta,n}} [h(\bar{x}_\Delta(s^-)) - h(x(s^-))]ds|^p\Big).$$

The Doob martingale inequality, martingale isometry and Lemma 4.4.2 give us

$$\begin{aligned}
H_{331} &\le 2^{p-1}3^{p-1}\bar{c}_p\lambda^{\frac{p}{2}}\Big(\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |h(\bar{x}_\Delta(s^-)) - h(x(s^-))|^2 d\widetilde{N}(s)\Big)^{\frac{p}{2}} \\
&\le 2^{p-1}3^{p-1}\bar{c}_p\lambda^{\frac{p}{2}}T^{\frac{p-2}{2}}K_n^p\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds,
\end{aligned}$$

where $\bar{c}_p$ is a positive constant. Moreover by the Hölder inequality and Lemma 4.4.2,

$$\begin{aligned}
H_{332} &\le 2^{p-1}3^{p-1}\lambda^p T^{p-1}\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |h(\bar{x}_\Delta(s^-)) - h(x(s^-))|^p ds \\
&\le 2^{p-1}3^{p-1}\lambda^p T^{p-1}K_n^p\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds,
\end{aligned}$$

where Lemma 4.4.2 has been used. Substituting $H_{331}$ and $H_{332}$ into $H_{33}$ yields

$$H_{33} \le 2^{p-1}3^{p-1}K_n^p(\bar{c}_p\lambda^{\frac{p}{2}}T^{\frac{p-2}{2}} + \lambda^p T^{p-1})\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds. \tag{4.43}$$

We now combine (4.40), (4.41) and (4.43) to have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1} |x_\Delta(t\wedge \varsigma_{\Delta,n}) - x(t\wedge \varsigma_{\Delta,n})|^p\Big) \\
&\le \zeta_1\Delta^{p\gamma}\tau + (\zeta_2+\zeta_3+\zeta_4)\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds \\
&\le \zeta_1\Delta^{p\gamma}\tau + (\zeta_2+\zeta_3+\zeta_4)\mathbb{E}\int_0^{t_1\wedge \varsigma_{\Delta,n}} |\bar{x}_\Delta(s) - x(s)|^p ds,
\end{aligned}$$

where

$$\zeta_1 = 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_p L_n^p K_3^p(\mu(n))^p,$$

$$\zeta_2 = 3^{p-1}T^{p-1}K_n^p,$$
$$\zeta_3 = 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}c_p(L_n^p(\mu(n))^p + K_n^p\sigma^p)$$

and

$$\zeta_4 = 2^{p-1}3^{p-1}K_n^p(\bar{c}_p\lambda^{\frac{p}{2}}T^{\frac{p-2}{2}} + \lambda^p T^{p-1}).$$

Meanwhile by elementary inequality and Lemma 4.5.1,

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big)\\
&\le \zeta_1\Delta^{p\gamma}\tau + 2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\int_0^T \mathbb{E}\Big(\mathbb{E}|\bar{x}_\Delta(s) - x_\Delta(s)|^p\Big|\mathcal{F}_{k(s)}\Big)ds\\
&+ 2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big)ds\\
&\le \zeta_1\Delta^{p\gamma}\tau + 2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\mathfrak{D}_1\Big(\Delta^{p/2}(\pi(\Delta))^p + \Delta\Big)\int_0^T \mathbb{E}|\bar{x}_\Delta(s)|^p ds\\
&+ 2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big)ds
\end{aligned}$$

So by Lemma 4.5.2, we have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big)\\
&\le \zeta_1\tau\Delta^{p\gamma} + 2^{p-1}\rho_3\mathfrak{D}_1T(\zeta_2+\zeta_3+\zeta_4)\Big([\Delta^{p/4}(\pi(\Delta))^p]\Delta^{p/4} + \Delta^{p(1/p)}\Big)\\
&+ 2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big)ds\\
&\le \Big(\zeta_1\tau + 2^{p-1}\rho_3\mathfrak{D}_1T(\zeta_2+\zeta_3+\zeta_4)(\Delta^{p/4}(\pi(\Delta))^p + 1)\Big)\Delta^{p(1/4\wedge\gamma\wedge1/p)}\\
&+ 2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\varsigma_{\Delta,n}) - x(t\wedge\varsigma_{\Delta,n})|^p\Big)ds.
\end{aligned}$$

Noting from (4.20) that $[\Delta^{1/4}(\pi(\Delta))]^p \le 1$, we have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\varsigma_{\Delta,n})-x(t\wedge\varsigma_{\Delta,n})|^p\Big)\\
&\le \Big(\zeta_1\tau+2^p\rho_3\mathfrak{D}_1T(\zeta_2+\zeta_3+\zeta_4)\Big)\Delta^{p(1/4\wedge\gamma\wedge1/p)}\\
&+2^{p-1}(\zeta_2+\zeta_3+\zeta_4)\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\varsigma_{\Delta,n})-x(t\wedge\varsigma_{\Delta,n})|^p\Big)ds.
\end{aligned}$$

By the Gronwall inequality, we obtain

$$\mathbb{E}\Big(\sup_{0\le t\le T}|x_\Delta(t\wedge\varsigma_{\Delta,n})-x(t\wedge\varsigma_{\Delta,n})|^p\Big)\le\mathcal{K}\Delta^{p(1/4\wedge\gamma\wedge1/p)},$$

where $\mathcal{K}=\varrho_1(p)e^{\varrho_2(p)}$ with

$$\varrho_1(p)=\zeta_1\tau+2^p\rho_3\mathfrak{D}_1T(\zeta_2+\zeta_3+\zeta_4)$$

and

$$\varrho_2(p)=2^{p-1}(\zeta_2+\zeta_3+\zeta_4).$$

Moreover, the required inequality (4.39) is deduced by setting $\Delta\to0$. □

The following gives the strong convergence theory of the truncated EM scheme.

**Theorem 4.5.5.** *Let Assumptions 4.3.1, 4.3.3, 4.4.1 and 4.3.2 hold. Then for any $p\ge2$, we have*

$$\lim_{\Delta\to0}\mathbb{E}\Big(\sup_{0\le t\le T}|x_\Delta(t)-x(t)|^p\Big)=0 \tag{4.44}$$

*and consequently*

$$\lim_{\Delta\to0}\mathbb{E}\Big(\sup_{0\le t\le T}|\bar{x}_\Delta(t)-x(t)|^p\Big)=0. \tag{4.45}$$

*Proof.* We only need to prove the theorem for $p\ge3$ as for $p\in[2,3)$ it follows from the case of $p=3$ and the Hölder inequality. Let $\vartheta_{\Delta,n}$, $\tau_n$ and $\varsigma_{\Delta,n}$ be the same as before and set

$$e_\Delta(t) = x_\Delta(t) - x(t).$$

For any arbitrarily $\delta > 0$, the Young inequality yields

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^p\Big) \qquad (4.46)\\
&= \mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^p 1_{\{\tau_n > T \text{ and } \vartheta_{\Delta,n} > T\}}\Big) + \mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^p 1_{\{\tau_n \le T \text{ or } \vartheta_{\Delta,n} \le T\}}\Big)\\
&\le \mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^p 1_{\{\varsigma_{\Delta,n} > T\}}\Big) + \frac{\delta}{2}\mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^{2p}\Big) + \frac{1}{2\delta}\mathbb{P}(\tau_n \le T \text{ or } \vartheta_{\Delta,n} \le T).
\end{aligned}$$

So for $p \ge 3$, Lemmas 4.3.5 and 4.5.2 give us

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^{2p}\Big) &\le 2^{2p}\mathbb{E}\Big(\sup_{0\le t\le T} |x(t)|^{2p} \vee \sup_{0\le t\le T} |x_\Delta(t)|^{2p}\Big)\\
&\le 2^{2p}(\rho_1 \vee \rho_3)^2. \qquad (4.47)
\end{aligned}$$

Also by Theorem 4.3.4 and Lemma 4.5.4,

$$\mathbb{P}(\varsigma_{\Delta,n} \le T) \le \mathbb{P}(\tau_n \le T) + \mathbb{P}(\vartheta_{\Delta,n} \le T). \qquad (4.48)$$

Moreover, by Lemma 4.5.4

$$\mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^p 1_{\{\varsigma_{\Delta,n} > T\}}\Big) \le \mathcal{K}\Delta^{p(1/4\wedge\gamma\wedge 1/p)}. \qquad (4.49)$$

Therefore, we substitute (4.47), (4.48) and (4.49) into (4.46) to have

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T} |e_\Delta(t)|^p\Big) &\le \frac{2^{2p}(\rho_1 \vee \rho_3)^2\delta}{2} + \mathcal{K}\Delta^{p(1/4\wedge\gamma\wedge 1/p)}\\
&\quad + \frac{1}{2\delta}\mathbb{P}(\tau_n \le T) + \frac{1}{2\delta}\mathbb{P}(\vartheta_{\Delta,n} \le T).
\end{aligned}$$

Given $\epsilon \in (0,1)$, we can select $\delta$ so that

$$\frac{2^{2p}(\rho_1 \vee \rho_3)^2\delta}{2} \leq \frac{\epsilon}{4}. \tag{4.50}$$

Similarly, for any given $\epsilon \in (0,1)$, there exists $n_o$ so that for $n \geq n_o$, we may select $\delta$ to have

$$\frac{1}{2\delta}\mathbb{P}(\tau_n \leq T) \leq \frac{\epsilon}{4} \tag{4.51}$$

and select $n(\epsilon) \leq n_o$ such that for $\Delta \in (0, \bar{\Delta}]$

$$\frac{1}{2\delta}\mathbb{P}(\vartheta_{\Delta,n} \leq T) \leq \frac{\epsilon}{4}. \tag{4.52}$$

Finally, we may select $\Delta \in (0, \bar{\Delta}]$ sufficiently small for $\epsilon \in (0,1)$ such that

$$\mathcal{K}\Delta^{p(1/4\wedge\gamma\wedge 1/p)} \leq \frac{\epsilon}{4}. \tag{4.53}$$

Combining (4.50), (4.51), (4.52) and (4.53), we establish

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|x_\Delta(t) - x(t)|^p\Big) \leq \epsilon.$$

Therefore, we obtain (4.44) and clearly, by Lemma 4.5.1, also get (4.45) by letting $\Delta \to 0$. □

## 4.6 Numerical examples

In this section, we analyse the strong convergence result established in Theorem 4.5.5 by comparing the truncated Euler-Maruyama (TEM) scheme with backward Euler-Maruyama (BEM) scheme for SDDE (4.3) without $\alpha_{-1}x(t)^{-1}$ term. It is already noted in [39] that the BEM scheme is not known to cope with $\alpha_{-1}x(t)^{-1}$ term but the TEM could cope with this term. Consider the following form of SDDE (4.2)

$$dx(t) = (\alpha_{-1}x(t^-)^{-1} - \alpha_0 + \alpha_1 x(t^-) - \alpha_2 x(t^-)^2)dt$$

$$+\varphi(x((t-1)^-))x(t^-)^{5/4}dB(t)+\alpha_3 x(t^-)dN(t), \tag{4.54}$$

with initial data $\xi(t)=0.2$. Here $\varphi(y)$ is a sigmoid-type function defined by

$$\varphi(y)=\begin{cases}\frac{1}{2}\frac{(1+(e^y-e^{-y}))}{(e^y+e^{-y})}, & \text{if } y\geq 0\\ \frac{1}{4}, & \text{Otherwise,}\end{cases} \tag{4.55}$$

Obviously, equation (4.55) meets all the conditions imposed on $\varphi(y)$ (see [39]). The coefficient terms $f(x)=\alpha_{-1}x^{-1}-\alpha_0+\alpha_1 x-\alpha_2 x^2$ and $g(x)=x^{5/4}$ are locally Lipschitz continuous and hence fulfil Assumption 4.3.5. Moreover, we easily observe

$$\sup_{1/u\leq x\leq u}(|f(x)|\vee g(x))\leq K_9 u^2, \quad u\geq 1,$$

where $K_9=\alpha_{-1}+\alpha_0+\alpha_1+\alpha_2+\alpha_3$. We can now use $\mu=K_9u^2$ with inverse $\mu^{-1}(u)=(u/K_9)^{1/2}$.

### 4.6.1 Numerical results

By choosing $\pi(\Delta)=\Delta^{-2/3}$, step size $\Delta=10^{-2}$ and the following coefficient values in Table 4.1, we obtain Monte Carlo simulated sample path of $x(t)$ to SDDE (4.54) in Figure 4.1 using the TEM scheme.

| $\alpha_{-1}$ | $\alpha_0$ | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ |
|---|---|---|---|---|
| 0.2 | 0.3 | 0.2 | 0.5 | 1 |

Table 4.1: Coefficient values including $\alpha_{-1}$

By similarly choosing $\pi(\Delta)=\Delta^{-2/3}$, step size $\Delta=10^{-2}$ and the coefficient values in Table 4.2 below, we also obtain Monte Carlo simulated sample paths of $x(t)$ to SDDE (4.54) in Figure 4.2 using the TEM and BEM schemes. Clearly, the sample paths of these two schemes are almost identical for the step size $\Delta=10^{-2}$.

| $\alpha_0$ | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ |
|---|---|---|---|
| 0.3 | 0.2 | 0.5 | 1 |

Table 4.2: Coefficient values excluding $\alpha_{-1}$

Finally, the log-log plot of the strong errors between the TEM and BEM numerical solutions for step sizes $10^{-2}$, $10^{-3}$, $10^{-4}$ and $10^{-5}$ is displayed in Figure 4.3 with a reference line of slope 1.0. Do note this simulated result of strong errors is not yet established theoretically.

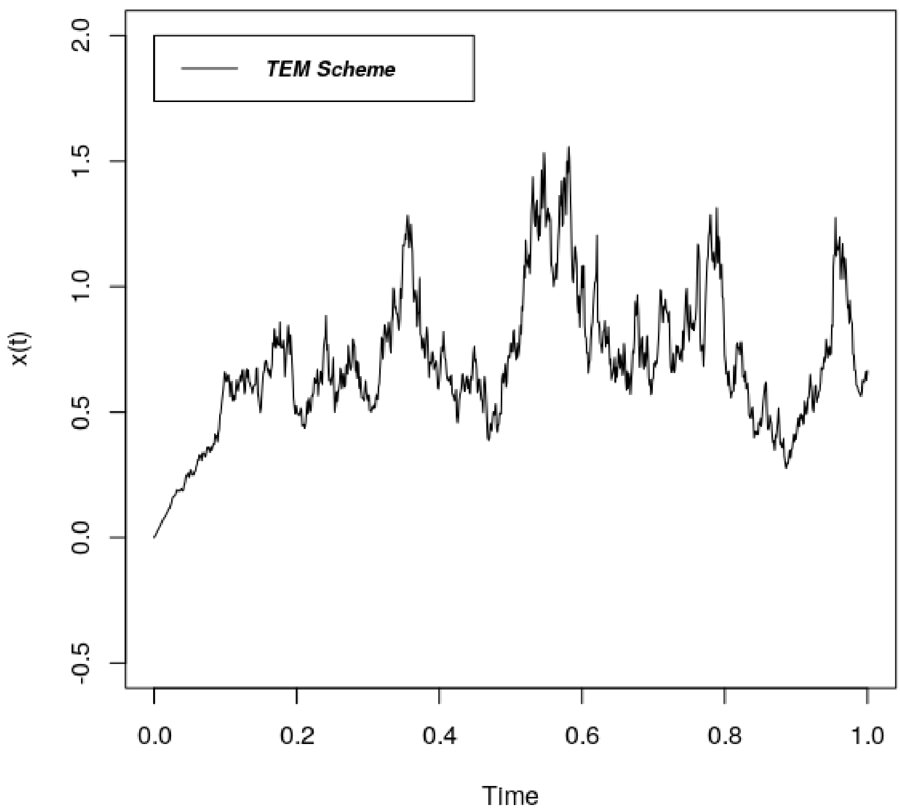


Figure 4.1: Simulated sample path of $x(t)$ when $\Delta = 0.01$

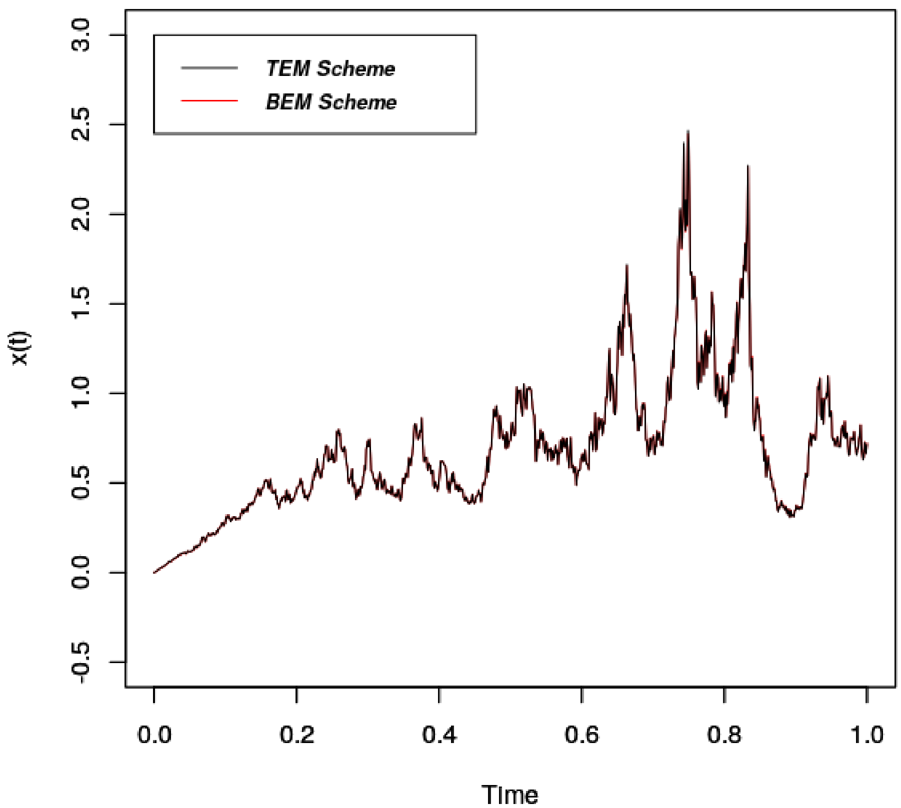


Figure 4.2: Convergence of TEM and BEM solutions when $\Delta = 0.01$

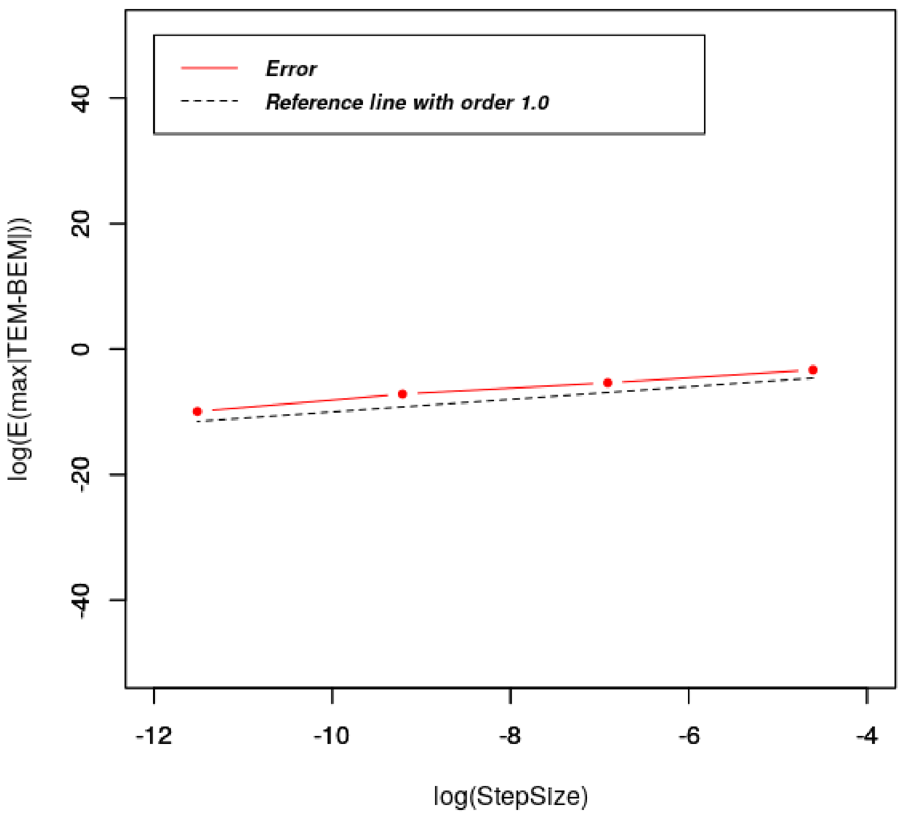


Figure 4.3: Strong errors between TEM and BEM schemes

## 4.7 Summary

In this chapter, we proposed the Poisson-jump Ait-Sahalia-type interest rate model with delay. In this case, we retained the delayed volatility function introduced in Chapter 3 and added Poisson-driven jumps to help explain interest rate dynamics against unexpected joint effects of extreme volatility and jump behaviour or information flows which commonly pervade most financial markets. We proved analytical properties such as existence of pathwise unique positive global solutions and finite moments of the exact solution to the proposed model.

Apparently, the proposed model is not analytically tractable. So, we employed the truncated EM scheme to approximate it. Then we proceeded to establish numerical properties such as boundedness of moments in the strong sense and finite time strong convergence order of the truncated EM approximate solutions to the exact solution under the local Lipschitz condition plus the Khasminskii-type condition. We established the strong convergence result of $\mathcal{K}\Delta^{p(1/4\wedge\gamma\wedge1/p)}$. Lastly, we performed some numerical simulations to support the theoretical findings.

# Chapter 5

# Numerical approximation of hybrid Poisson-jump Ait-Sahalia-type interest rate model with delay

## 5.1 Introduction

The aim of this chapter is to perform theoretical and numerical analyses of the proposed SDDE model (1.4) by drawing on applicable analytical and numerical tools and techniques developed in Chapters 3 and 4.

The rest of the chapter is organised as follows: In Section 5.2, we introduce the hybrid Poisson-jump Ait-Sahalia-type interest rate model with delay. We examine the existence and uniqueness of the solution to the proposed model and show that the solution will always be positive in Section 5.3. We also establish boundedness of moments of the exact solution in this section. In Section 5.4, we define the truncated EM scheme for the proposed model. We survey boundedness of moments of the truncated EM approximate solutions and employ the truncated EM techniques to study finite time strong convergence of the numerical solutions to the exact solution in Section 5.5. In Section 5.6, we implement some numerical examples to validate efficiency of the proposed scheme. Brief summary of the results established is then provided in the last section.

## 5.2 The hybrid Poisson-jump Ait-Sahalia-type interest rate model with delay

We let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. Consider the following scalar dynamics as equation of SDDE (1.4)

$$\begin{aligned} dx(t) &= f(x(t^-), r(t))dt + \varphi(x((t-\tau)^-)), r(t))g(x(t^-))dB(t) \\ &\quad + h(x(t^-), r(t))dN(t), \end{aligned} \tag{5.1}$$

such that $f(x,i) = \alpha_{-1}(i)x^{-1} - \alpha_0(i) + \alpha_1(i)x - \alpha_2(i)x^\rho$, $g(x) = x^\theta$, $h(x,i) = \alpha_3(i)x$, $\forall x \in \mathbb{R}_+$ and $i \in S$, where $\varphi(y,i) \in C(\mathbb{R}_+ \times \mathcal{S}; \mathbb{R}_+)$. For each $H \in C^{2,1}(\mathbb{R} \times \mathbb{R}_+ \times \mathcal{S}; \mathbb{R})$, define the jump-diffusion operator $\mathcal{L}H : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}$ by

$$\mathcal{L}H(x,y,t,i) = \mathcal{I}H(x,y,t,i) + \lambda(H(x+h(x),t,i) - H(x,t,i)) + \sum_{j=1}^{N} \gamma_{ij} H(x,t,j), \tag{5.2}$$

where $\mathcal{I}H : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}$ is the diffusion operator defined by

$$\mathcal{I}H(x,y,t,i) = H_t(x,t,i) + H_x(x,t,i)f(x) + \frac{1}{2}H_{xx}(x,t,i)\varphi(y,i)^2 g(x)^2, \tag{5.3}$$

with $H_t(x,t,i)$ and $H_x(x,t,i)$ as first-order partial derivatives with respect to $t$ and $x$, and $H_{xx}(x,t,i)$ as a second-order partial derivative with respect to $x$. Given the jump-diffusion operator, we could deduce the generalised Itô formula as

$$dH(x(t),t,r(t)) = \mathcal{L}H(x(t^-), x((t-\tau)^-), t, r(t))dt$$

$$
\begin{aligned}
&+ H_x(x(t^-), t, r(t))\varphi(x((t-\tau)^-), r(t))g(x(t^-))dB(t) \\
&+ (H(x(t^-)) + h(x(t^-)), t, r(t)) - H(x(t^-), t, r(t)))d\tilde{N}(t) \\
&+ \int_{\mathbb{R}} (H(x(t^-), t, i_0 + q(x(t^-), z)) - H(x(t^-), t, r(t))))M(dt, dz), \quad \text{a.s.}
\end{aligned}
\tag{5.4}
$$

Consult (e.g., see [56] and the references cited therein). We impose the following standing hypotheses which are very useful for the proofs.

**Assumption 5.2.1.** *The volatility function* $\varphi : \mathbb{R}_+ \times S \to \mathbb{R}_+$ *of* SDDE (5.1) *is Borel-measurable and bounded by a positive constant, that is*

$$\varphi(y, i) \leq \sigma, \tag{5.5}$$

$\forall y \in \mathbb{R}_+$ *and* $i \in \mathcal{S}$.

**Assumption 5.2.2.** *For any* $R > 0$*, there exists a constant* $L_R > 0$ *such that the volatility function* $\varphi : \mathbb{R}_+ \times S \to \mathbb{R}_+$ *of* SDDE (5.1) *satisfies*

$$|\varphi(y, i) - \varphi(\bar{y}, i)| \leq L_R |y - \bar{y}|, \tag{5.6}$$

$\forall (y, \bar{y}) \in [\frac{1}{R}, R]$ *and* $i \in \mathcal{S}$.

**Assumption 5.2.3.** *The parameters of* SDDE (5.1) *obey*

$$1 + \rho > 2\theta, \quad \rho, \theta > 1. \tag{5.7}$$

## 5.3 Analytical properties

In this section, we study the existence of pathwise uniqueness and finite moments of the exact solution to SDDE (5.1).

### 5.3.1 Existence and uniqueness of solution

One basic requirement of a financial model is the existence of a pathwise unique positive solution. The following lemma therefore reveals this requirement.

**Lemma 5.3.1.** *Let Assumptions 5.2.1 and 5.2.3 hold. Then for any given initial data*

$$\{x(t) : -\tau \leq t \leq 0\} = \xi(t) \in C([-\tau, 0] : \mathbb{R}_+), \quad r_0 \in \mathcal{S}, \tag{5.8}$$

*there exists a unique global solution $x(t)$ to* SDDE (5.1) *on $t \geq -\tau$ and $x(t) > 0$ a.s.*

*Proof.* Since the coefficient terms of SDDE (5.1) are locally Lipschitz continuous in $[-\tau, \infty)$, then there exists a unique positive maximal local solution $x(t) \in [-\tau, \tau_e)$ for any given initial data (5.8), where $\tau_e$ is the explosion time (e.g., see [56]). Let $n_0 > 0$ be sufficiently large such that

$$\frac{1}{n_0} < \min_{-\tau \leq t \leq 0} |\xi(t)| \leq \max_{-\tau \leq t \leq 0} |\xi(t)| < n_0.$$

For each integer $n \geq n_0$, define the stopping time

$$\tau_n = \inf\{t \in [0, \tau_e) : x(t) \notin (1/n, n)\}. \tag{5.9}$$

Obviously, $\tau_n$ is increasing as $n \to \infty$. Set $\tau_\infty = \lim_{n\to\infty} \tau_n$, whence $\tau_\infty \leq \tau_e$ a.s. In other words, to complete the proof, we need to show that

$$\tau_\infty = \infty \quad \text{a.s.}$$

We define a $C^2$-function $H : \mathbb{R}_+ \to \mathbb{R}_+$ for some $\phi \in (0, 1]$ by

$$H(x) = x^\phi - 1 - \phi \log(x). \tag{5.10}$$

From the operator (5.3) and by Assumption 5.2.1, we obtain

$$\mathcal{I}H(x, y, t, i) \leq \alpha_{-1}(i)\phi x^{\phi-2} - \alpha_0(i)\phi x^{\phi-1} + \alpha_1(i)\phi x^\phi - \alpha_2(i)\phi x^{\rho+\phi-1} - \alpha_{-1}(i)\phi x^{-2}$$

$$+\alpha_0(i)\phi x^{-1} - \alpha_1(i)\phi + \alpha_2(i)\phi x^{\rho-1} + \frac{\sigma^2}{2}\phi(\phi-1)x^{\phi+2\theta-2} + \frac{\sigma^2}{2}\phi x^{2\theta-2}.$$

By the Jump-diffusion operator in (5.2), we now have

$$\mathcal{L}H(x,y,t,i) \le \mathcal{I}H(x,y,t,i) + \lambda((1+\alpha_3(i))^{\phi} - 1)x^{\phi} - \lambda\phi\log(1+\alpha_3(i)).$$

For $\phi \in (0,1]$ and by Assumption 5.2.3, we observe $-\alpha_{-1}(i)\phi x^{-2}$ dominates and tends to $-\infty$ for small $x$ and for large $x$, $-\alpha_2(i)\phi x^{\rho+\phi-1}$ dominates and tends to $-\infty$. So there exists a constant $\mathcal{K}_0$ such that

$$\mathcal{L}H(x,y,t,i) \le \mathcal{K}_0.$$

So for any arbitrary $t_1 \ge 0$, the Itô formula gives us

$$\mathbb{E}[H(x(\tau_n \wedge t_1))] \le H(\xi(0)) + \mathcal{K}_0 t_1.$$

It then follows

$$\mathbb{P}(\tau_n \le t_1) \le \frac{H(\xi(0)) + \mathcal{K}_0 t_1}{H(1/n) \wedge H(n)}.$$

This implies $\mathbb{P}(\tau_\infty \le t_1) = 0$ and consequently, we must have

$$\mathbb{P}(\tau_\infty = \infty) = 1$$

as the required assertion. The proof is thus complete. □

### 5.3.2 Moment bounds

The following lemma shows the moment of the exact solution $x(t)$ to SDDE (5.1) is finite.

**Lemma 5.3.2.** *Let Assumptions 5.2.1 and 5.2.3 hold. Then for any $p > 2 \vee (\rho - 1)$,* *the solution* $x(t)$ *to* SDDE (5.1) *satisfies*

$$\sup_{0\le t<\infty}\Big(\mathbb{E}|x(t)|^p\Big)\le c_1, \tag{5.11}$$

*and consequently*

$$\sup_{0\le t<\infty}\Big(\mathbb{E}|\frac{1}{x(t)}|^p\Big)\le c_2, \tag{5.12}$$

*where* $c_1$ *and* $c_2$ *are constants.*

*Proof.* For every sufficiently large integer $n$, we define the stopping time by

$$\tau_n=\inf\{t\ge 0: x(t)\notin(1/n,n)\}.$$

We also define a Lyapunov function $H\in C^{2,1}(\mathbb{R}_+\times\mathbb{R}_+;\mathbb{R}_+)$ by $H(x,t)=e^t x^p$. By Assumption 5.2.1, we apply (5.1) to obtain

$$\mathcal{L}H(x,y,t,i)\le\mathcal{I}H(x,y,t,i)+\lambda e^t x^p[(1+\alpha_3(i))^p-1],$$

where

$$\mathcal{I}H(x,y,t,i)\le e^t\Big[x^p+px^{p-2}(\alpha_{-1}(i)-\alpha_0(i)x+\alpha_1(i)x^2-\alpha_2(i)x^{\rho+1}+\frac{(p-1)}{2}\sigma^2x^{2\theta})\Big].$$

Apparently, by Assumption 5.2.3, $-p\alpha_2(i)x^{\rho+p-1}$ dominates and tends to $-\infty$ for large $x$. So there exists a constant $\mathcal{K}_1$ such that

$$\mathcal{L}H(x,y,t,i)\le\mathcal{K}_1e^t.$$

By the the Itô formula, we have

$$\mathbb{E}[e^{t\wedge\tau_n}|x(t\wedge\tau_n)|^p]\le|\xi(0)|^p+\mathcal{K}_1e^t.$$

Applying the Fatou lemma and letting $n \to \infty$ yields

$$\mathbb{E}|x(t)|^p < \frac{|\xi(0)|^p + \mathcal{K}_1 e^t}{e^t}$$

and consequently, we obtain (5.14) as the required assertion. Moreover, by applying the operator (5.1) to the Lyapunov function $H(x,t) = e^t/x^p$, we compute

$$\mathcal{L}H(x,y,t,i) \leq \mathcal{I}H(x,y,t,i) + \lambda e^t x^{-p}[(1+\alpha_3(i))^{-p} - 1],$$

where Assumption 5.2.1 has been used and

$$\mathcal{I}H(x,y,t,i) \leq e^t \Big[x^{-p} - p x^{-(p+2)}(\alpha_{-1}(i) - \alpha_0(i)x + \alpha_1(i)x^2 - \alpha_2(i)x^{\rho+1} + \frac{(p+1)}{2}\sigma^2 x^{2\theta})\Big].$$

For $p > 2 \vee (\rho - 1)$, we note $-\alpha_{-1}(i)px^{-(p+2)}$ dominates and tends to $-\infty$ for small $x$. Moreover, we also note $p\alpha_2(i)x^{\rho-p-1}$ dominates and tends to 0 for large $x$. We then find a constant $\mathcal{K}_2$ such that

$$\mathcal{L}H(x,y,t,i) \leq \mathcal{K}_2 e^t.$$

So from the Itô formula, we can apply the Fatou lemma and let $n \to \infty$ to arrive at (5.12). □

## 5.4 Numerical method

Under this section, we recall the truncated EM method and apply it for convergent approximation of SDDE (5.1). To start with, let us also impose the following important condition on the initial data.

**Assumption 5.4.1.** *There is a pair of constant $\mathcal{K}_3 > 0$ and $\Upsilon \in (0,1]$ such that for all $-\tau \leq s \leq t \leq 0$, the initial data $\xi$ satisfies*

$$|\xi(t) - \xi(s)| \leq \mathcal{K}_3 |t-s|^{\Upsilon}. \tag{5.13}$$

We also need the following lemmas (see [39]).

**Lemma 5.4.2.** *For any $R > 0$, there exists a constant $K_R > 0$ such that the coefficient terms of* SDDE (5.1) *satisfy*

$$|f(x,i) - f(\bar{x},i)| \vee |g(x) - g(\bar{x})| \vee |h(x,i) - h(\bar{x},i)| \leq K_R|x - \bar{x}|, \tag{5.14}$$

$\forall x, \bar{x} \in [\frac{1}{R}, R]$ *and* $i \in \mathcal{S}$.

**Lemma 5.4.3.** *Let Assumptions 5.2.1 and 5.2.3 hold. For any $p \geq 2$, there exists $\mathcal{K}_4 = \mathcal{K}_4(p) > 0$ such that the coefficient terms of* SDDE (5.1) *satisfy*

$$xf(x,i) + \frac{p-1}{2}|\varphi(y,i)g(x)|^2 \leq \mathcal{K}_4(1 + |x|^2), \tag{5.15}$$

$\forall (x,y,i) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{S}$.

The truncated EM scheme for SDDE (5.1) is now defined in the following subsection.

### 5.4.1 The truncated EM method

Let extend the volatility function $\varphi(y,i)$ and the jump term $h(x,i)$ from $\mathbb{R}_+$ to $\mathbb{R}$ by setting $\varphi(y,i) = \varphi(0,i)$ and $h(x,i) = 0$ for $x < 0$. These extensions do not in any way affect above conditions and results. To define the truncated EM scheme for SDDE (5.1), we first choose a strictly increasing continuous function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\mu(r) \to \infty$ as $r \to \infty$ and

$$\sup_{1/r \leq x \leq r} (|f(x,i)| \vee g(x)) \leq \mu(r), \quad \forall r > 1. \tag{5.16}$$

Let $\mu^{-1}$ be the inverse function of $\mu$ and $\psi : (0,1) \to \mathbb{R}_+$ a strictly decreasing function such that

$$\lim_{\Delta \to 0} \psi(\Delta) = \infty \text{ and } \Delta^{1/4}\psi(\Delta) \leq 1, \quad \forall \Delta \in (0,1]. \tag{5.17}$$

Find $\Delta^* \in (0,1)$ such that $\mu^{-1}(\psi(\Delta^*)) > 1$ and $f(x,i) > 0$ for $0 < x < \Delta^*$. For a given step size $\Delta \in (0, \Delta^*)$, let us define the truncated functions

$$f_\Delta(x,i) = f\Big(1/\mu^{-1}(\psi(\Delta)) \vee (x \wedge \mu^{-1}(\psi(\Delta))), i\Big), \quad \forall (x,i) \in \mathbb{R} \times \mathcal{S}$$

and

$$g_\Delta(x) = \begin{cases} g\Big(x \wedge \mu^{-1}(\psi(\Delta))\Big), & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases}$$

Then for $x \in [1/\mu^{-1}(\psi(\Delta)), \mu^{-1}(\psi(\Delta))]$, we get

$$\begin{aligned} |f_\Delta(x,i)| = |f(x,i)| &\leq \max_{1/\mu^{-1}(\psi(\Delta)) \leq z \leq \mu^{-1}(\psi(\Delta))} |f(z,i)| \\ &\leq \mu(\mu^{-1}(\psi(\Delta))) = \psi(\Delta) \end{aligned}$$

and

$$g_\Delta(x) \leq \mu(\mu^{-1}(\psi(\Delta))) = \psi(\Delta).$$

We easily see that

$$|f_\Delta(x,i)| \vee g_\Delta(x) \leq \psi(\Delta), \quad \forall (x,i) \in \mathbb{R} \times \mathcal{S}. \tag{5.18}$$

The following lemma confirms that $f_\Delta$ and $g_\Delta$ nicely reproduce (5.15).

**Lemma 5.4.4.** *Let Assumption 5.2.1 and 5.2.3 hold. Then, for all $\Delta \in (0, \Delta^*)$ and $p \geq 2$, the truncated functions satisfy*

$$x f_\Delta(x,i) + \frac{p-1}{2} |\varphi(y,i) g_\Delta(x)|^2 \leq \mathcal{K}_5(1 + |x|^2) \tag{5.19}$$

$\forall (x,y,i) \in \mathbb{R} \times \mathbb{R} \times \mathcal{S}$, *where $\mathcal{K}_5$ is independent of $\Delta$. Consult [39].*

Let also recall the following useful lemma.

**Lemma 5.4.5.** *Given $\Delta > 0$, let $r_\Delta^k = r_\Delta(k\Delta)$ for $k \geq 0$. Then $\{r_\Delta^k, k = 0, 1, 2, \cdots\}$ is a discrete Markov chain with the one-step transition probability matrix*

$$P(\Delta) = (P_{ij}(\Delta))_{N \times N} = e^{\Delta\Gamma}.$$

The discrete Markovian chain $\{r_\Delta^k, k = 0, 1, 2, \cdots\}$ can be simulated as follows: compute the one-step transition probability matrix

$$P(\Delta) = (P_{ij}(\Delta))_{N \times N} = e^{\Delta\Gamma}.$$

Let $r_\Delta^0 = i_0$ and generate a random number $\varpi$ which is uniformly distributed in $[0, 1]$. Define

$$r_\Delta^1 = \begin{cases} i_1 & \text{if } i_1 \in \mathcal{S} - \{N\} \text{ such that } \sum_{j=i}^{i_1-1} P_{i_0, j(\Delta)} \leq \varpi_1 < \sum_{j=i}^{i_1} P_{i_0, j(\Delta)} \\ N, & \text{if } \sum_{j=i}^{N-1} P_{i_0, j(\Delta)} \leq \varpi_1, \end{cases}$$

where we set $\sum_{j=i}^{0} P_{i_0, j(\Delta)} = 0$ as usual. Generate independently a new random number $\varpi_2$ which is again uniformly distributed in $[0, 1]$ and then define

$$r_\Delta^2 = \begin{cases} i_2 & \text{if } i_2 \in \mathcal{S} - \{N\} \text{ such that } \sum_{j=i}^{i_2-1} P_{r_\Delta^1, j(\Delta)} \leq \varpi_2 < \sum_{j=i}^{i_2} P_{r_\Delta^1, j(\Delta)} \\ N, & \text{if } \sum_{j=i}^{N-1} P_{r_\Delta^1, j(\Delta)} \leq \varpi_2, \end{cases}$$

Repeating this procedure, a trajectory of $\{r_\Delta^k, k = 0, 1, 2, \cdots\}$ can be generated.

Given the discrete Markovian chain scheme, we now form the discrete-time truncated EM scheme for SDDE (5.1) by first letting $T > 0$ be arbitrarily fixed and the step size $\Delta \in (0, \Delta^*]$ be a fraction of $\tau$. Define $\Delta = \tau/M$ for some positive integer $M$. Define $t_k = k\Delta$ for $k = -M, -(M-1), \cdots, 0, 1, 2, \cdots$, set $X_\Delta(t_k) = \xi(t_k)$ for $k = -M, -(M-1), \cdots, 0$ and then compute

$$X_\Delta(t_{k+1}) = X_\Delta(t_k) + f_\Delta(X_\Delta(t_k), r_\Delta(t_k))\Delta + \varphi(X_\Delta(t_{k-M}), r_\Delta(t_k))g_\Delta(X_\Delta(t_k))\Delta B_k$$

$$+ h(X_\Delta(t_k), r_\Delta(t_k))\Delta N_k \tag{5.20}$$

for $k = 0, 1, 2, \cdots$, where $\Delta B_k = B(t_{k+1}) - B(t_k)$ and $\Delta N_k = N(t_{k+1}) - N(t_k)$. We have two versions of the continuous-time truncated EM solutions. The first one is defined by

$$\bar{x}_\Delta(t) = \sum_{k=-M}^{\infty} X_\Delta(t_k) 1_{[t_k, t_{k+1})}(t) \text{ and } \bar{r}_\Delta(t) = \sum_{k=-M}^{\infty} r_\Delta(t_k) 1_{[t_k, t_{k+1})}(t). \tag{5.21}$$

These are the continuous-time step processes $\bar{x}_\Delta(t)$ and $\bar{r}_\Delta(t)$ on $t \geq -\tau$, where $1_{[t_k, t_{k+1})}$ is the indicator function on $[t_k, t_{k+1})$. The second one is the continuous-time continuous process $x_\Delta(t)$ on $t \geq -\tau$ defined by setting $x_\Delta(t) = \xi(t)$ for $t \in [-\tau, 0]$ while for $t \geq 0$

$$\begin{aligned} x_\Delta(t) = \xi(0) &+ \int_0^t f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))ds + \int_0^t \varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) g_\Delta(\bar{x}_\Delta(s^-))dB(s) \\ &+ \int_0^t h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))dN(s). \end{aligned} \tag{5.22}$$

Apparently $x_\Delta(t)$ is an Itô process on $t \geq 0$ satisfying Itô differential

$$\begin{aligned} dx_\Delta(t) = &f_\Delta(\bar{x}_\Delta(t^-), \bar{r}_\Delta(t))dt + \varphi(\bar{x}_\Delta((t-\tau)^-), \bar{r}_\Delta(t)) g_\Delta(\bar{x}_\Delta(t^-))dB(t) \\ &+ h(\bar{x}_\Delta(t^-), \bar{r}_\Delta(t))dN(t). \end{aligned} \tag{5.23}$$

We observe $x_\Delta(t_k) = \bar{x}_\Delta(t_k) = X_\Delta(t_k)$, for all $k = -M, -(M-1), \cdots$.

## 5.5 Numerical properties

Let us now investigate the numerical properties of the truncated EM scheme. In the sequel, we let

$$k(t) = [t/\Delta]\Delta,$$

for any $t \in [0, T]$, where $[t/\Delta]$ denotes the integer part of $t/\Delta$. The following lemma affirms $x_\Delta(t)$ and $\bar{x}_\Delta(t)$ are close to each other in the strong sense.

### 5.5.1 Moment bounds

**Lemma 5.5.1.** *Let Assumption 5.2.1 hold. Then for any fixed $\Delta \in (0, \Delta^*]$, we have for $p \in [2, \infty)$*

$$\mathbb{E}\Big(|x_\Delta(t) - \bar{x}_\Delta(t)|^p \Big| \mathcal{F}_{k(t)}\Big) \leq \mathfrak{C}_1 \Big(\Delta^{p/2}(\psi(\Delta))^p + \Delta\Big)|\bar{x}_\Delta(t)|^p \tag{5.24}$$

*$\forall t \geq 0$, where $\mathfrak{C}_1$ denotes positive generic constants dependent only on $p$ and may change between occurrences.*

*Proof.* Fix any $\Delta \in (0, \Delta^*)$ and $t \in [0, T]$. Then for $p \in [2, \infty)$, we derive

$$\begin{aligned}
\mathbb{E}\Big(|x_\Delta(t) - \bar{x}_\Delta(t)|^p \Big| \mathcal{F}_{k(t)}\Big) &\leq 3^{p-1}\Big(\mathbb{E}\big(\Big|\int_{k(t)}^t f_\Delta(\bar{x}_\Delta(s), \bar{r}_\Delta(s))ds\Big|^p \Big| \mathcal{F}_{k(t)}\big) \\
&+ \mathbb{E}\big(\Big|\int_{k(t)}^t \varphi(\bar{x}_\Delta((s-\tau)), \bar{r}_\Delta(s))g_\Delta(\bar{x}_\Delta(s))dB(s)\Big|^p \Big| \mathcal{F}_{k(t)}\big) \\
&+ \mathbb{E}\big(\Big|\int_{k(t)}^t h(\bar{x}_\Delta(s), \bar{r}_\Delta(s))dN(s)\Big|^p \Big| \mathcal{F}_{k(t)}\big)\Big) \\
&\leq 3^{p-1}\Big(\Delta^{p-1}\mathbb{E}\big(\int_{k(t)}^t |f_\Delta(\bar{x}_\Delta(s), \bar{r}_\Delta(s))|^p ds \Big| \mathcal{F}_{k(t)}\big) \\
&+ C_p\Delta^{(p-2)/2}\mathbb{E}\big(\int_{k(t)}^t |\varphi(\bar{x}_\Delta((s-\tau)), \bar{r}_\Delta(s))g_\Delta(\bar{x}_\Delta(s))|^p ds \Big| \mathcal{F}_{k(t)}\big) \\
&+ \mathbb{E}\big(\Big|\int_{k(t)}^t h(\bar{x}_\Delta(s), \bar{r}_\Delta(s))dN(s)\Big|^p \Big| \mathcal{F}_{k(t)}\big)\Big) \\
&\leq 3^{p-1}\Big(\Delta^{p-1}\Delta(\psi(\Delta))^p + C_p\Delta^{(p-2)/2}\Delta(\sigma\psi(\Delta))^p \\
&+ \mathbb{E}\big(\Big|\int_{k(t)}^t h(\bar{x}_\Delta(s), \bar{r}_\Delta(s))dN(s)\Big|^p \Big| \mathcal{F}_{k(t)}\big)\Big).
\end{aligned}$$

Recalling the characteristic function's argument $\mathbb{E}|\Delta N_k|^p \leq \bar{C}\Delta$, $\forall \Delta \in (0, \Delta^*)$, in [45], we note

$$
\begin{aligned}
&\mathbb{E}\Big(|x_\Delta(t)-\bar{x}_\Delta(t)|^p\Big|\mathcal{F}_{k(t)}\Big)\\
&\le 3^{p-1}\Big(\Delta^{p-1}\Delta(\psi(\Delta))^p+C_p\Delta^{(p-2)/2}\Delta(\sigma\psi(\Delta))^p+|h(\bar{x}_\Delta(t),r(t))|^p\mathbb{E}|\Delta N_k|^p\Big)\\
&\le 3^{p-1}\Big(\Delta^{p-1}\Delta(\psi(\Delta))^p+C_p\Delta^{(p-2)/2}\Delta(\sigma\psi(\Delta))^p+\bar{C}\alpha_3(i)^p|\bar{x}_\Delta(t)|^p\Delta\Big),
\end{aligned}
$$

where $h(\cdot,\cdot)$ and $\bar{C}>0$ are independent of $N_k$ and $\Delta$ respectively. We now have

$$
\begin{aligned}
&\mathbb{E}\Big(|x_\Delta(t)-\bar{x}_\Delta(t)|^p\Big|\mathcal{F}_{k(t)}\Big)\\
&\le 3^{p-1}(1\vee C_p\sigma^p\vee\bar{C}\alpha_3(i)^p)\Big(\Delta^{p/2}(\psi(\Delta))^p+|\bar{x}_\Delta(t)|^p\Delta\Big)\\
&\le \mathfrak{C}_1\Big(\Delta^{p/2}(\psi(\Delta))^p+\Delta\Big)|\bar{x}_\Delta(t)|^p,
\end{aligned}
$$

where

$$\mathfrak{C}_1=3^{p-1}(1\vee C_p\sigma^p\vee\bar{C}\alpha_3^p)\text{ and }\alpha_3=\max_{i\in\mathcal{S}}\alpha_3(i).$$

Moreover, for $p\in(0,2)$, we obtain from the Jensen inequality that

$$
\begin{aligned}
\mathbb{E}\Big(|x_\Delta(t)-\bar{x}_\Delta(t)|^p\Big|\mathcal{F}_{k(t)}\Big)&\le\Big[\mathfrak{C}_1\Big(\Delta(\psi(\Delta))^2+\Delta\Big)|\bar{x}_\Delta(t)|^p\Big]^{p/2}\\
&\le 2^{p/2-1}\mathfrak{C}_1^{p/2}\Big(\Delta^{p/2}(\psi(\Delta))^p+\Delta^{p/2}\Big)(|\bar{x}_\Delta(t)|^p)^{p/2}\\
&\le\mathfrak{C}_2\Big(\Delta^{p/2}(\psi(\Delta))^p\Big)|\bar{x}_\Delta(t)|^p, \qquad (5.25)
\end{aligned}
$$

where $\mathfrak{C}_2=2^{p/2}\mathfrak{C}_1^{p/2}$. The proof is now complete. □

The following lemma reveals the finite moment of the truncated EM solutions.

**Lemma 5.5.2.** *Let Assumptions 5.2.1 and 5.2.3 hold. Then for any* $p\ge3$

$$\sup_{0\le\Delta\le\Delta^*}\sup_{0\le t\le T}(\mathbb{E}|x_\Delta(t)|^p)\le c_3,\quad\forall T>0, \qquad (5.26)$$

*where* $c_3:=c_3(T,p,\mathcal{K}_5,\xi)$ *and may change between occurrences.*

*Proof.* Fix any $\Delta \in (0, \Delta^*)$ and $T \geq 0$. For $t \in [0, T]$, we obtain from (5.2) and Lemma 5.4.4

$$\begin{aligned}
\mathbb{E}|x_\Delta(t)|^p - |\xi(0)|^p \leq\ & \mathbb{E}\int_0^t p|x_\Delta(s^-)|^{p-2}\Big(\bar{x}_\Delta(s^-) f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) \\
& + \frac{p-1}{2}|\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) g_\Delta(\bar{x}_\Delta(s^-))|^2\Big) ds \\
& + \mathbb{E}\int_0^t p|x_\Delta(s^-)|^{p-2}(x_\Delta(s^-) - \bar{x}_\Delta(s^-)) f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) ds \\
& + \lambda \mathbb{E}\Big(\int_0^t |x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|^p - |x_\Delta(s^-)|^p\Big) ds \\
\leq\ & \mathcal{J}_{11} + \mathcal{J}_{12} + \mathcal{J}_{13},
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{J}_{11} &= \mathbb{E}\int_0^t \mathcal{K}_5 p|x_\Delta(s^-)|^{p-2}(1 + |\bar{x}_\Delta(s^-)|^2) ds, \\
\mathcal{J}_{12} &= \mathbb{E}\int_0^t p|x_\Delta(s^-)|^{p-2}\Big(x_\Delta(s^-) - \bar{x}_\Delta(s^-)\Big) f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) ds
\end{aligned}$$

and

$$\mathcal{J}_{13} = \lambda \mathbb{E}\Big(\int_0^t |x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|^p - |x_\Delta(s^-)|^p\Big) ds.$$

The Young inequality gives us

$$\begin{aligned}
\mathcal{J}_{11} &\leq \mathcal{K}_5 \int_0^t \Big((p-2)\mathbb{E}|x_\Delta(s^-)|^p + 2^p(1 + \mathbb{E}|\bar{x}_\Delta(s^-)|^p)\Big) ds \\
&\leq \nu_1 \int_0^t (1 + \mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p) ds,
\end{aligned}$$

where $\nu_1 = \mathcal{K}_5(2^p \vee (p-2))$. By the triangle inequality, we have for $p \geq 3$

$$\mathcal{J}_{12} \leq p\mathbb{E}\int_0^t \Big(|x_\Delta(s^-) - \bar{x}_\Delta(s^-)| + |\bar{x}_\Delta(s^-)|\Big)^{p-2}$$

$$\begin{aligned}
&\times |x_\Delta(s^-) - \bar{x}_\Delta(s^-)||f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|ds \\
&\le 2^{(p-3)} p\mathbb{E} \int_0^t \Big( |x_\Delta(s^-) - \bar{x}_\Delta(s^-)|^{p-2} + |\bar{x}_\Delta(s^-)|^{p-2} \Big) \\
&\times |x_\Delta(s^-) - \bar{x}_\Delta(s^-)||f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|ds \\
&= \mathcal{J}_{121} + \mathcal{J}_{122},
\end{aligned}$$

where

$$\mathcal{J}_{121} = 2^{(p-3)} p\mathbb{E} \int_0^t |\bar{x}_\Delta(s^-)|^{p-2} |x_\Delta(s^-) - \bar{x}_\Delta(s)||f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|ds$$

and

$$\mathcal{J}_{122} = 2^{(p-3)} p\mathbb{E} \int_0^t |x_\Delta(s^-) - \bar{x}_\Delta(s^-)|^{p-1} |f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|ds.$$

We now obtain from (5.18) and (5.25)

$$\begin{aligned}
\mathcal{J}_{121} &\le 2^{(p-3)} p \int_0^t \mathbb{E}\Big\{ |\bar{x}_\Delta(s)|^{p-2} |f_\Delta(\bar{x}_\Delta(s), \bar{r}_\Delta(s))| \mathbb{E}\Big( |x_\Delta(s) - \bar{x}_\Delta(s)| \mathcal{F}_{k(s)} \Big) \Big\} ds \\
&\le 2^{(p-3)} p \mathfrak{C}_2 (\psi(\Delta)) \Delta^{1/2} (\psi(\Delta)) \int_0^t \mathbb{E}\Big\{ |\bar{x}_\Delta(s)| (|\bar{x}_\Delta(s)|^{p-2}) \Big\} ds \\
&\le 2^{(p-3)} p \mathfrak{C}_2 (\psi(\Delta)) \Delta^{1/2} (\psi(\Delta)) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\le 2^{(p-3)} \mathfrak{C}_2 (\psi(\Delta))^2 \Delta^{1/2} \int_0^t \Big( 1 + (p-1)\mathbb{E}|\bar{x}_\Delta(s)|^p \Big) ds \\
&\le \nu_2 + \nu_3 \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^p ds, \qquad (5.27)
\end{aligned}$$

where $\nu_2 = 2^{(p-3)} \mathfrak{C}_2 T$, $\nu_3 = 2^{(p-3)} \mathfrak{C}_2 (p-1)$ and $[(\psi(\Delta))\Delta^{1/4}]^2 \le 1$. We also have from (5.18)

$$\mathcal{J}_{122} \le 2^{(p-3)} p\psi(\Delta) \int_0^t \mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^{p-1} ds. \qquad (5.28)$$

We clearly observe that for $p \geq 3$ and $\varkappa \in (0, 1/4]$, $p\varkappa \leq (p-1)/2$ and hence

$$\Delta^{(p-1)/2-\varkappa p} \leq 1. \tag{5.29}$$

So for $p \geq 3$ and $\varkappa = 1/4$, we get from (5.28), Lemma 5.5.1, (5.29) and the Young's inequality that

$$\begin{aligned}
\mathcal{J}_{122} &\leq 2^{(p-3)} p \mathfrak{C}_1 \Big( \Delta^{(p-1)/2} (\psi(\Delta))^{p-1} (\psi(\Delta)) + \Delta(\psi(\Delta)) \Big) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\leq 2^{(p-3)} p \mathfrak{C}_1 \Big( \Delta^{(p-1)/2} (\psi(\Delta))^{p} + \Delta(\psi(\Delta)) \Big) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\leq 2^{(p-3)} p \mathfrak{C}_1 \Big( \Delta^{(p-2)/4} + \Delta(\psi(\Delta)) \Big) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p-1} ds \\
&\leq 2^{(p-2)} \mathfrak{C}_1 \int_0^t \Big( 1 + (p-1)\mathbb{E}|\bar{x}_\Delta(s)|^{p} \Big) ds \\
&\leq \nu_4 + \nu_5 \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p} ds,
\end{aligned}$$

where $\nu_4 = 2^{(p-2)}\mathfrak{C}_1 T$ and $\nu_5 = 2^{(p-2)}\mathfrak{C}_1(p-1)$. We now combine $\mathcal{J}_{121}$ and $\mathcal{J}_{122}$ to have

$$\begin{aligned}
\mathcal{J}_{12} &\leq \nu_2 + \nu_4 + (\nu_3 + \nu_5) \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p} ds \\
&\leq \nu_6 + \nu_7 \int_0^t \mathbb{E}|\bar{x}_\Delta(s)|^{p} ds,
\end{aligned}$$

where $\nu_6 = \nu_2 + \nu_4$ and $\nu_7 = \nu_3 + \nu_5$. Also we estimate $\mathcal{J}_{13}$ as

$$\begin{aligned}
\mathcal{J}_{13} &= \lambda \mathbb{E}\Big( \int_0^t |x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|^p - |x_\Delta(s^-)|^p \Big) ds \\
&\leq \lambda \mathbb{E}\Big( \int_0^t 2^{p-1}|x_\Delta(s^-)|^p + 2^{p-1}|h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))|^p - |x_\Delta(s^-)|^p \Big) ds \\
&\leq \lambda \mathbb{E}\Big( \int_0^t (2^{p-1} - 1)|x_\Delta(s^-)|^p + 2^{p-1}\alpha_3(i)^p |\bar{x}_\Delta(s^-)|^p \Big) ds
\end{aligned}$$

$$\leq \nu_8 \int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p) ds,$$

where $\nu_8 = \lambda((2^{p-1} - 1) \vee 2^{p-1}\alpha_3^p)$ and $\alpha_3 = \max_{i \in \mathcal{S}} \alpha_3(i)$. We combine $\mathcal{J}_{11}$, $\mathcal{J}_{12}$ and $\mathcal{J}_{13}$ to have

$$\begin{aligned}\mathbb{E}|x_\Delta(t)|^p &\leq |\xi(0)|^p + (\nu_1 T + \nu_6) \\ &\quad + \int_0^t \Big( (\nu_1 + \nu_8)\mathbb{E}|x_\Delta(s)|^p + (\nu_1 + \nu_7 + \nu_8)\mathbb{E}|\bar{x}_\Delta(s)|^p \Big) ds \\ &\leq \nu_9 + 2\nu_{10} \int_0^t \sup_{0 \leq u \leq s} \Big( \mathbb{E}|x_\Delta(u)|^p \Big) ds,\end{aligned}$$

where

$$\nu_9 = |\xi(0)|^p + \nu_1 T + \nu_6$$

and

$$\nu_{10} = (\nu_1 + \nu_8) \vee (\nu_1 + \nu_7 + \nu_8).$$

As this holds for any $t \in [0, T]$, we then have

$$\sup_{0 \leq u \leq t} (\mathbb{E}|x_\Delta(u)|^p) \leq \nu_9 + 2\nu_{10} \int_0^t \sup_{0 \leq u \leq s} \Big( \mathbb{E}|x_\Delta(u)|^p \Big) ds.$$

The Gronwall inequality gives us

$$\sup_{0 \leq u \leq T} (\mathbb{E}|x_\Delta(u)|^p) \leq c_3,$$

where $c_3 = \nu_9 e^{2\nu_{10} T}$ is independent of $\Delta$. The proof is thus complete. □

### 5.5.2 Finite time strong convergence

**Lemma 5.5.3.** *Suppose Assumptions 5.2.1, 5.2.3 and 5.4.1 hold and fix $T > 0$. Then for any $\epsilon \in (0,1)$, there exists a pair $n = n(\epsilon) > 0$ and $\bar{\Delta} = \bar{\Delta}(\epsilon) > 0$ such that*

$$\mathbb{P}(\varsigma_{\Delta,n} \le T) \le \epsilon \tag{5.30}$$

*as long as $\Delta \in (0, \bar{\Delta}]$, where*

$$\varsigma_{\Delta,n} = \inf\{t \in [0,T] : x_\Delta(t) \notin (1/n, n)\} \tag{5.31}$$

*is a stopping time.*

*Proof.* Let $H(\cdot)$ be the Lyapunov function in (5.11). Then for $t \in [0,T]$, the Itô formula gives us

$$\begin{aligned}\mathbb{E}(H(x_\Delta(t \wedge \varsigma_{\Delta,n})) - H(\xi(0))) = \mathbb{E}\int_0^{t\wedge\varsigma_{\Delta,n}} &\Big[H_x(x_\Delta(s^-))f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) \\ &+ \frac{1}{2}H_{xx}(x_\Delta(s^-))\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 g_\Delta(\bar{x}_\Delta(s^-))^2 \\ &+ \lambda\Big(H(x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))) - H(x_\Delta(s^-))\Big)\Big]ds.\end{aligned}$$

For $s \in [0, t \wedge \varsigma_{\Delta,n}]$, we can expand to have

$$\begin{aligned}&H_x(x_\Delta(s^-))f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) + \frac{1}{2}H_{xx}(x_\Delta(s^-))\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 g_\Delta(\bar{x}_\Delta(s^-))^2 \\ &+ \lambda\Big(H(x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))) - H(x_\Delta(s^-))\\ &\le \mathcal{L}(x_\Delta(s^-), x_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) + \mathcal{J}_{21} + \mathcal{J}_{22} + \mathcal{J}_{23},\end{aligned}$$

where $\mathcal{L}H$ is the operator in (5.2), which now takes the form

$$\begin{aligned}&\mathcal{L}(x_\Delta(s^-), x_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) \\ &= H_x(x_\Delta(s^-))f_\Delta(x_\Delta(s^-), \bar{r}_\Delta(s)) \\ &+ \frac{1}{2}H_{xx}\Big(x_\Delta(s^-))\varphi(x_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 g_\Delta(x_\Delta(s^-))^2\end{aligned}$$

$$+ \lambda(H(x_\Delta(s^-) + h(x_\Delta(s^-), \bar{r}_\Delta(s)) - H(x_\Delta(s^-)))\Big)$$

with $H$ independent of $t$ and

$$\begin{aligned}
\mathcal{J}_{21} &= H_x(x_\Delta(s^-))\Big(f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - f_\Delta(x_\Delta(s^-), \bar{r}_\Delta(s))\Big), \\
\mathcal{J}_{22} &= \frac{1}{2} H_{xx}(x_\Delta(s^-))\Big(\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 g_\Delta(\bar{x}_\Delta(s^-))^2 \\
&\quad - \varphi(x_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 g_\Delta(x_\Delta(s^-))^2\Big), \\
\mathcal{J}_{23} &= \lambda\Big(H(x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))) - H(x_\Delta(s^-) + h(x_\Delta(s^-), \bar{r}_\Delta(s)))\Big).
\end{aligned}$$

By Assumptions 5.2.1 and 5.2.3, we can find a constant $\mathcal{K}_6$ such that

$$\mathcal{L}(x_\Delta(s^-), x_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) \leq \mathcal{K}_6. \tag{5.32}$$

Recalling from the definition of $f_\Delta$ and $g_\Delta$, we note for $s \in [0, t \wedge \varsigma_{\Delta,n}]$

$$f_\Delta(x_\Delta(s^-), \bar{r}_\Delta(s)) = f(x_\Delta(s^-), \bar{r}_\Delta(s)) \text{ and } g_\Delta(x_\Delta(s^-)) = g(x_\Delta(s^-)).$$

So for $s \in [0, t \wedge \varsigma_{\Delta,n}]$, we obtain from Lemma 5.4.2 that

$$\mathcal{J}_{21} \leq K_n H_x(x_\Delta(s^-))|\bar{x}_\Delta(s^-) - x_\Delta(s^-)|.$$

Moreover, for $s \in [0, t \wedge \varsigma_{\Delta,n}]$ and any $\bar{x}_\Delta(s^-), x_\Delta(s^-) \in [1/n, n]$, we note from (5.16) that

$$g(\bar{x}_\Delta(s^-)) \vee g(x_\Delta(s^-)) \leq \mu(n).$$

So for $s \in [0, t \wedge \varsigma_{\Delta,n}]$, we now obtain from Assumptions 5.2.1 and 5.2.2, and Lemma 5.4.2

$$\mathcal{J}_{22} \leq \frac{1}{2} H_{xx}(x_\Delta(s^-))\Big[g(x_\Delta(s^-))^2\Big(\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 - \varphi(x_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2\Big)$$

$$
\begin{aligned}
&+ \varphi(x_\Delta((s-\tau)^-), \bar{r}_\Delta(s))^2 \Big( g(\bar{x}_\Delta(s^-))^2 - g(x_\Delta(s^-))^2 \Big) \Big] \\
&\leq H_{xx}(x_\Delta(s^-)) \Big[ L_n \sigma(\mu(n))^2 |\bar{x}_\Delta(s-\tau)^- - x_\Delta(s-\tau)^-| \\
&+ K_n \sigma^2 \mu(n)) |\bar{x}_\Delta(s^-) - x_\Delta(s^-)| \Big].
\end{aligned}
$$

Also for $s \in [0, t \wedge \varsigma_{\Delta,n}]$, we obtain from the Lyapunov function in (5.11) and the mean value theorem that

$$
\begin{aligned}
\mathcal{J}_{23} \leq \lambda \Big[ &\Big( x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) \Big)^\phi - \phi \log \Big( x_\Delta(s^-) + h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) \Big) \\
&- \Big( x_\Delta(s^-) + h(x_\Delta(s^-), \bar{r}_\Delta(s)) \Big)^\phi + \phi \log \Big( x_\Delta(s^-) + h(x_\Delta(s^-), \bar{r}_\Delta(s)) \Big) \Big] \\
&\leq \lambda \Big[ \Big( x_\Delta(s^-) + \alpha_3(i) \bar{x}_\Delta(s^-) \Big)^\phi - \Big( x_\Delta(s^-) + \alpha_3(i) x_\Delta(s^-) \Big)^\phi \\
&+ \phi \log \Big( x_\Delta(s^-) + \alpha_3(i) x_\Delta(s^-) \Big) - \phi \log \Big( x_\Delta(s^-) + \alpha_3(i) \bar{x}_\Delta(s^-) \Big) \Big] \\
&\leq n\lambda |x_\Delta(s^-) + \alpha_3(i) \bar{x}_\Delta(s^-) - x_\Delta(s^-) - \alpha_3(i) x_\Delta(s^-)| \\
&+ n\lambda\phi |x_\Delta(s^-) + \alpha_3(i) x_\Delta(s^-) - x_\Delta(s^-) - \alpha_3(i) \bar{x}_\Delta(s^-)| \\
&\leq n\lambda\alpha_3(1+\phi) |\bar{x}_\Delta(s^-) - x_\Delta(s^-)|,
\end{aligned}
$$

where $\alpha_3 = \max_{i \in \mathcal{S}} \alpha_3(i)$. Combining $\mathcal{J}_{21}$, $\mathcal{J}_{22}$ and $\mathcal{J}_{23}$ with (5.32), we now have

$$
\begin{aligned}
&\mathbb{E}(H(x_\Delta(t \wedge \varsigma_{\Delta,n})) \\
&\leq H(\xi(0))) + \mathcal{K}_6 T + K_n \mathbb{E} \int_0^{t \wedge \varsigma_{\Delta,n}} H_x(x_\Delta(s^-)) |\bar{x}_\Delta(s^-) - x_\Delta(s^-)| ds \\
&+ \mathbb{E} \int_0^{t \wedge \varsigma_{\Delta,n}} H_{xx}(x_\Delta(s^-)) \Big[ L_n \sigma(\mu(n))^2 |\bar{x}_\Delta(s-\tau)^- - x_\Delta(s-\tau)^-| \\
&+ K_n \sigma^2 \mu(n) |\bar{x}_\Delta(s^-) - x_\Delta(s^-)| \Big] ds + n\lambda\alpha_3(1+\phi) \mathbb{E} \int_0^{t \wedge \varsigma_{\Delta,n}} |\bar{x}_\Delta(s^-) - x_\Delta(s^-)| ds \\
&\leq K7 + \mathcal{K}_8 \mathbb{E} \int_{-\tau}^0 |\xi([s/\Delta]\Delta) - \xi(s)| ds + (\mathcal{K}_8 + \mathcal{K}_9) \int_0^T \mathbb{E} \Big( \mathbb{E} |\bar{x}_\Delta(s) - x_\Delta(s)|^p \Big| \mathcal{F}_{k(s)} \Big)^{1/p} ds
\end{aligned}
$$

where

$$\mathcal{K}_7 = H(\xi(0)) + \mathcal{K}_6 T,\ \mathcal{K}_8 = \max_{1/n \le x \le n}\{H_{xx}(x)\sigma(\mu(n))^2 L_n\}$$

and

$$\mathcal{K}_9 = \max_{1/n \le x \le n}\{H_x(x)K_n + H_{xx}(x)\sigma^2\mu(n)K_n + n\lambda\alpha_3(1+\phi)\}.$$

So by Lemmas 5.5.1 and 5.5.1, we now obtain

$$\begin{aligned}\mathbb{E}(H(x_\Delta(t\wedge\varsigma_{\Delta,n}))) &\le \mathcal{K}_7 + \mathcal{K}_3\mathcal{K}_8 T\Delta^\Upsilon + (\mathcal{K}_8+\mathcal{K}_9)\mathfrak{C}_1^{1/p}\Big(\Delta^{p/2}(\psi(\Delta))^p + \Delta\Big)^{1/p} \\ &\quad\times\int_0^T (\sup_{0\le u\le s}(\mathbb{E}|\bar{x}_\Delta(u)|^p))^{1/p} ds \\ &\le \mathcal{K}_7 + \mathcal{K}_3\mathcal{K}_8 T\Delta^\Upsilon + (\mathcal{K}_8+\mathcal{K}_9)\mathfrak{C}_1^{1/p}\Big(\Delta^{p/2}(\psi(\Delta))^p + \Delta\Big)^{1/p} c_3^{1/p} T.\end{aligned}$$

This implies

$$\mathbb{P}(\varsigma_{\Delta,n}\le T) \le \frac{\mathcal{K}_7 + \mathcal{K}_3\mathcal{K}_8 T\Delta^\Upsilon + (\mathcal{K}_8+\mathcal{K}_9)\mathfrak{C}_1^{1/p}\Big(\Delta^{p/2}(\psi(\Delta))^p + \Delta\Big)^{1/p} c_3^{1/p} T}{H(1/n)\wedge H(n)}. \tag{5.33}$$

For any $\epsilon \in (0,1)$, we may select sufficiently large $n$ such that

$$\frac{\mathcal{K}_7}{H(1/n)\wedge H(n)} \le \frac{\epsilon}{2} \tag{5.34}$$

and sufficiently small of each step size $\Delta \in (0, \bar{\Delta}]$ such that

$$\frac{\mathcal{K}_3\mathcal{K}_8 T\Delta^\Upsilon + (\mathcal{K}_8+\mathcal{K}_9)\mathfrak{C}_1^{1/p}\Big(\Delta^{p/2}(\psi(\Delta))^p + \Delta\Big)^{1/p} c_3^{1/p} T}{H(1/n)\wedge H(n)} \le \frac{\epsilon}{2}. \tag{5.35}$$

We can now combine (5.34) and (5.35) to obtain the required assertion. □

The following lemma shows the truncated EM solutions converges strongly to the exact solution in finite time.

**Lemma 5.5.4.** *Let Assumptions 5.2.1, 5.2.2, 5.2.3 and 5.4.1 hold. Set*

$$\vartheta_{\Delta,n} = \tau_n \wedge \varsigma_{\Delta,n},$$

*where* $\tau_n$ *and* $\varsigma_{\Delta,n}$ *are* (5.9) *and* (5.31). *Then for any* $p \geq 2$, $T > 0$, *we have for any sufficiently large* $n$ *and any* $\Delta \in (0, \Delta^*]$,

$$\mathbb{E}\Big(\sup_{0\leq t\leq T} |x_\Delta(t\wedge\vartheta_{\Delta,n}) - x(t\wedge\vartheta_{\Delta,n})|^p\Big) \leq \mathcal{C}\Big((\Delta + o(\Delta))(\psi(\Delta))^p) \vee \Delta^{p(1/4\wedge\Upsilon\wedge 1/p)}\Big) \tag{5.36}$$

*where* $\mathcal{C}$ *is a constant independent of* $\Delta$ *and consequently,*

$$\lim_{\Delta\to 0} \mathbb{E}\Big(\sup_{0\leq t\leq T} |x_\Delta(t\wedge\vartheta_{\Delta,n}) - x(t\wedge\vartheta_{\Delta,n})|^p\Big) = 0. \tag{5.37}$$

*Proof.* For $t_1 \in [0, T]$, we obtain from (5.1) and (5.23) that

$$\mathbb{E}\Big(\sup_{0\leq t\leq t_1} |x_\Delta(t\wedge\vartheta_{\Delta,n}) - x(t\wedge\vartheta_{\Delta,n})|^p\Big) \leq \mathcal{J}_{31} + \mathcal{J}_{32} + \mathcal{J}_{33}. \tag{5.38}$$

where

$$\begin{aligned}
\mathcal{J}_{31} &= 3^{p-1}\mathbb{E}\Big(|\int_0^{t_1\wedge\vartheta_{\Delta,n}} [f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - f(x(s^-), r(s))]ds|^p\Big),\\
\mathcal{J}_{32} &= 3^{p-1}\mathbb{E}\Big(\sup_{0\leq t\leq t_1} |\int_0^{t\wedge\vartheta_{\Delta,n}} [\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s))g_\Delta(\bar{x}_\Delta(s^-))\\
&\quad - \varphi(x((s-\tau)^-), r(s))g(x(s^-))]dB(s)|^p\Big)
\end{aligned}$$

and

$$\mathcal{J}_{33} = 3^{p-1}\mathbb{E}\Big(\sup_{0\leq t\leq t_1} |\int_0^{t\wedge\vartheta_{\Delta,n}} [h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - h(x(s^-), r(s))]dN(s)|^p\Big).$$

By the Hölder inequality, we compute

$$\mathcal{J}_{31} \leq \mathcal{J}_{311} + \mathcal{J}_{312},$$

where

$$\mathcal{J}_{311} = 3^{p-1}2^{p-1}T^{p-1}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |f_\Delta(\bar{x}_\Delta(s^-), r(s)) - f(x(s^-), r(s))|^p ds$$

and

$$\mathcal{J}_{312} = 3^{p-1}2^{p-1}T^{p-1}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - f_\Delta(\bar{x}(s^-), r(s))|^p ds.$$

It is clear from the definition of the truncated function $f_\Delta$ that

$$f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) = f(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s))$$

for $s \in [0, t_1 \wedge \vartheta_{\Delta,n}]$. So by Lemma 5.4.2,

$$\begin{aligned}
\mathcal{J}_{311} &= 3^{p-1}2^{p-1}T^{p-1}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |f(\bar{x}_\Delta(s^-), r(s)) - f(x(s^-), r(s))|^p ds \\
&\leq 3^{p-1}2^{p-1}T^{p-1}K_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.
\end{aligned}$$

Let $n = [T/\Delta]$ be the integer part of $T/\Delta$. Then

$$\begin{aligned}
\mathcal{J}_{312} &= 3^{p-1}2^{p-1}T^{p-1}\mathbb{E}\int_0^{T} |f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - f_\Delta(\bar{x}(s^-), r(s))|^p ds \\
&= 3^{p-1}2^{p-1}T^{p-1}\sum_{k=0}^{n}\mathbb{E}\int_{t_k}^{t_{k+1}} |f_\Delta(\bar{x}_\Delta(t_k), r(t_k)) - f_\Delta(\bar{x}(t_k), r(s))|^p ds,
\end{aligned}$$

with $t_{n+1}$ now set to be $T$. We now have from (5.18)

$$\mathcal{J}_{312} \leq 3^{p-1}2^{2(p-1)}T^{p-1}\sum_{k=0}^{n}\mathbb{E}\int_{t_k}^{t_{k+1}} [|f_\Delta(\bar{x}_\Delta(t_k), r(t_k))|^p$$

$$
\begin{aligned}
&+|f_\Delta(\bar{x}(t_k), r(s))|^p]1_{\{r(s)\neq r(t_k)\}}ds \\
&\le 3^{p-1}2^{2(p-1)}T^{p-1}\sum_{k=0}^{n}\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\mathbb{E}[((\psi(\Delta))^p + (\psi(\Delta))^p)1_{\{r(s)\neq r(t_k)\}}|r(t_k)]\Big]ds \\
&= 3^{p-1}2^{2(p-1)}T^{p-1}\sum_{k=0}^{n}\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\mathbb{E}[2(\psi(\Delta))^p|r(t_k)]\mathbb{E}[1_{\{r(s)\neq r(t_k)\}}|r(t_k)]\Big]ds,
\end{aligned}
\tag{5.39}
$$

where we use the fact that $\bar{x}(s)$ and $1_{r(s)\neq r(t_k)}$ are conditionally independent with respect to the $\sigma-$algebra generated by $r(t_k)$ in the last step. By the Markov property, we compute

$$
\begin{aligned}
\mathbb{E}[1_{\{r(s)\neq r(t_k)\}}|r(t_k)] &= \sum_{i\in\mathcal{S}} 1_{\{r(t_k)=i\}}\mathbb{P}(r(s)\neq i|r(t_k)=i) \\
&= \sum_{i\in\mathcal{S}} 1_{\{r(t_k)=i\}}\sum_{i\neq j}(\gamma_{ij}(s-t_k)+o(s-t_k)) \\
&\le (\max_{1\le i\le N}(-\gamma_{ij})\Delta + o(\Delta))\sum_{i\in\mathcal{S}}1_{\{r(t_k)=i\}} \\
&\le \bar{c}_1\Delta + o(\Delta).
\end{aligned}
\tag{5.40}
$$

where $\bar{c}_1 = \max_{1\le i\le N}(-\gamma_{ij})$. By Lemma 5.5.2, we note

$$
\begin{aligned}
\mathbb{E}\int_{t_k}^{t_{k+1}}|f_\Delta(\bar{x}_\Delta(t_k), r(t_k)) - f(\bar{x}(t_k), r(s))|^p ds &\le 2(\bar{c}_1\Delta + o(\Delta))\int_{t_k}^{t_{k+1}}(\psi(\Delta))^p ds \\
&\le 2(\bar{c}_1\Delta + o(\Delta))\Delta(\psi(\Delta))^p.
\end{aligned}
$$

This implies

$$
\mathbb{E}\int_0^T |f_\Delta(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - f(\bar{x}(s^-), r(s))|^p ds \le 2(\bar{c}_1\Delta + o(\Delta))(\psi(\Delta))^p
$$

and consequently

$$\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |f_\Delta(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - f(\bar{x}(s^-),r(s))|^p ds \le 2(\bar{c}_1\Delta + o(\Delta))(\psi(\Delta))^p.$$

Substituting this into $\mathcal{J}_{312}$ yields

$$\mathcal{J}_{312} \le 3^{p-1}2^{2p-1}T^{p-1}(\bar{c}_1\Delta + o(\Delta))(\psi(\Delta))^p.$$

We then combine $\mathcal{J}_{311}$ and $\mathcal{J}_{312}$ to obtain

$$\mathcal{J}_{31} \le \bar{c}_2(\bar{c}_1\Delta + o(\Delta))(\psi(\Delta))^p + \bar{c}_3\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds,$$

where

$$\bar{c}_2 = 3^{p-1}2^{2p-1}T^{p-1}$$

and

$$\bar{c}_3 = 3^{p-1}2^{p-1}T^{p-1}K_n^p.$$

Also by the Hölder and Burkholder-Davis Gundy inequalities, we have

$$\begin{aligned}\mathcal{J}_{32} \le 3^{p-1}T^{\frac{p-2}{2}}C_1(p)\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} \Big(&|\varphi(\bar{x}_\Delta((s-\tau)^-),\bar{r}_\Delta(s))g_\Delta(\bar{x}_\Delta(s^-))\\
&- \varphi(x((s-\tau)^-),r(s))g_\Delta(\bar{x}_\Delta(s^-)) + \varphi(x((s-\tau)^-),r(s))g_\Delta(\bar{x}_\Delta(s^-))\\
&- \varphi(x((s-\tau)^-),r(s))g(x(s^-))|^p\Big)ds\\
&\le \mathcal{J}_{321} + \mathcal{J}_{322},\end{aligned}$$

where

$$\mathcal{J}_{321} = 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} g_\Delta(\bar{x}_\Delta(s^-))^p|\varphi(\bar{x}_\Delta((s-\tau)^-),\bar{r}_\Delta(s))$$

$$-\varphi(x((s-\tau)^-),r(s))|^p ds$$

and

$$\mathcal{J}_{322}=2^{p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}}\varphi(x((s-\tau)^-),r(s))^p|g_\Delta(\bar{x}_\Delta(s^-))-g(x(s^-))|^p ds.$$

where $C_1(p)$ is a positive constant. For $s\in[0,t_1\wedge\vartheta_{\Delta,n}]$, we note from (5.16) that $\bar{x}_\Delta(s^-)\in[1/n,n]$ and $g_\Delta(\bar{x}_\Delta(s^-))\le\mu(n)$. So we now have

$$\begin{aligned}\mathcal{J}_{321}\le\ &2^{p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}}|\varphi(\bar{x}_\Delta((s-\tau)^-),\bar{r}_\Delta(s))\\&-\varphi(x((s-\tau)^-),r(s))|^p ds\\&\le\mathcal{J}_{323}+\mathcal{J}_{324},\end{aligned}$$

where

$$\begin{aligned}\mathcal{J}_{323}=\ &2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}}|\varphi(\bar{x}_\Delta((s-\tau)^-),r(s))\\&-\varphi(x((s-\tau)^-),r(s))|^p ds\end{aligned}$$

and

$$\begin{aligned}\mathcal{J}_{324}=\ &2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}}|\varphi(\bar{x}_\Delta((s-\tau)^-),\bar{r}_\Delta(s))\\&-\varphi(\bar{x}((s-\tau)^-),r(s))|^p ds\end{aligned}$$

By Assumption 5.2.2, we obtain

$$\mathcal{J}_{323}\le 2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^pL_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}}|\bar{x}_\Delta((s-\tau)^-)-x((s-\tau)^-)|^p ds.$$

Also as before, we compute

$$\begin{aligned}\mathcal{J}_{324} &= 2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\mathbb{E}\int_0^T|\varphi(\bar{x}_\Delta((s-\tau)^-),\bar{r}_\Delta(s))\\&-\varphi(\bar{x}((s-\tau)^-),r(s))|^pds\\&=2^{p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sum_{k=0}^{n}\mathbb{E}\int_{t_k}^{t_{k+1}}|\varphi(\bar{x}_\Delta((t_k-\tau)^-),\bar{r}_\Delta(t_k))\\&-\varphi(\bar{x}((t_k-\tau)^-),r(s))|^pds,\end{aligned}$$

where $n$ is the usual integer part of $T/\Delta$ with $t_{n+1}$ set to be $T$. By elementary inequality,

$$\begin{aligned}\mathcal{J}_{324} &\le 2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sum_{k=0}^{n}\mathbb{E}\int_{t_k}^{t_{k+1}}[|\varphi(\bar{x}_\Delta((t_k-\tau)^-),\bar{r}_\Delta(t_k))|^p\\&+|\varphi(\bar{x}((t_k-\tau)^-),r(s))|^p1_{\{r(s)\neq r(t_k)\}}]ds\\&=2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sum_{k=0}^{n}\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\mathbb{E}[|\varphi(\bar{x}_\Delta((t_k-\tau)^-),\bar{r}_\Delta(t_k))|^p\\&+|\varphi(\bar{x}((t_k-\tau)^-),r(s))|^p1_{\{r(s)\neq r(t_k)\}}|r(t_k)]\Big]ds\\&=2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sum_{k=0}^{n}\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\mathbb{E}[|\varphi(\bar{x}_\Delta((t_k-\tau)^-),\bar{r}_\Delta(t_k))|^p\\&+|\varphi(\bar{x}((t_k-\tau)^-),r(s))|^p]\mathbb{E}[1_{\{r(s)\neq r(t_k)\}}|r(t_k)]\Big]ds\end{aligned}$$

We note from (5.40) that

$$\mathbb{E}[1_{\{r(s)\neq r(t_k)\}}|r(t_k)]\le\bar{c}_1\Delta+o(\Delta).$$

By Assumption 5.2.1, we have

$$\mathbb{E}\int_{t_k}^{t_{k+1}}|\varphi(\bar{x}_\Delta((t_k-\tau)^-),\bar{r}_\Delta(t_k))-\varphi(\bar{x}((t_k-\tau)^-),r(s))|^pds$$

$$\leq 2(\bar{c}_1\Delta + o(\Delta)) \int_{t_k}^{t_{k+1}} \sigma^p ds$$
$$\leq 2\sigma^p(\bar{c}_1\Delta + o(\Delta))\Delta.$$

This means by Assumption 5.2.1, we have

$$\mathbb{E}\int_0^T |\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) - \varphi(\bar{x}((s-\tau)^-), r(s))|^p ds \leq 2\sigma^p(\bar{c}_1\Delta + o(\Delta))$$

and hence,

$$\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\varphi(\bar{x}_\Delta((s-\tau)^-), \bar{r}_\Delta(s)) - \varphi(\bar{x}((s-\tau)^-), r(s))|^p ds \leq 2\sigma^p(\bar{c}_1\Delta + o(\Delta)).$$

Inserting this into $\mathcal{J}_{324}$ yields

$$\mathcal{J}_{324} \leq 2^{2p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sigma^p(\bar{c}_1\Delta + o(\Delta)).$$

We obtain from $\mathcal{J}_{323}$ and $\mathcal{J}_{324}$

$$\mathcal{J}_{321} \leq 2^{2p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sigma^p(\bar{c}_1\Delta + o(\Delta))$$
$$+ 2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^pL_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta((s-\tau)^-) - x((s-\tau)^-)|^p ds.$$

Moreover, by Assumption 5.2.1 and Lemma 5.4.2

$$\mathcal{J}_{322} \leq 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^pK_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.$$

Combining $\mathcal{J}_{321}$ and $\mathcal{J}_{322}$, we have

$$\mathcal{J}_{32} \leq \bar{c}_4(\bar{c}_1\Delta + o(\Delta)) + \bar{c}_5\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta((s-\tau)^-) - x((s-\tau)^-)|^p ds$$
$$+ \bar{c}_6\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds,$$

where

$$\bar{c}_4 = 2^{2p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^p\sigma^p,$$
$$\bar{c}_5 = 2^{2(p-1)}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^pL_n^p$$

and

$$\bar{c}_6 = 2^{p-1}3^{p-1}T^{\frac{p-2}{2}}C_1(p)(\mu(n))^pK_n^p.$$

Furthermore, by elementary inequality

$$\begin{aligned}\mathcal{J}_{33} &= 3^{p-1}\mathbb{E}\Big(\sup_{0\le t\le t_1}|\int_0^{t\wedge\vartheta_{\Delta,n}}[h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(x(s^-),r(s))]d\widetilde{N}(s)\\ &\quad + \lambda\int_0^{t\wedge\vartheta_{\Delta,n}}[h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(x(s^-),r(s))]ds|^p\Big)\\ &\le \mathcal{J}_{331} + \mathcal{J}_{332},\end{aligned}$$

where

$$\mathcal{J}_{331} = 2^{p-1}3^{p-1}\mathbb{E}\Big(\sup_{0\le t\le t_1}|\int_0^{t\wedge\vartheta_{\Delta,n}}[h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(x(s^-),r(s))]d\widetilde{N}(s)|^p\Big)$$

and

$$\mathcal{J}_{332} = 2^{p-1}3^{p-1}\lambda^p\mathbb{E}\Big(\sup_{0\le t\le t_1}|\int_0^{t\wedge\vartheta_{\Delta,n}}[h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(x(s^-),r(s))]ds|^p\Big).$$

By the Doob martingale inequality and martingale isometry, we have

$$\begin{aligned}\mathcal{J}_{331} &\le 2^{p-1}3^{p-1}\lambda^{\frac{p}{2}}C_2(p)\mathbb{E}\Big(\sup_{0\le t\le t_1}|\int_0^{t\wedge\vartheta_{\Delta,n}}|h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(x(s^-),r(s))|^2d\widetilde{N}(s)\Big)^{\frac{p}{2}}\\ &\le 2^{p-1}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}}|h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(x(s^-),r(s))|^pds\\ &\le \mathcal{J}_{333} + \mathcal{J}_{334},\end{aligned}$$

where

$$\mathcal{J}_{333} = 2^{p-1}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-), r(s)) - h(x(s^-), r(s))|^p ds$$

and

$$\mathcal{J}_{334} = 2^{p-1}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - h(\bar{x}_\Delta(s^-), r(s))|^p ds$$

and $C_2(p)$ is a positive constant. By Lemma 5.4.2,

$$\mathcal{J}_{333} \leq 2^{p-1}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)K_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.$$

We also compute

$$\begin{aligned}
\mathcal{J}_{334} &= 2^{p-1}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\mathbb{E}\int_0^{T} |h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - h(\bar{x}_\Delta(s^-), r(s))|^p ds \\
&= 2^{p-1}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\sum_{k=0}^{n}\mathbb{E}\int_{t_k}^{t_{k+1}} |h(\bar{x}_\Delta(t_k), \bar{r}_\Delta(t_k)) - h(\bar{x}_\Delta(t_k), r(s))|^p ds \\
&\leq 2^{2(p-1)}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\sum_{k=0}^{n}\mathbb{E}\int_{t_k}^{t_{k+1}}[|h(\bar{x}_\Delta(t_k), \bar{r}_\Delta(t_k))|^p \\
&+ |h(\bar{x}_\Delta(t_k), r(s))|^p 1_{\{r(s)\neq r(t_k)\}} ds \\
&\leq 2^{2(p-1)}3^{p-1}\lambda^{p/2}T^{\frac{p-2}{2}}C_2(p)\sum_{k=0}^{n}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\mathbb{E}[|h(\bar{x}_\Delta(t_k), \bar{r}_\Delta(t_k))|^p \\
&+ |h(\bar{x}_\Delta(t_k), r(s))|^p|r(t_k))]\mathbb{E}[1_{\{r(s)\neq r(t_k)\}}|r(t_k)]\Big] ds,
\end{aligned}$$

where $n$, as usual, is the integer part of $T/\Delta$ with $t_{n+1}$ set to be $T$. By Lemma 5.5.2 and (5.40)

$$\mathbb{E}\int_{t_k}^{t_{k+1}} |h(\bar{x}_\Delta(t_k), \bar{r}_\Delta(t_k)) - h(\bar{x}_\Delta(t_k), r(s))|^p ds$$

$$
\begin{aligned}
&\leq (\bar{c}_1\Delta + o(\Delta)) \int_{t_k}^{t_{k+1}} 2\alpha_3(i)\mathbb{E}|\bar{x}_\Delta(t_k)|^p ds \\
&\leq 2\alpha_3(\bar{c}_1\Delta + o(\Delta))\Delta,
\end{aligned}
$$

where $\alpha_3 = \max_{i\in\mathcal{S}} \alpha_3(i)$. Consequently, we have

$$
\mathbb{E}\int_0^T |h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - h(\bar{x}_\Delta(s^-), r(s))|^p ds \leq 2\alpha_3(\bar{c}_1\Delta + o(\Delta))
$$

and then,

$$
\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - h(\bar{x}_\Delta(s^-), r(s))|^p ds \leq 2\alpha_3(\bar{c}_1\Delta + o(\Delta)). \tag{5.41}
$$

We substitute this into $\mathcal{J}_{334}$ to get

$$
\mathcal{J}_{334} \leq 2^{2p-1}3^{p-1}C_2(p)\alpha_3\lambda^{p/2}T^{\frac{p-2}{2}}(\bar{c}_1\Delta + o(\Delta)).
$$

It then follows from $\mathcal{J}_{333}$ and $\mathcal{J}_{334}$ that

$$
\begin{aligned}
\mathcal{J}_{331} \leq\ & 2^{2p-1}3^{p-1}C_2(p)\alpha_3\lambda^{p/2}T^{\frac{p-2}{2}}(\bar{c}_1\Delta + o(\Delta)) \\
&+ 2^{p-1}3^{p-1}C_2(p)K_n^p\lambda^{p/2}T^{\frac{p-2}{2}}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.
\end{aligned}
$$

By the the Hölder inequality,

$$
\begin{aligned}
\mathcal{J}_{332} &\leq 2^{p-1}3^{p-1}\lambda^p T^{p-1}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-), \bar{r}_\Delta(s)) - h(x(s^-), r(s))|^p ds \\
&= \mathcal{J}_{335} + \mathcal{J}_{336},
\end{aligned}
$$

where

$$
\mathcal{J}_{335} = 2^{p-1}3^{p-1}\lambda^p T^{p-1}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-), r(s)) - h(x(s^-), r(s))|^p ds
$$

and

$$\mathcal{J}_{336} = 2^{p-1}3^{p-1}\lambda^p T^{p-1}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(\bar{x}_\Delta(s^-),r(s))|^p ds.$$

So by Lemma 5.4.2,

$$\mathcal{J}_{335} = 2^{p-1}3^{p-1}\lambda^p T^{p-1}K_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.$$

Apparently, we see from (5.41) that

$$\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |h(\bar{x}_\Delta(s^-),\bar{r}_\Delta(s)) - h(\bar{x}_\Delta(s^-),r(s))|^p ds \le 2\alpha_3(\bar{c}_1\Delta + o(\Delta)).$$

This implies

$$\mathcal{J}_{336} \le 2^{p-1}3^{p-1}2\alpha_3\lambda^p T^{p-1}(\bar{c}_1\Delta + o(\Delta)).$$

We now have from $\mathcal{J}_{335}$ and $\mathcal{J}_{336}$

$$\begin{aligned}\mathcal{J}_{332} \le\ & 2^{p-1}3^{p-1}2\alpha_3\lambda^p T^{p-1}(\bar{c}_1\Delta + o(\Delta)) \\ & + 2^{p-1}3^{p-1}\lambda^p T^{p-1}K_n^p\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.\end{aligned}$$

We then combine $\mathcal{J}_{331}$ and $\mathcal{J}_{332}$ to have

$$\begin{aligned}\mathcal{J}_{33} \le\ & \bar{c}_7(\bar{c}_1\Delta + o(\Delta)) + \bar{c}_8\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds \\ & + \bar{c}_9\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds,\end{aligned}$$

where

$$\begin{aligned}\bar{c}_7 &= 2^{2p-1}3^{p-1}C_2(p)\alpha_3\lambda^{p/2}T^{\frac{p-2}{2}} + 2^{p-1}3^{p-1}2\alpha_3\lambda^p T^{p-1}, \\ \bar{c}_8 &= 2^{p-1}3^{p-1}C_2(p)K_n^p\lambda^{p/2}T^{\frac{p-2}{2}}\end{aligned}$$

and

$$\bar{c}_9 = 2^{p-1}3^{p-1}\lambda^p T^{p-1} K_n^p.$$

Substituting $\mathcal{J}_{31}$, $\mathcal{J}_{32}$ and $\mathcal{J}_{33}$ into (5.38), we get

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\vartheta_{\Delta,n}) - x(t\wedge\vartheta_{\Delta,n})|^p\Big)\\
&\le \bar{c}_2(\bar{c}_1\Delta + o(\Delta))(\psi(\Delta))^p + \bar{c}_4(\bar{c}_1\Delta + o(\Delta)) + \bar{c}_7(\bar{c}_1\Delta + o(\Delta))\\
&+ \bar{c}_5\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta((s-\tau)^-) - x((s-\tau)^-)|^p ds\\
&+ \bar{c}_6\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds\\
&+ \bar{c}_8\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds + \bar{c}_9\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds.
\end{aligned}$$

It then follows that

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\vartheta_{\Delta,n}) - x(t\wedge\vartheta_{\Delta,n})|^p\Big)\\
&\le \bar{c}_{10}(\bar{c}_1\Delta + o(\Delta))(\psi(\Delta))^p\\
&+ \bar{c}_5\mathbb{E}\int_{-\tau}^0 |\xi([s/\Delta]\Delta) - \xi(s)|^p ds + \bar{c}_{11}\mathbb{E}\int_0^{t_1\wedge\vartheta_{\Delta,n}} |\bar{x}_\Delta(s^-) - x(s^-)|^p ds,
\end{aligned}$$

where

$$\bar{c}_{10} = \bar{c}_2 \vee \bar{c}_4 \vee \bar{c}_7$$

and

$$\bar{c}_{11} = \bar{c}_5 \vee \bar{c}_6 \vee \bar{c}_8 \vee \bar{c}_9.$$

By elementary inequality, Assumption 5.4.1 and Lemma 5.5.1

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\vartheta_{\Delta,n})-x(t\wedge\vartheta_{\Delta,n})|^p\Big)\\
&\le \bar{c}_{10}(\bar{c}_1\Delta+o(\Delta))(\psi(\Delta))^p)\\
&+\bar{c}_5\Delta^{p\Upsilon}\tau+2^{p-1}\bar{c}_{11}\int_0^T\mathbb{E}\Big(\mathbb{E}|\bar{x}_\Delta(s)-x_\Delta(s)|^p\Big|\mathcal{F}_{k(s)}\Big)ds\\
&+2^{p-1}\bar{c}_{11}\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\vartheta\Delta,n)-x(t\wedge\vartheta\Delta,n)|^p\Big)ds\\
&\le \bar{c}_{10}(\bar{c}_1\Delta+o(\Delta))(\psi(\Delta))^p)\\
&+\bar{c}_5\Delta^{p\Upsilon}\tau+2^{p-1}\bar{c}_{11}\mathfrak{C}_1\Big(\Delta^{p/2}(\psi(\Delta))^p+\Delta\Big)\int_0^T\mathbb{E}|\bar{x}_\Delta(s)|^pds\\
&+2^{p-1}\bar{c}_{11}\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\vartheta_{\Delta,n})-x(t\wedge\vartheta_{\Delta,n})|^p\Big)ds
\end{aligned}
$$

So by Lemma 5.5.2 and noting that $(\Delta^{1/4}(\psi(\Delta)))^p\le 1$, we now have

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\vartheta_{\Delta,n})-x(t\wedge\vartheta_{\Delta,n})|^p\Big)\\
&\le \bar{c}_{10}(\bar{c}_1\Delta+o(\Delta))(\psi(\Delta))^p)\\
&+\bar{c}_5\Delta^{p\Upsilon}\tau+2^{p-1}\bar{c}_{11}c_3\mathfrak{C}_1\Big([\Delta^{p/4}(\psi(\Delta))^p]\Delta^{p/4}+\Delta^{p(1/p)}\Big)\\
&+2^{p-1}\bar{c}_{11}\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\vartheta_{\Delta,n})-x(t\wedge\vartheta_{\Delta,n})|^p\Big)ds\\
&\le \bar{c}_{10}(\bar{c}_1\Delta+o(\Delta))(\psi(\Delta))^p)\\
&+\Big(\bar{c}_5\tau+2^{p-1}\bar{c}_{11}c_3\mathfrak{C}_1(\Delta^{p/4}(\psi(\Delta))^p+1)\Big)\Delta^{p(1/4\wedge\Upsilon\wedge 1/p)}\\
&+2^{p-1}\bar{c}_{11}\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\vartheta_{\Delta,n})-x(t\wedge\vartheta_{\Delta,n})|^p\Big)ds\\
&\le \bar{c}_{10}(\bar{c}_1\Delta+o(\Delta))(\psi(\Delta))^p)+\bar{c}_{12}\Delta^{p(1/4\wedge\Upsilon\wedge 1/p)}\\
&+\bar{c}_{13}\int_0^{t_1}\mathbb{E}\Big(\sup_{0\le t\le s}|x_\Delta(t\wedge\vartheta_{\Delta,n})-x(t\wedge\vartheta_{\Delta,n})|^p\Big)ds,
\end{aligned}
$$

where $\bar{c}_{12} = \bar{c}_5\tau + 2^p\bar{c}_{11}c_3\mathfrak{C}_1$ and $\bar{c}_{13} = 2^{p-1}\bar{c}_{11}$. The Gronwall inequality gives us

$$\mathbb{E}\Big(\sup_{0\le t\le t_1}|x_\Delta(t\wedge\vartheta_{\Delta,n}) - x(t\wedge\vartheta_{\Delta,n})|^p\Big) \le \mathcal{C}\Big((\Delta + o(\Delta))(\psi(\Delta))^p) \vee \Delta^{p(1/4\wedge\Upsilon\wedge 1/p)}\Big),$$

as the required result in (5.36). where $\mathcal{C} = (\bar{c}_{10}(\bar{c}_1 \vee 1) \vee \bar{c}_{12})e^{\bar{c}_{13}}$. By letting $\Delta \to 0$, we get (5.37). $\square$

The strong convergence of the truncated EM approximate solutions is revealed in the following theorem.

**Theorem 5.5.5.** *Let Assumptions 5.2.1, 5.2.2, 5.2.3 and 5.4.1 hold. Then for any $p \ge 2$, we have*

$$\lim_{\Delta\to 0}\mathbb{E}\Big(\sup_{0\le t\le T}|x_\Delta(t) - x(t)|^p\Big) = 0 \tag{5.42}$$

*and consequently*

$$\lim_{\Delta\to 0}\mathbb{E}\Big(\sup_{0\le t\le T}|\bar{x}_\Delta(t) - x(t)|^p\Big) = 0. \tag{5.43}$$

*Proof.* Here, we only prove the theorem for $p \ge 3$. As for $p \in [2,3)$, it follows directly from the case of $p = 3$ and the Hölder inequality. Let $\tau_n$, $\varsigma_{\Delta,n}$ and $\vartheta_{\Delta,n}$, be the same as before. Set

$$e_\Delta(t) = x_\Delta(t) - x(t).$$

For any arbitrarily $\delta > 0$, the Young inequality gives us

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p\Big) &= \mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\tau_n>T \text{ and } \varsigma_{\Delta,n}>T\}}\Big)\\
&+ \mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\tau_n\le T \text{ or } \varsigma_{\Delta,n}\le T\}}\Big)\\
&\le \mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^p 1_{\{\vartheta_{\Delta,n}>T\}}\Big) + \frac{\delta}{2}\mathbb{E}\Big(\sup_{0\le t\le T}|e_\Delta(t)|^{2p}\Big)\\
&+ \frac{1}{2\delta}\mathbb{P}(\tau_n \le T \text{ or } \varsigma_{\Delta,n} \le T).
\end{aligned} \tag{5.44}$$

So for $p \geq 3$, Lemmas 5.3.2 and 5.5.2 give us

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|e_\Delta(t)|^{2p}\Big) \leq 2^{2p}\mathbb{E}\Big(\sup_{0\leq t\leq T}|x(t)|^{2p} \vee \sup_{0\leq t\leq T}|x_\Delta(t)|^{2p}\Big)$$
$$\leq 2^{2p}(c_1 \vee c_3)^2. \tag{5.45}$$

By Lemmas 5.3.1 and 5.5.3,

$$\mathbb{P}(\vartheta_{\Delta,n} \leq T) \leq \mathbb{P}(\tau_n \leq T) + \mathbb{P}(\varsigma_{\Delta,n} \leq T). \tag{5.46}$$

Also by Lemma 5.5.4,

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|e_\Delta(t)|^p 1_{\{\vartheta_{\Delta,n}>T\}}\Big) \leq \mathcal{C}\Big((\Delta + o(\Delta))(\psi(\Delta))^p) \vee \Delta^{p(1/4\wedge\Upsilon\wedge 1/p)}\Big). \tag{5.47}$$

Substituting (5.45), (5.46) and (5.47) into (5.44) yields

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|e_\Delta(t)|^p\Big) \leq \frac{2^{2p}(c_1 \vee c_3)^2\delta}{2} + \mathcal{C}\Big((\Delta + o(\Delta))(\psi(\Delta))^p) \vee \Delta^{p(1/4\wedge\Upsilon\wedge 1/p)}\Big)$$
$$+ \frac{1}{2\delta}\mathbb{P}(\tau_n \leq T) + \frac{1}{2\delta}\mathbb{P}(\varsigma_{\Delta,n} \leq T).$$

Given $\epsilon \in (0,1)$, we can select $\delta$ so that

$$\frac{2^{2p}(c_1 \vee c_3)^2\delta}{2} \leq \frac{\epsilon}{4}. \tag{5.48}$$

Similarly, for any given $\epsilon \in (0,1)$, there exists $n_o$ so that for $n \geq n_o$, we may select $\delta$ to have

$$\frac{1}{2\delta}\mathbb{P}(\tau_n \leq T) \leq \frac{\epsilon}{4} \tag{5.49}$$

and select $n(\epsilon) \leq n_o$ such that for $\Delta \in (0, \bar{\Delta}]$

$$\frac{1}{2\delta}\mathbb{P}(\varsigma_{\Delta,n} \leq T) \leq \frac{\epsilon}{4}. \tag{5.50}$$

Finally, we may select $\Delta \in (0, \bar{\Delta}]$ sufficiently small for $\epsilon \in (0,1)$ such that

$$\mathcal{C}\Big((\Delta + o(\Delta))(\psi(\Delta))^p) \vee \Delta^{p(1/4 \wedge \Upsilon \wedge 1/p)}\Big) \leq \frac{\epsilon}{4}. \tag{5.51}$$

Combining (5.48), (5.49), (5.50) and (5.51), we get

$$\mathbb{E}\Big(\sup_{0 \leq t \leq T} |x_\Delta(t) - x(t)|^p\Big) \leq \epsilon.$$

as the required result in (5.42). By Lemma 5.5.1, we also get (5.43) by setting $\Delta \to 0$. □

## 5.6 Numerical simulations

Let us now implement the truncated EM (TEM) scheme for SDDE (5.1). To illustrate the strong result established in Theorem 5.5.5, we compare the scheme with the backward EM (BEM) scheme. For justification regarding the choice of BEM scheme and its limitation, we refer the reader to consult [39]. Now consider the following form of SDDE (5.1)

$$\begin{aligned} dx(t) &= f(x(t^-), r(t))dt + \varphi(x((t-\tau)^-)), r(t))g(x(t^-))dB(t) \\ &\quad + h(x(t^-), r(t))dN(t), \end{aligned} \tag{5.52}$$

on $t \geq 0$ with initial values $\xi(t) = 0.2$ and $r_0 = 1$, where $r(t)$ is a Markovian chain defined on the state $\mathcal{S} = \{1, 2\}$ with the generator given by

$$\Gamma = (\gamma)_{2\times 2} = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}. \tag{5.53}$$

Moreover, let

$$f(x,i) = \begin{cases} 0.3x^{-1} - 0.2 + 0.1x - 0.5x^2, & \text{if } i = 1 \\ 0.2x^{-1} - 0.3 + 0.2x - 0.6x^2, & \text{if } i = 2, \end{cases} \tag{5.54}$$

$\forall(x, i) \in (\mathbb{R} \times \mathcal{S})$,

$$g(x) = x^{5/4}, \tag{5.55}$$

$\forall x \in \mathbb{R}$, and

$$h(x, i) = \begin{cases} x, & \text{if } i = 1 \\ 2x, & \text{if } i = 2, \end{cases} \tag{5.56}$$

$\forall(x, i) \in (\mathbb{R} \times \mathcal{S})$. The volatility process $\varphi(y, i)$ is a sigmoid-type function defined as follows:

for $i = 1$,

$$\varphi(y, i) = \begin{cases} \frac{1}{2}\frac{(1+(e^{y}-e^{-y}))}{(e^{y}+e^{-y})}, & \text{if } y \geq 0 \\ \frac{1}{4}, & \text{Otherwise,} \end{cases}$$

and for $i = 2$,

$$\varphi(y, i) = \begin{cases} \frac{1}{4}\frac{(1+(e^{y}-e^{-y}))}{(e^{y}+e^{-y})}, & \text{if } y \geq 0 \\ \frac{1}{8}, & \text{Otherwise,} \end{cases} \tag{5.57}$$

$\forall(y, i) \in (\mathbb{R} \times \mathcal{S})$. Obviously, all the assumptions imposed on $\varphi(y, i)$ are met(see [39]). We clearly see

$$\sup_{1/u \leq x \leq u} (|f(x, i)| \vee g(x)) \leq 3u^2, \quad u \geq 1.$$

We can now set $\mu = 3u^2$ with inverse $\mu^{-1}(u) = (u/3)^{1/2}$.

### 5.6.1 Numerical results

By selecting $\psi(\Delta) = \Delta^{-2/3}$ and a step size $\Delta = 10^{-2}$, we obtain Monte Carlo simulated sample path of $x(t)$ to SDDE (5.52) at a terminal time $T$ in Figure 5.1 using the TEM scheme. The strong convergence between TEM and BEM numerical solutions for the step size $10^{-2}$ is shown in Figure 5.2. In Figure 5.3, we observe the log-log plot of strong errors between TEM and BEM numerical solutions for step sizes $10^{-2}$, $10^{-3}$, $10^{-4}$ and $10^{-5}$ with a reference line of slope 1.0. Do note that Figure 5.2 and Figure 5.3 were obtained without the $\alpha_{-1}(r(t))x(t^{-})^{-1}$ drift term (see [39]).

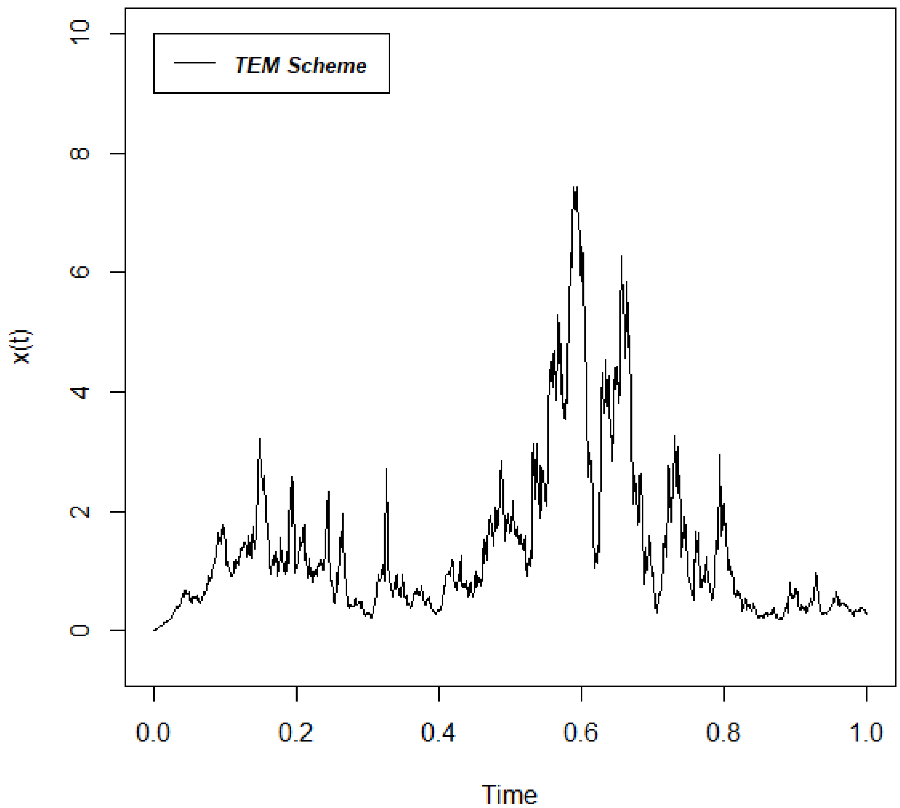


Figure 5.1: Simulated sample path of $x(t)$ when $\Delta = 0.01$

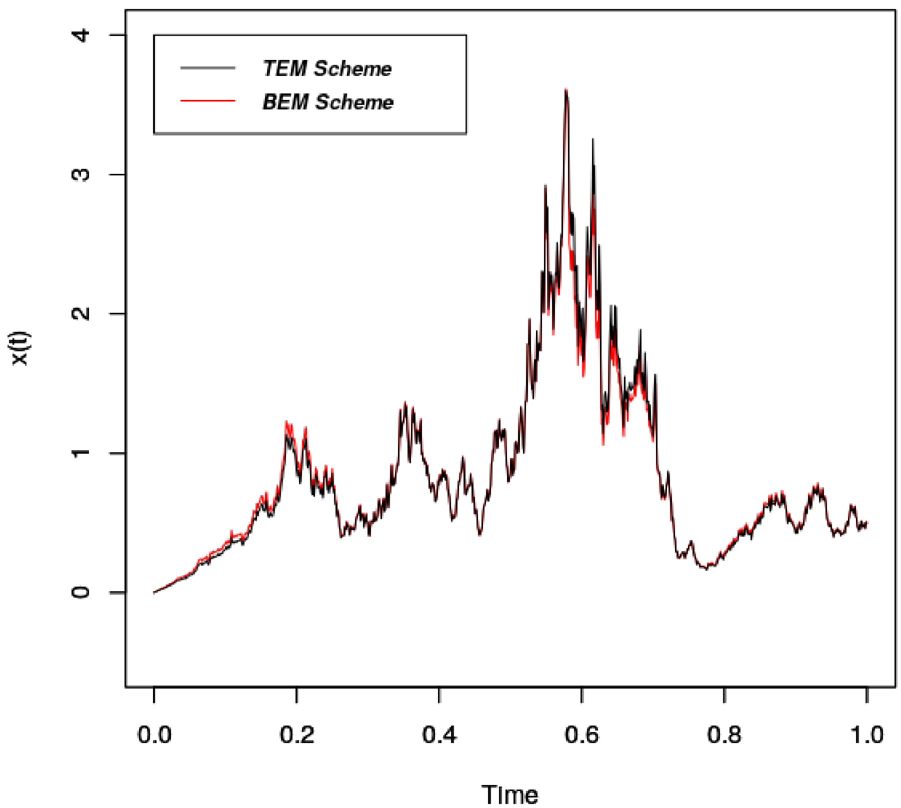


Figure 5.2: Convergence of TEM and BEM solutions when $\Delta = 0.01$

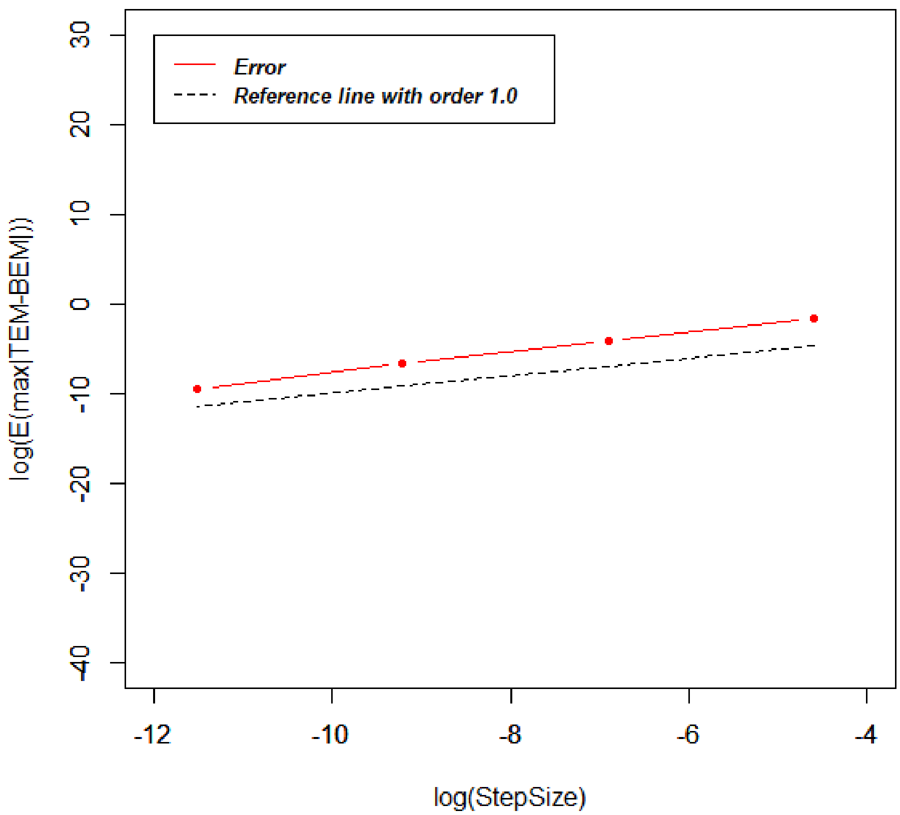


Figure 5.3: Strong errors between TEM and BEM schemes

## 5.7 Summary

In this chapter, we studied Ait-Sahalia-type interest rate model with inherent delayed volatility function and Poisson-driven jumps which is then modulated by finite state Markov chains. Because of the analytical intractability of this model, we employed relevant tools and techniques of the truncated EM scheme developed in the previous chapters to investigate its feasibility from viewpoint of financial applications.

We proved analytical properties such as existence of pathwise unique positive global solutions and finite moments of the exact solution to the proposed model. We drew on the truncated EM scheme to approximate the model. We then moved on to prove numerical properties such as finite moments and the strong convergence order in finite time of the truncated EM approximate solutions under the local Lipschitz condition plus the Khasminskii-type condition. We established $\mathcal{C}((\Delta + o(\Delta))(\psi(\Delta))^p) \vee \Delta^{p(1/4\wedge\Upsilon\wedge 1/p)})$ as the strong pathwise error. We concluded the chapter by performing numerical simulations to validate the theoretical findings.

# Chapter 6

# Applications in finance

In this chapter, we apply the strong convergence results obtained in the previous chapters to evaluate expected payoffs of some financial products such as a bond and a path-dependent barrier option.

### 6.0.1 Bond valuation

Let us apply the results to evaluate the price of a bond for maturity $T$ which is characterised by a unit amount available at time $T$. If we characterise the distribution of the bond price in terms of a chosen dynamics for a stochastic interest rate, we can compute the price at the end of time $T$. This is revealed in the following lemma.

**Lemma 6.0.1.** *If the short-term interest rate dynamics is described by SDDE* (3.1), (4.1) *or* (5.1) *then the payoff of a bond at the end of time $T$ is given by*

$$B(T) = \mathbb{E}\Big[\exp\Big(-\int_0^T x(t)dt\Big)\Big]. \tag{6.1}$$

*Using the step function $\bar{x}(t)$ defined in* (3.27), (4.24) *or* (5.21) *in Chapter 3, 4 or 5 respectively, the approximate payoff based on the truncated EM method becomes*

$$\bar{B}_\Delta(T) = \mathbb{E}\Big[\exp\Big(-\int_0^T |\bar{x}_\Delta(t)|dt\Big)\Big]. \tag{6.2}$$

*Then from Theorem 3.5.6, 4.5.5 or 5.5.5 in Chapter 3, 4 or 5 respectively, it follows that*

$$\lim_{\Delta \to 0} |B(T) - \bar{B}_\Delta(T)| = 0. \tag{6.3}$$

*Proof.* The proof we established here is similar to the one obtained in [15]. In general, it is remarkable to consider

$$\bar{B}_\Delta(T) = \mathbb{E}\Big[\exp\Big( - \Delta \sum_{k=-M}^{\infty} |X_\Delta(t_k)|\Big)\Big]$$

as a natural approximation of equation (6.1) based on (3.26), (4.23) or (5.20). This can be written conveniently as

$$\bar{B}_\Delta(T) = \mathbb{E}\Big[\exp\Big( - \int_0^T |\bar{x}_\Delta(t)|dt\Big)\Big],$$

where $\bar{x}(t)$ is (3.27), (4.24) or (5.21) respectively. Nothing that

$$\exp(-|x|) - \exp(-|y|) \leq |x - y|$$

and by the positivity of $x(t)$, we have

$$\begin{aligned}
|B(T) - \bar{B}_\Delta(T)| &= \mathbb{E}\Big[\exp\Big( - \int_0^T x(t)dt\Big) - \exp\Big( - \int_0^T |\bar{x}_\Delta(t)|dt\Big)\Big] \\
&\leq \mathbb{E}\Big|\int_0^T x(t)dt - \int_0^T \bar{x}_\Delta(t)dt\Big| \\
&\leq \mathbb{E}\int_0^T |x(t) - |\bar{x}_\Delta(t)||dt \\
&\leq \mathbb{E}\int_0^T |x(t) - \bar{x}_\Delta(t)|dt \\
&\leq T \sup_{0,T} \mathbb{E}|x(t) - \bar{x}_\Delta(t)|.
\end{aligned}$$

So by Theorem 3.5.6, 4.5.5 or 5.5.5 implies that $\sup_{0,T} \mathbb{E}|x(t) - \bar{x}_\Delta(t)| \to 0$ and hence, we obtain the required assertion. $\square$

### 6.0.2 Barrier option valuation

Theorem 3.5.6, 4.5.5 or 5.5.5 is very important where SDDE (3.1), (4.1) or (5.1) is to be numerically simulated to calibrate and price a path-dependent option. In this context, we may represent any of these SDDEs as an asset on which an option is to be written on. However, it is established in [15, 17] that the strong convergence of SDE models for asset price guarantees convergence in Monte Carlo simulations for option value. More specifically, suppose that we compute the expected payoff from Monte Carlo simulations based on (3.26), (4.23) or (5.20). Then the following lemma affirms that the expected payoff for a path-dependent barrier option from the truncated EM method will converge to the exact expected payoff as $\Delta \to 0$.

**Lemma 6.0.2.** *Consider a barrier option with a European payoff $P$. Let the asset price be the exact solution $x(T)$ to SDDE* (3.1)*,* (4.1) *or* (5.1) *at expiry date $T$. In this case, the asset price never exceeds a fixed barrier $\boldsymbol{B}$ and pays zero otherwise, where $\Lambda$ is the strike price and $\boldsymbol{B} > \Lambda$. The payoff at expiry date is*

$$P(T) = \mathbb{E}\Big[(x(T) - \Lambda)^+ 1_{\sup_{0\le t\le T} x(t) < \boldsymbol{B})}\Big]. \tag{6.4}$$

*The approximate payoff using the truncated EM scheme defined in* (3.27)*,* (4.24) *or* (5.21) *in Chapter 3, 4 or 5 respectively, becomes*

$$\bar{P}_\Delta(T) = \mathbb{E}\Big[(\bar{x}_\Delta(T) - \Lambda)^+ 1_{\sup_{0\le t\le T} \bar{x}_\Delta(t) < \boldsymbol{B})}\Big]. \tag{6.5}$$

*So from Theorem 3.5.6, 4.5.5 or 5.5.5 in Chapter 3, 4 or 5 respectively, we have*

$$\lim_{\Delta\to 0} |P(T) - \bar{P}_\Delta(T)| = 0. \tag{6.6}$$

The proof of this lemma can be established in the same way as in [15,17]. The proof presented in [15, 17] depends only on the strong convergence properties rather than the structure of the underlying asset price models. This justifies that we can apply the truncated EM method to evaluate a path-dependent barrier option.

# Chapter 7

# Conclusion

In this thesis, we proposed a modified version of the Ait-Sahalia-type interest rate model by incorporating three specifications, namely delayed volatility function, Poisson-driven jump and Markovian switching. These three specifications are introduced to help provide adequate descriptions of interest rates against collective effects of unexpected empirical phenomena such as volatility 'skews' and 'smiles', jump behaviour, market regulatory lapses, economic crisis, financial clashes, political instability, among others. However, the proposed model is not analytically tractable. This motivated the need to perform theoretical and numerical analyses to examine its feasibility from viewpoint of financial applications. We carried out these analyses by splitting the proposed model into three sets of stochastic interest rate models and studied them respectively in Chapters 3-5.

One of the major challenges is that the drift terms of these three stochastic interest rate models posses a reciprocal function which may explode to infinity in finite time around the origin and the drift and diffusion terms are of super-linear growth which may also explode during numerical simulations. We overcame this challenge by constructing a new implementable truncated EM method to deal with these terms by forcing numerical solutions to remain within a compact support. We did this by truncating the coefficients whenever they grow super-linearly (i.e. the numerical solutions are bounded from tending to 0 or $\infty$); otherwise, we apply the classical EM method. With this technique, finite time explosions of the numerical solutions to infinity during numerical simulations have been ultimately avoided. In fact, this numerical technique enables the reciprocal function in the drifts of the three stochastic interest rate models to cope well around the origin. This is one of the most significant results we achieved in this thesis. Another major challenge is the structure of the delayed volatility function. We overcame this challenge by allowing the delayed volatility function to be locally Lipschitz continuous and upper bounded. This contributed significantly to proving the main results of the thesis. We will consider if the proofs of these main results can be extended to cope with unbounded delayed volatility functions in the near future.

The main objective of the thesis is to establish the $L^p(p \geq 2)$ strong convergence theories for the truncated EM approximate solutions of the three stochastic interest rate models. Apparently, these results are very paramount for calibration and valuation of some financial products such as debt and path-dependent derivative instruments. The only known strong convergence result for Ait-Sahalia-type model is the one established under a monotone condition by Szpruch et al. in [20] based on implicit EM methods (i.e., backward and forward-backward EM methods). The existing explicit method for Ait-Sahalia-type model is the EM method where weak convergence or convergence in probability has been revealed in several literature. However, under sufficient conditions, we achieved this main objective of the thesis by establishing the $L^p(p \geq 2)$ finite time strong convergence theories of the numerical solutions to these three proposed models based on the explicit method for sufficiently small step size. These results are very essential in research and financial contexts.

In financial applications, preservation of positivity of numerical schemes for financial models is a desirable feature in the context of explaining pathwise movements of financial variables. One drawback is that the truncated EM method we developed falls short to preserve positive numerical paths of the proposed models due to the Brownian path taking all real values with positive probability. We retained this important qualitative property by applying the balanced-implicit method (BIM) techniques in [69, 70] to the truncated EM scheme. However, we have been confronted with the problem of non-adaptedness to filtration of the stochastic integral in the continuous-time continuous process. This is one of the major drawbacks yet to be overcome. Another drawback we also encountered is the failure to establish rates of convergence for the proposed numerical schemes. However, rates of convergence are crucial in providing useful insights into approximation quality and theoretical foundations for efficient variance reduction techniques in Monte Carlo simulations. We worked to address this problem but failed to prove the upper bound of the expected reciprocal function involving the approximate solutions. It would be desirable in a view of future work to think of, maybe, using forward Itô stochastic integral introduced in [67] and discrete case approach employed in [68] to address these two drawbacks respectively, and achieve the expected theoretical results.

The results obtained from this thesis have opened up new chapter for further research in connection with strong convergent approximation of Ait-Sahalia-type models and applications of this to other relevant fields. For instance, it would be interesting to study the strong convergent approximation of Ait-Sahalia-type model driven by fractional Brownian motion. It is well recognised that SDEs driven by fractional Brownian motion produce long range dependency and rough volatility trajectory or turbulence which are controlled by Hurst parameter (e.g., see [71, 72]). Using Levy process in place of fractional Brownian motion is also another useful approach to describe further unexpected empirical disasters against interest rates since the Levy process consists of linear drift, Brownian motion and Levy jump processes known as Levy triplet (e.g., see [73, 75]). Moreover, stochastic modelling of structured investment plans has been attracting a serious deal of attention in the worlds of research, finance and insurance lately. However, most of the existing literature employed constant interest rates to compute mean percentage returns of structured investment products (e.g., see [76] and references cited therein). With results from this thesis, it would also be interesting to examine the use of the three proposed Ait-Sahalia-type models to describe interest rates dynamics for computation of mean percentage returns of some of these structured investment products.

# Bibliography


[1] Brigo, D. and Mercurio, F., 2007. Interest rate models-theory and practice: with smile, inflation and credit. Springer Science and Business Media.

[2] McInerney, D. and Zastawniak, T., 2015. Stochastic interest rates. Cambridge University Press.

[3] Black, F. and Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of political economy, 81(3), pp.637-654.

[4] Merton, R.C., 1973. Theory of rational option pricing. Theory of Valuation, pp.229-288.

[5] Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of financial economics, 5(2), pp.177-188.

[6] Dothan, L.U., 1978. On the term structure of interest rates. Journal of Financial Economics, 6(1), pp.59-69.

[7] Brennan, M.J. and Schwartz, E.S., 1980. Analyzing convertible bonds. Journal of Financial and Quantitative analysis, 15(4), pp.907-929.

[8] Cox, J.C., Ingersoll Jr, J.E. and Ross, S.A., 1980. An analysis of variable rate loan contracts. The Journal of Finance, 35(2), pp.389-403.

[9] Cox, J.C., Ingersoll Jr, J.E. and Ross, S.A., 1985. A Theory of the Term Structure of Interest Rates. Econometrica: Journal of the Econometric Society, pp.385-407.

[10] Chan, K.C., Karolyi, G.A., Longstaff, F.A. and Sanders, A.B., 1992. An empirical comparison of alternative models of the short-term interest rate. The journal of finance, 47(3), pp.1209-1227.

[11] Nowman, K.B., 1997. Gaussian estimation of single-factor continuous time models of the term structure of interest rates. The journal of Finance, 52(4), pp.1695-1706.

[12] Lewis, A.L., 2000. Option Valuation Under Stochastic Volatility. Finance Press, California.

[13] Deng, S., Fei, C., Fei, W. and Mao, X., 2019. Generalized Ait-Sahalia-type interest rate model with Poisson jumps and convergence of the numerical approximation. Physica A: Statistical Mechanics and its Applications, 533, p.122057.

[14] Higham, D.J. and Kloeden, P.E., 2005. Numerical methods for nonlinear stochastic differential equations with jumps. Numerische Mathematik, 101(1), pp.101-119.

[15] Higham, D.J. and Mao, X., 2005. Convergence of Monte Carlo simulations involving the mean-reverting square root process. Journal of Computational Finance, 8(3), pp.35-61.

[16] Wu, F., Mao, X. and Chen, K., 2008. A highly sensitive mean-reverting process in finance and the Euler–Maruyama approximations. Journal of Mathematical Analysis and Applications, 348(1), pp.540-554.

[17] Mao, X., 2007. Stochastic differential equations and applications. 2nd ed. Chichester: Horwood Publishing Limited.

[18] Ait-Sahalia, Y., 1996. Testing continuous-time models of the spot interest rate. The review of financial studies, 9(2), pp.385-426.

[19] Cheng, S.R., 2009. Highly nonlinear model in finance and convergence of Monte Carlo simulations. Journal of Mathematical Analysis and Applications, 353(2), pp.531-543.

[20] Szpruch, L., Mao, X., Higham, D.J. and Pan, J., 2011. Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. BIT Numerical Mathematics, 51(2), pp.405-425.

[21] Dung, N.T., 2016. Tail probabilities of solutions to a generalized Ait-Sahalia interest rate model. Statistics and Probability Letters, 112, pp.98-104.

[22] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E., 2010. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 467(2130).

[23] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American statistical association, 96(453), pp.42-55.

[24] Hull, J. and White, A., 1987. The pricing of options on assets with stochastic volatilities. The journal of finance, 42(2), pp.281-300.

[25] Kind, P., Liptser, R.S. and Runggaldier, W.J., 1991. Diffusion approximation in past dependent models and applications to option pricing. The Annals of Applied Probability, 1(3), pp.379-405.

[26] Mao, X. and Sabanis, S., 2013. Delay geometric Brownian motion in financial option valuation. Stochastics An International Journal of Probability and Stochastic Processes, 85(2), pp.295-320.

[27] Mao, X., 2015. The truncated Euler–Maruyama method for stochastic differential equations. Journal of Computational and Applied Mathematics, 290, pp.370-384.

[28] Guo, Q., Mao, X. and Yue, R., 2018. The truncated Euler–Maruyama method for stochastic differential delay equations. Numerical Algorithms, 78(2), pp.599-624.

[29] Mao, X. and Sabanis, S., 2003. Numerical solutions of stochastic differential delay equations under local Lipschitz condition. Journal of computational and applied mathematics, 151(1), pp.215-227.

[30] Mao, X., 2016. Convergence rates of the truncated Euler–Maruyama method for stochastic differential equations. Journal of Computational and Applied Mathematics, 296, pp.362-375.

[31] Higham, D.J., Mao, X. and Stuart, A.M., 2002. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM Journal on Numerical Analysis, 40(3), pp.1041-1063.

[32] Itkin, A., 2015. To sigmoid-based functional description of the volatility smile. The North American Journal of Economics and Finance, 31, pp.264-291.

[33] Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of financial economics, 3(1-2), pp.125-144.

[34] Lin, B.H. and Yeh, S.K., 1999. Jump-Diffusion Interest Rate Process: An Empirical Examination. Journal of Business Finance and Accounting, 26(7-8), pp.967-995.

[35] Kou, S.G., 2002. A jump-diffusion model for option pricing. Management science, 48(8), pp.1086-1101.

[36] Wu, F., Mao, X. and Chen, K., 2008. Strong convergence of Monte Carlo simulations of the mean-reverting square root process with jump. Applied Mathematics and Computation, 206(1), pp.494-505.

[37] Wu, F., Mao, X. and Chen, K., 2009. The Cox–Ingersoll–Ross model with delay and strong convergence of its Euler–Maruyama approximate solutions. Applied Numerical Mathematics, 59(10), pp.2641-2658.

[38] Lee, M.K. and Kim, J.H., 2016. A delayed stochastic volatility correction to the constant elasticity of variance model. Acta Mathematicae Applicatae Sinica, English Series, 32(3), pp.611-622.

[39] Coffie, E. and Mao, X., 2021. Truncated EM numerical method for generalised Ait-Sahalia-type interest rate model with delay. Journal of Computational and Applied Mathematics, 383, p.113137.

[40] Deng, S., Fei, W., Liu, W. and Mao, X., 2019. The truncated EM method for stochastic differential equations with Poisson jumps. Journal of Computational and Applied Mathematics, 355, pp.232-257.

[41] Liu, W., Mao, X., Tang, J. and Wu, Y., 2020. Truncated Euler-Maruyama method for classical and time-changed non-autonomous stochastic differential equations. Applied Numerical Mathematics, 153, pp.66-81.

[42] Guo, Q., Liu, W. and Mao, X., 2018. A note on the partially truncated Euler–Maruyama method. Applied Numerical Mathematics, 130, pp.157-170.

[43] Zhang, W., 2021. The truncated Euler–Maruyama method for stochastic differential equations with piecewise continuous arguments driven by Lévy noise. International Journal of Computer Mathematics, 98(2), pp.389-413.

[44] Guo, Q., Liu, W., Mao, X. and Zhan, W., 2018. Multi-level Monte Carlo methods with the truncated Euler–Maruyama scheme for stochastic differential equations. International Journal of Computer Mathematics, 95(9), pp.1715-1726.

[45] Bao, J., Böttcher, B., Mao, X. and Yuan, C., 2011. Convergence rate of numerical solutions to SFDEs with jumps. Journal of Computational and Applied Mathematics, 236(2), pp.119-131.

[46] Benth, F.E., 2003. Option theory with stochastic analysis: an introduction to mathematical finance. Berlin: Springer Science and Business Media.

[47] Applebaum, D., 2009. Lévy processes and stochastic calculus. Cambridge university press.

[48] Elliott, R.J., Chan, L. and Siu, T.K., 2013. Option valuation under a regime-switching constant elasticity of variance process. Applied Mathematics and Computation, 219(9), pp.4434-4443.

[49] Hamilton, J.D., 1988. Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. Journal of Economic Dynamics and Control, 12(2-3), pp.385-423.

[50] Ratanov, N., 2016. Option pricing under jump-diffusion processes with regime switching. Methodology and Computing in Applied Probability, 18(3), pp.829-845.

[51] Bollen, N.P., Gray, S.F. and Whaley, R.E., 2000. Regime switching in foreign exchange rates: Evidence from currency option prices. Journal of Econometrics, 94(1-2), pp.239-276.

[52] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E., 2012. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. The Annals of Applied Probability, 22(4), pp.1611-1641.

[53] Wang, X. and Gan, S., 2013. The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients. Journal of Difference Equations and Applications, 19(3), pp.466-490.

[54] Liu, W. and Mao, X., 2013. Strong convergence of the stopped Euler–Maruyama method for nonlinear stochastic differential equations. Applied Mathematics and Computation, 223, pp.389-400.

[55] Mao, X. and Yuan, C., 2006. Stochastic differential equations with Markovian switching. Imperial college press.

[56] Yuan, C. and Mao, X., 2010. Stability of stochastic delay hybrid systems with jumps. European journal of control, 16(6), pp.595-608.

[57] Coffie, E., 2021. Delay stochastic interest rate model with jump and strong convergence in Monte Carlo simulations. arXiv preprint arXiv:2103.07651.

[58] Coffie, E., 2021. Numerical approximation of hybrid Poisson-jump Ait-Sahalia-type interest rate model with delay. arXiv preprint arXiv:2107.03712.

[59] Coffie, E., 2017. Sensitivity Analysis of Prices of Claims in Turbulent Markets (Master's thesis). University Of Oslo.

[60] Øksendal, B., 2013. Stochastic differential equations: an introduction with applications. Springer Science and Business Media.

[61] Glasserman, P., 2013. Monte Carlo methods in financial engineering (Vol. 53). Springer Science and Business Media.

[62] Fristedt, B.E. and Gray, L.F., 2013. A modern approach to probability theory. Springer Science and Business Media.

[63] Bosq, D. and Nguyen, H.T., 2013. A course in stochastic processes: stochastic models and statistical inference (Vol. 34). Springer Science and Business Media.

[64] Norris, J.R., 1998. Markov Chains (Vol. 2). Cambridge University Press.

[65] Mao, X. and Szpruch, L., 2013. Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. Stochastics An International Journal of Probability and Stochastic Processes, 85(1), pp.144-171.

[66] Kahl, C. and Jäckel, P., 2006. Fast strong approximation Monte Carlo schemes for stochastic volatility models. Quantitative Finance, 6(6), pp.513-536.

[67] Russo, F. and Vallois, P., 1993. Forward, backward and symmetric stochastic integration. Probability theory and related fields, 97(3), pp.403-421.

[68] Li, X., Mao, X. and Yin, G., 2019. Explicit numerical approximations for stochastic differential equations in finite and infinite horizons: truncation methods, convergence in p th moment and stability. IMA Journal of Numerical Analysis, 39(2), pp.847-892.

[69] Milstein, G.N., Platen, E. and Schurz, H., 1998. BIM methods for stiff stochastic systems. SIAM Journal on Numerical Analysis, 35(3), pp.1010-1019.

[70] Schurz, H., 1996. Numerical regularization for SDEs: Construction of non-negative solutions. Dynamic Systems and Applications, 5, pp.323-352.

[71] Gatheral, J., Jaisson, T. and Rosenbaum, M., 2018. Volatility is rough. Quantitative finance, 18(6), pp.933-949.

[72] Kijima, M. and Tam, C.M., 2013. Fractional Brownian motions in financial models and their Monte Carlo simulation. Theory and Applications of Monte Carlo Simulations, pp.53-85.

[73] Di Nunno, G., Øksendal, B.K. and Proske, F., 2009. Malliavin calculus for Lévy processes with applications to finance (Vol. 2). Berlin: Springer.

[74] Mao, X., Truman, A. and Yuan, C., 2006. Euler-Maruyama approximations in mean-reverting stochastic volatility model under regime-switching. Journal of Applied Mathematics and Stochastic Analysis, 2006.

[75] Rachev, S.T., Kim, Y.S., Bianchi, M.L. and Fabozzi, F.J., 2011. Financial models with Lévy processes and volatility clustering (Vol. 187). John Wiley and Sons.

[76] Feng, L., Huang, Z. and Mao, X., 2016. Mean percentage of returns for stock market linked savings accounts. Applied Mathematics and Computation, 273, pp.1130-1147.