Department of Mathematics and Statistics

Computational Methods for Various Stochastic Differential Equation Models in Finance

Yanli Zhou

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

November 2014

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

.....

Yanli Zhou

5 November 2014

Acknowledgements

The research reported in this thesis was carried out between November 2011 and November 2014, while I was a PhD student in the Department of Mathematics and Statistics, Curtin University. I really appreciate all kinds of help I have received from my supervisors, families, friends, and colleagues during this period of time.

I would like to express my heartfelt thanks to my supervisor, Prof. Yonghong Wu and his wife Mrs. Hui Wang. Professor Wu has guided my research during the past three years with remarkable patience and enthusiasm.

I would like to thank Prof. Xiangyu Ge, my master thesis supervisor in the Department of Statistics and Mathematics, Zhongnan University of Economics and Law. I have known Prof.Ge since I was a graduate student in the Department of Statistics and Mathematics, Zhongnan University of Economics and Law in 2008. It was him who led me to the road of research and to win a scholarship from China Scholarship Council for my PhD program.

I wish to acknowledge the help that I received from Prof. Tianhai Tian in the School of Mathematical Sciences, Monash University.

I also wish to thank John Curtin Distinguished Professor Kok Lay Teo, the previous head of Department of Mathematics and Statistics. He is also the Chair of my Thesis Committee. He has been very kind and most helpful.

I would like to give thanks to my fellow PhD students and friends that I have worked with in Australia, especially, A/Prof. Nikolai Dokuchaev, A/Prof. Guanglu Zhou, Dr. Honglei Xu, Dr. Qun Lin, Dr. Changjun Yu, Dr. Bin Li,

Qian Sun, Qinqin Chai, Shuang Li, Yan Zhang, Yufei Sun, and the department's academic visitors that I had the opportunity to meet. They are Prof. Shaoyong Lai, Prof. B.Wiwatanapataphee, Dr. Xinguang Zhang and A/Prof. Tieqiao Tang. They are always willing to help. I really appreciate all the unselfish helps I have received from them. In addition, I wish to thank my friends and teachers at Zhongnan University of Economics and Law, especially, A/Prof. Zhongming Ma, A/Prof. Guoguang Yan, A/Prof. Yongwei Chen, A/Prof. Yongsheng Jiang, Dr Songlin Liu, Dr. Xia Zhu, Dr. Chunxia Tang, Dr. Chaobin Hu, Dr. Qing Li. and Yang Wu. I really enjoyed the time that I spent with them.

I thank the staff in the Department of Mathematics and Statistics for making the work environment so friendly. The past and present administrative staff, in particular, Joyce Yang, Shuie Liu, Jeannie Darmago and Carey Ryken-Rapp deserve special thanks for providing such an effective and efficient administrative support to the department. They are always very kind and supporting whenever being approached.

I wish to acknowledge China Scholarship Council and Curtin University for providing me with the CSC-CIPRS scholarship for my PhD research.

Finally, on the personal note, I sincerely thank everyone in my family, especially my parents, my two younger sisters, for their love, encouragement, understanding and support throughout the entire period of my PhD candidature in Australia.

Abstract

Stochastic differential equations (SDEs) have been widely used as models for financial quantities such as interest rates, asset prices as well as their derivatives. Unlike deterministic models such as ordinary differential equation model, SDEs have solutions in the forms of continuous-time stochastic processes instead of a unique solution for each appropriate initial condition. Methods for the numerical solutions of stochastic differential equations are based on techniques for ordinary differential equations with generalization to provide support for stochastic dynamics.

This study consists of three parts. The first part of the research focuses on constructing an efficient numerical method for solving jump-diffusion stochastic differential equations under the Poisson random measure and a fixed time delay. A simplified Taylor method is established to give solutions with a weak convergence rate arbitrarily close to order β, and a corresponding convergence theorem for stochastic delay differential equations (SDDEs) with jumps is established and proved. A numerical example shows that the proposed numerical scheme gives stable results under certain accuracy requirement.

The second part of the research focuses on stochastic differential equations with fractional order for application in option pricing. We first formulate the stock price process by stochastic differential equations with fractional order, and then establish the European call option pricing formulae using the fractional order stochastic differential equation. Some comparisons have been made among the option pricing formula established by our proposed approach and two classical

models. We find that the new approach leads to a better result than the classic approach and the fractional Brownian motion approach, through simulating the stock prices by the Monte Carlo simulation method.

The third part of the research focuses on parameters calibration for stochastic models. First, we propose an effective algorithm for the estimation of parameters in SDE models, based on the implementation of the Bayesian inference and the Markov Chain Monte Carlo (MCMC) method. The importance sampling technique is used to increase the robustness of estimates. We have also examined the influence of different samples of the latent variables on the variation of estimates. Numerical results suggest that the method used in this work is robust for such variation. Then, we develop another algorithm for parameter estimation of the stiff SDE models. This approach is based on the usage of an implicit numerical scheme in the Monte Carlo simulation integrated with the particle swarm optimization (PSO). Experimental results show that when the SDE model is moderate stiff, estimates produced by the semi-implicit Milstein method have better accuracy than those obtained by the explicit Milstein method. Furthermore, the PSO algorithm can produce reliable estimates which are nearly independent of the implementation details compared with the genetic algorithm.

List of Publications

The following papers (which have been published or accepted for publication) were completed during my PhD candidature:

- Y. L. Zhou, Y. H. Wu, X. Y. Ge, and B. Wiwatanapataphee, "A Robust Weak Taylor Approximation Scheme for Solutions of Jump-Diffusion Stochastic Delay Differential Equations," in *Abstract and Applied Analysis*, vol. 2013, article ID 750147, 2013.

- Y. L. Zhou, Y. H. Wu, X. Y. Ge, and T. H. Tian, "Parameter Calibration of Interest Term Structure Models: an Implicit Numerical Method with Particle Swarm Optimization," *SIMULATION: Transactions of The Society for Modeling and Simulation International*, 2014.

- Y. S. Jiang, Y. L. Zhou, B. Wiwatanapataphee, and X. Y. Ge, "Nonexistence Results for the Schrödinger-Poisson Equations with Spherical and Cylindrical Potentials in R3," *Abstract and Applied Analysis*, vol. 2013, article ID 890126, 2013.

- T. H. Tian, Y. L. Zhou, Y. H. Wu, and X. Y. Ge,, "Estimation of the Parameters in Mean-reverting Stochastic Systems," *Mathematical Problems in Engineering*, vol. 2014, article ID 317059, 2014.

- S. Li, Y. L. Zhou, X. F. Ruan, and B. Wiwatanapataphee, "Pricing of American Put Option under a Jump-diffusion Process with Stochastic Volatility in an Incomplete Market," *Abstract and Applied Analysis*, vol. 2014, article ID 236091, 2014.

- Q. Li, Y. L. Zhou, X. Q. Zhao, and X. Y. Ge, "Fractional Order Stochastic Differential Equation with Application in European Option Pricing," *Discrete Dynamics in Nature and Society*, vol. 2014, article ID 621895, 2014.

The following papers were completed during PhD candidature and are currently under review:

- Y. L. Zhou, S. Li, Q. Li, and Y. H. Wu, "On the Convergence of Weak Taylor Approximation Scheme for solutions of Fractional Stochastic Differential Equations"

- S. Li, Y. L. Zhou, X. F. Ruan, and Y. H. Wu, "Equilibrium Approach of Asset and Option Pricing under Lévy Process and Stochastic Volatility"

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Background

During the past 30 years there has been an increasing demand for tools and methods of stochastic differential equations in various disciplines. In engineering, stochastic differential equations are used in filtering and control theory. In physics, stochastic differential equations are used to study the effects of random excitations on various physical phenomena. In biology, stochastic differential equations are employed to model the effects of stochastic variability in reproduction and environment on populations. One of the greatest demands has come from the growing area of mathematical finance, where stochastic differential equations are used for pricing and hedging of financial derivatives, such as options. They are also very important to much of modern finance theory and have been widely used to model the behaviour of key variables such as stock price, short term interest rate, asset returns and their volatility. Particularly, fluctuating global economic conditions require market uncertainty management to ensure the effectiveness of investment decision.

In mathematical finance, delays in the dynamics can represent memory or inertia in the financial system. For instance, successive price changes (or returns) are independently distributed, which is called inefficient market. Stochastic delay

differential equations is an effective instrument to model the memory effects. A lot of surveys of financial markets show that a large proportion of investors use past prices as a guide to make investment decisions. Such feedback trading strategies may lead to speculative asset bubbles and crashes. Because of the absence of this feedback behaviour in standard non-delay models, it is necessary to assume that aggregate demand is functional of past prices. In this case, price dynamics could be modeled by stochastic delay differential equations.

Fractional derivative has a history as long as that of classical calculus, but it is much less popular than it should be. In financial time series, long term memory indicates the correlation structure of a series at long lags. If a series shows long-term memory, there is persistent temporal dependence even between distant observations. The presence of long range memory in financial asset returns has important implications for many of the stochastic models in financial economics. For instance, optimal consumption or savings and portfolio investment decisions might become extremely sensitive to the investment period if the returns were long range dependent. To address the long memory effects in financial dynamics, the fractional stochastic differential equations may be a good alternative choice.

Stochastic modelling and simulation have become areas of intense research in recent years, as more sophisticated mathematical models of financial phenomena become available. Since only a small class of stochastic differential equations admits explicit solutions, efficient and effective numerical methods need to be constructed for the weak solutions of stochastic differential equations with/without jumps.

Parameter calibration is anther important issue in the field of mathematical finance. Given the widespread usage of SDEs in a diverse range of fields, any contribution that improves the performance of models based on SDEs provides obvious benefits to society. The explanatory and/or predictive power of these models depends crucially on the particularisation of the model SDE(s) to real

data through the specification of the SDE(s) and the choice of values for their parameters.

In econometrics, the estimates of optimal parameters are generally obtained by maximising the likelihood function of the sample. However, in the context of the estimation of the parameters in SDEs, a closed-form expression for the likelihood function is rarely available and hence exact maximum likelihood (EML) estimation used in classical econometric models is usually infeasible. Thus, further research is needed for the development of accurate and computationally feasible estimation procedures based on the maximum likelihood principle in the absence of a closed-form expression for the likelihood function.

1.2 Objectives

The main purpose of the work is to construct efficient discrete-time approximation methods to address the memory effects arising in financial markets, and develop new numerical algorithms for parameter estimation in nonlinear stochastic differential equations. More specifically, this thesis aims to

(i) develop a robust Taylor approximation scheme for the solutions of jump-diffusion stochastic delay differential equations and then examine the convergence of the numerical method in a weak sense.

(ii) construct a fractional order stochastic differential equation model to describe the effect of trend memory in financial derivative pricing.

(iii) develop a numerical algorithm to estimate unknown parameters in stochastic interest rate models by utilizing the Bayesian inference and Markov chain Monte Carlo method.

(iv) propose a new method by using an implicit method of nonlinear stochastic differential equation and particle swarm optimization algorithm to calibrate

parameters in the term structure models of interest rate.

1.3 Outlines of the Thesis

This thesis consists of six chapters. Chapter 1 presents a brief introduction of the research and gives the objectives of the study.

Chapter 2 reviews previous work relevant to the scope of this project. Some necessary background information and knowledge closely related to this research are also presented in this chapter.

Chapter 3 begins with an overview of the theory of weak solutions of stochastic delay differential equations in Section 3.2 and Section 3.3. In Section 3.4, we construct a robust Taylor approximation scheme and then examine the convergence of the method in a weak sense. A convergence theorem for the scheme is established and proved. Then, a scheme of high order is proposed for Monte Carlo simulation for jump-diffusion stochastic delay differential equations.

Chapter 4 begin with a review of fractional order ordinary differential equation used as an effective instrument for describing the memory effect in complex systems. A fractional order stochastic differential equation model is constructed for describing the effect of trend memory in financial pricing. We, then, derive a European option pricing formula based on the FSDE model and prove the existence of the trend memory (i.e., the mean value function) in the option pricing formula when the Hurst index is between 0.5 and 1. In addition, we carry out a comparison analysis among our proposed model, the classic Black-Scholes model, and the stochastic model with fractional Brownian motion. Numerical results suggest that our model leads to more accurate and lower standard deviation in the empirical study.

Chapter 5 concentrates on identifying parameters for stochastic models to generate accurate simulations. Utilizing the Bayesian inference and Monte Car-

lo Markov Chain method, we develop a numerical algorithm to estimate the unknown parameters in stochastic interest rate models. Section 5.2 gives the stochastic models for term structure of interest rates and numerical algorithms for simulating these stochastic models. Section 5.3 discusses the Bayesian inference and the Monte Carlo Markov Chain method and then propose a new effective method to estimate parameters in stochastic models. Section 5.4 reports the numerical results for parameters estimation in the stochastic models for the term structure of interest rates.

In Chapter 6, we develop a novel method by using implicit methods to solve SDEs, which is aimed at generating stable simulations for stiff SDE models. The particle swarm optimization method is used as an efficient searching method to explore the optimal estimate in the complex parameter space. Using the interest term structure model as the test system, numerical results show that the proposed new method is an effective approach for generating reliable estimates of unknown parameters in SDE models.

The last chapter provides summaries and a discussion of possible future research.

CHAPTER 2

Literature Review

2.1 General Overview

A differential equation that contains a random component, which in turn leads to a solution that is a random process, is known as a stochastic differential equation (SDE). Stochastic differential equation (SDE) can be defined as a deterministic differential equation perturbed by random disturbances that are not necessarily small. Stochastic delay differential equation models play a very important role in the study of many fields such as economics and finance, chemistry, biology, microelectronics, and control theories. In particular, SDEs are also central to much of modern finance theory and have been used profitably to model the behaviour of key variables such as stock prices, short term interest rate and their volatilities.

Models of this type can more accurately describe many phenomena in the real world by taking into account the effect of time delay and/or event-driven uncertainty. By time delay, it means a certain period of time is required for the effect of an action to be observed after the moment when the action takes place. This phenomenon, called memory effect, exists in most systems in almost any area of science; for example, a patient shows symptoms of an illness days or even weeks after he/she was infected. Similarly, random noises appear in almost all real world phenomena and systems, for example, the motion of molecules, the

price of assets in financial markets. Hence, study of SDDEs is an important undertaking in order to understand real world phenomena and systems precisely. By event-driven uncertainty, it means events such as corporate defaults, operational failures, market crashes or governmental macroeconomic announcements happens more and more frequently especially in the recent financial markets with increasing fluctuations, which cannot be properly modelled by purely continuous processes.

Alternatively, memory effects can be modeled by using fractional stochastic differential equations. Long-term memory means the correlation structure of a series at long lags. If long-term memory (or the biased random walk) exists in a series, there is persistent temporal dependence even between distant observations (Barkoulas and Balum [81]). During the last couple of decades, fractional Brownian motion and fractional stochastic differential equations are used and developed to describe the memory effects in many applications such as fiance and economics, engineering, biology and so on. Maheswaran underlies that the presence of long-term memory in stock returns has important implications for many of the paradigms in financial economics [122]. The existence of long-term memory volatility in asset returns has also important implications for pricing contingent claims in emerging markets by LeRoy [132].

Stochastic differential equations play an important role in modelling various phenomena arising in fields as diverse as finance, physics, chemistry, engineering, biology, neuroscience and others. These equations usually depend on parameters, which are often unknown. On the other hand knowledge of these parameters is critical for the study of the process at hand and hence their estimation based on the observational data on the process under study is of great importance in practical applications. The estimation of the parameters of SDEs from discretely-sampled data has received substantial attention in the financial econometrics literature, particularly in the last ten years [53, 78].

2.2 Numerical Methods for Stochastic Differential Equations

2.2.1 Stochastic Delay Differential Equations with Jumps

The dynamics of financial and economic quantities are often described by stochastic differential equations (SDEs). SDEs of jump-diffusion type, which capture the dynamics of the impact of event-driven uncertainty, receive much attention in financial and economic modelling, see Merton [116] or Cont and Tankov [111]. On the other hand, another broad class of dynamical systems of interest comprises memory. Phenomena involving memory effects and/or time delays are indeed ubiquitous. More specifically, differential equations with time delay play an important role in many fields such as physics, biology, economy, finance and so on. They have been the subject of extensive mathematical studies [88, 117, 134]. Under such a circumstance, one must know the whole past of the dynamical system as well as its present information, in order to predict its immediate future accurately. Several mathematical works have established a range of results on stochastic differential equations with random delay [54, 106, 119, 120, 143], with regard to the stability and convergence behavior of such stochastic dynamical systems.

Over the last couple of decades, a lot of work has been carried out to study differential equations with delay and/or random noises, for example, [68, 77, 101, 152, 161]. The best known and well studied theory and systems include the delay differential equations(DDEs) presented by Kolmanvskii & Myshkis [147] and their stochastic generalizations, and the Stochastic delay differential equations (SDDEs)established by Mohammed [127, 128], Mao [154, 155] and Mohammed & Scheutzow [131]. Other SDDEs theories of interest include, for instance, the so-called SDDEs with Markovian switching and Poisson jumps. These models have

been investigated in the literature [12, 72, 112].

Analytical solutions of SDDEs can hardly be obtained. It is thus important to develop and study discrete time approximation methods for solving SDDEs. Discrete-time approximations may be divided into two categories: weak approximations and strong approximations [107]. Some implicit and explicit numerical approximation methods for SDDEs in strong approximation sense were derived by Kuchler & Platen [144]. Weak numerical methods for SDDEs have been studied by Kuchler & Platen [145]. Monte Carlo simulation method has also been developed as a powerful simulation method for SDEs, while weak numerical approximations are required for Monte Carlo simulation [23, 24, 30, 60, 154].

The main motivation for considering weak approximations is the computation of the expectation of functionals of the solutions to stochastic differential equations. In mathematical finance, the fair pricing of options resorts to solving expected function of solutions of SDEs. Weak solutions are also used in the computation of Lyapunov exponents of systems described by stochastic functional differential equations, which has been presented by Milstein and Tretyakov in [61]. Lyapunov exponents for stochastic functional differential equations were studied by Mohammed and Scheutzow [129, 130]. Weak approximations for stochastic ordinary differential equations (without memory) are well-developed; and for more details, the reader is refered to the references due to Bally and Talay [146], Kloeden and Platen [107], Milstein and Tretyakov [62] and Kohatsu-Higa [3]. The earliest reference on weak approximation of numerical methods for stochastic differential equations with delay is Kuchler and Platen [145]. This paper however provides no rigorous justification of their statements. The first rigorous analysis was recently given by E. Buckwar and T. Shardlow in [43], which establishes weak convergence of order 1 for the Euler scheme. Recent results on weak convergence of the Euler scheme for a class of SFDEs were obtained independently by E. Clement, A. Kohatsu-Higa and D. Lamberton [1]. The results

in [1] offers the potential for developing higher-order weak convergence schemes for stochastic systems with memory.

A survey of uniqueness results for stochastic differential equations with jumps and regularity results for the corresponding harmonic functions are presented by Bass. Martingale problems are closely related to the notion of weak uniqueness for SDEs. Various work on the existence of solutions for stochastic differential equations are reported in [104] and [105]. Another approach that has been explored is the use of pseudo-differential operators [76, 151].

2.2.2 Fractional Stochastic Differential Equations

Time series incorporating memory structure has been widely used in biological, chemical, and physical system. Memory effects also exist in financial systems. For example, the decision will be effected spontaneously by the past experience of decision makers. Plenty of financial variables with long memory effects have been found [19, 44, 51, 160], such as the gross domestic product (GDP), interest rate, foreign exchange rates, stock price, and futures price. Garzareli et al. have proved the existence of memory effects in the stock price series by the conditional probability approach and measured the extent of long memory (autocorrelation) [2]. In [27] authors reported their results with time windows varying between tens of minutes to several days, and their data analysis results confirm long-range correlations in the volatility.

Memory effect is often measured by the autocorrelation function, and, recently, the Hurst index as an effective tool was introduced to measure the memory effect [66]. The Hurst index is often denoted by $H(0 < H < 1)$. In the case of $0 < H < 0.5$, time series has negative correlation and antipersistent behavior, which is called short-dependence memory. When $H = 0.5$, the time series has no dependence. However, in the case of $0.5 < H < 1$, time series has positive correlation and persistent behavior, which is long-dependence memory. The persistent

behavior was also called Joseph Effect by Mandelbrot and Wallis [17]. Cajueiro and Tabak [39,40] have also found that memory effect exists in financial markets.

A number of researchers used fractional Brownian motion to depict the characteristic of memory. Mandelbrot and Van Ness first found that long memory effects exist in stock returns and gave the definition of fractional Brownian motion [18]. Since then, describing the memory by the fractional Brownian motion in financial market becomes more and more popular. For instance, Bęben and Orłowski [91], Huang and Yang [16], Evertsz [26], Lo [13], and Wen et al. [50,162] have shown that the returns are of long-term (or short term) dependence in the markets. After Black and Scholes [47] developed the option pricing theory based on the classical stochastic differential equation, a large number of literatures studied the option price based on the fractional Brownian motion. For example, Necula [22], Rostek [125], and Hu and Øksendal [159] obtained the Black-Scholes option pricing formula under fractional Brownian motion. Ren et al. [156] have considered the option pricing model for $0.5 < H < 1$. In the case of $0 < H < 0.5$, the option pricing formula was studied by Wang et al. [150]. Chen et al. [48] established the mixed fractional version of Black-Scholes model with $0 < H < 1$ and gave the Itos formula correspondingly. Besides, a different approach called mixed fractional Brownian motion, where the stochastic process of the stock price is also transformed into a semimartingale, was again suggested by Cheridito [99].

Fractional Brownian motion becomes a suitable tool in different applications such as mathematical finance because of its self-similarity and long-range dependence properties. As for $H \neq 0.5$ the fractional Brownian motion is neither a Markov process nor a semimartingale, the usual stochastic calculus can not be applied to analyze it. After a pathwise integration theory for fractional Brownian motion was established ([133] and [86]), it was found that arbitrage exists in the market mathematical models driven by $B_H(t)$ have been found [89]. Hence, the fractional Brownian motion was no longer considered fit for mathematical mod-

eling in finance. After a new kind of integral based on the Wick product ([157] and [158]) called fractional Ito integral was developed, it was found that the corresponding Ito type fractional Black-Schools market has no arbitrage. In the work of Hu and Oksendal, a formula for the price of a European option at $t = 0$ is derived [15]. Sethi and Lehoczky [135] found that the Black-Scholes option pricing formula can be derived by using different types of stochastic integration calculus. The procedure for derivation of Ito integrals via Brownian motions is well-known. However, the derivation under the Stratonovich framework is not that prevalent.

However, the memory effects contain not only the noise memory effect but also the trend memory effect. Stochastic differential equation with fractional Brownian motion only describes the noise memory but cannot be used to study the trend memory effect of stock price. So we will describe the trend memory process by using the fractional derivative, which is another effective instrument to describe the memory effect. In particular, fractional calculus has been successfully applied in biology, physics, chemistry, and hydrology. Recently, the concept of fractal has been extended in financial mathematics [149]. This is due to the fact that fractional integral and derivatives can depict the memory and inherent process [70]. It has been realized that fractional derivative provides an excellent mathematical instrument for the description of complex process, irregular increment, memory properties, and intermediate process [36, 70, 85, 121].

2.3 Parameter Calibration

2.3.1 Parameter Estimation in Complex Systems

In natural systems, we can observe many fluctuating phenomena due to various resources of uncertainty and changes inside and outside of the systems. Stochastic differential equations have the ability to describe these fluctuations by adding a

noise term into deterministic models. Thus, Stochastic differential equations have been widely used in recent years to describe uncertainty property in complex systems arising from biological sciences, physical sciences, engineering, finance and economics. In the modern finance and economics theory, stochastic differential equations have been used as a fundamental tool to represent volatility such as the behaviour of short-term interest rate and asset prices.

In real applications, the parameters of the equations are unknown and need to be estimated. A major step in the development of financial models is the estimation of unknown parameters in the model. Parameter estimation in linear stochastic and nonlinear deterministic differential equations, even for data covered by additive observational noise, is well known, [73,82] for a review. However, the estimation of parameters in nonlinear stochastic differential equations is still under development. In most cases, we only have discretely sampled data on the equation, and thus it is a common practice to use the discretization of the original continuous time model for the parameter estimation. For financial models, it is very difficult to calculate the model parameters directly from financial data. Therefore the estimation of parameters in financial models has become an important research topic and received much attention in the last decade [53, 78, 95].

Parameter estimation in nonlinear SDEs driven by Wiener processes, when only discrete observation is available, is an inherently difficult problem and remains a challenge [74]. The main reason is that theoretically an unlimited number of solutions exist for a SDE, and obtaining the numerical solutions of SDEs arising from real world application is computationally demanding, particularly, when closed-form solutions for some SDEs do not exist. Various numerical methods such as the Euler methods and the high order Taylor schemes have been combined with a Monte Carlo approach to generate discrete-time trajectories of the state variables of SDEs [108]. By using these numerical methods, a large number of Wiener processes corresponding to different simulation trajectories need to

te generated, and hence achieving accurate results is computationally expensive, especially when the search space of parameters is complex. In addition to the previously remarkable results duo to Young [103], Soderstrom and Stoica [137], Sagara and Zhao [126], Soderstrom et al. [93, 102], and Haverkamp et al. [29], more efficient estimation methods are required.

2.3.2 Parameter Estimation Methods

The methods that were developed for parameter estimation of SDEs can be classified into three different categories: maximum likelihood estimation (MLE)/ simulated maximum likelihood (SML) [34, 59, 78], the methods of moments [7, 8, 42], and filtering (e.g. extended Kalman filter) [79]. Most of the methods are developed for modelling financial quantitatives, where the financial systems are characterised by long time horizons and often can be sampled at regular but relatively infrequent intervals (for example, on a daily basis).

The methods of moment are relatively simple and have been well illustrated in textbook and widely used in application, because they do not rely upon distributional assumptions. However, the accuracy of this method strongly depends on the frequency of financial observations. A generalized Method of Moments (GMM) due to Hansen [90], Ogaki [92], was originally developed for discrete-time stochastic models, and can be used to compute moment conditions from a discrete-time SDE model [52]. The main advantage of the GMM method is that it only requires some moment conditions rather than the full density. However, as the GMM does not make efficient use of all the information in the sample, this method may lead to a loss of efficiency. The Efficient Method of Moments (EMM), initially proposed by Gallant and Tauchen [10] and further developed by [9, 35, 141], introduces a natural progression for GMM that successfully addresses some of the problems mentioned above. Gallant and Tauchen [9] carries out an efficiency analysis of the EMM method compared to other methods of

moments using the Marron and Wand test [80] and concludes that the relative efficiency of EMM is uniformly higher than the other moment methods. The small-sample properties of EMM are studied by Chumacero, Michaelides and Ng, Andersen et al. [4, 67, 110] using Monte Carlo simulation. They conclude that EMM is more efficient than other methods of moments. A Simulated Method of Moments (SMM) was developed by Duffle and Singleton [32]. Based on the statistical technique of the random sampling instants, Duffle and Glynn [33] proposes a infinitesimal generator-based method. However, both methods cannot handle unobserved states.

Alternatively, Ljung and Soderstrom proposed the prediction error identification method in [87], and Shoji and Ozaki established the extended Kalman filter (EKF) and some other linearization methods for parameter estimation by approximating the conditional density in terms of conditional Gauss distributions [71]. The extended Kalmanfilter (EKF) is a well-known extension of the linear Kalman filter to nonlinear systems (see Jazwinski [6]). By adding the second-order terms to the moment equations, the second-order nonlinear filter (SNF) is defined in the work of [6]. In both cases the likelihood function is computed recursively using the prediction error decomposition [49]. For SDEs with a state dependent diffusion function, higher order filters are needed [109]. The Kalman-Bucy filter (Kalman and Bucy [118]) provides an exact solution to the filtering problem. A Gaussian likelihood function obtained from a Prediction Error Decomposition (PED) approach also can provide recursive residuals [49]. A maximum likelihood method for direct parameter estimation for SDEs is proposed based on the EKF and the PED method by Madsen and Melgaard [63], Melgaard and Madsen [64] and Bohlin and Graebe [136]. The EKF provides state estimates and the recursive residuals, while the PED is used to calculate QML estimates of the parameters using the Gaussian likelihood function. This approach may also be used for nonlinear stochastic systems.

The maximum likelihood estimator can be computed by maximising the likelihood function associated with a series of observations, provided both the invariant density and the transition density are known explicitly. Unfortunately, for many realistic and practically useful models, transition densities are not available in explicit form, which makes exact computation of the maximum likelihood estimator impossible. If the likelihood is a nonlinear function of the parameter of interest, computation of the maximum likelihood estimator is often far from straightforward (see e.g. Barnett [148]). In the paper of Bibby and Sorensen [14] a quasi maximum likelihood estimator has been proposed. Compared with the infeasible maximum likelihood estimator, the computationally feasible estimator becomes unbiased by paying the price that the resulting confidence region are no longer optimal. Experience shows that the loss in optimality is often rather small [25]. In the general multivariate case, we can not hope to obtain analytical solutions of the Fokker-Planck Equation and must resort to approximations and numerical procedures, for instance, matrix continued-fractions, finite differences, Monte Carlo methods, and so on, and for details, see Risken [65]; Press et al. [153]; Kloeden and Platen [107]. There is no restriction on the available data set to the MLE method, which is based on the availability of the closed-form expression for the transitional probability density function. When a closed-form expression is not available, the SML method is the approach to estimate the transitional probability density function by numerical simulations. As mentioned above, the computing time is the major obstacle in the application of the SML method. In addition, this method is technically complex and thus it is not widely used in financial and economic sectors. The likelihood-based methods are more reliable but has long been found that they are difficult to apply to SDEs due to its computational cost. For the complex SDEs, thousands of simulation trajectories or even more must be generated to ensure a low variance of the variable values. Consequently, a large number of competing estimation procedures have been proposed

in recent years.

In recent years, the Bayesian inference methods have been used to estimate unknown parameters in mathematical models [31,37,38,97,114]. Together with the Markov-chain Monte-Carlo (MCMC) and other methods, the Bayesian inference methods have also been used to infer stochastic models in financial mathematics [5, 20, 100]. The main advantage of these methods is the ability to infer the whole probability distribution of the parameters, rather than just a single estimate. In addition, the Bayesian methods can deal with noisy data and uncertain data. Another advantage of these methods is the capability to infer parameters in either deterministic models or stochastic models. However, the potential obstacle of these methods in application is that the samples are correlated and their performances heavily depend on prior hypothesises. A number of methods have been used to estimate the parameters in the single-factor continuous time models, including the generalized moment method [52] and the Gaussian estimation method [84]. However, our recent research work suggested that the accuracy of the estimates generated from these two methods is low, in particular, when the step size of observation time points is not small [140]. Thus in this work we will not test these methods again but concentrate on the proposed method that will generate accurate simulations of the stochastic model, which will lead to more accurate estimates of the model parameters.

There are two types of errors that should be considered for an efficient SML method. The first type of errors derived from the slow convergence rate of the Monte-Carlo approaches. To reduce the number of stochastic simulations, the variance-reduction methods have been designed to reduce the bias in the estimated moments of the solution. Two major types of variance-reduction methods include the importance sampler and the random number generation methods. The second error is the discretization difference between a numerical scheme and the original SDE. Currently the Euler-Maruyara method has been dominantly

used in the inference of stochastic models because of its simplicity and special property of this method, namely the numerical solution is a Gaussian random variable. However, this method do not have good stability property and the convergence rate is low. Thus in recent years, we have advocated to use high order methods and/or implicit methods to improve the stability property and to enhance the simulation accuracy. Another key issue in the implementation of the SML method is the optimization scheme for searching the optimal parameters. The machine-learning methods, such as the genetic algorithm (GA) and particle swarm optimization (PSO) algorithm, have been widely used in the SML methods for parameter inference [78, 98, 115]. The previous studies have shown that the GA has a better convergence rate than the Markov Chain Monte Carlo (MCMC) method [96], but our numerical results suggested that the GA could generate a wide range of estimates that all realize the observation data [142].

2.4 Concluding Remarks

As only a small class of jump diffusion SDDEs have explicit solutions, it is important to construct discrete-time approximations for stochastic differential equations of this type. The focus of this thesis is the numerical solution of SDDEs with jumps. We consider the Monte Carlo simulation, for which weak schemes are employed. Weak numerical schemes provide approximations of the expected functionals generated by the solution of a given SDE. These weak approximation schemes are appropriate for problems such as pricing of financial derivatives, the computation of risk and investors expected utilities. In addition, memory effects in financial and economic variables can also be depicted by fractional stochastic differential equations. Based on the previously notable research, constructing the fractional stochastic differential equations and its corresponding arbitrage strategy for derivatives pricing in mathematical finance becomes possible.

Whilst the analysis of stochastic processes and stochastic differential equations has received extensive treatment over a long period, the parameter estimation for such processes has received much less attention. It has recently been argued that because of the difficulty of obtaining consistent estimates of the parameters of nonlinear SDEs, parameter identification becomes one of the most pressing difficulties in the study of stochastic differential equations with discretely sampled data. In this work, we will develop some efficient and effective parameter identification algorithms for estimation of unknown parameters in nonlinear stochastic differential equations.

CHAPTER 3

A Robust Weak Taylor Approximation Scheme for Solutions of Jump-diffusion Stochastic Delay Differential Equations

3.1 General Overview

Stochastic delay differential equations with jumps have a wide range of applications, particularly, in mathematical finance. Solution of the underlying initial value problems is important to the understanding and control of many phenomena and systems in the real world. In this chapter, we develop a robust Taylor approximation scheme for solving stochastic delay differential equations with jumps and then examine the convergence of the method in a weak sense.

This chapter is organised as follows. In sections 3.2 and 3.3, we present a general stochastic delay differential equation with jumps and establish the conditions for the existence and uniqueness of solutions to the Jump-diffusion SDDEs, and present various lemmas to be used later for the proof of the convergence theorem. We then introduce, in Section 3.4, a general weak approximation scheme, where the simplified stochastic Taylor approximation scheme with order β is constructed, followed by a convergence theorem and its proof. In Section 3.5, we give a

numerical example to demonstrate the application and the convergence of the numerical scheme.

3.2 Stochastic Delay Differential Equations with Jumps

In this chapter, we extend a fully weak approximation method for SDDEs to a class of jump-diffusion SDDEs

$$\begin{aligned} d\boldsymbol{X}(t) =& \boldsymbol{a}(t, \boldsymbol{X}(t), \boldsymbol{X}(t-\gamma))dt + \boldsymbol{b}(t, \boldsymbol{X}(t), \boldsymbol{X}(t-\gamma))d\boldsymbol{W}_t \\ &+ \int_{\boldsymbol{\varepsilon}} \boldsymbol{c}(t, \boldsymbol{X}(t^-), \upsilon)p_\varphi(d\upsilon, dt) \end{aligned} \tag{3.1}$$

subject to the initial condition

$$X(\theta) = \chi(\theta) \quad \text{for} \quad \theta \in [-\gamma, 0] \tag{3.2}$$

where $t \in [0, T]$, γ is the time delay which is assumed to be constant at all time, $\boldsymbol{W}_t = \{(W_t^1, \ldots, W_t^m), t \in [0, T]\}$ is an $\mathcal{A}$-adapted m-dimensional standard Wiener process defined on the probability apace $(\Omega, \mathcal{A}, P)$, and p_φ denotes the Poisson random measure. Also here we denote by $\boldsymbol{X}(t^-)$ the almost sure left-hand limit of $\boldsymbol{X}(t)$. The coefficient $\boldsymbol{a}(t, x, x^r) : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $\boldsymbol{c}(t, x, \upsilon) : [0, T] \times \mathbb{R}^d \times \boldsymbol{\varepsilon} \to \mathbb{R}^d$ are d-dimensional vectors of Borel measurable functions. Further, $\boldsymbol{b}(t, x, x^r)$ defined on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ is a $d \times m$-matrix of Borel measurable functions.

Discrete-time approximations can be classified two major catalogs: strong approximations and weak approximations, as detailed in Kloeden & Platen (1999). A discrete-time approximation Y^Δ on a time discretization $(t)_\Delta$ with a maximum step size $\Delta > 0$, converges to the solution X of a given SDE with strong order α at time T, if there exists a positive constant C, independent of Δ, and a finite

number $\Delta_0 \in (0,1)$, such that

$$E(|X_T - Y_T^{\Delta}|) \leq C\Delta^{\alpha}$$

for all $\Delta \in (0, \Delta_0)$. From the definition of the strong error on the left hand side of the equation above, we notice that the strong numerical schemes provide pathwise approximations for the solution X of the given SDE. These methods therefore can be used for problems such as filtering, asset price simulation, hedge strategy simulation and other quantitative testing methods. Particularly, strong approximations are well suited for applications in the area of dynamic financial analysis.

On the other hand, a discrete-time approximation Y^{Δ} converges weakly with order β to X at time T, if for each $g \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ there exists a positive constant C, independent of Δ, and a finite number $\Delta_0 \in (0,1)$, such that

$$|E(g(X_T)) - E(g(Y_T^{\Delta}))| \leq C\Delta^{\beta}$$

for each $\Delta \in (0, \Delta_0)$. Here by $\mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$, we denote the set of $2(\beta+1)$ continuously differentiable functions with polynomial growth. This means that for any given $g \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ there exist constants $K > 0$ and $r \in N$, possibly depending on g, such that

$$|\partial_y^j g(y)| \leq K(1 + |y|^{2r})$$

for all $y \in \mathbb{R}^d$ and any partial derivative $\partial_y^j g(y)$ with the order of $j \leq 2(\beta+1)$. Weak schemes provide approximations of the probability functionals generated by the solution X of a given SDE. These schemes are suitable for problems such as financial derivative pricing, the evaluation of moments and risks and expected utilities of investors.

3.3 Existense and Uniqueness of Solutions of S-DDEs with Jumps

In this section, we give some basic concepts and definitions to be used in later sections, then establish the conditions for the existence and uniqueness of solutions to the Jump diffusion SDDEs (3.1), and then give some lemmas to be used later for the proof of the convergence theorem. In this work, we assume that the d-dimensional vector valued function for the initial condition, $\chi = \{\chi(s), s \in [-\gamma, 0]\}$, is right-continuous and has left-hand limits.

From equation (3.1), we have the following integral form of the jump diffusion equation SDDE,

$$\begin{aligned}\boldsymbol{X}(t) =& \boldsymbol{X}(0) + \int_0^t \boldsymbol{a}(\tau, \boldsymbol{X}(\tau), \boldsymbol{X}(\tau-\gamma))d\tau + \int_0^t \boldsymbol{b}(\tau, \boldsymbol{X}(\tau), \boldsymbol{X}(\tau-\gamma))d\boldsymbol{W}_\tau \\ &+ \sum_{i=1}^{p_\varphi(t)} \boldsymbol{c}(\tau_i, \boldsymbol{X}(\tau_i^-), \xi_i),\end{aligned} \tag{3.3}$$

where (τ_i, ξ_i), for $i \in \{1, 2, 3, \ldots, p_\varphi(t)\}$, denote a sequence of pairs of jump times and corresponding values generated by the Poisson random measure p_φ.

Definition 3.1. *Given a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$, a stochastic process given by $\boldsymbol{X} = \{\boldsymbol{X}(t), t \in [-\gamma, T]\}$ is known as a solution of the equation* (3.1) *subject to the initial condition* (3.2) *if $\boldsymbol{X}$ is $\underline{\mathcal{A}}$-adapted, the integrals in the equation are well-defined and the equalities* (3.3) *and* (3.2) *hold almost surely. Moreover, if any two solution processes $\boldsymbol{X}^{(i)} = \{\boldsymbol{X}^{(i)}(t), i \in 1, 2\}$ are indistinguishable on $[-\gamma, T]$ with the same initial segment χ and the same path on $[0, T]$, and*

$$P\left(\sup_{t\in[0,T]} \parallel X^{(1)}(t) - X^{(2)}(t) \parallel > 0\right) = 0 \tag{3.4}$$

where $\parallel . \parallel$ is the Euclidean norm, then if (3.1) *has a solution, it is a unique*

solution for this initial value problem.

To guarantee the existence of a unique solution of the jump diffusion SDDE (3.1), we assume that the coefficients of (3.1) satisfy the following Lipschitz conditions

$$
\begin{aligned}
&|\boldsymbol{a}(t,\boldsymbol{y}_1,\boldsymbol{z}_1)-\boldsymbol{a}(t,\boldsymbol{y}_2,\boldsymbol{z}_2)| \le C_1(|\boldsymbol{y}_1-\boldsymbol{y}_2|+|\boldsymbol{z}_1-\boldsymbol{z}_2|)\\
&|\boldsymbol{b}(t,\boldsymbol{y}_1,\boldsymbol{z}_1)-\boldsymbol{b}(t,\boldsymbol{y}_2,\boldsymbol{z}_2)| \le C_2(|\boldsymbol{y}_1-\boldsymbol{y}_2|+|\boldsymbol{z}_1-\boldsymbol{z}_2|)\\
&\int_{\varepsilon}|\boldsymbol{c}(t,\boldsymbol{y}_1,\upsilon)-\boldsymbol{c}(t,\boldsymbol{y}_2,\upsilon)|^2\varphi(d\upsilon) \le C_3(\boldsymbol{y}_1-\boldsymbol{y}_2)^2
\end{aligned}
$$

for $t\in[0,T]$ and $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\in\mathbb{R}^d$, as well as the growth conditions

$$
\begin{aligned}
&|\boldsymbol{a}(t,\boldsymbol{y},\boldsymbol{z})| \le D_1(1+|\boldsymbol{y}|+|\boldsymbol{z}|)\\
&|\boldsymbol{b}(t,\boldsymbol{y},\boldsymbol{z})| \le D_2(1+|\boldsymbol{y}|+|\boldsymbol{z}|)\\
&\int_{\varepsilon}|\boldsymbol{c}(t,\boldsymbol{y},\upsilon)|^2\varphi(d\upsilon) \le D_3(1+y^2)
\end{aligned}
$$

for $\boldsymbol{y},\boldsymbol{z}\in\mathbb{R}^d$, $t\in[0,T]$.

We denote by $\mathbb{C}=\mathbb{C}([-\gamma,0],\mathbb{R}^d)$ the Banach space of all d-dimensional continuous functions η on $[-\gamma,0]$ equipped with the supremums norm $||\eta||_{\mathbb{C}}=\sup_{s\in[-\gamma,0]}|\eta(s)|$. Furthermore, we suppose that the set $L_2(\Omega,\mathbb{C},\mathcal{A}_0)$ of the $\mathbb{R}^d$-valued continuous process $\eta=\{\eta(s),s\in[-\gamma,0]\}$ is $\mathcal{A}_0$-measurable with

$$
E(\|\eta\|_{\mathbb{C}}^2)=E\left(\sup_{s\in[-\gamma,0]}|\eta(s)|^2\right)<\infty \tag{3.5}
$$

Following the work of [127, 155], the following theorem can be established for the existence and uniqueness of a solution to the problem defined by (3.1) and (3.2).

Theorem 3.1. *Suppose that the Lipschitz conditions and the growth conditions are satisfied, and the initial condition χ is in $L_2(\Omega,\mathbb{C},\mathcal{A}_0)$. Then equation* (3.1)

subject to the initial condition (3.2) *admits a unique solution.*

Now we present some lemmas to be used later for the proof of the convergence theorem. Consider a right continuous process $Y^{\Delta_l} = \{Y^{\Delta_l}(t), t \in [-\gamma, T]\}$. Y^{Δ_l} is called a discrete time numerical approximation with maximum step size Δ_l, if it is obtained by using a time discretization t_{Δ_l}, and the random variable $Y^{\Delta_l}_{t_n}$ is $\mathcal{F}_{t_n}$-measurable for $n \in \{1, ...N\}$. Further, $Y^{\Delta_l}_{t_{n+1}}$ can be expressed as a function of $Y^{\Delta_l}_{t_{-l}}$,$Y^{\Delta_l}_{t_{-l+1}}$,...,$Y^{\Delta_l}_{t_n}$ and the discrete time t_n.

Because of dealing with the approximation of solutions of jump diffusion SDDEs, we introduce a concept of weak order convergence due to Kloeden & Platen [107].

Definition 3.2. *A discrete time approximation Y^{Δ_l} converges weakly towards X at time T with order $\beta > 0$ if for each $g \in \mathcal{C}_p$ there is a constant C, independent of Δ_l, such that*

$$|E(g(X(T))) - E(g(Y^{\Delta_l}(T)))| \leq C(\Delta_l)^{\beta} \tag{3.6}$$

where $\mathcal{C}_p$ denotes the set of all polynomials $g : \mathbb{R}^d \to \mathbb{R}$.

We now give some auxiliary results to prepare for the proof of the Weak Convergence Theorem to be presented.

Lemma 3.1. *For $n \in \{-l+1, \ldots, 0, 1, \ldots, N\}$ and $\boldsymbol{z} \in \mathbb{R}^d$, we have*

$$E\left(u(n, \boldsymbol{X}_{n^-}^{n-1,z}) - u(n-1, \boldsymbol{z}) + \int_{n-1}^{n} \int_{\boldsymbol{\varepsilon}} \boldsymbol{L}_{\upsilon}^{-1} u(\tau, \boldsymbol{X}_{\tau}^{n-1,z}) \varphi(d\upsilon) d\tau | \mathcal{A}_{n-1}\right) = 0 \tag{3.7}$$

for $(\tau, \boldsymbol{z}) \in [-\gamma, t] \times \mathbb{R}^d$ and $u(\tau, \boldsymbol{z}) = E(g(\boldsymbol{X}_T^{\tau,z}) | \mathcal{A}_\tau)$

The proof of the lemma for the case with no delay was established by Platen & Bruti-Liberati [45], and a similar procedure can be used for the proof of this lemma.

Lemma 3.2. *Given $p \in \{1, 2, 3, \ldots\}$ there is a bounded constant M satisfying*

$$E\left(|\boldsymbol{X}_{n^-}^{n-1,z} - \boldsymbol{z}|^{2q} \middle| \widetilde{\mathcal{A}}_{n-1}\right) \leq M(1 + |\boldsymbol{z}|^{2q})(\Delta_l)^q \tag{3.8}$$

for $q \in \{1, \ldots, p\}$ and $n \in \{-l + 1, \ldots, 0, 1, \ldots, N\}$.

The proof of (3.8) can be obtained by following that of a lemma for SDEs with jumps but with no delay in [107]. The following results are similar to what was given in Mikulevicius & Platen [113].

Lemma 3.3. *Given $p \in \{1, 2, 3, \ldots\}$, there is a finite constant M satisfying*

$$E\left(\sup_{-\gamma \leq t \leq T} |\boldsymbol{\zeta}(t)|^{2q}\right) \leq M(1 + |\boldsymbol{Y}_0|^{2q}). \tag{3.9}$$

for every $q \in \{1, \ldots, p\}$

Lemma 3.4. *Given $p \in \{1, 2, \ldots\}$, there is $r \in \{1, 2, 3, \ldots\}$ and a bounded constant M satisfying*

$$\begin{aligned} &\left|E\left(\left|\boldsymbol{F}_{\boldsymbol{p}}(\boldsymbol{\zeta}(z) - \boldsymbol{Y}_z^{\Delta_l})\right|^{2q} + \left|\boldsymbol{F}_{\boldsymbol{p}}(\boldsymbol{X}_z^{z,Y_z^{\Delta_l}} - \boldsymbol{Y}_z^{\Delta_l})\right|^{2q} \middle| \widetilde{\mathcal{A}}_z\right)\right| \\ &\qquad \leq M(1 + |\boldsymbol{Y}_z^{\Delta_l}|^{2r})(\Delta_l)^{qk} \end{aligned} \tag{3.10}$$

for each $q \in \{1, \ldots, p\}$, $k \in \{1, \ldots, 2(\beta + 1)\}$, $\boldsymbol{p} \in P_k = \{1, 2, ..., d\}^k$ and $z \in [-r, T]$, where $F_{\boldsymbol{p}}(\boldsymbol{y}) = \prod_{h=1}^k y^{p_h}$ for all $\boldsymbol{y} = (y^1, ..., y^d)^T \in \Re^d$ and $\boldsymbol{p} = (p_1, ..., p_k) \in P_k$.

The proof of the estimate (3.10) can be established by following Itô's formula for SDEs with jumps but with no delay as in [107].

Lemma 3.5. *For $\boldsymbol{p} \in P_k$, there exist $r \in \{1, 2, \ldots\}$ and a finite constant M satisfying*

$$\begin{aligned} &\left|E\left(\boldsymbol{F}_{\boldsymbol{P}}(\boldsymbol{\zeta}(t) - \boldsymbol{Y}_{n-1}^{\Delta_l}) - \boldsymbol{F}_{\boldsymbol{P}}(\boldsymbol{X}_t^{n-1,Y_{n-1}^{\Delta_l}} - \boldsymbol{Y}_{t_{n-1}}^{\Delta})\middle| \widetilde{\mathcal{A}}_t\right)\right| \\ &\qquad \leq M(1 + |\boldsymbol{Y}_{n-1}^{\Delta_l}|^r)(\Delta_l)^{\beta} \end{aligned} \tag{3.11}$$

for each $k \in \{1, \ldots, 2\beta + 1\}$, $n \in \{-l + 1, \ldots, 0, 1, \ldots, N\}$ *and* $t \in [t_{n-1}, t_n)$.

3.4 The Jump-adapted Weak Taylor Approximation Scheme

3.4.1 Multi-dimensional Ito Formula

For multi-factor financial models, the following generalization of the Ito formula is required. We consider a d-dimensional vector process $\boldsymbol{e} = \{e_t = (e_t^1, \ldots, e_t^d)^T, t \geq 0\}$ wherein all the components $e^k, k \in 1, 2, \ldots, d$ are predictable. Assume that

$$\int_0^T |e_z^k| dz < \infty$$

almost surely for all $k \in 1, 2, \ldots, d$. The $d \times m$-matrix process $\boldsymbol{F} = \{F_t = [F_t^{i,j}]_{i,j=1}^{d,m}, t0\}$ is supposed to have predictable elements $F^{i,j}$ with

$$\int_0^T (F_z^{i,j})^2 dz < \infty$$

almost surely for $i \in 1, 2, \ldots, d, j \in 1, 2, \ldots, m$ and all $T \in (0, \infty)$, see Protter (2005). Now we introduce a d-dimensional continuous-time stochastic vector process $\boldsymbol{X} = \{X_t = (X_t^1, X_t^2, \ldots, X_t^d)^T, t \geq 0\}$, where the kth component X^k is given by the following Ito differential

$$dX_t^k = e_t^k dt + \sum_{j=1}^{m} F_t^{k,j} dW_t^j$$

for $t \geq 0$ and a given initial value $\boldsymbol{X}_0 = (X_0^1, \ldots, X_0^d)^T \in \mathbb{R}^d$.

Now we consider a function $u : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ with continuous partial derivatives $\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x^k}$ and $\frac{\partial^2 u}{\partial x^k \partial x^i}$ for all $k, i \in 1, 2, \ldots, d, t \geq 0$ and $\boldsymbol{x} = (x^1, x^2, \ldots, x^d)^T$. The Ito formula for the scalar stochastic process $u = u(t, X_t^1, X_t^2, \ldots, X_t^d), t \geq 0$

is given by the following Ito derivative

$$\begin{aligned} &du(t, X_t^1, X_t^2, \ldots, X_t^d) \\ &\quad = \Big(\frac{\partial u}{\partial t} + \sum_{k=1}^{d} e_t^k \frac{\partial u}{\partial x^k} + \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{d} F_t^{i,j} F_t^{k,j} \frac{\partial^2 u}{\partial x^i \partial x^k}\Big) dt \\ &\quad + \sum_{j=1}^{m}\sum_{i=1}^{d} F_t^{i,j} \frac{\partial u}{\partial x^i} dW_t^j \end{aligned} \tag{3.12}$$

for $t \geq 0$ with initial value $u(0, X_0^1, X_0^2, \ldots, X_0^d)$, and its partial derivatives of the function u are evaluated at $(t, X_t^1, X_t^2, \ldots, X_t^d)$, which we suppressed in our notation.

Now, let us consider the Ito formula when the Wiener processes and Poisson jump measures drive the stochastic dynamical systems. Assume that $\boldsymbol{W} = \{W_t = (W_t^1, \ldots, W_t^m)^T, t \geq 0\}$ is an m-dimensional standard Wiener process and by $p_{\phi_r}^r(dv, dt)$, we denote the rth Poisson measure on $\boldsymbol{\varepsilon} \times [0, \infty)$ with the intensity measure

$$\mu_{\phi_r}^r(dv, dt) = \phi_r(d\mu)dt$$

$r \in \{m+1, m+2, \ldots, l\}, l \in \{m+1, m+2, \ldots\}$. Suppose that the underlying d-dimensional process X with the ith component X_t^i at time t has the following Ito differential

$$dX_t^i = a_t^i + \sum_{k=1}^{m} b_t^{i,k} dW_t^k + \sum_{r=m+1}^{l} \int_{\varepsilon} c^{i,r}(v, t-) p_{\phi_r}^r(dv, dt)$$

for $t \geq 0$ and $i \in 1, 2, \ldots, d$, where $a^i, b^{i,k}$ and $c^{i,r}$ are predictable processes and $\boldsymbol{\varepsilon} = \mathbb{R}\{0\}$ is the Poisson mark pace. Therefore, for a function $u : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ same as defined in the Ito formula without jumps, the Ito formula under the

Poisson measure has the form as following

$$
\begin{aligned}
&du(t, X_t^1, X_t^2, \ldots, X_t^d) \\
&\quad = \Big(\frac{\partial u}{\partial t} + \sum_{d=1}^{d} a_t^i \frac{\partial u}{\partial x^i} + \frac{1}{2}\sum_{i,j=1}^{d}\sum_{k=1}^{m} b_t^{i,k} b_t^{j,k} \frac{\partial^2 u}{\partial x^i \partial x^j}\Big) dt \\
&\quad + \sum_{k=1}^{m}\sum_{i=1}^{d} b_t^{i,k} \frac{\partial u}{\partial x^i} dW_t^k \\
&\quad + \sum_{r=m+1}^{l} \int_{\epsilon} (u(t, X_t^1, X_t^2, \ldots, X_t^d) - u(t-, X_{t-}^1, X_{t-}^2, \ldots, X_{t-}^d)) p_{\phi_r}^r (dv, dt)
\end{aligned}
\tag{3.13}
$$

for $t \geq 0$. This general versin of the Ito formula can be employed in the problems which may include Levy processes with infinite jump intensity as underlying factors.

3.4.2 Jump-adapted Numerical Approximation Scheme

In Monte Carlo simulations for functionals of jump diffusion SDDEs, one uses numerical approximations evaluated only at discretization time. Here, we first give a jump adapted weak approximation Taylor Scheme of order β, then study the basic properties of the discrete Taylor approximation in a weak order sense.

First we define the jump-adapted time discretization. Let $T > r > 0$. The jump adapted time discretization used throughout this paper is

$(t)_\Delta = \{t_i : i = \{-l, -l+1, ..., 0, 1, 2, ..., N\}\}$

where $\{t_i, i < 0\}$ represent the delay time, and the maximum step size Δ_l satisfies $\Delta_l \in (0, 1)$. We choose the time discretization in such a way that all jump times are at the nodes of the time discretization. If the discretization node t_i is not a jump time, then t_i is $\mathcal{A}_{t_{i-1}}$-measurable. Otherwise, t_i is $\mathcal{A}_{t_{i^-}}$-measurable. Also, throughout the paper, we denote the set of all multi-indices α by

$\mathcal{M}_m = \{(j_1, \ldots, j_l) : j_i \in \{0, 1, 2, \ldots, m\}, i \in \{1, 2, \ldots, l\} \text{for} l \in \mathcal{N}\} \cup \{v\}$,

where the element $\alpha = (j_1, j_2, \ldots, j_l)$ is called a multi-index of length $l = l(\alpha) \in \mathcal{N}$ and v has zero length. In the following, by a component $j \in \{0, 1, 2, \ldots, m\}$ of a multi-index we refer to the integration with respect to the jth Wiener process in a multiple stochastic integral. A component with $j = 0$ corresponds to integration with respect to time t.

Now define the following operators for the coefficient functions

$$\begin{aligned} L^0 =& \frac{\partial}{\partial t} + \sum_{i=1}^{d} a^i(t, x, x_l)\frac{\partial}{\partial x^i} + \sum_{i=1}^{d} a^i(t-l, x_l, x_{2l})\frac{\partial}{\partial x_l^i} \\ &+ \frac{1}{2}\sum_{i,\gamma=1}^{d}\sum_{j=1}^{m} b^{ij}(t, x, x_l)b^{\gamma j}(t, x, x_l)\frac{\partial^2}{\partial x^i \partial x^\gamma} \\ &+ \frac{1}{2}\sum_{i,\gamma=1}^{d}\sum_{j=1}^{m} b^{ij}(t-l, x_l, x_{2l})b^{\gamma j}(t-l, x_l, x_{2l})\frac{\partial^2}{\partial x_l^i \partial x_l^\gamma}, \end{aligned} \tag{3.14}$$

$$L^k = \sum_{i=1}^{d} b^{ik}(t, x, x_l)\frac{\partial}{\partial x^i} + \sum_{i=1}^{d} b^{ik}(t-l, x_l, x_{2l})\frac{\partial}{\partial x_l^i}. \tag{3.15}$$

A subset $\mathcal{A} \in \mathcal{M}$ is the hierarchical, and its corresponding remainder set $\bar{A}(\mathcal{M})$ is defined by $\bar{A}(\mathcal{M}) = \{\alpha \in \mathcal{M}_m \backslash \mathcal{A} : -\alpha \in \mathcal{A}\}$. For each $\beta \in 1, 2, 3, \ldots$, we can then define the hierarchical set $\Gamma_\beta = \{\alpha \in \mathcal{M}_m : l(\alpha) \leq \beta\}$. The weak Taylor method of order β is then constructed as follows

$$\boldsymbol{Y}_{(n+1)^-} = \boldsymbol{Y}_n + \sum_{\alpha \in \Gamma_\beta} f_\alpha(n, \boldsymbol{Y}_n, \boldsymbol{Y}_{n-l}) I_\alpha \tag{3.16}$$

and

$$\boldsymbol{Y}_{n+1} = \boldsymbol{Y}_{(n+1)^-} + \int_\varepsilon c(n, \boldsymbol{Y}_{(n+1)^-}, \upsilon) p_\varphi(d\upsilon, (n+1)) \tag{3.17}$$

where

$$f_\alpha(t, x, u) = \begin{cases} f(t, x) & \text{if} \quad l(\alpha) = 0 \\ L^{j_1} f_{-\alpha}(t, x, u) & \text{if} \quad l(\alpha) \geq 1, j_1 \in 0, 1, \ldots, m \end{cases} \tag{3.18}$$

The multiple stochastic integral is then defined recursively as follows

$$I_{\alpha,t} = \begin{cases} t & \text{if} \quad l = 0 \\ \int_0^t I_{\alpha-,z} dz & \text{if} \quad l \geq 1 \quad , \quad j_l = 0 \\ \int_0^t I_{\alpha-,z} dW_z^{jl} & \text{if} \quad l \geq 1 \quad , \quad j_l \in 1, \ldots, m \end{cases} \tag{3.19}$$

where $\alpha-$ is obtained from α by deleting its last component, while $-\alpha$ is obtained from α by deleting its first component.

Now, we give the weak convergence theorem of the Taylor approximation with order β.

Theorem 3.2. *Given $\beta \in \{1, 2, \ldots\}$, let $Y^{\Delta_l} = \{Y_n^{\Delta_l}, n \in [-l, \ldots, 0, 1, \ldots, N]\}$ be the results obtained from the Taylor Scheme 3.16-3.17 corresponding to $(t)_\Delta$ with maximum step size $\Delta_l \in (0, 1)$. Suppose that $E(|\boldsymbol{X}_\xi|^i) < \infty$ for $\xi \in (-\gamma, 0)$,$i \in \{1, 2, \ldots\}$, and $\boldsymbol{Y}_\xi^\Delta$ converges to $\boldsymbol{X}_\xi$ weakly with order $\beta \in \{1, 2, \ldots\}$. Assume that the coefficients, a^k, b^{kj}, c^k, are in the space $\mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$, for $j \in \{1, 2, \ldots, m\}$ and $k \in \{1, 2, \ldots, d\}$, and the coefficient functions f_α, with $f(t, \boldsymbol{y}) = \boldsymbol{y}$, satisfy the growth condition $|f_\alpha(t, \boldsymbol{y})| \leq M(1 + |\boldsymbol{y}|)$, with $M < \infty$, for all $t < T, \boldsymbol{y} \in \mathbb{R}^d$ and $\alpha \in \Gamma_\beta$. Then for any $g \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ there is a positive constant C, which does not depend on Δ, such that*

$$|E(g(X(T))) - E(g(Y^{\Delta_l}(T)))| \leq C(\Delta_l)^\beta. \tag{3.20}$$

3.4.3 Proof of Weak Convergence Theorem

In this subsection, we prove the convergence theorem 3.2 proposed in last subsection.

Proof. For $\beta \in \{1, 2, 3, \ldots\}$ and $g \in \mathcal{C}_p^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$, consider the Ito process below

$$\boldsymbol{X}_t^{z,y} = \boldsymbol{y} + \int_z^t \boldsymbol{a}(\boldsymbol{X}_u^{z,y}) du + \int_z^t \boldsymbol{b}(\boldsymbol{X}_u^{z,y}) d\boldsymbol{W}_u + \int_z^t \int_{\boldsymbol{\varepsilon}} \boldsymbol{c}(\boldsymbol{X}_{u-}^{z,y}, \upsilon) p_\varphi(d\upsilon, du) \tag{3.21}$$

Then we can get $u(0, X_0) = E(g(X_T^{0,X_0})) = E(g(X_T))$.

Define also the process $\zeta = \zeta(t), t \in (-\gamma, T)$ by

$$\zeta(t) = \zeta(t_n) + \sum_{\alpha \in \Gamma_\beta} I_\alpha(f_\alpha(n, \boldsymbol{\zeta}_n, \boldsymbol{\zeta}_{n-l})) + \int_n^t \int_\varepsilon c(n, \boldsymbol{\zeta}_{(n+1)^-}, \upsilon) p_\varphi(d\upsilon, (n+1)) \quad (3.22)$$

for $n \in \{-l, -l+1, \ldots, 0, 1, \ldots, N-1\}$, $t \in (t_n, t_{n+1}]$, $\zeta(0) = Y_0$ and $\zeta(t_n) = Y_{t_n}$ for $n \in \{-l, \ldots, 0, 1, 2, \ldots, N\}$.

By the definition of the functional u and the terminal condition of the stochastic process X, we have

$$H = |E(g(\boldsymbol{Y}_T^{\Delta_l})) - E(g(\boldsymbol{X}_T))| = |E(u(T, \boldsymbol{Y}_T^{\Delta_l}) - u(0, \boldsymbol{X}_0))|. \quad (3.23)$$

Since $\boldsymbol{Y}_0$ converges towards $\boldsymbol{X}_0$ weakly with order β, one has

$$H \le \left| E\Big(\sum_{n=-l+1}^{N} (u(n, \boldsymbol{Y}_n) - u(n, \boldsymbol{Y}_{n^-}) + u(n, \boldsymbol{Y}_{n^-}) - u(n-1, \boldsymbol{Y}_{n-1})) \Big) \right| + K(\Delta_l)^\beta. \quad (3.24)$$

By Lemma 3.1, we can write

$$\begin{aligned} H \le \Bigg| E\Bigg(\sum_{n=-l+1}^{N} \Big\{ & [u(n, \boldsymbol{Y}_n) - u(n, \boldsymbol{Y}_{n^-}) + u(n, \boldsymbol{Y}_{n^-}) - u(n-1, \boldsymbol{Y}_{n-1})] \\ & - [u(n, \boldsymbol{X}_{n^-}^{n-1, Y_{n-1}}) - u(n-1, \boldsymbol{Y}_{n-1})] \\ & + \int_{n-1}^{n} \int_{\boldsymbol{\varepsilon}} \boldsymbol{L}_\upsilon^{-1} u(z, \boldsymbol{X}_z^{n-1, Y_{n-1}}) \varphi(d\upsilon) dz] \Big\} \Bigg) \Bigg| + K(\Delta_l)^\beta. \end{aligned} \quad (3.25)$$

From the properties of stochastic integrals, we obtain

$$E\Big(\sum_{n=-l}^{N} [u(n, \boldsymbol{Y}_n) - u(n, \boldsymbol{Y}_{n^-})] \Big) = E\Big(\int_{-r}^{T} \int_{\boldsymbol{\varepsilon}} \boldsymbol{L}_\upsilon^{-1} u(z, \boldsymbol{\zeta}(z)) \varphi(d\upsilon) dz \Big). \quad (3.26)$$

Therefore, we have

$$H \le H_1 + H_2 + K(\Delta_l)^\beta, \quad (3.27)$$

where

$$H_1 = \left| E\Big(\sum_{n=-l+1}^{N} \Big\{ [u(n, \boldsymbol{Y}_{n^-}) - u(n, \boldsymbol{Y}_{n-1})] - [u(n, \boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}) - u(n, \boldsymbol{Y}_{n-1})] \Big\} \Big) \right| \tag{3.28}$$

and

$$\begin{aligned} H_2 = \Bigg| E\Bigg(\int_{-r}^{T} \int_{\boldsymbol{\varepsilon}} \Big\{ & [L_\upsilon^{-1} u(z, \boldsymbol{\zeta}(z)) - L_\upsilon^{-1} u(z, \boldsymbol{Y}_z)] \\ & - [L_\upsilon^{-1} u(z, \boldsymbol{X}_z^{z,Y_z}) - L_\upsilon^{-1} u(z, \boldsymbol{Y}_z)] \Big\} \varphi(d\upsilon) dz \Bigg) \Bigg|. \end{aligned} \tag{3.29}$$

In the following, we proceed to estimate H_1 and H_2 in step I and Step II repectively, and then complete the proof in Step III.

Step I: Let us assume that u is so smooth that the deterministic Taylor expansion may be applied. Hence, by expanding du in H_1, we get

$$\begin{aligned} H_1 = \Bigg| E\Bigg(\sum_{n=-l+1}^{N} \Bigg\{ & \Bigg[\sum_{k=1}^{2\beta+1} \frac{1}{k!} \sum_{\boldsymbol{P} \in P_k} (\partial_y^{\boldsymbol{P}} u(n, \boldsymbol{Y}_{n-1})) F_{\boldsymbol{P}}(\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1}) + R_n(\boldsymbol{Y}_{n^-}) \Bigg] \\ & - \Bigg[\sum_{k=1}^{2\beta+1} \frac{1}{k!} \sum_{\boldsymbol{P} \in P_k} (\partial_y^{\boldsymbol{P}} u(n, \boldsymbol{Y}_{n-1})) F_{\boldsymbol{P}}(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}} - \boldsymbol{Y}_{n-1}) \\ & + R_n(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}) \Bigg] \Bigg\} \Bigg) \Bigg|, \end{aligned} \tag{3.30}$$

where the remainder term is

$$R_n(\boldsymbol{Z}) = \frac{1}{2(\beta+1)!} \sum_{\boldsymbol{p} \in P_{2(\beta+1)}} \partial_y^{\boldsymbol{p}} u(n, \boldsymbol{Y}_{n-1} + \boldsymbol{\theta}_{\boldsymbol{p},n}(\boldsymbol{Z} - \boldsymbol{Y}_{n-1})) F_{\boldsymbol{p}}(\boldsymbol{Z} - \boldsymbol{Y}_{n-1}) \tag{3.31}$$

where $\boldsymbol{\theta}_{\boldsymbol{p},n}(\boldsymbol{Z})$ is a $d \times d$ diagonal matrix with $\theta_{\boldsymbol{p},n}^{k,k}(\boldsymbol{Z}) \in (0,1)$ for $k \in \{1, 2, 3, \ldots, d\}$, and $\boldsymbol{Z} = \boldsymbol{Y}_{n^-}$ and $\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}$, respectively.

Therefore, according to the properties of expectation and absolute value, we

get

$$\begin{aligned}
H_1 \leq & E\Bigg(\sum_{n=-l+1}^{N} \Bigg\{ \sum_{k=1}^{2\beta+1} \frac{1}{k!} \sum_{\boldsymbol{p}\in P_k} |(\partial_y^{\boldsymbol{p}} u(n, \boldsymbol{Y}_{n-1}))| \\
& \times (F_{\boldsymbol{p}}(\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1}) - F_{\boldsymbol{p}}(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}} - \boldsymbol{Y}_{n-1})) \\
& |R_n(\boldsymbol{Y}_{n^-})| + |R_n(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}})| \Bigg\} \Bigg). \\
\leq & E\Bigg(\sum_{n=-l+1}^{N} \Bigg\{ \sum_{k=1}^{2\beta+1} \frac{1}{k!} \sum_{\boldsymbol{p}\in P_k} |(\partial_y^{\boldsymbol{p}} u(n, \boldsymbol{Y}_{n-1}))| \\
& \times |E(F_{\boldsymbol{p}}(\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1}) - F_{\boldsymbol{p}}(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}} - \boldsymbol{Y}_{n-1})|\widetilde{\mathcal{A}}_{n-1})| \\
& + E(|R_n(\boldsymbol{Y}_{n^-})||\widetilde{\mathcal{A}}_{n-1}) + E(|R_n(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}})||\widetilde{\mathcal{A}}_{n-1}) \Bigg\} \Bigg).
\end{aligned} \tag{3.32}$$

By equation (3.31), the hölder inequality and Lemma 3.4, we get

$$\begin{aligned}
& E(|R_n(\boldsymbol{Y}_{n^-})||\widetilde{\mathcal{A}}_{n-1}) \\
& \leq M \sum_{\boldsymbol{p}\in P_{2(\beta+1)}} \Big[E\Big(|\partial_y^{\boldsymbol{p}} u(n, \boldsymbol{Y}_{n-1} + \boldsymbol{\theta}_{\boldsymbol{p},n}(\boldsymbol{Y}_{n^-})(\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1}))|^2 \Big| \widetilde{\mathcal{A}}_{n-1} \Big) \Big]^{\frac{1}{2}} \\
& \times \Big[E\Big(|F_{\boldsymbol{p}}(\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1})|^2 \Big| \widetilde{\mathcal{A}}_{n-1} \Big) \Big]^{\frac{1}{2}} \\
& \leq M \Big[E\Big(1 + |\boldsymbol{Y}_{n-1}|^{2r} + |\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1}|^{2r} \Big| \widetilde{\mathcal{A}}_{n-1} \Big) \Big]^{\frac{1}{2}} \\
& \times \Big[E\Big(|\boldsymbol{Y}_{n^-} - \boldsymbol{Y}_{n-1}|^{4(\beta+1)} \Big| \widetilde{\mathcal{A}}_{n-1} \Big) \Big]^{\frac{1}{2}} \\
& \leq M(1 + |\boldsymbol{Y}_{n-1}|^{2r})(\Delta_l)^{\beta+1}.
\end{aligned} \tag{3.33}$$

Similarly, by Lemma 3.2 and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&E(|R_n(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}})||\widetilde{\mathcal{A}}_{n-1}) \\
&\quad\le M\sum_{\boldsymbol{p}\in P_{2(\beta+1)}}\Big[E\Big(|\partial_y^{\boldsymbol{p}}u(n,\boldsymbol{Y}_{n-} \\
&\quad+\boldsymbol{\theta}_{\boldsymbol{p},n}(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}})(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}-\boldsymbol{Y}_{n-1}))|^2\Big|\widetilde{\mathcal{A}}_{n-1}\Big)\Big]^{\frac{1}{2}} \\
&\quad\times\Big[E\Big(|F_{\boldsymbol{p}}(\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}-\boldsymbol{Y}_{n-1})|^2\Big|\widetilde{\mathcal{A}}_{n-1}\Big)\Big]^{\frac{1}{2}} \\
&\quad\le M\Big[E\Big(1+|\boldsymbol{Y}_{n-1}|^{2r}+|\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}-\boldsymbol{Y}_{n-1}|^{2\gamma}\Big|\widetilde{\mathcal{A}}_{n-1}\Big)\Big]^{\frac{1}{2}} \\
&\quad\times\Big[E\Big(|\boldsymbol{X}_{n^-}^{n-1,Y_{n-1}}-\boldsymbol{Y}_{n-1}|^{4(\beta+1)}\Big|\widetilde{\mathcal{A}}_{n-1}\Big)\Big]^{\frac{1}{2}} \\
&\quad\le M(1+|\boldsymbol{Y}_{n-1}|^{2r})(\Delta_l)^{\beta+1}.
\end{aligned}\tag{3.34}$$

Now, from the Cauchy Schwarz inequality, Lemma 3.5, Lemma 3.3, and the inequalities (3.33) and (3.34), we obtain

$$\begin{aligned}
H_1&\le E\Big(K\sum_{n=-l+1}^{N}(1+|\boldsymbol{Y}_{n-1}|^{2\gamma})(\Delta_l)^{\beta}\Big) \\
&\le M(\Delta_l)^{\beta}\Big(1+E(\max_{-l\le n\le N}|\boldsymbol{Y}_n|^{2\gamma})\Big) \\
&\le M(\Delta_l)^{\beta}(1+|\boldsymbol{Y}_0|^{2\gamma})\le K(\Delta_l)^{\beta}
\end{aligned}\tag{3.35}$$

Step II: Now we estimate the term H_2 in inequality (3.27). By the jump

coefficient c and the smooth function u , applying the Taylor expansion yields

$$\begin{aligned}
H_2 =&\Bigg|E\Bigg(\int_{-r}^{T}\int_{\varepsilon}\Bigg\{\Bigg[\sum_{k=1}^{2\beta+1}\frac{1}{k!}\sum_{\boldsymbol{p}\in P_k}(\partial_y^{\boldsymbol{p}}L_{\upsilon}^{-1}u(z,\boldsymbol{Y}_z))F_{\boldsymbol{p}}(\boldsymbol{\zeta}(z)-\boldsymbol{Y}_{t_z})+R_n(\boldsymbol{\zeta}(z))\Bigg]\\
&-\Bigg[\sum_{k=1}^{2\beta+1}\frac{1}{k!}\sum_{\boldsymbol{p}\in P_k}(\partial_y^{\boldsymbol{p}}L_{\upsilon}^{-1}u(z,\boldsymbol{Y}_z))F_{\boldsymbol{p}}(\boldsymbol{X}_z^{z,Y_z}-\boldsymbol{Y}_{t_z})+R_n(\boldsymbol{X}_z^{z,Y_z})\Bigg]\Bigg\}\varphi(d\upsilon)dz\Bigg)\Bigg|\\
&\leq\int_{-r}^{T}\int_{\varepsilon}E\Bigg(\sum_{k=1}^{2\beta+1}\frac{1}{k!}\sum_{\boldsymbol{p}\in P_k}|(\partial_y^{\boldsymbol{p}}L_{\upsilon}^{-1}u(z,\boldsymbol{Y}_z))|\\
&\times|E(F_{\boldsymbol{p}}(\boldsymbol{\zeta}(z)-\boldsymbol{Y}_z)-F_{\boldsymbol{p}}(\boldsymbol{X}_z^{z,Y_z}-\boldsymbol{Y}_z)|\widetilde{\mathcal{A}}_z)|\\
&+E(|R_n(\boldsymbol{\zeta}(z))||\widetilde{\mathcal{A}}_z)+E(|R_n(\boldsymbol{X}_z^{z,Y_z})||\widetilde{\mathcal{A}}_z)\Bigg)\varphi(d\upsilon)dz
\end{aligned}\tag{3.36}$$

Similarly, we can estimate the reminders as follows

$$E(|R_n(\boldsymbol{\zeta}(z))||\widetilde{\mathcal{A}}_z)\leq M(1+|\boldsymbol{Y}_{t_z}|^{2r})(z-t_z)^{\beta+1},\tag{3.37}$$

$$E(|R_n(\boldsymbol{X}_z^{z,Y_z})||\widetilde{\mathcal{A}}_z)\leq M(1+|\boldsymbol{Y}_{t_z}|^{2r})(z-t_z)^{\beta+1}.\tag{3.38}$$

Then, by applying the hölder inequality, Lemma 3.5 and Lemma 3.3, inequalities (3.37) and (3.38) to estimate the inequality above, we get

$$\begin{aligned}
H_2&\leq M\int_{-r}^{T}\int_{\varepsilon}E(1+|\boldsymbol{Y}_{t_z}|^{2r})(\Delta_l)^{\beta}\varphi(d\upsilon)dz\\
&\leq M(\Delta_l)^{\beta}\int_{0}^{T}E(1+\max_{0\leq n\leq n_T}|\boldsymbol{Y}_{t_n}|^{2r})(z-t_z)dz\\
&\leq M(\Delta_l)^{\beta}
\end{aligned}\tag{3.39}$$

Step III: Finally, by the inequalities (3.27) and (3.35) as well as (3.39), we have

$$|E(g(X(T)))-E(g(Y^{\Delta_l}(T)))|\leq M(\Delta_l)^{\beta}\tag{3.40}$$

□

3.5 A Numerical Example

Here we give an illustrative example to demonstrate the application and the convergence of the proposed numerical scheme . We consider the following linear SDDE with Poisson jumps,

$$\begin{cases} dX_t = [(\mu - \nu\lambda)X_t + \alpha X_{t-1}]dt + [\sigma X_t + \beta X_{t-1}]dW_t + \nu X_t dN \\ X(t) = t + 1, t \in [-1, 0] \end{cases} \tag{3.41}$$

where μ, σ and ν are respectively the drift coefficient, the diffusion coefficient and the jump coefficient, λ is the jump intensity, and α and β are the delay coefficients.

By the method for solving linear stochastic differential equations in [46], the analytical solution for $t \in [0, 1]$ is obtained

$$X(t) = \Phi(t)\left(1 + \int_0^t \Phi(s)^{-1}(\alpha - \sigma\beta)sds + \int_0^t \Phi(s)^{-1}\beta s dW_s\right) \tag{3.42}$$

where

$$\Phi(t) = (\nu + 1)^{N(t)} exp\{(\mu - \nu\lambda - \frac{\sigma^2}{2})t + \sigma W_t\} \tag{3.43}$$

According to the weak Taylor approximation scheme (3.16)-(3.19) proposed in section 3, we now expand it with weak order 1 (well known as Euler scheme),

$$\begin{cases} Y_{(n+1)^-} = Y_n + ((\mu - \nu\lambda)Y_n + \alpha hn)h + (\sigma Y_n + \beta hn)\Delta W_n \\ Y_{n+1} = Y_{(n+1)^-} + \nu Y_n \Delta N_n \end{cases} \tag{3.44}$$

Here we have used the jump adapted time discretization, and h is the maximum step size.

For higher accuracy and efficiency, one needs to construct higher order nu-

merical schemes. We now give a Taylor scheme of weak order two below,

$$\begin{cases} Y_{(n+1)^-} = Y_n + ((\mu - \nu\lambda)Y_n + \alpha hn)h + (\sigma Y_n + \beta hn)\Delta W_n \\ \qquad + ((\mu - \nu\lambda)(\sigma Y_n + \beta hn) + \sigma((\mu - \nu\lambda)Y_n + \alpha hn))\dfrac{h}{2}\Delta W_n \\ \qquad + (\mu - \nu\lambda)((\mu - \nu\lambda)Y_n + \alpha hn)\dfrac{h^2}{2} \\ \qquad + \sigma(\sigma Y_n + \beta hn)\dfrac{(\Delta W_n)^2 - h}{2} \\ Y_{n+1} = Y_{(n+1)^-} + \nu Y_n \Delta N_n \end{cases} \tag{3.45}$$

Next, we study the convergence of the two numerical schemes presented above by using the weak errors measured by

$$\varepsilon(h) = |E(X(T)) - E(Y(T))|, \tag{3.46}$$

and compare the results obtained from these two schemes to the explicit eaxact solution. We estimate the weak errors $\varepsilon(h)$ by running a very large number of simulations. The exact number depends on the implemented scheme. We use the following parameters: $\alpha = 0.01$, $\beta = 0.01$, $\mu = 0.001$, $\sigma = 0.6$, $\nu = 0.002$ and $\lambda = 0.001$.

In figure 3.1, we give the sample paths under the two approximation schemes and the numerical explicit solution of the equation 3.41. We can see from the figure that the weak Taylor scheme path is closer to the analytical solution line than the Euler scheme.

Now we present the numerical errors generated by the two numerical schemes presented above. From table 3.1, we notice that, for all the step sizes used in the numerical experiments, the weak Taylor method is more accurate. Moreover, the errors of the weak order two Taylor method decrease faster than the Euler scheme.

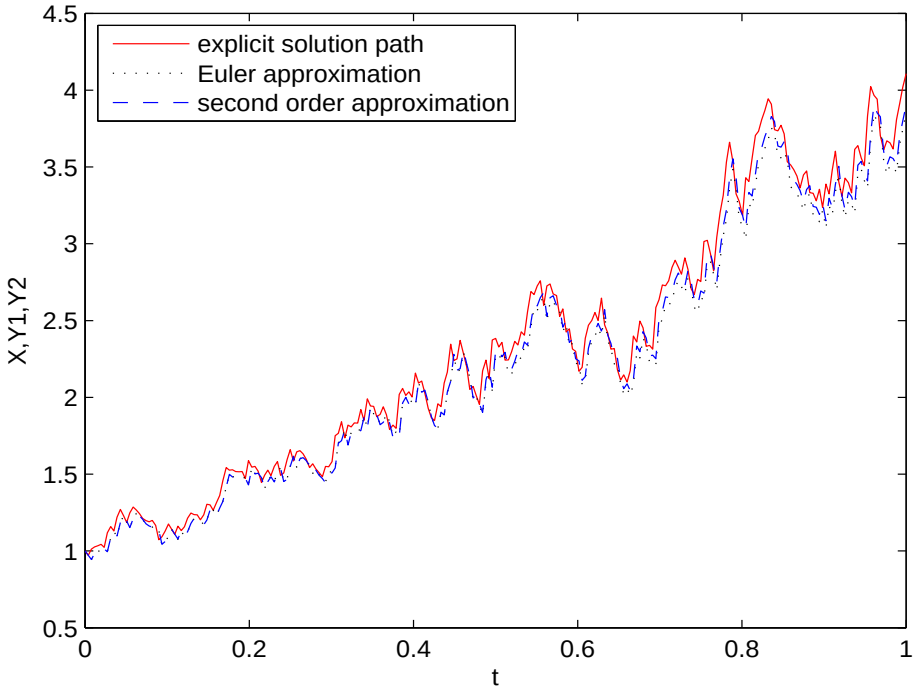

Figure 3.1: Sample paths of linear SDDE with jumps

Table 3.1: Convergence results for the linear SDDE with jumps

Stepsize $1/h$	2^8	2^{10}	2^{12}	2^{14}	2^{16}	2^{18}	2^{20}
WeakError1	0.00352	0.00201	0.00156	0.00114	0.00082	0.00060	0.00031
WeakError2	0.00229	0.00142	0.00061	0.00020	0.00009	0.00004	0.00002

3.6 Concluding Remarks

In this work, we have extended previous research on weak convergence to a more general class of stochastic differential equations involving both jumps and time delay. We proved that under the Poisson random measure and a fixed time delay, a simplified Taylor method gives weak convergence rate arbitrarily close to order β. The results of convergence analysis on a linear stochastic delay differential equation with jumps are reported.

CHAPTER 4

Fractional Stochastic Differential Equation with Application in European Option Pricing

4.1 General Overview

Memory effect is an important phenomenon in financial systems. A great deal of research work has been carried out to study the long memory in the financial markets. In recent years, fractional order ordinary differential equations are used as an effective instrument for describing the memory effect in complex systems. In this chapter, we establish a fractional order stochastic differential equation (FSDE) model to describe the effect of trend memory on financial pricing, and then, apply the model for European option pricing.

The rest of this chapter is organized as follows. In section 4.2, we construct a fractional stochastic differential equation model for application in financial market. Section 4.3 gives some basic concepts and theories on the fractional order ordinary differential equations and Hurst index and then establishes the fractional order stochastic differential equation for application in financial market. In Section 4.4, based on the proposed stochastic differential equation with fraction-

al order derivative, we give the corresponding Ito formula under the FSDE and then derive the fractional European option pricing formula. In Section 4.5, we conduct the empirical analysis of fractional order formula for the stock price process by using the Monte Carlo simulation method, and we also make a comparison analysis of the option pricing formula under the FSDE with the classic option pricing formula and the option pricing formula based on the fractional Brownian motion. The conclusions drawn from this study are presented in Section 4.5.

4.2 Fractional Stochastic Differential Equation Model

The fractional derivative is given as below:

$$d^{\alpha}X = a(X,t)dt^{\alpha} \tag{4.1}$$

where α is a fraction. This fractional differential equation is an appropriate mathematical approach to depict memory process of the increment. However, the fractional order derivative above only denotes the memory effect of a fixed process. Since the process in financial market has stochastic effect, we add stochastic process into fractional order ordinary differential equation. In this chapter, we propose a new model constructed by stochastic differential equation with fractional order. We denote the stochastic process of the asset price by fractional order stochastic differential equation as follows:

$$d^{\alpha}X = a(X,t)dt^{\alpha} + b(X,t)dW(t), \quad \alpha = 2H \tag{4.2}$$

In equation (4.2), $a(X,t), b(X,t)$ and $W(t)$ are respectively the drift coefficient, the diffusion coefficient and the standard Brownian motion, and H is the Hurst

index, which is an exponent describing the memory of the time series, and can be calculated by the R/SD analysis approach [66]. In the special case of $\alpha = 1$ (i.e., $H = 0.5$), the equation is reduced to the classic stochastic differential equation. Jumarie gave the Taylors series of fractional order, expressed dX in terms of fractional differential $d^{\alpha}X$ by using Taylors series of fractional order, and, hence, obtained the expression of $X(t)$, which involves the so-called Mittag-Leffler function [57, 58]. Momani and Odibat presented a numerical approach of differential equation of fractional order in [123]. Odibat proposed various algorithms to compute the functions of fractional derivative [163].

4.3 Fractional Differential Equations

In this section, we first give some preliminaries about the fractional order ordinary differential equation and then expand them to the field of the stochastic differential equations. Thus, based on these previous research results, we can construct the generalized fractional order stochastic differential equation.

4.3.1 Fractional Order Integration and Derivative

Now we introduce the definitions of fractional order integration and fractional order derivative. There exist several definitions of fractional derivatives, which are related to different applications. In our paper, we consider these two definitions, which are Riemann-Liouville integral and Caputo derivative [58].

Definition 4.1. *Suppose that $f(x)$ is a continuous function. Its Riemann-Liouville fractional integral of order α of function $f(x)$ is defined as follows:*

$$I^{\alpha}f(x) = \frac{1}{\Gamma(\alpha)}\int_0^x (x-t)^{\alpha-1}f(t)dt, \quad \alpha, x > 0, \tag{4.3}$$

where α is a fraction and $\Gamma(\alpha)$ is the Gamma function with

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} exp(-x) dx$$

Definition 4.2. *Consider the function of Definition 4.1. The Caputo fractional derivative of order α of function $f(x)$ is defined as*

$$\frac{d^\alpha f}{dx^\alpha} = D^\alpha f(x) = I^{m-\alpha} D^m f(x) = \frac{1}{\Gamma(m-\alpha)} \int_0^x (x-t)^{m-\alpha-1} f^{(m)}(t) dt \quad (4.4)$$

where α is a fraction, m is an integer and $m = [\alpha]$ is the value of α rounded up to the nearest integer, and $f^{(m)}$ is the ordinary derivative of f.

Based on the definitions above, the following equality holds [55, 56]:

$$f^\alpha(x) = \lim_{h \to 0} \frac{\Delta^\alpha f(x)}{h^\alpha} \quad (4.5)$$

In order to get the relations between the fractional derivative and ordinary derivative, we introduce the Taylor expansion of fractional order.

Proposition 4.1. *Assume that the continuous function $f(x)$ has fractional derivative of fractional order $k\alpha$, for any positive integer k at any $\alpha, 0 < \alpha < 1$; then the following equality holds:*

$$f(x+h) = \sum_{k=0}^{\infty} \frac{h^{k\alpha} f^{(k\alpha)}(x)}{\Gamma(1+k\alpha)}, \quad 0 < \alpha \leq 1 \quad (4.6)$$

where $f^{(k\alpha)}$ is the derivative of order $k\alpha$ of $f(x)$, which can be denoted by $D^{k\alpha} f$.

Lemma 4.1. *Assume that $m < \alpha < m+1, m \in N$; then,*

$$f^{(m)}(x+h) = \sum_{k=0}^{\infty} \frac{h^{k(\alpha-m)} D^{k(\alpha-m)} f^{(m)}(x)}{\Gamma(1+k(\alpha-m))}, \quad m < \alpha \leq m+1 \quad (4.7)$$

Let m be equal to 1 in equation (4.7), and take integration with respect to h;

we then have the following result:

$$f(x+h) = f(x) + hf'(x) + \sum_{k=1}^{\infty} \frac{h^{1+k(\alpha-1)}}{\Gamma(2+k(\alpha-1))} f^{(1+k(\alpha-1))}(x) \tag{4.8}$$

The proof of the lemma above can be found in [57].

By employing the fractional order Taylor formula and equation (4.5), we get the applications below. Given that m is an integer with $m \geq 1$, the following results hold:

$$f^{\alpha}(x) = \lim_{h\to 0} \frac{\Delta^{\alpha} f(x)}{h^{\alpha}} = \Gamma(1+\alpha) \lim_{h\to 0} \frac{\Delta f(x)}{h^{\alpha}}, \quad 0 < \alpha \leq 1 \tag{4.9}$$

$$f^{\alpha}(x) = \Gamma(1+(\alpha-m)) \lim_{h\to 0} \frac{\Delta f^{(m)}(x)}{h^{\alpha-m}}, \quad m < \alpha \leq m+1 (1 \geq m) \tag{4.10}$$

We then compare the two equations (4.8) and (4.9), when $1 < \alpha < 2$; thus, the relationship between fractional difference and finite difference is obtained as follows:

(1) Discrete form:

$$\Delta^{\alpha} f = \Gamma(1+\alpha)\Delta f, \quad 0 < \alpha \leq 1 \tag{4.11}$$

Continuous form:

$$d^{\alpha} f = \Gamma(1+\alpha) df, \quad 0 < \alpha \leq 1 \tag{4.12}$$

(2) Discrete form:

$$\Delta^{\alpha} f = \Gamma(1+\alpha)[\Delta f - f'(x)\Delta x], \quad 1 < \alpha < 2 \tag{4.13}$$

Continuous form:

$$d^{\alpha}f = \Gamma(1+\alpha)[df - f'(x)dx], \quad 1 < \alpha < 2 \tag{4.14}$$

For the purpose of constructing the fractional order stochastic differential equations in this section, now we give some results of the integral with respect to dt^{α} in Lemma 4.2 presented below. Its detailed proof can be obtained in [57, 58].

Lemma 4.2. *Let $f(t)$ denote a continuous function; then its integral with respect to dt^{α} is defined by the following equalities:*

$$\int_0^t f(\tau)(d\tau)^{\alpha} = \alpha \int_0^t (t-\tau)^{\alpha-1} f(\tau)d\tau, \quad 0 < \alpha \leq 1, \tag{4.15}$$

$$\int_0^t f(\tau)(d\tau)^{\alpha} = \alpha(\alpha-1) \int_0^t (t-\tau)^{\alpha-2} F(\tau)d\tau, \quad 1 < \alpha \leq 2 \tag{4.16}$$

where $F(t) = \int_0^t f(\tau)d\tau$; on making $f(t) = 1$, we can have the result: $\int_0^t f(\tau)(d\tau)^{\alpha} = t^{\alpha}$.

4.3.2 Memory Effect and the Hurst Index

Time series $X_t = X_1, X_2, \ldots, X_N$ is a stochastic process with X_t recorded at the discrete times $t = 0, 1, 2, \ldots, N$. A time series has the memory structure, if the lag period information affects the future changes. Time series displays long memory when the correlation between current and lag observations does not decay to zero quickly over time.

Let X_t be a stationary stochastic process with autocorrelation function $\rho(\tau)$, $\tau = 0, 1, 2, \ldots, m$, where τ denotes the time lag. If $\sum_{\tau=1}^{m} \rho(\tau) = \infty$, X_t is called a long memory process; if $\sum_{\tau=1}^{m} \rho(\tau) < \infty$, it is called a short memory process, and, otherwise, if $\rho(\tau) = 0$, for $\tau \neq 0$, X_t has no memory effect. The classical approach

to measure the stochastic memory process is the autocorrelation function. Now, the Hurst index is widely used as an effective substitute of the autocorrelation function to determine long-range or short-range dependence.

The memory effect can be described by the memory parameter, namely, the Hurst index. The Hurst index measures the smoothness of time series based on the asymptotic behavior of the rescale of the stochastic process. A key property of memory processes is self-similarity, which is denoted by the Hurst index.

Definition 4.3. *If a stochastic process $X = X_t, t = 1, 2, \ldots, N$ is self-similar with Hurst index H for any $a > 0$ and at any time t; then we denote it by $X_{at} \overset{d}{=} a^H X_t$, where the Hurst index describes the self-similarity of the stochastic process, and $\overset{d}{=}$ represents equality of the distribution.*

In the following lemma, some basic properties are given and the corresponding proofs can be obtained in [39, 66].

Lemma 4.3. *Suppose a time series $X_t = X_1, X_2, \ldots, X_N$ is self similar with strictly stationary increment; then this time series has the following properties.*

(1) The expectation of X_t is $E[X_t] = 0$ and, thus, $E[X_t^2] = \sigma^2$ for all $t = 1, 2, \ldots, N$.

(2) The covariance function $\gamma(s,t) = E([X_s - E(X_s)][X_t - E(X_t)]) = E[X_s X_t]$, which has the following result:

$$\gamma(s,t) = \frac{\sigma^2}{2}(|s|^{2H} - |s-t|^{2H} + |t|^{2H}) \tag{4.17}$$

(3) The autocovariance function of X_t is given by $\gamma(\tau), \tau = 1, 2, \ldots, n$ where τ is the lag period:

$$\gamma(\tau) = E[X_t X_{t+\tau}] = \frac{\sigma^2}{2}(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H}) \tag{4.18}$$

(4) If $\tau \neq 0$, then we get the relationship between the autocovariance function

and the Hurst index:

$$\begin{aligned} \gamma(\tau) = 0, \quad H = 0.5 \\ \gamma(\tau) < 0, \quad 0 < H < 0.5 \\ \gamma(\tau) > 0, \quad 0.5 < H < 1 \end{aligned} \tag{4.19}$$

which means $\gamma(\tau) > 0$ *in the case of* $0.5 < H < 1$*; similarly, in the case* $0 < H < 0.5, \gamma(\tau) < 0$*, and in the case* $H = 0.5, \gamma(\tau) = 0$*. According to the autocovariance function, we have that, in the case of* $0 < H < 0.5$*, the times series exhibit short range dependence; in the case of* $H = 0.5$*, the times series has no dependence, which is a perfect random walk; and in the case of* $0.5 < H < 1$*, time series has long-range dependence.*

The Hurst index is usually estimated by the R/S statistic approach. Given a stochastic process $X_t, t = 1, 2, \ldots, N$ of length N, we divide the time interval N into M contiguous subintervals of length n such that $M \times n = N$. For each subinterval, the average value is $X_n = E[X_t] = (1/n) \sum_{t=1}^{n} X_t$.

The running sum of the accumulated deviations from the mean is given as

$$X_{t,n} = \sum_{t=1}^{k} (X_t - E[X_t]), \quad k = 1, 2, \ldots, n. \tag{4.20}$$

The range over the time period n is

$$R(n) = max\{X_{k,n}\} - min\{X_{k,n}\}, \quad k = 1, 2, \ldots, n. \tag{4.21}$$

The standard deviation of $X_t, t = 1, 2, \ldots, N$ is

$$SD(n) = \sqrt{\frac{1}{n-1} \sum_{t=1}^{n} (X_t - E[X_t])^2} \tag{4.22}$$

The rescaled range is $(R/SD)(n) = R(n)/SD(n)$, and the relationship be-

tween the R/SD statistic and n is

$$(\frac{R}{SD})(n) = \frac{R(n)}{SD(n)} = \frac{1}{n}\sum_{i=1}^{n}\frac{R(i)}{SD(i)} \tag{4.23}$$

Thus, we can get the result:

$$E\Big[\frac{R(n)}{SD(n)}\Big] = \alpha n^H, \quad n \to \infty \tag{4.24}$$

where α is a constant and H is the Hurst index.

As a consequence, we can get the Hurst index for the observed time period by linear regression:

$$logE(\frac{R}{SD})(n) = Hlogn + log\alpha, \tag{4.25}$$

4.3.3 Solving Fractional Stochastic Differential Equations

Here, we generalize the classic stochastic differential equation to establish the fractional order stochastic differential equation based on the results presented before and then apply it to the option pricing in the next section.

Definition 4.4. *Assuming that a financial asset price is X, according to the fractional ordinary differential equation, and considering the stochastic process, we can get the FSDE as follows:*

$$d^\alpha X = a(X,t)dt^\alpha + b(X,t)dW(t), \quad \alpha = 2H \tag{4.26}$$

where $a(X,t)$ is the drift parameter, $b(X,t)$ is the diffusion parameter, $dW(t)$ is the Wiener process, $dW(t) = \epsilon\sqrt{dt}, \epsilon \sim N(0,1)$ (normal distribution), and dt and $dW(t)$ are uncorrelated, $dtdt = 0, dtdW(t) = 0, dW(t)dW(t) = dt$.

In a special case, suppose $a(X,t) = \mu X, \quad b(X,t) = \sigma X$, and then we have

the linear stochastic differential equation:

$$d^{\alpha}X = \mu X dt^{\alpha} + \sigma X dW(t), \quad \alpha = 2H \tag{4.27}$$

By using the results of (4.12), we can rewrite (4.26) in the following form of dX with respect to dt^{α}:

$$dX = \frac{a(X,t)}{\Gamma(1+\alpha)}(dt)^{\alpha} + \frac{b(X,t)}{\Gamma(1+\alpha)}dW(t), \quad 0 < \alpha \leq 1, 0 < H \leq 0.5 \tag{4.28}$$

$$dX = \frac{a(X,t)}{\Gamma(1+\alpha)}(dt)^{\alpha} + \frac{b(X,t)}{\Gamma(1+\alpha)}dW(t) + X'(t)dt, \quad 1 < \alpha \leq 2, 0.5 < H \leq 1 \tag{4.29}$$

where $X'(t)$ is the first order derivative of X about time t.

4.4 European Call Option Pricing Based on FSDE

In this section, the corresponding Ito formula and European call option pricing formula are derived based on the fractional order stochastic differential equation.

4.4.1 Ito Lemma Based on FSDE

Lemma 4.4. *Assume that the stock price X follows the fractional order stochastic differential equation as below:*

$$d^{\alpha}X = \mu dt^{\alpha} + \sigma dW(t), \quad \alpha = 2H \tag{4.30}$$

then, the function $f = f(X_t, t)$ is still an Ito stochastic process, and the following expressions hold.

When $0.25 < H \leq 0.5$,

$$df = [\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}]dt + \frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}(dt)^{2H} + \frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}dW(t) \quad (4.31)$$

When $0.5 < H \leq 1$,

$$\begin{aligned} df = &[\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} + X'(t)\frac{\partial f}{\partial X}]dt + \frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}(dt)^{2H} \\ &+ \frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}dW(t) \end{aligned} \quad (4.32)$$

Proof. According to the Ito formula, we notice that

$$\Delta f = \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial X}\Delta X + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(\Delta X)^2 + \frac{\partial^2 f}{\partial X \partial t}(\Delta X \Delta t) + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\Delta t)^2 \quad (4.33)$$

and the discrete form of $d^\alpha X = \mu X(dt)^\alpha + \sigma X dW(t)$ is $\Delta^\alpha X = \mu X(\Delta t)^\alpha + \sigma X \Delta W(t)$.

In this chapter, we only consider the case that $0.25 < H < 1$. There are two reasons for this consideration: first, the Hurst index H is much larger than 0 generally; second, when $0.25 < H < 1, \alpha = 2H > 0.5$, $(\Delta t)^{2\alpha}$ and $(\Delta t)^{\alpha+0.5}$ are infinitesimal. Hence, we do not need to consider the case of $0 < H \leq 0.25$.

(1) In the case of $0.25 < H \leq 0.5$, since $E(\epsilon) = 0, E(\epsilon^2) = 1$, we have

$$\begin{aligned} \Delta X &= \tfrac{\mu X}{\Gamma(1+\alpha)}(\Delta t)^\alpha + \tfrac{\sigma X}{\Gamma(1+\alpha)}\epsilon(\Delta t)^{0.5} \\ \Delta X \Delta t &= \tfrac{\mu X}{\Gamma(1+\alpha)}(\Delta t)^{\alpha+1} + \tfrac{\sigma X}{\Gamma(1+\alpha)}\epsilon(\Delta t)^{1.5} \to 0, \quad 0 < H \leq 0.5 \\ (\Delta X)^2 &= \tfrac{\mu^2 X^2}{\Gamma^2(1+\alpha)}(\Delta t)^{2\alpha} + \tfrac{\sigma^2 X^2}{\Gamma^2(1+\alpha)}\epsilon^2(\Delta t) + \tfrac{2\mu\sigma X^2}{\Gamma^2(1+\alpha)}\epsilon(\Delta t)^{\alpha+0.5} \to \tfrac{\sigma^2 X^2}{\Gamma^2(1+\alpha)}(\Delta t) \end{aligned} \quad (4.34)$$

According to the Ito formula presented above, we can get

$$\begin{aligned}\Delta f &= \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial X}\Delta X + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(\Delta X)^2 + \frac{\partial^2 f}{\partial X \partial t}(\Delta X \Delta t) + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\Delta t)^2 \\ &= \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial X}\Delta X + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(\Delta X)^2 \\ &= \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial X}[\frac{\mu X}{\Gamma(1+\alpha)}(\Delta t)^{\alpha} + \frac{\sigma X}{\Gamma(1+\alpha)}\epsilon(\Delta t)^{0.5}] + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}(\Delta t) \\ &= [\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}]\Delta t + \frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}(\Delta t)^{2H} + \frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}\epsilon(\Delta t)^{0.5}\end{aligned} \tag{4.35}$$

Thus, the differential form is given below:

$$df = [\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}]dt + \frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}(dt)^{2H} + \frac{\sigma X}{\Gamma(1+\alpha)}dW(t) \tag{4.36}$$

(2) In the case of $0.5 < H < 1$,

$$\begin{aligned}\Delta X &= \tfrac{\mu X}{\Gamma(1+\alpha)}(\Delta t)^{\alpha} + \tfrac{\sigma X}{\Gamma(1+\alpha)}\epsilon(\Delta t)^{0.5} + X'(t)\Delta t, \\ \Delta X \Delta t &= \tfrac{\mu X}{\Gamma(1+\alpha)}(\Delta t)^{\alpha+1} + \tfrac{\sigma X}{\Gamma(1+\alpha)}\epsilon(\Delta t)^{1.5} + X'(t)(\Delta t)^2 \to 0, \\ (\Delta X)^2 &= \tfrac{\mu^2 X^2}{\Gamma^2(1+\alpha)}(\Delta t)^{2\alpha} + \tfrac{\sigma^2 X^2}{\Gamma^2(1+\alpha)}\epsilon^2(\Delta t) + (X'(t))^2(\Delta t)^2 + \tfrac{2\mu\sigma X^2}{\Gamma^2(1+\alpha)}\epsilon(\Delta)^{\alpha+0.5} \\ &\quad + \tfrac{2\mu X X'(t)}{\Gamma^2(1+\alpha)}(\Delta t)^{\alpha+1} + \tfrac{2\sigma X X'(t)}{\Gamma^2(1+\alpha)}\epsilon(\Delta t)^{1.5} \to \tfrac{\sigma^2 X^2}{\Gamma^2(1+\alpha)}(\Delta t)\end{aligned} \tag{4.37}$$

$$\begin{aligned}\Delta f &= \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial X}\Delta X + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(\Delta X)^2 \\ &= \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial X}[\frac{\mu X}{\Gamma(1+\alpha)}(\Delta t)^{\alpha} + \frac{\sigma X}{\Gamma(1+\alpha)}\epsilon(\Delta)^{0.5} + X'(t)\Delta t] \\ &+ \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}(\Delta t) \\ &= [\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} + X'(t)\frac{\partial f}{\partial X}]\Delta t + \frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}(\Delta t)^{2H} \\ &+ \frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}\epsilon(\Delta t)^{0.5}\end{aligned} \tag{4.38}$$

Similarly, we obtain the differential form as follows:

$$df = [\frac{\partial f}{\partial t}+\frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}+X'(t)\frac{\partial f}{\partial X}]dt+\frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}(dt)^{2H}+\frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}dW(t) \tag{4.39}$$

□

To price a European option, we first introduce Lemma 4.5, which connects the fractional order stochastic differential equations to the partial differential equations.

Lemma 4.5. *$f(X(t),t)$ is the solution of the partial differential equations:*

$$\frac{\partial f}{\partial t} + rX\frac{\partial f}{\partial X} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} - rf = 0, \quad 0.25 < H \leq 0.5, \tag{4.40}$$

$$\frac{\partial f}{\partial t} + [X'(t) + rX]\frac{\partial f}{\partial X} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} - rf = 0, \quad 0.5 < H < 1, \tag{4.41}$$

$$f(X(T),T) = f(X(T)). \tag{4.42}$$

Proof. First, make portfolios $\Pi = \Delta X - f$ and $d\Pi = \Delta dX - df$.

(1) In the case of $0.25 < H \leq 0.5$,

$$\begin{aligned} d\Pi &= \Delta dX - df \\ &= \Delta[\frac{\mu X}{\Gamma(1+\alpha)}(dt)^{\alpha} + \frac{\sigma X}{\Gamma(1+\alpha)}dW(t)] - [(\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial f}{\partial X})dt \\ &+ (\frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X})(dt)^{2H} + \frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}dW(t)] \\ &= -[\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial f}{\partial X}]dt + \frac{\mu X}{\Gamma(1+\alpha)}(\Delta - \frac{\partial f}{\partial X})(dt)^{2H} \\ &+ \frac{\sigma X}{\Gamma(1+\alpha)}(\Delta - \frac{\partial f}{\partial X})dW(t) \end{aligned} \tag{4.43}$$

When $\Delta = \frac{\partial f}{\partial X}$, we can get the riskless asset portfolio

$$d\Pi = \Delta dX - df = -[\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2}]dt \tag{4.44}$$

And because the portfolio Π is riskless, according to the Bellman Equation,we have $d\Pi = r\Pi dt$, where r is the riskless rate. Thus, we get the equation $d\Pi = r\Pi dt = -[(\partial f/\partial t) + (\sigma^2 X^2/2\Gamma^2(1+\alpha))(\partial^2 f/\partial X^2)]dt$. Consequently, we obtain the first partial differential equation

$$\frac{\partial f}{\partial t} + rX\frac{\partial f}{\partial X} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} - rf = 0 \tag{4.45}$$

(2) In the case of $0.5 < H < 1$,

$$\begin{aligned}
d\Pi &= \Delta dX - df \\
&= \Delta[\frac{\mu X}{\Gamma(1+\alpha)}(dt)^\alpha + \frac{\sigma X}{\Gamma(1+\alpha)}dW(t) + X'(t)dt] \\
&- [(\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} + X'(t)\frac{\partial f}{\partial X})dt \\
&+ (\frac{\mu X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X})(dt)^{2H} + \frac{\sigma X}{\Gamma(1+\alpha)}\frac{\partial f}{\partial X}dW(t)] \\
&= -[\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} + X'(t)\frac{\partial f}{\partial X}]dt + \frac{\mu X}{\Gamma(1+\alpha)}(\Delta - \frac{\partial f}{\partial X})(dt)^{2H} \\
&+ \frac{\sigma X}{\Gamma(1+\alpha)}(\Delta - \frac{\partial f}{\partial X})dW(t)
\end{aligned} \tag{4.46}$$

When $\Delta = \partial f/\partial X$, we can also get the riskless asset portfolio

$$d\Pi = \Delta dX - df = -[\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} + X'(t)\frac{\partial f}{\partial X}]dt \tag{4.47}$$

And again because Π is riskless, we can get the equation

$$d\Pi = r\Pi dt = -[\frac{\partial f}{\partial t} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} + X'(t)\frac{\partial f}{\partial X}]dt \tag{4.48}$$

Similarly, the second partial differential equation can be obtained as below:

$$\frac{\partial f}{\partial t} + [X'(t) + rX]\frac{\partial f}{\partial X} + \frac{\sigma^2 X^2}{2\Gamma^2(1+\alpha)}\frac{\partial^2 f}{\partial X^2} - rf = 0 \tag{4.49}$$

□

4.4.2 European Call Option Based on FSDE

Before we proceed to price the European call option, we make the assumptions as below:

(1) r is the riskless rate and is a constant;

(2) the exchange of the stock is continuous and the stock can be divided;

(3) the tax of the stock exchange is free;

(4) the bonus of the stock cannot be paid within the duration of derivatives;

(5) no arbitrage exists in the market;

(6) the price of stock follows a fractional order stochastic differential equation

$$\frac{d^\alpha X}{X} = r(dt)^\alpha + \sigma dW(t), \quad \alpha = 2H;$$

(7) the strike price is K;

(8) the maturity is T,

where X is the price of the stock and r is the riskless interest rate; σ is the volatility of the price of stock; H is the Hurst parameter of the stock.

In the following work, we will derive the fractional option pricing formula based on the risk-neutral assumption. If the price of underlying asset is subject to the geometric Brownian motion and the return μ is equal to the riskless interest rate r (i.e., $\mu = r$), we have

$$\frac{d^\alpha X}{X} = r(dt)^\alpha + \sigma dW(t), \quad \alpha = 2H \tag{4.50}$$

(1) In the case $0.25 < H \leq 0.5$, according to Ito Lemma 4.4, we can get the

price of the stock as

$$d(lnX) = \frac{r}{\Gamma(1+\alpha)}(dt)^{2H} - \frac{\sigma^2}{2\Gamma^2(1+\alpha)}dt + \frac{\sigma}{\Gamma(1+\alpha)}dW(t), \quad 0.25 < H \leq 0.5 \tag{4.51}$$

Integrate (4.54) and use Lemma 4.2; then,we can get the solution of X,

$$X_T = Xexp\left(\frac{r}{\Gamma(1+\alpha)}(T^{2H}-t^{2H}) - \frac{\sigma^2}{2\Gamma^2(1+\alpha)}(T-t) + \frac{\sigma}{\Gamma(1+\alpha)}(W(T)-W(t))\right)$$

Therefore, the European call option pricing formula follows:

$$X_0 = Xe^{(r/\Gamma(1+\alpha))(T^{2H}-t^{2H})-r(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2)$$

where

$$d_1 = \big(\Gamma(1+\alpha)ln(X/K) + r(T^{2H}-t^{2H}) + \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H-t^H})^{-1}$$
$$d_2 = \big(\Gamma(1+\alpha)ln(X/K) + r(T^{2H}-t^{2H}) - \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H-t^H})^{-1}$$

Proof. The price of the European call option is given by $X_0 = e^{-r(T-t)}E[max(X_T - K), 0]$, where $E(\cdot)$ is the expectation of the option price based on risk-neutral, and the price of the asset X_T obeys the log normal distribution:

$$lnX_T - lnX \sim N\left(\frac{r}{\Gamma(1+\alpha)}(T^{2H}-t^{2H}) - \frac{\sigma^2}{2\Gamma^2(1+\alpha)}(T-t), \frac{\sigma^2}{\Gamma^2(1+\alpha)}(T-t)\right)$$

Let $Q = (lnX_T - m)/s$; obviously, $Q \sim N(0,1)$ and the probability density function $h(Q) = (1/\sqrt{2\pi})e^{-Q^2/2}$, where $m = E(lnX_T) = lnX + (r/\Gamma(1+\alpha))(T^{2H}-t^{2H}) - (\sigma^2/2\Gamma^2(1+\alpha))(T-t)$, and $s = \sigma\sqrt{T-t}/\Gamma(1+\alpha)$. Hence,

$$
\begin{aligned}
E[max(X_T - K, 0)] &= \int_{-\infty}^{+\infty} max(X_T - K, 0) h(X_T) dX(T) \\
&= \int_{K}^{+\infty} (X_T - K) h(X_T) dX_T + \int_{-\infty}^{K} 0 h(X_T) dX(T) \\
&= \int_{lnK}^{+\infty} (e^{X_T} - K) h(lnX_T) d(lnX_T) \\
&= \int_{lnK-m}^{+\infty} (e^{X_T} - K) h(\frac{lnX_T - m}{s}) d(\frac{lnX_T - m}{s}) \\
&= \int_{lnK-m}^{+\infty} (e^{X_T} - K) h(lnX_T) d(lnX_T) \\
&= \int_{lnK-m}^{+\infty} (e^{sQ+m} - K) h(Q) dQ \\
&= \int_{lnK-m}^{+\infty} e^{s^2/2+m} \frac{1}{\sqrt{2\pi}} e^{-(Q-s)^2/2} dQ - KN(\frac{m - lnK}{s}) \\
&= Xe^{(r/\Gamma(1+\alpha))(T^{2H} - t^{2H})} N(d_1) - KN(d_2)
\end{aligned}
\tag{4.52}
$$

where

$$
\begin{aligned}
d_1 &= \big(\Gamma(1+\alpha) ln(X/K) + r(T^{2H} - t^{2H}) + \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \\
d_2 &= \big(\Gamma(1+\alpha) ln(X/K) + r(T^{2H} - t^{2H}) - \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1}
\end{aligned}
$$

So we get the European option pricing formula as follows:

$$
\begin{aligned}
X_0 &= e^{-r(T-t)} E[max(X_T - K, 0)] \\
&= e^{-r(T-t)} [e^{(r/\Gamma(1+\alpha))(T^{2H} - t^{2H})} N(d_1) - KN(d_2)] \\
&= Xe^{(r/\Gamma(1+\alpha))(T^{2H} - t^{2H}) - r(T-t)} N(d_1) - Ke^{-r(T-t)} N(d_2)
\end{aligned}
\tag{4.53}
$$

□

(2) In the case of $0.5 < H < 1$, in a similar way, according to the Ito Lemma

4.4, the price of the stock is

$$d(lnX) = \frac{r}{\Gamma(1+\alpha)}(dt)^{2H} + (\frac{X'}{X} - \frac{\sigma^2}{2\Gamma^2(1+\alpha)})dt + \frac{\sigma}{\Gamma(1+\alpha)}dW(t) \quad (4.54)$$

Notice that $X'/Xdt = d(lnX) = lnX_{t+1} - lnX_t = m(t)$, $m(t)$ represents the daily logarithm returns of stock X, and $m(t) = \mu(t)dt$, $\mu(t)$ is the returns of one year; thus, $X'/X = \mu(t)$ and (52) can be written as follows:

$$d(lnX) = \frac{r}{\Gamma(1+\alpha)}(dt)^{2H} + (\mu(t) - \frac{\sigma^2}{2\Gamma^2(1+\alpha)})dt + \frac{\sigma}{\Gamma(1+\alpha)}dW(t) \quad (4.55)$$

By integrating (4.55) and employing Lemma 4.2, we get the solution of X:

$$\begin{aligned} X_T =& Xexp\bigg(\frac{r}{\Gamma(1+\alpha)}(T^{2H} - t^{2H}) + \int_t^T \mu(s)ds - \frac{\sigma^2}{2\Gamma^2(1+\alpha)}(T-t) \\ &+ \frac{\sigma}{\Gamma(1+\alpha)}(W(T) - W(t))\bigg) \end{aligned}$$

Consequently, the European call option pricing formula is obtained:

$$X_0 = Xe^{(r/\Gamma(1+\alpha))(T^{2H}-t^{2H})+\int_t^T \mu(s)ds - r(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2)$$

where

$$\begin{aligned} d_1 =& \big(\Gamma(1+\alpha)ln(X/K) + \int_t^T \mu(s)ds + r(T^{2H} - t^{2H}) \\ &+ \frac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \end{aligned}$$

$$\begin{aligned} d_2 =& \big(\Gamma(1+\alpha)ln(X/K) + \int_t^T \mu(s)ds + r(T^{2H} - t^{2H}) \\ &- \frac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \end{aligned}$$

Proof. Let $Q = (lnX_T - m)/s$, and it is obvious that $Q \sim N(0,1)$, and so the

probability density function $h(Q) = (1/\sqrt{2\pi})e^{-Q^2/2}$, where

$$m = \hat{E}(lnX_T) = lnX + \int_t^T \mu(s)ds + (r/\Gamma(1+\alpha))(T^{2H} - t^{2H}) - (\sigma^2/2\Gamma^2(1+\alpha))(T-t),$$

and $s = \sigma\sqrt{T-t}/\Gamma(1+\alpha)$.

Hence,

$$\begin{aligned} E[max(X_T - K, 0)] &= \int_{-\infty}^{+\infty} max(X_T - K, 0)h(X_T)dX(T) \\ &= \int_K^{+\infty} (X_T - K)h(X_T)dX_T \\ &= \int_{(lnK-m)/s}^{+\infty} (e^{sQ+m} - K)h(Q)dQ \\ &= \int_{(lnK-m)/s}^{+\infty} e^{s^2/2+m}\frac{1}{\sqrt{2\pi}}e^{-(Q-s)^2/2}dQ - KN(\frac{m - lnK}{s}) \\ &= Xe^{\int_t^T \mu(s)ds + (r/\Gamma(1+\alpha))(T^{2H} - t^{2H})}N(d_1) - KN(d_2) \end{aligned} \tag{4.56}$$

where

$$\begin{aligned} d_1 = &\big(\Gamma(1+\alpha)ln(X/K) + \int_t^T \mu(s)ds + r(T^{2H} - t^{2H}) \\ &+ \frac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \end{aligned}$$

$$\begin{aligned} d_2 = &\big(\Gamma(1+\alpha)ln(X/K) + \int_t^T \mu(s)ds + r(T^{2H} - t^{2H}) \\ &- \frac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \end{aligned}$$

Finally, the European option pricing formula is given as follows:

$$\begin{aligned} X_0 &= e^{-r(T-t)}E[max(X_T - K, 0)] \\ &= e^{-r(T-t)}[e^{\int_t^T \mu(s)ds + (r/\Gamma(1+\alpha))(T^{2H} - t^{2H})}N(d_1) - KN(d_2)] \\ &= Xe^{(r/\Gamma(1+\alpha))(T^{2H} - t^{2H}) + \int_t^T \mu(s)ds - r(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2) \end{aligned} \tag{4.57}$$

□

From the result we derived, the option price formula contains the mean value function $\int_t^T \mu(s)ds$ of the logarithmic returns of stock price, which is the effect of trend memory. Therefore, we proved that trend memory exists in the financial systems.

Now, we give the European call option pricing formula under the risk-neutral measure. Let the mean returns of stock be equal to the riskless rate r; by taking the expectation of the returns in the case $0.5 < H < 1$, we have $E[\mu(t)] = \mu = r$, where r is the riskless returns. Then, we simplify the mean value function $\int_t^T \mu(s)ds = \int_t^T \mu ds = \mu(T-t) = r(T-t)$ and have $\mu - r = 0$; thus, we get the option pricing formula

$$X_0 = Xe^{(r/\Gamma(1+\alpha))(T^{2H}-t^{2H})-r(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2)$$

where

$$\begin{aligned} d_1 &= \big(\Gamma(1+\alpha)ln(X/K) + r(T^{2H}-t^{2H}) + (r + \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)})(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \\ d_2 &= \big(\Gamma(1+\alpha)ln(X/K) + r(T^{2H}-t^{2H}) - (r - \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)})(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1} \end{aligned} \tag{4.58}$$

4.5 Numerical Analysis

To explain the memory effects in financial market, we make some comparisons in this section among our proposed European option pricing model and its underlying stock price equation and the well-known classic models, such as the Black-Scholes model (Black and Scholes (1973) [47]) and the Black-Scholes model under fractional Brownian motion (Necula (2002) [22], Hu and Øksendal (2003) [159]).

(1) The classical Black-Scholes model [47]:

$$dS = rSdt + \sigma SdW(t)$$

The European option pricing formula is

$$c = SN(d_{11}) - Ke^{-r(T-t)}N(d_{12})$$

where

$$\begin{aligned} d_{11} &= \frac{ln(S/K)+(r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \\ d_{12} &= \frac{ln(S/K)+(r-\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \end{aligned} \tag{4.59}$$

The classical Black-Scholes model was established under the assumption that the price process is a Markov process and that the price process is independent and has no memory effect; however, the memory effects exist in price process.

(2) The SDE with fractional Brownian motion [22, 159]:

$$dS = rSdt + \sigma SdW_H(t)$$

The European option pricing formula is

$$c = SN(d_{21}) - Ke^{-r(T-t)}N(d_{22})$$

where

$$\begin{aligned} d_{21} &= \frac{ln(S/K)+r(T-t)+(\sigma^2/2)(T^{2H}-t^{2H})}{\sigma(T^H-t^H)} \\ d_{22} &= \frac{ln(S/K)+r(T-t)-(\sigma^2/2)(T^{2H}-t^{2H})}{\sigma(T^H-t^H)} \end{aligned} \tag{4.60}$$

The fractional Brownian motion model has improved the Black-Scholes mod-

el by considering the memory effect of the asset price but only considered the memory effect of the noise.

(3) The fractional order SDE (FSDE model):

$$d^{\alpha}S = rS(dt)^{\alpha} + \sigma S dW(t), \quad \alpha = 2H$$

In the case of $0.25 < H < 0.5$, the European call option pricing formula is

$$c = Se^{(r/\Gamma(1+\alpha))(T-t)^{2H} - r(T-t)}N(d_{31}) - Ke^{-r(T-t)}N(d_{32})$$

where

$$d_{31} = \big(\Gamma(1+\alpha)ln(S/K) + r(T^{2H} - t^{2H}) + \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1}$$
$$d_{32} = \big(\Gamma(1+\alpha)ln(S/K) + r(T^{2H} - t^{2H}) - \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)}(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1}$$

When $H = 0.5$, the option formula is reduced to the classic option formula.

In the case of $0.5 < H < 1$, the European call option pricing formula is

$$c = Se^{(r/\Gamma(1+\alpha))(T^{2H} - t^{2H})}N(d_{41}) - Ke^{-r(T-t)}N(d_{42})$$

where

$$d_{41} = \big(\Gamma(1+\alpha)ln\tfrac{S}{K} + r(T^{2H} - t^{2H}) + (r + \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)})(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1}$$
$$d_{42} = \big(\Gamma(1+\alpha)ln\tfrac{S}{K} + r(T^{2H} - t^{2H}) - (r - \tfrac{\sigma^2/2}{2\Gamma(1+\alpha)})(T-t)\big)(\sigma\sqrt{T^H - t^H})^{-1}$$

In this work, our model takes a new memory effect into consideration, which is called the trend memory effect of the asset price.

The underlying asset price equations to be used to make comparison are given as follows:

(1) the SDE model: $dS = rSdt + \sigma SdW(t)$,

The stock price equation is $S_T = Sexp[(r - \sigma^2/2)(T-t) + \sigma\Delta W(t)]$;

(2) the SDE model with fractional Brownian motion: $dS = rSdt + \sigma SdW_H(t)$,

The stock price equation is $S_T = Sexp[r(T-t) - \sigma^2(T-t)^{2H} + \sigma\Delta W_H(t)]$;

(3) the fractional order SDE model: $d^\alpha S = rS(dt)^\alpha + \sigma SdW(t), \quad \alpha = 2H$,

when $0.25 < H \leq 0.5$, the stock price equation is

$$S_t = S_{t-1}exp\Big[\frac{r}{\Gamma(1+\alpha)}\Delta t^{2H} - \frac{\sigma^2}{2\Gamma^2(1+\alpha)}\Delta t + \frac{\sigma}{\Gamma(1+\alpha)}\epsilon\Delta W_t\Big] \tag{4.61}$$

However, when $0.5 < H < 1$, the stock price equation is

$$S_t = S_{t-1}exp\Big[\frac{r}{\Gamma(1+\alpha)}\Delta t^{2H} + \big(r - \frac{\sigma^2}{2\Gamma^2(1+\alpha)}\big)\Delta t + \frac{\sigma}{\Gamma(1+\alpha)}\epsilon\Delta W_t\Big] \tag{4.62}$$

To illustrate the validity of the proposed FSDE model, we simulate the three types of stochastic differential equations presented above by using the Monte Carlo simulation method and then make comparison of these three different models. The data used in the empirical analysis is the daily closing price index series of CSI300 index of China. The time range is from January 4, 2012, to October 27, 2012, with the initial value $S_0 = 2299$ (the price index on January 4, 2012) and the final value $S_T = 2445$. We choose the one-year bonds interest rate $r = 2.65\%$ in China as the riskless rate and the mean yield $\mu = 0.0266$. The Hurst parameter is $H = 0.6614$, which is estimated by R/S analysis approach. Given a 95 percent confidence interval, the simulation results are shown in Tables 4.1, 4.2, and 4.3.

Table 4.1: Simulation results of SDE model

Simulation times	Real value	Mean value	Standard deviation	Error rate	Confidence interval
10^2	2445	2351.9	198.6958	3.81%	[2313.0, 2390.8]
10^3	2445	2358.9	195.7074	3.53%	[2346.5, 2370.8]
10^4	2445	2357.0	200.5667	3.60%	[2353.1, 2360.9]
10^5	2445	2360.0	202.2802	3.48%	[2358.7, 2361.2]

From the results in Tables $4.1 - 4.3$, by using Monte Carlo simulation, we

Table 4.2: Simulation results of SDE with FBM

Simulation times	Real value	Mean value	Standard deviation	Error rate	Confidence interval
10^2	2445	2358.2	77.4242	3.55%	[2343.0, 2373.3]
10^3	2445	2360.8	76.3254	3.44%	[2356.1, 2365.6]
10^4	2445	2360.0	78.0423	3.48%	[2358.6, 2361.6]
10^5	2445	2361.2	78.5882	3.43%	[2360.8, 2361.7]

Table 4.3: Simulation results of FSDE

Simulation times	Real value	Mean value	Standard deviation	Error rate	Confidence interval
10^2	2445	2362.7	19.0566	3.37%	[2359.1, 2366.5]
10^3	2445	2368.4	21.1279	3.13%	[2367.2, 2369.8]
10^4	2445	2367.0	19.7396	3.19%	[2366.7, 2367.5]
10^5	2445	2369.5	20.4312	3.09%	[2369.5, 2369.7]

conclude that the error of our proposed FSDE model is smaller than that in the conventional SDE model and the SDE with FBM model. If we conduct the simulating process 100000 times, which is large enough for the error analysis, we obtain that the error rate of the SDE model is 3.48%, while the rate is 3.43% and 3.09% respectively for the SDE driven by fractional Brownian motion and the FSDE model.

In addition, the standard derivation of simulation in the FSDE model is also much lower than the SDE model and the SDE with FBM model, and the confidence interval is smaller than the two classic models. In the same way, when we simulate 100000 times, the standard deviation of SDE model is 202.2802, the SDE driven by fractional Brownian motion model is 78.5882, and FSDE model is 20.4312. Thus, we get the conclusion that the FSDE has about 10 times lower standard derivation than the SDE model and about 2.5 times lower standard derivation than the SDE driven by fractional Brownian motion.

4.6 Concluding Remarks

Because the fractional order ordinary differential equations can capture the memory effect in the financial system,we established the fractional order stochastic

differential equation by adding the stochastic process into the fractional ordinary differential equation. Based on this stochastic differential equation with fractional order, we apply the fractional order stochastic differential equation to the financial market. We constructed the stock price $d^{\alpha}S = \mu(S,t)dt^{\alpha} + \sigma(S,t)dW(t)$, where $\alpha = 2H$ and H is Hurst index, and derived the stock price process in the cases of $0.25 < H \leq 0.5$ and $0.5 < H < 1$, respectively, and the European call option pricing formula under the fractional order stochastic differential equation. From the European option pricing formula, we find that the trend memory in stock price process when Hurst index is between 0.5 and 1.

In addition, we made some comparisons in terms of the pricing option formula and its underlying stock price process among our proposed approach and the other two classical models. We find that the new approach leads to a better result than the classic approach and the fractional Brownian motion approach when we simulate the stock prices by Monte Carlo simulation.

CHAPTER 5

Estimation of Parameters in Mean-reverting Stochastic Systems

5.1 General Overview

Stochastic differential equation (SDE) is a very important mathematical tool to describe complex systems in which noise plays an important role. SDE models have been widely used to study the dynamic properties of various nonlinear systems in biology, engineering, finance and economics, as well as physical sciences. Since a SDE can generate unlimited numbers of trajectories, it is difficult to estimate model parameters based on experimental observations which may represent only one trajectory of the stochastic model. Although substantial research efforts have been made to develop effective methods, it is still a challenge to infer unknown parameters in SDE models from observations that may have large variations. Using an interest rate model as a test problem, in this work, we use the Bayesian inference and Markov Chain Monte Carlo method to estimate unknown parameters in SDE models.

A number of methods have been used to estimate the parameters in the single-factor continuous time models, including the generalized moment method [52] and Gaussian estimation methods [84]. However, our recent research work suggests

that the accuracy of the estimates generated from these two methods is low, in particular, when the stepsize of observation time points is not small [140]. Thus in this chapter we will concentrate on proposing a method that will give more accurate estimates of the model parameters, leading to more accurate simulation of the stochastic process. By utilizing the Bayesian inference and the MCMC method, we develop a numerical algorithm to estimate the unknown parameters in stochastic interest rate models. The remaining part of this chapter is organized as follows. Section 5.2 gives a class of stochastic models for term structure of interest rates and numerical algorithms for simulating these stochastic models. Section 5.3 discusses the Bayesian inference and the MCMC method, and proposes a new numerical algorithm for estimating parameters in the mean-reverting stochastic systems. Section 5.4 presents the numerical results for estimating the parameters in the stochastic models for the term structure of interest rates.

5.2 Stochastic Model and Direct Simulation Methods

We first introduce the general form of SDEs for interest rates, namely

$$dX = a(t, X)dt + b(t, X)dW(t), \tag{5.1}$$

where $a(t, X)$ is the drift term, $b(t, X)$ the diffusion term and $W(t)$ the Wiener process. The increment of Wiener process $\Delta W_n = W(t_{n+1}) - W(t_n)$ follows the Gaussian distribution $N(0, t_{n+1} - t_n)$.

Now we proceed to consider numerical methods for simulating the SDE. The widely used method in computational finance is the Euler-Maruyama method

whose strong convergence order is just 0.5, given by

$$X_{n+1} = X_n + ha(t_n, X_n) + b(t_n, X_n)\Delta W_n, \tag{5.2}$$

where X_n is the numerical solution at time point t_n, and $h = t_{n+1} - t_n$. Although this method is easy to implement, its stability property is not good enough to simulate SDEs with a relatively large diffusion term. In order to obtain stable simulations, a very small stepsize is required, which may lead to large computing time. To improve the stability property, the semi-implicit and fully implicit Euler method can be used to reduce the computing time [107, 139]. For example, the semi-implicit Euler method is given by

$$X_{n+1} = X_n + ha(t_{n+1}, X_{n+1}) + b(t_n, X_n)\Delta W_n. \tag{5.3}$$

In computational finance the Euler-Maruyama method with strong convergence order 0.5 is widely used because of its easy computer-based implementation. However, the drawback of this method is that its stability property is not satisfactory when simulating SDEs with relatively large diffusion term. One possible improvement is to use very small stepsize to get the stable simulations, but the computing time increases incredibly. Here we consider the implicit Milstein method to get reliable numerical results with good accuracy and stability property.

The Milstein scheme uses a higher order stochastic Taylor expansion and thus has a strong convergence order one, given by

$$\begin{aligned} X_{n+1} &= X_n + a(t_n, X_n)h + b(t_n, X_n)\Delta W_n \\ &\quad + \frac{1}{2}b(t_n, y_n)b'(t_n, y_n)((\Delta W_n)^2 - h). \end{aligned} \tag{5.4}$$

In order to improve the stability of the Milstein method, the semi-implicit

and fully implicit Milstein methods were proposed. The method, in which only the drift term is implicit, known as semi-implicit Milstein method as given below

$$\begin{aligned} X_{n+1} &= X_n + a(t_{n+1}, X_{n+1})h + b(t_n, X_n)\Delta W_n \\ &\quad + \frac{1}{2}b(t_n, y_n)b'(t_n, y_n)((\Delta W_n)^2 - h); \end{aligned} \tag{5.5}$$

while the method where both the drift term and the diffusion term are implicit is called fully implicit Milstein method [139], given by

$$\begin{aligned} X_{n+1} &= X_n + a(t_{n+1}, X_{n+1})h + b(t_{n+1}, X_{n+1})\Delta W_n \\ &\quad + \frac{1}{2}b(t_{n+1}, y_{n+1})b'(t_{n+1}, y_{n+1})((\Delta W_n)^2 + h). \end{aligned}$$

In this chapter we use the CIR (Cox, Ingersoll and Ross) model of the term structure of interest rate as the test system to examine the accuracy of the inference methods [11]. The CIR model was introduced to model the short term interest rate [28, 52], which is a linear mean-reversing stochastic differential equation [75]. The CIR model states that the short interest rate follows a square root diffusion process, which has the following continuous-time representation:

$$dX = \alpha(\beta - X)dt + \sigma\sqrt{X}dW(t), \tag{5.6}$$

where $\alpha, \beta, \sigma > 0$, α is the speed of adjustment (or mean reversion), β represents the long term value of the randomly moving interest rate, and σ is a constant volatility. This model implicates that both the drift and volatility change with the level of the short rate.

In this work we will use the Euler-Maruyama method (5.2) to generate samples of the interest rates. In fact, due to the linear feature of the drift term in the interest rate model, the semi-implicit method can be written in explicit form, and can also be used in the Bayesian inference method. For the benchmark model,

the Euler-Maruyama Scheme is

$$X_{n+1} = X_n + \alpha(\beta - X_n)h + \sigma\sqrt{X_n}\Delta W_n; \tag{5.7}$$

and the semi-implicit Euler scheme is given by

$$X_{n+1} = \frac{1}{1+\alpha h}\left(X_n + \alpha\beta h + \sigma\sqrt{X_n}\Delta W_n\right). \tag{5.8}$$

5.3 Parameter Estimation for the Term Structure Models

5.3.1 Bayesian Inference Approach

In this section, we establish a numerical algorithm for estimating parameters in stochastic models based on the Bayesian inference and MCMC method. In contrast with the classical approach in which the unknown parameters in a model have fixed quantity, the unknown parameters of the underlying model in the Bayesian paradigm are treated as a random variable with some prior beliefs. The heart of the Bayesian approach is the Bayes theorem which allows us to compute the conditional probability density function of the model parameters θ, assuming that the model parameters are continuous random variables, given the entire data set y

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}. \tag{5.9}$$

Since the probability $p(y)$ is independent of the model parameters, to maximize the joint probability density function, only the product $p(y|\theta)p(\theta)$ should be considered. Thus the posterior distribution $f(\theta|y)$ can be interpreted as our prior beliefs of the parameters $f(\theta)$ updated by the current information from the data. Because we have little prior knowledge of θ, we may simply use a "non-

informative" or "flat" prior.

In this work we use the Bayesian inference method derived by Joshi and Wilson [20] to infer the parameters in SDE models. It is assumed that the diffusion process $\{X_t\}$ is observed at time points $t_0, t_1, \ldots, t_n$ and the observation vector is $Y = \{y_0, y_1, \ldots, y_n\}$. Since the closed form of the transition densities of the diffusion processes is usually not available, the transition densities can be approximated by the densities of a numerical scheme such as the widely used Euler-Maruyama method (5.2). However, the observation time stepsize $\Delta t = t_{i+1} - t_i$ normally is quite large. To obtain more accurate approximation of the transition densities, a number of latent variables are introduced between every pair of consecutive observations,

$$t_i = \tau_{0,i} < \tau_{1,i} < \ldots < \tau_{M,i} = t_{i+1}.$$

The stepsize of the latent variables $\delta_\tau = \tau_{j+1,i} - \tau_{j,i}$ is small enough to ensure the accuracy and stability property of the Euler-Maruyama method. Then the transition density of the Euler scheme is

$$P_{Euler}\{X_{j+1,i}|X_{j,i}, \Theta\} = N(\mu_{Euler}, \sigma^2_{Euler}), \tag{5.10}$$

where $j = 0, 1, \ldots, M-1$ and

$$\mu_{Euler} = X_{j,i} + f(t_{j,i}, X_{j,i}, \Theta)\delta_\tau \tag{5.11}$$

$$\sigma_{Euler} = g(t_{j,i}, X_{j,i}, \Theta)\sqrt{\delta_\tau}. \tag{5.12}$$

Thus we have an inference problem with unknown parameter Θ using the latent variables $X = \{X_{j,i}\}$ for $i = 0, 1, \ldots, n$, $j = 1, \ldots, M-1$, and the observation data Y.

5.3.2 Markov Chain Monte Carlo Method

To generate the samples of the unknown parameters Θ_j, a grid-sampling method was used by dividing the potential area by a regular grid [20]. Although this is an effective approach to infer mathematical models with a small number of unknown parameters, it is difficult to use it for dealing with models with a larger number of unknown parameters. In this work we use the MCMC method to search the optimal model parameters. Since the closed-form posterior distribution for a complex model often cannot be obtained analytically, the MCMC method has been widely used to achieve the posterior distribution by simulation. There are a number of efficient MCMC algorithms, including the Metropolis algorithm, the Metropolis-Hastings (MH) algorithm and the Gibbs sampler method. In this work we use the MH algorithm to maximize the posterior distribution. The MH algorithm allows us to avoid the direct simulation from $\pi(\theta|y)$ by making use of a proposal distribution and computing the acceptance probability for a candidate sample. There are a number of important issues that are related to the implementation of the MCMC. For example, the selected initial estimate has influence on the generated sequence, and in particular, has strong influence on the initial sequence of simulations. Thus an important technique is the burn-in, which is designed to reduce the influence of initial iteration on the generated Markov Chain by discarding the first iteration sequences. Generally we discard the first half of simulated sequences and keep the remaining half of sequence to obtain the target distribution. Certainly this technique is convenient but is not the most efficient one because about a half of computing efforts are discarded. Although more specific methods have been designed to analyze the simulation output according to the dependence of simulation on the starting values [21], we typically go with the simple burn-in approach and accept the increased Monte Carlo error involved in discarding half of the simulations. To design a strategy

to complete the computation, we normally monitor the convergence of all the parameters and other quantities of interest separately. Our usual approach is, for each parameter, to calculate the variance of simulations from each chain (after the first half of the chain was discarded using the burn-in technique). Assuming we have J chains from different initial estimates and the length of each chain is G, let θ_{ij} be the j-th estimate in the sequence for parameter θ, the variance inside the chain is

$$W = \frac{1}{J(G-1)} \sum_{j=1}^{J} \sum_{g=1}^{G} (\theta_{ij} - \overline{\theta}_j)^2$$

And the variance between different chains is

$$B = \frac{1}{(J-1)G} \sum_{j=1}^{J} \left(\sum_{g=1}^{G} \theta_{gj} - \frac{1}{J} \sum_{j=1}^{J} \sum_{g=1}^{G} \theta_{gj} \right)^2$$

Based on these values we can calculate the value of R as

$$R = \sqrt{\frac{1}{G} \left(G - 1 + \frac{B}{W} \right)}$$

The value of R is always greater or equal to 1. When the variance inside the chain approaches the variance between the chains, the value of R approaches 1. We can accept that the chain is convergence when $R < 1.2$ [69]. Another important technique is thinning by rejecting certain part of the chain. If Θ_t is the current candidate of the model parameter and Θ^* is the newly generated one, let

$$\alpha = 1 \wedge \frac{P(\Theta^*|y)}{P(\Theta_t|y)}. \tag{5.13}$$

Generate a sample $r \sim U(0,1)$, and set $\Theta_{t+1} = \Theta^*$ if $r < \alpha$. Otherwise set $\Theta_{t+1} = \Theta_t$.

5.3.3 A New Parameter Estimation Algorithm

In the proposed Gaussian Modified Bridge Approximation with Importance sampling (GaMBA-I), the computing process is given below

Algorithm 5.1.

Step 1. Generate a sample of the unknown parameters Θ using the MCMC method or other methods.

Step 2. Sample the solution at the latent points X.

Step 3. Evaluate probability $P(Y, X|\Theta)$.

Step 4. Evaluate probability $P(X|Y, \Theta)$.

Step 5. Calculate probability

$$P(\Theta|Y) \propto \frac{P(Y, X|\Theta)P(\Theta)}{P(X|Y, \Theta)}.$$

When the importance sampling technique is used, the above probability is determined by a number of samples rather than a single sample as indicated above.

Step 6. Accept or reject the parameter sample Θ using the MCMC method or other methods.

The major step in this Bayesian inference method is the evaluation of the probabilities of the generated samples for the latent variables. The probability $P(Y, X|\Theta)$ is

$$\begin{aligned} P(Y, X|\Theta) \;\propto\; & \prod_{i=1}^{n} P(y_i|X_{M-1,i-1}, \Theta) \prod_{i=1}^{n} P(X_{1,i-1}|y_{i-1}, \Theta) \\ & \prod_{i=1}^{n} \prod_{j=1}^{M-1} P(X_{j,i-1}|X_{j-1,i-1}, \Theta) \end{aligned}$$

Here we assume that the probability for the initial observation y_0 is a constant. Each probability in the above expression can be approximated by the transition

density of the Euler method, given by

$$\begin{aligned} P(Y,X|\Theta) \;\approx\; & \prod_{i=1}^{n} P_{Euler}(y_i|X_{M-1,i-1},\Theta)\prod_{i=1}^{n} P_{Euler}(X_{1,i-1}|y_{i-1},\Theta) \\ & \prod_{i=1}^{n}\prod_{j=1}^{M-1} P_{Euler}(X_{j,i-1}|X_{j-1,i-1},\Theta) \end{aligned}$$

where P_{Euler} is the Euler density in Equation (5.10).

For the probability $P(X|Y,\Theta)$, we need to factorise it into

$$\begin{aligned} P(X|Y,\Theta) &= \prod_{i=0}^{n-1} P(X^{(i)}|y_i,y_{i+1},\Theta) \\ &= \prod_{i=0}^{n-1} P(X_{1,i},X_{2,i},\ldots,X_{M-1,i}|y_i,y_{i+1},\Theta) \\ &= \prod_{i=0}^{n-1}\prod_{j=1}^{M-1} P(X_{j,i}|X_{j-1,i},X_{M,i},\Theta) \end{aligned}$$

where $X_{0,i} = y_i$ and $X_{M,i} = y_{i+1}$. Using the Modified Brownian Bridge (MBB), the density of $P(X_j|X_{j-1},X_M,\Theta)$ can be approximated by

$$P_{MBB}(X_{j,i}|X_{j-1,i},X_{M,i},\Theta) \approx N_X(\mu_{MBB},\sigma^2_{MBB}), \tag{5.14}$$

where

$$\begin{aligned} \mu_{MBB} &= X_{j-1,i} + \left(\frac{X_{M,i}-X_{j-1,i}}{\tau_{M,i}-\tau_{j-1,i}}\right)\delta_\tau, \\ \sigma_{MBB} &= g(X_{j-1,i},\Theta)\sqrt{\frac{M-j}{M-j+1}\delta_\tau}. \end{aligned}$$

Thus, the solution at the latent points is sampled by using

$$X_{j,i} = \mu_{MBB} + \sigma_{MBB}N_j$$

where N_j is a sample of the standard Gaussian random variable $N(0,1)$.

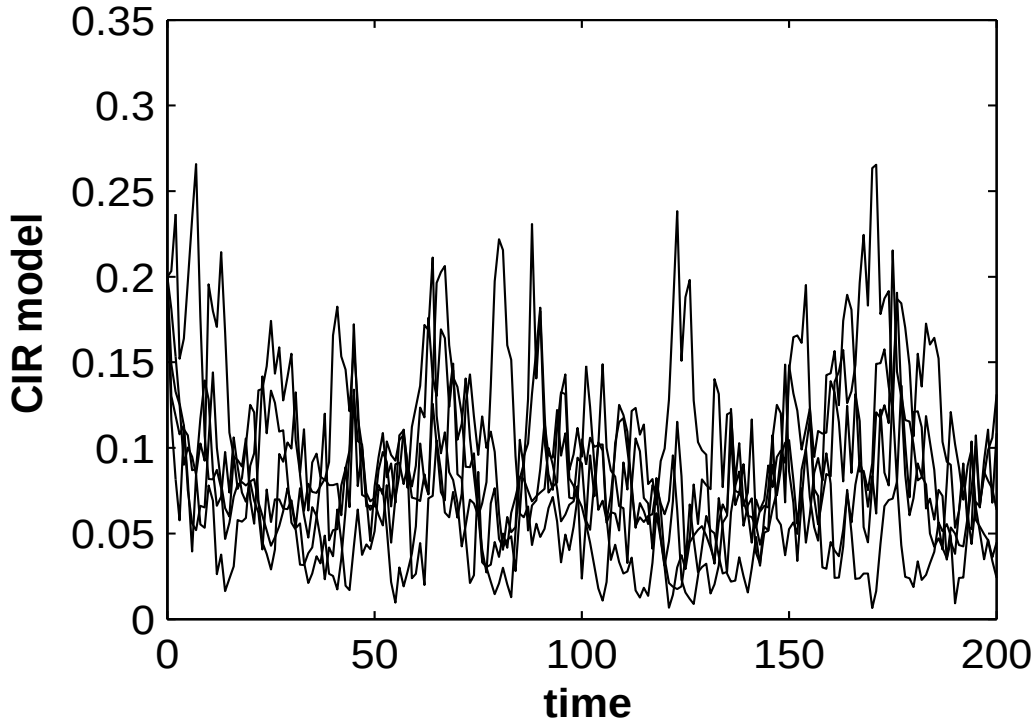

Figure 5.1: Five simulations of the CIR model with $\alpha = 0.2$, $\beta = 0.08$ and $\sigma = 0.1$.

When the importance sampling method is used, a number of samples of the latent variables are generated for $X_k \sim P_{MBB}(X_k|Y,\Theta)$ as described in Algorithm 1. Then we evaluate

$$P_{GaMBA}(\Theta|Y) \propto \frac{1}{K}\sum_{k=1}^{K}\frac{P_{Euler}(Y,X_k|\Theta)\cdot P(\Theta)}{P_{MBB}(X_k|\Theta,Y)} \tag{5.15}$$

5.4 Numerical Results

In this section, we use the numerical algorithm based on the Bayesian inference and MCMC method to estimate the parameters in the CIR model (5.6). Figure 5.1 gives 5 simulations of the CIR model with parameters $\alpha = 0.2$, $\beta = 0.08$ and $\sigma = 0.1$. When the volatility is not large, it shows that the values of short interest rate maintain positive. We use stepsize $h = 0.05$ in the numerical simulation to ensure the accuracy and stability property of simulations.

The estimated values of the parameters in the equation (5.6) for the mean-reverting test system and their standard deviations are given in Table 5.1, and more detailed simulation results of the Bayesian inference method are presented in Figure 5.2. In this test the size of the importance sampling is $K = 50$. For each parameter, we present the time series of the parameter values, the cumulative means and the histogram distribution. Compared with the exact values $(\alpha, \beta, \sigma) = (0.2, 0.08, 0.05)$, the Bayesian inference method gives estimates with good accuracy. For parameters β and σ, the histogram distributions are consistent with the cumulative means of the estimates. However, the histogram of parameter α is not symmetrical to the cumulative means.

Table 5.1: Estimated parameters and their standard errors

Parameters	$\alpha(0.2)$	$\beta(0.08)$	$\sigma(0.05)$
Estimated Values	0.1803	0.0792	0.0442
Standard Deviations	0.0109	0.0024	0.0013

In this study we tested the influence of the sample size in the importance sampling on the accuracy of the estimates. The sampling size was chosen as $K = 1, 10, 25, 50, 100, 200, 500$. Numerical results in Figure 5.3 show that the sampling size is important to improve the accuracy of the estimates but a larger sampling size does not necessary lead to much better accuracy, though numerical results in Figure 5.3 suggest that increasing sampling size can improve the accuracy slightly. Thus a reasonable size of the importance sampling is sufficient to generate estimates with adequate accuracy. This may be the reason that the sampling size is not very large in the previous studies [20].

We have also tested the influence of different samples of latent variables on the variation of estimates. In this test the simulated observations Y are kept unchanged. Figure 5.4 shows that the difference of sampling has certain influence on the variations of the estimates. The estimated model parameters vary in

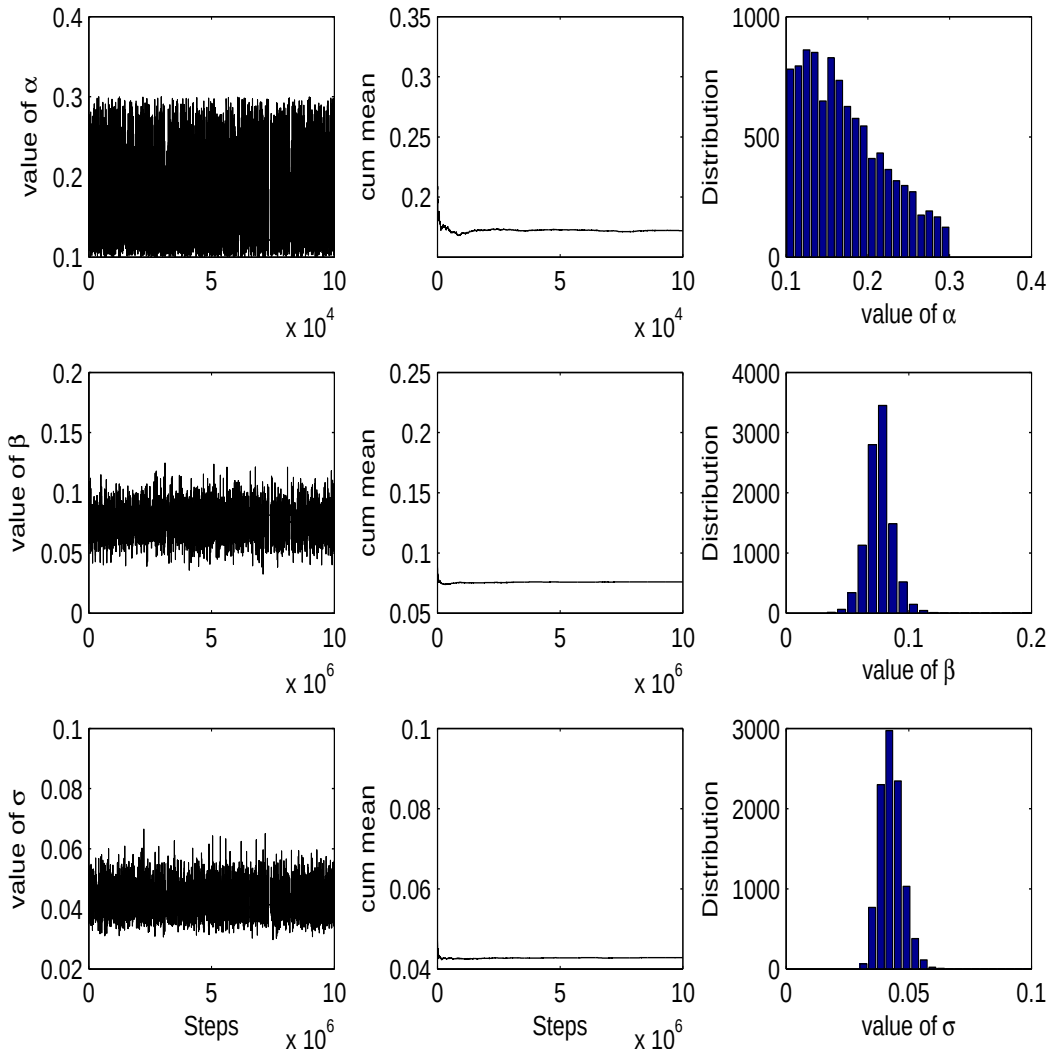

Figure 5.2: Simulation outputs for parameters α (the first row), β (the second row), and σ (the bottom row). Left column: the time series of the parameter values; middle column: the cumulative means of each parameter; right column: histogram distribution.

different tests. However, the variations in both the averaged parameter values and standard deviation are not large, which is consistent with the numerical results using the particle swarm optimization method to estimate model parameters. However, the variations of estimates are smaller than those obtained by using the genetic algorithm [138, 142].

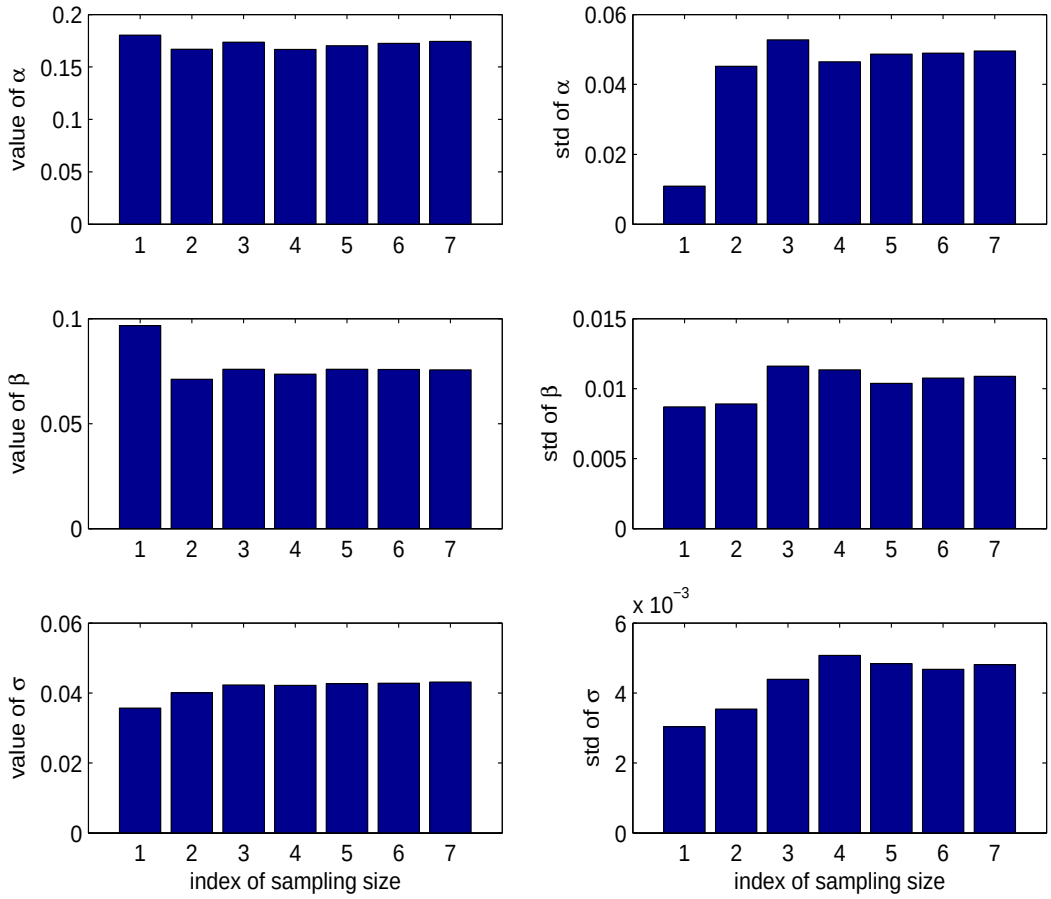

Figure 5.3: Estimated model parameters using different values of the importance sampling sizes. Left column: the estimated model parameters; right column: the standard deviation (std) of the estimates. The importance sampling size is $K = 1, 10, 25, 50, 100, 200, 500$ when index $= 1 \sim 7$.

5.5 Concluding Remarks

This work presents an effective algorithm for the estimation of parameters in stochastic differential equation models. The proposed approach is based on the Bayesian inference method and the Markov Chain Monte Carlo method. Compared with the grid method, the Markov Chain Monte Carlo based method can be used to infer stochastic models with a large number of unknown parameters. This method has been applied to an important stochastic model of the term structure of interest rate, which is a fundamental issue in the research area of financial mathematics. In addition, the importance sampling technique is used

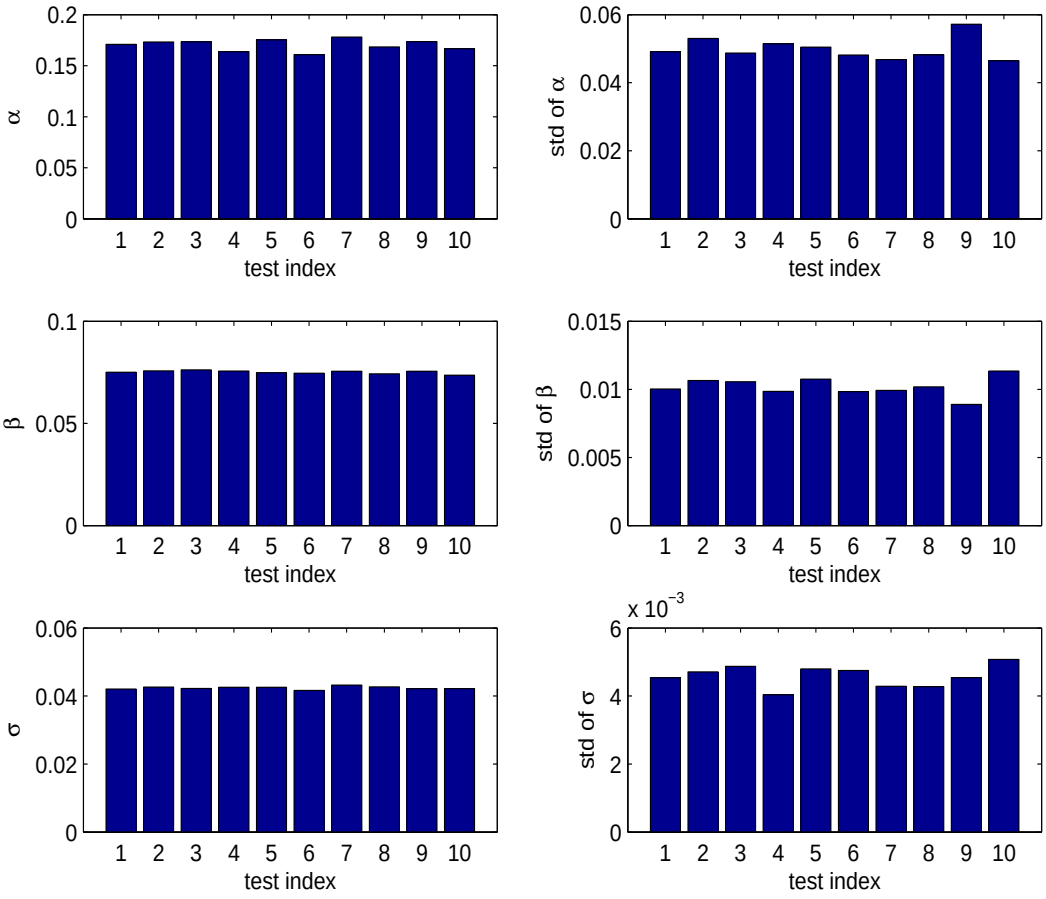

Figure 5.4: Variations of the estimated parameters when the sampling of the latent variables is different. Left column: the estimated model parameters; right column: the standard deviation (std) of the estimates.

to increase the robustness of estimates. We have also examined the influence of different samples of the latent variables on the variation of estimates. Numerical results suggest that the method used in this work is robust to such variation.

CHAPTER 6

Parameter Calibration of Term Structure Models: an Implicit Numerical Method with Particle Swarm Optimization

6.1 General Overview

Recently stochastic differential equations (SDEs) have been employed as a powerful tool to model the complex dynamics of a wide range of systems in biology, engineering, economics, finance and physical sciences. Compared with deterministic models, the key feature of a SDE model is its ability to generate a large number of different trajectories. However, this feature raises substantial challenges to the inference of unknown parameters in the SDE model, because experimental data actually represent only one simulation of the SDE models. To tackle the challenge, a number of methods have been proposed to infer reliable estimates. But these methods dominantly use explicit methods for solving SDEs, and thus are not appropriate to deal with experimental data with large variations. In this work, we develop a new method by using implicit methods to solve SDEs, which is aimed at generating stable simulations for stiff SDE models.

This chapter aims to address two issues: namely, establishing high order

implicit simulation methods and optimization methods. We will use the semi-implicit Milstein method to improve the accuracy and stability property of numerical simulations. In addition, we will test the efficiency of the particle swarm optimization algorithm. The remaining part of this chapter is organized as follows. Section 6.2 reviews the method of moment by using the stochastic models presented in chapter 5 for the parameter estimation. Section 6.3 presents the algorithms for generating parameter estimates, and the particle swarm optimization algorithm for searching the unknown parameters, and then proposes a new simulation method based on the particle swarm optimization. Section 6.4 reports the accuracy of the numerical results for the parameter estimates, and then carry out a validity test for the proposed algorithm. In section 6.5, we apply our proposed algorithm to the study of the financial case of US treasury bill data.

6.2 Method of Moment

A number of researchers concluded that the explicit Milstein method (5.4) can increase the estimation accuracy over the Euler method. In the following work we compare the accuracy of the Milstein method (5.4) and the semi-implicit Milstein method (5.5) for inferring the parameters of the CIR model. The application of the Milstein method to this model is straightforward, given by

$$X_{n+1} = X_n + \alpha(\beta - X_n)h + \sigma\sqrt{X_n}\Delta W_n + \frac{\sigma^2}{4}(\Delta W_n^2 - h) \tag{6.1}$$

Due to the linear feature of the drift term in the interest rate models, an explicit formula of the semi-implicit method can be obtained as follows,

$$X_{n+1} = \frac{1}{1+\alpha h}(X_n + \alpha\beta h + \sigma\sqrt{X_n}\Delta W_n + \frac{\sigma^2}{4}(\Delta W_n^2 - h)) \tag{6.2}$$

In Figure 6.1 we present 5 simulations of the CIR model with parameters $\alpha = 0.2$,

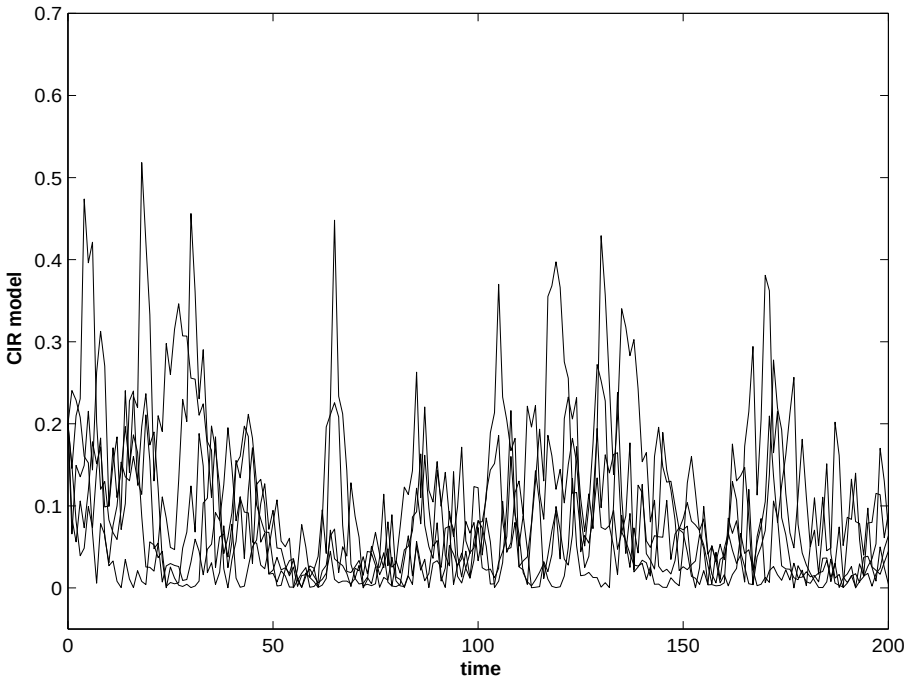

Figure 6.1: Five simulations of the CIR model.

$\beta = 0.08$ and $\sigma = 0.2$. When the volatility is not large, Figure 6.1 shows that the values of short rate maintain positive.

In the following, we shall estimate the parameter θ in the one-dimensional SDE

$$dX = a(X, \theta)dt + b(X, \theta)dW(t). \tag{6.3}$$

Now we sample X and get $(N + 1)$ observations $X_0, X_1, \ldots, X_n$ at the discrete time points $t_0, t_1, \ldots, t_N$, and thus the maximum-likelihood (ML) estimate of θ is generated by maximizing the likelihood function

$$L(\theta) = f_0(X_0|\theta) \prod_{k=0}^{N-1} f(X_{k+1}|X_k; \theta). \tag{6.4}$$

Equivalently, we can minimize the negative log-likelihood function to get the

estimate of θ

$$-\log L(\theta) = -\log[f_0(X_0|\theta)] - \sum_{k=0}^{N-1} \log[f(X_{k+1}|X_k;\theta)]. \tag{6.5}$$

where $f_0(X_0|\theta)$ is the density of the initial state X_0 and $f(X_{k+1}|X_k;\theta)$ is the value of the transitional probability density function (PDF) at (t_{k+1}, X_{k+1}) for the process starting at (t_k, X_k) and evolving to (t_{k+1}, X_{k+1}). Note that the Markovian property of Equation (6.3) ensures that the transitional PDF satisfies the Fokker-Planck equation. Unfortunately, in most cases the closed-form solution of the Fokker-Planck equation is not available and thus the exact maximum likelihood estimation is rare.

However, we can obtain the approximated transitional PDF by using the numerical solution of the original SDE (6.3). For instance, we use the Euler-Maruyama method to discretize (6.3) to yield

$$X_{k+1} = X_k + a(X_k, \theta)h + b(X_k, \theta)\Delta W_n, \tag{6.6}$$

where h is the stepsize of time discretization. Therefore the transitional PDF of X can be approximated by the normal distributed PDF with mean $X_k + a(X_k, \theta)h$ and variance $b^2(X_k, \theta)h$ such as

$$\frac{1}{b(X_k, \theta)\sqrt{2\pi h}} \exp\left[-\frac{(X_{k+1} - X_k - a(X_k, \theta)h)^2}{2b^2(X_k, \theta)h}\right]. \tag{6.7}$$

This is the simplest version of discrete maximum likelihood, namely the method of moment, when replacing the exact transitional PDF $f(X_{k+1}|X_k;\theta)$ in (6.5) with the approximated PDF above.

In this work, we focus on the CIR process (5.6). It has been established that the optimal values $\overline{\alpha}$ and $\overline{\beta}$ of the parameters α and β satisfy the equations

below [78]

$$\overline{\alpha}\left(\overline{\beta}\sum_{k=0}^{N-1}h-\sum_{k=0}^{N-1}X_k h\right) = X_N - X_0, \tag{6.8}$$

$$\overline{\alpha}\left(\overline{\beta}\sum_{k=0}^{N-1}\frac{h}{X_k}-\sum_{k=0}^{N-1}h\right) = \sum_{k=0}^{N-1}\frac{X_{k+1}-X_k}{X_k}, \tag{6.9}$$

and the optimal value $\overline{\sigma}$ of σ is

$$\overline{\sigma}^2 = \frac{1}{N}\sum_{k=0}^{N-1}\frac{(X_{k+1}-X_k-\overline{\alpha}(\overline{\beta}-X_k)h)^2}{X_k h}. \tag{6.10}$$

Although the Milstein variant has been proposed to increase the accuracy of the discrete maximum likelihood method [83, 96], it is difficult to derive an analytical expression of the parameter estimate from the transitional PDF. In that case the simulated maximum likelihood method is needed to estimate the transitional PDF from stochastic simulations.

6.3 Simulated Maximum Likelihood Method

6.3.1 Simulated Maximum Likelihood Function

We combine the simulated Milstein solution method and the particle swarm optimization method to construct an efficient algorithm for the parameter estimation. Parameter estimation in deterministic models can be achieved by the best fit of numerical simulations to experimental observations. However, this method is not feasible for SDE models because we can generate unlimited number of trajectories from a single SDE model. Here we use the simulated maximum likelihood (SML) method [11, 78] based on stochastic models. Given a sequence of $N+1$ observations $\{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_N\}$ at time points $\{t_0, t_1, \ldots, t_N\}$, we define the joint

transitional density or likelihood function of these observations as

$$f_0[(t_0, \mathbf{X}_0)|\theta] \prod_{i=1}^{N} f[(t_i, \mathbf{X}_i)|(t_{i-1}, \mathbf{X}_{i-1}), \ldots, (t_0, \mathbf{X}_0); \theta], \tag{6.11}$$

where $\theta = (\theta_1, \ldots, \theta_s)$ are the parameters in model (6.3) to be determined, $f_0[(t_0, \mathbf{X}_0)|\theta]$ is the density of the initial state, and

$$f[(t_i, \mathbf{X}_i)|(t_{i-1}, \mathbf{X}_{i-1}), \ldots, (t_0, \mathbf{X}_0); \theta]$$

is the transitional density starting from $(t_{i-1}, \mathbf{X}_{i-1})$ and evolving to $(t_i, \mathbf{X}_i)$. When the financial system is described by the stochastic model (6.3), the stochastic process $\mathbf{X}$ is Markov [41], and the transitional density can be simplified as

$$f[(t_i, \mathbf{X}_i)|(t_{i-1}, \mathbf{X}_{i-1}), \ldots, (t_0, x_0); \theta] = f[(t_i, \mathbf{X}_i)|(t_{i-1}, \mathbf{X}_{i-1}); \theta]. \tag{6.12}$$

An equivalent form of the maximum of the joint transitional density (6.11) is the minimum of the negative log-likelihood function (6.5) when time t is not explicitly presented in the formula.

Because the closed-form expression of the transitional density (6.12) is usually unavailable, we use a nonparametric kernel density function below

$$\overline{f}_M\left[(t, \mathbf{X})|(t_{i-1}, \mathbf{X}_{i-1}); \theta\right] = \frac{1}{MB} \sum_{j=1}^{M} K\left(\frac{\mathbf{X} - \mathbf{Y}_j}{B}\right) \tag{6.13}$$

in substitution for the transitional density. Here $\mathbf{Y}_1, \ldots, \mathbf{Y}_M$ are the M realizations of $\mathbf{X}_i$ at t_i given the initial condition $(t_{i-1}, \mathbf{X}_{i-1})$, and B is the kernel bandwidth and $K(\cdot)$ is a non-negative kernel function enclosing unit probability mass. In the case of SDE models with a single variable, the normal kernel is

widely used and the bandwidth can be chosen as

$$B = 0.9\sigma M^{-1/5},$$

wherein σ is the sample standard deviation of the M realizations [78]. For those multivariate stochastic models, we can either assume the independence of random variables or use the theory of multivariate density estimation [42].

6.3.2 Particle Swarm Optimization Algorithm

Another issue after setting up the object function is to choose a fast method to search for the optimal parameters. In the past decade, the particle swarm optimization (PSO) algorithm has been successfully applied in many research and application areas. It has been concluded, in many research papers, that PSO and the GA are able to arrive at solutions with the same quality. However, PSO offers a less expensive approach and there are fewer parameters to adjust than the GA [98, 124]. Numerical tests suggest that the PSO offers more computational saving for unconstrained nonlinear problems with continuous design variables whereas the computational saving is lower for constrained and mixed integer nonlinear problems [98].

The PSO algorithm is a population-based stochastic optimization technique developed by Dr.Eberhart and Dr.Kennedy in 1995 [115]. Unlike the genetic algorithm, the PSO algorithm, which is inspired by the social behaviour of bird flocking or fish schooling, has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. In this work, we use a PSO MATLAB toolbox downloaded from the MATLAB File Exchange Central [94] to estimate parameters in SDE models. This carefully-designed software system can be implemented in a wide range of optimization problems. Now we develop a

SML method to estimate the optimal value of unknown parameter θ in the SDEs model (6.3) by minimizing the log-likelihood function (6.5) over θ as follows.

6.3.3 A New Simulation Algorithm Based on Particle Swarm Optimization

Algorithm 6.1.

Step 1. Input the system states $\{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_N\}$ and time points $\{t_0, t_1, \ldots, t_N\}$.

Step 2. Take $\mathbf{X}_{i-1}$ at time t_{i-1} $(i = 1, \ldots, N)$ as the starting value and use the implicit Milstein method to generate M realizations $\mathbf{Y}_1, \ldots, \mathbf{Y}_M$ of $\mathbf{X}$ at t_i. A random seed is specified for generating samples of the Gaussian random variables.

Step 3. Use the nonparametric density (6.13) with the normal kernel or multivariate density functions to evaluate the transitional density (6.12).

Step 4. Steps 2 and 3 are repeated for each time point $t_0, \ldots, t_{N-1}$, and results are used to construct the log-likelihood function (6.5).

Step 5. Search the optimal kinetic rate by the particle swarm optimisation algorithm based on the minimum of $L(\theta)$ in (6.5).

Note that the same increments of the Wiener process should be used in numerical simulations with different values of parameter θ. In order to reduce the variation of the estimated parameters, the same random seeds (namely the same random samples) in Step 2 should be used in different candidate estimates of parameters.

6.4 Estimation of Parameters in the Interest Rate Models

6.4.1 Estimation Results

Using the methods discussed in the previous sections, now we estimate parameters in the CIR model. Given a set of parameters (namely exact parameters), we use the semi-implicit Milstein method with a very small stepsize ($h = 0.001$) to generate 20 trajectories. For each generated trajectory, we used the method of moment, the SML method with the explicit Milstein method, and the SML method with the semi-implicit Milstein method to estimate the model parameters. For each of these three types of methods, we used the PSO algorithm (the population size of 40 and 200 generations) to obtain 20 estimated sets of model parameters. Then we evaluated the mean and standard deviation (STD) of the errors of the estimates to the exact parameters.

Table 6.1: Estimation results of the parameters in the CIR model

	Moment method			Milstein			Semi-implicit		
	Mean	Bias	STD	Mean	Bias	STD	Mean	Bias	STD
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.1$, $\Delta = 5$)									
α	0.0501	0.1499	0.0103	0.2052	0.0052	6.3E-4	0.2091	0.0091	1.0E-3
β	-0.0118	0.0918	0.0193	0.0814	0.0011	1.1E-4	0.0787	0.0013	1.6E-6
σ	0.1067	0.0933	0.0148	0.0955	0.0045	3.3E-5	0.0933	0.0067	3.6E-5
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.2$, $\Delta = 5$)									
α	0.1013	0.0987	0.0253	0.2184	0.0184	0.0037	0.2116	0.0116	0.0028
β	0.0722	0.0078	0.0271	0.0812	0.0012	0.0005	0.0808	0.0008	0.0005
σ	0.3650	0.1650	0.2647	0.1835	0.0165	0.0006	0.1850	0.0150	0.0006
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.3$, $\Delta = 5$)									
α	0.0907	0.1093	0.0490	0.3269	0.1269	0.0158	0.3023	0.1023	0.0096
β	0.0746	0.0054	0.0230	0.0496	0.0304	0.0005	0.0534	0.0266	0.0006
σ	0.3319	0.1319	0.2128	0.2434	0.0566	0.0015	0.2375	0.0625	0.0007

Numerical results presented in Table 6.1 show that the method of moment cannot generate reliable estimates of the model parameters. There is significant

difference between the estimates of the same parameter in different implementations. In some cases the relative error is over 100%. Moreover, we even obtained negative coefficients, which is meaningless in finance, even though the strength of noise is small. However, our further tests suggested that the accuracy of the moment method relys upon the length of observation points Δ and the number of observation time points. When the moderate length and a relatively large number of observation points are chosen, the moment method could provide estimates with acceptable accuracy.

In order to examine how the noise affects the estimation accuracy, three values of the volatility control parameter σ were tested in each SDE model. As we can see from Table 6.1, the SML method with either the explicit Milstein method or the semi-implicit Milstein method provides reliable estimates of the parameters with small estimation errors and standard deviations when the fluctuations are small in the SDE model ($\sigma = 0.1$). Actually the SML method with the explicit Milstein method gave more accurate estimates. One possible explanation is that when it comes to the non-stiff SDEs, the explicit Milstein method can give better accuracy of simulations than the semi-implicit Milstein method. When the fluctuations in the SDE models are mild ($\sigma = 0.2$), the figures show that the semi-implicit Milstein method provides better parameter estimates in comparison with the explicit Milstein method. In this case, the better stability property of the semi-implicit method is more important than the slightly better accuracy property of the explicit method. When the noise in the interest rate models is large ($\sigma = 0.3$), the estimated parameters with acceptable accuracy still can be obtained by using the semi-implicit Milstein method. But, if the noise components in the SDE models are very large, then even the semi-implicit Milstein method could not produce reliable estimated model parameters. In this case, we may try a smaller stepsize in simulating the SDE models or use the fully implicit Milstein method [139] to guarantee the stability property of the numerical simulations.

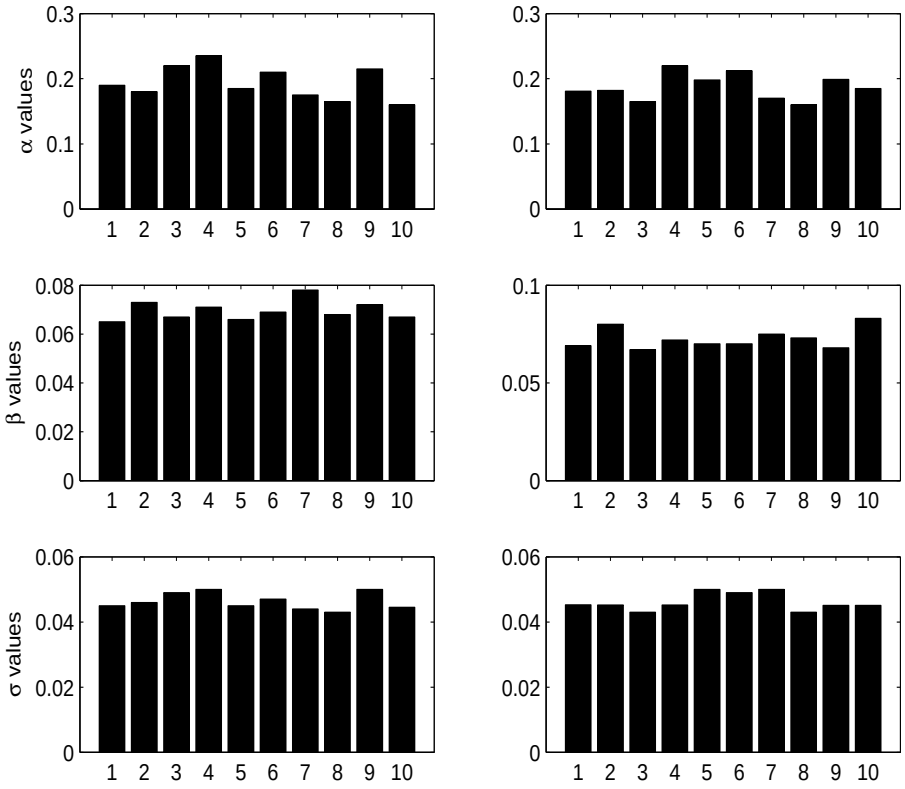

Figure 6.2: Ten sets of parameter estimates based on different random seeds for generating simulations of the CIR SDE model (left column) and different random seeds for generating the initial parameter values in the PSO algorithm (right column).

6.4.2 Test of the Estimation Algorithm

An important issue in the stochastic search methods is the robustness of estimation results. There are two possible resources of estimation variations in the proposed algorithm in this work: the variation of estimates due to different random seeds in Step 2 for simulating the SDE model; and the variation due to the different implementations of the PSO algorithm using different random samples.

The first variation is partially related to the convergence property of the Monte-Carlo simulation. The solution to this problem is to increase the number of stochastic simulations. We fixed one trajectory of the interest rate and

used 10 sets of random seeds in the Milstein methods to generate simulations of the stochastic model. Figure 6.2 (left column) shows that the estimates are relatively stable when the number of simulation is more than 5000. Further tests suggest that when the simulation number is $N = 10000$, we obtain very stable estimates. This result is consistent with our previous computing experience for estimating rate constants in discrete chemical reaction systems [142], which was implemented in a genetic algorithm (GA).

Our previous experiments suggested that the values of the initial parameters have significant influence on the final estimate when the genetic algorithm was applied as the stochastic searching method [138, 142]. The second variation is mainly from the influence of different initial model parameters in the PSO algorithm on the final estimated results. In this chapter, we use different random seeds in the PSO algorithm to generate initial model parameters for the CIR stochastic model. The estimated parameters are presented in the right column of Figure 6.2. Numerical results suggest that the PSO can produce reliable estimates nearly independent of the initial model parameters, which is a significant advantage of the PSO algorithm over the genetic algorithm in the parameter calibration of complex mathematical models. On the other hand, we may lose the opportunity to select the optimal estimate from a number of candidate estimates based on other criteria such as the robustness property of the mathematical model.

6.5 Application to US Treasury Bill Data

The classic one factor model of the term structure of the instantaneous interest rate r is given by

$$dr = \alpha(\theta - r)dt + \sigma r^{\gamma} dw \tag{6.14}$$

where dw is the standard Wiener process, α the speed of adjustment parameter, θ the mean interest rate, σ volatility control and γ the levels effect. These

parameters need to be estimated in the following work. Previously we discussed the CIR model, which is a special case of model (6.14) ($\gamma = 0.5$). Lots of recent empirical evidence suggests that γ should be estimated rather than imposed.

Estimating the parameters of the stochastic differential equation (SDE) (6.14) is a worthwhile and challenging task. Here we use the US 3-month Treasury Bill rate to give a reasonable approximation to the unobservable instantaneous short interest rate. The moment method is applied to this instance as a benchmark. The discrete equations of the one factor model for the Milstein method and semi-Milstein method of SML estimation are as follows:

$$r_{n+1} = r_n + \alpha(\theta - r_n)h + \sigma r_n^{\gamma}\Delta W_n + \frac{1}{2}\sigma^2\gamma r_n^{2\gamma-1}((\Delta W_n)^2 - h) \tag{6.15}$$

and

$$r_{n+1} = \frac{1}{1+\alpha h}(r_n + \alpha\theta h + \sigma r_n^{\gamma}\Delta W_n + \frac{1}{2}\sigma^2\gamma r_n^{2\gamma-1}((\Delta W_n)^2 - h)). \tag{6.16}$$

Table 6.2: Estimated parameter and standard error (in bracket) of the classic one-factor model (6.18)

Method Parameters	Moment method $\Delta = 0.01$	Milstein $M = 1000, \Delta = 0.01$	Semi-implicit $M = 1000, \Delta = 0.01$
α	0.0112(0.0094)	0.0107(0.0092)	0.0105(0.0056)
θ	0.0401(0.0115)	0.0399(0.0106)	0.0405(0.0058)
σ	0.0155(0.0060)	0.0143(0.0011)	0.0161(0.0027)
γ	0.6646(0.0923)	0.6680(0.0347)	0.6683(0.0508)

We use monthly US Treasury bill data from January 1985 to December 2007 with 276 observations in total to estimate the model parameters in the one-factor model (6.18). Figure 6.3 gives the evolution of monthly interest rate of US 3 month Treasury Bills. The estimated parameter values derived from the moment method, the explicit Milstein method and the semi-implicit Milstein method are shown in Table 6.2. In addition, the standard errors of estimates from the three

methods are also listed in Table 6.2 in brackets. Most of the estimates are broadly in agreement. Due to the large amount of available data, the Moment method still obtained estimates with good accuracy. However, the standard errors of the two Milstein methods are smaller than that of the moment method. We also tested the two Milstein methods using smaller stepsize in numerical simulation and more simulating numbers M in the inference methods. The improvement in reducing standard errors is not substantial.

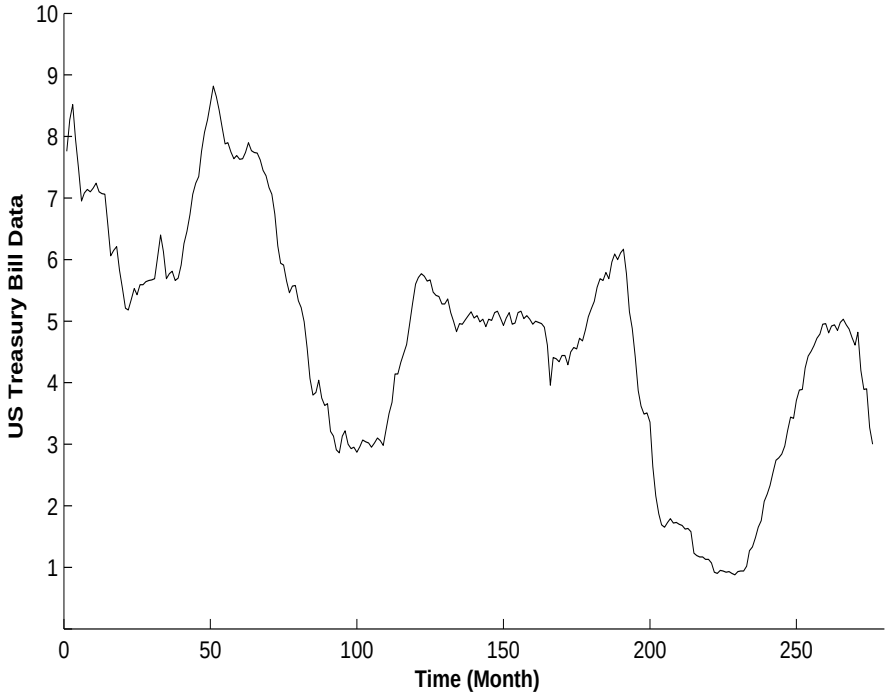

Figure 6.3: Interest rate series of the US 3-month Treasury Bills

Figure 6.3 gives the evolution of monthly interest rate of US 3 month Treasury Bills. As we can see from this picture, the estimates of the mean interest rate θ in Table 6.2 match the interest rate data very well. In particular, the estimate of the semi-implicit Milstein method is larger and also better than those of the other two methods. An interesting observation is that the level effect parameter γ obtained from these three methods are quite consistent to each other, compared with the other parameters. Moreover, the estimated standard error of γ is smaller than

the difference between the estimated values of γ and 0.5, which suggests that its value may be different from 0.5 in the CIR model. Therefore, the classic one-factor model (6.18) provides a better description than the CIR model and the value of parameter γ should be estimated from financial data rather than imposed in the interest term structure models.

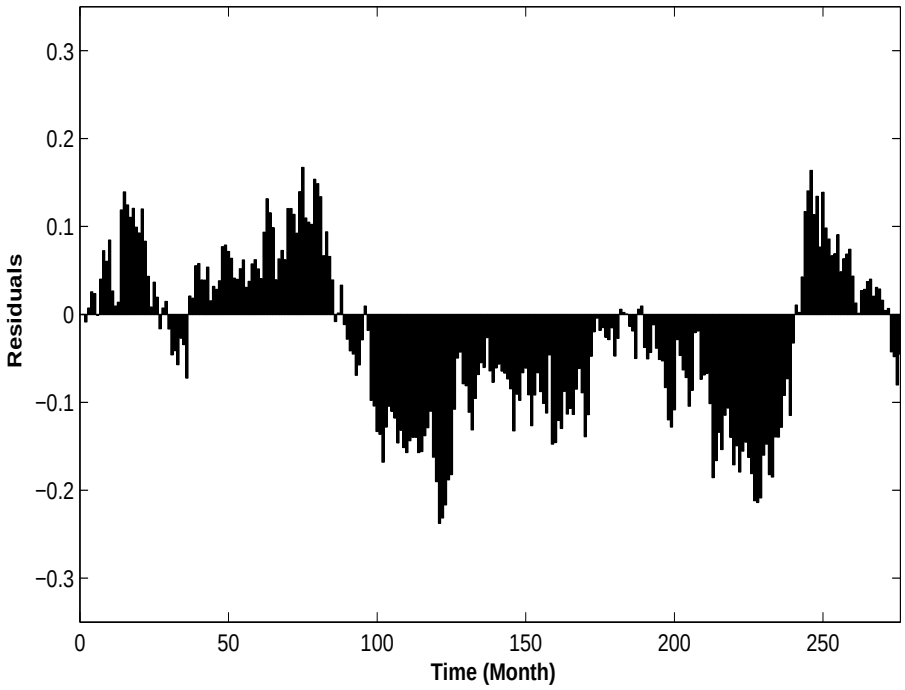

Figure 6.4: Residual of simulated and real data for 2000 simulations

To obtain some indications of the goodness of fit of the interest rate term structure model with levels effect, fitted observations were generated in the following manner. Taking the parameter estimates for the SML with semi-explicit Milstein method based on $\Delta = 0.01$, we simulated 5000 times between observations and calculated the mean of these trials. The deviations from the actual data were showed in Figure 6.4, where we can tell that most of the deviations are in a statistically acceptable range.

6.6 Concluding Remarks

In this work, we proposed an effective algorithm for parameter estimation of the stochastic differential equation models. This approach is based on the implicit numerical scheme in the Monte Carlo simulation integrated with the particle swarm optimization method. We conclude that the simulated maximum likelihood with the semi-implicit Milstein method can provide parameter estimates with both higher convergence order and better stability property than the widely used Euler-Maruyama method. Experimental results show that when the SDE model is moderate stiff, estimates produced by the semi-implicit Milstein method have better accuracy than those obtained by the explicit Milstein method. Furthermore, the PSO algorithm can produce reliable estimates which are nearly independent of the implementation details compared with the genetic algorithm. In the application to the stochastic model for the term structure of interest rate, we use the actual economic data to estimate the parameters. The model with the estimated parameters fits the real data very well with small errors.

CHAPTER 7

Summary and Suggestions for Future Research Directions

7.1 Summary of the Main Contributions

In this thesis, we consider the solution of various stochastic differential equation models for applications in finance. Various new algorithms and computational methods have been developed to solve these models and identify the model parameters. This involves establishment of a variety of novel techniques. The results obtained can be summarized in three aspects as follows.

(1) Development of a Taylor numerical method for jump-diffusion stochastic delay differential equations

We developed a robust Taylor approximation scheme for weak solutions of general stochastic delay differential equations with jumps. A convergence theorem for the scheme was constructed and proved.

Our numerical analysis shows that the high order weak approximation is efficient for jump-diffusion stochastic delay differential equation models by using Monte Carlo simulation.

(2) Development of a fractional stochastic differential equation model for application in option pricing

A fractional stochastic differential equation model was established for addressing the trend memory effect in financial pricing. Based on the model, its stochastic Ito formula and the European option pricing formula have been derived. The existence of the trend memory (i.e., the mean value function) in the option pricing formula when the Hurst index is between 0.5 and 1, has been established and proved.

We have conducted a comparison analysis among our proposed model, the classic Black-Scholes model, and the stochastic model with fractional Brownian motion. Numerical results suggest that our model leads to more accurate estimates and lower standard deviation in the empirical study.

(3) Development of computational algorithms for identifying parameters in stochastic differential equation models

We have developed a new method for estimating parameters in the mean-reverting stochastic systems, more specifically, We have established an effective algorithm for the estimation of parameters in SDE models based on the use of the Bayesian inference approach and the MCMC method. The importance sampling technique was used to increase the robustness of estimates. Numerical results suggested that the proposed method is robust to the variation of samples of the latent variables.

A novel parameter identification method has been proposed by using implicit simulation method and particle swarm optimization searching approach as detailed below

(a) A new method is established by using implicit methods to solve SDEs, which is aimed at generating stable simulations for stiff stochastic differential equation models. The particle swarm optimization method is used as an efficient searching method to explore the optimal estimate in the complex parameter space.

(b) Using the interest rate term structure model as the test system, numerical results showed that the proposed new method is an effective approach for

generating reliable estimates of unknown parameters in SDE models.

(c) By applying the proposed algorithm to the stochastic model of term structure of interest rate, we obtained the estimated parameters by using the actual economic data. The case study indicates that the model with the estimated parameters fits the real data very well with small errors.

7.2 Future Research Directions

In this thesis, our main work is in the development of computational algorithms for solving stochastic differential equations and fractional stochastic differential equations, and also in the development of numerical methods for estimating parameters in the stochastic models. It is observed that these algorithms are computationally very effective for solving all the problems under consideration. To make significant advancement, new and more efficient computational algorithms could be derived for solving existing stochastic problems and new unconventional problems arising in the study of real world practical problems. Further possible improvements and advancements may be made in the following directions:

(1) There is much scope for further work in the context of weak solution of jump-diffusion SDDEs. For example, it is clearly of great importance to extend the weak convergence theory to the case where coefficients in the equations are not globally Lipschitz, and to develop and analyse new methods that maintain good properties of convergence and stability.

(2) In this work, we develop fractional stochastic differential equation for addressing memory effects in the financial market. An interesting future work is to improve our model by connecting fractional ordinary differential equation with fractional Brownian motion, which can describe both the trend memory and the noise memory.

(3) We introduced the Gaussian Modified Bridge Approximation into the

Markov Chain Monte Carlo simulation and examine the accuracy and robustness of this approach. It is worth to note that the performance of the MCMC is related to a number of important factors, such as convergence criteria, burn-in, and thinning to reduce autocorrelation, and thus further efforts are needed to discuss these issues. In addition, it is still a challenging problem for estimating parameters in stiff stochastic models. Alternatively we may consider the implicit methods or high order methods rather than the explicit Euler method. Thus more effective calibration methods should be designed for estimating parameters in stiff SDEs.

(4) It is possible that the current approaches for estimating model parameters may fail to generate reliable estimates when the experimental observations have large variations and the length between the consecutive observations is large. Thus more effective calibration methods should be designed for estimating parameters in stiff SDEs.

(5) Efficiency is another major issue of the calibration methods for stochastic models. Since a large number of trajectories are needed to calculate the transitional probability distribution function, any improvement over the numerical efficiency will significantly reduce the total computational time. The variable-stepsize simulation methods can be used in stochastic simulations. In addition, the optimization methods for searching the optimal estimates from the space with complex error landscape are particularly important in the calibration of stochastic models.

Bibliography

[1] Clement E. Kohatsu-Higa A. and Lamberton D. A duality approach for the weak approximation of stochastic differential equations. *Annals of Applied Probability*, 16:1124–1154, 2006.

[2] Garzareli F. Cristelli M. Zaccaria A. and Pietronero L. Memory effects in stock price dynamics: Evidence of technical trading. *Scientific Reports*, 4:Article 4487, 2014.

[3] Kohatsu-Higa A. Weak approximations: A Malliavin calculus approach. *Mathematics of Computation*, 70:135–172, 2001.

[4] Michaelides A. and Ng S. Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators. *Journal of Econometrics*, 96(2):231–266, 2000.

[5] Sanford A.D. and Martin G.M. Simulation-based Bayesian estimation of an affine term structure model. *Computational Statistics & Data Analysis*, 49:527–554, 2005.

[6] Jazwinski A.H. *Stochastic processes and filtering theory.* New York: Academic Press, New York, 1970.

[7] Pereira Ló B. Haslam A.J. and Adjiman C.S. An algorithm for the estimation of parameters in models with stochastic differential equations. *Chemical Engineering Science*, 63:4820–4833, 2008.

[8] Gallant A.R. and Tauchen G. The relative efficiency of method of moments estimators. *Journal of Econometrics*, 92:149–172, 1999.

[9] Gallant A.R. and Tauchen G. The relative efficiency of method of moments estimators. *Journal of Econometrics*, 92:149–172, 1999.

[10] Gallant A.R. and Tauchen G.E. Which moments to match? *Econometric Theory*, 12(4):657–681, 1996.

[11] Hurn A.S. and Lindsay K.A. Estimating the parameters of stochastic differential equations. *Mathematics and Computers in Simulation*, 48:373–384, 1999.

[12] Svishchuk A.V. and Kazmerchuk Y.I. Stability of stochastic Ito equations delays, with jumps and Markov switchings, and their application in finance. *Theory of Probability and Mathematical Statistics*, 64:141–151, 2002.

[13] Lo A.W. Long term memory in stock market prices. *Econometrica*, 59:1279–1313, 1991.

[14] Bibby B. and Sorensen M. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, 1:17–39, 1995.

[15] Hu Y. Oksendal B. and Sulem A. Optimal consumption and portfolio in a Black-Scholes market driven by fractional Brownian motion. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 6(4), 2003.

[16] Huang B. and Yang C.W. The fractional structure in multinational stock returns. *Applied Economics Letters*, 2:67–71, 1995.

[17] Mandelbrot B.B. and Wallis J.R. Noah, joseph and operational hydrology. *Water Resources Research*, 4(5):909–918, 1968.

[18] Mandelbrot B.B. and Van Ness J.W. Fractional Brownian motions, fractional noises and applications. *Society for Industrial and Applied Mathematics*, 10(4):422–437, 1968.

[19] West B.J. and Picozzi S. Fractional Langevin model of memory in financial time series. *Physical Review EStatistical, Nonlinear, and Soft Matter Physics*, 65(3):Article ID 037106, March 2002.

[20] Joshi C. and Wilson S. Grid based Bayesian inference for stochastic differential equation models. *Technical Paper, Trinity College Dublin*, 2012.

[21] Liu C. and Rubin D.B. Model-based analysis to improve the performance of iterative simulations. *Statistica Sinica*, 12:751–767, 2002.

[22] Necula C. Option pricing in a fractional Brownian motion environment. *Mathematical Reports*, 6(56)(3):259–273, 2004.

[23] Tudor C. Approximation of delay stochastic equations with conatant retardation by usual Ito equations. *Romanian Journal of Pure and Applied Mathematics*, 34(1):55–64, 1989.

[24] Tudor C. and Tudor M. On approximation of solutions for stochastic delay equations. *Stud. Cerc. Mat.*, 39(3):265–274, 1987.

[25] Heyde C.C. *Quasi-likelihood and its application: A general approach to optimal parameter estimation.* Springer, NY, 1992.

[26] Evertsz C.J.G. Fractional geometry of financial time series. *Fractals*, 3:609–616, 1995.

[27] Liu Y. Gopikrishnan P. Cizeau P. Meyer M. Peng C.K. and Stanley H.E. Statistical properties of the volatility of price fluctuations. *Physical Review E*, 60(2):1390–1400, 1999.

[28] Jones C.S. Nonlinear mean reversion in the short-term interest rate. *Review of Financial Studies*, 16:793–843, 2003.

[29] Haverkamp B.R.J. Verhaegen M. Chou C.T. and Johansson R. Continuous-time subspace model identification method using laguerre filtering. *In: SYSID' 97-11th IFAC Symposium on System Identification*, 1997.

[30] Baker C.T.H. and Buckwar E. Introduction to the numerical analysis of stochastic delay differential equations. *Technical report, University of Manchester, Dept. of Mathematics*, Numberical Analysis Report 345, 1999.

[31] Komorowski M. Finkenstadt B. Harper C.V. and Rand D.A. Bayesian inference of biochemical kinetic parameters using the linear noise approximation. *BMC Bioinformatics*, 10:343, 2009.

[32] Duffle D. and Singleton K.J. Simulated moments estimation of markov models of asset prices. *Econometrica*, 61:929–952, 1993.

[33] Duffle D. and Glynn P. Estimation of continuous-time Markov processes sampled at random time intervals. *Econometrica*, 72(6):1773–1808, 2004.

[34] Florens-Zmirou D. Approximate discrete-time schemes for statistics of diffusion processes. *Statistics*, 20:547–557, 1989.

[35] Gallant A.R. Hsieh D. and Tauchen G.E. Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics*, 81:159–192, 1997.

[36] Monje C.A. Chen Y. Vinagre B. Xue D. and Feliu V. *Fractional-order systems and controls: Fundamentals and applications*. Springer, London, UK, 2010.

[37] Boys R.J. Wilkinson D.J. and Kirkwood T.B.L. Bayesian inference for a discretely observed stochastic kinetic model. *Statistics and Computing*, 18:125–135, 2008.

[38] Wilkinson D.J. Bayesian methods in bioinformatics and computational systems biology. *Brief Bioinform*, 8:109–116, 2007.

[39] Cajueiro D.O. and Tabak B.M. The Hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. *Physica A : Statistical Mechanics and its Applications*, 336(3-4):521–537, 2004.

[40] Cajueiro D.O. and Tabak B.M. Ranking efficiency for emerging markets. *Chaos, Solitons and Fractals*, 22(2):349–352, 2004.

[41] Gillespie D.T. Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*, 115:1716–1733, 2001.

[42] Scott D.W. *Multivariate density estimate: Theory, practiceand vasualization*. Wiley, NY, 1992.

[43] Buckwar E. and Shardlow T. Weak approximation of stochastic differential delay equations. *IMA Journal of Numerical Analysis*, 25:57–86, 2005.

[44] Panas E. Long memory and chaotic models of prices on the London metal exchange. *Resources Policy*, 27(4):485–490, 2001.

[45] Platen E. and Bruti-Liberati N. *Numerical solution of stochastic differential equations with jumps in finance*. Springer, Berlin, Germany, 2010.

[46] Lawrence C. Evans. An introduction to stochastic differential equations. *Lecture notes, University of California, Dept. of Mathematics*, 2006.

[47] Black F. and Scholes M. The pricing of option and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[48] Chen J. Ren F. and Qiu W. Option pricing of a mixed fractional-fractional version of the Black-Scholes model. *Chaos, Solitons and Fractals*, 21(5):1163–1174, 2004.

[49] Schweppe F. Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory*, 11:61–70, 1965.

[50] Wen F. and Yang X. Skewness of return distribution and coefficient of risk premium. *Journal of Systems Science and Complexity*, 22(3):360–371, 2009.

[51] Wen F. and Liu Z. A copula-based correlation measure and its application in chinese stock market. *International Journal of Information Technology and Decision Making*, 8(4):787–801, 2009.

[52] Chan K.C. Karolyi G.A. Longstaff F.A. and Sanders A.B. An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, 47:1209–1227, 1992.

[53] De Rossi G. Maximum likelihood estimation of the Cox-Ingersoll-Ross model using particle filter. *Computational Economics*, 36:1–16, 2010.

[54] Falin G. and Fricker C. On the virtual waiting time in an $M/G/1$ retrial queue. *Journal of Applied Probability*, 28(2):446–460, 1991.

[55] Jumarie G. Stochastic differential equations with fractional Brownian motion input. *International Journal of Systems Science*, 24(6):1113–1131, 1993.

[56] Jumarie G. Fractional Brownian motions via random walk in the complex plane and via fractional derivative, comparison and further results on their Fokker-Planck equations. *Chaos, Solitons and Fractals*, 22(4):907–925, 2004.

[57] Jumarie G. On the representation of fractional Brownian motion as an integral with respect to $(dt)^{\alpha}$. *Applied Mathematics Letters*, 18(7):739–748, 2005.

[58] Jumarie G. On the solution of the stochastic differential equation of exponential growth driven by fractional Brownian motion. *Applied Mathematics Letters*, 18(7):817–826, 2005.

[59] Durham G.B. and Gallant A.R. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business and Economic Statistics*, 20:297–316, 2002.

[60] Gilsing. On the stability of the Euler scheme for an affine stochastic differential equations with time delay. *Technical report, Humboldt University, Berlin*, Discussion Paper 20:373, 2001.

[61] Milstein G.N. and Tretyakov M.V. Numerical methods in the weak sense for stochastic differential equations with small noise. *SIAM Journal on Numerical Analysis*, 34:2142–2167, 1997.

[62] Milstein G.N. and Tretyakov M.V. *Stochastic numerics for mathematical physics*. Springer, Berlin, 2005.

[63] Madsen H. and Melgaard H. The mathematical and numerical methods used in CTLSM - a program for ML-estimation in stochastic, continuous time dynamical models. *Technical report, Institute of Mathematical Statistics and Operations Research*, 1991.

[64] Melgaard H. and Madsen H. CTLSM version 2.6 - a program for parameter estimation in stochastic differential equations. *Technical report, IMSOR. No. 1*, 1993.

[65] Risken H. *The Fokker-Planck equation (2nd ed.)*. Springer, Berlin, 1989.

[66] Hurst H.E. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.

[67] Andersen T.G. Chung H.J. and Screnseb B.E. Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics*, 91(1):61–87, 1999.

[68] Zhang H.M. and Gan S.Q. The split-step backward Euler method for linear stochastic delay differential equations. *Journal of Computational and Applied Mathematics*, 225(2):558–568, 2009.

[69] Gelman A. Carlin J.B. Stern H.S. and Rubin D.B. *Bayesian data analysis.* Chapman and Hall, London, 1995.

[70] Pollubny I. *Fractional differential equations.* Academic Press, San Diego, Calif, USA, 1999.

[71] Shoji I. and Ozaki T. Comparative study of estimation methods for continuous time stochastic processes. *Journalof Time Series Analysis*, 18:485–506, 1997.

[72] Luo J. Comparison principle and stability of Ito stochastic differential delay equations with Poisson jump and Markovian switching. *Nonlinear Analysis: Theory, Methods and Applications*, 64(2):253–262, 2006.

[73] Timmer J. Modeling noisy time series: physiological tremor. *International Journal of Bifurcation and Chaos*, 8:1505–1516, 1998.

[74] Timmer J. Parameter estimation in nonlinear stochastic differential equations. *Chaos Solitons Fractals*, 11:2571–2578, 2000.

[75] Cox J.C. Ingersoll J.E. and Ross S.A. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[76] He S.W Wang J.G. and Yan J.A. *Semimartingale theory and stochastic calculus.* CRC Press, Boca Raton, FL, 1992.

[77] Tan J.G. and Wang H.L. Mean-square stability of the EulercMaruyama method for stochastic differential delay equations with jumps. *International Journal of Computer Mathematics*, 88(2):421–429, 2011.

[78] Hurn A.S. Jeisman J.I. and Lindsay K.A. Seeing the wood for the tree: A critical evaluation of methods to estimate the parameters of stochastic differential equations. *Journal of Financial Economics*, 5:390–455, 2007.

[79] Nielsen J.N. and Madsen H. Applying the EKF to stochastic differential equations with level effects. *Automatica*, 37:107–112, 2001.

[80] Marron J.S. and Wand M.P. Exact mean integrated squared error. *The Annals of Statistics*, 20:712–736, 1992.

[81] Barkoulas J.T. and Baum C.F. Long-term dependence in stock returns. *Economics Letters*, 53(3):253–259, 1996.

[82] Beck J.V. and Arnold K.L. *Parameter estimation in engineering and science*. Wiley, New York, 1977.

[83] Hurn A.S. Lindsay K.A. and Martin V.L. On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations. *Journal of Time Series Analysis*, 24:45–63, 2003.

[84] Nowman K.B. Gaussian estimation of single-factor continuous time models of the term structure of interest rate. *Journal of Finance*, 52(4):1695–1706, 1997.

[85] Caponetto R. Dongola G. Fortuna L. and Petras I. *Fractional order systems: Modeling and control applications*. World Scientific, River Edge, NJ, USA, 2010.

[86] Decreusefond L. and Ustunel A.S. Stochastic analysis of the fractional Brownian motion. *Potential Analysis*, 10:177–214, 1999.

[87] Ljung L. and Soderstroim T. *Theory and practice of recursive identification.* MIT Press, Cambridge, 1983.

[88] Wen L. Lakshmikantham V. and Zhang B.G. *Theory of differential equations with unbounded delay (Mathematics and its Applications).* Kluwer Academic Publishers Group, Dordrecht, 1994.

[89] Rogers L.C.G. Arbitrage with fractional Brownian motion. *Mathematical Finance*, 7:95–105, 1997.

[90] Hausen L.P. Large sample properties of generalized method of moments. *Econometrica*, 50:1029–1054, 1982.

[91] Bęben M. and O rł owski A. Correlations in financial time series: established versus emerging markets. *The European Physical Journal B*, 20(4):527–530, 2001.

[92] Ogaki M. *Generalized method of moments: Econometric applications.* Elsevier, 1993.

[93] Soderstrom T. Fan H. Mossberg M. and Carlsson B. A bias-compensation scheme for estimating continuous time AR process parameters. *In: SYSID' 97-11th IFAC Symposium on System Identification*, 1997.

[94] MATLAB. Particle swarm toolbox. *http://www.mathworks.com/matlabcentral/fileexchange/25986-another-particle-swarm-toolbox.*

[95] Brandt M.W. and SantaClara P. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics*, 63:161–210, 2002.

[96] Elerian O. A note on the existence of a closed form conditional transition density for the Milstein scheme. *Working Paper, Oxford University*, 1998.

[97] Barndorff-Nielsen O.E. and Shephard N. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B*, 63:167–241, 2001.

[98] Hassan R. Cohanim B. DeWeck O.L. and Venter G. A comparison of particle swarm optimization and the genetic algorithm. *Proceeedings of the First AIAA Multidisciplinary Design Optimization Specialist Conference*, pages 18–21, 2005.

[99] Cheridito P. Mixed fractional Brownian motion. *Bernoulli*, 7:913–934, 2001.

[100] Gray P. Bayesian estimation of short-rate models. *Australian Journal of Management*, 30:1–22, 2005.

[101] Hu P. and Huang C.M. Stability of stochastic θ methods for stochastic delay integro-differential equations. *International Journal of Computer Mathematics*, 88(7):1417–1429, 2011.

[102] Soderstrom T. Stoica P. and Friedlander B. An indirect prediction error method for system identification. *Automatica*, 27(1):183–188, 1991.

[103] Young P. Parameter estimation for continuous-time models - a survey. *Automatica*, 17:23–39, 1981.

[104] Zanzotto P.A. On solutions of one-dimensional stochastic differential equations driven by stable Levy motion. *Stochastic Processes and their Applications*, 68(2):209–228, 1997.

[105] Zanzotto P.A. On stochastic differential equations driven by a Cauchy process and other stable Levy motions. *Annals of Probability*, 30(2):802–825, 2002.

[106] Caraballo T. Kloeden P.E. and Real J. Discretization of asymptotically stable stationary solutions of delay differential equations with a random stationary delay. *Dynamics and Differential Equations*, 18(4):863–880, 2006.

[107] Kloeden P.E. and Platen E. *Numerical solution of stochastic differential equations*. Springer, Berlin, Germany, 1999.

[108] Burrage K. Burrage P.M. and Tian T. Numerical methods for strong solutions of stochastic differential equations: An overview. *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 460:373–402, 2004.

[109] Maybeck P.S. *Stochastic models, estimation and control*. Academic Press, London, 1982.

[110] Chumacero R. Finite sample properties of the efficient method of moments. *Studies in Nonlinear Dynamics and Econometrics*, 2(2):35–51, 1997.

[111] Cont R. and Tankov P. *Financial modelling with jump processes (Financial Mathematics Series)*. Chapman & Hall/CRC, London, UK, 2004.

[112] Li R. and Chang Z. Convergence of numerical solution to stochastic differential delay equations with Poisson jump and Markovian switching. *Applied Mathematics and Computation*, 184(2):451–463, 2007.

[113] Mikulevicius R. and Platen E. Time discrete Taylor approximations for Ito processes with jump component. *Mathematische Nachrichten*, 138:93–104, 1988.

[114] Rogers S. Khanin R. and Girolami M. Bayesian model-based inference of transcription factor activity. *BMC Bioinformatics*, 8(S2), 2007.

[115] Eberhart R.C. and Kennedy J. A new optimizer using particles swarm theory. *Proceedings of Sixth International Symposium on Micro Machine and Human Science*, pages 39–43, 1995.

[116] Merton R.C. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 2:125–144, 1976.

[117] Driver R.D. *Ordinary and delay differential equations (Applied Mathematical Sciences 20)*. Springer-Verlag, New York, 1977.

[118] Kalman R.E. and Bucy R.S. New results in linear filtering and prediction problems. *Journal of Basic Engineering, Transactions ASMA, Series D*, 83:95–108, 1961.

[119] Kadiev R.I. Stability of solutions of stochastic differential equations with random delays. *Differential Equations*, 40(2):261–264, 287, 2004.

[120] Kadiev R.I. and Ponosov A.V. Stability of solutions of linear impulsive systems of Ito differential equations with aftereffect. *Differential Equations*, 43(7):879–885, 1004, 2007.

[121] Das S. *Functional fractional calculus for system identification and controls*. Springer, Berlin, Germany, 2008.

[122] Maheswaran S. Predictable short-term variation in asset prices: Theory and evidence. *working paper, Carlson School of Management, University of Minnesota*, 1990.

[123] Momani S. and Odibat Z. Numerical comparison of methods for solving linear differential equations of fractional order. *Chaos, Solitons and Fractals*, 31(5):1248–1255, 2007.

[124] Panda S. and Padhy N.P. Comparison of particle swarm optimization and genetic algorithm for TACTS-based controller design. *Applied Soft Computing*, 8:1418–1427, 2008.

[125] Rostek S. *Option pricing in fractional Brownian markets*. Springer Berlin Heidelberg, Berlin, Germany, 2009.

[126] Sagara S. and Zhao Z.Y. Numerical integration approach to on-line identification of continuous-time systems. *Automatica*, 26(1):63–74, 1990.

[127] Mohammed S.E.A. *Stochastic functional differential equations*. Pitman Advanced Publishing Program, 1984.

[128] Mohammed S.E.A. Lyapunov exponents and stochastic flows of linear and affine hereditary systems. *Diffusion processes and related problems in analysis, Vol II*, pages 141–169, 1990.

[129] Mohammed S.E.A. and Scheutzow M.K.R. Lyapunov exponents of linear stochastic functional differential equations driven by semimartingales I: The multiplicative ergodic theory. *Annales Henri Poincare*, 32:69–105, 1996.

[130] Mohammed S.E.A. and Scheutzow M.K.R. Lyapunov exponents of linear stochastic functional differential equations II: Examples and case studies. *Annals of Probability*, 25:1210–1240, 1997.

[131] Mohammed S.E.A. and Scheutzow M.L.R. Lyapunov exponents and stationary solutions for affine stochastic delay equations. *Stochastics and Stochastics Reports*, 29(2):259–283, 1990.

[132] LeRoy S.F. Efficient capital markets and martingales. *Journal of Economic Literature*, 27(4):1583–1621, 1989.

[133] Lin S.J. Stochastic analysis of fractional Brownian motion, fractional noises and applications. *SIAM Review*, 10:422–437, 1995.

[134] Diekmann O. Van Gils S.A. Verduyn Lunel S.M. and Walther H.O. *Delay equations: functional, complex, and nonlinear analysis (Applied Mathematical Sciences)*. Springer, New York, 1995.

[135] Sethi S.P. and Lehoczky J.P. A comparison of the Ito and Stratonovich formulations of problems in finance. *Journal of Economic Dynamics and Control*, 3:343–356, 1981.

[136] Bohlin T. and Graebe S.F. Issues in nonlinear stochastic grey-box identification. *International Journal of Adaptive Control and Signal Processing*, 9(6):465–490, 1995.

[137] Soderstrom T. and Stoica P. *System identification*. Prentice Hall, London, 1989.

[138] Tian T. Estimation of kinetic rates of MAP kinase activation from experimental data. *Proceedings of IJCBS, IEEE Press*, pages 457–462, 2009.

[139] Tian T. and Burrage K. Implicit taylor methods for stiff stochastic differential equations. *Applied Numerical Mathematics*, 38:167–185, 2001.

[140] Tian T. and Ge X. Calibration of stochastic differential equation models using implicit numerical methods and particle swarm optimization. *Proceedings of 2012 International Conference on Modelling, Identification and Control*, pages 1049–1054, 2012.

[141] Andersen T.G. and Lund J. Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics*, 77:343–378, 1997.

[142] Xu S. Gao J. Tian T. and Burrage K. Simulated maximum likelihood method for estimating kinetic rates in gene expression. *Bioinformatics*, 23:84–91, 2007.

[143] Crauel H. Doan T.S. and Siegmund S. Difference equations with random delay. *Journal of Difference Equations and Applications*, 15(7):627–647, 2009.

[144] Kuchler U. and Platen E. Strong discrete time approximation of stochastic differential equations with time delay. *Mathematics and Computers in Simulation*, 54:189–205, 2000.

[145] Kuchler U. and Platen E. Weak discrete time approximation of stochastic differential equations with time delay. *Mathematics and Computers in Simulation*, 59:497–507, 2002.

[146] Bally V. and Talay D. The law of the Euler scheme for stochastic differential equations I: Convergence rate of the distribution function. *Probability Theory and Related Fields*, 104:43–60, 1996.

[147] Kolmanovskii V. and Myshkis A. *Applied theory of fundamental differential equations*. Kluwer, Netherlands, 1992.

[148] Barnett V.D. Evaluation of the maximum-likehood estimator where the likelihood equation has multiple roots. *Biometrika*, 53:151–165, 1966.

[149] Chen W. Nonlinear dynamics and chaos in a fractional-order financial system. *Chaos, Solitons and Fractals*, 36(5):1305–1314, 2008.

[150] Wang X. Qiu W. and Ren F. Option pricing of fractional version of the Black-Scholes model with Hurst exponent H being in (1/3, 1/2). *Chaos, Solitons and Fractals*, 12(3):599–608, 2001.

[151] Hoh Walter. The martingale problem for a class of pseudo-differential operators. *Annals of Mathematics*, 300(1):121–147, 1994.

[152] Wang W.Q. and Chen Y.P. Mean-square stablity of semi-implicit Euler method for nonlinear neutral stochastic delay differential equations. *Applied Numerical Mathematics*, 61(5):696–701, 2011.

[153] Press W.H. Teukolsky S.A. Vetterling W.T. and Flannery B.P. *Numerical recipes in C (2nd ed.)*. CambridgeUniversityPress, London, 1992.

[154] Mao X. *Exponential stability of stochastic differential equations*. Marcel Dekker, New York, 1994.

[155] Mao X. *Stochastic differential equations and applications*. Horwood, Chichester, United Kingdom, 1997.

[156] Ren F. Wang X. and Liang J. A proof for frenchs empirical formula on option pricing. *Chaos, Solitons and Fractals*, 12(13):2441–2453, 2001.

[157] Duncan T.E. Hu Y. and Pasik-Duncan B. Stochastic calculus for fractional Brownian motion I. theory. *SIAM Journal on Control and Optimization*, 38:582–612, 2000.

[158] Hu Y. and Oksendal B. Fractional white noise calculus and application to finance. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 6(1), 2003.

[159] Hu Y. and Ø ksendal B. Fractional white noise calculus and applications to finance. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 6(1):1–32, 2003.

[160] Wen F. Gong X. Chao Y. and Chen X. The effects of prior outcomes on risky choice: Evidence from the stock market. *Mathematical Problems in Engineering*, 2014:Article ID 272518, 2014.

[161] Jacob N. Wang Y.T. and Yuan C.G. Numerical solutions of stochastic differential delay equations with jumps. *Stochastic Analysis and Applications*, 27(4):825–853, 2009.

[162] Wen F. He Z. and Chen X. Investors risk preference characteristics and conditional skewness. *Mathematical Problems in Engineering*, 2014:Article ID 814965, 2014.

[163] Odibat Z.M. Computational algorithms for computing the fractional derivatives of functions. *Mathematics and Computers in Simulation*, 79(7):2013–2020, 2009.