Springer Finance

Jianwei Zhu

Applications of Fourier Transform to Smile Modeling

Theory and Implementation

Second Edition

Springer

Springer Finance

Springer Finance

Springer Finance is a programme of books addressing students, academics and practitioners working on increasingly technical approaches to the analysis of financial markets. It aims to cover a variety of topics, not only mathematical finance but foreign exchanges, term structure, risk management, portfolio theory, equity derivatives, and financial economics.

Ammann M., Credit Risk Valuation: Methods, Models, and Application (2001)
Back K., A Course in Derivative Securities: Introduction to Theory and Computation (2005)
Barucci E., Financial Markets Theory. Equilibrium, Efficiency and Information (2003)
Bielecki T.R. and Rutkowski M., Credit Risk: Modeling, Valuation and Hedging (2002)
Bingham N.H. and Kiesel R., Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives (1998, 2nd ed. 2004)
Brigo D. and Mercurio F., Interest Rate Models: Theory and Practice (2001, 2nd ed. 2006)
Buff R., Uncertain Volatility Models – Theory and Application (2002)
Carmona R.A. and Tehranchi M.R., Interest Rate Models: An Infinite Dimensional Stochastic Analysis Perspective (2006)
Dana R.-A. and Jeanblanc M., Financial Markets in Continuous Time (2003)
Deboeck G. and Kohonen T. (Editors), Visual Explorations in Finance with Self-Organizing Maps (1998)
Delbaen F. and Schachermayer W., The Mathematics of Arbitrage (2005)
Elliott R.J. and Kopp P.E., Mathematics of Financial Markets (1999, 2nd ed. 2005)
Fengler M.R., Semiparametric Modeling of Implied Volatility (2005)
Filipović D., Term-Structure Models (2009)
Fusai G. and Roncoroni A., Implementing Models in Quantitative Finance: Methods and Cases (2008)
Jeanblanc M., Yor M. and Chesney M., Mathematical Methods for Financial Markets (2009)
Geman H., Madan D., Pliska S.R. and Vorst T. (Editors), Mathematical Finance – Bachelier Congress 2000 (2001)
Gundlach M. and Lehrbass F. (Editors), CreditRisk^{+} in the Banking Industry (2004)
Jondeau E., Financial Modeling Under Non-Gaussian Distributions (2007)
Kabanov Y.A. and Safarian M., Markets with Transaction Costs (2010)
Kellerhals B.P., Asset Pricing (2004)
Külpmann M., Irrational Exuberance Reconsidered (2004)
Kwok Y.-K., Mathematical Models of Financial Derivatives (1998, 2nd ed. 2008)
Malliavin P. and Thalmaier A., Stochastic Calculus of Variations in Mathematical Finance (2005)
Meucci A., Risk and Asset Allocation (2005, corr. 2nd printing 2007)
Pelsser A., Efficient Methods for Valuing Interest Rate Derivatives (2000)
Prigent J.-L., Weak Convergence of Financial Markets (2003)
Schmid B., Credit Risk Pricing Models (2004)
Shreve S.E., Stochastic Calculus for Finance I (2004)
Shreve S.E., Stochastic Calculus for Finance II (2004)
Yor M., Exponential Functionals of Brownian Motion and Related Processes (2001)
Zagst R., Interest-Rate Management (2002)
Zhu Y.-L., Wu X., Chern I.-L., Derivative Securities and Difference Methods (2004)
Ziegler A., Incomplete Information and Heterogeneous Beliefs in Continuous-time Finance (2003)
Ziegler A., A Game Theory Analysis of Options (2004)

Jianwei Zhu

Applications of Fourier Transform to Smile Modeling

Theory and Implementation

Springer

Dr. Jianwei Zhu
Lucht Probst Associates
Große Gallusstr. 9
60311 Frankfurt/Main
Germany
jianwei.zhu@l-p-a.com

ISBN 978-3-642-01807-7 e-ISBN 978-3-642-01808-4
DOI 10.1007/978-3-642-01808-4
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009937646

Mathematics Subject Classification (2000): 42A99, 60E10, 60G05, 60H05, 60J65, 60J75, 91B28, 91B70
JEL Classification: C02, C46, C63, E43, G12, G13, G24

Originally published as volume 493 in the series Lecture Notes in Economics and Mathematical Systems

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my wife Shiquan

Preface

This book addresses the applications of Fourier transform to smile modeling. Smile effect is used generically by financial engineers and risk managers to refer to the inconsistences of quoted implied volatilities in financial markets, or more mathematically, to the leptokurtic distributions of financial assets and indices. Therefore, a sound modeling of smile effect is the central challenge in quantitative finance. Since more than one decade, Fourier transform has triggered a technical revolution in option pricing theory. Almost all new developed option pricing models, especially in connection with stochastic volatility and random jump, have extensively applied Fourier transform and the corresponding inverse transform to express option pricing formulas. The large accommodation of the Fourier transform allows for a very convenient modeling with a general class of stochastic processes and distributions. This book is then intended to present a comprehensive treatment of the Fourier transform in the option valuation, covering the most stochastic factors such as stochastic volatilities and interest rates, Poisson and Lévy jumps, including some asset classes such as equity, FX and interest rates, and providing numerical examples and prototype programming codes. I hope that readers will benefit from this book not only by gaining an overview of the advanced theory and the vast large literature on these topics, but also by gaining a first-hand feedback from the practice on the applications and implementations of the theory.

This book has grown out of my early research manuscript “Modular Pricing of Options” that was published in 2000 in the series “Lecture Notes in Economics and Mathematical Systems”. From the point of view of a practitioner, the modular pricing approach is the one of the best applications of Fourier transform in the option valuation. By the modular pricing approach, each stochastic factor may be regarded as a module in a generalized pricing framework with Fourier transform. Following this building-block way for pricing, I address in this book stochastic volatilities, stochastic interest rates, Poisson jumps and Lévy jumps successively, and at the same time, expound the advanced option pricing theory step by step.

Additionally, I give detailed discussions on the numerical issues of stochastic volatility models, for instance, how to perform an efficient numerical integration of the inverse Fourier transform in a trading environment. Since stochastic volatility models are the representative examples of the applications of Fourier transform, almost all numerical methods for stochastic volatilities can be applied to other stochastic factors, for instance, Lévy jumps. In this book, Lévy jumps are referred to as the jumps with infinite activity, and are also the driving factor of the so-called Lévy pricing models in financial literature. This book delivers an illustrative and compact introduction to Lévy jumps. Exotic options and Libor Market Model (LMM) with stochastic factors are two other focuses of my book. Valuing exotic equity/FX option with stochastic factors, especially with stochastic volatility analytically, is generally not feasible, and therefore, requires some special treatments and modifications. The stochastic volatility LMMs are proposed to capture the smile effects in cap and swaption markets. Some of these extended LMMs use the Fourier transform extensively to arrive at some tractable solutions.

I organize this book in a self-contented way, and present the basic concepts of option pricing theory and the advanced techniques of asset modeling, especially with respect to Fourier transform and characteristic functions. The prerequisites for this book are some basic knowledge of stochastic theory and financial products, although the main results and theorems such as the Feynman-Kac theorem and the Girsanov theorem are briefly reviewed in the first chapter. This book is structured in 11 chapters. Chapter 1 summarizes the Black-Scholes model and explains the related smile effect of implied volatilities. Capturing smile effects is the main motivation of this book. At the same time, the key definitions and concepts in quantitative finance such as Itô process, the Feynman-Kac theorem and change of measure, are introduced. Chapter 2 addresses in details the fundamentals of a generalized pricing framework where characteristic functions play a key role in unifying all stochastic factors. Additionally, the Fourier transform as a pricing technique is expounded from the different aspects, both economically and mathematically. Afterwards, I begin to deal with various stochastic factors separately, namely, stochastic volatility and interest rate, Poisson jump and Lévy jump. Chapter 3 considers three analytical tractable stochastic volatility models: the Heston model (1993), the Schöbel and Zhu model (1999) and a double square root model, where the stochastic volatility or variance follows a mean-reverting square root process, a mean-reverting Ornstein-Uhlenbeck process and a mean-reverting double square root process, respectively. As an example, I derive the Heston model in details with two approaches: the PDE approach and the expectation approach, and show that the latter approach is much more powerful and can be applied to calculate the characteristic functions of broader stochastic processes. Next in Chapter 4, I discuss a more important problem for the practical applications of stochastic volatility models, namely, the numerical implementation of these models, for instance, the efficient numerical integration of inverse Fourier transform, the complex logarithms, the calculation of Greeks as well as quick calibration. All numerical methods and techniques for stochastic volatility are also relevant and valid for other stochastic factors. Chapter 5 focuses on the simulation of stochastic volatility models. The simulation of square root process, particularly the Heston model, has been a large challenge in computational finance. Chapter 5 presents some recently developed efficient simulation methods for the Heston model. Numerical examples show that the QE scheme of Andersen (2007) and the transformed volatility scheme of Zhu (2008) are two most efficient methods for simulating mean-reverting square root process. Additionally, we will discuss how to simulate extremes in a diffusive process and a multi-asset model with stochastic volatilities. Chapter 6 explains how to incorporate stochastic interest rates into an option pricing model in a short rate setup. Three representative short rate models of CIR (1985), Vasicek (1977) and Longstaff (1989) are examined where two cases, zero correlation and non-zero correlation, for the correlation between the stock returns and the interest rates are expounded.

Chapter 7 and Chapter 8 are dedicated to jump modeling. The Fourier transform is a very elegant and efficient technique when incorporating discontinuous jump events in an asset process. Chapter 7 deals with the traditional jump models where the jump mechanism is characterized by a compound Poisson process. A compound Poisson process is a mixed process with a jump counting described by a pure Poisson process and a random jump size described by a distribution. Dependently on the specification of jump sizes, we have different Poisson jump models, for instance, lognormal jumps, Pareto jumps and double exponential jumps. The affine jump-diffusion model of Duffie, Pan and Singleton (2000) is also reviewed briefly. Chapter 8 describes the Lévy jumps that are characterized by infinite activity. After an introduction to the Lévy process, two simple and popular Lévy jump models, the variance Gamma model and the normal inverse Gaussian model, are first discussed. Both of these processes belong to the class of stochastic clock models. Next I move to the so-called time-changed Lévy processes that are generated by applying a stochastic time change to the underlying Lévy processes, or equivalently, can be generated by a subordination. Two different cases of stochastic time-change, the zero correlation and the non-zero correlation between the stochastic time and the underlying Lévy process, are considered where the non-zero correlation is refereed to as the leverage effect, and can be dealt with conveniently by introducing a complex-valued leverage-risk neutral measure. Additionally, two particular Lévy models, the Barndorff-Nielsen and Shephard model (2001) and the alpha log-stable model are briefly examined.

After the stochastic volatility, the stochastic interest rate, the Poisson jump and the Lévy jumps have been discussed, Chapter 9 integrates all these stochastic factors into an unified pricing framework. I use two approaches for the integration, the modular approach and the time-change approach. While the modular approach has a simple, clear and transparent structure, and is also designed for practical implementations, the approach based on time-change is more technical and may embrace some nested factors. The variety of the resulting option valuation models by the modular pricing is immense. As numerical examples, the option prices of 108 different models are given to illustrate the large accommodation of the Fourier transform. Finally, I provide some criterions for choosing the "best" pricing models from the point of view of a practitioner.

Chapter 10 addresses the valuation of exotic options with stochastic factors, particularly with stochastic volatility. Generally, there are no closed-form pricing formulas for exotic options with stochastic volatility. However, for some special processes and model setups, I attempt to derive the tractable pricing solutions for some exotic options, for example, barrier options, correlation options and Asian options. In Chapter 11, I present five stochastic volatility Libor market models (LMM) that are applied to capture the smile effects in cap and swaption markets. All of these five models apply characteristic functions or moment-generating functions for pricing caps and swaptions to different extent. While characteristic functions do not find significant applications in the models of Andersen and Brotherton-Ratcliffe (2001), and Piterbarg (2003), the models of Zhang and Wu (2006), Zhu (2007) as well as Belomestny, Matthew and Schoenmakers (2007) have extensively and consequently used characteristic functions to arrive at the closed-form pricing formulas for caplets and swaptions.

This book is aimed at

- financial engineers and risk managers who are seeking for a simple and clearly structured book on the smile modeling with Fourier transform, and want to implement some pricing models with characteristic functions efficiently and robustly;
- graduate students who need a compact introduction to the option pricing theory with Fourier transform, and want to gain an overview of the vast literature on these advanced topics;
- researchers who want to know what is the most important for the smile modeling in practice, and want to get some feedbacks from practice on the applications of the option pricing models, particularly stochastic volatility models.

During the writing of this book and its early version, I have benefited from many discussions with my colleagues and large supports from friends. First, I would like to thank Professor Rainer Schöbel and Professor Gerd Ronning for their helpful comments and encouragement. During my research activities at the University of Tübingen, my mentor Professor Schöbel provided me with rich freedom in researching the interesting and exciting topics that are discussed in this book. The four years of my doctoral time have been a highly spiritual experience for me to enjoy the beauty and rigor of financial economics. In addition, I am grateful to Juri Hinz for some discussions and, to Professor Herbert Heyer for the lectures on stochastic calculus, both from the mathematical institute at the University of Tübingen.

It is my pleasure to express my thanks here to my colleagues at the university, Philipp Kellerhals, Stephan Heilig and Vera Klöckner, my colleagues Jügen Linde, Andreas Gaida and Reinhard Batin at Dresdner Bank, WestLB and Sal. Oppenheim, and my current colleagues Daniel Herzbach, Martin Kleine Kalvelage, Phillip Reinfeld and Geráld Wigger, as well as Professor Wolfgang Schmidt at HfB. I am grateful to Wolfram Boenkost, my current chef at LPA, for providing me with the challenging tasks to implement some sophisticated models and with the opportunities to do experiment with new ideas.

Finally, I would like to express my gratitude to my family for their unbroken and warm support. Especially, without the support of my wife, I would have never finished writing this book.

Frankfurt am Main, Germany

August 2009 *Jianwei Zhu*

Contents

1 Option Valuation and the Volatility Smile . . . 1
1.1 Stochastic Processes for Stocks . . . 1
1.1.1 Brownian Motion . . . 1
1.1.2 Stock Price as Geometric Brownian Motion . . . 3
1.1.3 Itô Process and Itô's Lemma . . . 4
1.2 The Black-Scholes Model . . . 5
1.2.1 Options and Dynamic Hedging . . . 5
1.2.2 Risk-Neutral Valuation . . . 7
1.2.3 Self-financing and No Arbitrage . . . 8
1.2.4 Equivalent Martingale Measures . . . 11
1.3 Volatility Quotations in Markets . . . 15
1.3.1 Implied Volatilities . . . 15
1.3.2 Market Quotations . . . 16
1.3.3 Special Case: FX Market . . . 17

2 Characteristic Functions in Option Pricing . . . 21
2.1 Constructing Characteristic Functions (CFs) . . . 22
2.1.1 A General Process for Stock Price . . . 22
2.1.2 Valuation of European-style Options via CFs . . . 23
2.1.3 Special Case: FX Options . . . 27
2.2 Understanding Characteristic Functions . . . 29
2.2.1 Properties of Characteristic Functions . . . 29
2.2.2 Economic Interpretation of CFs . . . 32
2.2.3 Examination of Existing Option Models . . . 35
2.2.4 Relationship between CF to PDE . . . 39
2.2.5 Advantages of CF and Modular Pricing . . . 42

3 Stochastic Volatility Models . . . 45
3.1 Introduction . . . 45
3.2 The Heston Model . . . 48
3.2.1 Model Setup and Properties . . . 48

3.2.2 PDE Approach to Pricing Formula 50
3.2.3 Expectation Approach to Pricing Formula 52
3.2.4 Various Representations of CFs 54
3.3 The Schöbel-Zhu Model 55
3.3.1 Model Setup and Properties 55
3.3.2 Derivation of CFs 58
3.3.3 Numerical Examples 60
3.4 Double Square Root Model 63
3.4.1 Model Setup and Properties 63
3.4.2 Numerical Examples 69
3.5 Other Stochastic Volatility Models 70
3.6 Appendices 73

4 Numerical Issues of Stochastic Volatility Models 77
4.1 Alternative Pricing Formulas with CFs 78
4.1.1 The Formula á la Black-Scholes 78
4.1.2 The Carr and Madan Formula 78
4.1.3 The Attari Formula 80
4.2 Risk Sensitivities 80
4.2.1 Delta and Gamma 81
4.2.2 Various Vegas 82
4.2.3 Curvature and Slope 84
4.2.4 Volga and Vanna 84
4.3 Direct Integration (DI) 86
4.3.1 The Gaussian Integration 86
4.3.2 Multi-Domain Integration 87
4.3.3 Strike Vector Computation 88
4.4 Fast Fourier Transform (FFT) 89
4.4.1 Algorithms of FFT 89
4.4.2 Restrictions 91
4.5 Direct Integration vs. FFT 92
4.5.1 Computation Speed 93
4.5.2 Computation Accuracy 94
4.5.3 Matching Market Data 94
4.5.4 Calculation of Greeks 95
4.5.5 Implementation 95
4.6 Logarithm of Complex Number 99
4.6.1 Definition 99
4.6.2 Three Algorithms Dealing with Branch Cut 101
4.6.3 When Main Argument Is Appropriate 103
4.7 Calibration to Market Data 104
4.7.1 General Procedure 104
4.7.2 Fixing Velocity Parameter 105
4.7.3 Fixing Spot Volatility 106
4.8 Markovian Projection 108

5 Simulating Stochastic Volatility Models . . . 113
5.1 Simulation Scheme . . . 114
5.1.1 Discretization . . . 114
5.1.2 Moment-Matching . . . 115
5.2 Problems in the Heston Model . . . 116
5.2.1 Negative Values in Paths . . . 116
5.2.2 Log-normal Scheme . . . 117
5.2.3 Transformed Volatility Scheme . . . 118
5.2.4 QE Scheme . . . 120
5.2.5 The Broadie-Kaya Scheme . . . 122
5.2.6 Some Other Schemes . . . 124
5.3 Simulation Examples . . . 125
5.4 Maximum and Minimum . . . 126
5.5 Multi-Asset Model . . . 130

6 Stochastic Interest Models . . . 135
6.1 Introduction . . . 135
6.2 The Cox-Ingosoll-Ross Model . . . 138
6.2.1 The Zero-Correlation Case . . . 138
6.2.2 The Correlation Case . . . 140
6.3 The Vasicek Model . . . 142
6.4 The Longstaff Model . . . 144
6.4.1 The Zero-Correlation Case . . . 145
6.4.2 The Correlation Case . . . 146
6.5 Correlations with Stock Returns: SI versus SV . . . 148

7 Poisson Jumps . . . 153
7.1 Introduction . . . 153
7.2 Simple Jumps . . . 158
7.3 Lognormal Jumps . . . 160
7.4 Pareto Jumps . . . 163
7.5 The Kou Model: An Equivalence to Pareto Jumps . . . 165
7.6 Affine Jump-Diffusions . . . 168

8 Lévy Jumps . . . 173
8.1 Introduction . . . 174
8.2 Stochastic Clock Models . . . 177
8.2.1 Variance-Gamma Model . . . 179
8.2.2 Normal Inverse Gaussian Model . . . 181
8.3 Time-Changed Lévy Process . . . 183
8.3.1 Uncorrelated Time-Change . . . 184
8.3.2 Correlated Time-Change . . . 188
8.4 The Barndorff-Nielsen and Shephard Model . . . 193
8.5 Alpha Log-Stable Model . . . 194
8.6 Empirical Performance of Various Lévy Processes . . . 196

8.7 Monte-Carlo Simulation . . . 198
8.7.1 Generating Random Variables . . . 198
8.7.2 Simulation of Lévy Process . . . 201

9 Integrating Various Stochastic Factors . . . 203
9.1 Stochastic Factors as Modules . . . 203
9.2 Integration Approaches . . . 206
9.2.1 Modular Approach . . . 206
9.2.2 Time-Change Approach . . . 212
9.3 Pricing Kernels for Options and Bonds . . . 217
9.4 Criterions for Model Choice . . . 218

10 Exotic Options with Stochastic Volatilities . . . 223
10.1 Forward-Starting Options . . . 224
10.2 Barrier Options . . . 226
10.2.1 Introduction . . . 226
10.2.2 Two Special Cases . . . 228
10.2.3 Numerical Examples . . . 233
10.3 Lookback Options . . . 235
10.3.1 Introduction . . . 235
10.3.2 Pricing Formulas with Stochastic Factors . . . 238
10.4 Asian Options . . . 244
10.4.1 Introduction . . . 244
10.4.2 The Black-Scholes World . . . 245
10.4.3 Asian Options in a Stochastic World . . . 249
10.4.4 Approximations for Arithmetic Average Asian Options . . . 252
10.4.5 A Model for Asian Interest Rate Options . . . 254
10.5 Correlation Options . . . 257
10.5.1 Introduction . . . 257
10.5.2 Exchange Options . . . 260
10.5.3 Quotient Options . . . 263
10.5.4 Product Options . . . 264
10.6 Other Exotic Options . . . 266
10.7 Appendices . . . 267

11 Libor Market Model with Stochastic Volatilities . . . 273
11.1 Introduction . . . 273
11.2 Standard Libor Market Model . . . 275
11.2.1 Model Setup . . . 275
11.2.2 Term Structure and Smile of Volatility . . . 280
11.3 Swap Market Model . . . 282
11.3.1 Model Setup . . . 282
11.3.2 Correlation Structure . . . 286
11.3.3 Convexity Adjustments for CMS . . . 291
11.4 Incorporating Stochastic Volatility . . . 294

11.4.1 The Andersen and Brotherton-Ratcliffe Model 295
11.4.2 The Piterbarg Model 298
11.4.3 The Wu and Zhang Model 302
11.4.4 The Zhu Model 305
11.4.5 The Belomestny, Matthew and Schoenmakers Model 312
11.5 Conclusive Remarks 316

References 319

Index 327

Chapter 1
Option Valuation and the Volatility Smile

In this chapter, we briefly present the basic concepts of option pricing theory. The readers who are familiar with these topics, can skip this chapter and begin with the next chapter directly. A Brownian motion is an elemental building-block in modeling the dynamics of stock returns, and correspondingly the geometric Brownian motion as an exponential function of Brownian motion is the simplest and most popular process for stock prices, on which the Black-Scholes model is based. Dynamic hedging in the Black-Scholes model is a self-financing trading strategy that ensures no arbitrage, and allows us to derive the Black-Scholes equation and formula. It is shown that no arbitrage implied in the dynamic hedging results in a risk-neutral process with which all financial derivatives may be valued. From the point of view of probability measures, a risk-neutral process can be regarded as a process derived via the change of the historical measure to a measure using the money market account as numeraire. Finally, we can verify that an equivalent martingale measure under certain conditions in turn implies no arbitrage. Therefore, no arbitrage, risk-neutral valuation and equivalent martingale measure are the key concepts in option pricing theory, and are essentially equivalent to each other, but from different points of view. Next, we focus on the practical challenge of option pricing: inconsistent implied volatilities, or the volatility smile, a challenge that arises from the weakness of the Black-Scholes model, and that we will tackle in this book. Implied volatilities display different smile patterns, and are extensively used by markets for quotations of option prices.

1.1 Stochastic Processes for Stocks

1.1.1 Brownian Motion

In financial markets there are a number of prices, rates and indices that display strong stochastic behavior in the course of time. Nowadays, a Brownian motion is

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_1,

a basic building-block in modeling financial variables, and extensively used to describe most periodic changes of random prices, for example, daily returns of stock prices. Brownian motion was first introduced by the Scottish botanist R. Brown in 1827 to document random movements of small particles in fluid. The French mathematician L. Bachelier (1900) applied Brownian motion in his lately honored PhD thesis "The Theory of Speculation" to analyze the price behavior of stocks and options. But first in 1905, A. Einstein delivered a formal mathematical analysis of Brownian motion and derived a corresponding partial differential equation. Afterwards Brownian motion became an important tool in physics. In honor of N. Wiener who made a great contribution to studying Brownian motion from the perspective of stochastic theory, a standard Brownian motion is also called a Wiener process.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, Q)$[1] be a complete probability space equipped with a probability measure Q. $\mathscr{F}_t$ is σ-algebra at time t and may be interpreted as a collection of all information up to time t.[2] A standard Brownian motion is defined as follows.

Definition 1.1.1. *Brownian motion: A standard Brownian motion* $\{W(t), t \geqslant 0\}$ *is a stochastic process satisfying the following conditions,*

1. *for* $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n < \infty$, $W(t_1), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1})$ *are stochastically independent;*
2. *for every* $t > s \geqslant 0$, $W(t) - W(s)$ *is normally distributed with a mean of zero and a variance of* $t - s$,
$$W(t) - W(s) \sim N(0, t-s);$$
3. *for each fixed* $\omega \in \Omega, W(\omega, t)$ *is continuous in* t;
4. $W(0) = 0$ *almost surely.*

Condition 1 says that non-overlapping increments of $W(t)$ are uncorrelated. Condition 2 states that any increment $W(t) - W(s)$ is distributed according to a Gaussian law with a mean of zero and a variance of $t - s$. As a Brownian motion $W(t)$ is defined on $\mathscr{F}_t$, the probability measure Q is then a Gaussian law. In other words, the dispersion of a standard Brownian motion is proportional to time length. A further implication of condition 2 is that a standard Brownian motion is homogeneous in the sense that the distribution of any increment depends only on time length $t - s$ and not on time point s. Condition 3 explains that a standard Brownian motion has continuous paths over time. The last condition just requires that a standard Brownian motion starts at zero almost surely, and implies the expected value of any $W(t)$ is equal to zero. Summing up, a standard Brownian motion belongs to a class of the processes with stationary independent increments.

Since a standard Brownian motion is normally distributed and defined on the entire real line, it may not be applied appropriately to model price levels that are usually strictly positive. A realistic concept for price movements is a geometric Brownian motion, an exponential function of a Brownian motion.

---

[1] For more detailed descriptions on probability space, σ-algebra and filtration, please refer to other mathematical textbooks, for example, Øksendal (2003).

[2] Throughout this book, for notation simplicity, we will not always express a conditional expectation by explicitly adding $\mathscr{F}_t$.

1.1.2 Stock Price as Geometric Brownian Motion

If instantaneous returns of an asset $S(t)$ are assumed to follow a fixed quantity plus a Brownian motion which describes the risky and unpredictable part of $S(t)$, the asset $S(t)$ itself follows a geometric Brownian motion,

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t), \tag{1.1}$$

where the parameters μ and σ are constant. Due to the randomness of $W(t)$, $S(t)$ is no longer a deterministic quantity, but governed by some stochastic dynamics. By intuition, if we are neglecting the term $\sigma dW(t)$ in (1.1), we obtain the following equation

$$\frac{dS(t)}{S(t)} = \mu dt, \quad \frac{dS(t)}{dt} = S(t)\mu, \tag{1.2}$$

which is an ordinary differential equation (ODE) and admits a deterministic solution

$$S(t) = S(0)e^{\mu t}.$$

Here the parameter μ can be interpreted as the instantaneous return of $S(t)$. By adding the random term $\sigma dW(t)$ to the ODE in (1.2), we obtain a stochastic differential equation (SDE) as in (1.1). There is no longer a deterministic solution to (1.1). However, in the sense of expectation we have

$$\mathbf{E}[S(t)] = S(0)e^{\mu t}.$$

This indicates that if an asset $S(t)$ is governed by a geometric Brownian motion, it has an expected return of μ which is also referred to as a drift term. Since the parameter σ scales $W(t)$ and then characterizes how large the instantaneous stock returns deviate from its drift, we call this parameter volatility. Additionally, specifying the stock price to be a geometric Brownian motion implies the following facts:

1. The increments $dS(t_1)$ and $dS(t_2), t_1 \neq t_2$, are independent, i.e., uncorrelated, since the increments $dW(t_1)$ and $dW(t_2)$ are independent.
2. The randomness of the instantaneous stock returns is normally distributed and has the same volatility.
3. As a result of normal distribution, stock returns have zero skewness and no excess kurtosis, which contradicts the most empirical studies that document the negative skewness and the excess kurtosis of historical stock returns.
4. The stock prices are always positive.
5. The random paths of the stock prices are continuous and have no jumps.

These implications indicate that a geometric Brownian motion can simply capture the main characteristics of stock prices, but also imposes some restrictions to their model behavior. For example, the stock prices exhibit no jumps and have a constant volatility. More complicated models for stock prices, as shown later, will relax these implications of the geometric Brownian motion by adding more realistic components.

1.1.3 Itô Process and Itô's Lemma

Geometric Brownian motion may be extended to a general stochastic process,

$$\begin{aligned} dS(t) &= S(t)\mu dt + S(t)\sigma dW(t) \\ &= a(S(t),t)dt + b(S(t),t)dW(t), \end{aligned} \tag{1.3}$$

where $a(S(t),t)$ and $b(S(t),t)$ could be deterministic functions or stochastic processes. If the following two conditions are satisfied,

$$\mathbf{P}\left[\int_0^t |a(S(u),u)|du < \infty, t \geqslant 0\right] = 1 \tag{1.4}$$

and

$$\mathbf{P}\left[\int_0^t b^2(S(u),u)du < \infty, t \geqslant 0\right] = 1, \tag{1.5}$$

the process $S(t)$ is called an Itô process. $\mathbf{P}[\cdot] = 1$ means the condition is fulfilled almost surely. In fact, the above two conditions ensure that the solution of the stochastic differential equation of an Itô process exists. In other words, we have

$$S(t) = S_0 + \int_0^t a(S(u),u)du + \int_0^t b(S(u),u)dW(u) < \infty.$$

If $a(S(t),t)$ and $b(S(t),t)$ are not functions of time t, namely $a(S(t),t) = a(S(t))$ and $b(S(t),t) = b(S(t))$, the Itô process $S(t)$ is also a time-homogenous process. In this case, conditions (1.4) and (1.5) are simplified to the Lipschitz condition

$$|a(x) - a(y)| + |b(x) - b(y)| \leqslant D|x - y|, \qquad x, y, D \in \mathbb{R}.$$

Given an Itô process

$$dX(t) = a(X(t),t)dt + b(X(t),t)dW(t),$$

a natural question with respect to Itô process is, if $Y(t) = f(X(t),t)$ is a function of the Itô process $X(t)$ and time t, what form should the process $Y(t)$ take. In fact, most financial derivatives are essentially a function of underlying assets and time, therefore it is very important and useful to derive the stochastic process of a financial derivative based on the process of underlying asset. Itô's lemma gives us an instructive way to derive a stochastic process, and then is a very useful tool to analyze and evaluate options. The simple version of Itô's lemma is following.

Theorem 1.1.2. *Itô's Lemma: If $X(t)$ is an Itô process, then the function*[3] $Y(t) = f(X(t),t) \in C^{2\times 1}[\mathbb{R}\times[0,\infty)\to\mathbb{R}]$ *is an Itô process of the following form,*

$$\begin{aligned}
dY(t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X(t)}dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial X^2(t)}(dX(t))^2 \\
&= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X(t)}dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial X^2(t)}b^2(X(t),t)dt \\
&= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial X^2(t)}b^2(X(t),t)\right)dt + \frac{\partial f}{\partial X(t)}dX(t).
\end{aligned} \tag{1.6}$$

Applying Itô's lemma, it is easy to obtain the process of log returns $x(t) = \ln S(t)$,

$$dx(t) = \big(\mu - \frac{1}{2}\sigma^2\big)dt + \sigma dW(t)$$

with $x(0) = \ln S(0)$.

1.2 The Black-Scholes Model

1.2.1 Options and Dynamic Hedging

The payoff of a European call option at maturity date with strike K is given by

$$C(T) = \max[S(T) - K, 0],$$

which depends directly on the underling stock price $S(T)$. The strong dependence of a call price on its underlying stock price implies that an option should have a "fair price" to best represent the co-movements between option and underlying stock. It then raises a question, what is the exact definition of the fair price for a call? Black-Scholes's answer to this question is that a fair price must be an arbitrage-free price in the sense that a risk-free portfolio comprising of options and underlying stocks must reward a risk-free return.

Now we consider a portfolio including a call and its underlying. Denote the price of a call at time t as $C(t)$, $C(t)$ should be a function of $S(t)$ and t. In order to obtain the stochastic process of $C(t)$, we apply Itô's lemma and obtain the following SDE for $C(S(t),t)$,

$$dC(t) = \Big(\frac{\partial C}{\partial t} + \mu S\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}\Big)dt + \sigma S\frac{\partial C}{\partial S}dW(t). \tag{1.7}$$

[3] The notation $C^{2\times 1}$ stands for the class of all functions that are twice differentiable with respect to first argument (here X) and once differentiable with respect to second argument (here t).

We construct a portfolio as follows: a long position of stock in amount of $\frac{dC}{dS}$ and a short position of call in amount of 1. The cost of this portfolio is

$$G(t) = \frac{dC}{dS} S(t) - C(t).$$

In a short time interval $(t, t+dt)$, the change of the portfolio in value is given by

$$dG(t) = \frac{dC}{dS} dS(t) - dC(t). \tag{1.8}$$

Inserting the SDEs (1.1) and (1.7) into the above equation yields

$$dG(t) = -\Big(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\Big) dt, \tag{1.9}$$

where the random term $dW(t)$ disappears, and therefore the portfolio is not exposed to the instantaneous risk linked to $dW(t)$. According to the definition of a fair price, this portfolio should have a risk-free return r, namely

$$G(t) = G(0)e^{rt},$$

which corresponds to an ODE

$$dG(t) = rG(t)dt = r\Big(\frac{dC}{dS} S(t) - C(t)\Big) dt. \tag{1.10}$$

By comparing this equation with the equation (1.9), we obtain a PDE

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC = 0. \tag{1.11}$$

This PDE is called the Black-Scholes PDE with which a call price $C(t)$ can be solved. Note t is the calendar time. Denoting T as the time to maturity and $\frac{\partial C}{\partial t} = -\frac{\partial C}{\partial T}$, we rewrite the Black-Scholes PDE as follows:

$$rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC = \frac{\partial C}{\partial T}. \tag{1.12}$$

Imposing an initial condition for $T = 0$,

$$C(S_0, T = 0) = \max[S_0 - K, 0],$$

and solving the PDE (1.12) yields the celebrated Black-Scholes formula,

$$C_0 = S_0 N(d_1) - Ke^{-rT} N(d_2) \tag{1.13}$$

with

$$d_1 = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}},$$

$$d_2 = \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}.$$

The above idea that we construct a dynamic risk-less portfolio at every time is called the dynamic hedging.

The put option with payoff at maturity

$$P(T) = \max[K - S(T), 0]$$

can be valued with a similar formula

$$P_0 = Ke^{-rT}N(-d_2) - S_0N(-d_1).$$

1.2.2 Risk-Neutral Valuation

In the context of the dynamic hedging as shown in (1.9) and (1.10), the drift term μ of the stock process $S(t)$ does not play a role any more. In fact, the PDE (1.11) implies that the drift terms of call and stock processes are equal to the risk-free interest rate r. Particularly we have a new dynamically hedged stock process and a call process respectively,

$$dS(t) = S(t)rdt + \sigma S(t)dW(t), \tag{1.14}$$

$$dC(t) = C(t)rdt + \sigma S(t)\frac{\partial C}{\partial S}dW(t). \tag{1.15}$$

Taking the expectation on both sides of the above SDE for a call option yields an ODE

$$dy(t) = ry(t)dt, \quad y(t) = \mathbf{E}[C(t)]. \tag{1.16}$$

This is equivalent to

$$y(T) = \mathbf{E}[C(T)] = C_0e^{rT}.$$

In other words, we have a more simple expression for the call price,

$$C_0 = e^{-rT}\mathbf{E}[C(T)]. \tag{1.17}$$

This means that the fair price of a call is simply the discounted present value of its expectation at maturity if we undertake a dynamic hedging strategy. In fact, the above procedure is a simple example of the Feynman-Kac theorem that may build up an equivalence relationship between the PDE (1.11) and the expectation solution (1.17). Instead of solving a PDE, we can calculate the expected value $\mathbf{E}[C(T)]$ and discount it with a risk-free interest rate, then obtain the price of a call. This approach to pricing a call option illustrated in (1.17) is the so-called risk-neutral valuation.

Additionally, the stock price given in (1.14) is also referred to as a risk-neutral stock process where the drift term is replaced by the risk-free interest rate r.

Since the principle of the risk-neutral valuation is strongly linked to the Feynman-Kac theorem, here is the best place to state it in more details.

Theorem 1.2.1. *The Feynman-Kac theorem: Assume that*

$$dX(t) = a(X,t)dt + b(X,t)dW(t)$$

is an Itô process starting at time $t = 0$, and $f(x)$ is a continuous function, $f(x,t) \in C^{2\times 1}[\mathbb{R} \times [0,\infty) \to \mathbb{R}]$. If $y(x,T)$ satisfies the following PDE,

$$\frac{\partial y(x,T)}{\partial T} = a(x,T)\frac{\partial y(x,T)}{\partial x} + \frac{1}{2}b^2(x,T)\frac{\partial^2 y(x,T)}{\partial x^2} - q(x)y(x,T) \tag{1.18}$$

with an initial condition

$$y(x_0, T = 0) = f(x_0, 0),$$

where $q(x) \in C(\mathbb{R})$ is lower bounded, then

$$y(x_0,T) = \boldsymbol{E}\Big[\exp\Big(-\int_0^T q(X(t))dt\Big) f(X(T))\Big|\mathscr{F}_0\Big]. \tag{1.19}$$

This means that the value of $y(x,T)$ at time T is equal to the expected value of $f(X(T))$ discounted by q. Vise versa, if the expected value of (1.19) exists, then the PDE (1.18) holds.

The Feynman-Kac theorem builds up a bridge between the deterministic world represented by the PDE (1.18) and the stochastic world represented by the expected value (1.19). If we have difficulty in calculating the expected value (1.19), we can at least obtain it by numerically solving the PDE, as the Feynman-Kac theorem states. In most cases, however, we know via the Feynman-Kac theorem that simulating a stochastic process and then computing the expected value is equivalent to solving a corresponding PDE. Mathematically, the Black-Scholes PDE (1.11) coincides with the risk-neutral valuation completely.

In most financial literature, it becomes a standard and convenient way to work directly with the risk-neutral process. Nevertheless, we should never forget that one of the important driving reasons for a risk-neutral process is the dynamic hedging. The other way to construct a risk-neutral process is to introduce the so-called market price of risk to compensate for the excess return to the risk-free return in incomplete markets.

1.2.3 Self-financing and No Arbitrage

In this section we generalize the concept of dynamic hedging, and briefly address how to construct an arbitrage-free portfolio within which options, or more precisely contingent claims, can be valued fairly. To this end, we assume a portfolio comprising n assets A_i that are governed by the following SDEs,

$$\frac{dA_i(t)}{A_i(t)} = \mu_i dt + \sigma_i dW_i(t), \qquad i = 1, 2, \cdots, n.$$

Next we denote a vector process $\phi(t) = (\phi_1(t), \phi_2(t), \cdots, \phi_n(t))$ as a trading strategy, the initial investment $G(0)$ of the desired portfolio with such trading strategy is simply

$$G_0(\phi) = \sum_{i=1}^{n} \phi_i(0) A_i(0).$$

To ensure a fair measurement, we require that this portfolio should behave like a closed fund without any cash in-flows and out-flows. More mathematically, if a trading strategy ϕ satisfies the following condition,

$$G(t;\phi) = G_0(\phi) + \sum_{i=1}^{n} \int_0^t \phi_i(u) dA_i(u), \tag{1.20}$$

we regard this trading strategy as a self-financing trading strategy .

On the other hand, a total integral over $dA_i(u)$ and $d\phi_i(u)$ yields

$$G(t;\phi) = G_0(\phi) + \sum_{i=1}^{n} \int_0^t \phi_i(u) dA_i(u) + \sum_{i=1}^{n} \int_0^t A_i(u) d\phi_i(u). \tag{1.21}$$

By a comparison of (1.20) and (1.21), we could arrive at

$$\sum_{i=1}^{n} \int_0^t A_i(u) d\phi_i(u) = 0, \tag{1.22}$$

which is the necessary condition for a self-financing strategy.

If for any contingent claim $H(T)$ at time T exists a self-financing strategy $\phi(t)$, so that

$$H(T) = H_0 + \sum_{i=1}^{n} \int_0^T \phi_i(u) dA_i(u),$$

then the market spanned by A_i is a complete market. Market completeness means that any contingent claim could be duplicated with the existing traded assets by using a self-financing strategy $\phi(t)$. In the given formulation of the market, the risk sources of market comes from n random variables $W_i(t)$, each of which controls the dynamics of the asset A_i, therefore each risk could be traded and well priced. Any contingent claim that is linked to one or more risks in this market can be duplicated perfectly via the traded assets and a given self-financing strategy. However, if the number of risk sources W is smaller than the number of assets A, some assets may share a common risk, and therefore some of these assets become redundant to the market. If the number of risk sources W is larger than the number of assets A, it is clear that some risks can not be adequately traded, the contingent claim linked to this non-tradable risk is then unattainable to the market. This market is then incomplete.

A typical incomplete market is the market described by a stochastic volatility model where the stochastic volatility is exposed to an additional risk sources. Therefore the number of the risk sources in a stochastic volatility model is larger than the number of the underlying asset. Consequently, stochastic volatility models are generally incomplete. To make the risk-neutral valuation applicable, we need introduce a market price of the volatility risk to obtain a risk-neutral process for stochastic volatility. We will return to this topic in Chapter 3.

Now we are at a stage to rigorously define an arbitrage-free strategy.

Definition 1.2.2. *No Arbitrage: An arbitrage-free strategy ϕ must be self-financing, and satisfy the following conditions,*

$$G_0 \leqslant 0 \Longrightarrow G(t;\phi) \leqslant 0, \quad \forall \quad t > 0 \tag{1.23}$$

and

$$G_0 = 0 \Longrightarrow G(t;\phi) = 0, \quad \forall \quad t > 0. \tag{1.24}$$

The first condition says that any negative initial investment can not generate a positive future value. In other words, one can not make money by borrowing money. The second condition gives a clear restriction: zero investment leads to zero future value. Both conditions exclude any form of "free lunch".

To show that the valuation of options in the Black-Scholes model is arbitrage-free, we assume a market consisting of three securities: stock S, bond B and call C. Namely, we have $A_1 = B$, $A_2 = S$, $A_3 = C$, and then construct a portfolio with an initial value of zero,

$$G_0(\phi) = \phi_1 B_0 + \phi_2 S_0 + \phi_3 C_0 = 0. \tag{1.25}$$

Additionally it follows

$$\begin{aligned} dG(t) &= \phi_1 dB(t) + \phi_2 dS(t) + \phi_3 dC(t) \\ &= (\phi_1 rB(t) + \phi_2 S(t)\mu + \phi_3 \mu_C)dt + (\phi_2 S(t)\sigma + \phi_3 \sigma_C)dW(t). \end{aligned}$$

Since we require that $G(t)$ is an arbitrage-free portfolio, it means $G(t) = 0, t \geqslant 0$, or $dG(t) = 0, t \geqslant 0$, this is equivalent to

$$\phi_1 rB(t) + \phi_2 S(t)\mu + \phi_3 \mu_C = 0, \tag{1.26}$$

$$\phi_2 S(t)\sigma + \phi_3 \sigma_C = 0. \tag{1.27}$$

We form a linear equation system via the above three equations and obtain

$$\begin{pmatrix} rB(t) & S(t)\mu & \mu_C \\ 0 & S(t)\sigma & \sigma_C \\ B(t) & S(t) & C(t) \end{pmatrix} \cdot \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Given the processes of S and C given in (1.1) and (1.7), it is easy to solve this equation system with the following solution,

$$\begin{aligned}\phi_1 &= -\frac{\frac{\partial C}{\partial S}S+C}{B}, \\ \phi_2 &= -\frac{\partial C}{\partial S}, \\ \phi_3 &= 1.\end{aligned} \tag{1.28}$$

Inserting this solution into the first equation immediately yields the Black-Scholes PDE,

$$\frac{\partial C}{\partial t}+rS\frac{\partial C}{\partial S}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}-rC=0.$$

This replication exercise illustrates that there is no arbitrage in the Black-Scholes model. So far, we have derived the option pricing formula within the Black-Scholes model with three different but linked concepts: dynamic hedging, risk-neutral pricing and no-arbitrage. While dynamic hedging and no-arbitrage have strong economic and financial implications, risk-neutral pricing is a simplified technical formulation of the former two concepts.

1.2.4 Equivalent Martingale Measures

As we have seen above, the principle of risk-neutral valuation just replaces the drift of the historical (real) asset process with the risk-free interest rate. For a better comparison, we recall both processes.

Historical asset process:

$$\frac{dS(t)}{S(t)}=\mu dt+\sigma dW(t).$$

Risk-neutral asset process:

$$\frac{dS(t)}{S(t)}=rdt+\sigma dW(t).$$

An implication of the risk-neutral process is that the discounted asset price is a martingale. Roughly speaking, martingale is a process with an identical conditional expected value at each time point, and therefore naturally linked to fair play. Here we will briefly show two results: (1) There exists an equivalent martingale measure between the historical asset process and its risk-neutral counterpart. In other words, the risk-neutral valuation implies an equivalent martingale measure. (2) For any self-financing portfolio, there exists also a martingale measure that ensures the absence of arbitrage, hence also guarantees the risk-neutral valuation. To this end, we first define a martingale and an equivalent martingale measure in more details.

Definition 1.2.3. *Martingale: A stochastic process $X(t)$ is a martingale based on a filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geqslant 0}$, if it satisfies the following three conditions*

1. *$X(t)$ is $\mathscr{F}_t$-measurable at any time t;*
2. *$\boldsymbol{E}[|X(t)|\,|\mathscr{F}_t]$ is well-defined, i.e., $\boldsymbol{E}[|X(t)|\,|\mathscr{F}_t] < \infty$;*
3. *$\boldsymbol{E}[X(s)|\mathscr{F}_t] = X(t), \quad s \geq t.$*

The first two conditions are purely technical and ensure the conditional expectation $\mathbf{E}[X(s)|\mathscr{F}_t]$ exists. The last condition is the essence of a martingale and states that any conditional expected value of $X(s)$ based on $\mathscr{F}_t$ is just $X(t)$, and hence keeps unchanged for any $s,t \leqslant s$. For example, a standard Brownian motion is a simple martingale.

A stochastic process is always associated with a measure that characterizes the distribution law of increments. We denote P and Q as the measures for the historical and the risk-neutral processes respectively. In fact, the measure Q of a risk-neutral process with respect to the measure P for any event is always continuous, this relation establishes an equivalence between two measures. Generally, given a measurable space (Ω,Σ), two measures P and Q are equivalent, denoted by $P \sim Q$, if

$$P(A) > 0 \Longrightarrow Q(A) > 0, \quad \forall A \in \Sigma.$$

and

$$P(A) = 0 \Longrightarrow Q(A) = 0.$$

Using two equivalent measures, we could define a Radon-Nikodym derivative,

$$M(t) = \frac{dQ}{dP}(t),$$

which enables us to change a measure to another. It follows immediately

$$\mathbf{E}^P[XM] = \int_\Omega X(\omega)MdP(\omega) = \int_\Omega X(\omega)dQ(\omega) = \mathbf{E}^Q[X].$$

This interchangeability of the expected values under two different measures confirms the important role of a Radon-Nikodym derivative as intermediary between two measures.

In most cases, we want to know more, and namely how to change a measure to another. The Girsanov theorem gives us some concrete instructions to change the measures for an Itô process.

Theorem 1.2.4. *The Girsanov theorem: Given a measurable space $(\Omega,\mathbb{F},P)$, and an Itô process $X(t)$,*

$$dX(t) = a(X,t)dt + b(X,t)dW(t).$$

Denote $M(t)$ as a (an exponential) martingale under the measure P,

$$M(t) = \exp\left(-\frac{1}{2}\int_0^t \gamma^2(u)du + \int_0^t \gamma(u)dW_M(u)\right)$$

with

$$\boldsymbol{E}^P[M(t)] = 1.$$

Additionally, W *and* W_M *are correlated with* $dWdW_M = \rho dt$.

Then we have the following results:

1. $M(t)$ *defines a Radon-Nikodym derivative*

$$M(t) = \frac{dP^*}{dP}(t).$$

2. *If we define*

$$dW^*(t) = dW(t) - \gamma(t) < dW(t), dW_M(t) >= dW(t) - \rho\gamma(t)dt,$$

it is then a new Brownian motion in $(\Omega, \mathbb{F}, P^*)$.

3. *The Itô process* $X(t)$ *may take a new form under* P^*,

$$\begin{aligned} X(t) &= a(X,t)dt + b(X,t)dW^*(t) \\ &= a(X,t)dt + b(X,t)[dW(t) - \rho\gamma(t)dt] \\ &= [a(X,t) - \rho\gamma(t)b(X,t)]dt + b(X,t)dW(t). \end{aligned} \tag{1.29}$$

Now we apply the Girsanov theorem to verify the equivalent measures between the historical stock process and the risk-neutral stock process. To this end, we construct a Radon-Nikodym derivative $M(t)$ as follows

$$M(t) = \frac{dQ}{dP}(t) = \exp\left(-\frac{1}{2}\int_0^t \gamma^2 du + \int_0^t \gamma dW(u)\right)$$

with

$$\gamma = \frac{\mu - r}{\sigma}.$$

Thus we have $W(t) = W_M(t)$. The term γ may be interpreted as an excess return measured in volatility. Therefore under the measure Q the Brownian motion $W^*(t)$ is equal to

$$dW^*(t) = dW(t) - \gamma dt = dW(t) - \frac{\mu - r}{\sigma}dt,$$

and we obtain

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu dt + \sigma[dW(t) - \frac{\mu - r}{\sigma}dt] \\ &= rdt + \sigma dW(t), \end{aligned}$$

that is identical to the risk-neutral process. In this sense, the measure Q is called the risk-neutral measure, and is equivalent to the historical statistical measure P.

Now we are able to show that for any self-financing portfolio $G(t)$, there exists also a martingale measure ensuring absence of arbitrage. At first we define a numeraire $N(t)$ that is any strictly positive and not-dividend paying process. Moreover, we assume

$$\frac{G(t)}{N(t)}$$

is a martingale. According to the property of a martingale, it follows

$$\frac{G_0}{N_0} = \mathbf{E}^{P_N}\left[\frac{G(t)}{N(t)}|\mathscr{F}_0\right], \quad \forall t, \tag{1.30}$$

where P_N represents the measure associated with numeraire $N(t)$, and is equivalent to the original measure of $G(t)$. Based on (1.30), we can check the conditions for absence of arbitrage. Firstly, given $N(t) > 0$, if $G_0 \leqslant 0$, then it follows immediately $G(t) \leqslant 0$. Secondly, if $G_0 = 0$, it is also easy to see $G(t) = 0$. Hence it is proven that (1.30) implies no arbitrage. Furthermore, the equation (1.30) implies the existence of the measure P_N associated with $N(t)$.

Generally we have different choices for $N(t)$. As a consequence, there exist different equivalent measures to the original measure. In other words, equivalent martingale measures are not unique. However, a natural numeraire asset may be a money market account and is defined by

$$H(t) = \exp\left(\int_0^t r(u)du\right), \tag{1.31}$$

where $r(t)$ denotes the instantaneous short rate of interest. Obviously, $H(t)$ is strongly linked to a zero-coupon bond,

$$H(t) = \frac{1}{B(t)} = 1/\exp\left(-\int_0^t r(u)du\right).$$

It can be verified that the measure under numeraire $H(t)$ is exactly the risk-neutral measure Q. To see this, we define

$$\frac{dQ}{dP_N} = \frac{H(t)N_0}{N(t)H_0} = \frac{H(t)N_0}{N(t)},$$

and obtain

$$\mathbf{E}^{P_N}\left[\frac{G(t)}{N(t)}|\mathscr{F}_0\right] = \mathbf{E}^{Q}\left[\frac{G(t)}{H(t)}|\mathscr{F}_0\right]\frac{1}{N_0} = \frac{G_0}{N_0}, \tag{1.32}$$

or

$$G_0 = \mathbf{E}^{Q}\left[\frac{G(t)}{H(t)}|\mathscr{F}_0\right] = \mathbf{E}^{Q}\left[B(t)G(t)|\mathscr{F}_0\right]. \tag{1.33}$$

This means that, if the money market account is used as a numeraire, the fair price of the arbitrage-free asset process $G(t)$ is just its present value. If $r(t)$ is a constant, we have an usual discounting formula,

$$G_0 = e^{-rt}\mathbf{E}^Q[G(t)|\mathscr{F}_0].$$

Since the money market account is unique, the risk-neutral measure Q is also unique. Finally we could conclude:

Theorem 1.2.5. *Equivalent martingale measure and no arbitrage: Denote $G(t)$ as the process of a self-financing portfolio, P is its historical measure. For any numeraire $N(t)$, there exists an equivalent P_N, such that*

$$\frac{G_0}{N_0} = \boldsymbol{E}^{P_N}\Big[\frac{G(t)}{N(t)}|\mathscr{F}_0\Big], \quad \forall \quad t,$$

and $G(t)$ is arbitrage-free under P_N. If $N(t)$ is the money market account, then P_N is the risk-neutral measure.

Summing up our discussions on no arbitrage (dynamic hedging in Black-Scholes model), risk-neutral valuation and equivalent martingale measure, we could establish a cycle relation as follows:

no arbitrage $\Rightarrow$ risk-neutral valuation $\Rightarrow$ equivalent martingale measure $\Rightarrow$ no arbitrage.

By this relation, we can also conclude that risk-neutral valuation implies no arbitrage, and no arbitrage implies equivalent martingale measure. The latter result is, however, difficult to be proven directly, see Harrison and Kreps (1979), Harrison and Pliska (1981), Geman, El Karoui and Rochet (1995), Musiela and Rutkowski (2005).

1.3 Volatility Quotations in Markets

1.3.1 Implied Volatilities

In the Black-Scholes formula there are five parameters determining the option price, they are the spot price S_0, the constant interest rate r, the strike K, the constant volatility σ and the maturity T. Except for the volatility σ, other four parameters are either specified by option contract or can be observed directly in markets. Therefore between the only unknown volatility parameter σ and option price there is a one-to-one correspondence. As long as the market price of a call or a put is given, we can calculate a volatility matching the option price perfectly. The volatility obtained from a given option market price is called the implied volatility. Mathematically, an implied volatility σ_{impl} is a quantity satisfying the following relation,

$$C^{BS}(\sigma_{impl};K,T,S_0,r) = C^{Market}.$$

As S_0 and r are given by markets, for each pair (K,T) that are determined by an option contract, we can find a corresponding implied volatility. However, if we retrieve the implied volatilities from a number of option prices with different strikes and maturities, they often are not same. This phenomena called the volatility smile contradicts the assumption in the Black-Scholes model that volatility is constant over all strikes and maturities, and at the same time, indicates that the constant volatility in the Black-Scholes model should be too restrictive to capture the true market movements. Implied volatility is then a “wrong” quantity generated by a “wrong” pricing tool with a right market price.

Generally, volatility smile has two basic patterns: symmetric smile and skew. Symmetric smile is referred to as the case where implied volatility increases if strike moves away from the stock spot price. Down-sloping skew, also called sneer, corresponds the case of decreasing implied volatilities with increasing strikes. Correspondingly, up-sloping skew stands for increasing implied volatilities with increasing strikes. For a fixed strike, implied volatilities over different maturities may display a certain term structure which is called the volatility term structure. Expressing all implied volatilities over strikes and maturities in a matrix form yields a smile matrix or a smile surface.

1.3.2 Market Quotations

Due to the one-to-one correspondence between option price and its implied volatility, implied volatilities should content the same market information as the option prices. However, option markets seem prefer to the implied volatilities for market quotations. Compared with quoting directly with prices, quoting with implied volatilities has the following advantages:

1. Quotations with implied volatilities clarify market situation and deliver a more clear and transparent picture about market sentiments because implied volatility standardizes option price in year, and is independent of the absolute values of spot stock prices. Usually it is quoted in terms of strike (or moneyness K/S_0) and maturity.
2. Market participants can compare the relative expenses of different options easily via implied volatility and can check the arbitrage conditions more conveniently.
3. Implied volatility allows us to gauge the historical movements of option prices consistently and transparently, and also allows us to measure the deviation between the underlying and option markets by comparing implied volatility and statistical volatility.
4. It is straightforward to estimate market volatilities for the options that are traded less actively and less liquidly. Using some interpolation techniques and models we can retrieve implied volatility for any strike and maturity based on smile surface.

However, oppositely to these advantages in valuing European-style options, implied volatilities can not be applied to soundly value exotic derivatives involving more than one strikes, for example, barrier options and clique options. Since implied volatilities are generated by a "wrong" pricing machine, there is no natural and reliable dynamics of the implied volatility. As a result, implied volatilities serve only as market quotations for options, not as market observations for true volatilities.

In stock option markets, a volatility surface is usually spanned by strike K (or moneyness K/S_0) and maturity T. In Figure (1.1), a volatility surface of DAX index illustrates strong down-sloping smiles for all maturities with the higher out-of-the-money (OTM) call volatilities and the lower in-the-money (ITM) call volatilities, reflecting the clear market fear of a further decline of DAX prices during the financial crisis in 2009.

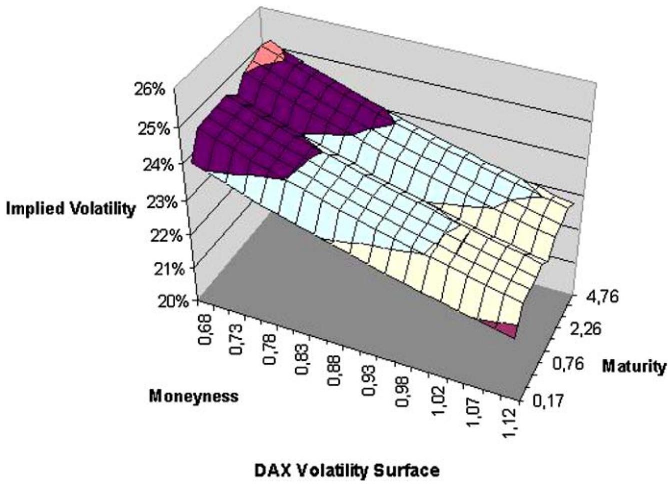

Fig. 1.1 The implied volatility surface of DAX index, as of March 12, 2009.

1.3.3 Special Case: FX Market

In FX option (foreign exchange option) markets, a volatility surface is given in terms of Delta and maturity. This special quotation convention is popular for traders and brokers, but not welcome for quants. We recall the risk-neutral FX process as follows:

$$\frac{dS(t)}{S(t)} = (r - r^*)dt + \sigma dW(t), \tag{1.34}$$

where r and r^* are the domestic and foreign interest rates respectively. Garman-Kohlhagen (1983) derived a pricing formula for FX options within the Black-Scholes framework, it reads

$$C_0 = S_0 e^{-r^*T} N(d_1) - K e^{-rT} N(d_2), \tag{1.35}$$

with

$$d_1 = \frac{\ln(S_0/K) + (r - r^* + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}},$$

$$d_2 = \frac{\ln(S_0/K) + (r - r^* - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}.$$

Delta is the first derivative of an option with respective to its underlying, and for FX options it is equal to

$$\Delta = e^{-r^*T} N(d_1),$$

which again is a function of stock price, strike, maturity and volatility. Therefore there is no clear functional relation between volatility and price if implied volatility is given in terms of Delta. For example, if a quoted implied volatility at the point of 0.15 Delta (call) is equal to 0.12, we have

$$\Delta(\sigma = 0.12) = e^{-r^*T} N(d_1, \sigma = 0.12) = 0.15.$$

FX option markets quote usually three most liquid options for a given maturity, they are -0.15 Delta put, ATM call and 0.15 Delta call where ATM means that strike K is equal to FX forward price, not spot price, namely $K = S_0 e^{(r-r^*)T}$. Let K_1,K_2 and K_3 correspond to the volatilities σ_{15P}, σ_{ATM} and σ_{15C} of the quoted options respectively, we have the following relations,

$$-0.15 = -e^{-r^*T} N\Big(-\frac{\ln(S_0/K_1) + (r - r^* + \frac{1}{2}\sigma_{15P}^2 T)}{\sigma_{15P}\sqrt{T}}\Big), \tag{1.36}$$

$$\Delta = -e^{-r^*T} N\Big(\frac{1}{2}\sigma_{ATM}\sqrt{T}\Big) \approx 0.5, \qquad K_2 = S_0 e^{(r-r^*)T}, \tag{1.37}$$

$$0.15 = e^{-r^*T} N\Big(\frac{\ln(S_0/K_3) + (r - r^* + \frac{1}{2}\sigma_{15C}^2 T)}{\sigma_{15C}\sqrt{T}}\Big). \tag{1.38}$$

Note that the Delta of short-maturity ATM option is approximately equal to 0.5. Given $(\sigma_{15P}, \sigma_{ATM}, \sigma_{15C})$, we can numerically calculate (K_1, K_2, K_3), and then establish a conventional volatility smile surface in terms of strike and maturity. However, If the quoted Delta volatilities have changed, the corresponding strikes change too. The translated FX volatility surface spanned by strike and maturity will be shifted correspondingly. Meanwhile, some data sponsors and brokers such as Reuters provide a large range of Delta volatilities for puts with -0.10 Delta, -0.15 Delta, -0.20 Delta, $\cdots$, -0.45 Delta, for calls with 0.45 Delta, 0.40 Delta, 0.35 Delta, $\cdots$, 0.10 Delta. Given so many Delta volatilities, we can therefore establish a fine FX volatility surface.

More interestingly, some data sponsors and brokers do not deliver directly Delta volatilities, but the so-called butterfly volatility σ_{BF}, the so-called risk-reversal volatility σ_{BF} and ATM volatility σ_{ATM}. The market conventions of the butterfly volatility σ_{BF} and the risk-reversal volatility σ_{BF} are parallel to the trading strategies named butterfly and risk-reversal, and given by

$$\sigma_{RR} = \sigma_{15C} - \sigma_{15P},$$

$$\sigma_{BF} = \frac{1}{2}(\sigma_{15P} + \sigma_{15C}) - \sigma_{ATM}.$$

Therefore, we can express the Delta volatilities with σ_{BF} and σ_{RR}, and have

$$\sigma_{15P} = \sigma_{ATM} + \sigma_{BF} - \frac{1}{2}\sigma_{RR}, \tag{1.39}$$

$$\sigma_{15C} = \sigma_{ATM} + \sigma_{BF} + \frac{1}{2}\sigma_{RR}. \tag{1.40}$$

In this sense, the quotations with the butterfly volatility σ_{BF} and the risk-reversal volatility σ_{BF} are equivalent to the quotation with two Delta volatilities. For more details on FX option markets, please see Wystup (2006).

Chapter 2
Characteristic Functions in Option Pricing

We consider in this chapter a general Itô process for stock prices. In a generalized valuation framework for options, the distribution function of stock price is analytically unknown. To express (quasi-) closed-form exercise probabilities and valuation formula, characteristic functions of the underlying stock returns (logarithms) are proven to be not only a powerful and convenient tool to achieve analytical tractability, but also a large accommodation for different stochastic processes and factors. In first section, we derive two important characteristic functions under Delta measure and forward measure respectively, under which two exercise probabilities can be calculated. The pricing formula for European-style options may be expressed in a form of inverse Fourier transform. As a result, we obtain a generalized principle for valuing options under the risk-neutral measure via characteristic functions. There is a corresponding extension for FX options.

Next in Section 2.2.1, we first demonstrate some nice properties of characteristic function. As conditional expected value, characteristic function shares all properties of integration operator and expectation operator. The most important property of characteristic function with respect to setting up a comprehensive option pricing framework is that if stochastic factors are mutually independent, the characteristic function of the sum of stochastic factors is just a product of the characteristic function of each stochastic factor. In Section 2.2.2, we follow an approach of Bakshi and Madan (2000) and interpret characteristic functions as Arrow-Debreu prices in a Fourier-transformed space. As shown in Section 2.2.3 by some examples, most popular pricing models, particularly, stochastic volatility models, admit analytical solutions for characteristic functions, and therefore, also analytical solutions for valuation formulas in terms of inverse Fourier transform. The Black-Scholes formula can be easily verified with the characteristic functions of normal distribution. Additionally, we can establish an equivalence between the Fourier transform approach and the traditional PDE approach where the Feynman-Kac theorem and Kolmogorov's backward equation play a central role. Finally, as one of many advantages of the Fourier transform approach, it allows us a modular valuation for options, with which all relevant independent stochastic factors such as interest rates, volatilities and jumps may be dealt with as a module in an unified valuation procedure.

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_2,

The approach of the modular pricing will be discussed later in details throughout this book, and provides us with large efficiency and rich flexibility in practical applications.

2.1 Constructing Characteristic Functions (CFs)

2.1.1 A General Process for Stock Price

We start with a rather general setting where the processes of two important financial factors: stock price process (risky asset process) and risk-free short rate process. Stock price $S(t)$ is governed by the following Itô process,

$$\frac{dS(t)}{S(t)} = a(u(t),t)dt + b(v(t),t)dW_1(t), \tag{2.1}$$

where $a(\cdot,t) \in \mathfrak{L} : \mathbb{R}^k \times [0,T] \to \mathbb{R}$, $b(\cdot,t) \in \mathfrak{L}^2 : \mathbb{R}^m \times [0,T] \to \mathbb{R}^n$ and dW_1 is a n-dimensional standard Brownian motion. $\mathfrak{L}$ and $\mathfrak{L}^2$ denote the set of functions $y(t)$ such that $\int_0^T |y(t)|\,dt < \infty$ almost surely and $\int_0^T y^2(t)dt < \infty$ almost surely, respectively. $u(t) \in \mathbb{R}^k$ and $v(t) \in \mathbb{R}^m$ could also follow Itô processes and control the process $S(t)$. Defining $x(t) \equiv \ln S(t)$ and applying Itô's lemma yield

$$dx(t) = \left(a(u(t),t) - \frac{1}{2}b^2(v(t),t)\right)dt + b(v(t),t)dW_1(t). \tag{2.2}$$

In fact, $a(\cdot,t)$ is irrelevant for the risk-neutral processes of $dS(t)$ and $dx(t)$.

For simplicity, we consider only a one-factor version of risk-free short rate process $r(t)$ which is defined by

$$dr(t) = c(r,t)dt + e(r,t)dW_2(t), \tag{2.3}$$

where $c(\cdot,t) \in \mathfrak{L} : \mathbb{R} \times [0,T] \to \mathbb{R}$, $e(\cdot,t) \in \mathfrak{L}^2 : \mathbb{R} \times [0,T] \to \mathbb{R}$ and dW_2 is a one-dimensional Brownian motion. All of the functions a, b, c and e satisfy the technical conditions[1] that guarantee the existence of a solution of the above two SDEs. Under the risk-neutral valuation, we have a simplified process for $x(t)$,

---

[1] The technical conditions sufficient for the existence of unique solution of stochastic differential equation are the Lipschitz and the growth conditions. A function $y(x,t)$ satisfies the Lipschitz condition in x if there is a constant k such that for any x_1 and x_2 and any time t

$$\| y(x_1,t) - y(x_2,t) \| \leq k \| x_1 - x_2 \| .$$

Similarly, $y(x,t)$ satisfies the growth condition in x if there is a constant k such that for any x and any t

$$\| y(x,t) \|^2 \leq k(1 + \| x \|^2).$$

$$dx(t) = \left(r(t) - \frac{1}{2}b^2(v(t),t)\right)dt + b(v(t),t)dW_1. \tag{2.4}$$

In this general setting, the dynamics of the stock price $S(t)$ are driven by a pure diffusion $W(t)$ as in the simple Black-Scholes model. The essential extensions here are a stochastic interest rate $r(t)$ and a stochastic volatility term $b(\cdot)$, that will be specified in a more correct modeling. As documented in a number of empirical studies, pure diffusive movements are not good enough to capture all statistical and implied characteristics of asset prices. Different types of jumps may be a good complement to the above setting, and will be discussed later.

To value FX options, we can adapt the process of $x(t)$ by utilizing the interest rate parity of FX rates, and obtain the following process for the FX rate $x(t)$,

$$dx(t) = \left(r(t) - r^*(t) - \frac{1}{2}b^2(v(t),t)\right)dt + b(v(t),t)dW_1 \tag{2.5}$$

with $r(t)$ and $r^*(t)$ as the domestic and foreign interest rate respectively. The foreign interest rate $r^*(t)$ is governed by a similar process as $r(t)$.

In the following, we use a simple functional form of $b(v(t),t)$ for general discussion, namely $b(v(t),t) = v(t)$. A popular choice for $b(\cdot)$ may be $b(v(t),t) = \sqrt{V(t)}$ as in the Heston model (1993). Another more complicated and still tractable form for $b(\cdot)$ may be the linear affine structure suggested by Dai and Singleton (2000).

2.1.2 Valuation of European-style Options via CFs

The general process for stock price given above includes the Black-Scholes world, stochastic interest rates, stochastic volatilities, and can be extended by adding random jumps. We know that the common property of the pricing formulas *à la* Black-Scholes is the (log-) normal distribution of exercise probabilities. Obviously, in the extended framework the exercise probabilities are no longer strictly (log-) normally distributed. However, they can be expressed by the Fourier inversion of the associated CFs which may often have closed-form solutions with different specifications of stochastic factors. Consequently, we can also express option pricing formula in a (quasi) closed-form within this new framework.

In coherence with the existing financial literature, we assume that a European-style option of the underlying stock S with strike price K can be valued by

$$\begin{aligned} C(S_0,T;K) &= \mathbf{E}^Q\left[\exp\left(-\textstyle\int_0^T r(t)dt\right)(S(T)-K)\cdot\mathbf{1}_{(S(T)>K)}\right] \\ &= \mathbf{E}^Q\left[\exp\left(-\textstyle\int_0^T r(t)dt\right)S(T)\cdot\mathbf{1}_{(S(T)>K)}\right] \\ &\quad -\mathbf{E}^Q\left[\exp\left(-\textstyle\int_0^T r(t)dt\right)K\cdot\mathbf{1}_{(S(T)>K)}\right], \end{aligned} \tag{2.6}$$

where Q denotes the risk-neutral (equivalent martingale) measure at the time T. According to Geman, El Karoui and Rochet (1995), Björk (1996) and others, in order to simplify the calculations in the above equation, we can change the numeraire. For the first term in the second equality, we choose the stock price $S(t)$ as numeraire and switch the measure Q to Q_1, and for the second term we use the zero-coupon bond to switch Q to the so-called T-forward measure Q_2.

We first consider the change of the risk-neutral measure Q to the new measure Q_1 associated with numeraire $S(t)$. According to the definition of equivalent measures and the Girsanov theorem given in Section 1.2.4, we should construct a Radon-Nikodym derivative using the corresponding numeraires,

$$\begin{aligned}\frac{dQ_1}{dQ}(t) &= \frac{S(t)H(0)}{H(t)S(0)} \\ &= \exp\left(-\int_0^t r(u)du\right)\frac{S(t)}{S_0} = g_1(t),\end{aligned} \tag{2.7}$$

where the numeraire underlying the risk-neutral measure is, as shown in Section 1.2.4, the money market account,

$$H(t) = \exp\left(\int_0^t r(u)du\right).$$

Obviously $H(t)$ is equal to 1 at time 0, namely $H(0) = 1$. It is easy to verify $g_1(t)$ is a martingale with an expected value of one, and has the Deléans exponential as follows:

$$M(t) = \exp\left(-\frac{1}{2}\int_0^t v(u)^2 du + \int_0^t v(u)dW(u)\right).$$

Applying Girsanov's theorem yields

$$dW^* = dW - v(t)dt.$$

The risk-neutral process $x(t)$ under the measure Q_1 is then changed to

$$\begin{aligned}dx(t) &= \left(r - \frac{1}{2}v(t)^2\right)dt + v(t)dW(t) \\ &= \left(r - \frac{1}{2}v(t)^2\right)dt + v(t)(dW^*(t) + v(t)dt) \\ &= \left(r + \frac{1}{2}v(t)^2\right)dt + v(t)dW^*(t),\end{aligned} \tag{2.8}$$

where we can see that the drift term becomes $r + \frac{1}{2}v(t)^2$ under the new measure. As shown in the Black-Scholes model, the risk sensitivity Delta, defined as the first derivative of call price to spot price, namely $\frac{\partial C}{\partial S}$, could be interpreted as an exercise probability under the measure Q_1. In this sense, we label Q_1 as the Delta measure.

The second change of measures takes place between two numeraires: the money market account $H(t)$ and the zero-coupon bond $B(t,T)$. Given the process of the risk-free short rate $r(t)$, a zero-coupon bond maturing at time T is defined by

$$B(t,T) = \mathbf{E}\Big[\exp(-\int_t^T r(u)du)\Big].$$

The Radon-Nikodym derivative for a change of the risk-neutral measure Q to a new measure Q_2 is given by

$$\begin{aligned}\frac{dQ_2}{dQ}(t) &= \frac{B(t,T)H(0)}{B(0,T)H(t)} \\ &= \exp\left(-\int_0^t r(u)du\right)\frac{B(t,T)}{B(0,T)} = g_2(t).\end{aligned} \tag{2.9}$$

It can be shown that $g_2(t)$ is a martingale with an expected value of one, and its Doléans exponential is

$$M(t) = \exp\left(-\frac{1}{2}\int_0^t \sigma_B^2(u)du - \int_0^t \sigma_B(u)dW(u)\right), \tag{2.10}$$

where $\sigma_B(u)$ is the volatility of $B(t,T)$. Dependently on the specification of the short-rate process, one can derive a pricing formula for $B(t,T)$. Since any risky asset under the measure Q_2 takes the zero-coupon bond $B(t,T)$ as numeraire, therefore an asset under the measure Q_2 can be regarded as the forward price of this asset for time T under the measure Q. In this sense, the measure Q_2 is called a forward $T-$ measure, or for short, forward measure. At time T, the Radon-Nikodym derivative $g_2(t)$ has a special simplified form

$$g_2(T) = \frac{\exp(-\int_0^T r(u)du)}{B(0,T)} = \frac{1}{H(T)B(0,T)}. \tag{2.11}$$

As a property of Radon-Nikodym derivative, $g_j(t), j=1,2$, define two density functions at t. We can also call $\{g_1(t)\}_{t\geqslant 0}$ the unit discounted stock process and $\{g_2(t)\}_{t\geqslant 0}$ the unit discounted bond process due to their martingale property.[2] Under these new measures, the option pricing representation (2.6) can be restated as

$$\begin{aligned}C_0 &= S_0\mathbf{E}^{Q_1}\left[\mathbf{1}_{(x(T)>\ln K)}\right] - B(0,T)K\mathbf{E}^{Q_2}\left[\mathbf{1}_{(x(T)>\ln K)}\right] \\ &= S_0F_1^{Q_1}(x(T)>\ln K) - B(0,T)KF_2^{Q_2}(x(T)>\ln K).\end{aligned} \tag{2.12}$$

This representation for an option pricing formula is similar to the Black-Scholes formula where the probabilities $F^{Q_j}, j=1,2$, in the latter case are two standard normal distributions. Now we express these probabilities by the Fourier transform. The characteristic functions of $F_j^{Q_j}$ are defined by

$$f_j(\phi) \equiv \mathbf{E}^{Q_j}\left[\exp(i\phi x(T))\right] \quad \text{for} \quad j=1,2. \tag{2.13}$$

[2] These two processes are also termed as likelihood processes because they imply two probabilities.

By using the above two Radon-Nikodym derivatives again, we obtain new expressions for the CFs $f_j(\phi), j = 1,2$, under the original risk-neutral measure Q:

$$f_1(\phi) \equiv \mathbf{E}^{Q_1}\left[\exp(i\phi x(T))\right] = \mathbf{E}^{Q}\left[g_1(T)\exp(i\phi x(T))\right] \tag{2.14}$$

and

$$f_2(\phi) \equiv \mathbf{E}^{Q_2}\left[\exp(i\phi x(T))\right] = \mathbf{E}^{Q}\left[g_2(T)\exp(i\phi x(T))\right]. \tag{2.15}$$

Given the CFs $f_j(\phi)$ of $x(T)$ under the two different measures Q_j, the density function $y_j(x(T))$ of $x(T)$ is simply the inverse Fourier transform,

$$y_j(x(T)) = \frac{1}{2\pi}\int_{\mathbb{R}} f_j(\phi)e^{-i\phi x(T)}d\phi, \quad j = 1,2. \tag{2.16}$$

Using the density function $y_j(x(T))$ and denoting $a = \ln K$, we can calculate the exercise probability F_j as follows:

$$\begin{aligned} F_j(x(T) > a) &= \int_a^\infty y_j(x)dx \\ &= \int_a^\infty \left(\frac{1}{2\pi}\int_{\mathbb{R}} f_j(\phi)e^{-i\phi x}d\phi\right)dx \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} f_j(\phi)\left(\int_a^\infty e^{-i\phi x}dx\right)d\phi, \end{aligned}$$

where we have changed the order of two integrations by applying the Fubini theorem. A further straightforward calculation yields

$$F_j = \frac{1}{2} + \frac{1}{2\pi}\int_{\mathbb{R}} f_j(\phi)\frac{e^{-i\phi a}}{i\phi}d\phi. \tag{2.17}$$

Note that the probability F_j must be real, the above expression can be rewritten as (see Stuart and Ord, 1994),

$$\begin{aligned} F_j &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}(f_j)\cos(\phi a) - \mathrm{Re}(f_j)\sin(\phi a)}{\phi}d\phi, \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_j(\phi)\frac{\exp(-i\phi\ln K)}{i\phi}\right)d\phi, \quad j = 1,2, \end{aligned} \tag{2.18}$$

where Im and Re denote the imaginary and real part of complex number respectively. The last expression of F_j allows us to carry out the integration over the half real line, instead of over the whole real domain.

This new expression for the probabilities F_j has some advantages: Firstly, we can show later that in the most existing and new option pricing models the CF $f_j(\phi)$ can be expressed by elementary functions, and have analytical solutions. Secondly, F_j contains only a single integral which can be calculated in an efficient way. Finally, these expressions are independent of the specifications of the processes of individual stochastic factors. Therefore, the above pricing approach offers us a comprehensive framework to derive the pricing formulas for European-style options with a minimal dependence upon the specifications of factors.

Summarizing the above results, we have the following principle for constructing characteristic functions for option pricing using the unique risk-neutral measure Q.

Theorem 2.1.1 *The Principle for Constructing Characteristic Functions 1: If the risk-neutral processes of stock price and interest short rate are of form in (2.4) and (2.3), respectively, and the drift and diffusion components of these processes satisfy all technical conditions so that the solutions to (2.4) and (1.21) exist, then the characteristic functions of the exercise probabilities F_1 and F_2 for a European-style call option*

$$C(S_0,T;K) = S_0F_1 - KB(0,T)F_2$$

have the expression of (2.14) and (2.15) respectively. This is independent of the specifications of the stock price process and the short-rate process.

An enhanced principle for constructing CFs will be given in the following section. In fact, one can extend the one-factor version of the short-rate process to a multi-factor version without having any effect on the above results. This principle allow us to construct much more complicated and interesting variants of option pricing formulas. Note that there are no particular constraints on the processes of $S(t)$, $v(t)$ and $r(t)$. Therefore, these factors can be specified rather arbitrarily. The following chapters will show the effectiveness of this method in developing new closed-form solutions.

In this manuscript, if not explicitly noted, expectations are always taken under the measure Q. Obviously, the principle of constructing CFs is also valid for European-style put options,

$$P_0 = KB(0,T)F_2 - S_0F_1 \tag{2.19}$$

with

$$F_j = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_j(\phi)\frac{\exp(-i\phi\ln K)}{i\phi}\right)d\phi, \quad j=1,2. \tag{2.20}$$

Expressing a probability via its characteristic function is equivalent to expressing a probability via a density function from many points of view. Firstly, both alternatives involve a single integral (for the one-dimensional case), and hence need a similar procedure to calculate the integral. Additionally, the one-to-one correspondence between a characteristic function and its distribution guarantees a unique form of the option pricing formula.

2.1.3 Special Case: FX Options

As argued by Garman and Kohlhagen (1983) in the Black-Scholes framework for European-style FX options, since no arbitrage in two-country economics is ensured by the interest rate parity, one needs a modification for the exchange rate risk-neutral process and therefore also for $g_1(t)$. We recall the risk-neutral process of exchange rate $S(t)$ given in (2.5),

$$\frac{dS(t)}{S(t)} = [r(t) - r^*(t)]dt + v(t)dW_1(t) \tag{2.21}$$

with $r(t)$ and $r^*(t)$ as the domestic and foreign interest rates respectively. The process of $r(t)$ has been given before in (2.3). We may propose a similar process for $r^*(t)$ and assume that the solution for a foreign zero-coupon bond price

$$B^*(t,T) = \mathbf{E}^Q\big[\exp(-\int_t^T r^*(u)du)\big]$$

exists. As indicated by the process for exchange rate in (2.21), the discounted risk-neutral exchange rate $S(t)$ by the money market account $H(t)$ is no longer a martingale. Instead by interest rate parity, $S(t)$ is a martingale only if it is discounted by the difference between domestic and foreign money market account, namely

$$\mathbf{E}^Q\big[S(t)\exp(-\int_0^t [r(u) - r^*(u)]du)\big] = \mathbf{E}^Q\big[S(t)\frac{H^*(t)}{H(t)}\big] = 1$$

with

$$H^*(t) = \exp\big(\int_0^t r^*(u)du\big).$$

This gives us a hint to define a Radon-Nikodym derivative as follows:

$$\frac{dQ_1^*}{dQ}(t) = \frac{S(t)H^*(t)}{S(0)H(t)}. \tag{2.22}$$

The FX options can then be valued similarly as the asset options,

$$\begin{aligned} C &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)(S(T) - K)\cdot \mathbf{1}_{(S(T)>K)}\right] \\ &= S_0\mathbf{E}^{Q_1^*}\left[\exp\left(-\int_0^T r^*(t)dt\right)\mathbf{1}_{(x(T)>\ln K)}\right] - B(0,T)K\mathbf{E}^{Q_2}\left[\mathbf{1}_{(x(T)>\ln K)}\right], \end{aligned}$$

where in the first term there is still the stochastic foreign discount factor. Therefore we can not consider the first term as a probability under measure Q_1^*, and face the same situation as computing F_2, this means, we need an additional foreign forward measure to eliminate the stochastic foreign discount factor. To this end, we define the following Radon-Nikodym derivative,

$$\frac{dQ_1}{dQ_1^*}(t) = \frac{B^*(t,T)H^*(0)}{B^*(0,T)H^*(t)} = \frac{B^*(t,T)}{B^*(0,T)H^*(t)}. \tag{2.23}$$

Particularly,

$$\frac{dQ_1}{dQ_1^*}(T) = \exp\left(-\int_0^T r^*(t)dt\right)\frac{1}{B^*(0,T)}.$$

Applying this Radon-Nikodym derivative yields immediately

$$C = S_0 B^*(0,T)\mathbf{E}^{Q_1}\left[\mathbf{1}_{(x(T)>\ln K)}\right] - B(0,T)K\mathbf{E}^{Q_2}\left[\mathbf{1}_{(x(T)>\ln K)}\right], \tag{2.24}$$

which is exactly the formula we want to obtain, in order to construct two CFs under appropriate measures. The final change of the risk-neutral measure Q to the new equivalent martingale measure Q_1 is already demonstrated by above two-step change and given by

$$\frac{dQ_1}{dQ}(t) = \frac{dQ_1^*}{dQ}\frac{dQ_1}{dQ_1^*}(t) = \frac{S(t)B^*(t,T)}{S(0)B^*(0,T)H(t)}. \tag{2.25}$$

Its form at time T is a bit simpler

$$\frac{dQ_1}{dQ}(T) = g_1(T) = \exp\left(-\int_0^T r(u)du\right)\frac{S(T)}{S_0 B^*(0,T)}, \tag{2.26}$$

which can be used to compute CF $f_1(\phi)$ and the corresponding first exercise probability F_1 for FX options. With F_2 keeping unchanged, the pricing formula for FX call options can be generally expressed as

$$C_0 = S_0 B^*(0,T)F_1 - KB(0,T)F_2,$$

that is an extension of the Garman and Kohlhagen formula in a general model setup.

2.2 Understanding Characteristic Functions

2.2.1 *Properties of Characteristic Functions*

The characteristic function of a random variable is essentially a Fourier transform mapping the real random variable into a frequency domain. Therefore, CFs $f_j(\phi)$ derived above share any property of Fourier transform, and also any property of integral operator. We now briefly examine some special features of CF with respect to probability.

The first useful property of CF lies in dealing with independent random variables. Let $X_1, X_2, \cdots, X_n$, be some independent, but not necessarily identical random variables, and $Y_n = \sum_{j=1}^n X_j$ is the sum of these variables, then the CF of Y_n is just the product of the CF of each X_j, namely

$$f(\phi;Y_n) = \prod_{j=1}^n f(\phi, X_j). \tag{2.27}$$

Since $X_j, j = 1,2,\cdots,n$, are not necessarily identically distributed, it is convenient to examine the joint distribution of X_j using the joint CF $f(\phi;Y_n)$. As shown later, if stock returns are modeled by a Brownian motion and a jump process whose increments are governed by Poisson law, we can calculate the distribution of stock returns by the joint CF which consists of the CF of Gaussian distribution and the CF of Poisson distribution. This allows us to deal with each stochastic factor in an asset process as a building block. From the point of view of engineering, each stochastic factor can be regarded as module in a whole stochastic process. The idea of modular pricing presented later in this book then utilizes this property of CF.

Now let y denote the inverse Fourier transform of CF f of a random variable x. As mentioned above, y is then the corresponding density function, and equal to

$$y(x) = \frac{1}{2\pi} \int_{\mathbb{R}} f(\phi) e^{-i\phi x} d\phi.$$

There is a Fourier transform relation between $y(x)$ and $f(\phi)$. We denote this relation with

$$f(\phi) \Longleftrightarrow y(x).$$

Some properties of the CF $f(\phi)$ and its inverse transform $y(x)$ can be given as follows:

- Linearity:
$$af_1(\phi) + bf_2(\phi) \Longleftrightarrow ay_1(x) + by_2(x).$$
- Multiplication:
$$f_1(\phi) f_2(\phi) \Longleftrightarrow \frac{1}{2\pi}(y_1 * y_2)(x).$$
- Convolution:
$$(f_1 * f_2)(\phi) \Longleftrightarrow y_1(x) y_2(x).$$
- Conjugation:
$$\overline{f(\phi)} \Longleftrightarrow \overline{y(-x)}.$$
- Scaling:
$$f(a\phi) \Longleftrightarrow \frac{1}{|a|} y(x/a), \quad a \in \mathbb{R}, \quad a \neq 0.$$
- Time reversal:
$$f(-\phi) \Longleftrightarrow y(-x).$$
  This feature can be considered as a special case of the scaling feature.
- Time shift:
$$f(\phi) \Longleftrightarrow y(-x).$$

Note that f_1 and f_2 here are two arbitrary CFs, not same as the first and second CFs given above in the context of option pricing.

The moments of a random variable can also be derived from its characteristic function. Thus, some important statistical parameters such as variance, skewness and kurtosis are automatically available if the characteristic function is known. Let $m_n^*(x(T))$ denote the n-th order non-central moment of $x(T)$, we can express $m_n^*(x(T))$ with the CF $f(\phi; x(T))$ by the following equation:

$$m_n^*(x(T)) = (-i)^n \left[\frac{\partial^n f(\phi; x(T))}{\partial \phi^n} \right]_{\phi=0}. \tag{2.28}$$

It must be emphasized that the CF $f(\phi; x(T))$ for calculating the moments should be derived from the original measure Q since we need the moments under actual measure,[3] that is

$$f(\phi; x(T)) \equiv \mathbf{E}^Q \left[\exp(i\phi x(T)) \right]. \tag{2.29}$$

Hence $f(\phi; x(T))$ is generally not identical to $f_j(\phi)$ as given above in (2.13). In fact, the moment-generating function of $x(T)$ is defined by

$$mgf(\eta) = \mathbf{E}^Q \left[\exp(\eta x(T)) \right] \tag{2.30}$$

for a positive real number $\eta, \eta > 0$. Obviously, a moment-generating function is very similar to a CF, and there is an interchange between them by setting $\phi = -i\eta$ or $\eta = i\phi$,

$$f(-i\eta) = mgf(\eta), \quad mgf(i\phi) = f(\phi).$$

In this sense, the CF $f(\phi)$ is equivalent to the moment-generating function, and therefore can be applied to calculate the moments of $X(T)$.

To compute skewness and kurtosis, the first four central moments $m_n(x(T))$ should be known and can be given by a recursive scheme (Stuard and Ord, 1994):

$$\begin{aligned} m_2 &= m_2^* - (m_1^*)^2, \\ m_3 &= m_3^* - 3m_2^* m_1^* + 2(m_1^*)^2, \\ m_4 &= m_4^* - 4m_3^* m_1^* + 6m_2^*(m_1^*)^2 - 3(m_1^*)^4. \end{aligned} \tag{2.31}$$

The skewness and kurtosis of $x(T)$ are then given by

$$skewness \equiv \frac{m_3}{(m_2)^{3/2}}, \qquad kurtosis \equiv \frac{m_4}{m_2^2}, \tag{2.32}$$

respectively.

As seen in the Black-Scholes model, $x(t)$ is Gaussian and therefore has a skewness of zero and a kurtosis of three. It is empirically evident and also supported by smile effect of implied volatilities that the actual distribution of stock returns $x(t)$ exhibits a negative skewness and an excess kurtosis. Therefore, in order to capture a leptokurtic distribution and to recover smile effect, a reasonable stock price process should imply

$$\frac{m_3}{(m_2)^{3/2}} < 0, \qquad \frac{m_4}{m_2^2} > 3.$$

These two statistics may be checked in advance by employing CF under the risk-neutral measure or the original statistical measure.

[3] Strictly speaking, the moments are calculated under the risk-neutral measure Q. To obtain the actual probability measure of stock returns, we must change the risk-neutral parameters to their empirical counterparts.

2.2.2 Economic Interpretation of CFs

At first glance, expressing the probabilities F_1 and F_2 by the Fourier inversion of CFs $f_1(\phi)$ and $f_2(\phi)$ respectively is more technical than economic. In fact, an economic interpretation of CFs could be exploited by its implicit spanning of state space. This aspect gains significance especially with respect to market completeness and Arrow-Debreu securities in a state-space approach. Ross (1976) has proved that contingent claims (options) enhances market efficiency in a state-space framework since creating options can complete (or span) markets in an uncertain world. Bakshi and Madan (2000) demonstrated that spanning via options and spanning via CFs are completely interchangeable. To see this, we note that the CF $f(\phi;x(T)) \equiv \mathbf{E}^Q[\exp(i\phi x(T))]$ under the risk-neutral measure Q can be rewritten as follows:

$$\begin{aligned} f(\phi;x(T)) &\equiv \mathbf{E}^Q[\exp(i\phi x(T))] \\ &= \mathbf{E}^Q[\cos(\phi x(T)) + i\sin(\phi x(T))]. \end{aligned} \tag{2.33}$$

Now let us imagine two "contingent claims" of $x(T)$ with payoff at time T: $\cos(\phi x(T))$ and $\sin(\phi x(T))$.[4] Thus, from the point of view of valuation, $f(x(T);\phi)$ is a "security" consisting of two trigonometric assets of $x(T)$ while the stock price is only an exponential asset of $x(T)$. To show the spanning equivalence via options and via CFs, the following two equations hold:

$$\begin{aligned} \cos(\phi x(T)) &= 1 - \int_0^\infty \phi^2 \cos(\phi \ln K) \max(0, x(T) - \ln K) d\ln K, \\ \sin(\phi x(T)) &= x(T) - \int_0^\infty \phi^2 \sin(\phi \ln K) \max(0, x(T) - \ln K) d\ln K. \end{aligned}$$

These two equations state that trigonometric functions can be expressed by options. Hence, it follows that the spanned security space of options and their underlying is the same one spanned by $\cos(\phi x(T))$ and $\sin(\phi x(T))$. We can consider two probabilities F_1 and F_2 as Arrow-Debreu prices in the space spanned by options and the underlying primitive assets; Analogously, CFs $f_1(\phi)$ and $f_2(\phi)$ may be interpreted as the Arrow-Debreu prices in the transformed space.[5]

Starting from this equivalence, we can price more general contingent claims on $x(T)$ by using CFs. As a useful feature of CFs, differentiation and translation of CFs can simplify considerably the construction of primitive and derivative assets on $x(T)$. To expound this interesting feature of CF in the context of option pricing, we define a discounted CF (also called the CF of the remaining uncertainty):[6]

---

[4] The contingent claim with payoff $\cos(\phi x_T)$ or $\sin(\phi x_T)$ is an "asset" with unlimited liability since $\cos(\phi x_T)$ and $\sin(\phi x_T)$ can become negative.

[5] It is well-known that normal Arrow-Debreu prices should (must) be positive and smaller than one. Since the transformed spanned space is the complex plane, we can impose the usual $\mathscr{L}_1$-norm on f_1 and f_2. It follows immediately that $|f_1| \leq 1$ and $|f_2| \leq 1$. Thus, f_1 and f_2 are well-defined.

[6] See Bakshi and Madan (2000).

$$f^*(\phi;x(T)) = \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)\exp(i\phi x(T))\right]. \tag{2.34}$$

Setting $\phi = 0$, we have

$$f^*(0;x(T)) = \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)\right] = B(0,T). \tag{2.35}$$

Thus the CF $f_2(\phi)$ in (2.13) can be expressed by

$$f_2(\phi) = \frac{f^*(\phi;x(T))}{f^*(0;x(T))} = \frac{1}{B(0,T)}f^*(\phi;x(T)). \tag{2.36}$$

Note that

$$f^*(\phi - i;x(T)) = \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)\exp((1+i\phi)x(T))\right],$$

$$f^*(-i;x(T)) = \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)\exp(x(T))\right] = S_0,$$

we immediately obtain an alternative expression for $f_1(\phi)$,

$$f_1(\phi) = \frac{f^*(\phi - i;x(T))}{f^*(-i;x(T))} = \frac{f^*(\phi - i;x(T))}{S_0}, \tag{2.37}$$

where $x(T) = \ln S(T)$. As indicated in (2.28), we can construct the moments of the random variable $x(T)$ by partially differentiating the CF with respect to ϕ. In the same manner, partially differentiating the discounted CF with respect to ϕ yields the present value of the moments of $x(T)$. Generally, we have the following result on the relationship between the n-th order polynomial of $x(T)$ and the n-th order partial differential of $f^*(\phi;x(T))$,

$$\begin{aligned}(-i)^n\left[\frac{\partial^n f(\phi;x(T))}{\partial\phi^n}\right]_{\phi=0} &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)x^n(T)\exp(i\phi x(T))\right]_{\phi=0}\\ &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)x^n(T)\right]\\ &= (-i)^n f^*_{\phi n}(0;x(T)).\end{aligned} \tag{2.38}$$

It follows

$$H_{x(T)} = \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)x(T)\right] = \frac{1}{i}f^*_{\phi}(0;x(T)), \tag{2.39}$$

where $H_{x(T)}$ is the present value of $x(T)$. Note that a real payoff function $L(x(T)) \in \mathbf{C}^\infty$ can be sufficiently approximated by a (Taylor-) polynomial series, and hence the expected discounted price of $L(x(T))$ can be approximated by a series of $f^*_{\phi n}(0;x(T))$. This approach states that $f^*_{\phi n}(0;x(T))$ forms a polynomial basis for the market of contingent claims on $x(T)$. Without loss of generality, we denote the spot price or the forward price of the contingent claim $L(x(T))$ by L_0 and expand $L(x(T))$ around zero,

$$\begin{aligned} G(T,\phi) &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)L(x(T))\exp(i\phi x(T))\right] \\ &= f^*(\phi;x(T))L(0) + \sum_{k=1}^{\infty}\frac{f^*_{\phi k}(\phi;x(T))}{(i)^k k!}\frac{\partial^k L}{\partial x^k}(0) \end{aligned} \tag{2.40}$$

and

$$\begin{aligned} G(T,0) = L_0 &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)L(x(T))\right] \\ &= f^*(0;x(T))L(0) + \sum_{k=1}^{\infty}\frac{f^*_{\phi k}(0;x(T))}{(i)^k k!}\frac{\partial^k L}{\partial x^k}(0). \end{aligned} \tag{2.41}$$

In many cases, for example, in the case of stock options, the spot price L_0, i.e., $G(T,0)$ is directly observable and does not have to be derived from the CFs. $G(T,0)$ plays an important role in valuation only for the case where no spot prices or forward prices of the contingent claim $L(x(T))$ are quoted in the markets. Using L_0 as numeraire, we can construct the first CF,

$$\begin{aligned} f_1(\phi) &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)\frac{L(x(T))}{L_0}\exp(i\phi x(T))\right] \\ &= \frac{G(T,\phi)}{G(T,0)}. \end{aligned} \tag{2.42}$$

To guarantee $f_1(\phi)$ to be a well-defined CF, a sufficient and necessary condition must be satisfied

$$\mathbf{E}^Q\left[\exp\left(-\int_0^T r(t)dt\right)\frac{L(x(T))}{L_0}\right] = 1. \tag{2.43}$$

Equation (2.43) implies that $L(x)$ is an arbitrage-free process and admits no free-lunch. The construction of CF $f_2(\phi)$ for pricing general contingent claims follows a similar way. Therefore, we are able to derive closed-form valuation formulas for a call on particular contingent claims in an unified framework. A simple example can be an option on interest rates if we specify $x(t)$ as $r(t)$. Summing up, we have a more general principle for constructing CFs:[7]

Theorem 2.2.1 *The Principle for Constructing Characteristic Functions 2: Assume a (primitive) asset return process $x(t)$ with $x(0) = 0$ follows an Itô process and satisfies the usual necessary technical conditions. In the market there is a basic security $L(x) \in \mathbf{C}^\infty$ on this return, which is strictly positive and invertible. Furthermore, the arbitrage-free condition (2.43) is satisfied. Then a European-style call option C on $L(x)$ with a strike price K and a maturity time of T can be valued by*

[7] This principle is essentially identical to the case 3 in Bakshi and Madan (2000).

$$C(x_0,T;K) = L_0F_1 - KB(0,T)F_2$$

with

$$F_j = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_j(\phi)\frac{\exp(-i\phi L^{-1}(K))}{i\phi}\right)d\phi, \quad j=1,2,$$

where

$$L_0 = G(T,0), \qquad B(0,T) = f^*(0;x(T)),$$

$$f_1(\phi) = \frac{G(T,\phi)}{G(T,0)}, \qquad f_2(\phi) = \frac{f^*(\phi;x(T))}{f^*(0;x(T))},$$

$$G(T,\phi) = f^*(\phi;x(T))L(0) + \sum_{k=1}^{\infty}\frac{f^*_{\phi k}(\phi;x(T))}{(i)^k k!}\frac{\partial^k L}{\partial x^k}(0).$$

This generalized principle for constructing CFs is useful in deriving a tractable valuation formula for options on securities with an unconventional payoff function $L(x)$ and provides us with a structural and amenable tool to analyze quantitative effects of uncertainty in economics and finance.

2.2.3 Examination of Existing Option Models

Since Black and Scholes (1973) derived their celebrated option pricing formula, an enormous number of variants of option pricing models have been developed. Merton (1973) examined an option valuation problem with stochastic interest rates using the assumption that the zero-coupon bond price is lognormally distributed. Geske (1979) gave a pricing formula for compound options. Merton (1976), Cox and Ross (1976) studied issues of option valuation where the underlying asset price follows a jump process and introduced for the first time a discontinuous price process in option pricing theory. Since the middle of 1980s, valuation with stochastic volatilities has drawn more and more the attention of financial economists. The models of Stein and Stein (1990), Heston (1993), BCC (1997), as well as Schöbel and Zhu (1999) are the most influential works on modeling stochastic volatility. Recently, pure jumps characterized by Lévy process are introduced to describe leptokurtic features of stock returns in the hope of recovering the smile of implied volatility. The representatives of option pricing models with Lévy process are the variance-Gamma model of Madan, Carr and Chang (1998), and the normal inverse Gaussian model of Barndorff-Nielsen (1998). Here, we give a brief overview of some important closed-form pricing solutions for options under the risk-neutral measure using the Fourier transform techniques. The detailed discussions for these models will be given in the following chapters.

(1). The Black-Scholes Model (1973). As mentioned in Chapter 1, it includes only one stochastic process

$$dS(t) = rS(t)dt + \sigma S(t)dW(t).$$

Let $x(t) = \ln S(t)$,

$$dx(t) = (r - \frac{1}{2}\sigma^2)dt + \sigma dW(t). \tag{2.44}$$

Hence,

$$g_1(T) = \exp(-rT + x(T) - x_0), \qquad g_2(T) = 1.$$

The two CFs can be calculated as follows:

$$\begin{aligned} f_1(\phi) &= \mathbf{E}^Q\left[g_1(T)e^{i\phi x(T)}\right] = \mathbf{E}^Q\left[\exp((1+i\phi)x(T) - rT - x_0\right] \\ &= \exp\left[i\phi(rT + x_0) - \frac{1+i\phi}{2}\sigma^2 T + \frac{(1+i\phi)^2}{2}\sigma^2 T\right] \\ &= \exp\left[i\phi\left(rT + x_0 + \frac{1}{2}\sigma^2 T\right) - \frac{\phi^2}{2}\sigma^2 T\right] \\ &= \exp\left(i\phi(x_0 + rT)\right) f_1^{BS}(\phi) \end{aligned} \tag{2.45}$$

with

$$f_1^{BS}(\phi) = \exp\left[i\phi\left(\frac{1}{2}\sigma^2 T\right) - \frac{\phi^2}{2}\sigma^2 T\right]. \tag{2.46}$$

and

$$\begin{aligned} f_2(\phi) &= \mathbf{E}^Q\left[g_2(T)e^{i\phi x(T)}\right] = \mathbf{E}^Q\left[\exp i\phi x(T)\right] \\ &= \exp\left[i\phi\left(rT + x_0 - \frac{1}{2}\sigma^2 T\right) - \frac{\phi^2}{2}\sigma^2 T\right] \\ &= \exp\left(i\phi(x_0 + rT)\right) f_2^{BS}(\phi) \end{aligned} \tag{2.47}$$

with

$$f_2^{BS} = \exp\left[i\phi\left(-\frac{1}{2}\sigma^2 T\right) - \frac{\phi^2}{2}\sigma^2 T\right]. \tag{2.48}$$

It is not hard to verify that

$$N(d_j) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_j(\phi)\frac{\exp(-i\phi\ln K)}{i\phi}\right)d\phi, \quad j = 1, 2.$$

$N(d_j)$ are the usual probability terms in the Black-Scholes formula.[8] Thus, we have shown that the Black-Scholes formula can be easily derived by using the Fourier inversion.

[8] Since the CF of a normal density function $n(\mu, \sigma)$ is $f(\phi) = \exp\left(i\phi\mu - \frac{1}{2}\sigma^2\phi^2\right)$, two probabilities $N(d_j), j = 1, 2$, in the Black-Scholes formula are Gaussian with cumulative distribution functions $N(rT + x_0 \pm \frac{1}{2}\sigma^2 T, \frac{1}{2}\sigma^2 T)$ respectively.

(2). The Heston Model (1993). This model is characterized by the following two processes:

$$\begin{aligned}\frac{dS(t)}{S(t)} &= r(t)dt + \sqrt{V(t)}dW_1(t)\\ &\text{or}\\ dx(t) &= [r(t) - \frac{1}{2}V(t)]dt + \sqrt{V(t)}dW_1(t),\\ dV(t) &= [\kappa\theta - (\kappa+\lambda)V(t)]dt + \sigma\sqrt{V(t)}dW_2(t)\end{aligned} \tag{2.49}$$

with

$$dW_1(t)dW_2(t) = \rho dt.$$

Here the squared volatilities follow a mean-reverting square-root process which was used in finance for the first time by CIR (1985b) to specify stochastic interest rates. Hence, we have

$$g_1(T) = \exp(-rT + x(T) - x_0), \qquad g_2(T) = 1.$$

The detailed derivation of this pricing formula for options will be given in the next chapter.

(3). The model of Bakshi, Cao and Chen (1997). This model includes stochastic volatility, stochastic interest rates and random jumps where the first two factors are specified by a mean-reverting square root process. Random jumps are expressed by Poisson process with lognormally distributed jump sizes.

$$\begin{aligned}\frac{dS(t)}{S(t)} &= [r(t) - \lambda\mu_J]dt + \sqrt{V(t)}dW_1(t) + J(t)dY(t)\\ &\text{or}\\ dx(t) &= [r(t) - \lambda\mu_J - \frac{1}{2}V(t)]dt + \sqrt{V(t)}dW_1(t) + \ln[1+J(t)]dY(t),\\ dV(t) &= \kappa_V[\theta_V - V(t)]dt + \sigma_V\sqrt{V(t)}dW_2(t),\\ dr(t) &= \kappa_r[\theta_r - r(t)]dt + \sigma_r\sqrt{r(t)}dW_3(t)\end{aligned} \tag{2.50}$$

with

$$\begin{aligned}dW_1(t)dW_2(t) &= \rho dt,\\ \ln[1+J(t)] &\sim N(\ln[1+\mu_J] - \frac{1}{2}\sigma_J^2, \sigma_J^2),\\ dY(t) &\sim \delta_1\lambda dt + \delta_0(1-\lambda)dt,\end{aligned}$$

where δ_n is an indicator function for the value n. Here the stock price process includes a jump component $J(t)dq(t)$. Since the jump has no impact on the Itô process, after taking the market price of the jump risk $\lambda\mu_J$ into account, we can construct the same $g_j(t), j=1,2$, as the ones without jumps. So we have

$$g_1(T) = \exp\left(-\int_0^T r(t)dt + x(T) - x_0\right),$$
$$g_2(T) = \exp\left(-\int_0^T r(t)dt\right)/B(0,T).$$

We will show that it is very convenient to derive the option pricing formula in this complicated model by applying the idea of modular pricing.[9] Scott (1997) developed a similar model incorporating stochastic interest rates and jumps.

(4). The Bates (1996) model for currency option. This model differs only slightly from the above one due to a different process for the underlying asset. Let $S(t)$ denote the exchange rate,

$$\begin{aligned} \frac{dS(t)}{S(t)} &= [r - r^* - \lambda\mu_J]dt + \sqrt{V(t)}dW_1(t) + J(t)dY(t) \\ &\text{or} \\ dx(t) &= [r - r^* - \lambda\mu_J - \frac{1}{2}V(t)]dt + \sqrt{V(t)}dW_1(t) + \ln[1+J(t)]dY(t), \\ dV(t) &= [\theta_V - \kappa_V V(t)]dt + \sigma_V\sqrt{V(t)}dW_2(t), \end{aligned} \tag{2.51}$$

with

$$\begin{aligned} dW_1(t)dW_2(t) &= \rho dt, \\ \ln[1+J(t)] &\sim N(\ln[1+\mu_J] - \frac{1}{2}\sigma_J^2, \sigma_J^2), \\ dY(t) &\sim \delta_1\lambda dt + \delta_0(1-\lambda)dt. \end{aligned}$$

Thus it follows

$$g_1(T) = \exp(-(r-r^*)T + x(T) - x_0), \qquad g_2(T) = 1.$$

The above models are some representatives in the vast option pricing literature. All of these models (except for the Black-Scholes model) have some common properties: First of all, stock prices (underlying assets) are characterized only by a Brownian motion with a correlated process for variance. Squared volatility (variance) follows a square-root process. Secondly, the process of short rates is one-dimensional and also follows a square root process. As shown later, a square root process allows for a nice derivation of characteristic functions, but, obviously, it is not the only way to describe volatilities and interest rates. The Ornstein-Uhlenbeck process, for example, may be an appealing alternative (Vasicek (1977), Stein and Stein (1991), Schöbel and Zhu (1999)). Furthermore, we can extend the one-dimensional stock

[9] The original derivation of the option pricing formula given by Bakshi, Cao and Chen is based on the PDEs, and rather complicated.

price process to a multi-dimensional case. We will face such multi-dimensional problems in pricing basket options and correlation options.

2.2.4 Relationship between CF to PDE

In this section, we will show that there is a relationship between the Fourier transform approach and the partial differential equation (PDE) approach to asset pricing. As shown in Chapter 1, a PDE called the Black-Scholes equation is established to derive option pricing formula, and can be applied to value any derivative within the Black-Scholes framework if the corresponding boundary and terminal conditions are imposed. In fact, the original works of Heston (1993), Bates (1996), BCC (1997) are based on a multi-dimensional PDE-approach with which particular option pricing formula may be given. However, a closed-form solution to a multi-dimensional PDE is generally not trivial, and we can not directly utilize some useful results of stochastic calculus if a PDE is applied to value options. Now we take stochastic volatility models, particularly the Heston model and the the BCC model, as examples to demonstrate the relationship between the CF and the PDE by applying Kolmogorov's backward equation.

Given $\mathbf{y}$ as a n-dimensional Itô process and a function $h \in \mathfrak{C}(\mathbb{R}^2)$, then the Kolmogorov's backward equation states:

$$\frac{\partial u}{\partial \tau} = \mathscr{K} u \quad \text{for} \quad u = \mathbf{E}^P\left[h(\tau, \mathbf{y})\right], \quad \tau \geqslant 0, \quad \mathbf{y} \in \mathbb{R}^n$$

with the boundary condition

$$u(0, \mathbf{a}) = h(0, \mathbf{a}), \quad \mathbf{y}(0) = \mathbf{a} \in \mathbb{R}^n,$$

and an operator $\mathscr{K}$ defined by[10]

$$\mathscr{K} = \sum_j \text{drift}_j \frac{\partial}{\partial y_j} + \frac{1}{2} \sum_{i,j} \text{cov}_{ij} \frac{\partial^2}{\partial y_i y_j}.$$

Now denote $\tau = T$ for the notation consistence. We have already known that CFs $f_1(\phi) = \mathbf{E}^{Q_1}\left[\exp(i\phi x(T))\right]$ and $f_2(\phi) = \mathbf{E}^{Q_2}\left[\exp(i\phi x(T))\right]$ are the CFs associated with the probability measures Q_1 and Q_2 respectively, and both twice differentiable, therefore belong to the function class $\mathfrak{C}(\mathbb{R}^2)$. Furthermore, assume $f_1(\phi)$ and $f_2(\phi)$ have the following functional form

$$f_1(\phi) = f_1(x_0 = \ln S_0, V_0, T; \phi), \quad \mathbf{a} = [x_0, V_0]$$

and

$$f_2(\phi) = f_2(x_0 = \ln S_0, V_0, T; \phi), \quad \mathbf{a} = [x_0, V_0].$$

[10] $\mathscr{K}$ is the generator of an Itô diffusion.

Applying the Girsanov theorem given in Section 2.1, $x(t)$ in the Heston model under the measures Q_1 follows a process

$$dx(t) = \left(r + \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_1^{Q_1},$$

and the stochastic variance $V(t)$ follows a mean-reverting square root process

$$dV(t) = \left[\kappa[\theta - V(t)] + \rho\sigma V(t)\right]dt + \sigma\sqrt{V(t)}dW_2^{Q_1}.$$

In this model setup, we have then $\mathbf{y} = \{x(t), V(t)\}$. Hence, a direct application of the operator $\mathscr{K}$ yields

$$\begin{aligned}\mathscr{K} f_1 = & \frac{1}{2}V\frac{\partial^2 f_1}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_1}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_j}{\partial V^2} \\ & + \left(r + \frac{1}{2}V\right)\frac{\partial f_1}{\partial x} + [\kappa(\theta - V) + \rho\sigma V]\frac{\partial f_1}{\partial V}.\end{aligned}$$

As a result, we obtain the corresponding Kolmogorov's backward equation under the measure Q_1,

$$\begin{aligned}\frac{\partial f_1}{\partial T} = & \frac{1}{2}V\frac{\partial^2 f_1}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_1}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_1}{\partial V^2} \\ & + \left(r + \frac{1}{2}V\right)\frac{\partial f_1}{\partial x} + [\kappa(\theta - V) + \rho\sigma V]\frac{\partial f_1}{\partial V},\end{aligned} \tag{2.52}$$

with which the CF $f_1(\phi)$ may be derived for initial condition.

If the risk-free interest rate r is constant as in the Heston model, we have shown that Q_2 is identical to Q, this means that $x(t)$ and $V(t)$ in the Heston model under the measures Q_2 are governed by the original risk-neutral processes respectively

$$dx(t) = \left(r - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_1^{Q_2}$$

and

$$dV(t) = \kappa[\theta - V(t)]dt + \sigma\sqrt{V(t)}dW_2^{Q_2}.$$

It follows the corresponding Kolmogorov's backward equation under the measure Q_2,

$$\begin{aligned}\frac{\partial f_2}{\partial T} = & \frac{1}{2}V\frac{\partial^2 f_2}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_2}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_2}{\partial V^2} \\ & + \left(r - \frac{1}{2}V\right)\frac{\partial f_2}{\partial x} + \kappa(\theta - V)\frac{\partial f_2}{\partial V}.\end{aligned} \tag{2.53}$$

This is the PDE to solve the CF $f_2(\phi)$.

Of course, if the instantaneous variance $V(t)$ is constant, the Heston model reduces to the Black-Scholes model, the CFs $f_j(\phi), j = 1, 2$, will be simplified to

$$f_j(\phi) = f_j(x_0, T; \phi), \quad \mathbf{a} = [x_0].$$

It is not hard to obtain the similar Black-Scholes equations for $f_j(\phi)$.

In the above mentioned BCC model where the instantaneous interest rate $r(t)$ is specified to follow a mean-reverting square root process, we shall extend $f_j(\phi)$ by

$$f_1(\phi) = f_1(x_0, V_0, r_0 T; \phi), \quad \mathbf{a} = [x_0, V_0, r_0]$$

and

$$f_2(\phi) = f_2(x_0, V_0, r_0, T; \phi), \quad \mathbf{a} = [x_0, V_0, r_0].$$

Since $r(t)$ is uncorrelated with $x(t)$ in the BCC model, the change of measure Q to Q_1 does not affect the process of $r(t)$, namely

$$dr(t) = \kappa_r[\theta_r - r(t)]dt + \sigma_r\sqrt{r(t)}dW_3^{Q_1}(t). \tag{2.54}$$

However, in this case, the change of measure Q to Q_2 will change also the drift term of the process of $r(t)$. More precisely, under the forward T-measure Q_2, we have an adjusted process for $r(t)$,

$$dr(t) = [\kappa_r(\theta_r - r(t)) + \sigma_r^2 b(t,T)]dt + \sigma_r\sqrt{r(t)}dW_3^{Q_2}(t) \tag{2.55}$$

with an exponential form for a zero-coupon price $B(t,T)$,

$$B(t,T) = e^{a(t,T)+b(t,T)r(t)}.$$

Based on these two different processes under two different measures, we can establish the corresponding PDEs for $f_j(\phi)$ in the BCC model as follows:

$$\begin{aligned}\frac{\partial f_1}{\partial T} &= \frac{1}{2}V\frac{\partial^2 f_1}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_1}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_1}{\partial V^2} \\ &+ \left(r + \frac{1}{2}V\right)\frac{\partial f_1}{\partial x} + [\kappa(\theta - V) + \rho\sigma V]\frac{\partial f_1}{\partial V} \\ &+ \frac{1}{2}\sigma_r^2 r\frac{\partial^2 f_1}{\partial r^2} + \kappa_r(\theta_r - r)\frac{\partial f_1}{\partial r}\end{aligned} \tag{2.56}$$

and

$$\begin{aligned}\frac{\partial f_2}{\partial T} &= \frac{1}{2}V\frac{\partial^2 f_2}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_2}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_2}{\partial V^2} \\ &+ \left(r + \frac{1}{2}V\right)\frac{\partial f_2}{\partial x} + \kappa(\theta - V)\frac{\partial f_2}{\partial V} \\ &+ \frac{1}{2}\sigma_r^2 r\frac{\partial^2 f_2}{\partial r^2} + [\kappa_r(\theta_r - r) + \sigma_r^2 b(t,T)]\frac{\partial f_2}{\partial r}.\end{aligned} \tag{2.57}$$

Therefore in the sense of Kolmogorov's backward equation, an equivalent relationship between the Fourier transform approach and the PDE approach can be established. If a process under a particular measure is given, then the PDE for its CF is just the corresponding Kolmogorov's backward equation.

2.2.5 Advantages of CF and Modular Pricing

In general, when we compare the Fourier transform approach with the traditional pricing method, a natural question arises: what are the advantages of the Fourier transform approach by using the CF to value European-style options? The answer to this question is at least fourfold:

1. The first gain is the tractability and easy manipulation of CFs which allows us to arrive at closed-form solutions in terms of CFs. In contrast, by applying the PDE approach it is difficult to make variable transformations to get closed-form pricing solutions. As shown above, due to the equivalence between the Fourier transform approach and the PDE approach, PDEs that are relevant to option pricing can be solved, at least numerically. However, if our focus is to derive the closed-form pricing solution for options, numerical solutions are no longer useful. In many cases, it is not easy to get a "right" guess for an analytical solution for CFs by using the PDE approach as in the Heston model. For example, with the volatility specified as a mean-reverting Ornstein-Uhlenbeck process, we will encounter problems to find analytical solutions if the PDE method is directly applied. In contrast, if we first calculate and manipulate the CFs to reduce dimensionality, then apply the PDEs as done in Scott (1997), Schöbel and Zhu (1999), we can indeed arrive at closed-form solutions for the CFs. As shown in the following chapters, a direct calculation of the CF may be applied broadly to almost all option pricing models with few restrictions, even in the present of jump components. Therefore, one can sometimes only solve a PDE numerically although a closed-form solution for the CF does exist.
2. Secondly, as demonstrated by the Heston model, the Fourier transform method provides us with a powerful tool to give a new form of option pricing formula. In fact, Fourier transform is often the only elegant and compact way to deal with other stochastic volatility models, as shown in Lewis (2000) and Zhu (2000). In a similar manner, stochastic interest rates can be incorporated into an option pricing model conveniently via CFs, as shown in BCC (1997) and, Dai and Singleton (2000). As shown later, the pricing formulas both for stochastic volatility models and for stochastic interest rate models share a common pricing kernel as long as they are specified to follow the same process. This point becomes especially more obvious by using Fourier analysis and was discussed in Madan and Bakshi (2000), Goldstein (1997), as well as Carr and Wu (2004). Consequently, we could incorporate the dynamics of interest rates into an option pricing model more flexibly and efficiently in the framework of Fourier transform. In fact, almost new researches on option pricing since 2000 are based on the technique of Fourier transform. The broad applications of Fourier transform revolutionized the valuation theory and practice.

3. Thirdly, CFs are always continuous in their domain even for a discontinuous distribution. This property is of great advantage when we consider to model jumps in an asset price process. The classical approach to dealing with discontinuous phonomania in stock price movements is to describe jump frequency with Poisson process and to describe jump size with some independent distributions. Additionally, more general jumps characterized by Lévy measures, for example, the jumps driven by a variance-gamma process (Madan, Carr and Chang (1999)), and a normal inverse Gaussian process (Barndorff-Nielsen (1998)), are applied to capture the dynamics of stock prices. We will show that CF is also a convenient and elegant way to arrive at closed-form option pricing formulas in all theses jumps models. Indeed, in these so-called Lévy jump models, there exists no general PDE for the relevant CF.
4. Finally, Fourier transform allows us to build up a flexible option pricing framework wherein each stochastic factor can be considered as a module. The idea of this building-block approach is referred to as the modular pricing of options in this book. In the light of the property of CF given by (2.27), the CFs in a comprehensive option pricing model are simply composed of the CFs of independent stochastic factors. Roughly and intuitively, we have

$$f(\phi) = f_{SV}(\phi) f_{SI}(\phi) f_{Jump}(\phi),$$

where SV, SI and $Jump$ represent the stochastic volatility, the stochastic interest rate and the jumps respectively, which again could be specified by various stochastic processes. Hence, the Fourier transform enables us to develop a number of option pricing models by simply combining different stochastic factors. In this sense, stochastic volatilities, stochastic interest rates and random jumps work as modules, and can be withdrawn from or inserted into a pricing formula by assembling the corresponding CFs. However, all pricing formulas for options have a common shell: the Fourier inversion given in (2.18).

From the point of view of financial engineering, the immediate gain of this modular concept is its convenient implementation. For example, we can define three basic classes for stochastic volatilities, stochastic interest rates and jumps. The particular specification of these stochastic factor may be implemented in the derived classes. Therefore, the implementation of a special new pricing model is just a simple assemble of three concrete classes where the corresponding CFs are implemented.

Chapter 3
Stochastic Volatility Models

In this chapter we discuss stochastic volatility models that play a crucial role in smile modeling not only on the theoretical side, but also on the practical side. Stochastic volatility models are extensively examined by financial researchers, and widely used in investment banks and financial institutions. In many trading floors, this type of option pricing model has replaced the Black-Scholes model to be the standard pricing engine, especially for exotic derivatives. Intuitively, it is a natural way to capture the volatility smile by assuming that volatility follows a stochastic process. Simple observations propose that the stochastic process for volatility should be stationary with some possible features such as mean-reverting, correlation with stock dynamics. In this sense, a mean-reverting square root process and a mean-reverting Ornstein-Uhlenbeck process adopted from the interest rate modeling are two ideal candidate processes for stochastic volatilities. Heston (1993) specified stochastic variances with a mean-reverting square root process and derived a pioneering pricing formula for options by using CFs. Stochastic volatility model with a mean-reverting Ornstein-Uhlenbeck process is examined by many researchers. Schöbel and Zhu (1999) extended Stein and Stein's (1991) solution for a zero-correlation case to a general non-zero correlation case. In this chapter, we will expound the basic skills to search for a closed-form formula for options in a stochastic volatility model. Generally, we have two approaches available: the PDE approach and the expectation approach, both are linked via the Feynman-Kac theorem. However, as demonstrated in following, the expectation approach utilizes the techniques of stochastic calculus and is then an easier and more effective way. To provide an alternative stochastic volatility model, we also consider a model with stochastic variances specified as a mean-reverting double square root process.

3.1 Introduction

Since the Black-Scholes formula was derived, a number of empirical studies have concluded that the assumption of constant volatility is inadequate to describe stock

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance, DOI 10.1007/978-3-642-01808-4_3,

returns, based on two findings: (1) volatilities of stock returns vary over time, but persist in a certain level (mean-reversion property). This finding can be traced back to the empirical works of Mandelbrot (1963), Fama (1965) and Blattberg and Gonedes (1974) with the results that the distributions of stock returns are more leptokurtic than normal; (2) volatilities are correlated with stock returns, and more precisely, they are usually inversely correlated. Furthermore, volatility smile provides a direct evident for inconsistent volatility pattern with moneyness in the Black-Scholes model. Black (1976a), Schmalensee and Trippi (1978), Beckers (1980) investigated the time-series property between stock returns and volatilities, and found an imperfect negative correlation. Bakshi, Cao and Chen (BCC, 1997) and Nandi (1998) also reported a negative correlation between the implied volatilities and stock returns. Moreover, Beckers (1983), Pozerba and Summers (1984) gave evidence that shocks to volatility persist and have a great impact on option prices, but tend to decay over time. These uncovered properties associated with volatility such as leptokurtic distribution, correlation, mean-reversion and persistence of shocks, should be considered in a suitable option pricing model.

In order to model the variability of volatility and to capture the volatility smile, several approaches have been suggested. One approach was the so-called constant elasticity of variance diffusion model developed by Cox (1975), and may be regarded as a representative of a more general model class labeled to local volatility model. Cox assumed that volatility is a function of the stock price with the following form:

$$v(S(t)) = aS(t)^{\delta-1}, \quad \text{with} \quad a > 0, \quad 0 \leqslant \delta \leqslant 1. \tag{3.1}$$

Since $v(S(t))$ is a decreasing function of $S(t)$, volatilities are inversely correlated with stock returns. However, this deterministic function can not describe other desired features of volatility. Derman and Kani (1994), Dupire (1994) and Rubinstein (1994) hypothesized that volatility is a deterministic function of the stock price and time, and developed a deterministic volatility function (DVF), with which they attempt to fit the observed cross-section of option prices exactly. This approach, as reported by Dumas, Fleming and Whaley (1998), does not perform better than an *ad hoc* procedure that merely smooths implied volatilities across strike prices and times to maturity.

A more general approach is to model volatility by a diffusion process and has been examined Johnson and Shanno (1987), Wiggins (1987), Scott (1987), Hull and White (1987), Stein and Stein (1991), Heston (1993), Schöbel and Zhu (1999) and Lewis (2000). The models following this approach are the so-called stochastic volatility models. Table (3.1) gives an overview of some representatives of stochastic volatility models.

There are no known closed-form option pricing formulas in the case of non-zero correlation between volatilities and stock returns for all models listed in Table 3.1 except for model (5) [Schöbel and Zhu, 1999], model (6) [Heston, 1993] as well as model (7). Models (1) and (3) perform no mean-reversion property and therefore can not capture the effects of shocks to volatility in the valuation of options. Models (1), (2), (3) (4) and (8) are not stationary processes and violate the feature of stationarity

(1):	$dv(t) = \kappa v dt + \sigma v dW(t)$	Johnson/Shanno (1987)
(2):	$d\ln v(t) = \kappa(\theta - \ln v)dt + \sigma dW(t)$	Wiggins (1987)
(3):	$dv(t)^2 = \kappa v^2 dt + \sigma v^2 dW(t)$	Hull/White (1987)
(4):	$dv(t)^2 = \kappa v^2(\theta - v)dt + \sigma v^2 dW(t)$	Hull/White (1987)
(5):	$dv(t) = \kappa(\theta - v)dt + \sigma dW(t)$	Stein and Stein (1991), Schöbel/Zhu (1999)
(6):	$dV(t) = \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dW(t)$	Heston (1993)
(7):	$dV(t) = \kappa(\theta - \sqrt{V(t)} - \lambda V(t))dt + \sigma\sqrt{V(t)}dW(t)$	suggested in this book
(8):	$dV(t) = \kappa(\theta - V^2(t))dt + \sigma V(t)^{3/2}dW(t)$	Lewis (2000)

Table 3.1 Overview of one-factor stochastic volatility models.

of volatility or variance. The so-called $3/2$ model of Lewis (2000) is unbounded, and as a result, option prices under the $3/2$ model are not martingales, but only local martingales, therefore not appropriate for practical applications. Consequently, only models (5) and (6) are reasonable and worthwhile to study in details.

As shown in Hull and White (1987), Ball and Roma (1994), if stock returns are uncorrelated with volatilities, the option price can be approximated very accurately by replacing the variance in the Black-Scholes formula by the expected average variance which is defined by

$$AV = \mathbf{E}\left(\frac{1}{T}\int_0^T v(t)^2 dt\right). \tag{3.2}$$

The mathematical background for this replacement is that the conditional distribution of terminal stock prices, given the expected average variance, is still lognormal. However, this argument fails in the presence of the correlation between stock returns and volatilities. Heston (1993) showed that the correlation between volatility and spot returns is necessary to create negative skewness and excess kurtosis in a leptokurtic distribution of spot returns. BCC (1997), Heston and Nandi (1997), and Nandi (1998) demonstrated that non-zero correlation is important in eliminating the smile effect of implied volatilities and leads to significant improvements both in pricing and in hedging options.

Some econometric models also attract attention when studying stochastic volatility. In fact, each model from (1) to (7) in Table (3.1) may correspond to an econometric counterpart by discretizating them on time points, and some could be referred to as the so-called autoregressive random variance models (ARV). Nelson and Foster (1994), Duan (1996) showed that some existing stochastic volatility models can be considered as the week limits of some generalized autoregressive conditional heteroscedasticity (GARCH) models.[1] The discrete-time versions of stochastic models play also an important role in empirical tests. For instances, Heston and Nandi (1997) suggested a GARCH option pricing model and derived a closed-form so-

[1] The crucial difference between ARV models and GRACH models is that ARV models include two innovations to generate data while GRACH models have only one innovation with conditional variances dependent upon the past information.

lution which allows for a correlation between stock returns and variance and even admits multiple lags in the GARCH process. However, since we focus on continuous time valuation in this book, econometric models will not be in consideration here.

3.2 The Heston Model

3.2.1 Model Setup and Properties

The Heston model (1993) is the first stochastic volatility model that systematically deal with the valuation of options with CFs and provides a quasi closed-form solution for options. Heston does not model stochastic volatility directly, but stochastic variance. In this section, we denote $V(t) = v^2(t)$ as variance. The risk-neutral Heston model consists of the following two processes[2]

$$\begin{aligned} \frac{dS(t)}{S(t)} &= rdt + \sqrt{V(t)}dW_1(t), \\ dV(t) &= \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dW_2(t), \\ dW_1(t)dW_2(t) &= \rho dt. \end{aligned} \tag{3.3}$$

The process specifying the variance is identical to the one that Cox-Ingersoll-Ross apply to model short interest rates, and is called mean-reverting square root process. The mean-reversion is a desired property for stochastic volatility or variance and is well documented by many empirical studies. The parameter θ marks the long-term level that variance gradually converges to. The parameter κ controls the speed of variance's reverting to θ. The parameter σ is refereed to as the volatility of variance since it scales the diffusion term of variance process.[3]

As noted by Heston, the volatility $v(t)$ in the Heston model initially follows an Ornstein-Uhlenbeck process with a mean-reverting level as zero, that is

$$dv(t) = -\beta v(t)dt + \delta dW_2(t). \tag{3.4}$$

[2] To obtain a risk-neutral process for variance, a market price of volatility risk λ must be introduced, which depends on the assumption about the risk aversion. A common way to determine λ is to assume $\lambda dt = \gamma Cov[dv, dC/C]$ where γ and C is the relative-risk aversion and consumption respectively. From the well-known Cox, Ingersoll and Ross (1985a) equilibrium model, one can get a consumption process [also see equation (8) in Heston]:

$$dC = \mu_c V(t)Cdt + \sigma_c\sqrt{V(t)}CdW_c(t).$$

Consequently, the risk premium is proportional to $V(t)$, $\lambda(V) = \lambda V$ with λ a constant. With the adjustment, the original process and adjusted process have the same structure.

[3] The true volatility of variance should be $\sigma\sqrt{V(t)}$ since it is the coefficient of the diffusion term.

So by applying Itô's lemma, the square of volatility $V(t) = v^2(t)$, namely the variance of the instantaneous stock return, follows a square-root process which is given by

$$dV(t) = \kappa[\theta - V(t)]dt + \sigma\sqrt{V(t)}dW_2(t) \tag{3.5}$$

with

$$\beta = \frac{\kappa}{2}, \quad \delta = \frac{\sigma}{2}, \quad \theta = \frac{\delta^2}{\kappa}. \tag{3.6}$$

If κ, θ and σ satisfy the following conditions,

$$2\kappa\theta > \sigma^2, \quad V_0 > 0, \tag{3.7}$$

it can be shown that variances $V(t)$ are always positive and the variance process given in (3.5) is then well-defined under above condition. The condition (3.7) is also referred to as Feller's condition for square root process. Let $\chi(x,d,n)$ denote a non-central chi-square distribution with d as the degrees of freedom and n as non-centrality parameter. The variance $V(t)$ conditional on $V(s)$ is distributed as a non-central chi-square distribution $F_{\chi^2}(aV(t),d,n)$ where

$$\begin{aligned}
d &= \frac{4\kappa\theta}{\sigma^2}, \\
n &= \frac{4\kappa e^{-\kappa(t-s)}}{\sigma^2\left(1 - e^{-\kappa(t-s)}\right)}, \\
a &= \frac{n}{e^{-\kappa(t-s)}} = \frac{4\kappa}{\sigma^2\left(1 - e^{-\kappa(t-s)}\right)}.
\end{aligned}$$

Based on the properties of a non-central chi-square distribution, $V(t)$ has the following two conditional moments,

$$\mathbf{E}[V(t)|\,V(s)] = \theta + (V(s) - \theta)e^{-\kappa(t-s)}, \tag{3.8}$$

$$\begin{aligned}
\mathbf{Var}[V(t)|\,V(s)] &= V(s)\frac{\sigma^2}{\kappa}e^{-\kappa(t-s)}\left[1 - e^{\kappa(t-s)}\right] \\
&\quad + \frac{\sigma^2\theta}{2\kappa}\left[1 - e^{\kappa(t-s)}\right]^2.
\end{aligned} \tag{3.9}$$

As mentioned above, the positive values of the square root process can be only guaranteed under the condition (3.7) which unfortunately is not always given in practical applications. Even under the condition (3.7), if we simulate a mean-reverting square root process, it is very possible to generate some negative paths due to the nature of Monte-Carlo simulation, and the negative $V(t)$ makes the computation of $\sqrt{V(t)}dW_2(t)$ impossible and leads to error. Hence, a satisfactory simulation of the mean-reverting square root process is a challenging issue in computational finance. We will return to this issue in Chapter 5 with more detailed discussions.

Again applying Itô's lemma, we obtain the following stochastic process for the volatility $v(t)$ from the variance $V(t)$,

$$dv(t) = \frac{1}{2}\kappa\big[(\theta - \frac{\sigma^2}{4\kappa})v^{-1}(t) - v(t)\big]dt + \frac{1}{2}\sigma dW_2(t). \tag{3.10}$$

Obviously, with the parameter setting given in (3.6), the above process is reduced to an Ornstein-Uhlenbeck process with zero mean level. Compared with the process of $V(t)$ given in (3.5), we can observe that the process of $v(t)$ perseveres the mean-reversion structure. However, the mean-reversion speed of $v(t)$ is $\frac{1}{2}\kappa$, and the volatility of volatility is $\frac{1}{2}\sigma$, equivalent to the half value of the corresponding variance process $V(t)$. More worthy to mention is the mean level of $v(t)$ which is $\theta_v = (\theta - \frac{\sigma^2}{4\kappa})v^{-1}(t) - v(t)$. Firstly, the mean level θ_v is not deterministic as for variance, but stochastic. This means we can not expect the volatility will converges to a given value in the Heston model. Secondly, θ_v has two terms $\theta v^{-1}(t)$ and $-\frac{\sigma^2}{4\kappa}v^{-1}(t)$. While the first term $\theta v^{-1}(t)$ is of a order of $v(t)$, the second term $\frac{\sigma^2}{4\kappa}v^{-1}(t)$ could be very large. These two effects together makes the mean level of volatility in the Heston model could be very small, even negative. The condition that θ_v keeps positive is $4\kappa\theta > \sigma^2$. Obviously, Feller's condition (3.7) for positive variance $V(t)$ is sufficient for a positive mean level of volatility $v(t)$ in the Heston model. For more discussions on the properties of the density of $v(t) = \sqrt{V(t)}$, please see Lipton and Sepp (2008).

3.2.2 PDE Approach to Pricing Formula

Now we demonstrate the original approach with which Heston has derived his pricing formula. Based on the risk-neutral processes in (3.5), we can apply the Feynman-Kac theorem and derive a two-dimensional pricing PDE for option as follows:

$$\begin{aligned}&\frac{1}{2}V\frac{\partial^2 C}{\partial x^2} + \rho\sigma V\frac{\partial^2 C}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 C}{\partial V^2} + \\ &(r - \frac{1}{2}V)\frac{\partial C}{\partial x} + \kappa(\theta - V)\frac{\partial C}{\partial V} - rC = \frac{\partial C}{\partial T},\end{aligned} \tag{3.11}$$

where $x(t) = \ln S(t)$ and T is the maturity of the option. To obtain a closed-form pricing formula for European-style options in the Heston model, we suggest that the searched formula takes a same form as in the Black-Scholes formula, that is

$$C_0 = S_0 F_1^{Q_1} - Ke^{-rT}F_2^{Q_2},$$

where $F_i^{Q_i}, i = 1, 2$, are two exercise probabilities under the respective measures. Inserting this equation into the PDE (3.11) yields two PDEs for $F_1^{Q_1}$ and $F_2^{Q_2}$ respectively. This first one is

$$\frac{1}{2}V\frac{\partial^2 F_1}{\partial x^2}+\rho\sigma V\frac{\partial^2 F_1}{\partial x\partial V}+\frac{1}{2}\sigma^2 V\frac{\partial^2 F_1}{\partial V^2}+$$
$$(r+\frac{1}{2}V)\frac{\partial F_1}{\partial x}+[\kappa(\theta-V)+\rho\sigma V]\frac{\partial F_1}{\partial V}=\frac{\partial F_1}{\partial T} \tag{3.12}$$

with a boundary condition

$$F_1(x_0,V_0)=\mathbf{1}_{(x_0>\ln K)}.$$

The second PDE is

$$\frac{1}{2}V\frac{\partial^2 F_2}{\partial x^2}+\rho\sigma V\frac{\partial^2 F_2}{\partial x\partial V}+\frac{1}{2}\sigma^2 V\frac{\partial^2 F_2}{\partial V^2}+$$
$$(r-\frac{1}{2}V)\frac{\partial F_2}{\partial x}+\kappa(\theta-V)\frac{\partial F_2}{\partial V}=\frac{\partial F_2}{\partial T} \tag{3.13}$$

with a boundary condition

$$F_2(x_0,V_0)=\mathbf{1}_{(x_0>\ln K)}.$$

This means that two exercise probabilities satisfy a similar PDE as the one for call price. According to Kolmogorov's backward equations, a moment-generating function must satisfy the same PDE as its probability, but with different boundary condition. Since the characteristic function is a special case of moment-generating function, we have two PDEs for the CFs $f_j(\phi), j=1,2$,

$$\frac{1}{2}V\frac{\partial^2 f_j}{\partial x^2}+\rho\sigma V\frac{\partial^2 f_j}{\partial x\partial V}+\frac{1}{2}\sigma^2 V\frac{\partial^2 f_j}{\partial V^2}+$$
$$(r-\frac{(-1)^j}{2}V)\frac{\partial f_j}{\partial x}+[\kappa(\theta-V)+a_j]\frac{\partial f_j}{\partial V}=\frac{\partial f_j}{\partial T}, \tag{3.14}$$

where $a_1=\rho\sigma V$ and $a_2=0$. The boundary conditions are

$$f_j(x_0,V_0,T=0)=\exp(i\phi x_0),\quad j=1,2.$$

We propose a guess for the solution for f_j which takes the following exponential form[4]

$$\begin{aligned}f_j(\phi,T)&=\exp\big(i\phi(x_0+rT)-s_{2j}(V_0+\kappa\theta T)+H_1(T)V_0+H_2(T)\big)\\&=\exp\big(i\phi(x_0+rT)\big)f_j^{Heston}(\phi)\end{aligned} \tag{3.15}$$

with

$$f_j^{Heston}(\phi)=\exp\big(-s_{2j}(V_0+\kappa\theta T)+H_1(T)V_0+H_2(T)\big). \tag{3.16}$$

By plunging this guess into the PDE in (3.14), we can arrive at two different ODEs for H_1 and H_2. Solving these ODEs yields the following functions,

[4] The proposed solution here is a bit different from the Heston's original one.

$$
\begin{aligned}
H_1(T) &= \frac{1}{\gamma_2}\big[\gamma_1 s_{2j}(1+e^{-\gamma_1 T}) - (1-e^{-\gamma_1 T})(2s_{1j}+\kappa s_{2j})\big], \\
H_2(T) &= \frac{2\kappa\theta}{\sigma^2}\ln\Big[\frac{2\gamma_1}{\gamma_2}\exp(\frac{1}{2}(\kappa-\gamma_1)T)\Big],
\end{aligned} \tag{3.17}
$$

$$
\begin{aligned}
\gamma_1 &= \sqrt{\kappa^2+2\sigma^2 s_{1j}}, \\
\gamma_2 &= 2\gamma_1 e^{-\gamma_1 T} + (\kappa+\gamma_1-\sigma^2 s_{2j})(1-e^{-\gamma_1 T}).
\end{aligned}
$$

For the first CF f_1, we have the parameters s_{11} and s_{21},

$$
\begin{aligned}
s_{11} &= -(1+i\phi)\big[\frac{\rho\kappa}{\sigma} - \frac{1}{2} + \frac{1}{2}(1+i\phi)(1-\rho^2)\big], \\
s_{21} &= (1+i\phi)\frac{\rho}{\sigma}.
\end{aligned}
$$

Similarly for the second CF f_2, we have s_{12} and s_{22},

$$
\begin{aligned}
s_{12} &= -i\phi\big[\frac{\rho\kappa}{\sigma} - \frac{1}{2} + \frac{1}{2}i\phi(1-\rho^2)\big], \\
s_{22} &= i\phi\frac{\rho}{\sigma}.
\end{aligned}
$$

Note the above expressions for CFs are not identical to Heston's original ones. It can be shown that the both expressions are equivalent.

3.2.3 Expectation Approach to Pricing Formula

As shown above, Heston's original method is to derive CFs via two-dimensional PDEs. But the derivation of a closed-form solution by solving a PDE is generally not a trivial task, and can not be always successful. Here we present an alternative and more general way to obtain the same closed-form solution for CFs in the Heston model. Namely, we calculate CF directly by reducing two-dimensional problem into one-dimensional problem. Zhu (2000) has shown this method in details. Now let us consider the first CF.

$$
\begin{aligned}
f_1(\phi) &= \mathbf{E}^Q[g_1(T)\exp(i\phi x(T))] = \mathbf{E}^Q[\exp(-rT-x_0+(1+i)\phi x(T))] \\
&= \exp[i\phi(x_0+rT)] \\
&\times \mathbf{E}^Q\Big[\exp\Big((1+i\phi)\big(-\frac{1}{2}\int_0^T V(t)dt + \int_0^T \sqrt{V(t)}dW_1(t)\big)\Big)\Big].
\end{aligned}
$$

Decomposing W_1 into two parts

$$
W_1 = \rho W_2 + \sqrt{1-\rho^2}W,
$$

where W is not correlated with W_1 and W_2. Now we can rewrite the above equation as

$$\begin{aligned}
f_1(\phi) &= \exp[i\phi(x_0+rT)]\mathbf{E}^Q\Big[\exp\Big((1+i\phi)\big(-\frac{1}{2}\int_0^T V(t)dt \\
&\quad + \rho\int_0^T \sqrt{V(t)}dW_2(t)+\sqrt{1-\rho^2}\int_0^T \sqrt{V(t)}dW(t)\big)\Big)\Big] \\
&= \exp[i\phi(x_0+rT)]\mathbf{E}^Q\Big[\exp\Big(-(1+i\phi)\frac{1}{2}\int_0^T V(t)dt \\
&\quad + (1+i\phi)\rho\int_0^T \sqrt{V(t)}dW_2(t)+\frac{1}{2}(1-\rho^2)(1+i\phi)^2\int_0^T V(t)dt\Big)\Big].
\end{aligned} \tag{3.18}$$

At the same time, we integrate $V(t)$ and obtain

$$V(T)-V_0=\kappa\theta T-\kappa\int_0^T V(t)dt+\sigma\int_0^T \sqrt{V(t)}dW_2(t).$$

Rearranging this equation yields

$$\int_0^T \sqrt{V(t)}dW_2(t)=\frac{1}{\sigma}\big[V(T)-V_0-\kappa\theta T+\kappa\int_0^T V(t)dt\big].$$

By inserting $\int_0^T \sqrt{V(t)}dW_2(t)$ into (3.18) we have

$$\begin{aligned}
f_1(\phi) &= \exp[i\phi(x_0+rT)]\mathbf{E}^Q\Big[\exp\Big(-(1+i\phi)\frac{1}{2}\int_0^T V(t)dt \\
&\quad + (1+i\phi)\frac{\rho}{\sigma}\big[V(T)-V_0-\kappa\theta T+\kappa\int_0^T V(t)dt\big] \\
&\quad + \frac{1}{2}(1-\rho^2)(1+i\phi)^2\int_0^T V(t)dt\Big)\Big] \\
&= \exp[i\phi(x_0+rT)-s_{21}(V_0+\kappa\theta T)] \\
&\quad \times \mathbf{E}^Q\Big[\exp\Big(-s_{11}\int_0^T V(t)dt+s_{21}V(T)\Big)\Big],
\end{aligned} \tag{3.19}$$

where s_{11} and s_{21} are already given above.

So far we have changed the two-dimensional expectation into a one-dimensional expectation where $S(T)$ has vanished. To obtain the final form of $f_1(\phi)$, we need only to calculate the following expectation

$$y(V_0,T)=\mathbf{E}^Q\Big[\exp\Big(-s_{11}\int_0^T V(t)dt+s_{21}V(T)\Big)\Big]. \tag{3.20}$$

According to the Feynman-Kac theorem, the expected value $y(V_0,T)$ should fulfill the following one-dimensional PDE,

$$\frac{\partial y}{\partial T} = -s_{11}Vy + \kappa(\theta - V)\frac{\partial y}{\partial V} + \frac{1}{2}\sigma^2 V \frac{\partial^2 y}{\partial V^2} \tag{3.21}$$

with a boundary condition

$$y(V_0, 0) = \exp(s_{21}V_0).$$

We can show easily that the solution to this PDE is

$$y = \exp(H_1(T)V_0 + H_2(T)), \tag{3.22}$$

where the functions $H_1(T)$ and $H_2(T)$ are given in (3.17). Plunging the solution of y into $f_1(\phi)$ yields the analytical solution of the first CF $f_1(\phi)$.

The above procedure shows the basic idea of the expectation approach: the dimension reduction via stochastic calculus. Finally we only need to solve a one-dimensional PDE instead of two two-dimensional PDE in a straightforward PDE approach. The derivation of $f_2(\phi)$ follows exactly the same way.

3.2.4 Various Representations of CFs

The CFs derived above are one of some representations in financial literature. In fact, the original CFs given in Heston (1993) take an other form. For readers's reference, we give the original Heston's representation as follows:

$$f_j^{Heston}(\phi) = \exp\left(H_1^*(T)V_0 + H_2^*(T)\right), \quad j = 1, 2, \tag{3.23}$$

with

$$H_1^*(T) = \frac{b_j + d_j - \rho\sigma i\phi}{\sigma^2}\left(\frac{1 - e^{d_jT}}{1 - g_j e^{d_jT}}\right), \tag{3.24}$$

$$H_2^*(T) = \frac{\kappa\theta}{\sigma^2}\left[(b_j + d_j - \rho\sigma i\phi)T - 2\ln\left(\frac{1 - g_j e^{d_jT}}{1 - g_j}\right)\right], \tag{3.25}$$

where

$$g_j = \frac{b_j + d_j - \rho\sigma i\phi}{b_j - d_j - \rho\sigma i\phi},$$

$$d_1 = \sqrt{(\rho\sigma i\phi - b_j)^2 + \sigma^2(\phi^2 - i\phi)},$$

$$d_2 = \sqrt{(\rho\sigma i\phi - b_j)^2 + \sigma^2(\phi^2 + i\phi)},$$

$$b_1 = \kappa - \rho\sigma, \quad b_2 = \kappa.$$

Straightforward but involved calculation can verify that $H_1^*(T)$ is equal to $H_1(T) + s_{2j}$, and $H_2^*(T)$ is equal to $H_2(T) + s_{2j}\kappa\theta T$ in (3.15).

A similar representation of CFs is given by Schoutens-Simons-Tistaert (2004) and Gatheral (2005), which reads

$$H_1^*(T) = \frac{b_j - d_j - \rho\sigma i\phi}{\sigma^2}\Big(\frac{1-e^{d_jT}}{1-q_je^{-d_jT}}\Big) \tag{3.26}$$

and

$$H_2^*(T) = \frac{\kappa\theta}{\sigma^2}\big[(b_j - d_j - \rho\sigma i\phi)T - 2\ln\big(\frac{1-q_je^{-d_jT}}{1-q_j}\big)\big] \tag{3.27}$$

with

$$q_j = 1/g_j.$$

To verify the equivalence between these two representations, we just utilize the following facts,

$$\begin{aligned}\frac{b_j + d_j - \rho\sigma i\phi}{\sigma^2}\Big(\frac{1-e^{d_jT}}{1-g_je^{d_jT}}\Big) &= \frac{b_j + d_j - \rho\sigma i\phi}{\sigma^2 g_j}\Big(\frac{1-e^{-d_jT}}{1-e^{-d_jT}/g_j}\Big)\\ &= \frac{b_j - d_j - \rho\sigma i\phi}{\sigma^2}\Big(\frac{1-e^{d_jT}}{1-q_je^{-d_jT}}\Big)\end{aligned}$$

and

$$\begin{aligned}d_jT - 2\ln\big(\frac{1-g_je^{d_jT}}{1-g_j}\big) &= d_jT - 2\ln\big(\frac{e^{d_jT}(e^{-d_jT}-g_j)}{1-g_j}\big)\\ &= d_jT - 2d_jT - 2\ln\big(\frac{(e^{-d_jT}/g_j - 1)}{1/g_1 - 1}\big)\\ &= -d_jT - 2\ln\big(\frac{(1-q_je^{-d_jT})}{1-q_j}\big).\end{aligned}$$

At a first glance the difference between two representations is very small. However, as discussed in Albrechter-Mayer-Schoutens-Tidtaert (2006), Roger-Kahl (2008), with the small modifications von $H_1^*(T)$ in (3.26) and $H_2^*(T)$ in (3.27) possible discontinuity of the CFs in the Heston model can be circumvented. We will discuss this topic in details in Chapter 4.

3.3 The Schöbel-Zhu Model

3.3.1 Model Setup and Properties

Besides the mean-reverting square root process in the Heston model, a mean-reverting Ornstein-Uhlenbeck process is an other popular process in financial modeling. The first short interest rate model, Vasicek model (1973), is based on a mean-reverting Ornstein-Uhlenbeck process. Generally, compared with a mean-reverting square root process, a mean-reverting Ornstein-Uhlenbeck process is Gaussian and has nice analytical tractability. Many financial economists applied also the mean-reverting Ornstein-Uhlenbeck process to model stochastic volatility, for example, Wiggins (1987), Stein and Stein (1991). Stein and Stein derived a quasi-closed form solution for option where the underlying asset $S(t)$ and its volatility $v(t)$ are not correlated. As we know, the volatility skew is mainly caused by the correlation between the underlying asset and its volatility. Therefore with respect to the correlation, Stein and Stein's solution is very restrictive and is not able to explain the most smile effects observed in option markets. Schöbel and Zhu (1999) extended the Stein and Stein's formulation to a general case and derived an analytic solution for options. Schöbel and Zhu model consists of the following two correlated processes,

$$\begin{aligned} \frac{dS(t)}{S(t)} &= rdt + v(t)dW_1(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma dW_2(t), \\ dW_1 dW_2 &= \rho dt. \end{aligned} \tag{3.28}$$

Differently from the Heston model, the modeling object is volatility $v(t)$, not variance. The volatility process is mean-reverting with a mean level of θ and a reverting speed parameter κ. The parameter σ is the volatility of volatility and controls the variation of $v(t)$. Therefore volatility specified to be a mean-reverting Ornstein-Uhlenbeck process satisfies almost all desired requirements for the stochastic volatility, which are well documented in empirical studies.

Additionally, conditional on $v(s)$,$v(t)$ is distributed according to Gaussian law, namely,

$$v(t) \sim N\big(m_1, m_2\big) \tag{3.29}$$

with

$$m_1 = \mathbf{E}[v(t)|\, v(s)] = \theta + (v(s) - \theta)e^{-\kappa(t-s)},$$

$$m_2 = \mathbf{Var}[v(t)|\, v(s)] = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa(t-s)}).$$

We observe that the conditional means of a mean-reverting square root process and a mean-reverting Ornstein-Uhlenbeck process are identical. However, the conditional variance m_2 is independent of the initial value $v(s)$, and keeps constant over all time interval $[s,t]$ irrespective of how large the volatility $v(s)$ is. Due to the Gaussian property, the volatility $v(t)$ can be negative. The probability for the volatility to be negative is $N(-m_1/\sqrt{m_2})$. Using the parameter values that match to market data, we find that this probability is very small.[5] But even in the case that the negative volatility occurs, the asset process $S(t)$ is still well-defined. From the point of mathematical

[5] The probability that $v(T)$ becomes negative is given by $P = N(-m_1/m_2)$. For example, setting $\kappa = 4, \theta = 0.2, \sigma = 0.1, T-t = 0.3$ and $v_0 = 0.2$ gives $P = 1.50 \times 10^{-9}$. Setting $\kappa = 1, \theta = 0.2, \sigma = 0.1, T-t = 0.5$ and $v_0 = 0.2$ gives $P = 1.87 \times 10^{-4}$. These probabilities are so small that it does not raise any serious problem at all in practical applications.

view, a negative volatility leads to a reflected Brownian motion and a sign change of the correlation term $S(t)v(t)\rho dt$. In this sense, negative volatility does not raise any severe issue for option pricing. In practical applications, the mean-reverting Ornstein-Uhlenbeck process can be easily and exactly simulated without any sophisticated techniques. A convenient simulation is a point that most researchers underestimate and is very important for practitioners.

Now we compare this stochastic volatility model with Heston's model. Looking at the processes (3.28) and (3.4) again, we find that the only difference between them is the mean-reversion parameter θ. While θ in (3.28) generally differs from zero, θ in (3.4) is zero. Since θ gives a level that a volatility approaches in a long run, process (3.4) does not seem to be very reasonable. Therefore, this restricted Heston model can be considered as a special case of Schöbel and Zhu model in terms of equations (3.5) and (3.6). Schöbel and Zhu model is reduced to the Heston model by setting the following parameters:

$$\kappa = \frac{\kappa_h}{2}, \quad \sigma = \frac{\sigma_h}{2}, \quad \theta = 0, \quad \theta_h = \frac{\sigma^2}{\kappa_h}, \tag{3.30}$$

where κ_h, σ_h and θ_h stand for the parameters in process (3.4). However, note that the parameters in process (3.5) are over-determined by (3.30), then for a wide range of the values of κ_h, σ_h and θ_h, the volatility process (3.4) can not be derived from (3.5). This means that (3.4) and (3.5) are not mutually consistent in many cases. Hence in this sense, the Heston model is not a special case of the Schöbel and Zhu model.

There are some confusing explanations about the above model setup in (3.28) in literature. For example, Ball and Roma (1994) discussed the difference between the absolute value process of $v(t)$ and the reflected process of $v(t)$, and claimed that the above model setup in (3.28) implied the absolute value process since "the volatility enters option pricing only as $v^2(t)$" in Stein and Stein's solution. Unfortunately, their claim is incorrect. If we are working with an absolute value process, we can expect the same option prices for $|v(t)| = v(t)$. But we can not get the same option prices for this case. Most rigorously, As shown in Schöbel and Zhu (1999), for arbitrary values of the long-term mean $\theta \neq 0$, the absolute value Ornstein-Uhlenbeck process, the reflected Ornstein-Uhlenbeck process and the unrestricted Ornstein-Uhlenbeck process are substantially different. Only for the special (symmetric) case $\theta = 0$ the reflecting barrier process coincides with the absolute value process. Furthermore, we also can verify that the density function of the above Ornstein-Uhlenbeck process process in (3.28) is identical to the one of the unrestricted Ornstein-Uhlenbeck process and does not satisfy the boundary condition which is necessary for an absolute value process.

3.3.2 Derivation of CFs

Schöbel and Zhu (1999) applied the expectation approach, as shown in the Heston model, to derive the analytic solutions for CFs. Readers can verify that the PDE approach is no longer valid for the stochastic volatility model in (3.28), which is a evidence that expectation approach is more general for PDE approach. Here we do not show the derivations in details and give directly the analytic forms of CFs.

$$\begin{aligned} f_1(\phi) &= \mathbf{E}^Q[g_1(T)\exp(i\phi x(T))] = \mathbf{E}^Q[\exp(-rT - x_0 + (1+i)\phi x(T))] \\ &= \exp\Big(i\phi(x_0 + rT) - \frac{\rho}{2\sigma}(1+i\phi)v_0^2 - \frac{1}{2}(1+i\phi)\rho\sigma T\Big) \\ &\times \mathbf{E}^Q\Big[\exp\Big(-s_{11}\int_0^T v^2(t)dt - s_{21}\int_0^T v(t)dt + s_{31}v^2(T)\Big)\Big] \\ &= \exp\big(i\phi(rT + x_0)\big) cf_1^{SZ}(\phi) \end{aligned} \tag{3.31}$$

with

$$\begin{aligned} cf_1^{SZ}(\phi) = \exp&\Big(-s_{31}v_0^2 - \frac{1}{2}(1+i\phi)\rho\sigma T\Big) \\ &\times \mathbf{E}^Q\Big[\exp\Big(-s_{11}\int_0^T v^2(t)dt - s_{21}\int_0^T v(t)dt + s_{31}v^2(T)\Big)\Big]. \end{aligned} \tag{3.32}$$

The parameters $s_{11}, s_{21}, s_{31}$are defined as follows:

$$\begin{aligned} s_{11} &= -\frac{1}{2}(1+i\phi)((1+i\phi)(1-\rho^2) - 1 + 2\rho\kappa\sigma^{-1}), \\ s_{21} &= \frac{\rho\kappa\theta}{\sigma}(1+i\phi), \\ s_{31} &= \frac{\rho}{2\sigma}(1+i\phi). \end{aligned}$$

We set

$$y(v_0, T) = \mathbf{E}^Q\Big[\exp\Big(-s_{11}\int_0^T v^2(t)dt - s_{21}\int_0^T v(t)dt + s_{31}v^2(T)\Big)\Big].$$

According to the Feynman-Kac theorem, y shall satisfy the following PDE,

$$\frac{\partial y}{\partial T} = -(s_{21}v^2 + s_{21}v)y + \kappa(\theta - v)\frac{\partial y}{\partial v} + \frac{1}{2}\sigma^2\frac{\partial^2 y}{\partial v^2} \tag{3.33}$$

with a boundary condition

$$y(v_0, 0) = \exp(s_{31}v_0^2).$$

This PDE has the following solution,

$$y(v_0,T) = \exp\left(\frac{1}{2}H_3(T)v_0^2 + H_4(T)v_0 + H_5(T)\right), \tag{3.34}$$

where

$$\begin{aligned}
H_3(T) &= \frac{\kappa}{\sigma^2} - \frac{\gamma_1}{\sigma^2\gamma_4}(\sinh(\gamma_1 T) + \gamma_2\cosh(\gamma_1 T)),\\
H_4(T) &= \frac{1}{\gamma_1\gamma_4\sigma^2}\\
&\quad\times[(\kappa\theta\gamma_1 - \gamma_2\gamma_3)(1-\cosh(\gamma_1 T)) - (\kappa\theta\gamma_1\gamma_2 - \gamma_3)\sinh(\gamma_1 T)],\\
H_5(T) &= -\frac{1}{2}\ln(\gamma_4) + \frac{\sinh(\gamma_1 T)}{2\gamma_1^3\gamma_4\sigma^2}[(\kappa\theta\gamma_1 - \gamma_2\gamma_3)^2 - \gamma_3^2(1-\gamma_2^2)]\\
&\quad+\frac{\gamma_3}{\gamma_1^3\gamma_4\sigma^2}(\kappa\theta\gamma_1 - \gamma_2\gamma_3)(\gamma_4 - 1) + \frac{T}{2\gamma_1^2\sigma^2}[\kappa\gamma_1^2(\sigma^2 - \kappa\theta^2) + \gamma_3^2]
\end{aligned}$$

with

$$\begin{aligned}
\gamma_1 &= \sqrt{2\sigma^2 s_{11} + \kappa^2},\\
\gamma_2 &= \kappa^2\theta - s_{21}\sigma^2,\\
\gamma_3 &= \frac{1}{\gamma_1}(\kappa - 2\sigma^2 s_{31}),\\
\gamma_4 &= \cosh(\gamma_1 T) + \gamma_2\sinh(\gamma_1 T).
\end{aligned}$$

The second CF $f_2(\phi)$ has a similar solution to $f_1(\phi)$ and takes the following form,

$$\begin{aligned}
f_2(\phi) &= \mathbf{E}^Q[g_1(T)\exp(i\phi x(T))] = \mathbf{E}^Q[\exp(-rT - x_0 + (1+i)\phi x(T))]\\
&= \exp\left(i\phi(x_0 + rT) - \frac{\rho}{2\sigma}i\phi v_0^2 - \frac{1}{2}i\phi\rho\sigma T\right)\\
&\quad\times\mathbf{E}^Q\left[\exp\left(-s_{12}\int_0^T v^2(t)dt - s_{22}\int_0^T v(t)dt + s_{32}v^2(T)\right)\right]\\
&= \exp\left(i\phi(rT + x_0)\right)cf_2^{SZ}(\phi)
\end{aligned} \tag{3.35}$$

with

$$\begin{aligned}
cf_2^{SZ}(\phi) &= \exp\left(-s_{32}v_0^2 - \frac{1}{2}i\phi\rho\sigma T\right)\\
&\quad\times\mathbf{E}^Q\left[\exp\left(-s_{12}\int_0^T v^2(t)dt - s_{22}\int_0^T v(t)dt + s_{32}v^2(T)\right)\right],
\end{aligned} \tag{3.36}$$

where s_{12}, s_{22}, s_{32} take the following forms,

$$s_{12} = -\frac{1}{2}i\phi(i\phi(1-\rho^2) - 1 + 2\rho\kappa\sigma^{-1}),$$
$$s_{22} = \frac{\rho\kappa\theta}{\sigma}i\phi,$$
$$s_{32} = \frac{\rho}{2\sigma}i\phi.$$

In order to compute $f_2(\phi)$, s_{k1} in $\gamma_j, j = 1,2,3,4$, should be replaced by s_{k2} correspondingly.

3.3.3 Numerical Examples

Here we give some numerical examples to demonstrate the properties of Schöbel-Zhu model. For comparison, we shall choose the suitable Black-Scholes (BS) option prices as benchmark. The first possible benchmark is the BS price using the expected average variance AV as input. As shown in Ball and Roma (1994), if $v(t)$ follows a mean-reverting Ornstein-Uhlenbeck process, we can calculate AV as follows:

$$AV = \frac{\sigma^2}{2\kappa} + \theta^2 + \frac{1}{T}\left(\frac{2\theta(v_0-\theta)}{\kappa}(1-e^{-\kappa T}) - \frac{\sigma^2 - 2\kappa(v_0-\theta)^2}{4\kappa^2}(1-e^{-2\kappa T})\right). \tag{3.37}$$

We denote the BS prices evaluated in this manner by BS_1. The second alternative as in Stein and Stein is to calculate the BS option prices according to the spot volatility $v(t)$ with κ, ρ and σ as nil. This calculated BS prices labeled to BS_2 are identical to these prices by using the BS formula with a constant volatility $v(t)$ being equal to θ. Heston (1993) used the BS prices with a volatility which equates the square root of variance of the spot stock returns up to maturity. His benchmark is then dependent on the correlation ρ and makes comparisons remarkably complicated.[6] It is difficult to say which benchmark is better. Evaluating BS_1 has an implication that volatility follows a stochastic process and is not constant. This is in contradiction with the BS's assumption. In this sense, BS_1 is not a logically consistent but only *ad hoc* BS value. In contrast, BS_2 implies that the volatility keeps constant and equates the initial value v_0. Therefore, both BS_1 and BS_2 might (not) be a right benchmark for comparison with the exact option prices. However, as shown by numerical examples, both benchmarks do not distinguishes each other remarkably.

[6] Since Heston's BS benchmark is a function of the correlation ρ, every model option value has a corresponding Heston's benchmark. Thus, we will lose some clarity in comparison if using Heston's benchmark although it is reasonable in some senses.

In Table (3.2), we choose the parameters suggested by Stein and Stein. In order to show the impact of correlation on the option prices, let ρ range from -1.0 to 1.0. Some observations are in order.

First of all, options with different moneyness have different sensitivity to the correlation ρ. The values of at-the-money (ATM) options do not change remarkably overall. However, the sensitivity of out-of-the-money (OTM) options to ρ is more conspicuous than of in-the-money (ITM) options. For example, in Panel C the relative changes of the OTM option prices due to the correlation ρ for $K = 120$ is about $\pm 14\%$ of the Stein and Stein value which is here 2.635.

Secondly, a comparison of Panels A, B and C shows that the mean-reversion level θ is important for the pricing of options. Keeping other parameters unchanged, the differential $(\theta - v)$ (mean-reversion level minus current volatility) has a great impact on the option values. From Panel A to C, $(\theta - v)$ is 0, -0.1 and 0.1 respectively, and the differences in option prices across these panels are mostly between 0.60\$ and 1.50\$. Since the expectation of the future spot price volatility approaches θ, the prices of options, especially the options with a long-term maturity, should be mainly affected by θ.

Thirdly, as expected, we find the price differences between BS_1 and the model values with $\rho = 0$ are smallest for all panels. This confirms Ball and Roma's finding. The good fit of these two values is not surprising since BS_1 is by nature an approximation for the exact option value with $\rho = 0$. Panel D in Table 2.3 presents all BS_1 values corresponding to Panels E-G. We can see that the BS_1 values in Panel D agree very well with the option values in Panel F. However, if $\rho \neq 0$, the price bias between BS_1 values and the exact option values are significant. Thus, BS_1 is not a suitable approximation for the correlation case. It is also not surprising that the BS_2 values match our model values closely for Panel A where $\theta = v$. Stein and Stein report an overall overvaluation relative to BS_2 due to stochastic volatility. This upward pricing bias should be caused by the zero correlation assumption between volatility and its underlying asset returns in Stein and Stein. For $\rho = 0$ the direction of the movement of $S(t)$ is not affected by stochastic volatility, and any stochastic volatility raises only the additional uncertainty of $S(t)$. Consequently, Stein and Stein and BS_1 values are greater than BS_2 values in Panel A.

Finally, ITM options and OTM options react to correlation just oppositely. Whereas ITM options ($K = 90, 95$) decrease in value with increasing correlation ρ, OTM option prices ($K = 105, 110, 115, 120$) go up. This finding is consistent with Hull's (1997, pages 492-500) excellent intuitive explanation of how correlation affects option prices. It is also empirically evident that stock returns are inversely correlated with the underlying volatilities. Panels A to C show that a negative correlation ρ leads to ITM (OTM) option prices in our model that are greater (less) than the corresponding BS option prices. This feature caused by negative correlation is useful for explaining volatility sneers. The smile pattern is most of the time not so obvious and appears to be more of a sneer. The monotonic downward sloping of the implied volatility with moneyness displayed in a sneer implies that the market ITM (OTM) options are undervalued (overvalued) by the BS formula. This pattern of pricing biases in the BS formula is reported by a number of empirical studies as in Nandi (1998), Heston and Nandi (1997), BCC (1997) and MacBeth and Merville (1979). Figure (3.1) shows the calibrated implied volatilities in Panel A where the downward sloping of implied volatility for $\rho < 0$ is significant. The good fit of our calibrated examples to the empirical picture supports that incorporating correlation between stock returns and volatilities is a promising way to improve the performance of option pricing. BCC (1997) and Nandi (1998) reported that taking stochastic volatility into account is of first-order importance in eliminating the volatility inconsistence and leads to significant improvements in hedging performance, but only in the presence of non-zero correlation. Their results should also be valid for our model.

In Table (3.3), we examine how the option prices vary with the mean-reverting level θ. The finding that option prices are very sensible to θ is confirmed. Since θ indicates the long-term level of volatility, this sensitivity can be considered as the sensitivity of option prices to their volatilities in the long run. Furthermore, it seems to be that θ is more important than the spot volatility $v(t)$ for the pricing of options in the framework of a mean-reverting process. If the true process of volatility performs mean-reversion, and option prices are evaluated using the BS formula regardless of BS_1 or BS_2, a significant pricing bias will occur although the magnitude of the pricing bias for BS_1 is smaller than that for BS_2. All prices in Panel F correspond to the case of Stein and Stein. The numbers in the first row in panels D, E, F and G are option prices under the restricted Heston model in the sense of equations (3.5) and (3.6). The implied zero level of the mean-reversion leads to an overall undervaluation of options compared with BS_2.

Table (3.4) demonstrates the impact of ρ on Delta Δ_S which is of first-order importance for hedging purposes whenever stochastic volatility models are used. Firstly, for the given parameters almost all of the Deltas are decreasing with correlation ρ except for a few deep-ITM and deep-OTM options across the three Panels H, I and J. Secondly, the changes of values of near-ATM options relative to the correlation ρ are more sensitive than these of deep-ITM and deep-OTM options. The differences between Δ_S in the Schöbel-Zhu model and the BS model (BS_1 and BS_2) for near-ATM options should not be neglected. For a negative correlation, using Δ_S of the BS formula and the Stein and Stein model seems to cause a severe underhedging for near-ATM options and some OTM options. Furthermore, the long-term level of volatility θ is also important for hedging. Keeping other parameters unchanged, the greater θ is, the smaller (greater) Δ_S will be for ITM (OTM) options. The sensitivity of Δ_S to θ is remarkable and can be studied more detailed by the second derivative $\Delta_{S\theta}$. We conclude that an unbiased estimate of θ is crucial for Delta-hedging.

Stochastic volatility models provide us with new insights into derivative security markets. In this section, we have derived a closed-form pricing formula for the general case where volatility is allowed to display arbitrary correlation with the underlying stock price. In comparison with the Heston model, this specification for volatility performs additional favorite properties such as a direct link to volatility instead of variance, easy estimation of parameters in a Gaussian framework. As discussed earlier, the negligibly small possibilities for negative volatilities in such a process do not raise a serious problem in option pricing. Additionally, since in a diffusion context negative volatilities only mean that upward moves of the driving Brownian motion become downward moves of the stock price and vice versa, we believe that this is not a severe theoretical restriction and suggest this new closed-form pricing formula as an alternative to Heston's solution: Not surprisingly, squared volatilities never become negative here either.

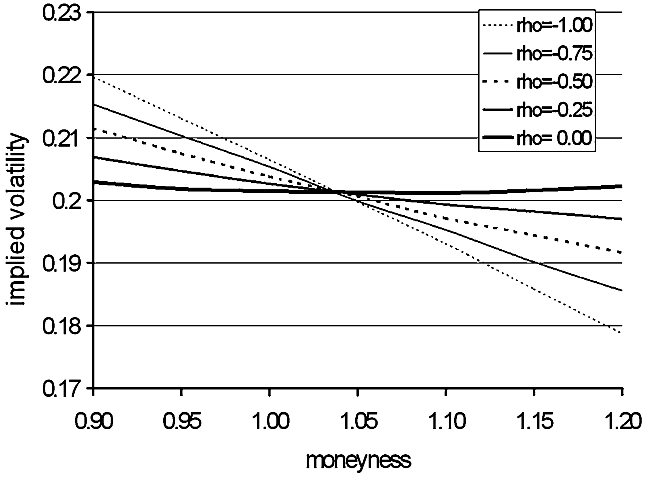

Fig. 3.1 Model volatility smile for negative and zero correlation, calculated from Panel A in Table (3.2)

3.4 Double Square Root Model

3.4.1 Model Setup and Properties

In this section, we consider a new variant of stochastic volatility model where squared volatility is specified as a double square root process. This is a process that was originally introduced by Longstaff (1989) to model stochastic interest rates, in order to explain non-linear term structure observed in yield curves. Longstaff showed that the non-linear model driven by double square root process outperforms the CIR model driven by square root process for the 1964-1986 period in explaining the behavior of yield curves. Since stochastic variance and stochastic interest rates share many same characteristics, it is interesting to adopt mean-reverting double square root to model the dynamics of stochastic variances $V(t)$. Heston (1993) considered also the stochastic volatility model with double square root process in

ρ \ K	90	95	100	105	110	115	120
BS_2	15.118	11.342	8.142	5.584	3.658	2.293	1.377
BS_1	15.152	11.391	8.203	5.650	3.722	2.349	1.422
-1.00	15.416	11.617	8.307	5.576	3.468	1.966	0.995
-0.75	15.355	11.562	8.275	5.585	3.525	2.063	1.110
-0.50	15.292	11.503	8.243	5.595	3.582	2.156	1.218
-0.25	15.225	11.443	8.210	5.606	3.638	2.246	1.321
0.00	15.155*	11.379*	8.176*	5.617*	3.694*	2.333*	1.420*
0.25	15.081	11.313	8.141	5.628	3.749	2.416	1.514
0.50	15.003	11.243	8.106	5.640	3.803	2.497	1.605
0.75	14.919	11.169	8.070	5.652	3.856	2.576	1.693
1.00	14.828	11.092	8.034	5.665	3.909	2.653	1.777

A: $\theta = 0.2, \kappa = 4, \sigma = 0.1, v = 0.2, T = 0.5, S = 100, r = 0.0953$

ρ \ K	90	95	100	105	110	115	120
BS_1	14.501	10.348	6.829	4.132	2.282	1.150	0.530
-1.00	14.730	10.600	6.976	4.057	1.977	0.739	0.179
-0.75	14.679	10.541	6.936	4.066	2.051	0.853	0.279
-0.50	14.626	10.480	6.894	4.075	2.121	0.957	0.371
-0.25	14.572	10.415	6.849	4.085	2.189	1.053	0.458
0.00	14.515*	10.346*	6.803*	4.093*	2.254*	1.144*	0.542*
0.25	14.456	10.271	6.753	4.102	2.316	1.230	0.621
0.50	14.495	10.191	6.701	4.110	2.376	1.311	0.698
0.75	14.330	10.103	6.645	4.117	2.433	1.389	0.773
1.00	14.261	10.005	6.587	4.125	2.489	1.463	0.845

B: $\theta = 0.1, \kappa = 4, \sigma = 0.1, v = 0.2, T = 0.5, S = 100, r = 0.0953$

ρ \ K	90	95	100	105	110	115	120
BS_1	16.111	12.666	9.711	7.264	5.305	3.786	2.646
-1.00	16.357	12.846	9.777	7.186	5.084	3.449	2.235
-0.75	16.298	12.800	9.755	7.198	5.133	3.531	2.340
-0.50	16.236	12.751	9.732	7.210	5.182	3.611	2.441
-0.25	16.172	12.702	9.710	7.223	5.230	3.690	2.540
0.00	16.106*	12.651*	9.687*	7.237*	5.279*	3.768*	2.635*
0.25	16.037	12.598	9.665	7.251	5.328	3.844	2.728
0.50	15.964	12.544	9.642	7.265	5.377	3.919	2.819
0.75	15.889	12.489	9.620	7.280	5.426	3.993	2.908
1.00	15.810	12.432	9.598	7.296	5.475	4.065	2.994

C: $\theta = 0.3, \kappa = 4, \sigma = 0.1, v = 0.2, T = 0.5, S = 100, r = 0.0953$

The numbers with * correspond to the model of Stein and Stein.

Table 3.2 The impact of the correlation ρ on option prices. The prices of ITM calls (OTM calls) decrease (increase) with increasing correlations ρ.

his appendix, but did not give an explicit solution for options. Following the method in his paper, i.e., the PDE approach, we find that it leads to a rather cumbersome procedure and can not decompose the PDEs into several ordinary differential equations which can be solved successively. Here we will give an explicit solution by applying stochastic calculus.

In fact, the generator of a double square root process is a Brownian motion with drift. To see this, we assume that volatility $v(t)$ is governed by the following process,

θ \ K	90	95	100	105	110	115	120
BS_2	14.515	10.374	6.867	4.175	2.322	1.180	0.550
0.0	14.194*	9.510*	5.277*	2.213*	0.654*	0.132*	0.018*
0.1	14.333	9.990	6.282	3.499	1.708	0.728	0.272
0.2	14.864	10.962	7.662	5.060	3.155	1.859	1.037
0.3	15.757	12.213	9.187	6.707	4.755	3.278	2.200

D: The BS_1 values using the expected average volatility as input
$\kappa = 4, \sigma = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$

θ \ K	90	95	100	105	110	115	120
0.0	14.189*	9.459*	5.141*	2.173*	0.763*	0.244*	0.075*
0.1	14.262	9.842	6.132	3.466	1.812	0.895	0.425
0.2	14.723	10.804	7.551	5.044	3.239	2.014	1.220
0.3	15.608	12.082	9.109	6.703	4.828	3.414	2.376

E: $\rho = 0.5, \kappa = 4, \sigma = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$

θ \ K	90	95	100	105	110	115	120
0.0	14.200*	9.527*	5.267*	2.170*	0.645*	0.150*	0.030*
0.1	14.351	9,997	6.254	3.451	1.679	0.731	0.292
0.2	14.871	10.952	7.632	5.022	2.124	1.845	1.040
0.3	15.754	12.198	9.161	6.676	4.727	3.259	2.192

F: $\rho = 0.0, \kappa = 4, \sigma = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$

θ \ K	90	95	100	105	110	115	120
0.0	14.225*	9.600*	5.372*	2.152*	0.504*	0.061*	0.004*
0.1	14.441	10.130	6.361	3.435	1.529	0.545	0.155
0.2	15.003	11.084	7.710	5.003	3.004	1.660	0.842
0.3	15.887	12.306	9.213	6.652	4.626	3.097	1.995

G: $\rho = -0.5, \kappa = 4, \sigma = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$
The numbers with * correspond to the restricted Heston model.

Table 3.3 The impact of the mean level θ on option prices. Option prices increase with increasing mean levels θ.

$$dv(t) = \mu dt + \varepsilon dW_2(t),$$

where μ and ε are constant, and $dW_1 dW_2 = \rho dt$. Applying Itô's lemma yields the dynamics of the variances,

$$\begin{aligned} dv^2(t) &= [\varepsilon^2 + 2\mu v(t)]dt + 2\varepsilon v(t)dW_2(t) \\ &= \kappa[\theta - v(t)]dt + \sigma v(t)dW_2(t) \end{aligned}$$

with

$$\kappa = -2\mu, \qquad \kappa\theta = \varepsilon^2, \qquad \sigma = 2\varepsilon.$$

Now let $V(t) = v^2(t)$ denote the squared volatility, we then arrive at a process for instantaneous variances $V(t)$ as follows:

$$dV(t) = \kappa\left[\theta - \sqrt{V(t)}\right]dt + \sigma\sqrt{V(t)}dW_2(t). \tag{3.38}$$

Since two square root terms arise in this process, it is thus referred to as double square root process. The necessary condition for generating process (3.38) is the

ρ \ K	90	95	100	105	110	115	120
BS$_2$	0.8755	0.7795	0.6582	0.5249	0.3950	0.2807	0.1890
BS$_1$	0.8731	0.7773	0.6571	0.5253	0.3968	0.2836	0.1922
-1.00	0.8751	0.7949	0.6895	0.5644	0.4299	0.3002	0.1883
-0.75	0.8708	0.7916	0.6825	0.5547	0.4204	0.2941	0.1881
-0.50	0.8751	0.7881	0.6751	0.5449	0.4113	0.2889	0.1883
-0.25	0.8752	0.7844	0.6673	0.5350	0.4027	0.2843	0.1887
0.00	0.8754*	0.7802*	0.6591*	0.5251*	0.3945*	0.2802*	0.1891*
0.25	0.8756	0.7757	0.6504	0.5153	0.3868	0.2765	0.1896
0.50	0.8759	0.7707	0.6413	0.5055	0.3795	0.2732	0.1899
0.75	0.8761	0.7649	0.6317	0.4957	0.3725	0.2701	0.1904
1.00	0.8761	0.7584	0.6216	0.4862	0.3659	0.2673	0.1907

H: $\theta = 0.2, \kappa = 4, \sigma = 0.1, v = 0.2, T = 0.5, S = 100, r = 0.0953$

ρ \ K	90	95	100	105	110	115	120
BS$_1$	0.9344	0.8401	0.6937	0.5166	0.3440	0.2047	0.1093
-1.00	0.9246	0.8490	0.7302	0.5684	0.3816	0.2048	0.0763
-0.75	0.9267	0.8479	0.7229	0.5554	0.3694	0.2025	0.0867
-0.50	0.9293	0.8467	0.7151	0.5424	0.3584	0.2012	0.0947
-0.25	0.9323	0.8456	0.7068	0.5293	0.3485	0.2003	0.1012
0.00	0.9359*	0.8445*	0.6978*	0.5163*	0.3395*	0.1996*	0.1065*
0.25	0.9403	0.8433	0.6881	0.5033	0.3311	0.1989	0.1111
0.50	0.9458	0.8421	0.6774	0.4903	0.3234	0.1982	0.1149
0.75	0.9530	0.8404	0.6656	0.4773	0.3163	0.1975	0.1182
1.00	0.9629	0.8379	0.6526	0.4645	0.3095	0.1967	0.1211

I: $\theta = 0.1, \kappa = 4, \sigma = 0.1, v = 0.2, T = 0.5, S = 100, r = 0.0953$

ρ \ K	90	95	100	105	110	115	120
BS$_1$	0.8225	0.7358	0.6373	0.5341	0.4334	0.3410	0.2606
-1.00	0.8317	0.7552	0.6646	0.5644	0.4606	0.3599	0.2682
-0.75	0.8300	0.7512	0.6584	0.5568	0.4532	0.3541	0.2652
-0.50	0.8282	0.7469	0.6519	0.5492	0.4459	0.3487	0.2626
-0.25	0.8264	0.7424	0.6452	0.5416	0.4389	0.3437	0.2604
0.00	0.8243*	0.7376*	0.6383*	0.5339*	0.4322*	0.3390*	0.2584*
0.25	0.8221	0.7324	0.6311	0.5263	0.4256	0.3347	0.2566
0.50	0.8196	0.7269	0.6237	0.5187	0.4193	0.3306	0.2550
0.75	0.8169	0.7210	0.6162	0.5112	0.4133	0.3267	0.2536
1.00	0.8137	0.7147	0.6084	0.5038	0.4074	0.3231	0.2522

J: $\theta = 0.3, \kappa = 4, \sigma = 0.1, v = 0.2, T = 0.5, S = 100, r = 0.0953$

The numbers with * correspond to the model of Stein and Stein.

Table 3.4 The impact of the correlation ρ on Delta Δ_S. Deltas decrease with increasing correlations ρ only for near-ATM options.

restriction $\sigma^2 = 4\kappa\theta$ with which some special expected values of $V(t)$ can be calculated analytically. Additionally, zero may be an accessible boundary under this restriction.[7]

Generally, we have two ways to adjust the above double square root process to a risk-neutral process by market price of risk. More precisely, one approach is

[7] When $0 < \sigma^2 < \kappa\theta$, zero value of $v(t)$ is inaccessible and no boundary condition can be imposed at zero. When $0 < \kappa\theta < \sigma^2$, zero is accessible. See Longstaff (1989).

to introduce a market price of volatility risk, the other is to utilize a market price of variance risk. By the first approach, the market price of risk is equal to $\lambda(t) = \lambda v(t) = \lambda\sqrt{V(t)}$, which implies a risk-neutral process as follows:

$$\begin{aligned} dV(t) &= \left[\kappa\theta - (\kappa-\lambda)\sqrt{V(t)}\right]dt + \sigma\sqrt{V(t)}dW_2(t) \\ &= \kappa^*\left[\theta^* - \sqrt{V(t)}\right] + \sigma\sqrt{V(t)}dW_2(t) \end{aligned} \tag{3.39}$$

with $\kappa^* = \kappa - \lambda$ and $\theta^* = \kappa\theta/\kappa^*$. Hence, the obtained risk-neutral process keeps the same structure as its original counterpart. In the second approach, we introduce an identical market price of risk as in the Heston model, i.e., $\lambda(t) = \lambda V(t)$. This interpretation of market price of risk λ leads to the following risk-neutral process,

$$\begin{aligned} dV(t) &= \left[\kappa\theta - \kappa\sqrt{V(t)} - \lambda V(t)\right]dt + \sigma\sqrt{V(t)}dW_2(t) \\ &= \left[\frac{1}{4}\sigma^2 - \kappa\sqrt{V(t)} - \lambda V(t)\right]dt + \sigma\sqrt{V(t)}dW_2(t), \end{aligned} \tag{3.40}$$

which has a mixed structure of double and single square root processes. Note the risk-neutral process of first approach is a special case of the one of second approach, we discuss the second approach for generality. By setting $\lambda = 0$, the price formula given below is reduced to the price formula with market price of volatility risk.

Compared with the single square root process (3.4) in the Heston model, the double square root process has some distinct features. Firstly, the mean level of variance here does not take an unambiguous value since the constant $\frac{1}{4}\sigma^2$ measures the long term level of $\kappa\sqrt{V(t)} + \lambda V(t)$. This mixed structure of the drift provides us with a weighted value of volatility and variance as mean level. Due to the restriction $\sigma^2 = 4\kappa\theta$, it seems that $\frac{1}{4}\sigma^2$ as long term value of $(\kappa\sqrt{V(t)} + \lambda V(t))$ is small in many cases. For example, for $\sigma = 0.5$ and $V_0 = 0.2$, the reasonable values of κ should not be larger than 1. This means that volatility shocks to stock price process are rather persistent and revert to mean level very slowly. Another feature of this model is that, as mentioned above, the generating process of double square root dynamics is a Brownian motion with drift and therefore is not stationary, and volatility itself does not perform mean-reversion, hence such specifications of volatility and variance are not tenable from *a priori* theoretical consideration. On the other hand, we have one parameter less to estimate empirically because of the above parameter restriction. Numerical examples show that this stochastic volatility model can still capture most essential features of volatility smile.

The double square process will be identical to single square process if we set some special parameter values,

$$\kappa_{DSR} = 0, \qquad \theta_{SR} = 0, \qquad \kappa_{SR} = \lambda_{DSR},$$

where the subscripts *DSR* and *SR* stand for double square root process and square root process respectively. Obviously, such parameter values are not realistic.

It is not hard to obtain the following two CFs for the double square root process:

$$
\begin{aligned}
f_1(\phi) &= \mathbf{E}^Q\left[\frac{S(T)}{S_0e^{rT}}\exp\left(i\phi\ln S(T)\right)\right]\\
&= \exp\left[i\phi(x_0+rT)-\frac{(1+i\phi)\rho}{\sigma}(V_0+\kappa\theta T)\right]\\
&\times\ \mathbf{E}^Q\left[\exp\left(-\frac{(1+i\phi)}{2}\left(1-(1+i\phi)(1-\rho^2)-\frac{2\lambda\rho}{\sigma}\right)\int_0^T V(t)dt\right.\right.\\
&+\ \left.\left.\frac{(1+i\phi)\kappa\rho}{\sigma}\int_0^T\sqrt{V(t)}dt+\frac{(1+i\phi)\rho}{\sigma}V(T)\right)\right]\\
&= \exp\left(i\phi(x_0+rT)-s_3(V_0+\kappa\theta T)\right)\\
&\times\ \mathbf{E}^Q\left[\exp\left(-s_1\int_0^T V(t)dt-s_2\int_0^T\sqrt{V(t)}dt+s_3V(T)\right)\right]\\
&= \exp\left(i\phi(x_0+rT)-s_3(V_0+\kappa\theta T)\right)\\
&\times\ \exp\left(H_6(T;s_1,s_2,s_3)V_0+H_7(T;s_1,s_2,s_3)\sqrt{V_0}+H_8(T;s_1,s_2,s_3)\right)\\
&= \exp(i\phi(x_0+rT))f_1^{DSR}(\phi)
\end{aligned}
\tag{3.41}
$$

with

$$
\begin{aligned}
f_1^{DSR}(\phi) = \exp\big(&-s_3(V_0+\kappa\theta T)+H_6(T;s_1,s_2,s_3)V_0\\
&+H_7(T;s_1,s_2,s_3)\sqrt{V_0}+H_8(T;s_1,s_2,s_3)\big),
\end{aligned}
\tag{3.42}
$$

where the parameters s_1, s_2 and s_3 are defined by

$$
\begin{aligned}
s_1 &= \frac{(1+i\phi)}{2}\left(1-(1+i\phi)(1-\rho^2)-\frac{2\lambda\rho}{\sigma}\right),\\
s_2 &= -\frac{(1+i\phi)\kappa\rho}{\sigma}, \qquad s_3 = \frac{(1+i\phi)\rho}{\sigma}.
\end{aligned}
$$

Thus, the whole problem is reduced to computing the expected value in the last equality, which can be solved by our standard method. The functions $H_6(T), H_7(T)$ and $H_8(T)$ are given in Appendix B. The explicit form of $f_1(\phi)$ has a complex structure, but is expressed in a closed-form manner with elementary functions.

Similarly, we have

$$
\begin{aligned}
f_2(\phi) &= \mathbf{E}^Q\left[\exp\left(i\phi\ln S(T)\right)\right]\\
&= \exp\left[i\phi(x_0+rT)-\frac{i\phi\rho}{\sigma}(V_0+\kappa\theta T)\right]\\
&\times\ \mathbf{E}^Q\left[\exp\left(-\frac{i\phi}{2}\left(1-i\phi(1-\rho^2)-\frac{2\lambda\rho}{\sigma}\right)\int_0^T V(t)dt\right.\right.\\
&+\ \left.\left.\frac{i\phi\kappa\rho}{\sigma}\int_0^T\sqrt{V(t)}dt+\frac{i\phi\rho}{\sigma}V(T)\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \exp\left(i\phi(\ln S_0 + rT) - s_3^*(V_0 + \kappa\theta T)\right) \\
&\times \mathbf{E}^Q\left[\exp\left(-s_1^*\int_0^T V(t)dt - s_2^*\int_0^T \sqrt{V(t)}dt + s_3^* V(T)\right)\right] \\
&= \exp\left(i\phi(x_0 + rT) - s_3^*(V_0 + \kappa\theta T)\right) \\
&\times \exp\Big(H_6(T; s_1^*, s_2^*, s_3^*)V_0 + H_7(T; s_1^*, s_2^*, s_3^*)\sqrt{V_0} + H_8(T; s_1^*, s_2^*, s_3^*)\Big) \\
&= \exp(i\phi(x_0 + rT)) f_2^{DSR}(\phi)
\end{aligned}
\tag{3.43}
$$

with

$$
\begin{aligned}
f_2^{DSR}(\phi) = \exp\big(&-s_3^*(V_0 + \kappa\theta T) + H_6(T; s_1^*, s_2^*, s_3^*)V_0 \\
&+ H_7(T; s_1^*, s_2^*, s_3^*)\sqrt{V_0} + H_8(T; s_1^*, s_2^*, s_3^*)\big),
\end{aligned}
\tag{3.44}
$$

where

$$
\begin{aligned}
s_1^* &= \frac{i\phi}{2}\left(1 - i\phi(1-\rho^2) - \frac{2\lambda\rho}{\sigma}\right), \\
s_2^* &= -\frac{i\phi\kappa\rho}{\sigma}, \qquad s_3^* = \frac{i\phi\rho}{\sigma}.
\end{aligned}
$$

Therefore, the option pricing formula for the model, where stochastic volatilities follow a double square root process, can be given analogously to the above section, and share all features and numerical implementation as the Heston model and the Schöbel-Zhu model.

3.4.2 Numerical Examples

We discuss this model briefly via some numerical examples. In order to compare this stochastic volatility model with the Black-Scholes model, we need to calculate suitable BS benchmark values, as discussed in the previous section. For this purpose we choose two possible BS prices:[8] The first one called BS_1 is computed according to the mean level of stochastic volatilities θ that is implied by $\theta = \sigma^2/(4\kappa)$. The other one denoted by BS_2 is the BS price by choosing the spot volatility as constant volatility. Each panel below has its own BS_2 as benchmark. Table (3.5) shows three data panels based on formulas (3.42) and (3.44) to demonstrate the impact of ρ on option values. We choose $\lambda = 1.0$ and $\kappa = 0.3$ which seems to be small compared with our previous two models.[9] and implies a θ being equal to 0.208. All panels in

[8] Because the expected average variance with a double square root process can not be given analytically, we can not calculate a benchmark according to the expected average variance, as with an Ornstein-Uhlenbeck process in the above subsection. Additionally, due to the restriction $\sigma^2 = 4\kappa\theta$, we can not get a good benchmark by letting $\sigma \to 0$, as done in Stein and Stein (1991).

[9] The Longstaff's (1989) empirical study shows that the values of the parameters λ and κ in a term structure model of interest rates indeed are of an order of 10^{-2} or 10^{-3} respectively.

Table (3.5) display the similar properties as shown in Table (3.2). Firstly, we find again that options with different moneyness react to the correlation ρ oppositely. For ITM options, their values decrease with the values of correlation. And OTM options display a wholly opposite relation with the correlation ρ to ITM options. Secondly, the finding in Subsection 3.3.3 that the bias between the long term mean θ and the spot volatility is important for the prices of options is confirmed. In Panel K we have $\theta = 0.208$ and spot volatility as $\sqrt{0.04} = 0.2$, then option prices in this panel are overall near these two BS benchmarks regardless of moneyness and correlation. In contrast to Panel K, the panels L and M are calculated with a spot volatility of 0.1 and 0.3 respectively. Consequently, the option prices in these panels deviate from BS_1 and BS_2 significantly. Finally, as reported in Subsection 3.3.3, options with different moneyness present different sensitivity to the correlation ρ. While OTM options are generally sensitive to correlation, ITM options do not change remarkably with ρ. Not surprisingly, we can calibrate the implied volatilities from the data with the negative correlation in Table (3.5) to capture the empirical volatility smile and sneer pattern.

Table (3.6) shows the impact of κ on option prices. Panel N gives all BS_1 values since every κ corresponds to a θ, and each BS_1 is then the benchmark for the option prices in the corresponding position in the panels O, P and Q. In the case of $\kappa = 0.1$, θ has an extraordinary high value of 0.625 and leads to unsuitable benchmarks. One obvious finding is that the option price is a decreasing function of κ regardless of the moneyness of options. This phenomenon can be explained as follows: For a fixed σ, the values of the explicit $\theta = \sigma^2/(4\kappa)$ go up with the falling values of κ, and imply a larger mean level of the weighted sum $\kappa\sqrt{V(t)} + \lambda v(t)$. Hence, option prices increase correspondingly. A comparison of panels O, P and Q shows that the true values of options result from a trade-off between long-term mean and spot volatility. By its very nature, the BS model can not reflect the dynamic effect of varying volatilities on the option prices. Based on the above observations most of which are identical to the models of square root process and Ornstein-Uhlenbeck process, we conclude that modeling stochastic volatility with a double square root process also constitutes a promising way in improving the pricing and hedging performance in practice.

3.5 Other Stochastic Volatility Models

In above sections, we have addressed three stochastic volatility models that admit analytical CFs under different measures, and therefore also analytical pricing formulas in form of the Fourier inversion. There are some extensions of these three stochastic volatility models. For example, Nögel and Mikhailov (2003) considered the Heston model with time-dependent parameters by iteratively solving the valuation PDEs at discrete time points. Popovici (2003) extended the Schöbel-Zhu model by introducing a time-deterministic function for mean level θ, and the new process for stochastic volatility reads

ρ \ K	90	95	100	105	110	115	120
$\mathrm{BS}_1(\theta)$	15.241	11.518	8.358	5.817	3.885	2.492	1.539
$\mathrm{BS}_2(V)$	15.118	11.342	8.142	5.584	3.658	2.293	1.377
-1.00	15.563	11.529	7.811	4.539	1.928	0.359	0.003
-0.75	15.435	11.407	7.753	4.646	2.305	0.891	0.274
-0.50	15.296	11.276	7.688	4.727	2.576	1.265	0.588
-0.25	15.144	11.128	7.605	4.786	2.789	1.569	0.873
0.00	14.977	10.956	7.501	4.826	2.987	1.831	1.132
0.25	14.791	10.753	7.371	4.848	3.150	2.062	1.371
0.50	14.582	10.505	7.206	4.853	3.292	2.271	1.595
0.75	14.353	10.184	6.993	4.839	3.420	2.465	1.806
1.00	14.188	9.681	6.681	4.813	3.546	2.655	2.013

K: $V_0 = 0.04, \lambda = 1.0, \kappa = 0.3, \sigma = 0.5, T = 0.5, S = 100, r = 0.0953$

ρ \ K	90	95	100	105	110	115	120
$\mathrm{BS}_2(V)$	14.223	9.668	5.684	2.765	1.081	0.3335	0.083
-1.00	14.814	10.524	6.498	3.018	0.612	0.008	0.00
-0.75	14.736	10.417	6.449	3.131	0.972	0.188	0.034
-0.50	14.647	10.314	6.378	3.197	1.228	0.411	0.141
-0.25	14.555	10.200	6.289	3.241	1.431	0.614	0.275
0.00	14.460	10.070	6.179	3.269	1.600	0.797	0.416
0.25	14.364	9.924	6.042	3.280	1.743	0.963	0.557
0.50	14.274	9.757	5.870	3.273	1.866	1.115	0.694
0.75	14.207	9.572	5.643	3.247	1.971	1.255	0.828
1.00	14.188	9.421	5.298	3.190	2.065	1.388	0.959

L: $V_0 = 0.01, \lambda = 1.0, \kappa = 0.3, \sigma = 0.5, T = 0.5, S = 100, r = 0.0953$

ρ \ K	90	95	100	105	110	115	120
$\mathrm{BS}_2(v)$	16.887	13.615	10.784	8.395	6.629	4.848	3.604
-1.00	16.784	13.082	9.701	6.704	4.159	2.147	0.772
-0.75	16.631	12.957	9.646	6.773	4.415	2.624	1.403
-0.50	16.465	12.824	9.591	6.846	4.644	2.997	1.853
-0.25	16.283	12.677	9.528	6.909	4.848	3.316	2.229
0.00	16.078	12.510	9.452	6.962	5.032	3.598	2.561
0.25	15.841	12.315	9.360	7.004	5.201	3.855	2.864
0.50	15.559	12.083	9.249	7.039	5.358	4.094	2.145
0.75	15.202	11.794	9.117	7.068	5.511	4.323	3.413
1.00	14.646	11.410	8.977	7.111	5.671	4.550	3.673

M: $V_0 = 0.09, \lambda = 1.0, \kappa = 0.3, \sigma = 0.5, T = 0.5, S = 100, r = 0.0953$

BS_2 values are calculated according to $V_0 = 0.04$.

V is the squared volatility.

Table 3.5 The impact of the correlation ρ on option prices. The prices of ITM calls (OTM calls) decrease (increase) with increasing correlations ρ.

κ \ K	90	95	100	105	110	115	120
θ =0.125	14.327	9.976	6.258	3.470	1.683	0.711	0.262
θ =0.156	14.573	10.480	7.017	4.344	2.479	1.304	0.633
θ =0.208	15.237	11.511	8.350	5.809	3.877	2.485	1.533
θ =0.313	17.150	13.928	11.131	8.760	6.794	5.197	3.925
θ =0.625	24.288	21.784	19.506	17.441	15.577	13.898	12.389

N: $\mathrm{BS}_1(\theta)$ are calculated according to $\theta = \sigma^2/(4\kappa)$.

κ \ K	90	95	100	105	110	115	120
$\mathrm{BS}_2(V)$	14.223	9.668	5.684	2.765	1.081	0.335	0.083
0.5	14.508	10.129	6.199	3.150	1.359	0.565	0.245
0.4	14.526	10.155	6.231	3.180	1.381	0.580	0.254
0.3	14.555	10.200	6.289	3.241	1.431	0.614	0.275
0.2	14.595	10.262	6.373	3.333	1.508	0.668	0.309
0.1	14.647	10.341	6.481	3.453	1.612	0.741	0.356

O: $V_0 = 0.01, \lambda = 1.0, \rho = -0.25, \sigma = 0.5, T = 0.5, S = 100, r = 0.0953$

κ \ K	90	95	100	105	110	115	120
$\mathrm{BS}_2(v)$	15.118	11.342	8.142	5.584	3.658	2.293	1.377
0.5	14.959	10.868	7.281	4.435	2.475	1.310	0.684
0.4	15.046	10.991	7.434	4.600	2.626	1.430	0.771
0.3	15.144	11.128	7.605	4.786	2.798	1.569	0.873
0.2	15.254	11.280	7.794	4.991	2.991	1.727	0.990
0.1	15.374	11.455	7.999	5.214	3.202	1.902	1.123

P: $V_0 = 0.04, \lambda = 1.0, \rho = -0.25, \sigma = 0.5, T = 0.5, S = 100, r = 0.0953$

κ \ K	90	95	100	105	110	115	120
$\mathrm{BS}_2(v)$	16.889	13.615	10.784	8.395	6.629	4.848	3.604
0.5	15.977	12.292	9.080	6.431	4.381	2.893	1.874
0.4	16.125	12.479	9.298	6.664	4.608	3.097	2.043
0.3	16.283	12.677	9.528	6.909	4.848	3.316	2.229
0.2	16.450	12.885	9.768	7.165	5.101	3.547	2.428
0.1	16.625	13.102	10.018	7.432	5.365	3.791	2.639

Q: $V_0 = 0.09, \lambda = 1.0, \rho = -0.25, \sigma = 0.5, T = 0.5, S = 100, r = 0.0953$
BS_2 values are calculated according to spot volatility,
V is the squared volatility.

Table 3.6 The impact of the velocity parameter κ on options prices. For the given model parameters, option prices increase with decreasing velocity parameters κ.

$$dv(t) = \kappa[\theta_1 + (\theta - \theta_1)e^{-\beta t}]dt + \sigma dW_2(t).$$

An analytical CF under the risk-neutral measure for this extended stochastic volatility model can be derived, as shown by Popovici. In the setting of affine jump-diffusion state processes, Duffie, Pan and Singleton (1999) set up a stochastic volatility model where the volatilities are driven by multiple mean-reverting square root processes, this model is essentially a multi-dimensional Heston model. Generally, extending a one-dimensional stochastic model to an uncorrelated multi-dimensional case is a straightforward task by applying the Fourier transform. For a more detailed discussion on stochastic volatility, please refer to Lewis (2000), Fouque, Papanicolaou and Sircar (2000), Shephard (2005), and Gatheral (2006).

3.6 Appendices

A: Derivation of the CFs with Ornstein-Uhlenbeck Process

The CF $f_1(\phi)$ can be calculated as follows:

$$
\begin{aligned}
f_1(\phi) &= \mathbf{E}\left[\exp(-rT - x_0 + (1+i\phi)x(T))\right] \\
&= \mathbf{E}\exp(-(rT + x_0) \\
&+ \ (1+i\phi)\left(x_0 + \int_0^T r dt - \frac{1}{2}\int_0^T v^2(t)dt + \int_0^T v(t)dW_1\right)\Bigg)\Bigg] \\
&= \exp(i\phi(rT + x_0)) \\
&\times \mathbf{E}\left[\exp\left((1+i\phi)\left(-\frac{1}{2}\int_0^T v^2(t)dt + \int_0^T v(t)dW_1\right)\right)\right] \\
&= \exp(i\phi(rT + x_0))\mathbf{E}\exp((1+i\phi) \\
&\times \ \left(-\frac{1}{2}\int_0^T v^2(t)dt + \rho\int_0^T v(t)dW_2 + \sqrt{1-\rho^2}\int_0^T v(t)dW\right)\Bigg)\Bigg].
\end{aligned}
$$

Note that dW s uncorrelated with dW_2, we have

$$
\begin{aligned}
f_1(\phi) &= \exp(i\phi(rT + x_0)) \\
&\times \mathbf{E}\left[\left(\exp\left((1+i\phi)\left(-\frac{1}{2}\int_0^T v^2(t)dt + \rho\int_0^T v(t)dW_2\right)\right)\right)\right. \\
&\times \ \left.\exp\left(\frac{1}{2}(1+i\phi)^2(1-\rho^2)\int_0^T v^2(t)dt\right)\right] \\
&= \exp(i\phi(rT + x_0)) \\
&\times \mathbf{E}\left[\exp\left(\frac{1}{2}(1+i\phi)(i\phi - \rho^2 - i\phi\rho^2)\int_0^T v^2(t)dt\right.\right. \\
&+ \ \left.\left.(1+i\phi)\rho\int_0^T v(t)dW_2\right)\right].
\end{aligned}
$$

Now we replace $\int_0^T v(t)dW_2$ by

$$
\frac{v^2(T)}{2\sigma} - \frac{v_0}{2\sigma} - \frac{\sigma}{2}T - \frac{\kappa\theta}{\sigma}\int_0^T v(t)dt + \frac{\kappa}{\sigma}\int_0^T v^2(t)dt
$$

and yield

$$
\begin{aligned}
f_1(\phi) &= \exp(i\phi(rT + \ln S_0)) \\
&\times \mathbf{E}\left[\exp\left(\frac{1}{2}(1+i\phi)(i\phi - \rho^2 - i\phi\rho^2)\int_0^T v^2(t)dt\right.\right. \\
&+ (1+i\phi)\rho\left(\frac{v^2(T)}{2\sigma} - \frac{v_0}{2\sigma} - \frac{\sigma}{2}T\right. \\
&- \left.\left.\left.\frac{\kappa\theta}{\sigma}\int_0^T v(t)dt + \frac{\kappa}{\sigma}\int_0^T v^2(t)dt\right)\right)\right].
\end{aligned}
$$

It follows by further calculation,

$$
\begin{aligned}
f_1(\phi) &= \exp(i\phi(rT + x_0)) \\
&\times \mathbf{E}\left[\exp\left([\frac{1}{2}(1+i\phi)(i\phi - \rho^2 - i\phi\rho^2 + \frac{2\rho\kappa}{\sigma}]\int_0^T v^2(t)dt\right.\right. \\
&= -\frac{\rho\kappa\theta}{\sigma}(1+i\phi)\int_0^T v(t)dt + \frac{\rho}{2\sigma}(1+i\phi)v^2(T) \\
&- \left.\left.\frac{\rho}{2\sigma}(1+i\phi)v_0 - \frac{\rho\sigma T}{2}(1+i\phi)\right)\right] \\
&= \exp\left(i\phi(rT + \ln S_0) - \frac{\rho}{2\sigma}(1+i\phi)v_0 - \frac{\rho\sigma T}{2}(1+i\phi)\right) \\
&\times \mathbf{E}\left[\exp\left(-s_1\int_0^T v^2(t)dt - s_2\int_0^T v(t)dt + s_3 v^2(T)\right)\right].
\end{aligned}
$$

The expansion of $f_2(x)$ follows the same way. What we need to do is to calculate the expected value

$$
\begin{aligned}
y(v_0, T) &= \mathbf{E}\left[\exp\left(-s_1\int_0^T v^2(t)dt - s_2\int_0^T v(t)dt + s_3 v^2(T)\right)\right] \\
&= \mathbf{E}\left[\exp\left(\int_0^T (-s_1 v^2(t) - s_2 v(t))dt\right)\exp(s_3 v^2(T))\right]
\end{aligned}
$$

for arbitrary complex numbers s_1, s_2 and s_3. and $s_1 v^2(t) + s_2 v(t)$ is lower bounded. According to the Feynman-Kac formula, y satisfies the following differential equation

$$
\frac{\partial y}{\partial \tau} = -(s_1 v^2 + s_2 v)y + \kappa(\theta - v)\frac{\partial y}{\partial v} + \frac{1}{2}\sigma^2\frac{\partial^2 y}{\partial v^2}
$$

with the boundary condition

$$
y(v_0, 0) = \exp(s_3 v_0^2).
$$

It can be shown that the above differential equation always has a solution of the form

$$\begin{aligned}
y(v_0,\tau) &= \exp\left(\frac{1}{2}\widetilde{D}(\tau)v_0^2 + H_4(\tau)v_0 + H_6(\tau) + s_3 v_0^2\right) \\
&= \exp\left(\frac{1}{2}(\widetilde{D}(\tau) + 2s_3)v_0^2 + H_4(\tau)v_0 + H_5(\tau)\right) \\
&= \exp\left(\frac{1}{2}H_3(\tau)v_0^2 + H_4(\tau)v_0 + H_5(\tau)\right)
\end{aligned}$$

with $H_3(\tau) = \widetilde{D}(\tau) + 2s_3$. Substituting this into the differential equation, we can obtain three differential equations that determine $H_3(\tau)$, $H_4(\tau)$ and $H_5(\tau)$:

$$\begin{aligned}
\frac{1}{2}(H_3)_\tau &= -s_1 - \kappa H_3 + \frac{1}{2}\sigma^2 H_3^2, \\
(H_4)_\tau &= -s_2 + \kappa\theta H_3 - \kappa H_4 + \sigma^2 H_4 H_3, \\
(H_5)_\tau &= \kappa\theta H_4 + \frac{1}{2}\sigma^2 H_4^2 + \frac{1}{2}\sigma^2 H_3,
\end{aligned}$$

where $H_3(0) = 2s_3, H_4(0) = 0$ and $H_5(0) = 0$. Solving these equations is straightforward but tedious.

$$\begin{aligned}
H_3(\tau) &= \frac{\kappa}{\sigma^2} - \frac{\gamma_1}{\sigma^2}\frac{\sinh(\gamma_1\tau) + \gamma_2\cosh(\gamma_1\tau)}{\gamma_4}, \\
H_4(\tau) &= \frac{(\kappa\theta\gamma_1 - \gamma_2\gamma_3)(1-\cosh(\gamma_1\tau)) - (\kappa\theta\gamma_1\gamma_2 - \gamma_3)\sinh(\gamma_1\tau)}{\gamma_1\gamma_4\sigma^2}, \\
H_5(\tau) &= -\frac{1}{2}\ln\gamma_4 + \frac{[(\kappa\theta\gamma_1 - \gamma_2\gamma_3)^2 - \gamma_3^2(1-\gamma_2^2)]\sinh(\gamma_1\tau)}{2\gamma_1^3\gamma_4\sigma^2} \\
&\quad + \frac{(\kappa\theta\gamma_1 - \gamma_2\gamma_3)\gamma_3(\gamma_4 - 1)}{\gamma_1^3\sigma^2\gamma_4} + \frac{\tau}{2\gamma_1^2\sigma^2}[\kappa\gamma_1^2(\sigma^2 - \kappa\theta^2) + \gamma_3^2]
\end{aligned}$$

with

$$\begin{aligned}
\gamma_1 &= \sqrt{2\sigma^2 s_1 + \kappa^2}, \quad \gamma_3 = \kappa^2\theta - s_2\sigma^2, \\
\gamma_2 &= \left(\frac{\kappa - 2\sigma^2 s_3}{\gamma_1}\right), \quad \gamma_4 = \cosh(\gamma_1\tau) + \gamma_2\sinh(\gamma_1\tau).
\end{aligned}$$

B: Derivation of the CFs with Double Square Root Process

Here we do the same thing as in Appendix A. Let y be the related expectation value, namely

$$y(V_0,T)=\mathbf{E}\left[\exp\left(-s_1\int_0^T V(t)dt-s_2\int_0^T\sqrt{V(t)}dt+s_3V(T)\right)\right],$$

we have the following PDE by applying the Feynman-Kac formula

$$\frac{\partial y}{\partial \tau}=-(s_1V+s_2\sqrt{V})y+(\kappa\theta-\kappa\sqrt{V}-\lambda V)\frac{\partial y}{\partial V}+\frac{1}{2}\sigma^2V\frac{\partial^2 y}{\partial V^2}$$

where τ denotes time to maturity. The boundary condition is

$$y(v_0,0)=\exp(s_3V_0).$$

For $4\kappa\theta=\sigma^2$, we think of a solution which takes the form of

$$y(v_0,\tau)=\exp(H_6(\tau)v_0+H_7(\tau)\sqrt{V_0}+H_8(\tau))$$

with $H_6(0)=s_3V_0$. Setting the corresponding derivatives into the above PDE, we obtain a system of ordinary differential equations:

$$\begin{aligned}
(H_6)_\tau &= -s_1-\lambda H_6+\frac{1}{2}\sigma^2H_6^2,\\
(H_7)_\tau &= -s_2-\kappa H_6-\frac{1}{2}\lambda H_7+\frac{1}{2}\sigma^2H_6H_7,\\
(H_8)_\tau &= \frac{1}{4}\sigma^2H_6-\frac{1}{2}\kappa H_7+\frac{1}{8}\sigma^2H_7^2.
\end{aligned}$$

This leads to the following solutions:

$$\begin{aligned}
H_6(\tau) &= \frac{\lambda}{\sigma^2}-\frac{2\gamma_1}{\sigma^2}\left(\frac{2\gamma_1\sinh(\gamma_1\tau)+\gamma_2\cosh(\gamma_1\tau)}{2\gamma_1\cosh(\gamma_1\tau)+\gamma_2\sinh(\gamma_1\tau)}\right),\\
H_7(\tau) &= \frac{2\sinh(\frac{1}{2}\gamma_1\tau)}{\gamma_1\gamma_4\sigma^2}\times\\
&\quad\left((\kappa\gamma_2-2\gamma_3)\cosh(\frac{1}{2}\gamma_1\tau)+(2\kappa\gamma_1-\gamma_2\gamma_3/\gamma_1)\sinh(\frac{1}{2}\gamma_1\tau)\right),\\
H_8(\tau) &= -\frac{1}{2}\ln\gamma_4+\frac{1}{4}\lambda\tau+\frac{\tau(\gamma_3^2-\kappa^2\gamma_1^2)}{2\gamma_1^2\sigma^2}+\frac{(\gamma_2\gamma_3-2\kappa\gamma_1^2)\gamma_3}{2\gamma_1^4\sigma^2}(\frac{1}{\gamma_4}-1)\\
&\quad+\frac{\sinh(\gamma_1\tau)(\kappa^2\gamma_1^2-\kappa\gamma_2\gamma_3-\gamma_3^2+0.5(\gamma_2\gamma_3/\gamma_1)^2)}{2\gamma_1^3\sigma^2\gamma_4},
\end{aligned}$$

where

$$\gamma_1=\frac{1}{2}\sqrt{\lambda^2+2\sigma^2s_1},\qquad \gamma_2=\lambda-\sigma^2s_3,$$

$$\gamma_3=\frac{1}{2}(\lambda\kappa+\sigma^2s_2),\qquad \gamma_4=\cosh(\gamma_1\tau)+\frac{\gamma_2}{2\gamma_1}\sinh(\gamma_1\tau).$$

Chapter 4
Numerical Issues of Stochastic Volatility Models

In this chapter we address various numerical issues associated with stochastic volatility models. The sophisticated numerical implementation of stochastic volatility models is crucial for a sound performance of the pricing engine and the model calibration, and includes some different aspects: the numerical integration of (inverse) Fourier transform, the computation of functions of complex number, especially the logarithm of complex number, the calculation of Greeks as well as the simulation of stochastic volatility process. Each of these issues has been already extensively discussed in financial literature. Here we present a comprehensive and compact treatment of these numerical issues from the point of view of practitioners. The discussion on the efficient simulations of stochastic volatility models is left in the next chapter.

Firstly we give some alternative pricing formulas for European-style options with CFs, which may be used to improve the efficiency and accuracy of the numerical integration of inverse Fourier transform in different pricing contexts. In Section 4.2, we discuss risk sensitivities in stochastic volatility models where Vega risk in the Black-Scholes model is translated into the parameter risks of stochastic volatility models. Some special risk sensitivities associated with stochastic volatility are proposed and discussed. In Sections 4.3 and 4.4, we examine two methods for the numerical integration of inverse Fourier transform in the option pricing formula, namely direct integration and FFT (Fast Fourier Transform). It will be shown in Section 4.5 that a sophisticated direct integration method, for instance, with the Gaussian integration and strike vector computation, may dominate the FFT in terms not only of efficiency but also of accuracy. Right computing the logarithm of complex number is another numerical problem, and will be discussed in Section 4.6. The issue with the logarithm of complex number is somehow underestimated in financial literature, but is very important for a correct implementation of the option pricing formula with CFs. We explain the phenomenon of the discontinuousness of the logarithms of complex number. Three algorithms are given for a sound treatment of this problem in the context of option valuation. The calibration of stochastic volatility models is of great importance with respect to their practical applications, and also extensively discussed in a large number of empirical studies. Some simple

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_4,

techniques may be used to improve the calibration in terms of speed and stability. As the last point, we briefly review the Markovian projection that is sometimes applied to calibrate the so-called local-stochastic volatility models.

4.1 Alternative Pricing Formulas with CFs

In order to compute the theoretical price of a European-style option, we need to compute the inverse Fourier transform numerically. In financial literature there are various pricing formulas expressed with CFs. In fact, the computation efficiency for options depends strongly on the choice of these alternative pricing formulas. For simplicity, we assume that the interest rate is constant in following.

4.1.1 The Formula á la Black-Scholes

We recall the general pricing formula expressed via CFs in (2.18),

$$C_0 = S_0 F_1 - Ke^{-rT} F_2 \tag{4.1}$$

where

$$F_j = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} I_j(\phi) d\phi, \quad j = 1, 2, \tag{4.2}$$

with

$$I_j(\phi) = \mathrm{Re}\left(f_j(\phi) \frac{\exp(-i\phi \ln K)}{i\phi} \right).$$

The pricing formula in (4.1) takes the similar form as the Black-Scholes formula where the exercise probabilities are given under two martingale measures explicitly. As shown before, Delta, the most important Greek, is equal to the first probability F_1 and the real possibility for exercising an option is equal to the second probability F_2. Therefore, both probabilities F_1 and F_2 are important not only for the valuation of plain-vanilla options, but also for the hedging decisions and the valuation of some exotic options. For example, the value of a digital option is determined by F_2. For many exotic structures embedded with digital options, it is convenient to calculate the option prices with (4.1).

4.1.2 The Carr and Madan Formula

An alternative pricing formula for a European-stlye call is suggested by Carr and Madan (1999) in order to apply fast Fourier transform (FFT). The formula of Carr and Madan is based on the CF under the risk-neutral measure Q,

$$f(\phi) = \mathbf{E}^Q[\exp(i\phi x(T))] \tag{4.3}$$

and takes the following form for call options

$$C_0 = \frac{\exp(-\alpha \ln K - rT)}{\pi} \int_0^\infty e^{-i\phi \ln K} I(\phi, \alpha) d\phi, \tag{4.4}$$

with an integrand function $I(\phi, \alpha)$

$$I(\phi, \alpha) = \frac{f(\phi - (\alpha + 1)i)}{\alpha^2 + \alpha - \phi^2 + (2\alpha + 1)i\phi}, \tag{4.5}$$

where α is a damping variable that controls the convergence speed of the integral in (4.4). Note that if the interest rate r is constant, the risk-neutral measure Q is identical to the second martingale measure Q_2. Therefore the CF $f(\phi)$ is also identical to $f_2(\phi)$.

An appropriate choice of α is somehow tricky. Carr and Madan provided a rule to determine the upper bound of α which makes the following expressions finite, namely

$$f(-(\alpha + 1)i) < \infty$$

or

$$\mathbf{E}^Q[S(T)^{\alpha+1}] < \infty.$$

According to Carr and Madan, one forth of the determined upper bound could be a good choice for α. However, as mentioned in Kilin (2007), it is difficult to choose an unique optimal damping variable α for all strikes and maturities in a calibration procedure if an option pricing model is calibrated to a whole smile surface.

For short maturity options whose values approach to theirs intrinsic values, the integrand $I(\phi, \alpha)$ oscillates highly over the real domain so that it is difficult to arrive at a satisfactory numerical integral. In order to overcome this problem, Carr and Madan suggested an alternative expression to calculate option price. Denote CP as the short maturity ITM option (call or put), the new proposed formula reads

$$CP_0 = \frac{e^{-rT}}{2\pi \sinh(\alpha \ln K)} \int_{-\infty}^{\infty} e^{-i\phi \ln K} I^*(\phi, \alpha) d\phi \tag{4.6}$$

with

$$I^*(\phi, \alpha) = \frac{1}{2}[\zeta(\phi - i\alpha) - \zeta(\phi + i\alpha)],$$

$$\zeta(x) = \frac{1}{1 + ix} - \frac{e^{rT}}{ix} - \frac{f(x - i)}{x^2 - ix}.$$

Here the value of α is again responsible for the control of the convergence of the integral, particularly, the steepness of the integrand $I^*(\phi, \alpha)$ near zero.

Compared to the pricing formula in (4.1), one can not obtain any exercise probability directly from Carr and Madan's formula given in (4.4) and (4.6).

4.1.3 The Attari Formula

Attari (2004) derived a pricing formula for options with CFs including only one integration. The call options may be valued with the following form,

$$C_0 = S_0 - e^{-rT} K\Big[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty I_A(\phi)d\phi\Big] \tag{4.7}$$

with

$$I_A(\phi) = \frac{[\mathrm{Re}(f(\phi)) + \phi^{-1}\mathrm{Im}(f(\phi))]\cos(\beta\phi) + [\mathrm{Im}(f(\phi)) - \phi^{-1}\mathrm{Re}(f(\phi))]\sin(\beta\phi)}{1+\phi^2} \tag{4.8}$$

and

$$\beta = \ln\Big(\frac{e^{-rT}K}{S_0}\Big).$$

The CF $f(\phi)$ is identical to the one in Carr and Madan's formula and evaluated under the original risk-neutral measure.[1] By the put-call parity, the pricing formula for puts is given by

$$P_0 = e^{-rT} K\Big[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty I_A(\phi)d\phi\Big]. \tag{4.9}$$

It is easy to find some nice features of Attari's formula. Firstly, it contains only one integral without damping variable. Secondly, the integrand $I_A(\phi)$ decays at a quadratic rate of ϕ and hence converges faster. Therefore, it should be a good choice for calibration procedure. Finally, as a result of single integral, both conventional direct integration and FFT can be applied to Attari's formula. On the other hand, as a property of the pricing formula with single integral, Attari's formula does not deliver directly any exercise probability and Greek.

4.2 Risk Sensitivities

In this section we present the standard sensitivities of a European call option in stochastic volatility models. In the Black-Scholes model, the classical sensitivities are Delta, Gamma, Vega and Theta. Vaga as a risk sensitivity for implied volatility is covered by a set of the parameters in a stochastic volatility model. As a result, we can not express the volatility risk in a single number, but in a few numbers. Generally, the risk sensitivities of exotic options are analytically not available, and must be calculated with finite difference method.

---

[1] More precisely, the CF $f(\phi)$ is calculated for $X(T) = \ln S(T) - \ln S_0 - rT$, not $X(T) = \ln S(T)$. Therefore, $f(\phi)$ is independent of $\ln S_0$ or S_0.

4.2.1 Delta and Gamma

If we use the conventional pricing formula (4.1) involving two exercise probabilities, Delta Δ as first exercise probability is a by-product of the price valuation. However, it is not this case if we use the pricing formulas of Carr and Madan, as well as of Attari. Generally, all pricing formulas including one integral do not provide Delta automatically. We have two approaches to obtaining Delta. The first approach is to drive a closed-form formula for Delta and Gamma from the corresponding pricing formula. The second is the numerical approach, namely finite difference method. Delta is the first derivative of option price with respect to spot value S_0. Note the spot value S_0 or its logarithm $x_0 = \ln S_0$ usually appears in a CF in an isolated term, and a CF may be rewritten as

$$f(\phi) = e^{i\phi \ln S_0} g(\cdot),$$

where $g(\cdot)$ denotes the rest terms of a given CF. It follows immediately

$$\Delta = \frac{\partial f(\phi)}{\partial S_0} = \frac{i\phi}{S_0} f(\phi) = \frac{i\phi}{S_0} e^{i\phi \ln S_0} g(\cdot).$$

In the following, we give the formulas for Delta with the formulations of Carr and Madan, as well as of Attari, respectively.

Carr and Madan's formula:

$$\Delta = \frac{\exp(-\alpha \ln K - rT)}{S_0 \pi} \int_0^\infty e^{-i\phi \ln K} [i\phi + (\alpha + 1)] I(\phi, \alpha) d\phi. \tag{4.10}$$

Attari's formula:

$$\Delta = 1 - \frac{e^{-rT} K}{S_0 \pi} \int_0^\infty \phi I_A'(\phi) d\phi, \tag{4.11}$$

with

$$I_A'(\phi) = \frac{[\mathrm{Re}(f(\phi)) + \phi^{-1}\mathrm{Im}(f(\phi))]\sin(\beta\phi) - [\mathrm{Im}(f(\phi)) - \phi^{-1}\mathrm{Re}(f(\phi))]\cos(\beta\phi)}{1 + \phi^2},$$

where we have utilized the fact that $f(\phi)$ in Attari's formula is independent of S_0. It is also possible to derive the corresponding formulas for Gamma, and however, is very cumbersome to derive other Greeks such as Theta. For such involved cases, the second approach of "shift and calculation" will be an easier alternative.

4.2.2 Various Vegas

Vega is the first derivative of option price with respect to the implied volatility. Since the implied volatility is dependent on strike and maturity, Vega is also a function of strike and maturity. Therefore, strictly speaking, Vega is a risk sensitivity only defined in the Black-Scholes model. In stochastic volatility models, the constant volatility is displaced by the stochastic volatility whose process is characterized by a set of model parameters. The volatility risk in terms of Vega in the Black-Scholes model is then distributed to the corresponding model parameters of the stochastic volatility. Note that Vega expresses essentially the risks associated with the parallel change of the constant volatility, hence it makes sense to formulate a similar Greek associated with the parallel change of the volatility surface. In the Heston model and the Schöbel-Zhu model, the spot value and the mean level of the stochastic volatility are responsible for the level of volatility dynamics. Hence, it is intuitive and reasonable in stochastic volatility models to define the so-called mean Vegas based in the spot volatility and the mean level.

The Heston Model:

Since the mean Vega is determined by the spot variance V_0 and the mean level θ, it is a gradient of two partial differentials

$$\mathscr{V} =< \frac{\partial C}{\partial V_0}, \frac{\partial C}{\partial \theta} > \tag{4.12}$$

or in terms of the volatility

$$\begin{aligned}\mathscr{V} &= < \frac{\partial C}{\partial V_0}\frac{\partial V_0}{\partial v_0}, \frac{\partial C}{\partial \theta}\frac{\partial \theta}{\partial \theta_v} > \\ &= < 2\frac{\partial C}{\partial V_0}v_0, 2\frac{\partial C}{\partial \theta}\theta_v >,\end{aligned}$$

where $v_0 = \sqrt{V_0}$ and $\theta_v = \sqrt{\theta}$, and we calculate the Vega with respect to volatility, not to variance, in analogy of the Black-Scholes model.

The cash amount of mean Vega labeled as mean cash Vega is the total differential,

$$\mathscr{V}_{cash} = \frac{\partial C}{\partial V_0}\Delta V_0 + \frac{\partial C}{\partial \theta}\Delta \theta$$

or

$$\mathscr{V}_{cash} = 2\frac{\partial C}{\partial V_0}v_0 \Delta v_0 + 2\frac{\partial C}{\partial \theta}\theta_v \Delta \theta_v. \tag{4.13}$$

The Schöbel-Zhu Model:

We have the following mean Vega,

$$\mathscr{V} = < \frac{\partial C}{\partial v_0}, \frac{\partial C}{\partial \theta} > . \tag{4.14}$$

The mean cash Vega is given by correspondingly,

$$\mathscr{V}_{cash} = \frac{\partial C}{\partial v_0} \Delta v_0 + \frac{\partial C}{\partial \theta} \Delta \theta. \tag{4.15}$$

In most cases, a simultaneous shift of v_0 and θ can nearly recover a parallel shift of the term structure of ATM volatility, and the above defined Vegas are rather close to the Vegas in the Black-Scholes model. However, as shown later, in the sense of dynamic hedging, the mean level θ is not relevant for the dynamic re-balance of the hedging portfolio, and only the change of spot volatility may be taken into account in the hedging decision. As a result, the Vegas associated only with spot volatility could be another alternative for a "right" volatility sensitivity. We define the spot Vegas as follows.

The Heston Model:

$$\mathscr{V} = \frac{\partial C}{\partial V_0} \tag{4.16}$$

or in terms of the volatility

$$\mathscr{V} = 2 \frac{\partial C}{\partial V_0} v_0.$$

The Schöbel-Zhu Model:

$$\mathscr{V} = \frac{\partial C}{\partial v_0}. \tag{4.17}$$

It is clear that spot Vega is only a part of mean Vega, therefore also smaller than mean Vega. In stochastic volatility models, the derivatives $\frac{\partial C}{\partial V_0}$ and $\frac{\partial C}{\partial v_0}$ can be given in an analytical form. Note that the coefficient functions $H_k(T)$ of the CFs both in the Heston model and in the Schöbel-Zhu model do not involve the spot volatility v_0 and variance V_0, the sensitivities $\frac{\partial f_j(\phi)}{\partial V_0}$ and $\frac{\partial f_j(\phi)}{\partial v_0}$ may be calculated easily as follows:

$$\frac{\partial f_j(\phi)}{\partial V_0} = [s_{2j} + H_1(T)] f_j(\phi), \quad j = 1, 2,$$

where $f_j(\phi)$ are given by (3.15) in the Heston model, and

$$\frac{\partial f_j(\phi)}{\partial v_0} = [-2 s_{3j} v_0 + H_3(T) v_0 + H_4(T)] f_j(\phi), \quad j = 1, 2,$$

where $f_j(\phi)$ are given by (3.32) and (3.36) in the Schöbel-Zhu model.

It follows immediately

$$\frac{\partial C}{\partial h} = S_0 \frac{\partial F_1}{\partial h} - e^{-rT} K \frac{\partial F_2}{\partial h}$$

with $h \in \{V_0, v_0\}$ and

$$\frac{\partial F_j}{\partial h} = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\partial f_j(\phi)}{\partial h}\frac{\exp(-i\phi\ln K)}{i\phi}\right)d\phi, \quad j = 1,2.$$

Numerical examples show that the spot Vegas are often smaller than the corresponding Vegas in the Black-Scholes model.

4.2.3 Curvature and Slope

As discussed in stochastic volatility models, the parameter σ, called often vol of vol, and the correlation parameter ρ account for the curvature and slope of a smile curve respectively. The parameter κ controls the speed of volatility's reversion to its mean level, and then plays a second-order role in managing volatility risk. It is straightforward to get the first derivative of a CF with respect to these parameters, namely $\partial f_j(\phi;h)/\partial h, h \in \{\sigma, \rho, \kappa\}$. These derivatives take a closed-form solution, however, have a very long functional forms (see Nagel, 1999). To save the space, we do not give them here.

The corresponding sensitivities share the same form as spot Vega and are given by

$$\frac{\partial C}{\partial h} = S_0\frac{\partial F_1}{\partial h} - e^{-rT}K\frac{\partial F_2}{\partial h} \tag{4.18}$$

with

$$\frac{\partial F_j}{\partial h} = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\partial f_j(\phi)}{\partial h}\frac{\exp(-i\phi\ln K)}{i\phi}\right)d\phi, \quad j = 1,2.$$

4.2.4 Volga and Vanna

Volga and Vanna are the risk sensitivities of second order, and are defined by the derivative of Vega with respect to spot vola and underlying price respectively. More precisely, we have

$$Volga = \frac{\partial \mathscr{V}}{\partial v_0} \quad \text{or} \quad \frac{\partial \mathscr{V}}{\partial V_0}, \tag{4.19}$$

and

$$Vanna = \frac{\partial \mathscr{V}}{\partial S}. \tag{4.20}$$

Obviously in stochastic volatility models, there are various definitions of Volga and Vanna based on the mean Vega and spot Vega. Let us focus on the versions of Volga and Vanna derived from the spot Vega, and refer to them as spot Volga and spot Vanna. We recall the PDE for call option in the Heston model,

$$\frac{1}{2}VS^2\frac{\partial^2 C}{\partial S^2}+\rho\sigma VS\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}\sigma^2 V\frac{\partial^2 C}{\partial V^2}+$$
$$rS\frac{\partial C}{\partial S}+\kappa(\theta-V)\frac{\partial C}{\partial V}-rC+\frac{\partial C}{\partial t}=0.$$

This is equivalent to

$$\begin{aligned}&\frac{1}{2}VS^2\Gamma+\rho\sigma VS\cdot Vanna+\frac{1}{2}\sigma^2 V\cdot Volga+\\&rS\Delta+\kappa(\theta-V)\mathscr{V}-rC+\Theta=0.\end{aligned} \tag{4.21}$$

where the pricing PDE is expressed via all relevant risk sensitivities. Note here $\mathscr{V}$ is the spot Vega. Assuming r is constant yields

$$C=\frac{1}{r}\left[\Theta+\frac{1}{2}VS^2\Gamma+\rho\sigma VS\cdot Vanna+\frac{1}{2}\sigma^2 V\cdot Volga+rS\Delta+\kappa(\theta-V)\mathscr{V}\right].$$

Therefore, the instantaneous price change of a call option can be given by the risk sensitivities,

$$\begin{aligned}\Delta_C=&\ \frac{1}{r}\left[\Theta\Delta_t+\frac{1}{2}VS^2\Gamma\Delta_S^2+\rho\sigma VS\cdot Vanna\cdot\Delta_v\Delta_S\right.\\&\left.+\ \frac{1}{2}\sigma^2 V\cdot Volga\cdot\Delta_V^2+rS\Delta\Delta_S+\kappa(\theta-V)\mathscr{V}\Delta_V\right],\end{aligned} \tag{4.22}$$

which implies a strategy how to hedge options with risk sensitivities, and also implies the following facts:

1. To some surprise, the risk sensitivities of the model parameters θ,κ,σ and ρ are not relevant for the dynamic hedging. However, their absolute values are important for the dynamic re-balance of the position of options. In the light of the numerical examples given in sections 3.3 and 3.4, the values of the mean reversion θ and the correlation ρ are the key factors to distinguish the stochastic volatility models from the Black-Scholes model.
2. The contribution of the change in volatility (variance) to the change in option price is
$$\kappa(\theta-V)\mathscr{V}\Delta_V,$$
whose sign is determined by the value $\theta-V$. This means that if the mean level of variance θ is larger than the spot variance V, the trader shall reduce Vega position to compensate for the positive price change as a response to the increasing tendency of variance. For the case $\theta<V$, we follow an analogy of reasoning and conclude that the trader shall accumulate the Vega position of his option portfolio.
3. The positions for Vanna and Volga shall be adjusted by $\rho\sigma$ and $\frac{1}{2}\sigma^2$ respectively. As the correlation coefficient ρ controls the co-movements of the stock prices and the stochastic variances, this adjustment for Vanna is very reasonable. The

volatility of variance σ, as shown before, is a parameter for the curvature of a smile curve, therefore the adjustment of $\frac{1}{2}\sigma^2$ for Volga is also plausible.

4. The mean Vegas given in (4.13) and (4.14) that may present a better parallel shift of ATM volatilities are not appropriate sensitivities in the above dynamic hedging. Instead, spot Vega is more important for the instantaneous change of the stochastic volatility under the assumption that other model parameters keep unchanged.
5. The similar hedging strategy in a Black-Scholes world is given by

$$\Delta_C = \frac{1}{r}\left[\Theta\Delta_t + \frac{1}{2}V_{BS}S^2\Gamma\Delta_S^2 + +rS\Delta\Delta_S\right],$$

where V_{BS} denotes the implied constant variance. It is clear that the volatility is not considered as a risk source for hedging in the Black-Scholes model. The additional terms in (4.22) to the Black-Scholes hedging present the addition efforts in hedging the volatility risks if the model parameters are assumed to keep unchanged in a short time. We conclude that the price change due to the stochastic volatility in the Heston model is given by

$$\Delta_C(\Delta_V) = \rho\sigma VS\cdot Vanna\cdot\Delta_v\Delta_S + \frac{1}{2}\sigma^2 V\cdot Volga\cdot\Delta_V^2 + \kappa(\theta - V)\mathscr{V}\Delta_V.$$

If we perform a Vanna-Volga hedging strategy in the Black-Scholes model, we do not have any knowledge about ρ, σ and θ. As a consequence, the hedging strategy based on the Black-Scholes's Vanna and Volga is too rough to reach an necessary hedging effectiveness.

4.3 Direct Integration (DI)

To evaluate the pricing formula given in (4.1), we must numerically calculate the integrals F_1 and F_2 directly. a direct integration may be also applied to evaluate Attari's formula. We have various methods to improve the numerical performance of the direct integration of inverse Fourier transform in terms of speed and accuracy.

4.3.1 The Gaussian Integration

There are many methods for numerical integration. The simple and usual methods are the Trapezoidal rule and the Simpson rule. To deal with the oscillating behavior of the integrands $I_j, j = 1,2$, in (4.1), an efficient and sophisticated numerical integration is required. As discussed below, the Gaussian integration method should be the best choice.

Let $y(x)$ be an arbitrary integrand function, we consider the following numerical integral

$$\int_a^b y(x)dx.$$

The Gaussian integration method allows us to integral the above arbitrary function with a number of weighted functions

$$\int_a^b y(x)dx = \frac{b-a}{2}\sum_{j=1}^{n} w_j y(x_j) + R_n \tag{4.23}$$

with

$$x_j = \left(\frac{b-a}{2}\right)q_j + \left(\frac{b+a}{2}\right).$$

Here w_j and q_j are some pre-determined Gaussian constants, and their values are dependent on how many n for abscissas points are chosen. The value R_n is the residual error of the Gaussian integration and equal to

$$R_n = \frac{(b-a)^{2n+1}(n!)^4}{(2n+1)[(2n)!]^3} y^{(2n)}(x^*)$$

for some x^* between a and b. Obviously, given a certain function $y(x)$ and an integration domain $[a,b]$, according to R_n, the larger the number of the abscissas points n is, the more accurate the Gaussian integration is. Davis and Polonsky (1965, in *Handbook of Mathematical Functions with Formulae, Graphs and Mathematical Tables* of Abramowitz and Stegun) gave the different w_j and q_j for n from 2 to 96. Even for $n = 24$ and $b - a = 10$, the residual error becomes extremely small and is of an order of 10^{-41},

$$R_{24} \approx 1.58 \times 10^{-41} y^{(2n)}(x^*).$$

For the case of $n = 96$, most personal computers can not calculate the exact value of R_n.

4.3.2 Multi-Domain Integration

In many circumstances it is sufficient to calculate $F_j, j = 1,2$, by applying the Gaussian integration up to a certain value in the positive real line due to the guaranteed convergence of CF. This means that in many cases we may simply set $a = 0$ and $b = b_{max}$ for a sufficient large b_{max}. However, due to the strong oscillation of the integrand for some strikes and maturities, it is difficult to determine an appropriate b_{max}. On the other hand, the Gaussian integration loses some accuracy for an erratic function. One way to overcome these problems is a multi-domain integration whereas we divide the whole real domain into several domains, say, $[a_0,a_1]$, $[a_1,a_2]$, $\cdots$, $[a_{k-1},a_k]$. We carry out the Gaussian integration from the first domain $[a_0,a_1]$, until the last domain does no longer contribute any significant absolute value, namely, until the following condition is satisfied,

$$|\int_{a_{k-1}}^{a_k} I_j(\phi)d\phi| < \varepsilon. \tag{4.24}$$

Here ε is a tolerance error that breaks down the integration procedure. Since the integrand $I_j(\phi)$ converges always to zero in the case of CF, the multi-domain integration will be terminated always for some k.

The advantages of a multi-domain integration is at least twofold. The first advantage is that we determine $a_k = b_{max}$ dynamically dependently on how fast the integrand converges. Note that the integrand in (4.1) behaves differently for different strikes and maturities, as well as for different parameter values of stochastic volatility models. The dynamic determination of b_{max} delivers more accurate and reliable results. Generally, the CFs of the short-maturity options oscillate much stronger than the ones of the long-maturity option, and therefore, the upper bound of the integration domain b_{max} for a short-maturity option is often larger than the one for a long-maturity option. This means that the number k of multi-domain integration for a short-maturity option is also larger than the number k for a long-maturity options. The multi-domain method then saves the computation time for long-maturity options and achieves the sufficient accuracy for short-maturity options. Secondly, dividing the whole positive real domain into some segments breaks down the potential strong oscillation of the integrand function, and improves the quality of the Gaussian integration. For practical applications, an equi-distance spacing of the domains is convenient, and a length $(a_j - a_{j-1})$ of 20 for all domains works very well.

4.3.3 Strike Vector Computation

As indicated above, computing the exercise probabilities F_j is time-consuming. To accelerate the calculation, especially to accelerate the calibration of a stochastic volatility model to market data, we have a simple and efficient trick. Note that in the integrand $I_j(\phi)$, the CF $f_j(\phi)$ is independent of the strike K. Therefore, if we calculate the exercise probabilities F_j for different strikes, as it is the case in a model calibration, we can compute $f_j(\phi)$ only once and utilize it for different strikes in the Gaussian integration. We refer to this trick as the strike vector computation (SVC). More mathematically, the SVC is equivalent to

$$\mathbf{F}_j(\phi, \mathbf{K}) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \mathbf{I}_j(\phi, \mathbf{K})d\phi, \quad j = 1, 2, \tag{4.25}$$

with

$$\mathbf{I}_j(\phi, \mathbf{K}) = \mathrm{Re}\left(f_j(\phi)\frac{\exp(-i\phi\mathbf{K})}{i\phi}\right),$$

where $\mathbf{F}_j, \mathbf{I}_j, \mathbf{K}$ denote the corresponding vectors of F_j, I_j, K, respectively. Since the main reason for the time-consuming computation of F_j is the complex function of

$f_j(\phi)$, the SVC saves time and resources remarkably, and generally accelerates the computation time $\frac{1}{2}N$ times for N strikes if we roughly estimate that the computation time for $f_j(\phi)$ and the integration time are equal. For a smile surface consisting of 20 strikes, according to the above estimation, we gain about 10 times efficiency.[2] Kilin (2007) used a same idea for his so-called caching technique where all calculated values of the characteristic functions are saved in cache (caching), and then re-used for other strikes. Additionally, SVC can be applied to any strike vector and has no restrictions to the form of strike vector. In contrast, as shown later, FFT can only be applied to the strikes whose logarithms $\ln K$ are equally spaced with a same length.

Similarly, since that the CFs $f(\phi)$ of the Carr-Madan's formula and the Attari's formula are also independent of strike K, SVC can also be applied to these two formulas.

For the calibration purpose, combining the multi-domain Gaussian integration, the SVC and Attari's formula seems to be the optimal strategy to perform the inverse numerical integrations of CFs in terms of speed and accuracy. The numerical examples in Kilin (2007) support also this strategy. However, for the pricing purpose in a trading environment where a few option prices and the related Greeks are often the first concern, the conventional pricing formula in (4.1) may be preferred to. Moveover, if the interest rate is not constant and specified as a stochastic process in an option pricing model such as in BCC(1997), Carr and Madan's formula, and Attari's formula will lose their validity, and must be modified. In contrast, the conventional pricing formula can be applied directly to stochastic interest rates.

4.4 Fast Fourier Transform (FFT)

In this section we discuss the application of Fast Fourier Transform (FFT) to option valuation and calibration. We will show briefly the algorithms of FFT, and expound the advantages and restrictions of FFT by using the formula of Carr and Madan from the point of view of practical applications.

4.4.1 Algorithms of FFT

Fast Fourier Transform (FFT) is an efficient algorithm to compute discrete Fourier transform in many scientific areas. Given a number of discrete points $\{x_0, x_1, \cdots, x_N\}$, FFT allows for a quick calculation of the corresponding CFs $f(x_j, \phi), 0 \leqslant j < N-1$. FFT is then designed to the cases where a large number of discrete CFs are simultaneously computed. For instance, digital signal processing and graph compressing are two typical applications of FFT.

[2] Five times efficiency is a conservative estimation. Kilin (2007) reported an essentially higher improvement.

A straightforward discrete Fourier transform of the pricing formula (4.4) is following

$$C_j = \frac{\exp(-\alpha \ln K_j - rT)}{\pi} \sum_{n=0}^{N-1} e^{-i \ln K_j u_n} I(u_n) w_n \eta,$$

$$= \sum_{n=0}^{N-1} e^{-i \ln K_j u_n} y_n \tag{4.26}$$

with

$$u_n = n\eta$$

and

$$y_n = \frac{\exp(-\alpha \ln K_j - rT)}{\pi} I(u_n) w_n \eta.$$

The weights w_h is the integral weights whose concrete forms depend on which numerical integral method is implemented. N is chosen to be large enough to make C_j converge to true option value. If we want to evaluate N different option prices $C_j, 0 \leqslant j < N-1$, directly, the arithmetical operations of an order $\mathscr{O}(N^2)$ have to be carried out. The basic idea of FFT is to divide the summation (4.26) into two subsequences, one is even-indexed, the other odd-indexed. In order to apply FFT, we divide the log strikes $k = \ln K$ equidistantly, namely

$$k_j = -k_{max} + \delta j, \quad j = 0, 1, \cdots, N,$$

with $k_0 = -k_{max}$ and $k_N = k_{max}$. δ is the distance of two log strikes and is equal to $2k_{max}/N$. Additionally, as pointed out by Carr and Madan, the following condition on η and δ must be satisfied

$$\delta\eta = \frac{2\pi}{N}. \tag{4.27}$$

We rewrite C_j as follows:

$$C_j = \sum_{n=0}^{N-1} e^{-ik_j u_n} y_n$$

$$= \sum_{n=0}^{N-1} e^{-(2\pi i) jn/N} e^{\delta n \pi i} y_n.$$

Denoting $M = N/2$ and setting $y_n^* = e^{\delta n \pi i} y_n$ yields

$$
\begin{aligned}
C_j &= \sum_{m=0}^{\frac{1}{2}N-1} y^*_{2m} e^{-(2\pi i)j2m/N} + \sum_{m=0}^{\frac{1}{2}N-1} y^*_{2m+1} e^{-(2\pi i)j(2m+1)/N} \\
&= \sum_{m=0}^{M-1} y^*_{2m} e^{-(2\pi i)jm/M} + e^{-(2\pi i)j/N} \sum_{m=0}^{M-1} y^*_{2m+1} e^{-(2\pi i)jm/M} \\
&= \begin{cases} C^e_j + e^{-(2\pi i)j/N} C^o_j, & \text{if} \quad j < M \\ C^e_{j-M} - e^{-(2\pi i)j/N} C^o_{j-M}, & \text{if} \quad j \geqslant M \end{cases},
\end{aligned} \tag{4.28}
$$

where C^e_j and C^o_j denote the discrete Fourier transforms on the even-indexed and odd-indexed points respectively. Due to the fact that $C^e_j = C^e_{j-M}$ and $C^o_j = C^o_{j-M}$, we are able to calculate C_j for $j \geqslant M = N/2$ via C^e_j and C^o_j. Furthermore, calculating C_j by computing two sub-sequences of length $N/2$ allows us a recursion algorithm that breaks down C^e_j and C^o_j further into two subsequences. Repeating this procedure leads finally to a sub-sequence of length of 1. This recursion algorithm is the essence of FFT that reduces the computation time $\mathscr{O}(N^2)$ of a conventional discrete Fourier transform to $\mathscr{O}(N \log_2 N)$. The difference between $\mathscr{O}(N^2)$ and $\mathscr{O}(N \log_2 N)$ is immense for a large N, however, is not significant for a small N.

4.4.2 Restrictions

We have briefly illustrated that FFT is an efficient algorithm if we calculate a large number of option prices with same maturity simultaneously. This seems to be a good news for financial engineers. But if we put the FFT into a real practical application, we will encounter more problems than what FFT promises us.

Firstly, we can only evaluate option prices for a restricted set of log strikes with the FFT. The log strikes must be spaced equidistantly as

$$
k_j = -k_{max} + \delta j, \quad j = 0, 1, \cdots, N.
$$

These equidistant log strikes cause large inconvenience to deal with smile surface and do not match to market strikes. No market liquid options are traded and quoted with the equidistant log strikes. Consequently, in order to apply the FFT, we need interpolate the quoted implied volatilities through the log strikes. The quality of any interpolation in turn depends on the method of interpolation and market data, and leads to a distortion of market information. For example, the implied market volatilities in FX markets are quoted in terms of Delta, and the corresponding implied strikes change if the implied volatilities change. An additional interpolation of implied volatilities through log strikes makes the option valuation more involved, less transparent and less accurate. The best way is then to choose the strikes with which market liquid options are traded and quoted directly.

Secondly, a by-product of the recursion algorithm is that the number N of grid points in the FFT must be a power of two, namely $N = 2^n$, where n is an integer and presents the efficiency of FFT. The computational improvement of FFT compared to conventional Fourier transform is given by $N/\log_2 N = N/n = 2^n/n$. For a number of n larger than 6, we have to evaluate more than 64 option prices per maturity. For $n = 10$, the number of calculated option prices is over 1000. This is to many in the practical applications.

Thirdly, the condition

$$\delta\eta = \frac{2\pi}{N}$$

implies a strong restriction on a free choice of the grids of integration domain, and is equivalent to

$$\eta = \frac{2\pi}{N\delta} = \frac{\pi}{k_{max}},$$

in which we can clearly observe that the grid length η of the integration domain is determined by the upper bound of the log strike k_{max}. We have only one freedom in building up the grids for log strikes and the integration domain in the FFT algorithm. If we build a fine grid for the integration, call options are calculated "at strikes spacings that are relatively large, with few strikes lying in the desired region near the stock price" since k_{max} becomes larger for a smaller η, as Carr and Madan (1999) observed. Chourdakis (2005) applied the so-called fractional FFT to overcome the restriction on η and δ. However, the restriction that log strikes must be spaced equidistantly holds still. Chourdakis (2005) and Kilin (2007) have shown the fractional FFT is faster than the standard FFT since the fractional FFT allows for a sparser grid of log strikes, and this effect is sometimes very important for accelerating the computation. Wu (2006) reported that if the models takes extreme parameter values or the bounds are set too tight, the FFT and the fractional FFT break down completely. During the calibration of a pricing model, the parameter values vary very largely and the proposed extreme values by optimization routine may cause a complete breakdown of calibration. Hence, he appealed for an additional concern for model calibration when FFT or the fractional FFT are applied.

Additionally, as the cost of the advantages of FFT, most advanced integration methods can not be combined with FFT. Particularly, the Gaussian integration method can not be applied for FFT since the grid points in the Gaussian quadrature are located according to a pre-defined scheme. Due to the equidistant spacing of grid in the integration domain, only the simple integration rules can be applied to FFT, for example, the Trapezoid rule or the Simpson rule. Therefore, the advantages of FFT are somehow traded off by the limited application of advanced numerical integration methods.

4.5 Direct Integration vs. FFT

We have discussed two numerical integration methods for the option valuation in details, the direct integration (DI) and the fast Fourier transform (FFT), that compete with each other. In this section we summarize the advantages and disadvantages of DI and FFT with respect to the following aspects.

4.5.1 Computation Speed

The most convincing argument for FFT in the literature is its ability to accelerate the speed of valuation time if we calculate option prices with CFs. In order to apply the FFT soundly and to obtain accurate prices, we have to employ a large number of strikes. Carr and Madan (1999) set N_{FFT} equal to $2^{12} = 4096$ for their numerical experiments. Furthermore, they used $\eta = 0.25$ for the integration spacing. According to the restriction $\delta\eta = \frac{2\pi}{N_{FFT}}$, we obtain $\delta = 0.00613$ and $k_{max} = 12.5664$ correspondingly if the spot price S_0 is assumed to be 1. The value 12.5664 for the upper bound of $\ln K$ is extremely large, and corresponds to an absolute value of 286751 for the upper bound of K. Correspondingly, the lower bound of strike approaches to zero. This extreme large band for strikes indicates how inflexible the strike spacing in the FFT pricing method is.

In the real trading, nobody puts so many call prices for actual calibration, and needless to say, for pricing. From the point of view of a practitioner, we usually use about 15 strikes for each maturity to calibrate an option pricing model. This means that only 15 of 4096 calculated option prices are really used, and for these 15 option prices we still need a computation effort of 4096×12.

In contrast to the FFT, if we want to calculate 15 option prices with the DI, a very conservative estimate for the computation effort is $24 \times 10 \times 15/2$ where a Gaussian quadrature with 24 abscissas points, 10 sub-domains with a length of 20, and SVC are employed.[3] The ratio of the computation speeds of FFT and DI in this realistic scenario is ca. 27 times in favor of DI, and confirms the Kilin's results. Kilin (2007) reported that the calibrations of some stochastic volatility models and Lévy models with a sophisticated DI are 31 times faster than the calibrations with the standard FFT, and 16 times than with the fractional FFT. The numerical examples of Carr and Madan (1999) was based on 160 option prices using the variance-gamma model, which are computed with a standard FFT, a simple DI and the analytic solution. In their study, the computation effort with the FFT is still 4096×12 whereas the computation time with a not optimized DI (without SVC) will increase explosively, at least 16×10 times. A corresponding comparison results in a ratio of ca. 5 times in favor of the FFT with the Carr and Madan's examples. Hence such a numerical comparison and its conclusion is questionable and unfair for DI.

Formally let N_P denote the number of option prices that we want to compute, and N the whole number of grid points in the (multi-domain) Gaussian integration, the speed ratio of the above described DI compared to the standard FFT can be estimated as

$$\frac{N_{FFT} \log_2 N_{FFT}}{\frac{1}{2} N_P N}.$$

[3] This means the upper bound for the integration domain is 200, a very large number for b_{max}.

If we want to evaluate only a few option prices for a maturity, which is often the case in the trading, FFT is extremely uneconomic. The numerical examples in Chourdakis (2005) show that the fractional FFT is almost 2 times faster than the standard FFT due to the sparse spacing of log strikes.

4.5.2 Computation Accuracy

Generally, the computation accuracy of DI and FFT mainly depends on the number of grid points in the corresponding integration methods, and which integration method is used. The residual error R_n of the DI with the Gaussian quadrature has been given above. The global error ε of the Gaussian quadrature is the maximum of R_n and ε where ε is the tolerance error to break the multi-domain integration, namely

$$\varepsilon = \max[R_n, \varepsilon].$$

Usually we always have $R_n < \varepsilon$ in the Gaussian quadrature. Therefore, the global error ε in the DI is usually equal to ε that is controlled by user.

For the case of the Simpson's rule, the error of the FFT in each integration interval is given by

$$R = c\eta^5 y^{(4)}(x^*)$$

for a constant c. The value of c depends on which version of the Simpson's rule is employed. If η takes a value of 0.25, the integration error R should be larger than $10^{-5}y^{(4)}(x^*)$. Due to the strong restriction $\eta k_{max} = \pi$, η can not be set to very small arbitrarily in the FFT, and this in turn limits the computation accuracy of the FFT. Therefore, we conclude that the computation accuracy of the FFT should not be better than the computation accuracy of the DI.

4.5.3 Matching Market Data

Another important criterion to chose a better numerical algorithm for pricing option prices is how sensible and flexible the applied numerical algorithm depends on market data. In other words, a better numerical algorithm should match the structure of market data better. This point is underestimated and and even completely ignored in financial literature, and also hardly discussed by practitioners.

The requirement for a pricing algorithm to match the structure of market data is backed by two facts. Firstly, the quoted market data are rare. Secondly, to retrieve the complete market data, especially a volatility smile surface, we need some extensive numerical techniques to fill the gaps in market data. For example, as discussed early, liquid options of all asset classes are traded and quoted in some few given strikes. If we calibrate an advanced pricing model to market data, we should not expect that the calibration quality is largely dependent on the interpolation and extrapolation quality of a smile surface. In contrast, in the modern financial industry, the advanced pricing models are calibrated to market data in order to price other illiquid and exotic products. Therefore, the advanced pricing models themselves become the models to interpolate market data, and illiquid plain-vanilla options are then priced via the calibrated parameters, instead of via the ad-hoc interpolation techniques, for example, the linear or cubic spline interpolation methods. In other words, the ad-hoc interpolation techniques are redundant if a pricing model is calibrated soundly. For example, if we have calibrated a stochastic volatility model to liquid Euro-Stoxx options, other exotic Euro-Stoxx derivatives such as barrier options can be priced by the calibrated stochastic volatility model, and we do not need any ad-hoc interpolation method.

Due to its inflexible spacing of log strikes, and due to many grid points and a large band for log strikes, FFT obviously can not satisfy the requirement for a matching of the pricing algorithm to the structure of market data, and must be supported by some extensive interpolation and extrapolation methods for smile surface. In contrast, the valuation of options with DI allows for flexible strikes and imposes no restrictions on the particular spacing of strikes. Hence, any liquid option can be included in a model calibration with DI. With respect to this aspect, DI is preferred to FFT.

4.5.4 Calculation of Greeks

The efficient calculation of risk sensitivities (Greeks) is the next aspect to compare DI and FFT, and unfortunately, is also ignored by most financial literature. In a real trading environment, risk sensitivities are equally important as prices. Therefore, more attentions should be paid on the efficient calculation of risk sensitivities, if a numerical pricing method is implemented. Generally we will achieve a higher efficiency with DI in computing a moderate number of prices and Greeks than with FFT. This implies that we shall void the FFT pricing method for real trading and hedging.

4.5.5 Implementation

Finally we consider the issue of the implementation of FFT and DI. Open source codes for both the Gaussian quadrature and FFT are available. The implementation of DI is straightforward whereas FFT involves a recursion algorithm and requires more care. The decisive difference between FFT and DI lies in the dampening parameter α of Carr and Madan's formula. As mentioned in Kilin (2007), it is impossible to find a single value of α that leads to an acceptable pricing error for all possible model parameters. Some suggestions for a better control of α (see Lee (2004), Lord and Kahl (2006)) rest on a fine grid of FFT which again increases significantly the upper bound of log strikes, and finally in turn slows down the calculation and calibration.

Summing up the aspects discussed above, FFT is not necessarily an efficient numerical method for option pricing although it is powerful in many scientific applications. As shown here, FFT is remarkably slower than a sophisticated direct integration. Thus, FFT is a good algorithm for a "wrong" application in the context of option valuation and calibration. This conclusion is contradictory to what Carr and Madan have claimed.

Finally, I give prototype programming codes for calculating European-style options in the following on readers's behalf. The implemented DI is a combination of the above described multi-domain integration and strike vector computation.

```
\\This function implements the exponents of CFs
\\ of stochastic volatility models.
Complex CF_Vola(int index, double phi, ProcessType type) {
    ....
}

//This function performs strike vector integration, two CFs version
double[] CFs(int index, double phi, double[] strikes)
{
    //If index is 0, we go to one CF version
    if ( index == 0 )
       return CFs(phi, strikes);

    Complex value = new Complex(0);
    value = CF_Vola(index, phi);

    double[] x = new double[strikes.Length];
    Complex y, v;
    for (int i = 0; i < strikes.Length; i++)
    {
        y = new Complex(0, Log(spot / strikes[i]) * phi, 0);
        v = value + y;
        x[i] = Exp(v.Real) * Sin(v.Imaginary) / phi;
    }
    return x;
}

// This function performs strike vector integrations, one CF version
// according to Attari's formulation
double[] CFs(double phi, double[] strikes)
{
    Complex value= new Complex(0);
    double r = interestRate;
    Complex lns = new Complex(0, ( Log(spot)+ r*t) * phi);

    // Note Attari formulation is not valid for stochastic interest rate
    // CF_Rate is not integrated
    value = CF_Vola(2,  phi)
    value = value.Exp();

    double real = value.Real;
    double imag = value.Imaginary;

    double[] x = new double[strikes.Length];
    double beta, v;
    for (int i = 0; i < strikes.Length; i++)
    {
        beta = Log( Exp(-r * t) * strikes[i] / spot) * phi;
        v = (real + imag / phi) * Cos(beta) + (imag - real / phi) * Sin(beta);
```

```
        x[i] = v/(1+phi*phi);
    }

    return x;
}

// This function performs strike vector integration
// for interval between first and end
double[] Integrals(int index, double first, double end, double[] strikes)
{
    int N = 10;
    double[] x, w;
    x = new double[N + 1];
    w = new double[N + 1];
    int size = strikes.Length;
    double[][] v = new double[2 * N + 1][];

    x[1] = 0.076526521133497;
    x[2] = 0.0227785851141645;
    x[3] = 0.373706088715419;
    x[4] = 0.519867001950827;
    x[5] = 0.636053680726515;
    x[6] = 0.746331906460150;
    x[7] = 0.839116971822218;
    x[8] = 0.912234428251325;
    x[9] = 0.963971927277913;
    x[10] = 0.993128599185094;

    w[1] = 0.152753387130725;
    w[2] = 0.149172986472603;
    w[3] = 0.142096109318382;
    w[4] = 0.131688638449176;
    w[5] = 0.118194531961518;
    w[6] = 0.101930119817240;
    w[7] = 0.083276741576704;
    w[8] = 0.062672048334109;
    w[9] = 0.040601429800386;
    w[10] = 0.017614007139152;

    double xm = 0.5 * (end + first);
    double xr = 0.5 * (end - first);
    double dx;
    double[] ss = new double[size];
    int j;

    // set initial value for first call
    if (iteration == 1)
    {
        v[0] = new double[size];
        v[0] = CFs(index, 0.000000001, strikes);
    }

    for (j = 1; j <= N; j++)
    {
        dx = xr * x[N - j + 1];
        v[j] = new double[size];
        v[j] = CFs(index, xm - dx, strikes);
    }

    for (j = N + 1; j <= 2 * N; j++)
    {
        dx = xr * x[j - N];
        v[j] = new double[size];
        v[j] = CFs(index, xm + dx, strikes);
    }

    for (int i = 0; i < size; i++)
```

```
    {
        for (j = 1; j <= N; j++)
        {
            ss[i] += w[j] * (v[N - j + 1][i]
                  + v[N + j][i]);
        }
        ss[i] *= xr;
    }
        return ss;
}

// This function calculates two exercises probabilities
// for index = 1 or 2 with multi-domain integration
double[] prob(int index, double[] strikes)
{
    int size = strikes.Length;
    double[] sum1 = new double[size], sum = new double[size];
    iteration = 0;
    int num = 0;
    while (( iteration <= 100) && (num < size) )
    {
        iteration += 1;
        //step is the length of domain on which Gaussian integration is performed
        sum = Integrals(index, step * (iteration - 1), step * iteration, strikes);
        for (j = 0; j < size; j++)
        {
            if (Abs(sum[j]) >= exact && (sum[j] != double.NaN) )
                sum1[j] += sum[j];
            else num += 1;
        }
    }

    for (int j = 0; j < size; j++)
        sum1[j] = 0.5 + sum1[j] / PI;

    return sum1;
}

// This function calculates option prices with f_1 and f_2
// like the Black-Scholes formula
double[] price(double[] strikes)
{
    int size = strikes.Length;
    double[] value = new double[size];

    double[] p1 = prob(1, strikes);
    double[] p2 = prob(2, strikes);

    for (int i=0; i< size; i++)
    {
        if ( isCall )
            value[i] = spot * p1[i] - df * strikes[i] * p2[i];
        else
            value[i] = -spot * (1.0 - p1[i]) + df * strikes[i] * (1.0 - p2[i]);
    }
    return value;
}

// This function calculates option prices with Attari's formula
double[] price_Attari(double[] strikes)
{
    int size = strikes.Length;
    double[] value = new double[size];
    double[] sum1 = new double[size], sum = new double[size];
    iteration = 0;
    int num = 0;
    while ((iteration <= 100) && (num < size))
```

```
    {
        iteration += 1;
        // we call the function Integrals with index =0.
        sum = Integrals(0, step * (iteration - 1), step * iteration, strikes);
        for (int j = 0; j < size; j++)
        {
            if (Abs(sum[j]) >= exact && (sum[j] != double.NaN))
                sum1[j] += sum[j];
            else
                num += 1;
        }
    }

    for (int i = 0; i < size; i++)
    {
        if ( isCall )
            value[i] = spot - df * strikes[i] * (0.5 + sum1[i] / PI);
        else
            value[i] =  m_df * strikes[i] * (0.5 - sum1[i] / Math.PI);
    }

    return value;
}
```

4.6 Logarithm of Complex Number

In this section we discuss some issues associated with the logarithm of complex number. In the Heston model and the Schöbel-Zhu model as well as in some Lévy jump models, the CFs content the logarithm of complex number that must be handled carefully in the implementation. Schöbel and Zhu (1999) have first mentioned the latent problem of the discontinuity of the logarithm of complex number in the context of option pricing. After then the logarithm of complex number becomes a serious part of the robust and accurate numerical implementation of the advanced option pricing models. Meanwhile many researchers provide some new, partially surprising, insights on the logarithm of complex number.

4.6.1 Definition

Denote $c = a + i \cdot b$ as a complex number. Its exponential form is given by

$$c = z_1 e^{iz_2}$$

with

$$z_1 = |c| = \sqrt{a^2 + b^2},$$

$$z_2 = z_0 + 2\pi k, \quad -\pi < z_0 \leqslant \pi,$$

where k is an arbitrary integer. $z_0 = arg(c) \in [-\pi, \pi)$ is called the main argument, and z_1 is the modulus or radius of the complex number c. Due to the Euler's equation

$\cos(z)+i\sin(z)=e^{iz}$, and the well-known properties of the trigonometric functions $\cos(z+2\pi k)=\cos(z)$ and $\sin(z+2\pi k)=\sin(z)$, we have

$$c = z_1 e^{iz_2} = z_1 e^{iz_0}.$$

Obviously a complex number c is uniquely determined by its modulus z_1 and main argument z_0. However, the logarithm of c does not uniquely depend on z_1 and z_2, and is defined by

$$\begin{aligned} y &= \ln c = \ln[z_1 e^{iz_2}] \\ &= \ln z_1 + iz_2 = \ln z_1 + i(z_0 + 2\pi k). \end{aligned} \tag{4.29}$$

It turns out that y is a multi-valued function since we can choose any k to recover $e^y = c$, and the main argument z_0 defines only one of all possible values. If we consider the logarithm of c in an isolated computation, it is not important which k is chosen for y, and in fact we usually use the main argument, i.e. $k=0$ for y. This approach is also used in all commercial mathematical softwares, for example, MathLab and Mathematica. However, this simplified approach is fatal in the implementation of some stochastic volatility models involving the logarithm of complex number since the fixed main argument for the logarithm of complex number will lead to a discontinuity of the integrand functions $I(\phi), I_j(\phi), I_A(\phi)$ in the respective pricing expressions, and therefore, also leads to wrong option prices. In some worst cases, the main argument causes even negative prices. In other words, if we compute the CFs, we must be careful to choose an appropriate value of k to make the integrand smooth over the real axis. Moreover, it is impossible to find a value of k prior to calculation to ensure a continuous integrand. An appropriate value of k can be found only in a context of computation where the past path of the integrand is known. The careless usage of the main argument in the logarithm of complex number is sometimes referred to as branch cut in literature.

As pointed out by Kahl and Jäckel (2006), the discontinuity problem of complex logarithm is strongly related to the discontinuity problem of complex power function. To see this, we consider a power function of c,

$$c^u = z_1^u e^{iz_2 u} = z_1^u[\cos(uz_2) + i\sin(uz_2)].$$

Obviously, uz_2 is not necessarily equal to the main argument of c^u. If $uz_2 \neq arg(c^u)$, any branch cut leads to jumps in the values of c^u. Since the Taylor's expansion of a logarithm function is the series of power functions, it is not surprising that both logarithm functions and the power functions of complex number suffer from the branch cut problem. In particular, the square root function of complex number has also branch cut. In the context of option pricing, the branch cut of complex logarithm seems to raise more serious problems than the branch cut of square root function.

4.6.2 Three Algorithms Dealing with Branch Cut

We here introduce three simple algorithms to deal with the logarithm of complex number. All algorithms can be applied efficiently for practical applications.

Approximation Algorithm:

This method is based on the first order Taylor's expansion for the imaginary part of the logarithm of complex number, and applied by Schöbel and Zhu (1999). Consider $y = \ln(a+ib)$ as the function of a and b. Given $y_0 = \ln(a_0 + ib_0)$, we have the following first order approximation,

$$\begin{aligned}\Delta_y &\approx \frac{\partial y}{\partial a}\Delta_a + \frac{\partial y}{\partial b}\Delta_b \\ &= \frac{\Delta_a}{a_0 + ib_0} + \frac{\Delta_b i}{a_0 + ib_0} \\ &= \frac{(a_0\Delta_a + b_0\Delta_b) + i(a_0\Delta_b - b_0\Delta_a)}{a_0^2 + b_0^2}.\end{aligned} \tag{4.30}$$

Taking the imaginary part of Δ_y yields

$$\operatorname{Im}(\Delta_y) \approx \frac{a_0\Delta_b - b_0\Delta_a}{a_0^2 + b_0^2}.$$

This is equivalent to

$$\operatorname{Im}(y) = z_0 + 2\pi k \approx \operatorname{Im}(y_0) + \frac{a_0\Delta_b - b_0\Delta_a}{a_0^2 + b_0^2} = \operatorname{Im}^*(y).$$

The integer k may be estimated by

$$k = round\Big(\frac{\operatorname{Im}^*(y) - z_0}{2\pi}\Big), \tag{4.31}$$

where $round(x)$ denotes the nearest integer to x. The above estimation for k works very well and quickly for a small increment of $|y - y_0|$. Of course, we can use a second order of Taylor's expansion for the estimation of k and achieve more accuracy.

Iteration Algorithm:

Priestly (1990) give an iteration algorithm to find a right k. Let k_0 be the argument of y_0. Priestly's algorithm is simple and reads:

1. If $\operatorname{Im}(y - y_0) < -\pi$, set $k = k_0 + 1$, until $\operatorname{Im}(y - y_0) > -\pi$.
2. If $\operatorname{Im}(y - y_0) > \pi$, set $k = k_0 - 1$, until $\operatorname{Im}(y - y_0) < \pi$.
3. Else $k = k_0$.

Finally y is replaced by $y+2\pi ki$. In many cases, the iteration method requires more calculations than the approximation method. Both two algorithms can be generally applied to any complex logarithm, and should be better implemented in a class of complex number.

Rotation Count Algorithm:

Rotation count algorithm is suggested by Kahl and Jäckel (2005) and is especially tailored to deal with the logarithm of complex number in the Heston model. Note that the logarithm terms in the original Heston's CFs take the following form,

$$\ln\Big(\frac{1-g_j e^{d_j T}}{1-g_j}\Big), \quad j=1,2,$$

where d_j and g_j are defined in (3.16).

Now we consider a particular complex number as follows:

$$c = ge^{dT}-1 = |g|e^{iz+d}-1$$

with $g=|g|e^{iz}$ and d as arbitrary complex numbers. The rotation count algorithm to calculate $\ln(c)$ consists of the following three steps:

1. Calculate the argument of ge^d with

$$n = round\Big(\frac{z+\mathrm{Im}(d)+\pi}{2\pi}\Big).$$

   If $\frac{z+\mathrm{Im}(d)+\pi}{2\pi} \in [-\pi,\pi)$, then n is equal to 0.
2. Calculate $|c| = |ge^d-1|$ and $z^* = arg(ge^d-1)+2n\pi$.
3. Calculate $\ln(c) = |c|e^{iz^*}$.

Kahl and Jäckel observed that the subtraction of 1 from ge^{dT} does not change the argument of c. Therefore, in the context of the Heston model, the argument of g_j that may cause the discontinuity in the integrand, is canceled out in $\frac{1-g_j e^{d_j T}}{1-g_j}$. Additionally, Kahl and Jäckel (2005), Lord and Kahl (2008) assumed that the discontinuity e^{dT} does not have an impact on the characteristic function.[4] Given this premise, two successive applications of the rotation count algorithm to $1-g_j e^{d_j T}$ and $1-g_j$ respectively should deliver a right logarithm appearing in the Heston model. In more details, we should carry out the following calculations:

1. Apply the rotation count algorithm to $\alpha = g_j e^{d_j T}-1$ and obtain $\alpha = |\alpha|e^{iz^*}$.
2. Apply the rotation count algorithm to $\beta = g_j - 1$ and obtain $\beta = |\beta|e^{iz^{**}}$.
3. Evaluate

$$\ln\Big(\frac{1-g_j e^{d_j T}}{1-g_j}\Big) = \frac{|\alpha|}{|\beta|}e^{i(z^*-z^{**})}.$$

[4] As mentioned by Lord and Kahl (2008), this premise is not proven completely, and partially contradicts the argument given in Albrechter, Mayer,Schoutens and Tistaert (AMST, 2006).

The rotation count algorithm is a procedure of self-adaption, and in contrast to the approximation algorithm and the iteration algorithm, does not need the information of the neighbored point on the path. On the other hand, the rotation count algorithm can be applied only to the special structure of complex number as shown above. Lord and Kahl (2008) provided a reformulation of the Schöbel-Zhu's formula so that it contents the Heston's CFs. Aa a result of this reformulation, the rotation count algorithm could be applied to the Schöbel-Zhu model.

4.6.3 When Main Argument Is Appropriate

We have three algorithms to deal with the potential discontinuity of logarithm of complex number. The approximation algorithm and the iteration algorithm are two general approaches for any case while the rotation count algorithm is based on the special structure of a complex number c, and provides us with a hint that the internal structure of the calculated complex number may be the trouble for discontinuity. In other words, the special algebraical structure of a complex number leads it often cross the negative real axis, and the resulting branch cut then makes trouble. This conjecture has confirmed by an observation that the alternative formulation of the Heston's CFs in (3.26) does never cause any discontinuity in practical applications even if the main argument is used, as found in AMST (2006). This seems surprising. To illustrate this fact, let us recall the complex logarithm in (3.26),

$$y_j = \ln\left(\frac{1-q_j e^{-d_j T}}{1-q_j}\right) \tag{4.32}$$

with

$$q_j = 1/g_j = \frac{b_j - d_j - \rho\sigma i\phi}{b_j + d_j - \rho\sigma i\phi},$$

$$d_1 = \sqrt{(\rho\sigma i\phi - b_j)^2 + \sigma^2(\phi^2 - i\phi)},$$

$$d_2 = \sqrt{(\rho\sigma i\phi - b_j)^2 + \sigma^2(\phi^2 + i\phi)},$$

$$b_1 = \kappa - \rho\sigma, \quad b_2 = \kappa.$$

In contrast to Kahl and Jäckel (2005), AMST found that the discontinuity stems from that $e^{d_j T}$ in the original CFs in the Heston model is a spiral with the exponentially growing radius. Instead of the term $e^{d_j T}$, the new formulation in (4.32) uses the term $e^{-d_j T}$. This small modification affects significantly the behavior of the argument of y_j. It may be proven (see AMST (2006), Lord and Kahl (2008)) that under usual conditions y_j never crosses the negative real axis, and therefore, the main argument is appropriate to generate a smooth path of complex logarithm.

However, with respect to the term $e^{-d_j T}$, we observe that the formulation of CFs in (3.15) still encounters the path discontinuity although we can verify $e^{-\gamma_1 T} =$

$e^{-d_j T}$. If the observations of AMST (2006), Lord and Kahl (2008) are true in the context of the Heston model, to some surprise, only the original Heston's formula and the formulation of CFs in (3.15) have the trouble of the discontinuity of the logarithm of complex number. Not to forget that the above discussions are based only on the Heston model, not on other stochastic volatility models and Lévy models. For a general and cautious treatment of the logarithm of complex number that may appear in any pricing formula with CFs, we recommend the application of the model-free algorithms, namely, the approximation algorithm and the iteration algorithm.

4.7 Calibration to Market Data

Model calibration is a key step to apply stochastic volatility models in practice. By calibrating a particular stochastic volatility model to a market volatility surface, we can find the model parameters consistent to market, and then use them to value exotic structures.

4.7.1 General Procedure

Finding the best model parameters to fit market data is not a task of finding the roots of an equation system, but a task of minimizing the errors between the model prices and the market prices in a given norm. In the Heston model and the Schöbel-Zhu model, we have five parameters $\Phi = \{V_0(v_0), \kappa, \theta, \sigma, \rho\}$ for estimation. We consider an implied volatility surface of N strikes and M maturities, a calibration procedure is equivalent to the following optimization problem,

$$\Phi^* = arg \min_{\Phi} \Big(\frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} \mid C_{ij}^{Model}(\Phi) - C^{Maket}(\sigma_{ij}^{BS}) \mid \Big),$$

where σ_{ij}^{BS} is the implied volatility of an option with i-th strike and j-th maturity, and $|\cdot|$ stands for an arbitrary norm. Applying the usual L_2 norm, namely a measure for mean square errors, the above optimization problem is rewritten as

$$\Phi^* = arg \min \Big(\frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} [C_{ij}^{Model}(\Phi) - C^{Maket}(\sigma_{ij}^{BS})]^2 \Big).$$

Since the absolute level of an option price is determined by the underlying asset price S_0, we can standardize the calibration error by dividing the option prices by S_0 for a more transparent comparison between different calibration sessions. There are a number of various strategies to perform the calibration:

1. Note that the deep OTM call option prices are very small and will lose their weights in the total errors for optimization, we can use put option prices for $K > S_0$ and call option prices for $K < S_0$ in a calibration.
2. Instead of estimating price errors, we can use the implied volatilities to measure the distance between model and market. However, this approach is not numerically efficient because we have to calculate the implied model volatilities at each optimization step, which is a very time-consuming, and not stable, especially for some extreme parameter values proposed by optimization routine.
3. Instead of estimating absolute price errors, we can use relative price errors for calibration. One drawback of the relative price errors lies in the latent overweights of deep-ITM and deep-OTM options in error function.
4. We can control the error function by adding the pre-defined weights to $C_{ij}^{Model}(\Phi) - C^{Maket}(\sigma_{ij}^{BS})$. If we want achieve a more perfect fitting in the area of ATM, we can give the near-ATM option prices some larger weights. Gammas and the Black-Scholes Vegas may be used as weights to control the calibration procedure. But the question is how the subject manipulation of the error function affects the valuation of exotic options, for example, barrier options if a model is calibrated more intensively in a strike range than other strike range.
5. To avoid unreasonable parameters values, we should add some constraints for the parameters in an optimization routine, or impose some penalties on undesired values in an error function.

According to my experiences, the five parameters of the stochastic volatility model exhibit different calibration stabilities with respect to the parameters constraints and error penalties. The following relation illustrates roughly the calibration stability of the five parameters in a standard market environment,

$$v_0(V_0) > \rho > \sigma > \theta > \kappa.$$

This means that spot volatility or variance displays the strongest stability while the reversion velocity parameter κ oscillates largely in a calibration.

4.7.2 Fixing Velocity Parameter

In practical calibrations, we can often observe that if we fix the reversion velocity parameter κ to be a constant, the calibration can accelerated remarkably, in many cases up to 3 times faster. This phenomena is perhaps due to the low sensitivity of option price to κ, and the related high instability of κ in a calibration. We find the strong evidences that a fixed velocity parameter improves the calibration stability and the plausibility of the estimated mean level θ. For an illustration, we have listed the estimated parameters of the Heston model in Table (4.1), that are calibrated to the FX USD-EUR option market as of July 3, 2008, according to two different calibration strategies. In the case of the unfixed κ, the value of θ seems to be too larger and is less realistic we compare it with the 4 year ATM FX implied volatility. The spot volatility keeps constant nearly for both cases.

The additional merit of a fixed velocity parameter is that we can set a rather large value for κ so that Feller's condition $2\kappa\theta > \sigma^2$ for the positive variances $V(t)$ could be satisfied in the Heston model. In Table (4.1), it can be easily verified that the condition for the positive variances is only fulfilled in the case of fixed $\kappa = 2$, but not in the case of unfixed velocity parameter. Therefore, fixing velocity parameter is not only a strategy to accelerate the calibration, but also a strategy to overcome possible negative values of variances in the Heston model.

Parameters	Velocity not fixed	Velocity fixed
V_0	0.0107	0.01047
θ	0.411826	0.012566
σ	0.105143	0.266236
ρ	−0.166304	−0.153689
κ	0.003909	2

Table 4.1 The calibrated Heston model to FX USD-EUR option market on July 3, 2008. The maximal maturity is 4 years. For the case of fixed velocity, κ is fixed to 2.

Parameters	Velocity not fixed	Velocity fixed
v_0	0.100266	0.095022
θ	0.0713392	0.073202
σ	0.094877	0.166265
ρ	−0.158667	−0.155823
κ	0.569760	2

Table 4.2 The calibrated Schöbel-Zhu model to FX USD-EUR option market on July 3, 2008. The maximal maturity is 4 years. For the case of fixed velocity, κ is fixed to 2.

In Table (4.2), we also give the calibrated parameters for the Schöbel-Zhu model based on the market data as of July 3, 2008. The different values of κ do not change the values of v_0, θ and ρ significantly, but affect the values of σ strongly, which are still in a reasonable range. This finding indicates that the strategy of fixing velocity parameter does not essentially destroy the quality of calibration.

4.7.3 Fixing Spot Volatility

As mentioned above, the spot volatility performs a strong calibration stability. Moreover, a further observation of Table (4.1) and Table (4.2) confirms that spot volatility (or the square root of spot variance) is very near to the short-term implied volatility, for example, 1 month implied volatility. This is a hint for a calibration strategy to fix the spot volatility with which we can potentially achieve a more stable and fast calibration.

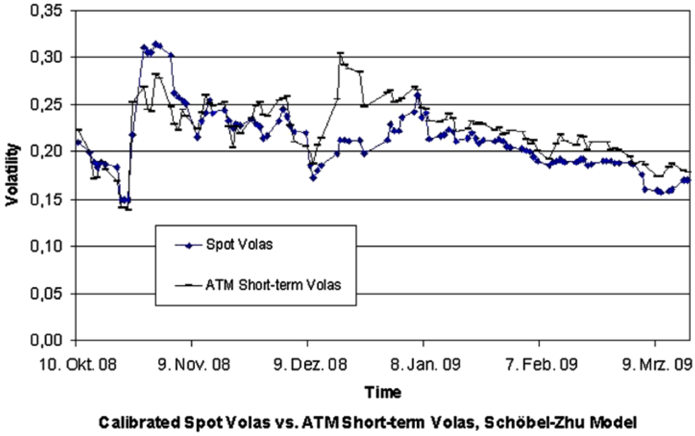

Fig. 4.1 Comparison of calibrated spot volatility v_0 and ATM short-term volatility in the Schöbel-Zhu model for EUR/USD FX volatility surface.

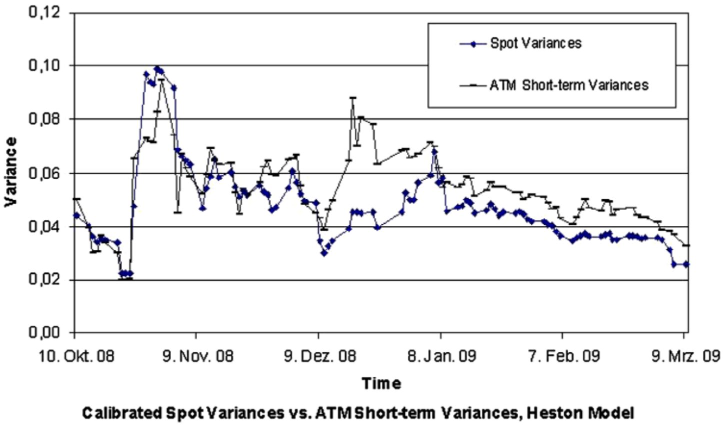

Fig. 4.2 Comparison of calibrated spot variance V_0 and ATM short-term variance in the Heston model for EUR/USD FX volatility surface.

By fixing spot volatility, the freedom of mean level is somehow restricted given by a market volatility surface, and the estimated mean level could take a more plausible value than in the case of unfixed spot volatility. For instance, we can set the value of spot volatility to be the average of some ATM short-term volatilities for a given smile surface. Figure (4.1) and Figure (4.2) give two pictures about how large the calibrated spot volatilities (variances) deviate from the corresponding ATM short-term volatilities (variances) in the Schöbel-Zhu model (the Heston model) between October 2008 and March 2009. We can observe that in most time the calibrated spot volatilities and variances are rather close to their ATM short-term counterparts except for the time period between the middle December 2008 and the middle January 2009. The test time period in Figure (4.1) and Figure (4.2) begins directly after the Lehman Brothers crisis and undergoes a strong turbulence in the FX markets, which is reflected in the erratic jumps of the calibrated volatilities and the ATM short-term volatilities.

4.8 Markovian Projection

Above we have discussed the calibration methods based on parameter optimization, and model parameters are optimized in terms of mean square price errors over all strikes and maturities. In recent literature, the so-called Markovian projection method (see Antonov and Misirpashaev (2006), Piterbarg (2006)) is applied to calibrate stochastic volatility model, particularly the Heston-like models. The basic idea of the Markovian projection method is to mimic the marginal distribution at certain time t of an Itô process with a simple and deterministic distribution by matching their first and second moments. The resulting new mimicking process is often Markovian, therefore, this mimicking approach is referred to as the Markovian projection. The Gyöngy theorem (1986) delivers the essential theoretical background of the Markovian projection.

Theorem 4.8.1. *Gyöngy theorem: Assume an Itô process $x(t)$ of the following form*

$$dx(t) = \alpha(x,t)dt + \beta(x,t)dW(t). \tag{4.33}$$

Define $a(t,x)$ and $b(t,x)$ by

$$a(t,x) = \boldsymbol{E}[\alpha(x,t)|\,x(t) = x]$$

and

$$b^2(t,x) = \boldsymbol{E}[\beta(x,t)|\,x(t) = x],$$

then the new process given by

$$dy(t) = a(t,y)dt + b(t,y)dW(t), \quad y_0 = x_0, \tag{4.34}$$

is a week solution to $x(t)$ so that both processes have the same one-dimensional distributions.

The new process of $y(t)$ is known as the local volatility process in financial literature, and $b(t,x)$ as local volatility or effective volatility. It is also known, as shown in Dupire (1997), that European-style option prices are uniquely determined by the terminal distribution at maturity T. Therefore, if the terminal distribution at time T of the stock price process $x(T)$ in a stochastic volatility model is accurately mimicked by another simple one $y(T)$, the European-style option prices with maturity T of stochastic volatility model is then also accurately approximated by the (Black-Scholes) prices using the new simple (Gaussian) distribution. The process $y(t)$ is then labeled as the Markovnian projection of the process $x(t)$.

In the following, we first consider the Heston model that is often used as a volatility-driving process, to illustrate the Markovian projection in literature. Recall the Heston model as follows:

$$\begin{aligned} dx(t) &= \big(r - \frac{1}{2}V(t)\big)dt + \sqrt{V(t)}dW_1(t), \\ dV(t) &= \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dW_2(t). \end{aligned}$$

These two SEDs admit obviously unique solution. Define $a(t,x)$ and $b(t,x)$ by

$$a(t,x) = \mathbf{E}[r - \frac{1}{2}V(t)|\, x(t) = x]$$

and

$$b^2(t,x) = \mathbf{E}[V(t)|\, x(t) = x],$$

the new process of the following form:

$$dy(t) = a(t,y)dt + b(t,y)dW_1(t), \quad y_0 = x_0, \tag{4.35}$$

has a consistent one-dimensional distribution at time t as $x(t)$ in the sense of first and second conditional moments. In this special case, since r is constant, the expected value of $\mathbf{E}[r - \frac{1}{2}V(t)|x(t) = x_0]$ can be simplified to

$$a(t,x) = r - \frac{1}{2}\mathbf{E}[V(t)|\, x(t) = x] = r - \frac{1}{2}b^2(t,x).$$

It becomes clear that mimicking the log return process $x(t)$ in the stochastic volatility model is essentially equivalent to calculating expected value of $V(t)$ conditional on $x(t)$.

Generally, computing the conditional expected first and second moments is not trivial, and in practice, raises a big numerical challenge, especially in more complicated stochastic volatility models, for example, local-stochastic volatility models where local volatility and stochastic volatility are combined. One method is to compute this conditional expected values in a Kolmogorov's forward PDE for $(x(t),V(t))$, and is numerically extensive. Piterbarg (2006) proposed various methods such as the Gaussian approximation, parameter averaging as well as local volatility approximation, to calculate the conditional expected first and second moments. Due to its analytical tractability, the Gaussian approximation is simple and may be easily implemented. We take again the Heston model for example. Since W_1

and W_2 are both Gaussian, it follows

$$\begin{aligned} b^2(t,x) &= \mathbf{E}[V(t)| \quad x(t)=x] \\ &= \mathbf{E}[V(t)] + \frac{\mathbf{Cov} < x(t),V(t) >}{\mathbf{Var}(x(t))}(x(t)-\mathbf{E}[x(t)]), \end{aligned}$$

therefore it is straightforward to obtain the expected value of $V(t)$ conditional on $x(t)$,

$$b^2(t,x) = A_1^2(t) + A_2(t)[x(t) - x_0 - rt + \frac{1}{2}\mathbf{E}[V(t)]] \tag{4.36}$$

with

$$\begin{aligned} A_1^2(t) &= \int_0^t \mathbf{E}[\kappa(\theta - V(s))]ds = \kappa\theta t - \kappa\int_0^t \mathbf{E}[V(s))]ds \\ &= \kappa\theta t - \kappa\int_0^t [\theta + (V(0)-\theta)e^{-\kappa s}]ds \\ &= (\theta - V(0))(1-e^{-\kappa t}) \end{aligned}$$

and

$$A_2(t) = \frac{\rho\sigma\int_0^t \mathbf{E}[\sqrt{V(s)}]A_1(s)ds}{\int_0^t A_1^2(s)ds}.$$

Another simple example of the Markovian projection is to mimic some skew models with the displaced diffusions. Since the displaced diffusion model admit an analytical pricing formula for European-style options, the Markovian projection with the displaced diffusion allows for an potential convenience in estimating parameters. Consider the following process

$$dx(t) = \beta(t)dW(t), \tag{4.37}$$

where $\beta(t)$ is some function to capturing volatility skew (not smile). The new mimicking process $y(t)$ is proposed to have displaced diffusion

$$y(t) = [1 + \Delta y(t)q(t)]\sigma(t)dW(t), \tag{4.38}$$

where $\Delta y(t) = y(t) - y(0)$ and $\sigma(t)$ is a deterministic function to fit possible term structure of volatility in $x(t)$. $q(t)$ controls the degree of skew. By the Gyöngy theorem, the matching condition is given by

$$\sigma(t)^2[1+\Delta y(t)q(t)]^2 \approx \mathbf{E}[\beta(t)^2|x(t)=x] = b^2(t,x).$$

Minimizing the L_2 distance between $\sigma(t)^2[1+\Delta y(t)q(t)]^2$ and $b^2(t,x)$ yields

$$\sigma(t)^2 = \mathbf{E}[\beta(t)^2] \tag{4.39}$$

and

$$q(t) = \frac{\mathbf{E}[\beta^2(t)\Delta y^2(t)]}{2\mathbf{E}[\beta(t)^2]\mathbf{E}[\Delta y^2(t)]}. \tag{4.40}$$

The quality of the Markovian projection via a displaced diffusion depends on the form of $\beta(t)$ and get worse for highly heterogeneous cases. Furthermore, it can not capture volatility smile. We will discuss an application of the Markovian projection to a more complicated local-stochastic model of Piterbarg (2003) in the context of Libor Market Model in Chapter 11.

The main drawback of the Markovian projection, however, is its limitation to a fixed time. In other words, with the Markovian projection method, we can only approximate the terminal distribution of the original process by other projected process, and estimate the model parameters of the projected model for a fixed maturity. As a result, it is not clear if the mimicking process with estimated parameters exhibits the same (or at least similar) dynamics at any time as the original process. The answer seems to be a no. As mentioned by Piterbarg (2006), "the prices of securities dependent on sample of $S(t)$ at multiple times, such as barriers, American options and so on, are different between the original model and the projected model." As discussed above, the goal of a calibration is not only to target the prices of European-style options, but also allows us to value exotic derivatives soundly with the calibrated model. The potential discrepancy between the original and projected processes is difficult to gauge, and therefore restricts the Markovian projection to wide applications in practice.

Chapter 5
Simulating Stochastic Volatility Models

In this chapter we address how to simulate stochastic volatility model. After a stochastic volatility model is calibrated to market smile surface, we can use the calibrated stochastic volatility parameters to simulate the prices of the underlying asset, and then to value exotic products.

The mean-reverting square root process is a very popular process and widely used to model stochastic interest rates, for instance, by Cox, Ingorsoll and Ross (1985), Duffie and Kan (1996), and to model stochastic variances by Heston (1993). Therefore, an efficient and accurate simulation of the mean-reverting square root process is of a great importance in finance. On the other hand, due to the non-trivial distribution and the negative values in the simulation of the mean-reverting square root process, its robust simulation remains the middle point of computational finance since long time. Hence we will pay more attentions on the numerical issues of the mean-reverting square root process. In particular, the negative simulated values and the related calculation of the square root of negative values are a big challenge for practitioners. In this chapter, some approaches for an efficient and robust simulation of the mean-reverting square root process are addressed. According to numerical experiments, the so-called QE scheme proposed by Andersen (2007) and the transformed volatility scheme proposed by Zhu (2008) perform efficiently and robustly. Both schemes can be easily implemented. Additionally, from the viewpoint of practical applications, the simulation of the maximum and minimum of asset prices is often necessary in valuing exotic options, for example, barrier options and lookback options. Basket options and correlation options are a class of actively traded exotic options. To value them, we need to set up a multi-asset model. In the last chapter, we consider a multi-asset model with stochastic volatilities, and propose an intuitive positive semi-definite correlation matrix for the multi-asset model. A corresponding simulation method for the multi-asset case with stochastic volatilities is also given.

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_5,

5.1 Simulation Scheme

5.1.1 Discretization

The simplest discretization scheme is an Euler scheme. Given a time partition $\{t_0 = 0, t_1, t_2, \cdots, t_{H-1}, t_H\}$, according to the Euler scheme, the Heston model and the Schöble-Zhu model can be discretized respectively as follows:

The Heston Model:

$$\begin{aligned} S(t_{h+1}) &= S(t_h)[1 + r(t_h)\Delta_h + \sqrt{V(t_h)\Delta_h} Z_1(t_h)], \\ V(t_{h+1}) &= V(t_h) + \kappa(\theta - V(t_h))\Delta_h + \sigma\sqrt{V(t_h)\Delta_h} Z_2(t_h). \end{aligned} \tag{5.1}$$

The Schöble-Zhu model:

$$\begin{aligned} S(t_{h+1}) &= S(t_h)[1 + r(t_h)\Delta_h + v(t_h)\sqrt{\Delta_h} Z_1(t_h)], \\ v(t_{h+1}) &= v(t_h) + \kappa(\theta - v(t_h))\Delta_h + \sigma\sqrt{\Delta_h} Z_2(t_h). \end{aligned} \tag{5.2}$$

Here Δ_h is the time step between t_{h+1} and t_h. $Z_1(t_h)$ and $Z_2(t_h)$ are two correlated standard normally distributed random variables. They may be drawn at time t_h according to the following methods,

$$\begin{aligned} Z_1(t_h) &= Ran_1, \quad Ran_1 \sim N(0,1), \\ Z_1(t_h) &= \rho Ran_1 + \sqrt{1-\rho^2} Ran_2, \quad Ran_2 \sim N(0,1). \end{aligned}$$

The random variables Ran_1 and Ran_2 are two independent standard normally distritbuted and may be created subsequently by a same random generator. Usually, the most efficient random generating algorithms produce only uniform random variables for interval $[0,1]$. Hence, we need to produce the Gaussian random variables from the uniform random variables. To this end, some efficient and quick algorithms are available including the Box-Mueller algorithm and the inverse-transform algorithm. For detailed treatments on this topic, see Glasserman (2004).

To avoid possible negative values of $S(t_h)$ in a simulation, we may simulate $X(t) = \ln S(t)$ instead of $S(t)$. The corresponding discretization of $X(t)$ is given by for the Heston model,

$$X(t_{h+1}) = X(t_h) + [r(t_h) - \frac{1}{2}V(t_h)]\Delta_h + \sqrt{V(t_h)\Delta_h} Z_1(t_h),$$

and for the Schöble-Zhu model,

$$X(t_{h+1}) = X(t_h) + [r(t_h) - \frac{1}{2}v^2(t_h)]\Delta_h + v(t_h)\sqrt{\Delta_h} Z_1(t_h).$$

The simulated value of $S(t_{h+1})$ is then equal to $\exp(X(t_{h+1})$.

If we simulate $S(t)$ directly, another discretization scheme called the Milstein scheme is available and improves the Euler's scheme by adding a term

$$\frac{1}{2}Diff(S(t),t)Diff^{'}(S(t(,t)[Z_1(t)^2-1]\Delta_t,$$

where $Diff$ and $Diff^{'}$ denote the diffusion term and its derivation to $S(t)$ of the process $S(t)$. The mean-reverting square-root process $V(t)$ has also a Milstein discretization. The corresponding discretization for the Heston model and the Schöble-Zhu model are given as follows:

The Heston Model:

$$\begin{aligned}
S(t_{h+1}) &= S(t_h)[1+r(t_h)\Delta_h+\sqrt{V(t_h)\Delta_h}Z_1(t_h)+\frac{1}{2}V(t_h)(Z_1^2(t_h)-1)\Delta_h],\\
V(t_{h+1}) &= V(t_h)+\kappa(\theta-V(t_h))\Delta_h+\sigma\sqrt{V(t_h)\Delta_h}Z_2(t_h)\\
&\quad+\frac{1}{4}\sigma^2(Z_2^2(t_h)-1)\Delta_h.
\end{aligned} \tag{5.3}$$

The Schöble-Zhu model:

$$\begin{aligned}
S(t_{h+1}) &= S(t_h)[1+r(t_h)\Delta_h+v(t_h)\sqrt{\Delta_h}Z_1(t_h)+\frac{1}{2}v^2(t_h)(Z_1^2(t_h)-1)\Delta_h],\\
v(t_{h+1}) &= v(t_h)+\kappa(\theta-v(t_h))\Delta_h+\sigma\sqrt{\Delta_h}Z_2(t_h).
\end{aligned} \tag{5.4}$$

However, the above defined Milstein scheme has a convergence of the same week order as the Euler scheme, namely a week order of one.

5.1.2 *Moment-Matching*

Note the stochastic volatility $v(t)$ follows a mean-reverting Ornstein-Uhlenbeck process and is Gaussian, we can utilize this property to arrive at a more efficient discretization. According to (3.29), conditional on $v(t_h)$, $v(t_{h+1})$ is distributed by a Gaussian law,

$$v(t_{h+1}) \sim N\big(m_1(t_h),m_2(t_h)\big) \tag{5.5}$$

with

$$m_1(t_h) = \theta+(v(t_h)-\theta)e^{-\kappa\Delta_h},$$

$$m_2(t_h) = \frac{\sigma^2}{2\kappa}(1-e^{-2\kappa\Delta_h}).$$

Therefore, $v(t)$ can be simulated as follows

$$v(t_{h+1}) = m_1(t_h)+\sqrt{m_2(t_h)}Z_2(t_h). \tag{5.6}$$

This simulation of $v(t)$ does not have a discretization error. The only discrepancy from the true process is due to the random variables $Z_2(t_h)$. Additionally, it allows for large time steps Δ_h without losing simulation quality. Hence, we need only a few number of time steps to simulate $v(t)$ and achieve remarkable efficiency.

5.2 Problems in the Heston Model

This section focuses on a special problem of the Heston model: how to achieve a robust and accurate simulation for a mean-reverting square root process. More precisely, the problem is that the possible negative samples for $V(t)$ and the related squared root $\sqrt{V(t)}$ break down the simulation. This is a problem that many financial engineers have tried to solve since more than a decade. Meanwhile, many approaches are proposed to deal with this problem with some successes. In the following, we first discuss the reasons why the negative values of $V(t)$ in a simulation occur. Then we present some simulation schemes which are especially tailored to do an accurate, efficient and robust simulation of a mean-reverting square root process. According to my own experiences, the scheme labeled to QE suggested by Andersen (2007) and the transformed volatility scheme suggested by Zhu (2008) are two of the most efficient simulation algorithms for the Heston model. Of course, these schemes can be applied to the CIR (1985) interest rate model without any restriction.

5.2.1 Negative Values in Paths

As discussed early, Feller's conditions of the Heston model for the positivity of variance $V(t)$ are the following restrictions of parameters,

$$2\kappa\theta > \sigma^2, \quad V_0 > 0.$$

An unconstrained calibration of the Heston model does not guarantee the above condition and leads often to a prior possible negative values in the simulation of $V(t)$. This makes the calculation of the square root $\sqrt{V(t_h)\Delta_h}$ impossible and breaks down the simulation. Therefore, a constrained calibration with an embedded parameter restriction is necessary, and however, sacrifices some fitting quality.

But even the required parameter restriction is satisfied in the Heston model, it is still possible to generate some negative values of $V(t_h)$ in paths in practical simulations. If the generated $Z_2(t_h)$ is negative enough, so that

$$\kappa(\theta - V(t_h))\Delta_h + \sigma\sqrt{V(t_h)\Delta_h}Z_2(t_h) < 0,$$

we obtain $V(t_{h+1}) < 0$. More mathematically, even when $V(t_n) > 0$ and the Feller condition is satisfied, the probability of $V(t_{h+1})$ becoming negative in a simple Euler scheme is given by

$$P(V(t_{h+1}) < 0) = N\Big(\frac{-(1-\kappa\Delta_n)V(t_n) - \kappa\theta\Delta_n}{\sigma\sqrt{V(t_n)\Delta_n}}\Big),$$

since the Euler scheme practically replaces the non-central chi-square distribution of $V(t)$ by a normal distribution. This finding is sometimes more frustrating than the parameter constraints. This simple and naive approach to avoiding the computation of $\sqrt{V(t_h)\Delta_h}$ is to truncate $V(t_h)$ to $V^+(t_h) = \max[V(t_h),0]$. The standard Euler scheme then becomes

$$\begin{aligned} S(t_{h+1}) &= S(t_h)[1 + r(t_h)\Delta_h + \sqrt{V^+(t_h)\Delta_h}Z_1(t_h)], \\ V(t_{h+1}) &= V(t_h) + \kappa(\theta - V(t_h))\Delta_h + \sigma\sqrt{V^+(t_h)\Delta_h}Z_2(t_h), \end{aligned} \tag{5.7}$$

which is called the (partial) truncation scheme. Numerical examples show that the truncation scheme is not robust and generates some largely biased paths. Therefore, the truncation scheme is not applicable for pricing options.

Approximating a mean-reverting squared-root process with a process or a distribution without the term $\sqrt{V(t_h)}$ seems to be the only way to overcome this problem encountered in the Heston model.

5.2.2 Log-normal Scheme

The idea of the log-normal scheme is straightforward and approximates the conditional non-central chi-square distribution of $V(t)$ by a log-normal distribution, and the resulting paths are then always positive. The approximation is performed by matching first two moments of non-central chi-square distribution and log-normal distribution. Let $y(t) = \ln Y(t)$ follow a Brownian motion with drift,

$$dy(t) = \mu dt + \nu dW(t).$$

$Y(t)$ is then log-normally distributed with the following mean and variance respectively,

$$E_Y(s) = \mathbf{E}[Y(t)|Y(s)] = Y(s)e^{(\mu+\frac{1}{2}\nu^2)(t-s)},$$

$$\begin{aligned} Var_Y(s) &= \mathbf{Var}[Y(t)|Y(s)] \\ &= Y^2(s)\big(e^{\nu^2(t-s)} - 1\big)e^{(2\mu+\nu^2)(t-s)} = \big(e^{\nu^2(t-s)} - 1\big)E_Y^2. \end{aligned}$$

Setting E_Y and Var_Y equal to the true conditional mean and variance of $V(t)$ yields

$$V(s)e^{(\mu+\frac{1}{2}\nu^2)(t-s)} = E_V(s) = \theta + (V(s) - \theta)e^{-\kappa(t-s)},$$

$$\left(e^{\nu^2(t-s)}-1\right)E_V^2(s) = Var_V(s) = \frac{\sigma^2\theta}{2\kappa}\left[1-e^{\kappa(t-s)}\right]^2 + V(s)\frac{\sigma^2}{\kappa}e^{-\kappa(t-s)}\left[1-e^{\kappa(t-s)}\right],$$

which can be rearranged to

$$E_V(s) = V(s)e^{(\mu+\frac{1}{2}\nu^2)(t-s)},$$

$$Var_V(s) = \left(e^{\nu^2(t-s)}-1\right)E_V^2.$$

We can solve μ and ν from the above equations and obtain

$$\nu^2(s) = \frac{1}{t-s}\ln\left(\frac{Var_V(s)}{E_V^2(s)}+1\right), \tag{5.8}$$

$$\mu(s) = \frac{1}{t-s}\ln\frac{E_V(s)}{V(s)} - \frac{1}{2}\nu^2. \tag{5.9}$$

Based on the approximated log-normal distribution, a new discretization scheme is given by

$$V(t_{h+1}) = V(t_h)\exp\left(\mu(t_h)\Delta_h + \nu(t_h)\sqrt{\Delta_h}Z_2(t_h)\right). \tag{5.10}$$

5.2.3 Transformed Volatility Scheme

As indicated above, the negative values in paths are caused by two reasons. First one is the unfortunate parameter restriction that allows a negative variance $V(t_h)$. Second one lies in the nature of the simulation that can not exclude the negative variance $V(t_h)$. Both problems break down the calculation of $\sqrt{V(t_h)}$. If the term $\sqrt{V(t_h)}$ vanishes, we do not need to care the negative values of $V(t_h)$ any more. For this reason, it is natural to consider $v(t) = \sqrt{V(t_h)}$. As mentioned in Chapter 3, we can derive the process $v(t)$ by Itô's lemma, and recall it,

$$dv(t) = \frac{1}{2}\kappa\left[(\theta - \frac{\sigma^2}{4\kappa})v^{-1}(t) - v(t)\right]dt + \frac{1}{2}\sigma dW_2(t).$$

It is straightforward to discretize $v(t)$ and obtain a total new Euler scheme for the Heston model based on $v(t)$,

$$\begin{aligned} X(t_{h+1}) &= X(t_h) + [r(t_h) - \frac{1}{2}v^2(t_h)]\Delta_h + v(t_h)\sqrt{\Delta_h}Z_1(t_h), \\ v(t_{h+1}) &= v(t_h) + \frac{1}{2}\kappa\left[(\theta - \frac{\sigma^2}{4\kappa})v^{-1}(t_h) - v(t_h)\right]\Delta_h + \frac{1}{2}\sigma\sqrt{\Delta_h}Z_2(t_h). \end{aligned} \tag{5.11}$$

The above discretization is very similar to the discretization of the Schöbel-Zhu model.

The advantages of this scheme based on the transformed volatility are at least twofold. Firstly, the transformed volatility process $v(t)$ coincides completely with the original Heston model and at the same time avoids any calculation of squared root. The above transform is theoretically exact and uses no approximation. The only error of the simulation comes from the discretization and the generated random numbers. Secondly, any parameter constellation is allowed in the process of $v(t)$. Negative values of $v(t)$ lead still positive variance $V(t)$, and can be similarly interpreted as in the Schöbel-Zhu model.

The single drawback of the transformed volatility scheme is the mean level $\theta_v = (\theta - \frac{\sigma^2}{4\kappa})v^{-1}(t)$ has a term $v^{-1}(t)$ which makes the mean level stochastic over time. The naive Euler scheme can not capture the erratic behavior of $v^{-1}(t)$ in time interval $[t, t+\Delta]$.[1] This erratic behavior of $v^{-1}(t)$ will be amplified by the negative value of θ_v for the parameter constellation $4\kappa\theta > \sigma^2$. In this case, θ_v often jumps from a large positive value to a large negative value in a short time, and vice versa. The value of $\theta_v(t_h)$ evaluated with the start value $v(t_h)$ in time interval $[t, t+\Delta]$ is no longer appropriate for the entire time interval. In order to achieve accurate simulation, we suggest a robust approximation for θ_v by moment-matching.

We rewrite the transformed volatility process with a mean level θ_v as follows:

$$dv(t) = \kappa_v(\theta_v - v(t))dt + \sigma_v dW_2(t) \tag{5.12}$$

with

$$\kappa_v = \frac{1}{2}\kappa, \quad \sigma_v = \frac{1}{2}\sigma.$$

Note $\mathbf{E}[v(t)^2] = \mathbf{E}[V(t)]$, we have

$$\mathbf{E}[V(t+\Delta)] = \mathbf{Var}[v(t+\Delta)] + \mathbf{E}[v(t+\Delta)]^2, \tag{5.13}$$

where

$$\mathbf{E}[V(t+\Delta)] = \theta + [V(t) - \theta]e^{-\kappa\Delta},$$

$$\mathbf{Var}[v(t+\Delta)] = \frac{\sigma_v^2}{2\kappa_v}(1 - e^{-2\kappa_v\Delta}) = \frac{\sigma^2}{4\kappa}(1 - e^{-\kappa\Delta}),$$

$$\mathbf{E}[v(t+\Delta)] = \theta_v + [v(t) - \theta_v]e^{-\kappa_v\Delta} = \theta_v + [v(t) - \theta_v]e^{-\frac{1}{2}\kappa\Delta}.$$

The single unknown variable is θ_v that can be solved as follows:

$$\theta_v^* = \frac{\beta - v(t)e^{-\frac{1}{2}\kappa\Delta}}{1 - e^{-\frac{1}{2}\kappa\Delta}} \tag{5.14}$$

with

[1] By Itô's lemma, we can derive the stochastic process of v^{-1} which is involved in a term $v^{-3}(t)$. This indicates that the dynamics of $v^{-1}(t)$ is very bad-behaved.

$$\beta = \sqrt{[\mathbf{E}[V(t+\Delta)] - \mathbf{Var}[v(t+\Delta)]]^{+}},$$

where β is set to zero if $\mathbf{E}[V(t+\Delta)] < \mathbf{Var}[v(t+\Delta)]$.

The matched θ_v^* can be applied to any parameter constellation and produces good simulations even for the case of $2\kappa\theta < \sigma^2$. The moment-matching of θ_v^* improves the accuracy of simulations significantly. The transformed volatility scheme with θ_v^* is then given by

$$\begin{aligned} X(t_{h+1}) &= X(t_h) + [r(t_h) - \frac{1}{2}v^2(t_h)]\Delta_h + v(t_h)\sqrt{\Delta_h}Z_1(t_h), \\ v(t_{h+1}) &= v(t_h) + \frac{1}{2}\kappa\left[\theta_v^*(t_h) - v(t_h)\right]\Delta_h + \frac{1}{2}\sigma\sqrt{\Delta_h}Z_2(t_h). \end{aligned} \tag{5.15}$$

Some advanced simulation techniques, for example, martingale correction, may be coupled with the above scheme for potential improvements.[2]

5.2.4 QE Scheme

Andersen (2007) proposed a two-segment approximation for non-central chi-square distribution, one with a quadratic function, one with exponential function. Therefore this moment-matching scheme is referred to as the QE (quadratic-exponential) scheme. In more details, Andersen represented the segment of the large value of $V(t)$ with a quadratic function of a standard Gaussian variable, and described the segment of the small values of $V(t)$ with a distribution of exponential form.

Firstly, let us consider the segment of the large values of $V(t), V(t) > g$, where g is defined rather arbitrarily. To preserve the non-negative values of $V(t)$, it is reasonable to approximate $V(t)$ with a quadratic function of a random variable. Particularly, Andersen suggested the following form

$$V(t+1) \approx V^*(t+1) = a(b+Z)^2, \tag{5.16}$$

where Z is a standard Gaussian variable. It is easy to verify that $V(t+1)$ is approximated by a non-central chi-square distribution with one degree of freedom and non-centrality parameter b^2, multiplied with a,

$$V^*(t+1) \sim F_{\chi^2}(x/a; 1, b^2).$$

Therefore it follows immediately

$$E_V^* = a(1+b^2),$$

[2] There is no Milstein scheme for $v(t)$ since the diffusion term is constant. Generally, a second-order discretization method does not improve the simulation of Ornstein-Uhlenbeck process.

$$Var_V^* = 2a^2(1+2b^2).$$

By setting the above mean and variance to the correspond mean and variance of $V(t+1)$, we obtain

$$E_V(t) = a(1+b^2),$$
$$Var_V(t) = 2a^2(1+2b^2).$$

These two equations imply a quadratic equation for b or a. Denote $\gamma = Var_V^2(t)/E_V^2(t)$ and assume $\gamma \leqslant 2$, we can solve a and b,

$$a = \frac{E_V(t)}{1+b^2}, \tag{5.17}$$

$$b = 2\gamma^{-1} - 1 + \sqrt{2\gamma^{-1}}\sqrt{2\gamma^{-1} - 1}. \tag{5.18}$$

The next step is to consider the segment of the small values of $V(t), V(t) \leqslant g$. Andersen realized the density of $V(t)$ will be very large around zero for small $V(t)$. By imitating the density function of a non-central chi-square distribution, he suggested the following approximated density function for $V(t)$,

$$p(V(t) = x) = p\delta(0) + q(1-p)e^{-px},$$

where $\delta(\cdot)$ is a Dirac delta-function. Both p and q are non-negative constants. For $0 \leqslant p \leqslant 1$ and $q \geqslant 0$, the above density function is well-defined. By integrating the density function, we obtain a cumulative distribution function for small $V(t)$,

$$F(x) = P(V(t) < x) = p + (1-p)(1-e^{-qx}), \quad x \geqslant 0. \tag{5.19}$$

According the above approximated density function and distribution function, it is easy to calculate the mean and variance which take the following form,

$$E_V^*(t) = \frac{1-p}{q}, \tag{5.20}$$

$$Var_V^*(t) = \frac{1-p^2}{q^2}. \tag{5.21}$$

Moment-matching conditions lead to the equation system

$$E_V(t) = \frac{1-p}{q}, \quad Var_V(t) = \frac{1-p^2}{q^2}.$$

Solving this equation system yields

$$p = \frac{\gamma-1}{\gamma+1}, \tag{5.22}$$

$$q = \frac{1-p}{E_V(t)}. \tag{5.23}$$

To make the parameters p and q sense, and also make the proposed density function well-defined, we must require $\gamma \geqslant 1$.

Differently from other schemes, the QE scheme does not demand a discretization of the process $dV(t)$, but delivers an approximated distribution for $V(t)$ based on two segments. Therefore, we have the following algorithm to draw the random variables $V(t)$:

1. Given $V(t_h)$, compute $E_V(t_h)$ and $Var_V(t_h)$ for $V(t_{h+1})$, and then γ.
2. Draw a uniform random number U.
3. If $\gamma \leqslant g$:
   (a). Compute a and b.
   (b). Compute $Z = N^{-1}(U)$ to obtain a Gaussian random number.
   (c). Compute $V(t_{h+1}) = a(b+Z)^2$.
4. If $\gamma > g$:
   (a). Compute p and q.
   (b). Compute $V(t_{h+1}) = F^{-1}(U; p, q)$ for small $V(t_{h+1})$.

In the above algorithms we used the inverse transform method to generate the desired random numbers from the uniform random number U. The inverse function of $F(x; p, q)$ reads

$$F(u; p, q) = \begin{cases} 0, & 0 \leqslant u \leqslant p, \\ q^{-1} \ln\left(\frac{1-p}{1-u}\right), & p < u \leqslant 1. \end{cases}$$

For practical applications, $g = 1.5$ should be a good choice.

5.2.5 *The Broadie-Kaya Scheme*

Broadie and Kaya (2006) proposed a simulation scheme which is unbiased by construction. However, their unbiased scheme is achieved at a price of large complexity and slow speed. As mentioned in Section 3.2, $V(t+\Delta)$ can be expressed by the stochastic integral of the Heston SDE,

$$V(t_{h+1}) - V(t_h) = \kappa\theta\Delta_h - \kappa \int_{t_h}^{t_{h+1}} V(u)du + \sigma \int_{t_h}^{t_{h+1}} \sqrt{V(u)}dZ_2(u),$$

or equivalently,

$$\int_{t_h}^{t_{h+1}} \sqrt{V(u)}dZ_2(u) = \frac{1}{\sigma}\left[V(t_{h+1}) - V(t_h) - \kappa\theta\Delta + \kappa \int_{t_h}^{t_{h+1}} V(u)du\right]. \quad (5.24)$$

On the other hand, the log return $X(t)$ of $S(t)$ can be rewritten via Z_2 and another independent Brownian motion Z by the Cholesky decomposition,

$$X(t_{h+1}) = X(t_h) + r(t_h)\Delta_h - \frac{1}{2}\int_{t_h}^{t_{h+1}} V(u)du$$
$$+ \rho\sqrt{V(t_h)\Delta_h}Z_2(t_h) + \sqrt{1-\rho^2}\sqrt{V(t_h)\Delta_h}Z(t_h).$$

Inserting the expression of $\int_{t_h}^{t_{h+1}} \sqrt{V(u)}dZ_2(u)$ into the above discretization equation yields the Broadie-Kaya Scheme,

$$X(t_{h+1}) = X(t_h) + r(t_h)\Delta_h + \sqrt{1-\rho^2}\sqrt{V(t_h)\Delta_h}Z(t_h)$$
$$+ \frac{\rho}{\sigma}[V(t_{h+1}) - V(t_h) - \kappa\theta\Delta_h] + \left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right)\int_{t_h}^{t_{h+1}} V(u)du. \tag{5.25}$$

The readers familiar with the expectation approach to deriving the characteristic functions in the Heston model in Section 3.2 will be also familiar with the technique used here to arrive at the final expression of $X(t_{h+1})$ conditional on $V(t_{h+1})$ and $\int_{t_h}^{t_{h+1}} V(u)du$. However, the discretization (5.25) is not ready for simulation yet since $V(t_{h+1})$ and $\int_{t_h}^{t_{h+1}} V(u)du$ are not known. Broadie and Kaya suggested the following steps to complete the simulation of $X(t)$:

1. Simulate $V(t_{h+1})$ according to the known non-central chi-square distribution. Particularly Broadie and Kaya used the acceptance-rejection technique to generate a gamma distribution, and therefore also the non-central chi-square distributed $V(t_{h+1})$ (see Glasserman (2004)).
2. Simulate $\int_{t_h}^{t_{h+1}} V(u)du$ conditional on $V(t_{h+1})$ and $V(t_h)$.
3. Simulate $X(t_{h+1})$ conditional on $V(t_{h+1})$ and $\int_{t_h}^{t_{h+1}} V(u)du$.

While the steps 1 and 3 are somehow straightforward, the step 2 is numerically much more involved since the distribution of $\int_{t_h}^{t_{h+1}} V(u)du$ is not known analytically. To obtain the distribution of $\int_{t_h}^{t_{h+1}} V(u)du$ conditional on $V(t_{h+1})$ and $V(t_h)$, Broadie and Kaya calculated the CF of $\int_{t_h}^{t_{h+1}} V(u)du$ as follows:

$$f(\phi) = \mathbf{E}\left[\exp\left(i\phi\int_{t_h}^{t_{h+1}} V(u)du\right)|V(t_h), V(t_{h+1})\right]$$

whose solution is analytically available and contains two modified Bessel functions of the first kind. Therefore, we can numerically compute the corresponding sample of $\int_{t_h}^{t_{h+1}} V(u)du$ by using the inversion of the conditional distribution function. The total procedure to simulate $\int_{t_h}^{t_{h+1}} V(u)du$ requires extensive numerical computations and hence is time-consuming. As reported by Lord, Koekkoek and van Dijk (2008), Broadie and Kaya's Scheme is several times slower than other simulation schemes. Due to this obvious drawback, the Broadie-Kaya Scheme is unsuitable for practical applications.

5.2.6 Some Other Schemes

In the past decade there are a number of financial literature discussing efficient and robust, or shortly better, simulation schemes of the Heston-like models. The Lognormal scheme, the transformed volatility scheme and the QE scheme as well as the Broadie-Kaya Scheme are some representative methods with which we can achieve for a better simulation. The Kahl-Jäckel scheme supposed Kahl and Jäckel (2006) and the full truncation scheme supposed by Lord, Koekkoek and van Dijk (2008), are of particular interest in practical and theoretical interest.

Kahl-Jäckel's scheme is a combination of an implicit Milstein scheme of the variance process $V(t)$ and a central discretization of underlying process $X(t)$. Recall a simple Milstein scheme for the variance process $V(t)$,

$$V(t_{h+1}) = V(t_h) + \kappa(\theta - V(t_h))\Delta_h + \sigma\sqrt{V(t_h)\Delta_h}Z_2(t_h) + \frac{1}{4}\sigma^2(Z_2^2(t_h) - 1)\Delta_h.$$

The implicit Milstein scheme replaces the term $\kappa(\theta - V(t_h))\Delta_h$ with $\kappa(\theta - V(t_{h+1}))\Delta_h$. Hence, we can express $V(t_{h+1})$ as follows:

$$V(t_{h+1}) = \frac{1}{1+\kappa\Delta_h}[V(t_h) + \kappa\theta\Delta_h + \sigma\sqrt{V(t_h)\Delta_h}Z_2(t_h) + \frac{1}{4}\sigma^2(Z_2^2(t_h) - 1)\Delta_h]. \tag{5.26}$$

It can be proven that $V(t_{h+1})$ will be positive if $V(t_h) > 0$ and $4\kappa\theta > \sigma^2$, which relaxes the original positivity condition. However, if $4\kappa\theta < \sigma^2$ which is often the case in practice, the truncation is still applied to $V(t_{h+1})$ to ensure positive paths, namely, we have to replace $V(t_{h+1})$ by $V^+(t_{h+1})$. On the side of $X(t)$, Kahl and Jäckel replaced $V(t_h)$ and $\sqrt{V(t_h)}$ by two ad hoc central discretizations,

$$\begin{aligned} V(t_h) &\approx \frac{1}{2}[V(t_{h+1}) + V(t_h)], \\ \sqrt{V(t_h)} &\approx \frac{1}{2}[\sqrt{V(t_{h+1})} + \sqrt{V(t_h)}] \end{aligned}$$

and finally arrived at the following scheme for $X(t)$,

$$\begin{aligned} X(t_{h+1}) = X(t_h) &+ r(t_h)\Delta_h - \frac{\Delta_h}{4}[V(t_{h+1}) + V(t_h)] + \rho\sqrt{V(t_h)\Delta_h}Z_2(t_h) \\ &+ \frac{1}{2}[\sqrt{V(t_{h+1})} + \sqrt{V(t_h)}](Z_1(t_h) - \rho Z_2(t_h))\Delta_h \\ &+ \frac{1}{4}\rho\sigma(Z_2^2(t_h) - 1)\Delta_h. \end{aligned} \tag{5.27}$$

Recent comparison studies by Andersen (2007) as well as Lord, Koekkoek and van Dijk (2008) can not confirm a good performance of the Kahl-Jäckel scheme. As reported by Andersen (2007), the main reason for the large biases of the Kahl-Jäckel scheme could be due to the special discretization of $X(t)$.

Another interesting and simple scheme is the so-call full truncation scheme suggested by Lord, Koekkoek and van Dijk (2008). Contradict to partial truncation scheme given in (5.7), full truncation scheme replaces $V(t_h)$ in drift term with the truncated variance $V^+(t_h)$, and takes the following form,

$$V(t_{h+1}) = \kappa(\theta - V^+(t_h))\Delta_h + \sigma\sqrt{V^+(t_h)\Delta_h}Z_2(t_h). \tag{5.28}$$

As demonstrated by ATM calls as example in Lord, Koekkoek and van Dijk, the full truncation outperforms the partial truncation scheme and even the log-normal scheme significantly. Unfortunately, this result is not tested with deep ITM and deep OTM options with which we can gauge better the ability of a certain scheme for simulating the tail distribution of the Heston model reliably.

5.3 Simulation Examples

To demonstrate the quality of some simulation schemes, we compare the simulated prices of European-style call options with the analytic prices of the Heston model. Particularly, we simulate call options with three various schemes: the log-normal scheme, the transformed volatility (TV) scheme, and Andersen's QE scheme. As shown in Andersen (2007), QE scheme outperforms other existing schemes and could be considered as a benchmark method for a mean-reverting square root process. The log-normal scheme is also widely used in practical applications and could be competitive to other schemes. For a systematic comparison, we consider three cases.

1. Case 1: $\kappa \geqslant \sigma^2/(2\theta)$.
2. Case 2: $\sigma^2/(2\theta) > \kappa \geqslant \sigma^2/(4\theta)$.
3. Case 3: $\kappa < \sigma^2/(4\theta)$.

The first case with $\kappa \geqslant \sigma^2/(2\theta)$, reported in Table (5.2), is equivalent to the condition for positive values $V(t)$. Therefore, the square root process of $V(t)$ behaves soundly in this case, and the simulations should usually not encounter any problem. In the second case reported in Table (5.3), the parameter restriction for positive values of $V(t)$ is no longer satisfied, and this case becomes more challenging for the most existing schemes. However, this case does not raise serious issues for the transformed volatility process $v(t)$ in (5.15) because the mean level θ_v is almost positive. The most challenging case is the third case where κ is smaller than $\sigma^2/(4\theta)$ and is far away from $\sigma^2/(2\theta)$. In this case, the most probability masses of $V(t)$ concentrate on the near of zero. This case is then a stress test for an efficient simulation scheme.

Table (5.1) gives the data in three test settings. We simulate European call prices with a spot price of 100, a maturity of 6 years, and strikes ranging from 70 to 130. All simulations are run with a number of paths 20000. This is a moderate number from the point of theoretical point, but is more compatible for practical applications.

The number of time steps per year is 32 and therefore is 192 for a maturity of 6 years. The parameters for the mean-reverting square root process are representative for equity options markets. Table (5.2), Table (5.3) and Table (5.4) give the numerical results using three different simulation schemes, as well as the corresponding analytical prices. For a detailed comparison, we provide also the standard deviations for the simulated prices and the relative price differences which are defined by

$$RPD = \frac{P_{Simulation} - P_{Analytic}}{P_{Analytic}}.$$

Relative price difference is preferred to absolute price difference in the comparison since the former eliminates the effect of underlying spot price and strike on option price.

From these three test cases, we observe the following points:

1. The numerical results of TV and QE schemes are very close to each other through all strikes and scenarios, not only in terms of prices, but also in terms of standard deviations. This delivers an strong evidence that both schemes work very well for all parameter constellations, even for the critical test case 3.
2. The log-normal scheme is competitive to TV and QE schemes in case 1 where the square root process is good–behaved, it delivers also acceptable prices for ATM options, but produces strongly biased prices for ITM and OTM options in cases 2 and 3. The log-normal scheme fails to pass the test case 3. Both TV and QE schemes outperform log-normal scheme clearly.
3. TV and QE produce more accurate prices for ITM options than OTM options in terms of the relative price differences.

The above numerical results and findings deliver the strong evidences that the transformed volatility scheme and the QE scheme can produce highly accurate simulations for the Heston model for any parameter constellation.

Call Option:	S_0 = 100	T = 6Y	r = 0.04	
Simulation:	Paths = 20000 TS = 32(per year)			
Process:	V_0 = 0.0225	θ = 0.04	σ = 0.3	ρ = -0.5
Case 1:	κ = 2			
Case 2:	κ = 0.8			
Case 3:	κ = 0.4			

Table 5.1 Test data with three different κ for the simulation examples in tables 5.2, 5.3 and 5.4.

5.4 Maximum and Minimum

Many exotic options involve the maximum or the minimum of underlying asset up to maturity or between two dates. For example, a down-and-out option is dependent

Strikes	TV	QE	LogN	Analytic
K=70	47.2726	47.2696	46.9707	47.1518
SDev	0.3162	0.3164	0.3388	
RPD	0.0025	0.0025	-0.0038	
K=80	40.9035	40.9014	40.6333	40.8003
SDev	0.3056	0.3059	0.3287	
RPD	0.0025	0.0025	-0.0041	
K=90	35.0888	35.0875	34.9108	34.9894
SDev	0.2926	0.2929	0.3162	
RPD	0.0028	0.0028	-0.0023	
K=100	29.8402	29.8408	29.8015	29.7543
SDev	0.2778	0.2781	0.3019	
RPD	0.0029	0.0029	0.0016	
K=110	25.1831	25.1850	25.3062	25.1049
SDev	0.2615	0.2618	0.2695	
RPD	0.0031	0.0032	0.0080	
K=120	21.1083	21.1134	21.4188	21.0302
SDev	0.2442	0.2446	0.2695	
RPD	0.0037	0.0039	0.0185	
K=130	17.5730	17.5824	18.0684	17.5020
SDev	0.2265	0.2268	0.2525	
RPD	0.0040	0.0040	0.0320	

Table 5.2 Test Case 1: $\kappa = 2 \geqslant \sigma^2/(2\theta)$. SDev stands for the standard deviations, RPD for the relative price difference.

Strikes	TV	QE	LogN	Analytic
K=70	47.4219	47.4210	46.8853	47.2812
SDev	0.2834	0.2835	0.3229	
RPD	0.0029	0.0030	-0.0084	
K=80	40.8673	40.8679	40.4694	40.7576
SDev	0.2730	0.2731	0.3130	
RPD	0.0026	0.0027	-0.0070	
K=90	34.7920	34.7933	34.6501	34.6872
SDev	0.2604	0.2605	0.3006	
RPD	0.0030	0.0031	-0.0011	
K=100	29.2214	29.2244	29.4391	29.1296
SDev	0.2459	0.2461	0.2861	
RPD	0.0032	0.0032	0.0106	
K=110	24.2162	24.2208	24.4391	24.1311
SDev	0.2300	0.2304	0.2703	
RPD	0.0035	0.0037	0.0298	
K=120	19.8193	19.8238	20.8727	19.7210
SDev	0.2128	0.2129	0.2533	
RPD	0.0049	0.0052	0.0584	
K=130	16.0099	16.0158	17.4480	15.9076
SDev	0.1951	0.1952	0.2359	
RPD	0.0064	0.0068	0.0968	

Table 5.3 Test Case 2: $\sigma^2/(2\theta) \leqslant \kappa = 0.8 < \sigma^2/(2\theta)$. SDev stands for the standard deviations, RPD for the relative price difference.

Strikes	TV	QE	LogN	Analytic
K=70	47.3616	47.3750	46.7192	47.2115
SDev	0.2559	0.2529	0.3066	
RPD	0.0032	0.0035	-0.0104	
K=80	40.6052	40.6039	40.1897	40.4726
SDev	0.2463	0.3433	0.2939	
RPD	0.0032	0.0032	-0.0070	
K=90	34.2329	34.2112	34.2369	34.0975
SDev	0.2347	0.2317	0.2816	
RPD	0.0039	0.0033	0.0041	
K=100	28.3237	28.2745	28.8903	28.1628
SDev	0.2214	0.2184	0.2672	
RPD	0.0057	0.0040	0.0258	
K=110	22.9460	22.8548	24.1725	22.7535
SDev	0.2065	0.2036	0.2512	
RPD	0.0134	0.0045	0.0623	
K=120	18.1958	18.0601	20.0733	17.9555
SDev	0.1905	0.1877	0.2341	
RPD	0.0133	0.0058	0.1179	
K=130	14.1228	13.9508	16.5605	13.8427
SDev	0.1738	0.1711	0.2164	
RPD	0.0202	0.0078	0.1963	

Table 5.4 Test Case 3: $\kappa = 0.4 < \sigma^2/(4\theta)$. SDev stands for the standard deviations, RPD for the relative price difference.

on the minimum of the prices over the entire maturity whereas a up-and-in option is associated with the maximum of the prices over the entire maturity. Simulating the maximum and the minimum of an asset in a given time period is very usual in financial engineering. If we simulate a stochastic process $X(t), 0 \leqslant t \leqslant T_M$, and generate a path $(X_0, X_1, \cdots, X_M)$, a rough estimation of simulated maximum and minimum of $X(t)$ should be

$$\hat{X}_{max} = \max[X_0, X_1, \cdots, X_M]$$

and

$$\hat{X}_{min} = \min[X_0, X_1, \cdots, X_M].$$

A simple and straightforward strategy is to increase number of time steps M to achieve the true extremum. Unfortunately, this strategy often fails due to two reasons. Firstly, in an usual Euler scheme, the simulated maximum and minimum converge very slowly, at a half speed of the convergency of $X(t)$ itself. Secondly, in order to get a satisfactory maximum and minimum, we have to increase the number of time steps and simulation paths. So this simple strategy becomes very resource-expensive and time-consuming.

On the other hand, given the simulated X_h and X_{h+1}, it is known that the maximum and the minimum of $X(t)$ in the time period $[t_h, t_{h+1}]$ follow a distribution called a Rayleigh distribution. For simplicity, let $W(t)$ be a standard Wiener process with $W(0) = 0$, given $W(1) = a$, the maximum y of $W(t)$ in time interval $[0, 1]$

follows a Rayleigh distribution as follows:

$$F_{max}(y) = 1 - e^{-2y(y-a)}. \tag{5.29}$$

In the same manner, the maximum z of $W(t)$ in time interval $[0,1]$ follows another Rayleigh distribution,

$$F_{min}(z) = 1 - e^{-2y(y+a)}. \tag{5.30}$$

Therefore, the extremum of any Brownian bridge with a fixed start value and a fixed end value is distributed according to the Rayleigh law. By the inverse transform method, we can simulate the Rayleigh distribution. Denote U as an uniform random variable, the simulated maximum is

$$Y = F_{max}^{-1}(U,a) = \frac{1}{2}a + \frac{\sqrt{a^2 - 2\ln U}}{2}. \tag{5.31}$$

The simulated minimum Z is given correspondingly by

$$Z = F_{min}^{-1}(U,a) = \frac{1}{2}a - \frac{\sqrt{a^2 - 2\ln U}}{2}. \tag{5.32}$$

Note that the above results are only applicable for a standard Brownian motion, or more precisely, a Brownian bridge, we need some modifications for a Brownian motion with drift. Assume that the local volatility of $X(t)$ is a constant σ_h, the simulated maximum $\hat{X}_{max}$ and minimum $\hat{X}_{min}$ are given by

$$\hat{X}^h_{max} = \frac{1}{2}(X_{h+1} - X_h) + \frac{\sqrt{(X_{h+1} - X_h)^2 - 2\sigma_h^2 \Delta t \ln U}}{2} \tag{5.33}$$

and

$$\hat{X}^h_{min} = \frac{1}{2}(X_{h+1} - X_h) - \frac{\sqrt{(X_{h+1} - X_h)^2 - 2\sigma_h^2 \Delta t \ln U}}{2}. \tag{5.34}$$

If we simulate $S(t)$ directly, the simulated maximum $\hat{S}_{max}$ and minimum $\hat{S}_{min}$ are given by

$$\hat{S}^h_{max} = S_h \exp\big[\frac{1}{2}\ln(S_{h+1}/S_h) + \frac{\sqrt{\ln(S_{h+1}/S_h)^2 - 2\sigma_h^2 \Delta t \ln U}}{2}\big] \tag{5.35}$$

and

$$\hat{S}^h_{min} = S_h \exp\big[\frac{1}{2}\ln(S_{h+1}/S_h) - \frac{\sqrt{\ln(S_{h+1}/S_h)^2 - 2\sigma_h^2 \Delta t \ln U}}{2}\big]. \tag{5.36}$$

The global extremum is then the extremum of all local extremes. More about the simulation of the maximum and the minimum of an asset see Beaglehoe, Dybvig and Zhou (1997), and Glassermann (2004).

5.5 Multi-Asset Model

So far we have only discussed the case of a single asset. For basket options and multi-asset options, we need a multi-asset model which is equipped with some mean-reverting Ornstein-Uhlenbeck processes

$$\begin{aligned} \frac{dS_j(t)}{S_j(t)} &= \mu_j dt + v_j(t) dW_j(t), \\ dv_j(t) &= \kappa_j(\theta_j - v_j(t))dt + \sigma_j dW_j^*(t), \end{aligned} \tag{5.37}$$

or with some mean-reverting squared root processes

$$\begin{aligned} \frac{dS_j(t)}{S_j(t)} &= \mu_j dt + \sqrt{V_j(t)} dW_j(t), \\ dV_j(t) &= \kappa_j(\theta_j - V_j(t))dt + \sigma_j \sqrt{V_j(t)} dW_j^*(t), \end{aligned} \tag{5.38}$$

where $dW_i(t)dW_j(t) = \rho_{ij}dt$ and $dW_j(t)dW_j^*(t) = \rho_j dt$ describe the correlations between two assets and the correlation between asset and its volatility respectively. The correlations ρ_{ij} can be estimated statistically from the historical data of S_i and S_j. The volatility correlation ρ_j can be obtained from the calibration of a single asset model to the respective market option prices.

What we need to specify additionally in the above multi-asset model is the cross-correlation between $i-$th asset and $j-$the volatility, namely $dW_i dW_j^* = \zeta_{ij}dt$, and the volatility cross-correlation between $i-$th volatility and $j-$the volatility, namely $dW_i^* dW_j^* = \zeta_{ij}^* dt$. Generally, it is difficult to estimate robustly these correlations because the implied volatility processes are unobservable and depend on strikes. Ad-hoc assumptions and specifications are then necessary and should be a better way in many cases. The entire correlation matrix Σ in the above multi-asset model is given by

$$\Sigma = \left(\begin{array}{c|cccccccc} & S_1 & S_2 & \cdots & S_N & v_1 & v_2 & \cdots & v_N \\ \hline S_1 & 1 & \rho_{12} & \cdots & \rho_{1N} & \rho_1 & \zeta_{12} & \cdots & \zeta_{1N} \\ S_2 & \rho_{12} & 1 & \cdots & \rho_{2N} & \zeta_{21} & \rho_2 & \cdots & \zeta_{2N} \\ \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots \\ S_N & \rho_{1N} & \rho_{2N} & \cdots & 1 & \zeta_{N1} & \zeta_{2N} & \cdots & \rho_N \\ v_1 & \rho_1 & \zeta_{21} & \cdots & \zeta_{N1} & 1 & \zeta_{12}^* & \vdots & \zeta_{1N}^* \\ v_2 & \zeta_{12} & \rho_2 & \cdots & \zeta_{N2} & \zeta_{12}^* & 1 & \vdots & \zeta_{2N}^* \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & 1 & \vdots \\ v_N & \zeta_{1N} & \zeta_{2N} & \vdots & \rho_N & \zeta_{1N}^* & \zeta_{2N}^* & \vdots & 1 \end{array} \right).$$

Intuitively we have two simple approaches to determining the correlations ζ_{ij} and ζ_{ij}^*. The first one is somehow too simplified, and just set both ζ_{ij} and ζ_{ij}^* equal to zero. The generated correlation matrix Σ is very sparse, and only the upper-link block in Σ is occupied entirely with elements. The drawback of this approach is that the matrix Σ can not be assured to be positive semi-definite. The application of Cholesky-decomposition and the generation of $2N$-dimensional Gaussian random variables are sometimes impossible.

The other approach is to set $dW_i dW_j^* = \zeta_{ij}dt = \rho_{ij}\rho_j dt$ and $dW_i^* dW_j^* = \zeta_{ij}^* dt = \rho_i\rho_j\rho_{ij}dt$. The correlation matrix Σ becomes

$$\Sigma = \begin{array}{c|c} & \begin{matrix} S_1 & S_2 & \cdots & S_N & v_1 & v_2 & \cdots & v_N \end{matrix} \\ \hline \begin{matrix} S_1 \\ S_2 \\ \vdots \\ S_N \\ v_1 \\ v_2 \\ \vdots \\ v_N \end{matrix} & \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1N} & \rho_1 & \rho_{12}\rho_2 & \cdots & \rho_{1N}\rho_N \\ \rho_{12} & 1 & \cdots & \rho_{2N} & \rho_{21}\rho_1 & \rho_2 & \cdots & \rho_{2N}\rho_N \\ \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{1N} & \rho_{2N} & \cdots & 1 & \rho_{N1}\rho_1 & \rho_{2N}\rho_2 & \cdots & \rho_N \\ \rho_1 & \rho_{21}\rho_1 & \cdots & \rho_{N1}\rho_1 & 1 & \rho_1\rho_2\rho_{12} & \vdots & \rho_1\rho_N\rho_{1N} \\ \rho_{12}\rho_1 & \rho_2 & \cdots & \rho_{N2}\rho_2 & \rho_1\rho_2\rho_{12} & 1 & \vdots & \rho_2\rho_N\rho_{2N} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & 1 & \vdots \\ \rho_{1N}\rho_N & \rho_{2N}\rho_N & \vdots & \rho_N & \rho_1\rho_N\rho_{1N} & \rho_2\rho_N\rho_{2N} & \vdots & 1 \end{pmatrix} \end{array}. \tag{5.39}$$

It can be shown that the resulting matrix Σ is always positive semi-definite, and then is a well-defined correlation matrix. Furthermore, this approach should coincide with the following method to generate random variables:

1. Step1: Generate N-dimensional correlated Gaussian random variables for $W_j, 1 \leqslant j \leqslant N$ with the correlation coefficients $\{\rho_{ij}\}, 1 \leqslant i, j \leqslant N$,
2. Step 2: Generate N-dimensional uncorrelated Gaussian random variables for $W_j^*, 1 \leqslant j \leqslant N$,
3. Step 3: Simulate every $S_j(t)$ and $v_j(t)$ using the calibrated correlation coefficient ρ_j.

To verify that the above simulation agrees with the correlation structure of the second approach, we denote the vector $(W_1, W_2, \cdots, W_N)$ as a N-dimensional standard normally distributed random variable, i.e.,

$$(W_1, W_2, \cdots, W_N)^\top \sim N(0, \Sigma_S),$$

with

$$\Sigma_S = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1N} \\ \rho_{21} & \ddots & \cdots & \rho_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N1} & \rho_{N2} & \cdots & 1 \end{pmatrix}, \quad \rho_{ij} = \rho_{ji},$$

which is exactly the upper-link block of the whole correlation matrix Σ. Step 1 above includes the following Cholesky-decomposition,

$$\Sigma_S = U \Lambda U^\top, \tag{5.40}$$

where U is an eigenvector matrix with the eigenvectors of Σ_S as column vector. Additionally, U is an unitary matrix,

$$UU^\top = I.$$

I is a diagonal matrix with the diagonal elements being one. Λ is the eigenvalue matrix with the eigenvalues of Σ as diagonal elements,

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix}.$$

Note that the eigenvalues λ_j should not be negative. On the other hand, it exists an orthogonal matrix A such that

$$\Sigma = AA^\top = U\Lambda U^\top,$$

which is equivalent to

$$A = U\Lambda^{\frac{1}{2}} = U \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_N} \end{pmatrix}.$$

Assume that $\{\lambda_j\}_{1\leqslant j\leqslant N}$ is a non-increasing sequence. The first few dominate components are often called the principal components. By using whole eigenvalues, we can rewrite the Brownian motion W_j with N independent standard Brownian motions Z,

$$W_j = \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} Z_l, \quad 1 \leqslant j \leqslant N, \quad 1 \leqslant l \leqslant N. \tag{5.41}$$

Therefore we can rewrite the system of stock prices via N independent standard Brownian motions Z,

$$\frac{dS_j(t)}{S_j(t)} = \mu_j dt + v_j(t) \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l, \quad 1 \leqslant j \leqslant N. \tag{5.42}$$

On the other hand, we denote the N independent standard Brownian motions as Z^* generated in Step 2, that are also uncorrelated with Z. The dynamics of stochastic volatilities may be expressed as follows:

$$dv_j(t) = \kappa_j(\theta_j - v_j(t))dt + \sigma_j \big[\rho_j \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l + \sqrt{1-\rho_j^2} Z_j^*\big].$$

Given the new formulations of the multi-asset model, it is easy to calculate the correlation of the diffusion terms. Note

$$< \sum_{l=1}^{N} \sqrt{\lambda_l} u_{il} dZ_l, \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l >= \rho_{ij},$$

it follows immediately

$$\begin{aligned} < W_i, W_j^* > &= < \sum_{l=1}^{N} \sqrt{\lambda_l} u_{il} dZ_l, \rho_j \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l + \sqrt{1-\rho^2} Z_j^* > \\ &= < \sum_{l=1}^{N} \sqrt{\lambda_l} u_{il} dZ_l, \rho_j \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l >= \rho_j \rho_{ij} \end{aligned}$$

and

$$\begin{aligned} < W_i^*, W_j^* > &= < \rho_i \sum_{l=1}^{N} \sqrt{\lambda_l} u_{il} dZ_l + \sqrt{1-\rho_i^2} Z_i^*, \rho_j \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l + \sqrt{1-\rho_j^2} Z_j^* > \\ &= \rho_i \rho_j < \sum_{l=1}^{N} \sqrt{\lambda_l} u_{il} dZ_l, \sum_{l=1}^{N} \sqrt{\lambda_l} u_{jl} dZ_l >= \rho_i \rho_j \rho_{ij}. \end{aligned}$$

Therefore, we have shown the above described simulation procedure implies the special correlation matrix given in (5.39). The proposed simulation does not need to decompose the whole large matrix Σ, but only a smaller matrix that corresponds to the upper-link block of Σ, and therefore requires much less computing time and is more efficient.

Chapter 6
Stochastic Interest Models

The aim of this chapter is not to present a complete overview of stochastic interest rate models, but to show to which extend stochastic interest rates can be incorporated into a pricing formula for European-style stock options. To this end, we focus on only three typical one-factor short rate models, namely, the Vasicek model (1977), the CIR model (1985) and the Longstaff model (1989), which are again specified by a mean-reverting Ornstein-Uhlenbeck process, a mean-reverting square root process and a mean-reverting double square root process, respectively. In turn, these three processes correspond to these ones in stochastic volatility models discussed in Chapter 3. Since stochastic short rate appears in a risk-neutral stock process as drift, it becomes impossible for a square root process to incorporate a correlation between the short rate and the stock diffusion term with CFs. Therefore, we propose a modification of the stock price process so that the CIR model and the Longstaff model may be embedded into an option pricing formula. However, the mean-reverting Ornstein-Uhlenbeck process can be nested with the stock price process for a non-zero correlation without any modification. The extension of a one-factor stochastic interest rate models into a multi-factor case is straightforward, and some multi-factor models can be incorporated into the valuation of stock options in analogy to the one-factor model if the independence conditions are satisfied.

6.1 Introduction

Nowadays stochastic interest rate models form a much larger theoretical field than stochastic volatility models, and even have a longer history because interest rates are the most important factor in economics. Interest rates are classified in short-term, medium-term and long-term rates according to the time to maturity of underlying bonds. Here we deal with only instantaneous short-term interest rates, or short rates, which obviously possesses some special properties: mean-reversion, stationarity, decreasing current shocks to future rates. All of these properties show that modeling the dynamics of instantaneous short rates should follow the similar way as modeling

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_6,

stochastic volatilities. In fact, the most short-rate models are formulated by using the following stochastic differential equation:

$$dr(t) = \kappa[\theta - r(t)^a]dt + \sigma r(t)^b dW, \tag{6.1}$$

with $a, b \geqslant 0$. The parameters κ and θ have the same meaning as in stochastic volatility models, and represent the reversion velocity and mean level respectively. The general model (6.1) is referred to as the time-homogenous one-factor model since the parameters in this process are time-independent and the resulting bond prices are determined solely by the short rate $r(t)$. With a special choice of distinct values for the parameters a and b, we can obtain different short-rate models which are given partially in Table (6.1).

(1):	$dr(t) = \mu dt + \sigma dW(t)$	Merton (1978)	$a = 0$	$b = 0$
(2):	$dr(t) = \kappa(\theta - r)dt + \sigma dW(t)$	Vasicek (1977)	$a = 1$	$b = 0$
(3):	$dr(t) = \sigma r dW(t)$	Dothan (1978)	$a = 0$	$b = 1$
(4):	$dr(t) = \kappa(\theta - r)dt + \sigma r dW(t)$	Courtadon (1982)	$a = 1$	$b = 1$
(5):	$dr(t) = \kappa(\theta - r)dt + \sigma\sqrt{r} dW(t)$	CIR (1985)	$a = 1$	$b = 0.5$
(6):	$dr(t) = \kappa(\theta - \sqrt{r})dt + \sigma\sqrt{r} dW(t)$	Longstaff (1989)	$a = 0.5$	$b = 0.5$

Table 6.1 Overview of one-factor short interest rate models.

Short rates are not a directly traded good, hence the above models are not complete under the original statistical measure P. However, as long as the risk-neutral process of short rates is established by introducing a market price of risk, we can calculate the underlying zero-coupon bond price by applying the local expectations hypothesis, which postulates that the current bond price is the expected value of the discounted terminal value by rolling over the short rates from now to the terminal time.[1] Then a zero-coupon bond with one dollar face value can be expressed by

$$\begin{aligned} B(0,T) &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt \cdot B(T,T)\right)\right] \\ &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\right] \quad \text{for} \quad B(T,T) = 1. \end{aligned} \tag{6.2}$$

To obtain the pricing formula for $B(0,\tau)$, we can employ the Feynman-Kac formula again. The corresponding PDE reads

$$\frac{\partial B}{\partial \tau} = -rB + \kappa[\theta - r^a]\frac{\partial B}{\partial r} + \frac{1}{2}\sigma r^{2b}\frac{\partial^2 B}{\partial r^2}, \tag{6.3}$$

with $\tau = T - t$ and the boundary condition

[1] This hypothesis is more explicit in a discrete-time setting. There are other two expectations hypothesis on short-rates: return-to-maturity expectations hypothesis and yield-to-maturity expectations hypothesis. For a detailed discussion see Ingersoll (1987).

$$B(T,T) = B(\tau = 0) = 1.$$

The standard method to solve this PDE is to suggest an exponential solution structure, that is

$$B(\tau; r_0) = \exp(a_1(\tau) + a_2(\tau) r_0^a + a_3(\tau) r_0^{2b}). \tag{6.4}$$

With this guess, a closed-form solution for the zero-coupon bond price is given for all models except for model (4). With $B(0,T)$ having been known, the term structure of interest rates is automatically available. Among the listed models, the Vasicek model and the CIR model are termed as affine models of the term structure and play a central role in modeling stochastic interest rates. Duffie and Kan (1996) provide the necessary and sufficient conditions on this representation in a multivariate setting.

Not all models in Table (6.1) are reasonable with respect to the nature of interest rates and only gain their significance in a historical context. The models (1) and (3) have an unlimited variance if time goes to infinity whereas the model (4) is stationary only for $2\kappa > \sigma^2$. Additionally, the models (1) and (3) do not perform the feature of mean-reversion. The positive value of interest rate is guaranteed by all model except for the Vasicek model. However, for most parameters that are consistent with empirical values, the Vasicek model raises only a negligibly small probability of negative interest rates. Model (6) is also called "double square root process" model and has a closed-form solution for the zero-coupon bond price only for the special case $4\kappa\theta = \sigma^2$. In the following three sections, we will only concentrate on models (1), (5) and (6), all of which are popular in interest rate theory, and incorporate them into our option pricing models.

There are some further developments and extensions in modeling interest rates:

1. Time-inhomogeneous processes are proposed to specify interest rate dynamics and lead to an inversion of the term structure of interest rates (Hull and White, 1990. 1993);
2. Starting with the current term structure of spot rates or future rates, Ho and Lee (1986), Heath, Jarrow and Morton (1992) attempted to model directly the dynamics of the term structure. A further development of this approach is Libor Market Model (LMM), and also called the BGM model in honor of a contribution of Brace, Gatarek and Musiela (1997). We will devote ourself in the last chapter to address the smile modeling in LMM using the Fourier transform.
3. An interest rate process may be specified in a way that all parameters are matched as good as possible to the given current term structure of interest rates. (Brown and Dybvig, 1986).

These approaches are strongly associated with the term structure and not directly with the dynamics of short interest rates. Hence they are difficult to be incorporated into the valuation of asset (equity-like) options, and will not considered here.

6.2 The Cox-Ingosoll-Ross Model

Modeling interest rates as a square root process was for the first time suggested by CIR (1985b) in an equilibrium model. The advantage of this modeling is that the interest rates can never subsequently become negative if their initial values are nonnegative.[2] Formally, the interest rate dynamics are given by

$$dr(t) = \kappa[\theta - r(t)]dt + \sigma\sqrt{r(t)}dW_3. \tag{6.5}$$

6.2.1 The Zero-Correlation Case

At first, we consider the case $dW_3 dW_1 = 0$, that is, the short rates are uncorrelated with the stock returns in spite of some empirical objections to this simplified assumption. We will relax this assumption later. The market price of risk to $r(t)$ is determined endogenously by employing equilibrium arguments. Therefore, we regard (6.5) as risk-neutralized. As discussed in the above section, process (6.5) satisfies many properties that a short rate model should have. Following the principle of constructing CFs, we have the following two Radon-Nikydom derivatives in the presence of stochastic interest rates,

$$g_1(T) = \exp\left(-\int_0^T r(t)dt\right)\frac{S(T)}{S_0} \tag{6.6}$$

and

$$g_2(T) = \exp\left(-\int_0^T r(t)dt\right)\frac{1}{B(0,T)}. \tag{6.7}$$

The first CF may be calculated as follows:

$$\begin{aligned}
&f_1(\phi)\\
&= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\frac{S(T)}{S_0}\exp\left(i\phi x(T)\right)\right]\\
&= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt - x_0 + (i\phi+1)x(T)\right)\right]\\
&= \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt + i\phi x_0 + (i\phi+1)\left(-\frac{1}{2}\int_0^T v^2 dt + \int_0^T v dW_1\right)\right)\right].
\end{aligned}$$

Since the interest rate is independent of dW_1 and v, we calculate the expected value associated with $r(t)$ separately. It follows

[2] Two key assumptions in the original work of CIR are (i) the utility function is logarithmic and thus the interest rate has a linear relationship to the state variable in a production economy, and (ii) the state variable (technological change) follows a square-root process. Consequently, interest rates follow also square-root processes. In this sense, the result about interest rates is state-variable dependent, not derived from the equilibrium-coincide logic.

$$f_1(\phi) = \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt\right)\right]$$
$$\times\,\mathbf{E}\left[\exp\left(i\phi x_0 + (i\phi+1)\left(-\frac{1}{2}\int_0^T v^2 dt + \int_0^T v dW_1\right)\right)\right]. \tag{6.8}$$

The second expectation in the above equation should be computed according to the specifications of volatility process $v(t)$. In the Black-Scholes world, we can immediately obtain

$$f_1(\phi) = \exp\left(i\phi x_0 + \frac{1}{2}i\phi(i\phi+1)v^2T\right)$$
$$\times\,\mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt\right)\right]. \tag{6.9}$$

The remaining expectation $\mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt\right)\right]$ can be calculated by using the formula (3.20) in Chapter 3. Note the term $\exp\left(\frac{1}{2}i\phi(i\phi+1)v^2T\right)$ is the CF associated with constant volatility. By denoting it as $f_1^{BS}(\phi)$, we can represent the CF $f_1(\phi)$ as follows:

$$f_1(\phi) = \exp(i\phi x_0)f_1^{BS}(\phi)f_1^{CIR}(\phi) \tag{6.10}$$

with

$$f_1^{CIR}(\phi) = \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt\right)\right]. \tag{6.11}$$

Following the same steps, we arrive at the following expression for $f_2(\phi)$,

$$f_2(\phi) = \exp\left(i\phi x_0 + \frac{1}{2}i\phi(i\phi-1)v^2T\right)$$
$$\times\,\mathbf{E}\left[\exp\left((i\phi-1)\int_0^T r(t)dt\right)/B(0,T)\right]$$
$$= \exp(i\phi x_0)\,f_2^{BS}(\phi)f_2^{CIR}(\phi) \tag{6.12}$$

with

$$f_2^{CIR}(\phi) = \mathbf{E}\left[\exp\left((i\phi-1)\int_0^T r(t)dt\right)/B(0,T)\right]. \tag{6.13}$$

In the same manner, the expected value $\mathbf{E}\left[\exp\left((i\phi-1)\int_0^T r(t)dt\right)\right]$ can be evaluated by applying the formula (3.20).

So far we have inserted the stochastic interest rates into the CFs, and therefore, are able to price options with stochastic interest rates for the case of zero correlation. A favorite property of this case is that the expected values $\mathbf{E}[\cdot]$ are purely associated with the interest rate $r(t)$, and are not nested with the stochastic volatility v. This allows us to specify a stochastic v arbitrarily and to proceed in the same way as shown in Section 3.2 as long as the interest rate and the volatility are not mutually correlated. Unfortunately, when we let both stochastic factors be correlated, we can not derive the tractable CFs even if the volatility v is stochastically independent of stock returns.

6.2.2 The Correlation Case

To extend the zero correlation case to the non-zero correlation case, we need a modification of the stock price process that is given by

$$\frac{dS(t)}{S(t)} = r(t)dt + v\sqrt{r(t)}dW_1, \tag{6.14}$$

where we allow $dW_1 dW_3 = \rho dt$. The term $\sqrt{r(t)}$ as a stochastic part of volatility is firstly introduced in Scott (1997), and as an advantageous feature, the modified process guarantees positive interest rates. We can justify this model by the following observations: (i) Interest rates are inversely correlated with stock returns, as reported in Scott (1997), this is in line with the relationship between volatilities and stock returns; (ii) Both interest rates and volatilities share similar properties such as mean-reversion, and both processes can be specified by the same class of SDEs. Thus, the stock price process given in (6.14) is the simplest and parsimonious process that allows for the homogeneous stochastic interest rate and stochastic volatility, both of which are driven by a single factor "$r(t)$". In fact, process (6.14) can be extended to a process driven by the affine structure of the factor "$r(t)$",

$$\frac{dS(t)}{S(t)} = [a_1 \cdot r(t) + b_1]dt + v\sqrt{a_2 \cdot r(t) + b_2}dW_1. \tag{6.15}$$

However, we concentrate still on the simple process (6.14) to illustrate how to incorporate the correlation between the interest rate and the stock returns into the option valuation.

With $x(t) = \ln S(t)$, we have

$$dx(t) = r(t)(1 - \frac{1}{2}v^2)dt + v\sqrt{r(t)}dW_1. \tag{6.16}$$

Consequently, the CFs can be calculated as follows:

$$\begin{aligned} f_1(\phi) &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\frac{S(T)}{S_0}\exp\left(i\phi x(T)\right)\right] \\ &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right) - x_0 + (1+i\phi)x(T)\right] \end{aligned}$$

$$
\begin{aligned}
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2\right)\int_0^T r(t)dt\right)+i\phi x_0+v\int_0^T\sqrt{r(t)}dW_1\right]\\
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2\right)\int_0^T r(t)dt\right)+i\phi x_0\right.\\
&+\ \left.(1+i\phi)\left(\rho v\int_0^T\sqrt{r(t)}dW_3+\sqrt{1-\rho^2}v\int_0^T\sqrt{r(t)}dW\right)\right]\\
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2+\frac{1}{2}(1+i\phi)^2(1-\rho^2)v^2\right)\int_0^T r(t)dt\right)\right.\\
&+\ \left.\frac{(1+i\phi)\rho v}{\sigma}[r(T)-r_0-\kappa\theta T+\kappa\int_0^T r(t)dt]+i\phi x_0\right]\\
&= \mathbf{E}\left[\exp\left(i\phi x_0+\frac{(1+i\phi)\rho v}{\sigma}r(T)-\frac{(1+i\phi)\rho v}{\sigma}(r_0+\kappa\theta T)\right.\right.\\
&+\ \left.\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2+\frac{1}{2}(1+i\phi)^2(1-\rho^2)v^2+\frac{(1+i\phi)\rho v\kappa}{\sigma}\right)\int_0^T r(t)dt\right]\\
&= \mathbf{E}\left[\exp\left(i\phi x_0-s_1\int_0^T r(t)dt+s_2r(T)-s_2(r_0+\kappa\theta T)\right)\right]\\
&= \exp(i\phi x_0)f_1^{CIR-Corr}(\phi)
\end{aligned}
\tag{6.17}
$$

with

$$
\begin{aligned}
s_1 &= -i\phi(1-\frac{1}{2}v^2)+\frac{1}{2}v^2-\frac{1}{2}(1+i\phi)^2(1-\rho^2)v^2-\frac{(1+i\phi)\rho v\kappa}{\sigma},\\
s_2 &= \frac{(1+i\phi)\rho v}{\sigma}.
\end{aligned}
$$

By using again formula (3.20) in Section 3.2, we obtain the final closed-form solution for $f_1(\phi)$ in the case of correlation. Since the modified stock process $S(t)$ has incorporated stochastic volatility, the CF $f_1(\phi)$ for non-zero correlation case does not have the term $f_1^{BS}(\phi)$.

In the same fashion, the second CF $f_2(\phi)$ is given by

$$
\begin{aligned}
&f_2(\phi)\\
&= \mathbf{E}\left[\frac{\exp\left(-\int_0^T r(t)dt\right)}{B(0,T)}\exp\left(i\phi x(T)\right)\right]\\
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)+\frac{i\phi v\rho\kappa}{\sigma}-\frac{1}{2}\phi^2v^2(1-\rho^2)-1\right)\int_0^T r(t)dt\right.\right.\\
&+\ \left.\left.\frac{i\phi v\rho}{\sigma}r(T)\right)\right]\times\exp\left(-\frac{i\phi v\rho}{\sigma}(\kappa\theta T+r_0)+x_0-\ln B(0,T)\right)\\
&= \mathbf{E}\left[\exp\left(x_0-\ln B(0,T;r_0)-s_1^*\int_0^T r(t)dt+s_2^*r(T)-s_2^*(\kappa\theta T+r_0)\right)\right]\\
&= \exp(i\phi x_0)f_2^{CIR-Corr}(\phi)
\end{aligned}
\tag{6.18}
$$

with

$$
\begin{aligned}
s_1^* &= -i\phi(1-\frac{1}{2}v^2) - \frac{i\phi v\rho\kappa}{\sigma} + \frac{1}{2}\phi^2 v^2(1-\rho^2)+1,\\
s_2^* &= \frac{i\phi v\rho}{\sigma}.
\end{aligned}
$$

Given these two CFs, we can evaluate option prices if stochastic interest rates are correlated with the stock returns. This model is certainly of interest when concerning the strong impact of the short-rates on market liquidity and, hence, also on stock prices, especially stock indices.

6.3 The Vasicek Model

As shown above, the drawback of modeling interest rates as a square root process in option pricing is that we need an alternative stock price process (6.16) for the case of non-zero correlation, otherwise the closed-form solution is not available. In order to overcome this drawback, as suggested by Rabinovitch (1989), we can introduce the Vasicek model (1977) where short interest rates are governed by a risk-neutralized mean-reverting Ornstein-Uhlenbeck process as follows:

$$dr(t) = \kappa[\theta - r(t)]dt + \sigma dW_3. \tag{6.19}$$

At the same time, stock price process is still modeled by the usual geometric Brownian motion. The correlation is permitted, i.e., $dW_1 dW_3 = \rho dt$. The analytical forms of the two CFs in this case will be more complicated than their counterparts in the CIR model. We calculate $f_1(\phi)$ and $f_2(\phi)$ in details.

$$
\begin{aligned}
&f_1(\phi)\\
&= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\frac{S(T)}{S_0}\exp\left(i\phi x(T)\right)\right]\\
&= \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt + i\phi x_0 + (i\phi+1)\left(-\frac{1}{2}\int_0^T v^2 dt + \int_0^T v dW_1\right)\right)\right]\\
&= \exp\left(i\phi x_0 - \frac{1}{2}(i\phi+1)v^2 T\right)\\
&\times \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt + (i\phi+1)\rho v\int_0^T dW_3 + (i\phi+1)\sqrt{1-\rho^2}v\int_0^T dW\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \exp\left(i\phi x_0 - \frac{(i\phi+1)}{2}v^2T\right.\\
&- \left.\frac{(i\phi+1)v\rho}{\sigma}(r_0+\kappa\theta T) + \frac{1}{2}(i\phi+1)^2v^2(1-\rho^2)T\right)\\
&\times \mathbf{E}\left[\exp\left(\left(i\phi+\frac{(1+i\phi)v\rho\kappa}{\sigma}\right)\int_0^T r(t)dt + \frac{(1+i\phi)v\rho}{\sigma}r(T)\right)\right]\\
&= \exp\left(i\phi x_0 - \frac{(i\phi+1)}{2}v^2T\right.\\
&- \left.\frac{(i\phi+1)v\rho}{\sigma}(r_0+\kappa\theta T) + \frac{1}{2}(i\phi+1)^2v^2(1-\rho^2)T\right)\\
&\times \mathbf{E}\left[\exp\left(-s_1\int_0^T r(t)dt + s_2 r(T)\right)\right]\\
&= \exp(i\phi x_0) f_1^{Vasicek-Corr}(\phi) \qquad (6.20)
\end{aligned}
$$

with

$$s_1 = -i\phi - \frac{(1+i\phi)v\rho\kappa}{\sigma}, \qquad s_2 = \frac{(1+i\phi)v\rho}{\sigma}.$$

To compute the expected value in the last equality of (6.20), we can use the Feynman-Kac formula to obtain a PDE as follows:

$$\frac{\partial y}{\partial \tau} = -s_1 r y + \kappa(\theta - r)\frac{\partial y}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 y}{\partial r^2}, \qquad (6.21)$$

subject to the boundary condition

$$y(r_0, 0) = \exp(s_2 r_0).$$

By applying the standard method, we obtain a general result as follows:

$$\mathbf{E}\left[\exp\left(-s_1\int_0^T r(t)dt + s_2 r(T)\right)\right] = \exp(J(T)r_0 + K(T)) \qquad (6.22)$$

with

$$
\begin{aligned}
J(T) &= \frac{s_1 - (s_2\kappa + s_1)e^{-\kappa T}}{-\kappa},\\
K(T) &= \left(\frac{\sigma^2 s_1^2}{2\kappa^2} - s_1\theta\right)T + \left(\frac{\sigma^2 s_1}{\kappa^3} - \frac{\theta}{\kappa}\right)(s_2\kappa + s_1)(e^{-\kappa T} - 1)\\
&\quad - \frac{\sigma^2}{4\kappa^3}(s_2\kappa + s_1)^2(e^{-2\kappa T} - 1).
\end{aligned}
$$

The calculation of $f_2(\phi)$ follows the same steps,

$$
\begin{aligned}
& f_2(\phi) \\
&= \mathbf{E}\left[\frac{\exp\left(-\int_0^T r(t)dt\right)}{B(0,T)}\exp\left(i\phi x(T)\right)\right] \\
&= \exp\left(i\phi x_0 - \frac{i\phi}{2}v^2 T - \frac{i\phi v\rho}{\sigma}(r_0+\kappa\theta T) - \frac{1}{2}\phi^2 v^2(1-\rho^2)T - \ln B(0,T)\right) \\
&\times \mathbf{E}\left[\exp\left(\left(i\phi + \frac{i\phi v\rho\kappa}{\sigma} - 1\right)\int_0^T r(t)dt + \frac{i\phi v\rho}{\sigma}r(T)\right)\right] \\
&= \exp\left(i\phi x_0 - \frac{i\phi}{2}v^2 T - s_2^*(r_0+\kappa\theta T) - \frac{1}{2}\phi^2 v^2(1-\rho^2)T - \ln B(0,T)\right) \\
&\times \mathbf{E}\left[\exp\left(-s_1^*\int_0^T r(t)dt + s_2^* r(T)\right)\right] \\
&= \exp(i\phi x_0) f_2^{Vasicek-Corr}(\phi)
\end{aligned} \tag{6.23}
$$

with

$$
s_1^* = -i\phi - \frac{i\phi v\rho\kappa}{\sigma} + 1, \qquad s_2^* = \frac{i\phi v\rho}{\sigma}.
$$

As long as both CFs are known analytically, the closed-form option pricing formula is given by the Fourier inversion for the case where stochastic interest rates are correlated with stock returns, and is identical to Rabinovitch's solution (1989).This model may provide us with some new insights into the impact of stochastic interest rates on option prices. Scott (1997) developed a jump-diffusion model with stochastic interest rates which are separated in two state variables. A difficult issue in the practical implementation of his model is to identify these two variables. In Subsection 6.5, we will study this model in more details and attempt to answer the following question: Which correlation with stock returns is more important for a correct valuation of options? The correlation of stock returns with interest rates or with volatilities?

By setting $\rho = 0$, two CFs are simplified and essentially similar to (6.11) and (6.13), namely,

$$
f_j(\phi) = \exp(i\phi x_0) f_j^{BS} f_j^{Vasicek}(\phi), \quad j = 1,2, \tag{6.24}
$$

which allows us to extend the pricing models by incorporating stochastic volatilities.

6.4 The Longstaff Model

Longstaff (1989) set up an alternative general equilibrium model of the term structure within the CIR theoretical framework by assuming that the state variable for technological change follows a random walk. The resulting process for short interest rates is a mean-reverting double square root process that has been applied to model stochastic volatility in Chapter 3. Longstaff incorrectly claimed that the square root of the interest rate $r(t)$ follows a process reflected at $r(t) = 0$. Beaglehole and Tenney (1992) showed by simulation that Longstaff's solution for the zero-coupon bond price is not identical to the one in the reflecting case of the double square root process. As shown in Subsection 3.4.1, a double square root process is generated by a Brownian motion with drift, and is not imposed by any boundary condition.

Since interest rates are governed by the square of a state variable following a Brownian motion with drift, one can easily obtain a double square root process by applying Itô's lemma. Thus, the dynamics of the equilibrium interest rates are given by

$$dr(t) = \kappa\left[\theta - \sqrt{r(t)}\right]dt + \sigma\sqrt{r(t)}dW_3(t), \tag{6.25}$$

where the restriction $4\kappa\theta = \sigma^2$ is imposed on this model and a zero value of $r(t)$ is accessible, as discussed in Subsection 3.4.1. Process (6.25) shares most of the properties of the CIR square root process except that interest rates revert to the square root of itself in a double square root process. This implausible asymptotic property raises doubts to this model. However, a double square root process has two potential advantages: (i) only two parameters are required to be estimated because of the parameter restriction, and (ii) the yields of the zero-coupon bond derived from this model are nonlinear, more precisely, are a function with the terms r and $\sqrt{r}$. As reported by Longstaff (1989), this so-called nonlinear term structure outperforms the CIR square root process in describing actual Treasury Bill yields for the period of 1964-1986 where he used the GMM to obtain the parameter estimates. Here, we adopt his double square root process to specify the interest rates. Similarly to the CIR model, with a market price of risk proportional to $r(t)$, we have the following risk-neutral process,

$$dr(t) = \left[\kappa\theta - \kappa\sqrt{r(t)} - \lambda r(t)\right]dt + \sigma\sqrt{r(t)}dW_3(t). \tag{6.26}$$

As for the CIR model, to incorporate Longstaff interest rate model into the valuation of European-style options, we regard two cases: zero-correlation case and correlation case.

6.4.1 The Zero-Correlation Case

Two CFs in this case are simple and take the same form as in CIR model. For the first CF, we have

$$\begin{aligned} f_1(\phi) &= \exp\left(i\phi x_0 + \frac{1}{2}i\phi(i\phi+1)v^2T\right) \\ &\quad \times \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt\right)\right] \\ &= \exp(i\phi x_0) f_1^{BS}(\phi) f_1^{Longstaff}(\phi) \end{aligned} \tag{6.27}$$

with

$$f_1^{Longstaff} = \mathbf{E}\left[\exp\left(i\phi\int_0^T r(t)dt\right)\right]. \tag{6.28}$$

For the second CF, we have

$$\begin{aligned} f_2(\phi) &= \exp\left(i\phi x_0 + \frac{1}{2}i\phi(i\phi-1)v^2T\right) \\ &\quad \times \mathbf{E}\left[\exp\left((i\phi-1)\int_0^T r(t)dt\right)/B(0,T)\right] \\ &= \exp\left(i\phi x_0\right) f_2^{BS}(\phi) f_2^{Longstaff}(\phi) \end{aligned} \tag{6.29}$$

with

$$f_2^{Longstaff}(\phi) = \mathbf{E}\left[\exp\left((i\phi-1)\int_0^T r(t)dt\right)/B(0,T)\right]. \tag{6.30}$$

The two expected values appearing in the above CFs may be computed according to the results for double square root process in Appendix B in Chapter 3.

6.4.2 The Correlation Case

As in Section 6.2, the stock prices are assumed to have the same dynamics as given by process (6.14) and allowed to be correlated with the interest rates, that is $dW_1dW_3 = \rho dt$. Therefore, the corresponding CFs are calculated as follows:

$$\begin{aligned} f_1(\phi) &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\frac{S(T)}{S_0}\exp\left(i\phi x(T)\right)\right] \\ &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right) - x_0 + (1+i\phi)x(T)\right] \\ &= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2) - \frac{1}{2}v^2\right)\int_0^T r(t)dt\right) + i\phi x_0 + v\int_0^T \sqrt{r(t)}dW_1\right] \end{aligned}$$

$$
\begin{aligned}
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2\right)\int_0^T r(t)dt\right)+i\phi x_0\right.\\
&+\ \left.(1+i\phi)\left(\rho v\int_0^T \sqrt{r(t)}dW_3+\sqrt{1-\rho^2}v\int_0^T \sqrt{r(t)}dW\right)\right]\\
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2+\frac{1}{2}(1+i\phi)^2(1-\rho^2)v^2\right)\int_0^T r(t)dt\right)\right.\\
&+\ \frac{(1+i\phi)\rho v}{\sigma}\left(r(T)-r_0-\kappa\theta T+\lambda\int_0^T r(t)dt+\kappa\int_0^T \sqrt{r(t)}dt\right)\\
&+\ \left.i\phi x_0\right]\\
&= \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)-\frac{1}{2}v^2+\frac{1}{2}(1+i\phi)^2(1-\rho^2)v^2\right.\right.\right.\\
&+\ \left.\frac{(1+i\phi)\rho v\lambda}{\sigma}\right)\int_0^T r(t)dt+\frac{(1+i\phi)\rho v\kappa}{\sigma}\int_0^T \sqrt{r(t)}dt\\
&+\ \left.\frac{(1+i\phi)\rho v}{\sigma}r(T)-\frac{(1+i\phi)\rho v}{\sigma}(r_0+\kappa\theta T)+i\phi x_0\right]\\
&= \mathbf{E}\left[\exp\left(i\phi x_0-s_3(r_0+\kappa\theta T)\right.\right.\\
&-\ \left.\left.s_1\int_0^T r(t)dt-s_2\int_0^T \sqrt{r(t)}dt+s_3 r(T)\right)\right]\\
&= \exp(i\phi x_0)f_1^{Longstaff-Corr}(\phi)
\end{aligned}
\tag{6.31}
$$

with

$$
\begin{aligned}
s_1 &= -i\phi(1-\frac{1}{2}v^2)+\frac{1}{2}v^2-\frac{1}{2}(1+i\phi)^2(1-\rho^2)v^2-\frac{(1+i\phi)\rho v\lambda}{\sigma},\\
s_2 &= -\frac{(1+i\phi)v\rho\kappa}{\sigma}, \qquad s_3=\frac{(1+i\phi)v\rho}{\sigma},
\end{aligned}
$$

and

$$
\begin{aligned}
&f_2(\phi)\\
&= \mathbf{E}\left[\frac{\exp\left(-\int_0^T r(t)dt\right)}{B(0,T)}\exp\left(i\phi x(T)\right)\right]\\
&= \exp\left(i\phi x_0+\frac{i\phi v\rho}{\sigma}(\kappa\theta T+r_0)-\ln B(0,T)\right)\\
&\times \mathbf{E}\left[\exp\left(\left(i\phi(1-\frac{1}{2}v^2)+\frac{i\phi v\rho\lambda}{\sigma}-\frac{1}{2}\phi^2v^2(1-\rho^2)-1\right)\int_0^T r(t)dt\right.\right.\\
&+\ \left.\left.\frac{i\phi v\rho\kappa}{\sigma}\int_0^T \sqrt{r(t)}dt+\frac{i\phi v\rho}{\sigma}r(T)\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \exp\left(i\phi x_0 - s_3^*(\kappa\theta T + r_0) - \ln B(0,T)\right) \\
&\times \mathbf{E}\left[\exp\left(-s_1^*\int_0^T r(t)dt - s_2^*\int_0^T \sqrt{r(t)}dt + s_3^* r(T)\right)\right] \\
&= \exp(i\phi x_0) f_2^{Longstaff-Corr}(\phi)
\end{aligned}
\tag{6.32}
$$

with

$$
\begin{aligned}
s_1^* &= -\left(i\phi(1-\frac{1}{2}v^2) + \frac{i\phi v\rho\lambda}{\sigma} - \frac{1}{2}\phi^2 v^2(1-\rho^2) - 1\right), \\
s_2^* &= -\frac{i\phi v\rho\kappa}{\sigma}, \qquad s_3^* = \frac{i\phi v\rho}{\sigma}.
\end{aligned}
$$

By using the results in Appendix B in Chapter 3 again, we can immediately obtain the closed-form solutions for the CFs and the pricing formula for the zero-coupon bond price $B(0,T)$. Hence, the closed-form formula for options in this case is derived. We have shown at the same time that the stochastic interest rate model and the stochastic volatility model share the same expectation functional.

6.5 Correlations with Stock Returns: SI versus SV

In contrast to stochastic volatilities, few attention is paid to the impact of the correlation between stochastic interest rates and stock returns on the pricing issue of options. In this chapter, we have established some option pricing models focusing on this impact, and then are able to compare the effects of the correlation between stock returns with stochastic interest rates (SI) and stochastic volatility (SV) respectively. Here we expound these models with respect to the correlation.

Incorporating correlation between stock returns and stochastic interest rates into option valuation is justified by the empirical evidence that this correlation is statistically significantly negative. For example, Scott (1997) reported that the monthly correlation between the S&P 500 index returns and the rates on three-month Treasury bills over the period 1970 to 1987 is -0.158. If the crash in October 1987 is included, then this correlation over the period 1979 to 1990 is -0.096. A negative correlation can be explained in economic terms: If interest rates increase, then market liquidity becomes lower and opportunity costs will be higher. As a consequence, stock prices will also move downwards. For the same reason, stock prices tend to go up if interest rates become lower. Additionally, from the point of view of real economics, higher interest rates lead to larger financing costs for future projects and reduce investor's incentive for interesting. Therefore, stock markets react to higher interest rates generally negatively.

A negative correlation changes the terminal distribution of stock prices in a way that both tails of the distribution are thinner than a lognormal distribution if given a same mean and a same standard deviation. This is because stock prices tend to grow slowly if spot stock price is high in a framework of risk-neutral pricing, and vice versa. Consequently, call and put options will be overvalued by the Black-Scholes formula regardless of ITM or OTM. By contrast, a positive correlation between stock returns and SI causes the fatter tails of the true distribution of stock prices. Thus, in this case, the Black-Scholes formula undervalues call and put options regardless of moneyness. In Chapter 3, we have discussed stochastic volatility models in details and found that the Black-Scholes formula overprices OTM call options and undervalues ITM call options if a negative correlation between stock returns and SV is present. A positive correlation results in an opposite relation in a stochastic volatility model. The choice of two alternative correlations in a practical application is then dependent on which pattern of stock price distribution fits the true distribution better. This requires more empirical investigations and perhaps changes from case to case.

Rabinovitch (1989) provided an option pricing formula in a Gaussian model where interest rates follow a mean-reverting Ornstein-Uhlenbeck process and the correlation between stock returns and interest rates is allowed. In a single-factor model of interest rates, the zero-coupon bond pricing formula generally takes the following form,

$$B(0,T) = \exp[A(T)r_0 + C(T)].$$

Based on this result, a call option can be valued by the following formula,[3]

$$C = S_0 N(d_1) - KB(0,T)N(d_2) \tag{6.33}$$

with

$$d_{1,2} = \frac{\ln(S_0/K) - \ln B(0,T) \pm \frac{1}{2}\eta^2(T)}{\eta(T)} \tag{6.34}$$

and

$$\eta^2(T) = v^2 T + \int_0^T [\sigma^2 A(t)^2 - 2\rho v \sigma A(t)]dt. \tag{6.35}$$

Thus we can take stochastic interest rates and their correlation with stock returns ρ into account when the variance $v^2 T$ in the Black-Scholes formula is replaced by the term $\eta^2(T)$. Note that the function $A(T)$ in the Vasicek model is deterministic, the equation (6.35) can be integrated. Formula (6.33) is simple and easy to implement. If we use $A(T)$ in the Vasicek model, we will obtain the same pricing model for which a Fourier transform pricing formula is given by the CFs (6.20) and (6.23). Another alternative model considering such non-zero correlation was developed by Scott (1997). In his model, stochastic interest rates and stochastic volatilities share a common state variable which is correlated with stock returns. Obviously, this implies the unappealing feature that correlations of stock returns with interest rates and with volatilities are same. In addition, it is difficult to identify the two state variables specified in his model.

[3] It is assumed that the volatility v is a bounded deterministic function (see Rabinovich (1989), or Musiela and Rutkowski (2005)). Hence, stochastic volatilities can not be incorporated into this pricing formula.

Based on the CFs (6.20) and (6.23), we present some numerical examples which demonstrate the effects of correlation between stock returns and interest rates on option prices. Here we calculate only one BS benchmark price by setting the discounter factor to be the corresponding zero-coupon bond price. The BS values given in Panel U are benchmarks for the SI option prices in the corresponding position across Panel V to Panel X in Table (6.3). Table (6.2) shows the impact of ρ on stock prices. The first finding is that the prices of call options are an increasing function of correlation regardless of moneyness. In fact, the prices of put options are also increasing with correlation and are not reported in Table (6.2) to save space; Secondly, we observe in Panel R that the BS option values best fit to the model values if we set $\rho = 0$ and $\theta_r = r$ (the long-term level is equal to the spot rate). In connection with the first finding, the Black-Scholes formula overvalues (undervalues) options in presence of negative (positive) correlation between stochastic interest rates and stock returns regardless of moneyness. This result is different from what we gain from the SV models presented in Chapter 3. The overvaluation or undervaluation due to correlations in the SV models depends on moneyness; Thirdly, we find that long-term level of interest rate θ_r is important for pricing options, especially for long-term options. Across the panels R to T, options prices change significantly with θ_r. Table (6.3) also supports this finding. Hence, this model performs the same sensitivity of option prices to long-term level as the stochastic volatility models. In Table (6.3), we see that option prices are also an increasing function of θ_r. This is not surprising because option prices go up with increasing interest rates even in the BS model. The option values in Panel V are generally very close to the option values in Panel W where the correlation between the stock returns and interest rates is zero. This confirms that the SI option prices can be very closely approximated by the BS option prices in the case of zero-correlation. In this respect the SI models resemble the SV models.

The SI model differs from the SV model mainly in that ITM options in the SI model increase in value with the correlation ρ. Given a negative correlation, what we need to know to choose between these two alternative models is whether the actual market prices of ITM options are undervalued by the BS formula. If this is the case, we prefer the SV model, and otherwise we prefer the SI model. An empirical study made by Rubinstein (1978) shows that the relative pricing biases of the Black-Scholes formula do not display a persistent pattern over time and change from period to period. If his study presents a true picture of options markets, then the correlated SI model might provide us with a potential application for pricing options.

ρ \ K	90	95	100	105	110	115	120
BS	15.114	11.338	8.138	5.581	3.656	2.291	1.376
–1.00	14.918	11.048	7.774	5.184	3.274	1.961	1.115
–0.75	14.968	11.125	7.871	5.290	3.376	2.048	1.183
–0.50	15.019	11.201	7.967	5.395	3.477	2.135	1.251
–0.25	15.071	11.276	8.061	5.497	3.575	2.220	1.319
0.00	15.122	11.351	8.154	5.597	3.672	2.305	1.387
0.25	15.174	11.425	8.245	5.696	3.768	2.389	1.455
0.50	15.226	11.498	8.335	5.793	3.862	2.472	1.523
0.75	15.277	11.571	8.426	5.889	3.955	2.555	1.590
1.00	15.329	11.643	8.510	5.986	4.046	2.636	1.657

R: $\theta_r = 0.0953, r = 0.0953, \kappa_r = 4, \sigma_r = 0.1, v = 0.2, T = 0.5, S = 100$

ρ \ K	90	95	100	105	110	115	120
BS	14.189	10.494	7.414	4.997	3.215	1.977	1.165
–1.00	13.971	10.183	7.036	4.600	2.845	1.666	0.927
–0.75	14.028	10.265	7.137	4.706	2.943	1.748	0.989
–0.50	14.084	10.347	7.237	4.811	3.041	1.830	1.051
–0.25	14.141	10.428	7.334	4.913	3.136	1.910	1.113
0.00	14.198	10.508	7.430	5.014	3.231	1.990	1.175
0.25	14.255	10.586	7.524	5.113	3.324	2.070	1.238
0.50	14.312	10.664	7.617	5.210	3.415	2.149	1.300
0.75	14.369	10.742	7.708	5.306	3.506	2.227	1.362
1.00	14.426	10.818	7.798	5.400	3.595	2.304	1.425

S: $\theta_r = 0.05, r = 0.0953, \kappa_r = 4, \sigma_r = 0.1, v = 0.2, T = 0.5, S = 100$

ρ \ K	90	95	100	105	110	115	120
BS	15.416	11.617	8.381	5.779	3.808	2.400	1.451
–1.00	15.228	11.334	8.021	5.383	3.423	2.064	1.183
–0.75	15.276	11.409	8.117	5.489	3.526	2.153	1.253
–0.50	15.326	11.483	8.211	5.593	3.627	2.241	1.323
–0.25	15.375	11.557	8.304	5.695	3.726	2.329	1.393
0.00	15.425	11.629	8.396	5.795	3.824	2.415	1.462
0.25	15.475	11.702	8.486	5.894	3.920	2.500	1.532
0.50	15.525	11.773	8.575	5.991	4.015	2.585	1.601
0.75	15.575	11.845	8.652	6.086	4.108	2.668	1.670
1.00	15.625	11.915	8.748	6.179	4.200	2.751	1.739

T: $\theta_r = 0.11, r = 0.0953, \kappa_r = 4, \sigma_r = 0.1, v = 0.2, T = 0.5, S = 100$

BS values are calculated by setting the discounter factor to be zero-coupon bond price.

Table 6.2 The impact of correlation ρ on option prices in SI model. Option prices increase with increasing correlations ρ.

θ_r \ K	90	95	100	105	110	115	120
0.03	13.061	9.039	5.756	3.349	1.775	0.858	0.380
0.06	13.724	9.642	6.253	3.713	2.012	0.996	0.451
0.09	14.392	10.259	6.770	4.101	2.272	1.150	0.534
0.12	15.064	10.888	7.307	4.513	2.554	1.323	0.628

U: BS benchmark values are calculated according to θ_r.

θ_r \ K	90	95	100	105	110	115	120
0.03	13.155	9.197	5.962	3.564	1.961	0.994	0.466
0.06	13.809	9.796	6.453	3.930	2.206	1.142	0.547
0.09	14.468	10.398	6.964	4.318	2.472	1.307	0.640
0.12	15.132	11.016	7.493	4.729	2.760	1.489	0.744

V: $\rho = 0.5, \kappa_r = 4, \sigma_r = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$

θ_r \ K	90	95	100	105	110	115	120
0.03	13.070	9.055	5.777	3.371	1.794	0.872	0.388
0.06	13.733	9.657	6.273	3.735	2.032	1.010	0.461
0.09	14.400	10.273	6.790	4.123	2.292	1.166	0.544
0.12	15.070	10.901	7.326	4.535	2.575	1.339	0.640

W: $\rho = 0.0, \kappa_r = 4, \sigma_r = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$

θ_r \ K	90	95	100	105	110	115	120
0.03	12.987	8.910	5.584	3.169	1.622	0.749	0.314
0.06	13.657	9.522	6.086	3.532	1.852	0.878	0.377
0.09	14.334	10.148	6.609	3.920	2.105	1.024	0.452
0.12	15.013	10.786	7.153	4.333	2.382	1.187	0.538

X: $\rho = -0.5, \kappa_r = 4, \sigma_r = 0.1, v = 0.15, T = 0.5, S = 100, r = 0.0953$

Table 6.3 The impact of mean level θ_r on options prices in SI model. Option prices increase with increasing mean levels θ_r.

Chapter 7
Poisson Jumps

In this chapter we focus on Poisson jump models that are very popular in financial modeling since Merton (1976) first derived an option pricing formula based on a stock price process generated by a mixture of a Brownian motion and a Poisson process. This mixed process is also called the jump-diffusion process. The requirement for a jump component in a stock price process is intuitive, and supported by the big crashes in stock markets: The Black Monday on October 17, 1987 and the recent market crashes in the financial crisis since 2008 are two prominent examples. To model jump events, we need two quantities: jump frequency and jump size. The first one specifies how many times jumps happen in a given time period, and the second one determines how large a jump is if it occurs. In a compound jump process, jump arrivals are modeled by Poisson process and jump sizes may be specified by various distributions. Here, we consider three representative jump models that are distinguished from each other solely by the distributions of jump sizes, they are the simple deterministic jumps, the log-normal jumps and the Pareto jumps. Additionally, we will show that Kou's jump model (2002) with weighted double-exponential jumps is equivalent to the Pareto jump model. The CFs of all these jump models can be derived to value European-style options. Finally, we present the affine jump diffusion model of Duffie, Pan and Singleton (2000), which extends a simple jump-diffusion model to a generalized case. The solution of CFs in the affine jump-diffusion model has an exponential affine form, and may be derived via two ODEs.

7.1 Introduction

A Brownian motion is not the only way to specify a stock price process, at least not a complete way. Another fundamental continuous-time process in stochastic the-

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance, DOI 10.1007/978-3-642-01808-4_7,

ory, the Poisson process,[1] is perhaps a good alternative to describe some abnormal events in financial markets. Merton (1976) first derived an option pricing formula based on a stock price process generated by a mixture of a Brownian motion and a Poisson process. In his interpretation, the normal price changes are, for example, due to a temporary non-equilibrium between supply and demand, changes in capitalization rates, changes in the economic outlook, all of which have a marginal impact on prices. Therefore, this normal component is modeled by a Brownian motion. The "abnormal" component is produced by the irregular arrival of important new information specific to a firm, an industry or a country, and has a non-marginal effect on prices. This component is then modeled by a Poisson process. The best known abnormal event in finance history is the great crash of 1987. Empirical distributions of stock returns display obvious leptokurtosis, and the implied volatilities for options on the most widely used stock market indices perform remarkable "smile" pattern. It seems that adding a random jump to the stock price process is important in deriving more realistic option valuation formulas and in fitting the empirical leptokurtic distribution of stock returns. In this sense, the models with a jump component compete with stochastic volatility models in smile modeling. Jorion (1988) reported that 96% of the total exchange risk and 36% of the total stock risk are caused by the respective jump components. He also concluded that a jump-diffusion process outperforms a GARCH model, a discrete version of some stochastic volatility models, in describing the exchange rate process. Bates (1996) showed that in many cases it is sufficient to reduce the volatility smile by using a jump-diffusion model. Bakshi, Cao and Chen (1997) arrived at a similar result. All of these studies support the argument that jump processes are important in option pricing theory. We define the Poisson process formally.

Definition 7.1.1. *A Poisson process is an adapted counting process $Y(t)$ with the following properties:*

1. *For $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n < \infty$, $Y(t_1), Y(t_2) - Y(t_1), \cdots, Y(t_n) - Y(t_{n-1})$ are stochastically independent;*
2. *For every $t > s \geqslant 0$, $Y(t) - Y(s)$ is Poisson distributed with the parameter λ, i.e., $P(Y(t) - Y(s) = n) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^n}{n!}$;*
3. *For each fixed $\omega \in \Omega, Y(\omega, t)$ is continuous in t;*
4. *$Y(0) = 0$ almost surely.*

When compared with a Brownian motion, we can see that the only difference between both processes is the probability law: One is governed by the normal distribution which is suitable for the description of continuous events, and the other one is governed by the Poisson distribution which is good for counting discontinuous events. Both processes have the properties known as independent and stationary increments, and then belong to the class of Lévy process.[2]

The first two central moments of $Y(t)$ are identical and given as follows

[1] The Poisson process is a continuous-time but not continuous-path process due to its jump property.

[2] We will address Lévy process more in details in the next chapter.

$$\mathbf{E}[Y(t)] = Var[Y(t)] = \lambda t. \tag{7.1}$$

Consequently, both $Y(t) - \lambda t$ and $Y(t)^2 - \lambda t$ are martingales with an expectation of zero. The process $Y(t) - \lambda t$ is referred to as the compensated Poisson process in stochastic literature and plays a key role in constructing a risk-neutral jump-diffusion processes. Since the Poisson process is Markovian, we compute the covariance between $Y(t)$ and $Y(s), t > s$, as follows:

$$\begin{aligned} Cov[Y(t), Y(s)] &= \mathbf{E}[\{Y(t) - \lambda t\}\{Y(s) - \lambda s\}] \\ &= \mathbf{E}[Y(t)Y(s)] - \lambda^2 ts \\ &= \mathbf{E}[\{Y(t-s) + Y(s)\}Y(s)] - \lambda^2 ts \\ &= \mathbf{E}[Y(s)Y(s)] - \lambda^2 ts \\ &= \lambda s + \lambda^2 s^2 - \lambda^2 ts = \lambda s[1 - \lambda(t-s)]. \end{aligned} \tag{7.2}$$

An even more interesting property of the Poisson process is that the probability that a jump occurs only once in an infinitesimal time-increment is just the parameter λ times dt, that is

$$\lim_{t \to 0} \frac{1}{t} P(Y(t) = 1) = \lambda \qquad \text{or} \qquad P(dY(t) = 1) = \lambda dt. \tag{7.3}$$

At the same time, the probability that a jump occurs more than once in the time interval dt can be regarded as zero, i.e., $P(Y(t) \geqslant 2)$ is of the order $o(t)$. Because of these two special features, we can express the Poisson process in the continuous-time limit case as follows:

$$\text{Jump probability in } dt = dY(t) \tag{7.4}$$

with

$$dY(t) \sim \delta_1 \lambda dt + \delta_0 (1 - \lambda) dt,$$

where δ_n denotes the Dirac indicator function with $\delta_n = \mathbf{1}(dY(t) = n) = n$ for $n = 0, 1$. Hence, a jump process is strongly connected with a binomial distribution. Cox, Ross and Rubinstein (1979) even showed that a binomial model can converge to a jump process under special conditions.[3] Now by adding a jump size J that is not correlated with $Y(t)$, we are able describe a jump event not only with jump frequency characterized by the parameter λ, but also with jump size, and obtain

$$\text{Jumps in } dt = JdY(t)$$

or

---

[3] It is also known that the Black-Scholes formula can be derived in a binomial setup. As shown as in Cox, Ross and Rubinstein (1979), one can obtain the option pricing formulas for the Brownian process and the Poisson process respectively in a continuous-time limit by setting different up- and down probabilities.

$$\text{Jumps in } t = \int_0^t JdY(s) = \sum_{k=1}^{Y(t)} J_k, \tag{7.5}$$

where J_k is the jump size conditional on k-th jump event. Such a jump process $JdY(t)$ combining jump counting and jump size is call compound jump process.

Assuming that the compound Poisson process and the Brownian motion are mutually stochastically independent, we can apply these two components to describe stock price dynamics, and obtain the following mixed process,

$$\frac{dS(t)}{S(t)} = r(t)dt + v(t)dW_1(t) + JdY(t). \tag{7.6}$$

Due to the new jump component $JdY(t)$, the process $g_1(t)$ in (2.7), which is necessary for constructing the CFs, is no longer a martingale with an expected value of one. To validate the risk-neutral pricing approach, the stock price process (7.6) should be modified by using the martingale property of the compensated Poisson process:

$$\frac{dS(t)}{S(t)} = [r(t) - \lambda \mathbf{E}[J]]\, dt + v(t)dW_1(t) + JdY(t). \tag{7.7}$$

Due to the drift compensator $\lambda \mathbf{E}[J]$, one can easily verify that the required martingale property of $g_j(t), j = 1, 2$, in (2.7) and (2.9) in the above jump-diffusion process is also satisfied. In order to value options via Fourier transform in a jump-diffusion process, we want to calculate the corresponding CFs under two the relevant measures Q_1 and Q_2. Generally these CFs may be solved, as in a pure diffusion setting, by PDE. For example, we take again the Heston model for illustration where $V(t) = v^2(t)$ follows a mean-reverting square root process. Extending the PDEs given in (2.52) and (2.53) in Chapter 2 with jump components yields the following two PDEs, respectively,

$$\begin{aligned}\frac{\partial f_1}{\partial T} &= \frac{1}{2}V\frac{\partial^2 f_1}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_1}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_1}{\partial V^2} + \left(r + \frac{1}{2}V\right)\frac{\partial f_1}{\partial x} \\ &\quad + [\kappa(\theta - V) + \rho\sigma V]\frac{\partial f_1}{\partial V} + \lambda \int_{\mathbb{R}} [f_1(x+q) - f_1(x)]g(q)dq\end{aligned} \tag{7.8}$$

and

$$\begin{aligned}\frac{\partial f_2}{\partial T} &= \frac{1}{2}V\frac{\partial^2 f_2}{\partial x^2} + \rho\sigma V\frac{\partial^2 f_2}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 f_2}{\partial V^2} + \left(r + \frac{1}{2}V\right)\frac{\partial f_2}{\partial x} \\ &\quad + \kappa(\theta - V)\frac{\partial f_2}{\partial V} + \lambda \int_{\mathbb{R}} [f_2(x+q) - f_2(x)]g(q)dq,\end{aligned} \tag{7.9}$$

where

$$q = \ln(J+1)$$

and $g(q)$ is the probability density function of jump log-size q. The boundary conditions are given by

$$f_j(\phi) = e^{i\phi x_0}, \quad j = 1.2.$$

Due to the term $\lambda \int_{\mathbb{R}}[f_2(x+q) - f_2(x)]g(q)dq$, the equations (7.8) and (7.9) are called a partial integro-differential equation (PIDE). For some special specifications of J or q, closed-form solutions are available and, in particular take the following exponential affine form,

$$f_j(\phi) = e^{A(T)x_0 + B(T)V_0 + C(T)\lambda + D(T)}, \tag{7.10}$$

where $B(T)$ and $C(T)$ capture the effects of stochastic variances and random jumps respectively. As discussed later, the above simple jump-diffusion model can be extended to a general affine jump-diffusion model as in Duffie, Pan and Singleton (2000). In this chapter, we do not intend to search for the solutions for CFs $f_j(\phi)$ directly via the PIDEs (7.8) and (7.9). Instead, we derive the solutions for the CFs $f_j(\phi)$ via the techniques of stochastic calculus.

In spite of this risk-neutral process, there is a consensus that the risks associated to jumps cannot be hedged away in the Black-Scholes's sense. In other words, one can not form a portfolio protecting against *any* price change at *any* time. However, the martingale property implies that trading using the Black-Scholes's hedge is a "fair game" over a long time in an expectation sense even when jumps happen.[4] *"If an investor follows a Black-Scholes hedge where he is long the stock and short the option,* $\cdots\cdots$*, the large losses occur just frequently enough so as to, on average, offset the almost steady "excess" return"* (Merton, 1976). Additionally, there are some assumptions imposed on the option pricing model, in order to overcome the hedge problem in connection with jump risks. For example, suppose that the capital asset pricing model (CAPM) holds for asset returns, and that jumps occurring in the stock prices are completely firm-specific,[5] then the jump component is uncorrelated with market movements and represents unsystematic risks.

In the following sections we study three cases. The first one termed as simple jumps is a case where jump size J is constant, the second one termed as lognormal jumps means that the jump size J is lognormally distributed and independent of $Y(t)$ and $W_1(t)$. Finally, we consider a case where the jump size is governed by a Pareto distribution, and show that the jump model equipped with two independent Pareto jumps agrees with the model of Kou (2002), known as double-exponential jump model. We will incorporate all three different types of jumps into the option valuation theory by using CFs.

---

[4] In their seminal work, Cox and Ross (1976) argued that risk-neutral pricing is valid for the jump process without any additional restriction and a riskless portfolio can be found by the Black-Scholes trading strategy. We think that their arguments could be understood only in the here mentioned long run sense.

[5] See also the interpretation of jumps at the beginning of this section.

7.2 Simple Jumps

In this subsection we at first briefly review a simple jump model that is studied by Cox, Ross and Rubinstein (1979) and is an extension of the work of Cox and Ross (1976). Assume that the stock price follows a process of the following form

$$\frac{dS(t)}{S(t)} = qdt + JdY(t), \tag{7.11}$$

which is formally equivalent to the following description

$$\begin{cases} dS(t) = S(t)(u-1)dt & \text{if jumps occur in } dt \\ \\ dS(t) = S(t)qdt & \text{if jumps do not occur in } dt \end{cases}.$$

Most of the time, stock prices grow with rate q; Occasionally, jumps come with a intensity rate λ and a jump size J. Taking the growth rate q into account, we have $q + J\lambda = (u-1)\lambda$. The necessary and sufficient condition for a risk-neutral process of (7.11) is $q + J\lambda = r$, which leads to $(u-1)\lambda = r$.[6]

In the continuous-time case, jumps occur according to a Poisson process with intensity λ, and stock prices are log-Poisson distributed. For $q < 0$ and $J > 0$,[7] the stock price distribution generated by these simple jumps has a left tail which is thinner than (a right tail fatter than) the counterpart of a corresponding lognormal one. With this setup, a European call option price may be given by

$$C = S_0 f(a;b) - Ke^{-rT}\Psi(a;b/u) \tag{7.12}$$

with

$$a = \left\{ \frac{\ln(K/S_0) + qT}{\ln u} \right\}^+, \qquad b = \frac{(r+q)uT}{u-1}.$$

The symbol $\{x\}^+$ denotes the smallest nonnegative integer that is greater than x. The function $f(z;y)$ takes the following form

$$\Psi(z;y) = \sum_{k=z}^{\infty} \frac{e^{-y}y^k}{k!}.$$

By setting $q = 0$ we obtain the identical formula for options with a so-called birth process (Cox and Ross,1976) that describes two states only: jump or nothing. Since the jump process is closely connected with binomial distribution in continuous-time limit, formula (7.12) can be derived in a binomial setup (Cox, Ross and Rubinstein, 1979).

---

[6] This condition is identical to that in Cox and Ross (1976), which allows them to eliminate λ from the option pricing formula.

[7] The original assumption of Cox, Ross and Rubinstein (1979) is $q < 0$ and $J > 0$, which is unrealistic in that jumps can only be positive. Also see Hull (1997).

In the light of the jump-diffusion process given by (7.7), we introduce the dynamics of stock prices with pure jumps as follows:

$$\frac{dS(t)}{S(t)} = [r(t) - \lambda J]\,dt + v(t)dW_1(t) + JdY(t). \tag{7.13}$$

Since here we only need to show what terms the jumps contribute to the CFs, we let the interest rates and volatilities to be constant. By using Itô's lemma for Poisson process,[8] we obtain the process for $x(t) = \ln S(t)$,

$$\begin{aligned}
dx(t) &= \left(r - \frac{1}{2}v^2 - \lambda J\right) x_S S dt + v dW_1 \\
&\quad + \mathbf{E}_x\left[dY(t)\{\ln(S(t)(1+J)) - \ln S(t)\}\right] \\
&= \left(r - \frac{1}{2}v^2 - \lambda J\right) dt + v dW_1 + dY(t)\mathbf{E}_x\left[\ln(1+J)\right] \\
&= \left(r - \frac{1}{2}v^2 - \lambda J\right) dt + v dW_1 + \ln(1+J)dY(t),
\end{aligned} \tag{7.14}$$

where $\mathbf{E}_x$ stands for the expectation operator working only with the probability law of x. Recalling the principle for constructing CFs in Section 2.1, we find that this principle is essentially independent of the specification of the underlying process. The important point in constructing a CF is that the discounted process $\left(\exp(-\int_0^t r(u)du)S(t)/S_0\right)_{t>0}$ must be a martingale, which is obviously satisfied by (7.13).

Now we start calculating the CFs for the Poisson process as follows:[9]

$$\begin{aligned}
f_1(\phi) &= \mathbf{E}\left[\frac{S(T)}{S_0 e^{rT}}\exp\left(i\phi x(T)\right)\right] \\
&= \mathbf{E}\left[\exp(-rT - x_0 + (1+i\phi)x(T))\right] \\
&= \mathbf{E}\left[\exp(-rT + i\phi x_0 + (1+i\phi)\left(r - \frac{1}{2}v^2 - \lambda J\right)T\right. \\
&\quad \left. + \ (1+i\phi)v dW_1 + (1+i\phi)\ln(1+J)\int_0^T dY(t)\right] \\
&= \exp\left(i\phi(x_0 + rT) - (1+i\phi)\left(\frac{1}{2}v^2 + \lambda J\right)T + \frac{1}{2}(1+i\phi)^2 v^2 T\right. \\
&\quad + \ \lambda T \exp[(1+i\phi)\ln(1+J)] - \lambda T) \\
&= \exp\left(i\phi(x_0 + rT)\right) f_1^{BS}(\phi) f_1^{JSimple}(\phi)
\end{aligned} \tag{7.15}$$

[8] For more about this topic see Merton (1990), Malliaris and Brock (1991).

[9] Here we use the well-known result on the CF of a Poisson process $y(t)$, that is

$$f(\phi; ky) = \mathbf{E}\left[e^{i\phi ky}\right] = \exp\left[\lambda t(e^{i\phi k} - 1)\right],$$

where k is an arbitrary real number.

with

$$f_1^{JSimple}(\phi) = \exp\big(-\frac{1}{2}(1+i\phi)\lambda JT + \lambda T \exp[(1+i\phi)\ln(1+J)] - \lambda T\big). \quad (7.16)$$

In a similar fashion, we obtain

$$\begin{aligned}
f_2(\phi) &= \mathbf{E}\left[\exp\left(i\phi x(T)\right)\right] \\
&= \mathbf{E}\left[\exp(i\phi rT + i\phi x_0 - i\phi\left(\frac{1}{2}v^2 + \lambda J\right)T + i\phi\ln(1+J)\int_0^T dY(t)\right] \\
&= \exp\left(i\phi rT + i\phi x_0 - i\phi\left(\frac{1}{2}v^2 + \lambda J\right)T + \frac{1}{2}(1+i\phi)^2 v^2 T\right. \\
&\quad \left. + \lambda T \exp[i\phi\ln(1+J)] - \lambda T\right) \\
&= \exp\left(i\phi(x_0 + rT)\right) f_2^{BS}(\phi) f_2^{JSimple}(\phi) \qquad (7.17)
\end{aligned}$$

with

$$f_2^{JSimple}(\phi) = \exp\big(-\frac{1}{2}i\phi\lambda JT + \lambda T \exp[i\phi\ln(1+J)] - \lambda T\big). \quad (7.18)$$

The most appealing attribute to $f_j(\phi)$ is that they are continuous functions of ϕ although the Poisson distribution is not continuous. The closed-form pricing formula for the European call options is then given in the standard way where the probabilities F_j are expressed by the Fourier inversion. If we omit the diffusion parts from these two CFs, that is, $v = 0$ and $r - \lambda J = q$, we obtain the same option pricing formula as given in (7.12). Obviously, this new formula is more easily computable and unified with other pricing formulas in a same Fourier transform framework. Due to the deterministic specification of jump size J, the simple jump model presented above finds only a limited application in practice, and does not have large flexibility to capture smile features of implied volatilities. In the following two subsections, we discuss two more complicated and realistic jump models.

7.3 Lognormal Jumps

In the line with Merton (1976), we go back to the process (7.7), the so-called jump-diffusion model, and further assume that jump size J is no longer deterministic but lognormally distributed. The Brownian motion $w_1(t)$, the Poisson term $Y(t)$ and the jump size J are mutually stochastically independent. Formally, we have the following specification,[10]

$$\ln(1+J) \sim N\big[\ln(1+\mu_J) - \frac{1}{2}\sigma_J^2, \sigma_J^2\big], \qquad \mu_J \geqslant -1. \quad (7.19)$$

[10] Compare with Scott (1997) and BCC (1997).

This means that J has a mean of μ_J and $\ln(1+J)$ has a standard deviation of σ_J. Because of the normal distribution of the random variable $\ln(1+J)$, this jump itself is log normally distributed, and is referred to as lognormal jump. Lognormal jump model is the most popular jump model in financial literature, and is widely used to examine the empirical properties of stock prices and the smile features of implied volatilities.

We obtain immediately

$$dx(t) = [r(t) - \lambda\mu_J - \frac{1}{2}v^2(t)]dt + v(t)dW_1 + \ln(1+J)dY(t), \tag{7.20}$$

where it is clear that the abnormal stock returns governed by the term $\ln(1+J)dY(t)$ have a normally distributed jump size. Hence, the abnormal stock returns could be both positive and negative, and are modeled more realistically than in a simple jump model. Merton (1976) gave a pricing formula in terms of an infinite sum,

$$C = \sum_{n=0}^{\infty} \frac{\exp[-\lambda T(1+\mu_J)](\lambda T(1+\mu_J))^n}{n!} C_n, \tag{7.21}$$

with C_n as the Black-Scholes price with an instantaneous variance $v^2 + \frac{n}{T}\sigma_J^2$ and the risk-free rate $r - \lambda\mu_J + \frac{n}{T}\ln(1+\mu_J)$. This formula is obviously not a closed-form solution since it includes an infinite number of Black-Scholes formulas. Again, we can express formula (7.21) in an alternative way with CF, and for convenience, let the interest rate and volatility be constant for the time being.

We perform the calculation of the first CF as follows:

$$\begin{aligned}
f_1(\phi) &= \mathbf{E}\left[\frac{S(T)}{S_0 e^{rT}} \exp\left(i\phi x(T)\right)\right] \\
&= \mathbf{E}\Bigg[\exp\Bigg(i\phi(rT + x_0) - (1+i\phi)\lambda T\mu_J - (1+i\phi)\frac{1}{2}v^2 T \\
&+ \frac{1}{2}(1+i\phi)^2 v^2 T + (1+i\phi)\ln(1+J)\int_0^T dY(t)\Bigg)\Bigg] \\
&= \mathbf{E}\left[\exp\left(\lambda T \exp((1+i\phi)\ln(1+J)) - \lambda T\right)\right] \\
&\times \exp\left(i\phi(rT + x_0) - (1+i\phi)\lambda T\mu_J - (1+i\phi)\frac{1}{2}v^2 T + \frac{1}{2}(1+i\phi)^2 v^2 T\right) \\
&= \exp\Bigg(i\phi(rT + x_0) - (1+i\phi)\lambda T\mu_J - (1+i\phi)\frac{1}{2}v^2 T \\
&+ \frac{1}{2}(1+i\phi)^2 v^2 T + \lambda T \cdot \mathbf{E}\left[\exp((1+i\phi)\ln(1+J))\right] - \lambda T\Bigg)
\end{aligned}$$

$$= \exp\Bigg(i\phi(rT + x_0) - (1+i\phi)\lambda T\mu_J + \frac{1}{2}i\phi(1+i\phi)v^2T$$
$$+\ \lambda T[(1+\mu_J)^{(1+i\phi)}e^{\frac{1}{2}i\phi(1+i\phi)\sigma_J^2} - 1]\Bigg),$$
$$= (i\phi(x_0 + rT))\, f_1^{BS}(\phi) f_1^{LogN}(\phi) \tag{7.22}$$

with

$$f_1^{LogN}(\phi) = \exp\big(-(1+i\phi)\lambda T\mu_J + \lambda T[(1+\mu_J)^{(1+i\phi)}e^{\frac{1}{2}i\phi(1+i\phi)\sigma_J^2} - 1]\big), \tag{7.23}$$

where, since J and $dY(t)$ are independent of each other,[11] we can set the expectation operator directly in the front of $\exp[(1+i\phi)\ln(1+J)]$.

In a similar manner, we may calculate

$$f_2(\phi) = \mathbf{E}[\exp(i\phi x(T))]$$
$$= \exp\Bigg(i\phi(rT + \ln S_0) - i\phi\lambda T\mu_J + \frac{1}{2}i\phi(1+i\phi)v^2T$$
$$+\ \lambda T[(1+\mu_J)^{i\phi}e^{\frac{1}{2}i\phi(i\phi-1)\sigma_J^2} - 1]\Bigg)$$
$$= (i\phi(x_0 + rT))\, f_2^{BS}(\phi) f_2^{LogN}(\phi) \tag{7.24}$$

with

$$f_2^{LogN}(\phi) = \exp\big(-i\phi\lambda T\mu_J + \lambda T[(1+\mu_J)^{i\phi}e^{\frac{1}{2}i\phi(i\phi-1)\sigma_J^2} - 1]\big). \tag{7.25}$$

The closed-form pricing formula is then given in the usual way by applying the the Fourier inversion. Additionally, we can easily do comparative statics with respect to the jump parameters λ and μ_J while it is rather cumbersome to do this with Merton's formula. If σ_J is approaching zero, then the jump size will converge to the deterministic one and the lognormal jump becomes a simple jump which is exactly the case addressed above.

[11] The detailed calculation is as follows:

$$\mathbf{E}[\exp((1+i\phi)\ln(1+J))]$$
$$= \exp\left((1+i\phi)\ln(1+\mu_J) - \frac{1}{2}(1+i\phi)\sigma_J^2 + \frac{1}{2}(1+i\phi)^2\sigma_J^2\right)$$
$$= (1+\mu_J)^{(1+i\phi)}\exp\left(\frac{1}{2}i\phi(1+i\phi)\sigma_J^2\right).$$

7.4 Pareto Jumps

An other way to enhance the variants of jumps is to specify the jump size by an alternative distribution. In contrast to the above section, the logarithm of the random jump size may be suggested to be exponentially distributed with a mean of μ as follows:

$$\ln(1+J) \sim Ex[\mu], \tag{7.26}$$

where the probability density function (pdf) of an exponential distribution takes the following form

$$pdf_{Ex}(y;\mu) = \frac{1}{\mu}\exp\left(-\frac{y}{\mu}\right), \qquad 0 < y < \infty, \tag{7.27}$$

that is defined only in positive real line.

Duffie, Pan and Singleton (1999) applied this type of distribution to model the jumps occurring in a volatility process. The mean and standard deviation of the exponentially distributed variable are identical and are given by

$$\mathbf{E}[y] = \mu, \qquad Std[y] = \mu. \tag{7.28}$$

Therefore, an exponential distribution is entirely characterized by a single parameter μ. We can immediately derive from (7.27) that J has a probability density function of the following form

$$\begin{aligned} pdf_J(y) &= \frac{1}{\mu(1+y)}\exp\left(-\frac{\ln(1+y)}{\mu}\right) \\ &= \frac{1}{\mu}(1+y)^{-\frac{1}{\mu}-1}, \qquad 0 < y < \infty, \end{aligned} \tag{7.29}$$

which means that the jump size J is Pareto distributed.[12] Hence, we call this a Pareto jump. The interchangeability between an exponential distribution and its Pareto counterpart is similar to the interchangeability between a normal distribution and its lognormal counterpart. One of the important implications of the Pareto distribution and the associated exponential distribution is that the random jump size J and its logarithm $\ln(1+J)$ can only be positive and have positive means. Given $0 < \mu < 1$, the mean of J exists and is equal to

$$\mathbf{E}[J] = \frac{\mu}{1-\mu}. \tag{7.30}$$

Therefore, if we specify the abnormalities in the dynamic movements of stock prices using a single positive Pareto jump, the jumps can only be positive. To incorporate possible negative jumps into stock price process, we may introduce two Parato jumps with different expected jump sizes and jump intensities, one is re-

[12] For more details see Stuard and Ord (1994), or Johnson, Kotz and Balakrishnan (1994).

sponsible for positive surprises, the other for negative surprises. More precisely, we specify the following dynamics for stock returns:

$$
\begin{aligned}
dx(t) &= [r-\mu-\frac{1}{2}v^2]dt + vdW_1 \\
&+ \ln(1+J_1)dY_1(t) - \ln(1+J_2)dY_2(t), \\
&= [r-\mu-\frac{1}{2}v^2]dt + vdW_1 \\
&+ y_1 dY_1(t) - y_2 dY_2(t),
\end{aligned} \tag{7.31}
$$

with

$$
\begin{aligned}
\mu &= \lambda_1\mathbf{E}[J_1] - \lambda_1\mathbf{E}[J_2] = \frac{\lambda_1\mu_{J_1}}{1-\mu_{J_1}} - \frac{\lambda_2\mu_{J_2}}{1-\mu_{J_2}}, \\
\ln(1+J_1) &\sim \frac{1}{\mu_{J_1}}\exp\left(-\frac{\ln(1+J_1)}{\mu_{J_1}}\right), \qquad 0 < J_1 < \infty, \\
\ln(1+J_2) &\sim \frac{1}{\mu_{J_2}}\exp\left(-\frac{\ln(1+J_2)}{\mu_{J_2}}\right), \qquad 0 < J_2 < \infty.
\end{aligned}
$$

The jump $Y_1(t)$ with jump intensity λ_1 characterizes the positive abnormal movements in stock prices while $Y_2(t)$ with jump intensity λ_2 characterizes the negative abnormal movements. Their joint impact on the drift of the risk-neutral process is then $\mu = \lambda_1\mathbf{E}[J_1] - \lambda_2\mathbf{E}[J_2]$. Together we use 4 parameters to specify two jumps.

Note, if y is exponentially distributed with a mean μ, for any real or complex number k, the moment-generating function of y is given by

$$
\mathbf{E}[e^{ky}] = \frac{1}{1-k\mu}.
$$

Based on the risk-neutralized process, applying the moment-generating function of y, we can calculate the two CFs that are relevant for option pricing as follows:

$$
\begin{aligned}
& f_1(\phi) \\
&= \mathbf{E}\left[\frac{S(T)}{S_0 e^{rT}}\exp(i\phi x(T))\right] \\
&= \mathbf{E}\Bigg[\exp\Bigg(i\phi(rT+x_0) - (1+i\phi)\mu T - (1+i\phi)\frac{1}{2}v^2 T + \frac{1}{2}(1+i\phi)^2 v^2 T \\
&+ \ (1+i\phi)\ln(1+J_1)\int_0^T dY_1(t) - (1+i\phi)\ln(1+J_2)\int_0^T dY_2(t)\Bigg)\Bigg] \\
&= \exp\left(i\phi(rT+x_0) - (1+i\phi)\mu T - (1+i\phi)\frac{1}{2}v^2 T + \frac{1}{2}(1+i\phi)^2 v^2 T\right) \\
&\times \mathbf{E}[\exp[(\lambda_1 T\exp((1+i\phi)\ln(1+J_1)) - \lambda_1 T) \\
&- (\lambda_2 T\exp((1+i\phi)\ln(1+J_2)) - \lambda_2 T)]].
\end{aligned}
$$

Computing the expected value under exponential distribution yields

$$\begin{aligned} f_1(\phi) &= \exp\bigg(i\phi(rT+x_0)-(1+i\phi)\mu T-(1+i\phi)\frac{1}{2}v^2T+\frac{1}{2}(1+i\phi)^2v^2T \\ &+\ \lambda_1 T\cdot\mathbf{E}\left[\exp((1+i\phi)\ln(1+J_1))\right]-\lambda_1 T \\ &-\ \lambda_2 T\cdot\mathbf{E}\left[\exp((1+i\phi)\ln(1+J_2))\right]-\lambda_2 T) \\ &= \exp\bigg(i\phi(rT+x_0)-(1+i\phi)\mu T+\frac{1}{2}i\phi(1+i\phi)v^2T \\ &+\ \lambda_1 T[(1-(1+i\phi)\mu_{J_1})^{-1}-1]-\lambda_2 T[(1-(1+i\phi)\mu_{J_2})^{-1}-1]\big) \\ &= (i\phi(x_0+rT))\, f_1^{BS}(\phi) f_1^{Pareto}(\phi) \end{aligned} \tag{7.32}$$

with

$$\begin{aligned} f_1^{Pareto}(\phi) &= \exp\big(-(1+i\phi)\mu T+\lambda_1 T[(1-(1+i\phi)\mu_{J_1})^{-1}-1] \\ &-\ \lambda_2 T[(1-(1+i\phi)\mu_{J_2})^{-1}-1]\big). \end{aligned} \tag{7.33}$$

We have the similar result for the second CF,

$$\begin{aligned} f_2(\phi) &= \mathbf{E}\left[\exp\left(i\phi x(T)\right)\right] \\ &= \exp\bigg(i\phi(rT+x_0)-i\phi\mu T+\frac{1}{2}i\phi(1+i\phi)v^2T \\ &+\ \lambda_1 T[(1-i\phi\mu_{J_1})^{-1}-1]-\lambda_2 T[(1-i\phi\mu_{J_2})^{-1}-1]\big) \\ &= (i\phi(x_0+rT))\, f_2^{BS}(\phi) f_2^{Pareto}(\phi) \end{aligned} \tag{7.34}$$

with

$$\begin{aligned} f_2^{Pareto}(\phi) &= \exp\big(-i\phi\mu T+\lambda_1 T[(1-i\phi\mu_{J_1})^{-1}-1] \\ &-\ \lambda_2 T[(1-i\phi\mu_{J_2})^{-1}-1]\big). \end{aligned} \tag{7.35}$$

Given these two CFs, the option pricing formula with Pareto jumps may be correspondingly derived.

7.5 The Kou Model: An Equivalence to Pareto Jumps

Kou (2002) proposed a jump-diffusion model with jumps that may be regarded as two weighted Pareto jumps. Under the risk-neutral measure, Kou's model reads

$$\begin{aligned} \frac{dS(t)}{S(t)} &= (r-\lambda\mathbf{E}[U-1])dt+vdW(t)+d\Big(\sum_{i=1}^{Y(t)}(U_i-1)\Big) \\ &= (r-\lambda\mathbf{E}[J])dt+vdW(t)+JdY(t), \end{aligned} \tag{7.36}$$

where it is assumed that $u = \ln U$ has an asymmetric double-exponential distribution with a density function

$$pdf_u(z) = p\eta_1 e^{-\eta_1 z}\mathbf{1}_{(z\geqslant 0)} + q\eta_2 e^{\eta_2 z}\mathbf{1}_{(z<0)},$$

with $\eta_1 > 1$ and $\eta_2 > 0$. The parameters p and q with $p+q=1$ are considered as the probabilities for positive and negative abnormalities, that correspond to the positive and negative abnormalities respectively in the above discussed Pareto jump model. In other words, we have a weighted jump,

$$u = pu^+ + q(-u^-) = pu^+ - qu^-,$$

with

$$\mathbf{E}[u] = p\frac{1}{\eta_1} - q\frac{1}{\eta_2},$$

$$\mathbf{E}[U] = p\mathbf{E}[e^{u^+}] + q\mathbf{E}[e^{-u^-}] = p\frac{\eta_1}{\eta_1 - 1} + q\frac{\eta_2}{\eta_2 + 1},$$

where u^+ and u^- are two exponential random variables with means $1/\eta_1$ and $1/\eta_2$ respectively, and are weighted according to the Bernoulli law.

To show that the Kou model is equivalent to the Pareto jump model, we denote

$$\mu_1 = \frac{1}{\eta_1}, \quad \mu_2 = \frac{1}{\eta_2}, \quad z^- = -z\mathbf{1}_{(z<0)},$$

then $z^- > 0$ is exponentially distributed with mean $\mu_2 = \frac{1}{\eta_2}$. With these notations, the expected value of U can be rewritten as

$$\mathbf{E}[U] = p\frac{1}{1-\mu_1} + q\frac{1}{1+\mu_2}.$$

It follows immediately,

$$pdf_u(z) = p\frac{1}{\mu_1}\exp\Big(-\frac{z}{\mu_1}\Big)\mathbf{1}_{(z\geqslant 0)} + q\frac{1}{\mu_2}\exp\Big(-\frac{z^-}{\mu_1}\Big)\mathbf{1}_{(z^->0)}. \tag{7.37}$$

Therefore, the stock price process $S(t)$ can be restated as

$$\begin{aligned}\frac{dS(t)}{S(t)} &= (r - \lambda\mathbf{E}[J])dt + vdW(t)\\ &+ pJ^+dY(t) + q(-J^-)dY(t),\end{aligned} \tag{7.38}$$

with $J^+ + 1 = e^{u^+}$ and $J^- + 1 = e^{-u^-}$. This implies that $u^+ = \ln(J^+ + 1)$ and $u^- = -\ln(J^- + 1)$ are exponentially distributed. The process for stock return $x(t)$ takes the following form,

$$
\begin{aligned}
dx(t) &= \left(r - \lambda \mathbf{E}[J] - \frac{1}{2}v^2\right)dt + vdW(t) \\
&\quad + p\ln(J^+ + 1)dY(t) - q\ln(J^- + 1)dY(t) \\
&= \left(r - \lambda \mathbf{E}[J] - \frac{1}{2}v^2\right)dt + vdW(t) \\
&\quad + \ln(J^+ + 1)dY_1(t) - \ln(J^- + 1)dY_2(t),
\end{aligned} \tag{7.39}
$$

where two new Poisson processes are defined by

$$
dY_1(t) = pdY(t), \quad dY_2(t) = qdY(t)
$$

with two arrival intensity parameters $\lambda_1 = p\lambda$ and $\lambda_2 = q\lambda$. By a comparison with the processes (7.31) and (7.39), the new formulation of Kou's jump model is obviously equivalent to the Pareto jump model discussed above. Finally we can verify

$$
\begin{aligned}
\mathbf{E}[U] &= p\mathbf{E}[e^{u^+}] + q\mathbf{E}[e^{-u^-}] = p\mathbf{E}[1 + J^+] + q\mathbf{E}[1 + J^-] \\
&= p\mathbf{E}[J^+] + q\mathbf{E}[J^-] + 1 = \mathbf{E}[J] + 1,
\end{aligned} \tag{7.40}
$$

from which it follows

$$
\mathbf{E}[J] = \mathbf{E}[U] - 1 = p\frac{\mu_1}{1 - \mu_1} + q\frac{\mu_2}{1 + \mu_2}.
$$

To emphasis that a negative exponentially distributed random variable is incorporated into this jump model, we set $\mu_2^* = -\mu_2$, therefore we have

$$
\mu = \lambda \mathbf{E}[J] = \frac{\lambda_1 \mu_1}{1 - \mu_1} - \frac{\lambda_2 \mu_2^*}{1 - \mu_2^*}, \tag{7.41}
$$

that is identical to the jump correction term in the Pareto jump model. The Kou model can be re-formulated in the following log-normal form,

$$
\begin{aligned}
dx(t) &= \left(r - \mu - \frac{1}{2}v^2\right)dt + vdW(t) \\
&\quad + \ln(J^+ + 1)dY_1(t) - \ln(J^- + 1)dY_2(t) \\
&= \left(r - \mu - \frac{1}{2}v^2\right)dt + vdW(t) \\
&\quad + u^+ dY_1(t) - u^- dY_2(t),
\end{aligned} \tag{7.42}
$$

that is mathematically identical to the process given by (7.31) in the Pareto jump model.

In spite of the clear similarity of the Pareto jump model and the Kou model, the interpretations of these two jump models are different and subtle. While the Pareto jump model is equipped with positive and negative jumps, both of which can happen simultaneously, and overlap each other, the positive and negative jumps in Kou's model are weighted according to Bernoulli's law, and can not occur simultaneously. Although we may define two Poisson processes $Y_1(t)$ and $Y_2(t)$ in the Kou model, but they are driven by a same randomness. As a result, conditional on a jump event occurring, either a positive or a negative surprise can happen in the Kou model.

Summing up, by setting

$$\mu_1 = \frac{1}{\eta_1}, \quad \mu_2^* = -\frac{1}{\eta_2}, \quad \lambda_1 = p\lambda, \quad \lambda_2 = q\lambda, \tag{7.43}$$

we have established an equivalence between the Pareto jump model and Kou's model. Kou's original option pricing formula is rather complicated and is expressed in terms of a sum of Hh functions, a special function of mathematical physics. The numerical computation of European-style option prices is then very extensive. In contrast, as an equivalence of Kou's model to the Pareto jumps is established, a pricing formula for Kou's model similar to (7.33) and (7.35) is available, and shares the same computation method as other jump models. The application of the Fourier transform in Kou's model then gains many advantages in terms of efficiency, tractability and simplicity.

7.6 Affine Jump-Diffusions

In this chapter, we have actually discussed jumps in a jump-diffusion model where the underlying stock prices $S(t)$ are driven jointly by a diffusion, i.e., a Brownian motion $W(t)$, and a jump component $JdY(t)$. In fact, the jump-diffusion as an enhanced setting may be applied to model any financial quantity. Duffie, Pan and Singleton (DPS, 2000) used affine structures of a multi-dimensional jump-diffusion $X(t)$ as background-driving factors to model the dynamics of all pricing-relevant rates, such as stock returns, interest rates, volatilities, and foreign exchanges. As in Bakshi and Madan (1999), Duffie, Pan and Singleton considered a discounted CF (see Section 2.2.2) as follows

$$f^*(\phi, T-t) = \mathbf{E}\Big[\exp\big(-\int_t^T r(X(s),s)ds\big)e^{i\phi X(T)}|\mathscr{F}_t\Big], \tag{7.44}$$

where interest rate $r(X(t),t)$ is affine in $X(t)$. By setting $\phi = 0$, we obtain the expression for zero-coupon bond. An important insight of DPS for affine jump-diffusion is that the discounted CF $f^*(\phi)$ under some technical regularity conditions admits a solution of exponential affine form,

$$f^*(\phi, T-t) = e^{\alpha(t)+\beta(t)X(t)}, \tag{7.45}$$

where $\alpha(t)$ and $\beta(t)$ are determined by two ODEs, and have the closed-form solutions for many particular processes.

In the line of the jump-diffusion model discussed above, we propose that $X(t) \in \mathbb{R}^n$ is governed by the following n-dimensional process under the risk-neutral measure,

$$dX(t) = \mu(X)dt + v(X)dW(t) + qdY(t), \tag{7.46}$$

where the interest rate $r(X) \in \mathbb{R}$, the drift $\mu \in \mathbb{R}^n$, the variance $v(X)v(X)^\top \in \mathbb{R}^{n\times n}$ and the intensity $\lambda(X) \in \mathbb{R}^n$ all take an affine functional form of X. In details,

- $\mu(x) = A_0 + A_1 \cdot x, \quad A = (A_0, A_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$.
- $(v(x)v(x)^\top)_{ij} = (B_0)_{ij} + (B_1)_{ij} \cdot x, \quad B = (B_0, B_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}$.
- $\lambda(x) = c_0 + C_1 \cdot x, \quad C = (c_0, C_1) \in \mathbb{R} \times \mathbb{R}^n$.
- $r(x) = d_0 + D_1 \cdot x, \quad D = (d_0, D_1) \in \mathbb{R} \times \mathbb{R}^n$.

Therefore, based on the factor process $X(t)$, the coefficients (A,B,C,D), called the characteristics of affine jump-diffusion, completely determine all dynamics of the underlying rates that are relevant for pricing. Because the drift of $X(t)$ is specified explicitly, $r(X)$ is only responsible for discounting. Additionally, the counting mechanism for jumps of $X(t)$ has an stochastic intensity $\lambda(x)$.[13]

Now we denote $f_Y(z;q)$ as an "extended" Fourier transform or a moment-generating function of the jump component $qdY(t)$,[14]

$$\begin{aligned} f_Y(z;q) &= \mathbf{E}\left[e^{zqY(1)}\right] = \exp\left[\lambda(\mathbf{E}[e^{zq}] - 1)\right] \\ &= \exp\left[\lambda f_q(z) - 1)\right] \end{aligned} \tag{7.47}$$

where $f_q(z)$ may be again regarded as an "extended" Fourier transform of q. With the special affine structure characterized by (A,B,C,D), the functions $\alpha(t)$ and $\beta(t)$ of the exponential affine solution of the discounted CF $f^*(\phi,t)$ satisfy two complex-valued ODEs,

$$\beta'(t) = D_1 - A_1\beta(t) - \frac{1}{2}B_1 v^2(t) - C_1[f_q(\beta(t)) - 1], \tag{7.48}$$

$$\alpha'(t) = d_0 - A_0\beta(t) - \frac{1}{2}B_0 v^2(t) - c_0[f_q(\beta(t)) - 1], \tag{7.49}$$

with boundary conditions

$$\beta(T) = i\phi, \quad \alpha(T) = 0.$$

These two ODEs may be considered as an extended version of the Feynman-Kac theorem at the presence of Poisson jumps. In fact, the above two ODEs are the results of the PIDE given in (7.9) under the risk-neutral measure. Note that the ODE for $\beta(t) \in \mathbb{R}^n$ in (7.48) is an equation system consisting of n one-dimensional ODEs, while the ODE for $\alpha(t)$ in (7.49) is only one-dimensional. If the closed-form solutions for these ODEs are not available for some special specifications of $X(t)$, we would solve them numerically. To ensure the solutions $\alpha(t)$ and $\beta(t)$ solvable, the coefficients (A,B,C,D) are required to be well-behaved, which means in turn that the following integrability conditions must be satisfied:

[13] A poisson process with stochastic intensity is also called Cox process.

[14] As discussed in Chapter 8, "extended" Fourier transform essentially is the generalized Fourier transform.

1. $\mathbf{E}\big[|g^*(\phi,T)|\big] < \infty, \quad g^*(\phi,t) = \exp\big(-\int_0^t r(X(s),s)ds\big) f^*(\phi,T-t)$.
2. $\mathbf{E}\big[\int_0^T |\gamma_1(t)|dt\big] < \infty, \quad \gamma_1(t) = g^*(\phi,t)[f_q(\beta(t))-1]\lambda(X)$.
3. $\mathbf{E}\big[\big(\int_0^T \gamma_2^2(t)|dt\big)^{1/2}\big] < \infty, \quad \gamma_2(t) = g^*(\phi,t)\beta(t)v(X)$.

In the above ODEs, the term $f_q(\beta(t))$ is responsible for the effects of jump component in the final solutions. Therefore, the concrete specification of q determines the concrete solution forms of $\alpha(t)$ and $\beta(t)$. In the line with the above discussions, we have the following choices for q in a one-dimensional setting.[15]

1. q is constant: q is constant as in a simple jump in Section 7.2, and in particular, may be equal to 1, $f(z)$ reduces to the CF of a Poisson distribution,
$$f_Y(z;q=1) = \exp\big[\lambda(e^z-1)\big], \quad f_q = e^z.$$
2. q is Gaussian: Note that q is equivalent to $\ln(1+J)$ as in lognormal jumps. Hence, q may be assumed to normally distributed to $N(\mu_q,\sigma_q)$, the "extended" Fourier transform can be computed as
$$f_Y(z;q) = \exp\big[\lambda(f_q(z)-1)\big], \quad f_q(z) = e^{\mu_q z + \frac{1}{2}\sigma_q^2 z^2}.$$
3. q is exponentially distributed: q is again equivalent to $\ln(1+J)$ as in Pareto jump. Consequently, q is exponentially distributed with a mean of μ_q. The "extended" Fourier transform takes a form of
$$f_Y(z;q) = \exp\big[\lambda(f_q(z)-1)\big], \quad f_q(z) = \frac{1}{1-z\mu_q}.$$

Since the CF $f^*(\phi,t)$ is the so-called discounted CF, the pricing formula based on $f_1(\phi)$ and $f_2(\phi)$ as well as the alternative pricing formulas given in Chapter 4, can not be applied directly. To value a European-style option, we define the underlying stock price via $X(t)$, namely
$$S(T) = e^{\xi X(T)}$$
for a deterministic vector $\xi \in \mathbb{R}^n$. A call option is then valued as follows:

$$\begin{aligned}
C_0 &= \mathbf{E}^Q\Big[\exp\big(-\int_0^T r(X(s),s)ds\big)[e^{\xi X(T)}-K]\mathbf{1}_{(\xi X(T)>K)}|\mathscr{F}_0\Big] \\
&= \mathbf{E}^Q\Big[\exp\big(-\int_0^T r(X(s),s)ds\big)e^{\xi X(T)}\mathbf{1}_{(\xi X(T)>K)}|\mathscr{F}_0\Big] \\
&\quad - K\mathbf{E}^Q\Big[\exp\big(-\int_0^T r(X(s),s)ds\big)e^{0X(T)}\mathbf{1}_{(\xi X(T)>K)}|\mathscr{F}_0\Big].
\end{aligned}$$

Define a function

[15] In the previous sections in this chapter, we always specify the distribution of J instead of q. Since $q = \ln(1+J)$, we have
$$\mu_J = \mathbf{E}[J] = \mathbf{E}[e^q] - 1.$$

$$\Omega_{u_1,u_2}(y;X_0,T) = \mathbf{E}^Q\Big[\exp\big(-\int_0^T r(X(s),s)ds\big)e^{u_1X(T)}\mathbf{1}_{(u_2X(T))<y}|\mathscr{F}_0\Big],$$

the call option may be expressed as

$$C_0 = \Omega_{\xi,-\xi}(-\ln K;X_0,T) - K\Omega_{0,-\xi}(-\ln K;X_0,T).$$

Duffie, Pan and Singleton derived a more detailed form for $\Omega_u(y;X_0,T)$ based on the inverse Fourier transform of $f^*(\phi,t)$, which is given by

$$\Omega_{u_1,u_2}(y;X_0,T) = \frac{1}{2}f^*(-iu_1,T) - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}[f^*(-iu_1+u_2z,T)e^{-izy}]}{z}dz. \quad (7.50)$$

An illustrative example proposed by Duffie, Pan and Singleton is the Heston model added by two correlated jumps. Denote $X(t) = (G(t) = \ln S(t), V(t))$, the detailed formulation of this special affine jump-diffusion model reads

$$\begin{aligned} d\begin{pmatrix} G(t) \\ V(t) \end{pmatrix} &= \begin{pmatrix} r-\lambda m - \frac{1}{2}V(t) \\ \kappa[\theta - V(t)] \end{pmatrix} dt \\ &+ \sqrt{V(t)}\begin{pmatrix} 1 & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} + \begin{pmatrix} q_1 dY_1 \\ q_2 dY_2 \end{pmatrix}, \end{aligned} \quad (7.51)$$

where $W(t) = (W_1(t), W_2(t))$ is an independent two-dimensional Brownian motion. The jumps are driven jointly by two independent Poisson processes Y_1 and Y_2, and two correlated jumps size $q = (q_1, q_2)$. Denote $f_q(z_1,z_2)$ is the joint transform of q. In order to ensure that $G(t)$ is a martingale, m is required to be equal to $\mathbf{E}[J] = \mathbf{E}[e^{q_1}] - 1$, that in turn is equal to $f_q(1,0) - 1$.

The quasi closed-form solutions for $f^*(\phi,t)$ takes then the following form,

$$f^*(\phi,T) = e^{-rT}e^{i\phi(G(0)+rT)}f_2^{Heston}(\phi)f^{Jump}(\phi), \quad (7.52)$$

where $f_2^{Heston}(\phi)$ is the second CF in the Heston model. As given in (3.16), $f_2^{Heston}(\phi)$ has the the following solution,

$$f_j^{Heston}(\phi) = \exp\big(s_{22}(V_0 + \kappa\theta T) + H_1(T)V_0 + H_2(T)\big).$$

The detailed formulations for $s_{22}, H_1(T)$ and $H_2(T)$ can be found in Section 3.2. Denote $\beta(t) = s_{22} + H_1(t)$, $f^{Jump}(\phi)$ may be calculated as follows:

$$f^{Jump}(\phi) = \exp\big(-i\phi\bar{\lambda}mT + \bar{\lambda}\int_0^T \big[f_q(\phi,\beta(t)) - 1\big]dt\big). \quad (7.53)$$

Therefore, the final solution of $f^*(\phi,T)$ depends on the concrete form of $f_q(z_1,z_2)$ that in turn depends on the concrete specification of the joint distribution of q_1 and q_2. In the jump-diffusion models discussed in the previous sections, q_1 is specified concretely while q_2 is left unspecified and is equal to nil. Therefore $f_q(z_1,z_2)$ reduces to $f_{q_1}(z_1)$, and $\beta(t)$ does not affect the jump transform $f_{q_1}(z_1)$.

To incorporate some dependence between q_1 and q_2, Duffie, Pan and Singleton, for instance, suggested the following nested transform,

$$f_q(z_1,z_2) = \frac{1}{\bar{\lambda}}[\lambda_1 f_{q_1}(z_1) + \lambda_2 f_{q_2}(z_2) + \lambda_3 f_{q_1,q_2}(z_1,z_2)], \tag{7.54}$$

with

$$\bar{\lambda} = \lambda_1 + \lambda_2 + \lambda_3,$$
$$f_{q_1}(z_1) = \exp(\mu_1 z_1 + \frac{1}{2}\sigma_1^2 z_1^2),$$
$$f_{q_2}(z_2) = \frac{1}{1 - z_2\mu_2},$$
$$f_{q_1,q_2}(z_1,z_2) = \frac{\exp(\mu_3 z_1 + \frac{1}{2}\sigma_3^2 z_1^2)}{1 - z_2\mu_4 - \rho_J\mu_4 z_1}.$$

These complicated specifications of jump sizes and their implied correlation may be simply divided into three parts. The first one is a normal (lognormal) jump given by $f_{q_1}(z_1)$ for stock returns $G(t=1)$ (or $S(t=1)$), the second one is an exponential (Pareto) jump given by $f_{q_2}(z_2)$ for $V(t=1)$ (or $e^{V(t)}$). The last transform $f_{q_1,q_2}(z_1,z_2)$ determines the marginal joint distribution of the jumps in $G(t)$ and $V(t)$. Conditional on a realization q_V of jumps in $V(1)$, the jump size in $G(1)$ is normally distributed with a mean of $\mu_3 + \rho_J q_V$ and a variance of σ_3^2. With $f_q(z)$ having been specified, we can integrate the term $\int_0^T f_q(\phi,\beta(t))dt$ to obtain $f^{Jump}(\phi)$.

Chapter 8
Lévy Jumps

Lévy processes are referred to as a large class of stationary processes with independent identical increments. Brownian motion and Poisson process can be regarded as two special cases of Lévy process, and have only finite activity in a finite time interval. In this chapter, we only consider Lévy processes with infinity activity in a finite time interval. With respect to jump event modeling in finance, compound Poisson jumps discussed in Chapter 7 are appropriate for capturing rare and large jumps such as market crashes, while Lévy processes with infinity activity may be better suited to describe many small jumps such as discontinuous information flows and discrete trading activities. To distinguish Poisson jump dynamics from jump dynamics with infinity activity, we label the latter jumps as Lévy jumps in this book.

In first section, we introduce Lévy process and present the basic concepts of Lévy processes, such as the Lévy measure and the Lévy-Khintchine theorem. In Section 8.2, two popular Lévy models, the variance-gamma model and the normal inverse Gaussian model, are examined. It can be shown that both Lévy processes may be constructed by the subordination of a Brownian motion to gamma process and inverse Gaussian process, respectively, that are called subordinators. A subordinator may be intuitively interpreted as a stochastic time-change. Therefore the option pricing models with the above Lévy processes are also referred to as stochastic clock models. In Section 8.3, we consequently apply stochastic time-change to general Lévy processes, even in the presence of a dependence between the stochastic time and the Lévy process itself. To deal with such dependence conveniently, a complex-valued measure is introduced. The resulting time-changed Lévy processes are not necessarily Lévy processes, but have much richer spectra of distributions, and therefore may describe asset dynamics better. In Section 8.4 and 8.5, we briefly discuss two special Lévy processes: the Barndorff-Nielsen and Shephard model and the alpha log-stable model. Next, we summarize some empirical test results on various Lévy processes and time-changed Lévy processes, their empirical performances with respect to pricing and hedging are an important aspect for practical applications of Lévy jump models. Finally in Section 8.7, some algorithms for generating Lévy random numbers and some methods for simulating Lévy processes are given.

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_8,

8.1 Introduction

A Lévy process is a natural extension of a Wiener process and a Poisson process, and is defined as follows:

Definition 8.1.1. *Lévy process: A Lévy process* $\{L(t), t \geqslant 0\}$ *is a stochastic process satisfying the following conditions,*

1. *For* $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n < \infty$, $L(t_1), L(t_2) - L(t_1), \cdots, L(t_n) - L(t_{n-1})$ *are stochastically independent;*
2. *Its increments are time-homogenous, it means that* $L(t+s) - L(t), s > 0$ *is independent of* t*;*
3. *For each fixed* $\omega \in \Omega, L(\omega, t)$ *is continuous in* t*;*
4. $L(0) = 0$ *almost surely.*

By comparing the above definition of a Lévy process with the ones of a Wiener process and a Poisson process, we can see that a Lévy process does not explicitly define a concrete distribution of its increments. The conditions 1, 3, and 4 are the same as for a Wiener process or for a Poisson process. Therefore, Lévy process contents Wiener process and Poisson process as special cases and more other processes, some of which are the focus of this chapter. In fact, as discussed later, Wiener process and Poisson process are two important ingredients of Lévy process. So far, we have discussed how to model financial processes by using Wiener processes and Poisson processes. Hence, applying Lévy processes to valuing financial derivatives is a self-understanding choice.

In more details, a Lévy process is a linear combination of a deterministic drift, a Wiener process and a jump process (particularly Poisson process) where the deterministic drift, the Wiener process can be verified uniquely. The construction of jumps in a Lévy process is more involved and consists of many possible types that could be characterized by the so-called Lévy measure. In this sense, different Lévy measures generate different Lévy processes. A standard Lévy process can be expressed in the following form,

$$L(t) = \mu t + \sigma W(t) + \int_0^t \int_{\mathbb{R}/\{0\}} [x - h(x)] \nu(x) dx ds, \tag{8.1}$$

where μ is a deterministic increment, $W(t)$ is a Wiener process and $\nu(x)$ is a Lévy measure. Roughly speaking, Lévy measure is a measure on $\mathbb{R}$, and satisfies the following two conditions:

$$\nu(\{0\}) = 0 \tag{8.2}$$

and

$$\int_{\mathbb{R}} \min(1, x^2) \nu(x) dx < \infty. \tag{8.3}$$

Note that the definition of Lévy measure $\nu(x)$ does not requires $\int_{\mathbb{R}} \nu(x) dx < \infty$, an usual requirement for density measure. In fact, many Lévy measures are not integrable.

The function $h(x)$ is an arbitrary truncation function. One of possible forms could be

$$h(x) = x\mathbf{1}_{|x|<1}.$$

The truncation function $h(x)$ ensures that the integral part in Lévy process is integrable. Generally, we do not need a truncation function to define Poisson process via the above formulation. By setting $h(x) = 0$, we obtain

$$\int_{\mathbb{R}/\{0\}} x\nu(x)dx = \int_{\mathbb{R}/\{0\}} x\lambda g(x)dx = \lambda \int_{\mathbb{R}/\{0\}} xg(x)dx.$$

Here $g(x)$ and λ can be considered as size density function and Poisson intensity parameter, respectively. $\lambda \int_{\mathbb{R}/\{0\}} xg(x)dx$ is the cumulative jumps in a small interval dt, and therefore, can be rewritten as

$$\lambda \int_{\mathbb{R}/\{0\}} xg(x)dx = JdY,$$

which is the formulation for compound Poisson jump used before. In this simplified form, Lévy measure becomes a joint measure for jump size and arrival rate, and Lévy process reduces to a jump-diffusion process

$$dL(t) = \mu dt + \sigma dW(t) + JdY.$$

In the previous chapter on Poisson jumps, we have observed that various random jump processes are rather moderate in the sense: (1) they do not have infinitely large jumps and (2) they do not have too "many small changeful" jumps so that the quadratic variation does not exists. To ensure that a Lévy process is a "reasonable" process, infinite large jumps and too "many small changeful" jumps should be eliminated from Lévy process. Mathematically, we have the following definitions for these two irrational cases:

1. "Too large" jumps: Lévy process $L(t)$ is regarded to have "too large" jumps, if

$$\int_{|x|>1} |x|\nu(x)dx = \infty.$$

2. Too "many small changeful" jumps: Lévy process $L(t)$ is regarded to have too "many small changeful" jumps, if

$$\int_{|x|<1} x^2\nu(x)dx = \infty.$$

It is easy to see in this two cases that a Lévy process does not allow that the large jumps ($|x| > 1$) have infinite variation, at the same time, does not allow that the small jumps ($|x| < 1$) have infinite quadratic variation. In fact, introducing a truncation function $h(t)$ in a Lévy process is to eliminate these two irrational cases from Lévy process, and is equivalent to the following condition,

$$\int_{\mathbb{R}/\{0\}} (x^2 \wedge 1)\nu(x)dx < \infty. \tag{8.4}$$

This means that a Lévy process must have a finite quadratic variation.

A Lévy process is considered to have an infinite activity, if

$$\int_{\mathbb{R}/\{0\}} \nu(x)dx = \infty,$$

otherwise, to have a finite activity. If we do not take the Wiener process into account, and

$$\int_{\mathbb{R}/\{0\}} (|x| \wedge 1)\nu(x)dx = \infty,$$

a Lévy process exhibits an infinite variation, otherwise, exhibits a finite variation. For instance, a Poisson process has a finite activity and a finite variation, this is the reason why we think the Poisson process is a rather moderate Lévy process.

The Lévy-Khintchine theorem describes a Lévy process in terms of characteristic function. Since the characteristic function is the key to establish the distribution of Lévy random variable at a certain time, the Lévy-Khintchine theorem is important in the context of the option valuation with Lévy process.

Theorem 8.1.2. *Lévy-Khintchine theorem: Any Lévy process*

$$L(t) = \mu t + \sigma W(t) + \int_0^t \int_{\mathbb{R}/\{0\}} [x - h(x)]\nu(x)dxds$$

has the following characteristic function

$$f(\phi,t) = \boldsymbol{E}[e^{i\phi L(t)}] = e^{\Psi(\phi,t)}, \tag{8.5}$$

where

$$\Psi(\phi,t) = i\phi\mu t - \frac{1}{2}\phi^2\sigma^2 t + \int_{\mathbb{R}/\{0\}} [e^{i\phi x} - 1 - i\phi h(x)]\nu(x)dx.$$

The term $\Psi(\phi,t)$ is called the Lévy exponent.

According to the Lévy-Khintchine theorem, any Lévy process is characterized uniquely by the triple (μ, σ, ν), that is referred to as the Lévy triple. Additionally, The characteristic function of a Lévy process has the following properties:

1. The characteristic function of a Lévy process comprises the characteristic function of a Brownian motion with drift and the characteristic function of jump components, namely, $f(\phi,t) = f_B(\phi) f_{Jump}(\phi)$ with

   $$f_B(\phi) = i\phi\mu t - \frac{1}{2}\phi^2\sigma^2 t$$

   and

   $$f_{Jump}(\phi) = \int_{\mathbb{R}/\{0\}} [e^{i\phi x} - 1 - i\phi h(x)]\nu(x)dx.$$

2. Lévy process is infinitely divisible,

$$f(\phi,L,t) = f^n(\phi,L/n,t).$$

This property implies the self-similarity or the stability property of a Lévy process.

There are various option pricing models based on Lévy processes which in turn are described by various Lévy measures, for instance, the variance-gamma process. Since a Lévy measure is generally a function, we can actually propose an infinite number of Lévy processes. For practical applications, only Lévy measures allowing for an analytical closed-form solution for the Lévy exponent could be of interest. As long as the characteristic function of a particular Lévy process is analytically known, we can apply the Fourier inversion to obtain the corresponding distribution, and therefore also the pricing formula for European-style options. Carr, Geman, Madan and Yor (2001) discussed various Lévy processes and their applications in option pricing. Cont and Tankov (2004) provided a detailed mathematical treatment of Lévy process in the connection with option theory. Schoutens, Simons and Tistaret (2004) compared different Lévy processes and the associated model risks. Carr and Wu (2004) extended stochastic time Lévy process with non-zero correlation between stochastic time and the underlying Lévy process. Huang and Wu (2004) examined the specification problems with respect to time-changed Lévy processes. Wu (2006) presented an excellent overview on the financial modeling with Lévy process. For more detailed discussions on Lévy process, please see Bertoin (1996), Sato (1999) and Schoutens (2003).

In most financial literature, however, Lévy processes are implicitly referred to the ones characterized only by infinite activity, even excluding Poisson jumps discussed perviously. To avoid any confusion in this chapter, we refer to all jumps excluding Poisson jumps in Lévy process as Lévy jumps in following.

8.2 Stochastic Clock Models

In this section, we discuss the so-called stochastic clock model, a commonplace for a range of Lévy models that are generated by subordinating a Brownian motion to another process. A subordinator is a process $L(t)$ such that for all $t,t > 0$, we have $L(t) > 0$. This means that a subordinator just maps a positive time to another positive time, and virtually serves as a clock changing time stochastically. So any model underlying this mechanism can be classified to a stochastic clock model.

The intuition behind the subordination of a Brownian motion with drift, which is usually applied to model the dynamics of asset returns, is that the stochastic clock may be considered as a cumulative measure of economic activity. If the clock runs faster than usually, i.e., $L(t) > t$, larger variances in a stochastic clock model could be generated than in the corresponding Gaussian model. In contrast, if the clock runs slower than usually, i.e., $L(t) < t$, we could expect smaller variances in a stochastic clock model. In a similar manner, negative skewness and excess kurtosis could be also generated by applying a sound subordinator. As discussed in the section on Poisson jumps, diffusion processes such as Brownian motion do not content large movements enough so that the fat tail of the distribution of most asset returns can not be properly recovered. The fact that the generated subordinated processes in stochastic clock models may be mathematically equivalent to pure jump Lévy process with infinite activity, supports the potential capacity of stochastic clock models in capturing leptokurtic feature of asset returns.

In the Black-Scholes world, the CF of log return $x(T)$ is given by

$$f(\phi) = \mathbf{E}[e^{i\phi x(T)}].$$

Now we use a subordinator $L(T)$ to randomize T. Denoting $\widetilde{T} = L(T)$ and calculating $f(\phi)$ under the Gaussian law leads to

$$f(\phi) = \mathbf{E}[e^{i\phi\mu\widetilde{T} - \frac{1}{2}\phi^2\sigma^2\widetilde{T}}] = e^{\Psi_B(\phi)\widetilde{T}}, \tag{8.6}$$

with $\Psi_B(\phi) = i\phi\mu - \frac{1}{2}\phi^2\sigma^2$ as the exponent of the CF of $x(1)$. Therefore, the expected value

$$\mathbf{E}[e^{\Psi_B(\phi)\widetilde{T}}]$$

may be considered as the moment-generating function of $\widetilde{T}$ with a parameter $\Psi_B(\phi)$. On the other hand, we assume the moment-generating function of the subordinator $L(T)$ takes the following form

$$\mathbf{E}[e^{uL(T)}] = e^{\Psi_L(u)}, \tag{8.7}$$

where $\Psi_L(u)$ is the exponent of the moment-generating function of the Lévy subordinator. By setting $u = \Psi_B(\phi)$, we obtain the final form of $f(\phi)$ under the joint distribution law of the Brownian motion and the Lévy subordinator as follows

$$f(\phi) = e^{\Psi_L(\Psi_B(\phi))}, \tag{8.8}$$

which is comprised of two nested moment-generating functions. If both $e^{\Psi_L(u)}$ and $\Psi_B(\phi)$ have analytical solutions, the CF $f(\phi)$ of the time-changed Brownian motion has also an analytical one. This result is especially important if we search for tractable option pricing formulas of stochastic clock models. The procedure that subordinating a Brownian motion $B(t)$ to $L(t)$ and deriving the corresponding CF is called the Bochner procedure.

Note that $\Psi_L(u)$ is the exponent of a moment-generating function, not of characteristic function of $L(t)$. To keep the notations consistent, we can also express the moment-generating function$f(\phi)$ with the exponent of a characteristic function. Denote the CF of $L(T)$ as

$$\mathbf{E}[e^{i\phi L(T)}] = e^{\Psi_L(\phi)},$$

here $\Psi_L(u)$ is then the desired exponent. It follows immediately

$$f(\phi) = e^{\Psi_L(-i\Psi_B(\phi))}. \tag{8.9}$$

In following subsections, we discuss two popular stochastic clock models: the variance-gamma model and the normal inverse Gaussian model. The implied distributions in both models are two special cases of the generalized hyperbolic distribution. For more detailed applications of the generalized hyperbolic distribution in finance, see Eberlein, Keller and Prause (1998).

8.2.1 Variance-Gamma Model

In this section we consider a popular Lévy jump model called the variance-gamma model. This model was initialized by Madan and Milne (1991), and developed by Madan, Carr and Chang (1998). The basic idea of the variance-gamma model is to randomize the deterministic time length of a Brownian motion by a gamma distribution. More strictly, the Brownian motion is subordinated to a gamma process. Since the original Brownian motion has a variance of one in a time unit, the variance of the changed Brownian motion is then gamma-distributed. Therefore, this model is referred to as the variance-gamma model.

Denote $B(t)$ as a Brownian motion with drift,

$$B(t) = B(t;\mu,\sigma) = \mu t + \sigma W(t)$$

and denote $\gamma(t;\alpha,\beta)$ as a gamma process with a mean of α and a variance of β in an unit time length. The variance-gamma process is constructed as

$$VG(t;\mu,\sigma,\beta) = B(\gamma(t;1,\beta);\mu,\sigma), \tag{8.10}$$

where it is clear that the time t in the Brownian motion is replaced by $\gamma(t;1,\beta)$. This gamma process has a mean of one, and insures that the time-changed Brownian motion still has an expected time length of t. Since the gamma distribution is defined in the positive real axis, the variance-gamma process is well-defined.

In the variance-gamma model, the risk-neutral asset price $S(t)$ is represented as a martingale-adjusted form,

$$S(t) = S_0 \exp(rt - mt + VG(t;\mu,\sigma,\beta)), \tag{8.11}$$

where m will compensate for the dynamics of $VG(t;\mu,\sigma,\beta)$ in terms of expectation, and is equal to

$$m = \ln \mathbf{E}[\exp VG(1;\mu,\sigma,\beta)].$$

Define $x(t) = \ln S(t)$, we may express $S(t)$ in a form of stochastic process,

$$dx(t) = (r-m)dt + dVG(\mu,\sigma,\beta). \tag{8.12}$$

We apply the Bochner's procedure to obtain the CF of the variance-gamma model. The CF of the gamma process $\gamma(t;1,\beta)$ is known as

$$f_{gamma}(\phi) = (1 - i\phi\beta)^{-\frac{t}{\beta}}. \tag{8.13}$$

It follows then

$$f(\phi) = f_{gamma}(-i\Psi_B(\phi)) = \left(1 - i\phi\mu\beta + \frac{1}{2}\sigma^2\phi^2\beta\right)^{-\frac{t}{\beta}} \tag{8.14}$$

with

$$\Psi_B(\phi) = i\phi\mu - \frac{1}{2}\phi^2\sigma^2.$$

The term m adjusting the variance-gamma process to be a martingale is then identical to $f(-i;t=1)$, and is equal to

$$m = \frac{1}{\beta}\ln[1 - \mu\beta - \frac{1}{2}\sigma^2\beta].$$

With this adjustment of the drift, we can easily verify that the discounted asset price in the variance-gamma process is a martingale, namely,

$$e^{-r(T-t)}\mathbf{E}[S(T)|\mathscr{F}_t] = S(t)\exp(-m(T-t) + VG(T-t;\mu,\sigma,\beta)) = S(t).$$

The Lévy measure of the variance-gamma process is given by

$$\nu_{VG} = \frac{e^{cx}}{\beta|x|}\exp(-\frac{|x|}{\beta}\sqrt{2\beta^{-1} + c\mu}) \tag{8.15}$$

with

$$c = \frac{\mu}{\sigma^2}.$$

From the above Lévy measure we can derive the following properties of the variance-gamma model:

1. The variance-gamma process exhibits infinite activity, or equivalently exhibits infinite jump frequency, since it exists $\int_{\mathbb{R}} \nu_{VG}(x)dx = \infty$.
2. The variance-gamma process exhibits finite variation due to the following property

$$\int_{|x|<1} x\nu_{VG}(x)dx < \infty.$$

3. The first four central moments of the variance-gamma process are given by

$$\begin{aligned} m_1 &= \mu t, \\ m_2 &= (\mu^2\beta + \sigma^2)t, \end{aligned}$$

$$\begin{aligned}
m_3 &= (2\mu^3\beta^2 + 3\mu\beta\sigma^2)t, \\
m_4 &= (3\beta\sigma^4 + 12\sigma^2\beta^2\mu^2 + 6\mu^4\beta^3)t \\
&\quad + (3\sigma^4 + 6\beta\sigma^2\mu^2 + 3\mu^4\beta^2)t^2.
\end{aligned}$$

It follows immediately that the parameter μ controls the sign of the skewness. Therefore, a negative μ should be expected for modeling an asset process. Moveover, if $\mu = 0$, the excess kurtosis is equal to $3(1+\beta)$.

Now we calculate the CFs for the option pricing formula.

$$\begin{aligned}
f_1(\phi) &= \mathbf{E}[g_1(T)e^{i\phi x(T)}] = \mathbf{E}[e^{-(x_0+rT)+(1+i\phi)x(T)}] \\
&= e^{i\phi(x_0+rT)-(1+i\phi)mT} \times \left(\frac{1}{1-(1+i\phi)\mu\beta - \frac{1}{2}\sigma^2\beta(1+i\phi)^2}\right)^{T/\beta} \\
&= e^{i\phi(x_0+rT)} f_1^{VG}(\phi) \qquad (8.16)
\end{aligned}$$

and

$$\begin{aligned}
f_2(\phi) &= \mathbf{E}[e^{i\phi x(T)}] = e^{i\phi(x_0+rT)-i\phi mT} \times \left(\frac{1}{1-i\phi\mu\beta + \frac{1}{2}\sigma^2\beta\phi^2}\right)^{T/\beta} \\
&= e^{i\phi(x_0+rT)} f_2^{VG}(\phi). \qquad (8.17)
\end{aligned}$$

It is very convenient to derive the pricing formula based on the above two CFs within the framework given in Chapter 2. The numerical Fourier inversion of the both CFs share the same implementation as in stochastic volatility models, and present no large challenge. However, the analytical formula given by Madan, Carr and Chang (1998) is based on a modified Bessel function, and therefore is difficult to implement. We do not discuss this complicated formula here.

8.2.2 Normal Inverse Gaussian Model

The normal inverse Gaussian (NIG) model suggested by Barndorff-Nielsen (1998) is very similar to the variance-gamma model, except for that the deterministic time length of a Brownian motion in the NIG model is randomized by an inverse Gaussian distribution. Therefore, the time-changed Brownian motion in the NIG process is distributed according to inverse Gaussian law. Denote $IG(t;\alpha,\theta)$ as an inverse Gaussian distribution for a time length of t where α is mean and θ is variance. An inverse Gaussian random variable can be originated to be the first passage time for a fixed level $b, b>0$, of a Brownian motion with a drift κ. So we have

$$\alpha = \frac{b}{\kappa}, \quad \theta = \frac{b}{\kappa^3}.$$

In the context of the normal inverse Gaussian model, we set b to be the time length t of the subordinated Brownian motion $B(t;\mu,\sigma)$. The CF of an inverse Gaussian is given by

$$\Psi(\phi) = \exp(t[\kappa - \sqrt{\kappa^2 - 2i\phi}]). \tag{8.18}$$

The NIG process $NIG(t;\mu,\sigma,\kappa)$ can be expressed as

$$NIG(t;\mu,\sigma,\kappa) = B(IG(t;t/\kappa,t/\kappa^3);\mu,\sigma). \tag{8.19}$$

This means that the time t is in fact replaced by the first passage time of an other Brownian motion hitting t. Correspondingly, to construct a risk-neutral process for asset $S(t)$, we need an adjustment for drift and arrive at the following form,

$$S(t) = S_0 \exp(rt - mt + NIG(t;\mu,\sigma,\kappa))$$

or

$$dx(t) = (r-m)dt + dNIG(t;\mu,\sigma,\kappa)) \tag{8.20}$$

with

$$m = \ln \mathbf{E}[\exp NIG(1;\mu,\sigma,\kappa)].$$

According to the Bochner's procedure, it is easy to obtain the CF of the NIG process,

$$f_{NIG}(\phi) = f_{IG}(-i\Psi_B(\phi)) = \exp\left(t[\kappa - \sqrt{\kappa^2 + \sigma^2\phi^2 - 2\mu i\phi}]\right). \tag{8.21}$$

It follows immediately

$$m = \kappa - \sqrt{\kappa^2 - \sigma^2 - 2\mu}.$$

The Lévy measure implied in the NIG process is given by

$$\nu_{NIG}(x) = \sqrt{2/\pi}(\kappa^2\sigma^{-1} + c^2\sigma)\frac{e^{cx}K(|x|)}{|x|} \tag{8.22}$$

with

$$c = \frac{\mu}{\sigma^2},$$

where $K(\cdot)$ is the modified Bessel function of the second kind (see Carr-Geman-Madan-Yor (2001)). From the given Lévy measure we can conclude

1. The NIG process has infinite activity since it exists $\int_{\mathbb{R}} \nu_{NIG}(x)dx = \infty$. This property is identical to the one of variance-gamma process.
2. The NIG process has also infinite variation due to the property

$$\int_{|x|<1} x\nu_{NIG}(x)dx = \infty.$$

The two CFs of $x(T)$ can be calculated as follows:

$$\begin{aligned}
f_1(\phi) &= \mathbf{E}[g_1(T)e^{i\phi x(T)}] = \mathbf{E}[e^{-(x_0+rT)+(1+i\phi)x(T)}] \\
&= e^{i\phi(x_0+rT)-(1+i\phi)mT} \exp\left(T[\kappa - \sqrt{\kappa^2 - \sigma^2(1+i\phi)^2 - 2\mu(1+i\phi)}]\right) \\
&= e^{i\phi(x_0+rT)} f_1^{NIG}(\phi)
\end{aligned} \tag{8.23}$$

and

$$\begin{aligned}
f_2(\phi) &= \mathbf{E}[e^{i\phi x(T)}] \\
&= e^{i\phi(x_0+rT)-i\phi mT} \exp\left(T[\kappa - \sqrt{\kappa^2 + \sigma^2\phi^2 - 2\mu i\phi}]\right) \\
&= e^{i\phi(x_0+rT)} f_2^{NIG}(\phi).
\end{aligned} \tag{8.24}$$

The valuation of European-style options with the NIG process follows then the same way as with the variance-gamma process.

8.3 Time-Changed Lévy Process

We has seen that the variance-gamma process and the NIG process result from a subordination of a Brownian motion to a gamma process and an inverse Gaussian process, respectively. In this section, we discuss how to construct a process by subordinating a general Lévy process with a (general) subordinator.

As indicated above, a subordinator is defined as a Lévy process that changes a positive time to another stochastic positive time. Since a clock generates always an increasing time, the new cumulative stochastic time must be also increasing. Generally speaking, a subordinator is not necessarily a Lévy process so that such a stochastic clock can be defined. For example, a mean-reverting square root process that has been discussed early in details in the Heston model, may be used to map a positive time to another positive time if a suitable parameter set is chosen. As in stochastic clock models such as the variance-gamma model, the stochastic clock here may be interpreted as a driver for instantaneous business activity. A higher activity rate generates higher volatility in a time-changed process, and a smaller activity rate leads to smaller volatility in a time-changed process. Hence, a stochastic time-change is often called instantaneous activity rate, in order to emphasis the similar functionality of the stochastic time-change in a time-changed Lévy process to the stochastic volatility in a diffusion process with respect to smile modeling.

As we can see in the CF of the mean-reverting square root process given in the Heston model, a mean-reverting square root process is just an Itô process, not a Lévy process since its CF is not infinitely divisible, a necessary condition that Lévy process must satisfy. This indicates that a subordinator may not be necessarily a Lévy process. Except for the mean-reverting square root process, a so-called Gamma

Ornstein-Uhlenback process, first introduced by Barndorff-Nielsen and Schephard (2001), is a Lévy process and can also serve as a subordinator.

In the following, two cases will be examined with respect to the correlation between the Lévy process and the subordinator. The first case is a case with uncorrelated time-change where the changed Lévy process and the substantiator are independent of each other. Due to the independence, the distribution law of the subordinator can be different from the distribution law of the underlying Lévy process. In the second case, the changed Lévy process is correlated with its subordinator. Due to the very nature of correlation or dependence, the distribution type of the subordinator is usually identical to the distribution type of the Lévy process.

8.3.1 Uncorrelated Time-Change

Denote $L(t)$ as a Lévy process again, particularly it could be the variance-gamma process or the NIG process. Due to the infinite divisibility, the CF of a Lévy process $L(t)$ can be expressed in the following form

$$f(\phi) = \mathbf{E}[e^{i\phi L(T)}] = e^{\Psi_L(\phi)} = e^{\Psi_L^*(\phi)T}, \tag{8.25}$$

which is essentially an extension of (8.6), and is identical to (8.9). $\Psi_L(\phi)$ may be regarded as the CF exponent of $L(1)$,

$$\mathbf{E}[e^{i\phi L(1)}] = e^{\Psi_L^*(\phi)}.$$

For a particular subordinator $Y(t)$ that is uncorrelated with $L(t)$, we define a time-changed Lévy process as follows

$$X(t) = L(Y(t)).$$

By applying the Bochner's procedure, we may obtain the CF of the generated process $X(t)$,

$$\begin{aligned}\mathbf{E}[e^{i\phi X(t)}] = &= \mathbf{E}[e^{i\phi X(Y(t))}] = \mathbf{E}[e^{\Psi_L^*(\phi)Y(t)}] \\ &= \mathbf{E}[e^{i(-i\Psi_L^*(\phi))Y(t)}] = e^{\Psi_Y(-i\Psi_L^*(\phi))},\end{aligned} \tag{8.26}$$

with $\Psi_Y(\phi_Y)$ as the CF exponent of the subordinator process $Y(t)$. Note that we are using $\phi_Y = -i\Psi_L^*(\phi)$ as the characteristic parameter in the application of the CF $e^{\Psi_Y(\phi_Y)}$ of $Y(t)$. As ϕ_Y is defined in the real domain, $-i\Psi_L^*(\phi)$ is generally not real-valued, but complex-valued. For example, let $L(t)$ be a Brownian motion with drift. Its CF exponent is $i\phi\mu + \frac{1}{2}\sigma^2\phi^2$. The resulting ϕ_Y is then $\phi\mu - \frac{1}{2}\sigma^2\phi^2 i$. In this sense, we have extended the definition domain of ϕ_Y to the complex plane unconsciously. In fact, if a characteristic parameter is defined in the complex plane, the Fourier transform is referred to as the generalized Fourier transform. We will discuss the generalized Fourier transform more in details later.

Based on the time-changed process $X(t)$, the corresponding asset process $S(t)$ is assumed to take the following form,

$$S(t) = S_0 \exp(rt - m(t) + X(t)) \tag{8.27}$$

with a drift compensator,

$$m(t) = \ln \mathbf{E}[\exp(X(t)].$$

In the following, we will consider two cases: In the first case, the subordinator $Y(t)$ is a mean-reverting square root process. In the second case, the Lévy process $L(t)$ is time-changed by a Gamma Ornstein-Uhlenbeck process. Finally, we will follow the Carr and Wu's approach (2004) to relax the restriction of the zero-correlation between $L(t)$ and $Y(t)$, and explore the correlation case between $L(t)$ and $Y(t)$ by introducing the generalized Fourier transform where a leverage-neutral measure and the corresponding complex-valued change of measure play a key role in constructing the generalized time-changed Lévy process.

Time-Changed By Square Root Process

We address first briefly how we may enhance the variance-gamma model and the NIG model if we use a mean-reverting square root process to stochastically change the time in the variance-gamma model and the NIG model again, as suggested by Carr, Geman, Madan and Yor (2001). The stochastic clock defined by a mean-reverting square root process $y(t)$ is

$$Y(t) = \int_0^t y(u)du, \tag{8.28}$$

where

$$dy(t) = \kappa_y(\theta_y - y(t))dt + \sigma_y\sqrt{y(t)}dW(t).$$

The CF of $y(t)$ is known in the Heston model and has the following form,

$$\Psi_Y(\phi) = H_1(t)y_0 + H_2(t).$$

The functions $H_1(t)$ and $H_2(t)$ are defined in the Heston model. Carr, Geman, Madan and Yor constructed a so-called stochastic volatility Lévy process[1] as follows:

$$X(t) = L(Y(t)). \tag{8.29}$$

From the CF of $X(t)$, especially of $Y(t)$, we can easily verify that $X(t)$ is no longer a Lévy process since the resulting CF is not infinitely divisible.

[1] The name stochastic volatility Lévy process is a bit misleading. The term "stochastic volatility" is generic for a mechanism generating volatility smile.

Two particular processes of $L(t)$, the variance-gamma process and the NIG process can be employed to $Y(t)$ to demonstrate how to enhance the existing Lévy processes.

Variance-gamma process:

The new process $XVG(T) = X_{VG}(t)$ can be expressed as

$$XVG(t) = VG(Y(t); \mu, \sigma, \beta = 1).$$

The corresponding CF of $X_{VG}(t)$ takes a nested form

$$f_{XVG}(\phi; \mu, \sigma, y_0, \kappa_y, \theta_y, \sigma_y) = \exp(\Psi_Y(-i\Psi^*_{VG}(\phi))).$$

Here β is set to 1 in the variance-gamma process. As argued by Carr, Geman, Madan and Yor, the effective β in the time-changed process $XVG(t)$ should be set to $1/y(0)$.

NIG Process:

Subordinating the NIG process to $Y(t)$ yields the new process $XNIG(t) = X_{NIG}(t)$,

$$XNIG(t) = NIG(Y(t); \mu, 1, \kappa)$$

and the corresponding CF of $X_{NIG}(t)$

$$f_{XNIG}(\phi; \mu, \kappa, y_0, \kappa_y, \theta_y, \sigma_y) = \exp(\Psi_Y(-i\Psi^*_{NIG}(\phi))).$$

Here $y(0)$ may be suggested to be equal to the parameter σ in the NIG model.

Time-Changed By Gamma-OU Process

The gamma Ornstein-Uhlenback process, suggested by Barndorff-Nielsen and Schephard (2001), takes the following form,

$$dy(t) = -\kappa_y y(t)dt + dZ(t), \tag{8.30}$$

where $Z(t)$ is the so-called background driving Lévy process for $y(t)$.[2] Here we consider a special construction where $Z(t)$ is a compound Poisson process

$$Z(t) = \sum_{n=1}^{N(t)} x_n,$$

with $N(t)$ as a Poisson counting process and x_n as an independent and identical exponentially distributed sequence. Hence, the gamma Ornstein-Uhlenback process

[2] As shown in Barndorff-Nielsen and Schephard (2001), the background driving Lévy process could be any Lévy process that satisfies the technical condition $\mathbf{E}[\ln(1 + |Z(1)|)] < \infty$.

can be alternatively expressed as

$$dy(t) = -\kappa_y y(t)dt + xdN(t). \tag{8.31}$$

This presentation of the compound Poisson process agrees with the one in Chapter 7.

It can be shown that $y(t)$ then follows a marginal gamma distribution. Due to the formal similarity between the process (8.30) and the Gaussian Ornstein-Uhlenbeck process, the process given in (8.30) is referred to as gamma Ornstein-Uhlenback process, in short, gamma-OU process. Particularly, we assume that $y(t)$ follows a marginal gamma process with a mean of α_y and a variance of β_y, and examine how to use $Y(t) = \int_0^t y(u)du$ to subordinate another Lévy process $L(t)$. As shown in Schoutens, Simons and Tistaert (2004), the CF of $Y(t)$ is given by

$$\begin{aligned} f_{GammaOU}(\phi) &= \exp\big(i\phi\gamma y_0 - \frac{i\phi\kappa_y\alpha_y t}{i\phi - \kappa_y c} + \frac{\kappa_y\alpha_y c}{i\phi - \kappa_y c}\ln\big(\frac{c}{c - i\phi\gamma}\big)\big) \\ &= e^{\Psi_Y(\phi)} \end{aligned} \tag{8.32}$$

with

$$\gamma = \frac{1}{\kappa_y}(1 - e^{-\kappa_y t}), \quad c = \frac{\alpha_y}{\beta_y}.$$

Variance-gamma process:

The new process $XVG(T) = X_{VG}(t)$ can be constructed as

$$XVG(t) = VG(Y(t); \mu, \sigma, \beta).$$

The corresponding CF of $X_{VG}(t)$ takes the following form

$$f_{XVG}(\phi; \mu, \sigma, \beta, y_0, \kappa_y, \alpha_y, \beta_y) = \exp(\Psi_Y(-i\Psi^*_{VG}(\phi))).$$

NIG Process:

Subordinating the NIG process to the gamma-OU process $Y(t)$ yields the new process $XNIG(t) = X_{NIG}(t)$,

$$XNIG(t) = NIG(Y(t); \mu, \sigma, \kappa)$$

and the corresponding CF of $X_{NIG}(t)$

$$f_{XNIG}(\phi; \mu, \sigma, \kappa, y_0, \kappa_y, \alpha_y, \beta_y) = \exp(\Psi_Y(-i\Psi^*_{NIG}(\phi))).$$

8.3.2 Correlated Time-Change

Generalized Fourier Transform

As shown above, two subordinators, the mean-reverting square root process and the gamma Ornstein-Uhlenbeck process, are not correlated with the underlying Lévy process. In this section, we relax the assumption of zero correlation and accommodate a non-zero correlation between the subordinator $Y(t)$ and the underlying Lévy process $L(t)$. In the context of pure jumps, we refer to this correlation as a leverage effect that generates the asymmetry of the distribution of the time-changed Lévy process.

To deal with leverage effect properly, we extend formally the real-valued parameter ϕ in a CF $\mathbf{E}[e^{-i\phi x(t)}]$ to the complex plane although we have done it through this chapter unconsciously. Denote $\mathscr{D} \subseteq \mathbb{C}^n$ as the domain where the CF with ϕ is well-defined. Such a CF $\mathbf{E}[e^{-i\phi x(t)}]$ with $\phi \in \mathscr{D}$, is called the generalized Fourier transform according to Tichmarsh (1975). Recall the expression for the CF of the time-changed Lévy process $X(t)$ in (8.26),

$$\mathbf{E}[e^{i\phi X(t)}] = e^{\Psi_Y(-i\Psi_L(\phi))}, \tag{8.33}$$

that is only valid if the stochastic time $Y(t)$ is independent of $L(x)$, but no longer valid if a non-zero correlation between $Y(t)$ and $L(x)$ is present. To evaluate the CF in the case with the correlation, we explain an approach of Carr and Wu (2004) in following.

Denote $\Psi(\phi)$ as the characteristic exponent of a Fourier transform such that $\mathbf{E}[e^{i\phi x(t)}] = e^{\Psi(\phi)}$.[3] Next, we introduce the Laplace transform that is defined by

$$g_x(\tau) \equiv \mathbf{E}[e^{-\tau x(t)}]$$

with τ as the Laplace parameter.

Carr and Wu (2004) showed that evaluating the generalized Fourier transform of the time-changed Lévy process $X(t)$ under the risk-neutral measure Q is equivalent to evaluating the Laplace transform of the random time $Y(t)$ under a complex-valued measure Q_ϕ. That in turn "can be used to reduce the the problem of finding the characteristic function of $X(t)$ in the original economy into the problem of finding it in an artificial economy that devoid of the leverage effect". More mathematically, a new equivalent complex-valued measures Q_ϕ with respect to Q may be defined by

$$M_\phi(t) = \frac{dQ_\phi}{dQ}(t) \tag{8.34}$$

with a complex-valued martingale

[3] Carr and Wu used $-\Psi(\phi)$ as the characteristic exponent to introduce complex-valued measure. As a consequence, some results and equations here seem different from the original ones in Carr and Wu (2004).

$$M_\phi(t) = \exp\left(i\phi X(t) - Y(t)\Psi_L(\phi)\right). \tag{8.35}$$

This new measure Q_ϕ is called the leverage-neutral measure. It can be shown that $M_\phi(t)$ is a well-defined complex-valued Q-martingale, and is dependent on the value of ϕ. Under the leverage-neutral measure Q_ϕ, we have

$$f_{X(t)}(\phi) \equiv \mathbf{E}^Q\left[e^{i\phi X(t)}\right] = \mathbf{E}^{Q_\phi}\left[e^{-Y(t)[-\Psi_L(\phi)]}\right] \equiv g^\phi_{Y(t)}(-\Psi_L(\phi)), \tag{8.36}$$

where $g^\phi_{Y(t)}(-\Psi_L(\phi))$ denotes the Laplace transform of $Y(t)$ under the complex-valued measure Q_ϕ. By introducing a new notation $\Psi^*_L(\phi) = -\Psi_L(\phi)$, the above relation may be expressed as

$$\mathbf{E}^Q\left[e^{i\phi X(t)}\right] = \mathbf{E}^{Q_\phi}\left[e^{-Y(t)\Psi^*_L(\phi)}\right] \Rightarrow f_{X(t)}(\phi) = g^\phi_{Y(t)}(\Psi^*_L(\phi)).$$

This means that through this special change of measure, we in fact change a Fourier transform under Q to a Laplace transform under Q_ϕ. In many cases where a leverage effect between the random time and the original Lévy process is present, it is more convenient to compute the Laplace transform $g^\phi_{Y(t)}(\Psi^*_L(\phi))$ under the measure Q_ϕ for $f_{X(t)}(\phi)$. In more details, the calculation of $g^\phi_{Y(t)}(\Psi^*_L(\phi))$ may be performed in analogy to the zero bond pricing,

$$g^\phi_{Y(t)}(\Psi^*_L(\phi)) = \mathbf{E}^{Q_\phi}\left[e^{-\int_0^T \Psi^*_L(\phi)y(t)dt}\right]. \tag{8.37}$$

Obviously, if $y(t)$ is a diffusion process, the Feynman-Kac theorem may be directly applied to solve the above expected value, as we have done repeatedly in stochastic volatility models and short rate models.

Similarly to the Doléans exponential of a diffusion process, for any complex value γ, the Doléans exponential of time-changed Lévy process $X(t)$ is given by

$$\mathscr{E}(\gamma X(t)) = \exp\left(\gamma X(t) + Y(t)\Psi^*_L(-i\gamma)\right),$$

which has an expected value of one. The complex-valued martingale $M_\phi(t)$ is then a special case of $\mathscr{E}(\gamma X(t))$, namely

$$M_\phi(t) = \mathscr{E}(i\phi X(t)).$$

Because correlation or dependence is a very nature of the processes of a same type, in order to specify a concrete measure for a correlation or dependence, it is reasonable to assume that the Lévy process $L(t)$ and the time-change process $y(t)$ are governed by a same type of distribution laws or Lévy measures. Therefore, if a Lévy process is purely continuous, i.e., is a Brownian motion, non-zero correlation can only be generated by a continuous component. Similarly, if a Lévy process is purely jump, non-zero correlation can only be generated by a pure jump. In particular, a correlated time change of a Lévy process with finite (infinite) activity may be only achieved by a subordinator with finite (infinite) activity. In following, we will take a mean-reverting square root process and a gamma-OU process as two special subordinators to show how to build up a leverage effect between $L(t)$ and $y(t)$, and how to deal with such leverage effect by applying the technique of the leverage-neutral measure.

Time-Changed By Square Root Process

We now consider a case where a mean-reverting square root process serves as a subordinator. As indicated above, a correlated time change of a Brownian motion can be only generated by another Brownian motion. Hence, we apply a mean-reverting square root process to time-change another Brownian motion that drives the dynamics of asset price $S(t)$.

We recall the Heston model again, and denote $L(t) = W_1(t)$ and $y(t) = V(t)$,

$$\begin{aligned} \frac{dS(t)}{S(t)} &= rdt + \sqrt{y(t)}dL(t), \\ dy(t) &= \kappa(\theta - y(t))dt + \sigma\sqrt{y(t)}dW_2(t), \\ dL(t)dW_2(t) &= \rho dt. \end{aligned} \tag{8.38}$$

The new time-changed Lévy process $X(t)$ is then given by

$$X(t) = L(Y(t)) = \int_0^t \sqrt{y(s)}dW_1(s).$$

By the relation given in (8.36), the CF of $X(t)$ under the risk-neutral measure Q is equivalent to the Laplace transform of $Y(t)$ under the complex-valued leverage-neutral measure Q_ϕ,

$$f_{X(t)}(\phi) = \mathbf{E}^Q\left[e^{i\phi X(t)}\right] = \mathbf{E}^{Q_\phi}\left[e^{-\frac{1}{2}Y(t)\phi^2}\right],$$

where $-\frac{1}{2}\phi^2$ is the characteristic exponent of a standard Brownian motion $L(t)$. The Radon-Nikodym derivative between the risk-neutral measure Q and the corresponding leverage-neutral measure Q_ϕ takes the following form,

$$\frac{dQ_\phi}{dQ}(t) = \exp\left(i\phi X(t) + \frac{1}{2}\phi^2 \int_0^t \sqrt{y(s)}ds\right).$$

By setting $\phi = -i$, we obtain the conventional Radon-Nikodym derivative

$$\frac{dQ_{-i}}{dQ}(t) = \exp\left(X(t) - \frac{1}{2}\int_0^t \sqrt{y(s)}ds\right).$$

Now we apply an extended Girsanov theorem under a complex-valued measure to the process $y(t)$, and obtain the following new Brownian motion,

$$dW^{Q_\phi}(t) = -i\phi\sqrt{y(t)} < dW_1(t), dW_2(t) > +dW_2(t).$$

It follows immediately a complex-valued process under Q_ϕ,

$$dy(t) = [\kappa\theta - \kappa y(t) + i\phi\rho\sigma y(t)]dt + \sigma\sqrt{y(t)}dW^{Q_\phi}(t), \tag{8.39}$$

where the drift term includes a complex value $i\phi\rho\sigma y(t)$. Under the measure Q_ϕ, computing

$$\mathbf{E}^{Q_\phi}\left[e^{-\frac{1}{2}Y(t)\phi^2}\right] = \mathbf{E}^{Q_\phi}\left[e^{-\frac{1}{2}\phi^2\int_0^t\sqrt{y(s)}ds}\right]$$

is then straightforward by using the Feynman-Kac theorem. More simply, we may directly apply the results given in (3.16) to obtain the final analytical solution for the CF of $X(t)$.

Obviously, in the light of time change, the Heston model may be interpreted as a time changed Brownian motion with drift where a mean-reverting square root process serves as a subordinator. This is a new insight that we gain from the leverage-neutral measure in the connection with a correlated time-changed Brownian motion. A further question is then, weather this time-change approach may be extended to other stochastic volatility models, for example, to the Schöble and Zhu model where stochastic volatility is specified to be a mean-reverting Ornstein-Uhlenbeck process? Since a mean-reverting Ornstein-Uhlenbeck process is Gaussian and may take negative values, it is not appropriate to be a subordinator. However, if we relax and even ignore the condition that a subordinator shall be a positive function, the above used technique may still be applied to other stochastic volatility models.

Time-Changed By Gamma-OU Process

To illustrate a leverage effect between a jump Lévy process $L(t)$ and its subordinator, and to generate the related asymmetry in the time-changed pure Lévy process, we consider a simplified Pareto model in Section 7.4 with a single Poisson counting process. To void possible notational conflicts, we rewrite the asset process driven by a compound Poisson process $q_1N(t)$ in the following form,

$$d\ln S(t) = [r - \mu - \frac{1}{2}v^2]dt + vdW + q_1dN(t) \tag{8.40}$$

with identical exponentially distributed jump sizes q_1. Therefore we have here

$$dL(t) = q_1dN(t) \quad \text{or} \quad L(t) = \sum_{n=1}^{N(t)} q_1(n).$$

On the other hand, we specify the subordinator $y(t)$ as a gamma-OU process,

$$dy(t) = [a - \kappa_y y(t)]dt + q_2dN(t), \tag{8.41}$$

where $N(t)$ is an identical Poisson process to the one specified in the asset price $S(t)$, the jump size q_2 is exponentially distributed, and has a joint distribution with q_1. The joint distribution $F(dq_1, dq_2)$ determines the dependence of q_1 and q_2. Since the both processes $L(t)$ and $y(t)$ share the same Poisson component, the leverage effect comes solely from the correlated jump sizes.

Applying the time change to $L(t)$ by $Y(t)$ yields

$$X(t) = L(Y(t)) = \sum_{n=1}^{N(Y(t))} q_1(n).$$

Again by the relation given in (8.36), we should evaluate the following Laplace transform under the leverage-neutral measure Q_ϕ,

$$\mathbf{E}^{Q_\phi}\left[e^{-\Psi_L^*(\phi)Y(t)}\right]$$

with

$$\Psi_L^*(\phi) = \lambda\left(1 - \frac{1}{1 - i\phi\mu_1}\right),$$

where μ_1 is the mean of q_1.[4] The Radon-Nikodym derivative for the change from Q to Q_ϕ is then given by

$$\frac{dQ_\phi}{dQ}(t) = \exp\left(i\phi X(t) + \Psi_L^*(\phi)\int_0^t y(s)ds\right).$$

To obtain a concrete form of Q_ϕ, we consider the joint characteristic function $f(\phi_1, \phi_2)$ of q_1 and q_2, which is defined by

$$f(\phi_1, \phi_2) = \mathbf{E}\left[e^{i\phi_1 q_1 + i\phi_2 q_2}\right] = \int_{\mathbb{R}^+}\int_{\mathbb{R}^+} e^{i\phi_1 q_1 + i\phi_2 q_2} dF(dq_1, dq_2).$$

By setting

$$dF^*(dq_2) = \int_{\mathbb{R}^+} e^{i\phi_1 q_1} dF(dq_1, dq_2),$$

the joint characteristic function $f(\phi_1, \phi_2)$ can be rewritten as

$$f(\phi_1, \phi_2) = \int_{\mathbb{R}^+} e^{i\phi_2 q_2} dF^*(dq_2) = f^*(\phi_2),$$

which may be regarded as the CF of q_2 with the new distribution function $F^*(dq_2)$. Hence, $F^*(dq_2)$ defines a new measure that is identical to Q_{ϕ_1} under which the expected value of q_2 may be calculated as if it is independent of q_1. Formally, the leverage-neutral measure Q_{ϕ_1} implies the following equation

$$dF^{Q_{\phi_1}}(dq_2) = \int_{\mathbb{R}^+} e^{i\phi q_1} dF(dq_1, dq_2). \tag{8.42}$$

[4] For a detailed calculation of $\Psi_L^*(\phi)$, please see Section 7.4.

It follows

$$
\begin{aligned}
f_{X(t)}(\phi) &= \mathbf{E}^{Q_{\phi_1}}\left[e^{-\Psi_L^*(\phi)Y(t)}\right] \\
&= \int_{\mathbb{R}^+} e^{-\Psi_L^*(\phi)Y(t)} dF^{Q_{\phi_1}}(dq_2).
\end{aligned}
$$

Carr and Wu (2002) showed that the CF of $X(t)$ still admits an exponential-affine solution as follows:

$$
f_{X(t)}(\phi) = e^{\alpha(t)+\lambda\beta(t)}. \tag{8.43}
$$

Applying the ODEs (7.48) and (7.49) of the affine jump-diffusion model given in Section 7.6, $\alpha(t)$ and $\beta(t)$ are determined by the solutions of the ODEs,

$$
\beta^{'}(t) = -\Psi_L^*(\phi) + \kappa_y\beta(t) + \lambda[f_2^{Q_{\phi_1}}(\beta(t)) - 1],
$$

$$
\alpha^{'}(t) = -a\beta(t),
$$

where $f_2^{Q_{\phi_1}}(\beta(t))$ is the marginal moment-generating function of q_2 under Q_{ϕ_1},

$$
f_2^{Q_{\phi_1}}(\beta(t)) = \int_{\mathbb{R}^+} e^{\beta(t)q_2} dF^{Q_{\phi_1}}(dq_2).
$$

8.4 The Barndorff-Nielsen and Shephard Model

Barndorff-Nielsen and Schephard (BNS, 2001) proposed an asset price which is a mixture of a Brownian motion and a Lévy jump. Let $x(t) = \ln S(t)$, their mixed process of $x(t)$ is given by

$$
dx(t) = \left(r - \lambda - \frac{1}{2}\beta v^2(t)\right)dt + v(t)dW(t) + \rho dZ(t), \tag{8.44}
$$

$$
dv^2(t) = -\kappa v^2(t)dt + dZ(t), \tag{8.45}
$$

where $dW(t)$ and $dZ(t)$ are not correlated. $Z(t)$ is not a Brownian motion, instead a Lévy process describing the jumps of instantaneous asset returns. Additionally, stochastic variances $v^2(t)$ is also driven by $Z(t)$ only, and follows a gamma OU process. This is different from the most stochastic volatility models such as the Heston model. The parameter ρ is usually negative and links the inverse movements of the jumps in volatility and the jumps in price, and should not be interpreted as a correlation coefficient, but a leverage effect between volatility and asset price. In fact, some practical calibrations deliver a larger negative number for ρ (see Schoutens, Simons and Tistaert (2004)). Particularly, $Z(t)$ in the BNS model is specified to follow a gamma process with a mean of α and a variance of β in an unit time interval. The density function of a gamma distribution is

$$\nu(z) = \frac{c^{\alpha c}}{\Gamma(\alpha c)} e^{-cz} z^{\alpha c - 1}, \quad c = \frac{\alpha}{\beta}, \quad z > 0,$$

with $\Gamma(\cdot)$ as a gamma function. The discounted asset prices are martingale only if

$$\lambda = \kappa \int_{\mathbb{R}} e^{\rho z} \nu(dz) = \frac{\kappa \alpha \rho}{c - \rho}.$$

The CFs in the BNS model can be derived as follows:

$$\begin{aligned} f_1(\phi) &= \mathbf{E}[e^{-(x_0+r)T+(1+i\phi)x(T)}] \\ &= \exp\Big(i\phi x_0 + i\phi(r-\lambda)T - \frac{1}{2}\gamma(\phi^2 - i\phi)v_0^2\Big) \\ &\times \exp\Big(\frac{\alpha}{c - g_2(1+i\phi)}\big[\kappa g_2(1+i\phi)T + c\ln\big(\frac{c - g_1(1+i\phi)}{b-(1+i\phi)\rho}\big)\big]\Big) \\ &= e^{i\phi(x_0+rT)} f_1^{BNS}(\phi) \end{aligned} \tag{8.46}$$

and

$$\begin{aligned} f_2(\phi) &= \mathbf{E}[e^{i\phi x(T)}] \\ &= \exp\Big(i\phi x_0 + i\phi(r-\lambda)T - \frac{1}{2}\gamma(\phi^2 + i\phi)v_0^2\Big) \\ &\times \exp\Big(\frac{\alpha}{c - g_2(\phi)}\big[\kappa g_2(\phi)T + c\ln\big(\frac{c - g_1(\phi)}{b - i\phi\rho}\big)\big]\Big) \\ &= e^{i\phi(x_0+rT)} f_2^{BNS}(\phi) \end{aligned} \tag{8.47}$$

with

$$g_1(u) = iu\rho - \frac{1}{2}\gamma(u^2 + iu),$$

$$g_2(u) = iu\rho - \frac{1}{2}\kappa(u^2 + iu).$$

The function γ is already defined before in (8.32). Note that the BNS's original pricing formula for European-style options is based on $f_2(\phi)$ that is essentially the CF under the original risk-neutral measure. In fact, as mentioned in BNS, $Z(t)$ can also be driven by other Lévy pure jumps. Possible candidates for $Z(t)$ could be the normal inverse Gaussian process and the variance-gamma process given above.

8.5 Alpha Log-Stable Model

Alpha stable distribution is essentially a distribution family with four parameters. Specifying these parameters to particular values leads to various well-known distributions such as normal distribution, Lévy distribution and Cauchy distribution.

Denote the alpha stable distribution as $L(x,\alpha,\beta,\nu,c)$ with four parameters. These parameters can be roughly interpreted as follows:

- $\alpha \in (0,2]$ is a measure for kurtosis, and determines the weight of the tails of the distribution.
- $\beta \in [-1,1]$ is a measure for skewness. Positive and negative values of β generate positive and negative skew respectively. If β is equal to zero, the distribution is symmetric.
- $\mu \in \mathbb{R}$ is responsible for location, and characterizes the mean of the distribution.
- $c \in \mathbb{R}_+$ scales the distribution, and expresses the dispersion of the distribution.

Let $x \sim L(\alpha,\beta,\nu,c)$, the CF of x is given by

$$f(\phi) = \exp\left(i\phi\mu - |c\phi|^\alpha(1 - i\beta sign(\phi)A\right) \tag{8.48}$$

with

$$A = \begin{cases} \tan\left(\frac{1}{2}\alpha\pi\right) & \text{if} \quad \alpha \neq 1 \\ -\frac{2}{\pi}\ln|\phi| & \text{if} \quad \alpha = 1 \end{cases}.$$

Although the parameters α and β may be considered as the measures for kurtosis and skewness, respectively, the moments $\mathbf{E}[L^k], k \geqslant \alpha$, of the alpha stable distribution are not defined, and hence the kurtosis and skewness do not exist. If increments of a process are characterized by alpha stable law, it will, as expected, exhibit the properties such as self-similarity and infinite-divisibility. Except for the case of $\alpha = 2$ where alpha-stable distribution reduces to the normal distribution, the alpha stable distribution is leptokurtic and heavy-tailed. Due to its ability of capturing leptokurtic property of assets returns, the alpha stable process has been recognized by many economists and mathematicians to model and analyze empirical financial data (see Mandelbrot (1967), Voit (2003)), and recently to price options (see Popava and Ritchken (1998), McCulloch (1996), and, Hurst, Platen and Rachev (1999)). However, the undesired properties of the alpha stable process is its infinite moments, therefore also infinite price moments. As a result, call price can not be soundly evaluated under an alpha stable process without any restriction.

A simple version of the restricted alpha stable processes is the so-called finite moment log-stable process. This special alpha stable process initialized by McCulloch (1996) and developed by Carr and Wu (2003a) is specified with the maximum negative skewness $\beta = -1$. The intention of forcing the maximum negative skewness to an alpha stable process, as argued by Carr and Wu, is twofold: Firstly, for the case of $\beta = -1$, the process still retains the property of unrestricted alpha stable process allowing for fat tails, even for long maturities, and keeping asymmetric structure of the distribution unchanged. Especially, it generates a down-sloping skew as observed in typical equity option markets.[5] Secondly, although the moments of returns in the case of $\beta = -1$ are infinite, the prices have finite moments. Consequently, option prices with finite moment log-stable process are bounded.

[5] The down-sloping skew of implied volatilities in equity markets is typical, but is not the single smile pattern only. In many cases, we can often observe strong symmetric smile for short-term volatilities.

More in details, under the risk-neutral measure, the SDE of a finite moment log-stable process for stock price reads

$$dx(t) = (r-m)dt + dL(t;\alpha,\sigma), \quad \alpha \in (1,2), \tag{8.49}$$

where alpha stable distribution $L(t;\alpha,\sigma)$ has a location parameter μ equal to zero, and a scaling parameter c equal to σ. Namely, $L(t;\alpha,\sigma)$ is equivalent to $L(t,\alpha,-1,0,\sigma)$. The finite moment logs-table process given above is rather parsimonious, and characterized by two free parameters α and σ only. Since σ is a scaling parameter, The SDE of (8.49) can be rewritten as

$$dx(t) = (r-m)dt + \sigma dL(\alpha,1).$$

The drift compensator m enforcing the stock process to be a martingale can be easily verified as

$$m = \ln[f(\phi=-i)] = -\sigma^{\alpha}\sec\left(\frac{\alpha\pi}{2}\right). \tag{8.50}$$

Finally we can derive the CFs under two different martingale measures Q_1 and Q_2,

$$\begin{aligned} f_1(\phi) &= \mathbf{E}[g_1(T)e^{i\phi x(T)}] = \mathbf{E}[e^{-(x_0+rT)+(1+i\phi)x(T)}] \\ &= \exp\left(i\phi(x_0+rT) - (1+i\phi)mT - ((1+i\phi)\sigma)^{\alpha}T\sec\left(\frac{\alpha\pi}{2}\right)\right) \\ &= e^{i\phi(x_0+rT)}f_1^{ALS}(\phi) \end{aligned} \tag{8.51}$$

and

$$\begin{aligned} f_2(\phi) &= \mathbf{E}[e^{i\phi x(T)}] \\ &= \exp\left(i\phi(x_0+rT) - i\phi mT - (i\phi\sigma)^{\alpha}T\sec\left(\frac{\alpha\pi}{2}\right)\right) \\ &= e^{i\phi(x_0+rT)}f_2^{ALS}(\phi). \end{aligned} \tag{8.52}$$

The option pricing formula is then given correspondingly.

8.6 Empirical Performance of Various Lévy Processes

In the above sections we have examined a number of Lévy processes that are applied to capture the non-normality of assets returns, and consequently used to price European-style options. We have seen that various Lévy processes can be constructed in different ways. One powerful and illustrative way is the subordination of a Lévy process by another process. This procedure may be illustrated by random time-changing, and corresponds to nesting the CFs of two involved processes. With this method, we have formulated the variance-gamma process and the NIG process as well as their four extended processes subordinated to a mean-reverting square-root process and a gamma OU process respectively. Another way to construct Lévy process is the direct specification of the distribution of the increments of the process, for example, the gamma OU process in BNS (2001) and the finite moment logstable process in Carr and Wu (2003a).

In this section we summarize some empirical results for these Lévy processes. Particularly, comparative testings with respect to pure diffusion models and Poisson jump models are carried out in some empirical studies. Some stylized results are given in following.

1. The variance-gamma process and the NIG process have the similar performance in pricing options. The test results based both on the statistic inference and on the calibration (i.e. pure optimization) deliver the mean-square errors of a same order. The equivalent performance should not be surprising since the marginal distributions of both processes belong to the generalized hyperbolic distribution, see Carr, Geman, Madan and Yor (2001).
2. Whenever the variance-gamma process and the NIG process are subordinated to a same process, for example, to a mean-reverting square root process, they exhibit the similar pricing performance, as shown in Schoutens, Simons and Tistaert (2004).
3. With respect to the valuation of exotic options, as reported by Schoutens, Simons and Tistaert (2004), the paths dynamics of stochastic volatility models such as the Heston model are more changeful than theses of Lévy jump models. This is indicated by the fact that the prices of lookback options and down-and-in barrier options with the Heston model are significantly higher than the prices with Lévy jump models although all models are equally good at the valuation of European-style options. This result seems somehow surprising since it is always argued in financial literature that Lévy jumps are born to generate fat tails and negative skews in a distribution, and therefore would be also able to produce more extreme events. One conclusion that we can draw from the large price deviations between stochastic volatility models and Lévy jump models in the valuation of path-dependent options, is that a good fitting of the terminal distribution for a given maturity to market data for various models does not imply a same local behavior of prices in these models. So far, to my best knowledge, there is no literature discussing the hedging performance of different Lévy jump models, and carrying out the corresponding comparative studies to stochastic volatility models, in order to answer what type models may be preferred to in the valuation of exotic options.
4. An empirical study of Huang and Wu (2004) investigated what type of jump structure best describes the return and process movements within time changed Lévy processes by using the market data of S&P index options. Their main result is that the variance-gamma model and the finite moment logstable model outperform Poisson jump models. In other words, a Lévy process incorporating hight frequency jumps may be better than the one incorporating low frequency jumps in capturing the behavior of options. Therefore, Huang and Wu recommended that an appropriate pricing model should incorporate Lévy jump components.
5. The finite moment log-stable model with the maximum negative skew is designed to capture the down-sloping pattern (smirk) of volatility. However. due to its in-

flexible skew structure, it is impossible for finite moment log-stable model to generate a symmetric volatility smile that happens often for short maturity options in an indecisive market.

8.7 Monte-Carlo Simulation

8.7.1 Generating Random Variables

The much broader applications of a pricing model in practice is to value and hedge exotic structures consistently to plain-vanilla options. For this purpose, an efficient Monte-Carlo simulation for Lévy jump models is an important issue for practitioners. Glasserman (2004), Schoutens, Simons and Tistaert (2004) have discussed how to simulate a pure jump process. As a nature of Lévy process, the key to this simulation is to produce random increments at each time step by generating the desired random variables according to the marginal distribution law. For example, if we want to simulate a NIG process, we need to generate the NIG random variables at each time step, this is in analog to simulating a Brownian motion where normal random variables are generated. As indicated in the construction of some Lévy jump processes, the pure jump random variables are a composition of two or more random variables of different distribution laws. Therefore, it makes sense here to address first how to generate random variables of some basic distributions, they are random variables of Poisson distribution, gamma distribution, inverse Gaussian distribution, and alpha stable distribution.

Poisson Distribution

A Poisson distribution is defined on the set of non-negative integers, and its density function for $N = k$ is given by

$$p(N=k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0,1,2,\cdots, \tag{8.53}$$

where λ is the mean and controls the intensity of events recorded in an unit time. It is easy to see that

$$p(N=k+1) = p(N=k)\frac{\lambda}{k+1}. \tag{8.54}$$

The cumulative Poisson distribution function $F(k)$ is then the sum of $p(N=j)$ over all integers up to k,

$$F(k) = \sum_{j=0}^{k} p(N=j) = \sum_{j=0}^{k} e^{-\lambda}\frac{\lambda^j}{j!}. \tag{8.55}$$

Given the cumulative distribution function, we can apply the inverse transform method to generate Poisson random variables.[6] Namely, we first generate an uniform random variable U, and then find a x such that

$$x = k, \quad \text{if} \quad F(k) \leqslant U < F(k+1).$$

This idea implies the following detailed algorithm for generating Poisson random variables.

1. Set $j = 0, p = p(j = 0) = e^{-\lambda}, F = p$,
2. Generate an uniform random variable $U \sim Unif[0,1]$,
3. While $U > F$, then set $p = p\frac{\lambda}{j+1}, F = F + p, j = j + 1$,
4. Finally set $x = j$.

where in Step 3 we have utilized the property (8.54) of Poisson distribution to calculate the cumulative distribution $F(k)$ such that k is the critical integer satisfying the condition $F(k) \leqslant U < F(k+1)$.

Gamma distribution

As in the variance-gamma model, we assume that x has a gamma distribution $\gamma(\alpha,\beta)$ with a mean of α and a variance of β. The value $b = \beta/\alpha$ is the scaling parameter for the gamma distribution. This means that the random number $y, y \sim \gamma(\alpha/b, \beta/b^2)$ and the random number $x, x \sim \gamma(\alpha,\beta)$ have the same distribution. Therefore, we need only to consider a gamma random variable $x, x \sim \gamma(\alpha,\alpha)$ with a scaling parameter of one. As discussed in Fishman (1996), Glasserman (2000), a two-case strategy based on the so-called ratio-of-uniforms method may be an efficient way to generate gamma distributed random variables, the two cases that must be considered are $\alpha > 1$ and $\alpha \leqslant 1$.

Case 1: $a \leqslant 1$.

1. Set $a = (\alpha + e)/e$ with $e = \exp(1)$,
2. Generate $U_1 \sim Unif[0,1]$ and $U_2 \sim Unif[0,1]$, then compute $d = aU_1$,
3. If $d \leqslant 1$, then compute $z = d^{1/\alpha}$. If $U_2 < \exp(-z)$, accept z,
4. Otherwise if $d > 1$, then compute $-\ln((a-d)/alpha)$. If $U_2 \leqslant z^{\alpha-1}$, accept z,
5. If no appropriate z is found, repeat steps 2, 3, 4.

Case 2: $a > 1$.

1. Set $a = (\alpha - \frac{1}{6\alpha})/(\alpha - 1), c = \frac{2}{\alpha-1}$,
2. Generate $U_1 \sim Unif[0,1]$ and $U_2 \sim Unif[0,1]$, then compute $d = a\frac{U_2}{U_1}$ and $z = (\alpha-1)d$,
3. Compute $y = cU_1 + d - (c+2) + 1/d$. If $y \leqslant 0$, accept z,
4. Otherwise compute $y = c\ln U_1 - \ln d + (d-1)$. If $y \leqslant 0$, accept z,
5. If no appropriate z is found, repeat steps 2, 3, 4.

[6] More about inverse transform method for generating random variables, see Glasserman (2004).

Inverse Gaussian Distribution

Denote $IG(\alpha,\theta)$ as an inverse Gaussian distribution. As explained in the normal inverse Gaussian model, if x is distributed according to inverse Gaussian law, x can be interpreted as the first passage time for a fixed level b of a Brownian motion with a drift κ. The relationships among these four parameters have been already given in the normal inverse Gaussian model. Furthermore, the following fact holds for x,

$$\frac{(\kappa x-b)^2}{x} \sim \chi_1^2.$$

This implies that given a χ_1^2 distributed random variable, we may obtain a corresponding x. The algorithms of Michael, Schucany and Hass (1976) for generating inverse Gaussian random variables is based on this idea, and can be implemented as follows:

1. Set $a=\kappa^{-1}$, $c=(ab)^2$,
2. Generate a standard normal variable $X \sim N[0,1]$, and compute $X=X^2$,
3. Compute $y=a(b+\frac{1}{2}aX)+\sqrt{aX(b+aX/4)}$, and $p=b/(b+\kappa y)$,
4. Generate an uniform random variable $U \sim Unif[0,1]$,
5. If $U>p$, then return c/y, otherwise return y.

Alpha-Stable Distribution

Simulating the alpha-stable distributed random variables efficiently is more challenging. For an alpha stable random variable $y \sim L(\alpha,\beta,\nu,c)$, there is a corresponding alpha stable random variable $x \sim L(\alpha,\beta,0,1)$ such that

$$y=\nu+cx.$$

Hence, we only need to consider how to simulate x with a location parameter of 0 and a scaling parameter of 1. Here we give an algorithm initially developed by Chambers, Mallows and Stuck (1976) and then augmented by Weron (1996, 1996a), that can be easily implemented as follows:

1. Generate $U \sim Unif[0,1]$ and set $V=\frac{\pi}{2}(U-\frac{1}{2})$,
2. Generate an exponential random variable $E \sim Exp(1)$ with a mean of one,
3. If $\alpha \neq 1$, then calculate

   $$c_1 = \arctan(\beta \tan(\pi\alpha/2)/\alpha$$

   and

   $$c_2 = [1+\beta^2\tan^2(\pi\alpha/2)]^{1/(2\alpha)},$$

   and finally compute and return

$$x = \frac{c_2 \sin(\alpha(V+c_1))}{\cos(V)^{1/\alpha}} \Big(\frac{\cos(V-\alpha(V+c_1))}{E}\Big)^{(1-\alpha)/\alpha},$$

4. If $\alpha = 1$, compute and return

$$x = \frac{2}{\pi}\Big[(\frac{1}{2}\pi + \beta V)\tan(V) - \beta \ln\big(\frac{\frac{1}{2}\cos(V)E\pi}{\frac{1}{2}\pi + \beta V}\big)\Big].$$

8.7.2 Simulation of Lévy Process

A Lévy process is derived by its independent increments that are distributed according to certain probability law. If the increments of a Lévy process can be generated, the simulation of this Lévy process is straightforward. Given a time partition $\{t_0 = 0, t_1, t_2, \cdots, t_{h-1}, t_H\}$, the Euler scheme of an asset process $S(t)$ with the Lévy dynamics has a similar form as the one in the Black-Scholes model,

$$S(t_{h+1}) = S(t_h)\exp\big((r-m)\Delta_h + \Delta L(t_h)\big).$$

According to the concrete formulation of $L(t)$, we can generate $\Delta L(t_h)$ at each time step t_h by applying above discussed methods, for example, in the alpha log-stable model. However, as a result of subordination, most Lévy jump processes and time-changed Lévy processes are driven by the dynamics that must be generated by more than one single distributions. More effort and care are then required in the detailed implementation.

Variance-Gamma Process

In this case, we have $L(t) = VG(t, \mu, \sigma, \beta)$. Generally there are two approaches to simulating $VG(t, \mu, \sigma, \beta)$. The first one is to simulate the variance-gamma process as it is constructed, and can be implemented with the following steps,

1. Generate a gamma distributed random variable G with a mean of one and a variance of β,
2. Generate a standard normally distributed random variable W,
3. Compute $\Delta VG(t_h) = \Delta_h G\mu + \sqrt{\Delta_h G}\sigma W$.

The second approach is based on the fact, as shown in Madan, Carr and Chang (1998), that a variance-gamma process can be decomposed by two independent gamma processes, namely $VG(t, \mu, \sigma, \beta) = \gamma_1(t, \alpha_1, \beta_1) - \gamma_2(t, \alpha_2, \beta_2)$ with

$$\alpha_1 = \frac{1}{2}\sqrt{\mu^2 + \frac{2\sigma^2}{\beta}} + \frac{\mu}{2}, \quad \beta_1 = \alpha_1^2 \beta,$$

$$\alpha_2 = \frac{1}{2}\sqrt{\mu^2 + \frac{2\sigma^2}{\beta}} - \frac{\mu}{2}, \quad \beta_2 = \alpha_2^2 \beta.$$

In this representation of the variance-gamma process, γ_1 and γ_2 may be interpreted as up-movement jumps and down-movement jumps. The corresponding simulation is then straightforward.

1. Generate a gamma distributed random variable G_1 with mean of α_1 and variance of β_1,
2. Generate a gamma distributed random variable G_2 with mean of α_2 and variance of β_2,
3. Compute $\Delta VG(t_h) = \Delta_h(G_1 - G_2)$.

NIG Process

For the NIG model, we have $L(t) = NIG(t, \mu, \sigma, \kappa)$. According to the construction of the NIG process, the simulation of NIG increments is following:

1. Generate an inverse Gaussian distributed random variable G with a mean of $1/\kappa$ and a variance of $1/\kappa^3$,
2. Generate a standard normally distributed random variable W,
3. Compute $\Delta NIG(t_h) = \Delta_h G\mu + \sqrt{\Delta_h G}\sigma W$.

Time-changed Lévy Process

Two stochastic processes, the square-root process and the gamma O-U process, are usually employed to time-change a Lévy process. The simulation of a mean-reverting square root process has been already discussed in details in Chapter 4. The simulation of a gamma O-U process process is straightforward since its increments are Gamma distributed. To explain how to simulate a time-changed Lévy process, we take the variance-gamma process time-changed by a square root process for illustration.

1. Generate the increment of a mean-reverting square process $\Delta Y(t_h)$ at time t_h,
2. Generate a gamma distributed random variable G_1 with a mean of α_1 and a variance of β_1,
3. Generate a gamma distributed random variable G_2 with a mean of α_2 and a variance of β_2,
4. Compute $\Delta VG(t_h) = \Delta Y(t_h)(G_1 - G_2)$,

where the time length Δ_h is replaced by the increment $\Delta Y(t_h)$ of the mean-reverting square process $y(t)$.

Chapter 9
Integrating Various Stochastic Factors

In this chapter we summarize the different option pricing models discussed previously, and build up an unified comprehensive option pricing framework. We have generally two approaches to integrating various stochastic factors: the modular approach and the time-change approach. While the modular approach has a simple, clear and transparent structure, and is also designed for practical implementations, the approach based on time-change is more technical and may embrace some nested factors. To show the large accommodation of the new pricing framework with CFs, we use four volatility specifications, four interest rate specifications, four Poisson jump specifications and four Lévy jump specifications to develop $4 \times 4 \times 4 \times 4 = 256$ different option pricing models. In fact, the number of possible pricing models by assembling different stochastic factors via CFs is much larger. Additionally, it can be shown that the pricing kernels for option pricing and bond pricing are identical if volatility and interest rate are specified to follow a same stochastic process. Finally, from the various points of view of practical applications, I give some criterions for choosing "ideal" models.

9.1 Stochastic Factors as Modules

In the previous chapters we have discussed four different streams to improve the Black-Scholes model by incorporating stochastic volatilities, stochastic interest rates, Poisson jumps as well as Lévy jumps into an underlying asset process, and have also examined in an isolated manner how each stochastic factor can be added into option pricing formulas in terms of CF, and consequently changes the prices of European-style options. Now we are in a stage to integrate them into an unified option pricing framework.

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_9,

Stochastic Interests		**Stochastic Volatilities**		**Poisson Jumps**		**Lévy Jumps**
deterministic (BS)		deterministic (BS)		No jumps (BS)		No jumps (BS)
Mean-reverting square root process (Cox, Ingersoll and Ross,1985)		Mean-reverting square root process* (Heston,1993)		Simple jumps (Cox and Ross,1976)		Variance Gamma (Madan, Carr and Chang,1988)
Mean-reverting Ornstein-Uhlenbeck process (Vasicek,1977)	×	Mean-reverting Ornstein-Uhlenbeck process (Schöbel and Zhu,1999)	×	Lognormal jumps (Merton,1976)	×	Normal inverse Gaussian (Barndorff-Nielsen,1998)
Mean-reverting double square root process (Longstaff, 1989)		Mean-reverting double square root process* (first in this study, 2000)		Double-Pareto jumps (first in this study,2000)		Time-changed Levy process (CGMY, 2001)
				Double-exponential jumps (Kou, 2002)		Alpha log-stable process (Carr and Wu, 2002)

*: Squared volatilities (variances) are specified.

Table 9.1 Stochastic factors as modules. Stochastic interest rates, stochastic volatilities, Poisson jumps and Lévy jumps are integrated into an unified pricing framework.

In a comprehensive option pricing model, as shown in the previous chapters, we have four types of stochastic factors to specify stock price dynamics where jumps are classified as Poisson jumps and Lévy jumps for expositional reason. Since each stochastic factor may be specified in some alternative ways, the number of possible option pricing models is theoretically equal to the product of the number of possible specifications of each factor. Up to now, the most option pricing models are separately developed, examined and implemented. It is certainly tedious and inefficient to build up and discuss every possible model and to derive every possible option pricing formula. On the other hand, each possible option pricing model may be a good candidate to match steadily changing market environments, and to meet the needs of financial practitioners. The specification issue of option pricing model becomes much more important after various factors are separately successfully incorporated into option valuation formulas. In fact, instead of dealing with single stochastic factor, we only need to establish a flexible framework as illustrated in Table (9.1), to perceive new interesting models. As indicated in Chapter 2, we refer to this building-block method as the modular pricing approach. With the help of Fourier transform and the numerical techniques discussed in Chapter 4, the modular pricing will work very efficiently and creates a clear and transparent structure for the valuation of options.

Table (9.1) embraces many classical specifications for each stochastic factor in financial literature and may be regarded as a summing-up of the most existing option pricing models, and at the same time provides some new models by specifying factors in various ways. In details, we have the following particular specifications as modules.

1. Stochastic Interest Rate (SI): We adopt four alternative specifications for short rates, namely, deterministic as in the Black-Scholes model, a mean-reverting square-root process as in the CIR model (1985b), a mean-reverting Ornstein-Uhlenbeck process as in the Vasicek model (1977), as well as a double square root process in the Longstaff model (1989). We consider here only one-factor short rate models for simplicity. Extending them into multi-factor cases and more generally affine term structure models as in Longstaff and Schwartz (1992a, 1992b), Duffie and Kan (1996), is straightforward, as long as the corresponding zero-coupon bond pricing formula is analytically tractable.
2. Stochastic Volatility (SI): Similarly, the specification of volatility follows the same alternatives: a constant volatility, a mean-reverting square-root process, a mean-reverting Ornstein-Uhlenbeck process and a double square root process. Again, the deterministic case corresponds to the Black-Scholes model. In Heston's model, the most popular stochastic volatility model, variance instead of volatility follows a square-root process, and a closed-form solution is available. Schöbel and Zhu (1999) extended Stein and Stein's (1991) formulation with non-zero correlation, and derived a closed-form pricing formula for options in the case of a mean-reverting Ornstein-Uhlenbeck process. Specifying volatility as a double square root process is studied here for the first time and offers us an alternative way to value options with stochastic volatility.

3. Poisson Jump (PJ): We distinguish Poisson jumps in four cases: no jump as in the Black-Scholes model; simple jumps with constant jump size as in Cox and Ross (1977); lognormal jumps as in Merton (1979), as well as Pareto jumps that is discussed for the first time in this book to capture the abnormalities in stock return process. Duffie, Pan and Singleton (1999) applied a similar formulation to model the abnormalities in volatility. Additionally, Kou's jump model is also an alternative and may be regarded as an equivalence to a Pareto jump model.
4. Lévy Jump (LJ): In phrase of this book, Lévy jumps are labeled as all jumps characterized by infinite activity, and exclude Poisson jumps. Intuitively, this type of jumps are driven by many small jumps. We have three representative specifications for Lévy jumps: the variance-gamma process as in Madan and Milne (1991) and Madan, Carr and Chang (1998), the normal inverse Gaussian model as in Barndorff-Nielsen (1998), as well as the alpha logstable model as in Carr and Wu (2003a). One of the appealing features of the above Lévy models is simple and analytical tractability. Besides, all time-changed Lévy models may be included in the modular pricing framework. To save space, we do not take them into account here.

Nevertheless, there are a number of other possible modelings for each given factor. As long as the CF under a particular specification is analytically available, it may be added to the list in Table (9.1), and embedded into the unified pricing framework.

9.2 Integration Approaches

To integrate four different stochastic factors into an unified framework, we have two approaches. The first one is the modular approach, the second one is the time-change approach. While the modular approach has a simple, clear and transparent structure, and is also designed for practical implementations, the approach based on time-change is more technical and may embrace some nested factors. Additionally, the time-change approach provides some interesting insights for the mechanism of constructing a comprehensive asset process.

9.2.1 Modular Approach

In Chapter 7, we have discussed some jump-diffusion models that take a general process form given in (7.7). Recall the process (7.7) and rewrite it in a logarithm form,

$$dx(t) = \left[r(t) - \frac{1}{2}v^2(t) - \lambda \mathbf{E}[J]\right] dt + v(t)dW_1(t) + JdY(t),$$

with which stochastic interest rates, stochastic volatilities and Poisson jumps may be incorporated in a valuation framework. Now, we extent this process by adding a

Lévy jump component and obtain the following process,

$$dx(t) = \left[r(t) - m - \frac{1}{2}v^2(t) \right] dt + v(t)dW_1(t) + JdY(t) + dL(t), \qquad (9.1)$$

where $L(t)$ represents the Lévy jumps discussed in Chapter 8. The drift compensator m adjusts the above process to be a martingale, and is equal to

$$m = \lambda \mathbf{E}[J] + \ln \mathbf{E}[e^{L(1)}].$$

The extended process then includes all stochastic factors that we want to unify in a single valuation framework, and is a natural development of the traditional jump-diffusion process. In fact, this new process is a mixture of Itô process and Lévy process if stochastic volatilities are present. Obviously, if $v(t)$ is specified to be stochastic and also correlated with $W_1(t)$, then $x(t)$ is unlikely to be a Lévy process, but indeed an Itô process. As a whole, $x(t)$ includes a diffusive part driven by Itô process, and a jump part driven by Lévy process. If $v(t)$ reduces to a constant, $x(t)$ becomes a Lévy type, and particularly may be characterized by the Lévy-Khintchine theorem.

To make the new option valuation framework tractable, we assume

1. Either SV or SI is correlated with the underlying stock returns, i.e., with $W_1(t)$. If SI is nested with $W_1(t)$, then SV has to be constant. If a non-zero correlation between SV and $W_1(t)$ is present, SI must be uncorrelated with $W_1(t)$. Additionally, the stock price process has to be modified to (6.14) for the cases that SI follows mean-reverting square root or double square root process.
2. SV, SI, PJ and LJ are mutually stochastically independent.

These two conditions are not restrictive because the most of the existing option pricing models are not beyond this framework.

In the light of the property of CF given in (2.27), the general CFs of an option pricing formula in the extended framework can be expressed by

$$f_j(\phi) = e^{i\phi x_0} f_j^{SV}(\phi) \times f_j^{SI}(\phi) \times f_j^{PJ}(\phi) \times f_j^{LJ}(\phi), \quad j = 1, 2. \qquad (9.2)$$

The scheme in (9.2) provides us with the basic idea of the modular pricing in establishing more sophisticated option pricing models. Now we are able to assemble various CFs by putting different stochastic modules together, and have a least the following choices.

Stochastic Volatilities:

$$f_1^{SV} = \begin{cases} (2.46), & v(t) \quad \text{as constant (BS);} \\ (3.16), & V(t) = v(t)^2 \quad \text{as square-root process (Heston);} \\ (3.32), & v(t) \quad \text{as Ornstein-Uhlenbeck process (SZ);} \\ (3.42), & V(t) = v(t)^2 \quad \text{as double square root process;} \end{cases}$$

and

$$f_2^{SV} = \begin{cases} (2.48), & v(t) \quad \text{as constant (BS);} \\ (3.16), & V(t) = v(t)^2 \quad \text{as square-root process (Heston);} \\ (3.36), & v(t) \quad \text{as Ornstein-Uhlenbeck process (SZ);} \\ (3.44), & V(t) = v(t)^2 \quad \text{as double square root process.} \end{cases}$$

Stochastic Interests (Zero-Correlation Case):

$$f_1^{SI} = \begin{cases} \exp(i\phi rT), & r(t) \quad \text{as constant (BS);} \\ (6.11), & r(t) \quad \text{as square root process (CIR);} \\ (6.24), & r(t) \quad \text{as Ornstein-Uhlenbeck process (Vasicek);} \\ (6.28), & r(t) \quad \text{as double square root process (Longstaff);} \end{cases}$$

and

$$f_2^{SI} = \begin{cases} \exp(i\phi rT), & r(t) \quad \text{as constant (BS);} \\ (5.15), & r(t) \quad \text{as square root process (CIR);} \\ (6.24), & r(t) \quad \text{as Ornstein-Uhlenbeck process (Vasicek);} \\ (6.30), & r(t) \quad \text{as double square root process (Longstaff).} \end{cases}$$

Stochastic Interests (Correlation Case):

$$f_1^{SI} = \begin{cases} (6.17), & r(t) \quad \text{as square root process (CIR);} \\ (6.20), & r(t) \quad \text{as Ornstein-Uhlenbeck process (Vasicek);} \\ (6.31), & r(t) \quad \text{as double square root process (Longstaff);} \end{cases}$$

and

$$f_2^{SI} = \begin{cases} (6.18), & r(t) \quad \text{as square root process (CIR);} \\ (6.23), & r(t) \quad \text{as Ornstein-Uhlenbeck process (Vasicek);} \\ (6.32), & r(t) \quad \text{as double square root process (Longstaff).} \end{cases}$$

Poisson Jumps:

$$f_1^{PJ} = \begin{cases} nil, \quad \text{not specified (BS);} \\ (7.16), \quad JdY(t) \quad \text{as simple jump (Cox and Ross);} \\ (7.23), \quad JdY(t) \quad \text{as lognormal jumps (Merton);} \\ (7.33), \quad JdY(t) \quad \text{as Pareto jump first in this study;} \\ (7.33) \text{ and } (7.43), \quad JdY(t) \quad \text{as double exponential jump (Kou),} \\ \qquad \text{equivalent to Pareto jump;} \end{cases}$$

and

$$f_2^{PJ} = \begin{cases} nil, & \text{not specified (BS);} \\ (7.18), & JdY(t) \quad \text{as simple jump (Cox and Ross);} \\ (7.24), & JdY(t) \quad \text{as lognormal jumps (Merton);} \\ (7.35), & JdY(t) \quad \text{as Pareto jump first in this study;} \\ (7.35) \text{ and } (7.43), & JdY(t) \quad \text{as double exponential jump (Kou),} \\ & \text{equivalent to Pareto jump.} \end{cases}$$

Lévy Jumps:

$$f_1^{LJ} = \begin{cases} nil, & \text{not specified (BS);} \\ (8.16), & L(t) \quad \text{as variance-gamma (MCC);} \\ (8.23), & L(t) \quad \text{as normal inverse Gaussian (Barndorff-Nielsen);} \\ (8.51), & L(t) \quad \text{as alpha log-stable (Carr and Wu);} \end{cases}$$

and

$$f_2^{LJ} = \begin{cases} nil, & \text{not specified (BS);} \\ (8.17), & L(t) \quad \text{as variance-gamma (MCC);} \\ (8.24), & L(t) \quad \text{as normal inverse Gaussian (Barndorff-Nielsen);} \\ (8.52), & L(t) \quad \text{as alpha log-stable (Carr and Wu).} \end{cases}$$

By using the above integration scheme, each stochastic factor can be arbitrarily inserted into or withdrawn from the pricing formula depending on the users. Practitioners will benefit from this modular pricing by building a flexible pricing formula to adapt the changing market environments. Now we list some new special models for options to illustrate the power of the modular pricing. An interesting new one is a model including SV, SI and lognormal PJ where first two factors are characterized by a mean-reverting Ornstein-Uhlenbeck process. This model is completely parallel to the BCC model (1997) which specifies SV and SI as mean-reverting square-root processes. In accordance with the integration scheme, we immediately obtain the following CFs for this model:

$$f_1(\phi) = e^{i\phi x_0} f_1^{SV}(3.32) \times f_1^{SI}(6.24) \times f_1^{PJ}(7.23) \tag{9.3}$$

and

$$f_2(\phi) = e^{i\phi x_0} f_2^{SV}(3.36) \times f_2^{SI}(6.24) \times f_2^{PJ}(7.24). \tag{9.4}$$

To give another example, an option valuation model with SV as a double square root process, and with RJ as simple jump as well as with constant interest rate may has the following CFs:

$$f_1(\phi) = e^{i\phi(x_0+rT)} f_1^{SV}(3.42) \times f_1^{PJ}(7.16) \tag{9.5}$$

and

$$f_2(\phi) = e^{i\phi(x_0+rT)} f_2^{SV}(3.44) \times f_2^{PJ}(7.18). \tag{9.6}$$

We can also conveniently combine stochastic volatility and Lévy jump, such a model is certainly of interest for practical applications due to its strong ability for smile modeling. A particular example may be a combination of the Schöbel-Zhu model and the variance-gamma model, whose CFs are simply composed of the following ingredients,

$$f_1(\phi) = e^{i\phi(x_0+rT)} f_1^{SV}(3.32) \times f_1^{LJ}(8.16) \tag{9.7}$$

and

$$f_2(\phi) = e^{i\phi(x_0+rT)} f_2^{SV}(3.36) \times f_2^{LJ}(8.17). \tag{9.8}$$

By the modular pricing, more new special options pricing formulas may be continuously listed in this way. We do not deal with all of the possible models. In Table (9.3), 108 of 256 possible models are illustrated and the corresponding option prices are computed for different moneynesses. In a practical implementation, every element in the CFs $f_j(\phi)$ can be programmed as a module or a procedure. By assembling these modules, we can efficiently compute the probabilities F_j in all possible option pricing models via the Fourier inversion where a step-by-step Gaussian quadrature routine is applied. The detailed numerical techniques have already discussed in Chapter 4. Here I give again some prototype programming codes for the modular pricing based on the codes given in Section 4.5. The strike vector integration is implemented in codes to calculate a range of option prices.

```
\\This function implements the exponents of CFs
\\of stochastic volatility models.
Complex CF_Vola(int index, double phi, ProcessType type)
{
    ....
}

\\This function implements the exponents of CFs
\\of stochastic interest models.
protected Complex CF_Rate(int index, double phi, ProcessType type)
{
    ....
}

\\This function implements the exponents of CFs of Poisson jumps.
Complex CF_Poisson(int index, double phi, JumpType type)
{
    ....
}

\\This function implements the exponents of CFs of Poisson jumps.
Complex CF_Levy(int index, double phi, LevyType type)
{
    ....
}

//This function performs strike vector integration, two CFs version
double[] CFs(int index, double phi, double[] strikes)
{
    //If index is 0, we go to one CF version
    if ( index = 0 )
       return CFs(phi, strikes);

    Complex value = new Complex(0);
```

```
        value = CF_Vola(index, phi) + CF_Rate(index, phi)
              + CF_Poisson(index, phi) + CF_Levy(index, phi);

        double[] x = new double[strikes.Length];
        Complex y, v;
        for (int i = 0; i < strikes.Length; i++)
        {
            y = new Complex(0, Log(spot / strikes[i]) * phi, 0);
            v = value + y;
            x[i] = Exp(v.Real) * Sin(v.Imaginary) / phi;
        }
        return x;
    }

    //This function performs strike vector integrations, one CF version
    //according to Attari's formulation
    double[] CFs(double phi, double[] strikes)
    {
        Complex value= new Complex(0);
        double r = interestRate;
        Complex lns = new Complex(0, ( Log(spot)+ r*t) * phi);

        // Note Attari formulation is not valid for stochastic interest rate
        // CF_Rate is not integrated
        value = CF_Vola(2,  phi) +  CF_Jump(2, phi) + CF_Levy(2, phi);
        value = value.Exp();

        double real = value.Real;
        double imag = value.Imaginary;

        double[] x = new double[strikes.Length];
        double beta, v;
        for (int i = 0; i < strikes.Length; i++)
        {
           beta = Log( Exp(-r * t) * strikes[i] / spot) * phi;
           v = (real + imag / phi) * Cos(beta) + (imag - real / phi) * Sin(beta);
           x[i] = v/(1+phi*phi);
        }

        return x;
    }
```

The option prices of the various models given in Table (9.3) are computed with the parameters given in Table (9.2), and serve as a theoretical illustration only. We choose the values of model parameters so that all ATM option prices are close to each other as possible. Keeping ATM option prices at same level, we can easily conclude which models are able generate more skewness than the Black-Scholes model. For example, both the Heston model encoded as $(SR \times No \times No \times No)$ and the Schöbel-Zhu model encoded as $(OU \times No \times No \times No)$ with negative correlation produce the higher prices for ITM options and the lower prices for OTM options than the Black-Scholes model, this indicates that both stochastic volatility models can generate a down-sloping smile. The alpha log-stable model of Carr and Wu (2003a) is labeled as $(BS \times No \times No \times ALS)$, and exhibits the similar price pattern as stochastic volatility models due to its maximal pre-setting of skewness parameter. Not surprisingly, adding additional stochastic factor to the Black-Scholes model produces higher prices than the pure Black-Scholes prices, for example, a model with constant volatility and the variance-gamma encoded as $(BS \times No \times No \times VG)$.

	$S = 100$	$T = 1$	r_0=0.05	v_0=0.2
SV:	κ_{SR}=3	θ_{SR}=0.04	σ_{SR}=0.1	ρ_{SR}=-0.5
	κ_{OU}=3	θ_{OU}=0.195	σ_{OU}=0.1	ρ_{OU}=-0.5
SI:	κ_{SR}=2	θ_{SR}=0.05	σ_{SR}=0.1	ρ_{SR}=0
	κ_{OU}=2	θ_{OU}=0.05	σ_{SR}=0.1	ρ_{SR}=0
Poisson:	λ_{LogN}=0.2	μ_{LogN}=0.15	σ_{LogN}=0.05	
	$\lambda_{Pareto1}$=0.5	$\mu_{Pareto1}$=0.2	$\lambda_{Pareto2}$=0.3	$\mu_{Pareto2}$=0.25
Lévy:	β_{VG}=0.1	μ_{VG}=0.15	σ_{VG}=0.05	
	κ_{NIG}=5	μ_{NIG}=0.2	σ_{NIG}=0.15	
	α_{ALS}=1.8	σ_{ALS}=0.03		

Table 9.2 Parameter values for the selected models. The theoretical values of options are given in Table 9.3.

9.2.2 Time-Change Approach

Huang and Wu (2004) proposed a time-changed Lévy process to unify some option pricing models, in order to examine the relative performance of jump component and diffusion component in capturing volatility smile. In details, their formulation for the underlying return process takes the following form,

$$x(t) = x_0 + rt + \left(vW_{Y^d(t)} - \frac{1}{2}v^2 Y^d(t)\right) + \left(L_{Y^j(t)} - mY^j(t)\right), \tag{9.9}$$

where the volatility v and the interest rate r are assumed to be constant, $W(t)$ and $L(t)$ denote a standard Brownian motion and a compensated pure Lévy jump process respectively. Following the time-change approach to Lévy process in Section 8.3, Huang and Wu applied two subordinators $Y^d(t)$ and $Y^j(t)$ to time-change the Brownian motion $W(t)$ and the pure jump process $L(t)$, respectively.

As discussed in Section 8.3, stochastic volatility models, particularly, the Heston model, may be created via applying stochastic time change to Brownian motion, and the most Lévy models in financial literature are resulted from stochastic time change, with or without leverage effect. In this sense, the time-change approach is a feasible method to unify the most option pricing models in this book and in the vast financial literature. So far, as shown in Chapter 8 and the assumption here, stochastic interest rates are never taken into account in the time-change approach. However, with some slight extensions, we may incorporate independent stochastic interest rates discussed in Chapter 6 into the process in (9.9). For discussion simplicity, we neglect stochastic interest rates here.

Similarly to the procedure in Section 8.3, the generalized Fourier transform is given by

SV	SI	Poisson	Lévy	K=80	K=90	K=95	K=100	K=105	K=110	K=120
BS	No	No	No	24.59	16.70	13.35	10.45	8.02	6.04	3.25
SR	No	No	No	24.69	16.80	13.41	10.47	7.98	5.95	3.09
OU	No	No	No	24.78	16.88	13.46	10.47	7.93	5.84	2.92
BS	SR	No	No	24.98	17.07	13.69	10.76	8.29	6.27	3.40
SR	SR	No	No	25.07	17.17	13.76	10.78	8.26	6.18	3.24
OU	SR	No	No	25.17	17.25	13.81	10.79	8.21	6.08	3.08
BS	OU	No	No	24.98	17.10	13.73	10.82	8.36	6.34	3.46
SR	OU	No	No	25.07	17.19	13.80	10.84	8.33	6.26	3.31
OU	OU	No	No	25.17	17.28	13.85	10.84	8.28	6.16	3.15
BS	No	LogN	No	24.70	16.95	13.67	10.84	8.45	6.50	3.69
SR	No	LogN	No	24.80	17.03	13.72	10.84	8.41	6.41	3.55
OU	No	LogN	No	24.89	17.11	13.76	10.83	8.35	6.31	3.40
BS	SR	LogN	No	25.08	17.31	14.00	11.14	8.72	6.72	3.85
SR	SR	LogN	No	25.18	17.40	14.06	11.15	8.68	6.64	3.71
OU	SR	LogN	No	25.27	17.47	14.10	11.14	8.62	6.54	3.56
BS	OU	LogN	No	25.09	17.34	14.05	11.19	8.78	6.79	3.91
SR	OU	LogN	No	25.18	17.42	14.10	11.20	8.74	6.71	3.77
OU	OU	LogN	No	25.27	17.50	14.14	11.20	8.69	6.61	3.62
BS	No	Pareto	No	24.68	16.91	13.62	10.77	8.37	6.39	3.55
SR	No	Pareto	No	24.78	16.99	13.67	10.77	8.32	6.31	3.41
OU	No	Pareto	No	24.87	17.07	13.70	10.76	8.26	6.20	3.25
BS	SR	Pareto	No	25.07	17.27	13.95	11.07	8.63	6.62	3.70
SR	SR	Pareto	No	25.16	17.36	14.01	11.08	8.60	6.54	3.56
OU	SR	Pareto	No	25.25	17.43	14.05	11.07	8.54	6.44	3.41
BS	OU	Pareto	No	25.07	17.30	14.00	11.13	8.70	6.69	3.77
SR	OU	Pareto	No	25.25	17.39	14.05	11.14	8.66	6.61	3.63
OU	OU	Pareto	No	25.25	17.46	14.09	11.13	8.60	6.51	3.48
BS	No	No	VG	24.74	17.01	13.73	10.88	8.49	6.51	3.67
SR	No	No	VG	24.84	17.09	13.78	10.89	8.45	6.43	3.53
OU	No	No	VG	24.93	17.16	13.81	10.89	8.39	6.33	3.38
BS	SR	No	VG	25.13	17.37	14.06	11.19	8.75	6.74	3.82
SR	SR	No	VG	25.22	17.45	14.12	11.20	8.72	6.66	3.69
OU	SR	No	VG	25.30	17.53	14.16	11.20	8.67	6.57	3.54
BS	OU	No	VG	25.13	17.40	14.11	11.24	8.81	6.80	3.88
SR	OU	No	VG	25.22	17.48	14.16	11.26	8.78	6.73	3.75
OU	OU	No	VG	25.30	17.55	14.20	11.25	8.73	6.64	3.61
BS	No	LogN	VG	24.87	17.26	14.04	11.26	8.90	6.94	4.09
SR	No	LogN	VG	24.95	17.33	14.08	11.26	8.86	6.87	3.96
OU	No	LogN	VG	25.04	17.39	14.11	11.24	8.80	6.77	3.83
BS	SR	LogN	VG	25.24	17.61	14.37	11.55	9.16	7.17	4.25
SR	SR	LogN	VG	25.33	17.69	14.42	11.56	9.12	7.10	4.12
OU	SR	LogN	VG	25.41	17.75	14.45	11.55	9.07	7.01	3.99
BS	OU	LogN	VG	25.25	17.64	14.42	11.61	9.22	7.23	4.31
SR	OU	LogN	VG	25.33	17.71	14.46	11.61	9.18	7.16	4.19
OU	OU	LogN	VG	25.41	17.77	14.49	11.60	9.13	7.07	4.06
BS	No	Pareto	VG	24.85	17.22	13.99	11.19	8.81	6.84	3.94
SR	No	Pareto	VG	24.93	17.29	14.03	11.19	8.77	6.76	3.82
OU	No	Pareto	VG	25.02	17.35	14.05	11.17	8.71	6.67	3.68
BS	SR	Pareto	VG	25.22	17.57	14.32	11.49	9.08	7.07	4.10
SR	SR	Pareto	VG	25.31	17.65	14.36	11.49	9.04	6.99	3.98
OU	SR	Pareto	VG	25.39	17.71	14.39	11.48	8.99	6.90	3.85
BS	OU	Pareto	VG	25.23	17.60	14.36	11.54	9.13	7.13	4.16
SR	OU	Pareto	VG	25.31	17.67	14.40	11.54	9.10	7.06	4.04
OU	OU	Pareto	VG	25.39	17.73	14.43	11.53	9.05	6.97	3.91

continuing in the next page

continued from the previous page

SV	SI	Poisson	Lévy	K=80	K=90	K=95	K=100	K=105	K=110	K=120
BS	No	No	NIG	24.75	17.02	13.73	10.89	8.49	6.51	3.66
SR	No	No	NIG	24.84	17.10	13.79	10.90	8.45	6.43	3.53
OU	No	No	NIG	24.93	17.17	13.82	10.89	8.40	6.33	3.38
BS	SR	No	NIG	25.13	17.38	14.07	11.19	8.75	6.74	3.82
SR	SR	No	NIG	25.22	17.46	14.12	11.21	8.72	6.66	3.68
OU	SR	No	NIG	25.31	17.53	14.16	11.21	8.68	6.57	3.54
BS	OU	No	NIG	25.13	17.41	14.11	11.25	8.82	6.80	3.88
SR	OU	No	NIG	25.22	17.49	14.17	11.26	8.79	6.73	3.75
OU	OU	No	NIG	25.31	17.56	14.20	11.26	8.74	6.64	3.61
BS	No	LogN	NIG	24.87	17.27	14.05	11.26	8.90	6.94	4.09
SR	No	LogN	NIG	24.96	17.34	14.09	11.26	8.86	6.87	3.96
OU	No	LogN	NIG	25.04	17.40	14.12	11.25	8.80	6.77	3.83
BS	SR	LogN	NIG	25.25	17.62	14.38	11.56	9.16	7.17	4.25
SR	SR	LogN	NIG	25.34	17.69	14.42	11.56	9.13	7.10	4.12
OU	SR	LogN	NIG	25.42	17.76	14.45	11.55	9.08	7.01	3.99
BS	OU	LogN	NIG	25.25	17.65	14.42	11.61	9.22	7.23	4.31
SR	OU	LogN	NIG	25.34	17.72	14.46	11.62	9.19	7.16	4.18
OU	OU	LogN	NIG	25.42	17.78	14.49	11.61	9.14	7.07	4.05
BS	No	Pareto	NIG	24.85	17.23	13.99	11.19	8.81	6.84	3.94
SR	No	Pareto	NIG	24.94	17.30	14.03	11.19	8.77	6.76	3.81
OU	No	Pareto	NIG	25.02	17.36	14.06	11.18	8.72	6.67	3.68
BS	SR	Pareto	NIG	25.23	17.58	14.32	11.49	9.08	7.07	4.10
SR	SR	Pareto	NIG	25.32	17.65	14.37	11.50	9.04	6.99	3.98
OU	SR	Pareto	NIG	25.40	17.72	14.40	11.49	8.99	6.91	3.85
BS	OU	Pareto	NIG	25.23	17.61	14.37	11.54	9.14	7.13	4.16
SR	OU	Pareto	NIG	25.32	17.68	14.41	11.55	9.10	7.06	4.04
OU	OU	Pareto	NIG	25.40	17.74	14.44	11.54	9.05	6.97	3.91
BS	No	No	ALS	25.91	17.69	14.10	10.94	8.24	5.98	2.69
SR	No	No	ALS	26.00	17.80	14.19	10.99	8.23	5.92	2.54
OU	No	No	ALS	26.09	17.90	14.26	11.02	8.21	5.84	2.38
BS	SR	No	ALS	26.30	18.08	14.47	11.28	8.54	6.24	2.87
SR	SR	No	ALS	26.39	18.19	14.56	11.34	8.54	6.19	2.73
OU	SR	No	ALS	26.49	18.29	14.64	11.37	8.52	6.11	2.57
BS	OU	No	ALS	26.30	18.10	14.51	11.34	8.60	6.31	2.94
SR	OU	No	ALS	26.39	18.21	14.60	11.39	8.60	6.26	2.80
OU	OU	No	ALS	26.48	18.30	14.67	11.42	8.59	6.19	2.65
BS	No	LogN	ALS	26.00	17.91	14.40	11.31	8.66	6.44	3.16
SR	No	LogN	ALS	26.09	18.00	14.47	11.34	8.64	6.37	3.02
OU	No	LogN	ALS	26.18	18.09	14.53	11.36	8.61	6.29	2.88
BS	SR	LogN	ALS	26.39	18.29	14.76	11.64	8.95	6.69	3.34
SR	SR	LogN	ALS	26.48	18.39	14.83	11.68	8.94	6.64	3.21
OU	SR	LogN	ALS	26.57	18.47	14.90	11.70	8.92	6.56	3.07
BS	OU	LogN	ALS	26.39	18.31	14.80	11.69	9.01	6.76	3.41
SR	OU	LogN	ALS	26.47	18.41	14.87	11.73	9.00	6.70	3.28
OU	OU	LogN	ALS	26.56	18.49	14.93	11.75	8.98	6.63	3.14
BS	No	Pareto	ALS	25.98	17.87	14.35	11.24	8.58	6.34	3.02
SR	No	Pareto	ALS	26.07	17.97	14.42	11.28	8.56	6.28	2.89
OU	No	Pareto	ALS	26.16	18.06	14.48	11.30	8.53	6.20	2.74
BS	SR	Pareto	ALS	26.38	18.26	14.71	11.58	8.88	6.60	3.21
SR	SR	Pareto	ALS	26.47	18.35	14.79	11.62	8.87	6.54	3.08
OU	SR	Pareto	ALS	26.55	18.44	14.85	11.64	8.84	6.47	2.94
BS	OU	Pareto	ALS	26.37	18.28	14.75	11.63	8.94	6.67	3.28
SR	OU	Pareto	ALS	26.46	18.37	14.82	11.67	8.93	6.61	3.15
OU	OU	Pareto	ALS	26.55	18.46	14.88	11.69	8.90	6.54	3.01

Table 9.3 Theoretical values of European options in 108 different models by applying the modular approach.

$$
\begin{aligned}
f(\phi) &= \mathbf{E}^Q[e^{i\phi x(t)}] \\
&= e^{(x_0+rT)i\phi} \\
&\times \mathbf{E}^Q\Big[\exp\Big(i\phi[vW_{Y^d(t)} - \frac{1}{2}v^2Y^d(t)] + i\phi[L_{Y^j(t)} - mY^j(t)]\Big)\Big] \\
&= e^{(x_0+rT)i\phi}e^{-\psi_W^*(\phi)Y_d(t)-\psi_L^*(\phi)Y_j(t)},
\end{aligned} \tag{9.10}
$$

with

$$
\psi_W^*(\phi) = \frac{1}{2}v^2[i\phi + \phi^2],
$$

and $\psi_L^*(\phi)$ is the characteristic exponent whose form depends on the concrete specification of the pure Lévy process $L(t)$.

The complex-valued Radon-Nikodym derivative changing the risk-neutral measure Q to the leverage-neutral measure Q_ϕ is defined by

$$
\begin{aligned}
\frac{dQ_\phi}{dQ} = \exp\Big(&i\phi[vW_{Y^d(t)} - \frac{1}{2}v^2Y^d(t)] + \psi_W^*(\phi)Y_d(t) \\
&+ i\phi[L_{Y^j(t)} - mY^j(t)] + \psi_L^*(\phi)Y_j(t)\Big).
\end{aligned} \tag{9.11}
$$

Now we show how to incorporate the various stochastic factors into a single valuation framework by the time-change approach, and check three cases: stochastic volatilities, Poisson jumps and Lévy jumps.

Stochastic Volatilities:

1. The Black-Scholes model:

$$
y_d(t) = 1, \quad Y_d(t) = t.
$$

2. The Heston model:

$$
Y_d(t) = \int_0^t \sqrt{y_d(s)}ds,
$$

   where $y_d(t)$ follows under Q_ϕ a mean-reverting square root process as follows

$$
dy_d(t) = [\kappa\theta - \kappa y_d(t) + i\phi\rho\sigma y_d(t)]dt + \sigma\sqrt{y(t)}dW^{Q_\phi}(t).
$$

3. The Schöbel and Zhu model: In this case, a strictly positive stochastic time can not be defined. But if we relax the strictly positive condition for stochastic time, the technique of complex-valued measure is still applicable.

$$
Y_d(t) = \int_0^t y_d(s)ds,
$$

   where $y_d(t)$ follows under Q_ϕ a mean-reverting Ornstein-Uhlenbeck process as follows

$$dy_d(t) = [\kappa\theta - \kappa y_d(t) + i\phi\rho\sigma y_d(t)]dt + \sigma dW^{Q_\phi}(t).$$

4. The double square root model: This is similar to the Heston model.

$$Y_d(t) = \int_0^t \sqrt{y_d(s)}ds,$$

where $y_d(t)$ follows under Q_ϕ a mean-reverting double square root process as follows

$$dy_d(t) = [\frac{1}{4}\sigma^2 - \kappa\sqrt{y_d(t)} + (i\phi\rho\sigma - \lambda)y_d(t)]dt + \sigma\sqrt{y(t)}dW^{Q_\phi}(t).$$

Poisson Jumps:

If the Lévy process $L(t)$ itself is a compound jump process, the stochastic time-change $Y_j(t)$ is trivial,

$$y_j(t) = 1, \quad Y_j(t) = t.$$

Lévy Jumps:

If the Lévy process $L(t)$ is specified as a Lévy process with infinite activity, the stochastic time-change is not necessary,

$$y_j(t) = 1, \quad Y_j(t) = t.$$

However, if $L(t)$ is specified as a Brownian motion, we can time-change $L(t)$ with gamma process or inverse Gaussian process to obtain Lévy jumps, as discussed in Section 8.2. In this case, we have

$$Y_j(t) = \int_0^t y_j(s)ds,$$

where $y_j(t)$ is a gamma process or an inverse Gaussian process.

Time-changed Lévy Jumps:

This case is essentially identical to the one discussed in Section 8.3. If the stochastic time change $Y_j(t)$ is not correlated with $L(t)$, any subordinator $Y_j(t)$ may be applied to $L(t)$. If there is a dependence between $Y_j(t)$ and $L(t)$, they should be of the same type of Lévy process.

Summing up, although we can embrace different pricing models into a single framework in (9.9) by the time-change approach, but it is only of theoretical consideration, and lacks a structural transparency and clear instructions to implementation.

9.3 Pricing Kernels for Options and Bonds

In the previous chapters, we have amplified how to incorporate stochastic factors, especially SV and SI, into a comprehensive option pricing model. To some surprise, similar functions are repeatedly used in computing CFs as long as SV and SI are specified to follow a same stochastic process. More interestingly, these functions display the same structure as the zero-coupon bond pricing formulas. This indicates that the pricing kernels for options are connected with the pricing kernels for bonds. Additionally, as shown in Carr and Wu (2002), time-change approach in fact applies also the zero-coupon bond pricing analogy to compute the CF under the so-called leverage-neutral measure. To see this, we summarize the similarity between these two kernels case by case:

1. Mean-reverting square-root process $z(t)$:
   The key function in pricing both options and bonds is
   $$y(z_0,T) = \mathbf{E}\left[\exp\left(-s_1\int_0^T z(t)dt + s_2 z(T)\right)\right]. \tag{9.12}$$
   Setting $s_1 = 1$ and $s_2 = 0$ yields the CIR (1985) bond pricing formula while $y(z_0,T)$ plays a central role in deriving the CFs for options in the Heston model, as shown in Section 3.2 and Section 6.2.
2. Mean-reverting Ornstein-Uhlenbeck process $z(t)$:
   In this case, the following function is crucial for pricing options and bonds:
   $$y(z_0,T) = \mathbf{E}\left[\exp\left(-s_1\int_0^T z^2(t)dt - s_2\int_0^T z(t)dt + s_3 z^2(T)\right)\right]. \tag{9.13}$$
   Once again, setting $s_1 = 0$, $s_2 = 1$ and $s_3 = 0$ reduces $y(z_0,T)$ to the well-known bond pricing formula of Vasicek (1977) while $y(z_0,T)$ is a key function to arrive at a closed-form solution for CFs in the option pricing formula as in Schöbel and Zhu (1999).
3. Mean-reverting double square root process $z(t)$:
   In this case, we have a function that takes the following form:
   $$y(z_0,T) = \mathbf{E}\left[\exp\left(-s_1\int_0^T z(t)dt - s_2\int_0^T \sqrt{z(t)}dt + s_3 z(T)\right)\right]. \tag{9.14}$$
   We obtain the Longstaff (1989) bond pricing formula by setting $s_1 = 1$, $s_2 = 0$ and $s_3 = 0$, and the closed-form solution for CFs and options by using the general formula of $y(z_0,T)$.

All three functions $y(z_0,T)$ correspond to a PDE in accordance with the Feynman-Kac formula and are available analytically. This makes the modular pricing feasible. While zero-coupon bond prices naturally are the expected value of the exponential function of $r(t)$ (here, $z(t)$) by the argument of the local expectations hypothesis, the CFs also are the expected value of the exponential function of $r(t)$ or $v(t)$ by

Fourier transform. In the light of spanning security market, the pricing kernels for bonds and options are nested in different spanned market spaces. While the pricing kernels for bonds with an exponential form are located in the original market space, the pricing kernels for options with the function $y(z_0, T)$ are generated in the transformed space. Both spaces are interchangeable by applying simple operators such as differentiation or translation (see Bakshi and Madan, 2000). The last step in option pricing in our new framework is the Fourier inversion which transports the valuation of options from the transformed space into the original one. Thus, at the end of whole pricing procedure, option prices return to the market space spanned by underlying assets, and options are valued in the original market spaces.

Regardless of whether in the original market space or in the Fourier transformed one, Kolmolgov's backward equation works, this means, the probabilities F_j and the corresponding CFs satisfy the same backward equation. This conjunction can be seen in Section 2.2, and establishes the mathematical basis for the identical pricing kernel for options and bonds. Thus, it is no longer surprising that as long as we can obtain a closed-form formula for a zero-coupon bond by applying a particular process, we can also obtain a closed-form formula for options in a stochastic volatility model by applying the same process.

9.4 Criterions for Model Choice

In this chapter, we have discussed how to efficiently and flexibly build up a large number of option pricing models by combining different stochastic factors. To show the power of the modular approach, 256 possible variants of option pricing models are given by 4 alternative volatility factors, 4 alternative interest rate factors, 4 alternative Poisson jumps and 4 alternative Lévy jumps. Of course, adding new specification for each factor creates new potential pricing models. Two natural questions now are: (1) which factor is more important for option valuation? (2) what conditions shall a good pricing model satisfy? These two questions are related, and in general, the second question is more extensive than the first one.

The first question is also referred to as the specification problem of option pricing models, and is intensively addressed in financial literature. Specification issue is examined mostly with European-style options. For example, Bates (1996), Bakshi, Cao and Chen (1997) examined the role of stochastic volatility and Poisson jumps, and found that stochastic volatility is the most important factor in explaining the smile effect, and however, adding Poisson jumps improves the performance remarkably. Recently, some studies including Carr, Geman, Madan and Yor (2001), Carr and Wu (2003a), Huang and Wu (2004) documented the need for Lévy jumps, namely infinite activity jumps, in order to capture the behavior of equity index options. Most of these models are included in Table (9.3). Distinguishing a key factor from four stochastic factors is essentially an issue of empirical tests.

Now we pay more attentions on the second question from the point of view of practical applications. Generally, it is difficult and even impossible to choose the best model from so a large variety of models. However, if we apply an option pricing model to a real trading environment, we will naturally raise some requirements and expectations that a pricing model should satisfy. As a practitioner, I summarize the following points for an applicable pricing model.

1. Analytical Tractability: Analytical tractability is the first of all requirements that a pricing model should fulfill. A pricing model must admit an (semi-) analytical closed-form valuation for European-style options. It is not only used to value European-style options quickly and accurately, but also, more importantly, allows an efficient calibration of the model to market prices, in the most cases, the volatility surfaces. Without a closed-form solution for options, any pricing model does not have a chance to be applied in practice.
2. Fitting to Market Prices: A good model shall and can be fitted consistently to market quoted prices. Since most liquid plain-vanilla options are valued by the Black-Scholes (-like) models and quoted with implied volatilities, fitting to market prices is in most cases equivalent to fitting to volatility smile. From the point of view of practitioner, there are two aspects for fitting. Firstly, a model should be calibrated efficiently and accurately to whole smile surface, instead to a smile at one maturity only. This requires that a model can recover not only the long-term smile pattern, but also the short-term smile patten. Secondly, a good model should be flexible enough to capture different smile pattern, namely, symmetric, down-sloping and even up-sloping smile. Table (9.4) gives the market FX volatilities of USD/EUR, GBP/EUR and JPY/EUR, as of February 10. 2009. We can easily observe that FX volatility surface of each foreign exchange rate presents different smile pattern: While the volatility smile surface of USD/EUR is nearly symmetric, the smiles (sneers) of GBP/EUR and JPY/EUR slope up and down respectively. In these cases, the models that can only generate the decreasing or increasing skewness, can not recover all FX option markets. According to my experience, stochastic volatility models are the only model class that can deal with all three smile patterns firmly.
3. Hedging Performance: Hedging performance is an other important criterion to test a pricing model. Any pricing model delivers Delta and Gamma, but not necessarily provides Vega since volatility risk is replaced by other risks of parameters capturing volatility smile. For example, the variance-gamma model expresses volatility risk via the parameters μ, σ, β. Whether a pricing model provides a better and more robust hedge performance than the Black-Scholes model, is an issue of statistical testing and practical application. Usually, we may use out-of-sample or in-the-sample empirical testings for check hedging performance. However, there is a consensus that a good hedging performance is conditional on a good fitting to smile surface.
4. Convenient Simulation: Good fitting to market is the minimal requirement for an applicable model, and a perfect recovery of smile is not our initial goal. The most sophisticated models are used to value and hedge exotic derivatives and structures in accordance with volatility smile. However, the advanced models generally do not admit the analytical solutions for most exotic options. As a result, Monte-Carlo simulation is most time the only way to value exotic options.

Therefore, quick, efficient and robust simulation of a pricing model is crucial in practical applications. Particularly, a robust simulation should be feasible for barrier options, or other path-dependent options.

5. Easy Setup of Correlation Structure: A wide class of options traded in markets are the so-called correlation products, or basket options. To value such correlation products properly, we need a multi-dimensional extension of particular advanced one-dimensional model, equipped with a simple and reasonable correlation structure. But for the most models discussed in financial literature, it is difficult to propose simple correlated setup. In this sense, diffusive models are superior to pure jump models.
6. Parameter Parsimoniousness: Parameter parsimonious of a model is especially important for trading and risk management. Generally, every parameter in a pricing model is associated with a price uncertainty to it. This is the so-called model risk. As model parameters are estimated via optimizations or statistical methods every day, we have to gauge the price sensitivity with respect to each parameter. Too many parameters make a pricing model impossible to hedge the underlying asset risk, and make risk management unmanageable. A good model is then a balanced tradeoff of accurate fitting and parameter parsimoniousness.
7. Economic Interpretation: Finally, model parameters shall and can be interpreted in an economic and financial manner. Complicated models with abstract factors can not be understood and accepted by traders and risk managers, at least not for production purpose. The lack of an economic interpretation and a transparent structure often leads to a declining acceptance in practice. In contrast, the models with intuition may be more popular, even with some theoretical drawbacks. Hedging and risk management based on abstract factor models are less reliable and more dangerous.

Applying these criterions to the discussed models, stochastic volatility models, especially the Heston model, may be in most cases preferred to other option pricing models, at least from the point of view of a practitioner.

USD/EUR	ATM	DeltaCall15	DeltaPut15
1M	0.1870	0.2020	0.2037
2M	0.1840	0.2010	0.2014
3M	0.1840	0.2033	0.2034
4M	0.1823	0.2033	0.2063
5M	0.1801	0.2015	0.2063
6M	0.1775	0.1991	0.2048
9M	0.1730	0.1955	0.2027
1Y	0.1725	0.1956	0.1995
2Y	0.1660	0.1873	0.1997
3Y	0.1505	0.1718	0.1914

Symmetric smile. The DeltaCall volatilities are nearly equal to the DeltaPut volatilities.

GBP/EUR	ATM	DeltaCall15	DeltaPut15
1M	0.1970	0.2240	0.2051
2M	0.1860	0.2154	0.1957
3M	0.1790	0.2099	0.1896
4M	0.1700	0.2022	0.1813
5M	0.1682	0.2011	0.1798
6M	0.1685	0.2020	0.1804
9M	0.1640	0.1986	0.1766
1Y	0.1610	0.1967	0.1742
2Y	0.1588	0.1894	0.1696
3Y	0.1500	0.1807	0.1609

Up-sloping smile. The DeltaCall volatilities are larger than the DeltaPut volatilities.

JPY/EUR	ATM	DeltaCall15	DeltaPut15
1M	0.2600	0.2442	0.3193
2M	0.2575	0.2361	0.3260
3M	0.2575	0.2339	0.3342
4M	0.2468	0.2247	0.3262
5M	0.2474	0.2237	0.3299
6M	0.2513	0.2245	0.3375
9M	0.2450	0.2140	0.3380
1Y	0.2475	0.2147	0.3471
2Y	0.2325	0.2073	0.3240
3Y	0.2351	0.2099	0.3266

Down-sloping smile. The DeltaCall volatilities are smaller than the DeltaPut volatilities.

Table 9.4 Different volatility smile pattern in FX markets, as of February 10. 2009, Source: Reuters and SocGen. For the interpretation of Delta volatility, please see Section 1.3.

Chapter 10
Exotic Options with Stochastic Volatilities

Exotic option is a common name for a number of options either with an unconventional payoff structure or with a complicated probability structure (i.e., path-dependent options). There is a long list of financial derivatives belonging to this class: barrier options, Asian options, correlation options, spread options, exchange options, clique options etc. Most of them are generated in the course of the expansion of the financial derivative business since the 1970s, and are referred to as second generation options although some exotic options, for example, barrier options, are as old as standard European-style options, and are traded in over-the-counter (OTC) markets. Recently, this situation has somehow changed. The American Stock Exchange trades quanto options while the New York Mercantile Exchange provides spread options.

With financial risks being understood better, exotic options are more and more widely employed by financial institutions, big corporations and fund managers. Some exotic options have already become commonplace in risk management due to their case-orientated properties, and embedded and repackaged into bonds, swaps and other instruments. Most investment certificates tailored to private investors usually constitute bonds and exotic options. Generally speaking, options with a complicated exercise probability are path-dependent and therefore difficult to price. If stochastic volatility, stochastic interest rate and random jump are present, Monte-Carlo simulation is in most cases the single available and efficient way to price exotic derivatives. However, in the Black-Scholes world, the closed-form pricing formulas for most exotic options have been derived.[1]

In this chapter, we attempt to incorporate stochastic volatilities, stochastic interest rates and jumps into the pricing formulas for exotic options with Fourier transform. Since it is impossible to deal with exotic option in the unified valuation framework given in Chapter 9, we have to price every particular exotic option in a modified model. Additionally, we do not handle all possible variants of exotic options and concentrate mainly on some typical representatives such as barrier options and Asian options. We will first treat forward-starting options in Section 10.1, and

[1] See Zhang (1997) for a comprehensive collection.

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_10,

then deals with barrier options in Section 10.2. Lookback options follow in Section 10.3. Next, we consider Asian options in Section 10.4. Section 10.5 addresses correlation options including exchange options, quotient options and product options. In Section 10.6, we briefly discuss how to value some exotic options with unconventional payoff structures. In most cases we only take stochastic volatilities as an example to show how to value exotic options with Fourier transform. However, it must be pointed out that not all stochastic factors can be embedded in the considered cases. For instance, it is unknown to value barrier option analytically with Poisson and Lévy jumps. Because square root process and double square root process share the same properties in many respects, we do not consider double square root process in the chapter.

10.1 Forward-Starting Options

Forward-starting option is an option, as indicated by its name, that starts at a future time U, and matures at a further future time $T, T > U$. To void potential uncertainty that the stock price may volatile strongly during the forward period between now and time U, and to minimize the related risk that option may move in opposite to the interest of issuer or investor if the strike has been already fixed now, hence most forward-starting options are designed with a relative strike k, that means that the payoff of a forward-starting call option at maturity is given by

$$\max\left[S(T) - kS(U), 0\right].$$

As $S(U)$ is unknown now and then stochastic, the absolute strike $kS(U)$ is also stochastic. This feature distinguishes the forward-starting option from its plain-vanilla variant, and makes the valuation more involved, especially in stochastic volatility models. Kruse and Nögel (2005) examined how to value forward-starting options in the Heston model where the change of measure again provides a powerful tool to simplify the calculation of the final pricing formula remarkably.

By applying the risk-neutral valuation, a forward-starting call option can be priced as follows:

$$\begin{aligned} FWC(0;U,T) &= \mathbf{E}^Q\left[\exp\left(-\int_0^T r(u)du\right)\max\left[S(T) - kS(U), 0\right]\right] \\ &= \mathbf{E}^Q\left[\frac{1}{H(T)}\max\left[S(T) - kS(U), 0\right]\right] \end{aligned} \tag{10.1}$$

with $H(T) = \exp\left(\int_0^T r(u)du\right)$ as the money market account. Instead of working with the risk-neutral measure Q, we may value the option under the measure associated with the numeraire $S(t)$. As discussed in Chapter 2, the measure associated with the numeraire $S(t)$ has already referred to as the Delta measure Q_1. We recall the radon-Nikodym derivative between Q and Q_1,

$$\frac{dQ}{dQ_1}(t) = \frac{H(t)S_0}{S(t)},$$

we have immediately

$$\begin{aligned} FWC(0;U,T) &= \mathbf{E}^{Q_1}\left[\frac{S_0}{S(T)}\max\left[S(T)-kS(U),0\right] \mid \mathscr{F}_0\right] \qquad (10.2)\\ &= S_0\mathbf{E}^{Q_1}\left[\max\left[1-k\frac{S(U)}{S(T)},0\right] \mid \mathscr{F}_0\right], \end{aligned}$$

where we use the filtration $\mathscr{F}$ to emphasize the expectation conditional on a certain filtration. At the same time, under the measure Q_1 the processes of $x(t) = \ln S(t)$ and $V(t)$ are governed by the new SDEs,

$$\begin{aligned} dx(t) &= [r(t) + \frac{1}{2}V(t)]dt + \sqrt{V(t)}dW_1(t), \qquad (10.3)\\ dV(t) &= \kappa^*(\theta^* - V(t))dt + \sigma\sqrt{V(t)}dW_2(t),\\ dW_1(t)dW_2(t) &= \rho dt. \end{aligned}$$

with

$$\kappa^* = \kappa - \rho\sigma, \qquad \theta^* = \frac{\kappa\theta}{\kappa - \rho\sigma}.$$

Now we rewrite $FWC(0;U,T)$ in a form of iterated expectations,

$$\begin{aligned} FWC(0;U,T) &= \mathbf{E}^{Q_1}\left[\frac{S_0}{S(T)}\max\left[S(T)-kS(U),0\right] \mid \mathscr{F}_0\right]\\ &= S_0\mathbf{E}^{Q_1}\left[\mathbf{E}^{Q_1}\left[\frac{S(U)}{S(T)}\max\left[S(T)-kS(U),0\right] \mid \mathscr{F}_U\right]\frac{1}{S(U)} \mid \mathscr{F}_0\right],\\ &= S_0\mathbf{E}^{Q_1}\left[FWC(U;U,T) \mid \mathscr{F}_U\right]\frac{1}{S(U)} \mid \mathscr{F}_0\Big], \qquad (10.4) \end{aligned}$$

where $FWC(U;U,T)$ is no longer forward-starting conditional on $\mathscr{F}_U$, and may be valued in a conventional way. Assuming that $r(t)$ is constant for simplicity, we can value $FWC(U;U,T)$ as follows:

$$\begin{aligned} FWC(U;U,T) &= S(U)F_1(x(U)=0, K=k, V(U)=V)\\ &\quad - kS(U)e^{(T-U)}F_2(x(U)=0, K=k, V(U)=V), \end{aligned}$$

with the corresponding CFs

$$f_j^*(\phi) = f_j(\phi; x(U)=0, V(U)=V, \tau = T-U), \quad j = 1,2,$$

where $f_j(\phi)$ are the CFs of the corresponding standard European-style option. Obviously, the only uncertainty in $FWC(U;U,T)$ is $V(U)$. It follows then

$$\begin{aligned} FWC(0;U,T) &= S_0 \mathbf{E}^{Q_1}\left[F_1(x(U)=0, K=k, V(U)=V) \mid \mathscr{F}_0\right] \\ &\quad - kS_0 e^{(T-U)} \mathbf{E}^{Q_1}\left[F_2(x(U)=0, K=k, V(U)=V) \mid \mathscr{F}_0\right]. \end{aligned} \tag{10.5}$$

Given the conditional distribution of $V(t)$ as a non-central chi-squared distribution in the Heston model, the conditional expected values above may be computed with

$$F_j(V_0) = \mathbf{E}^{Q_1}\left[F_j(V(U)=V) \mid \mathscr{F}_0\right] = \int_{\mathbb{R}} F_j(V(U)) g(V(U) \mid V_0) dV(U),$$

that can be computed numerically. In the Schöbel-Zhu model, since the volatility is distributed according to a Gaussian law, the above expected values can be calculated more efficiently. The final pricing formula for a forward-starting option takes then the following form

$$FWC(0;U,T) = S_0 F_1(V_0) - kS_0 e^{(T-U)} F_2(V_0). \tag{10.6}$$

As shown by the above procedure, the pricing formula for forward-starting options in (10.5) is not restricted to the Heston model only, but also applicable to other stochastic volatility models, and even to more complicated models. However, if other stochastic factors are present, the final expected values in (10.5) should be involved in a multi-dimensional integration correspondingly.

10.2 Barrier Options

10.2.1 Introduction

Barrier options are path-dependent and can be classified as knock-in or knock-out based on the ratio of asset spot price to barrier level. For knock-out barrier options, if the prices of the underlying asset during the option's lifetime reach a certain predetermined barrier, then the option becomes worthless. Although barrier options have been traded in OTC marketplace since the 1960s, their innovative features had not been uncovered among financial investors until the late 1980s. Meanwhile, barrier options are very popular in FX markets. Hedging strategies using barrier options are normally associated with the purpose of protecting against extreme downside or upside risk. A possible application may be portfolio insurance strategy for a well-endowed investor who only need to hedge against big unexpected price risks.[2] For this purpose, buying knock-out calls is a cheaper and more efficient hedging tool than using plain vanilla options. Additionally, the knock-out or knock-in components embedded in bonds may serve to reduce the costs related to costly bond covenants (Cox and Rubinstein, 1985).

[2] The so-called stop-loss strategy for portfolio insurance can be duplicated by knock-out barrier options.

In the framework of the Black-Scholes model, a closed-form solution for barrier options can be derived. Merton (1973) was the first to give a closed-form pricing formula for European knock-out options. Rubinstein and Reiner (1991), and Rich (1994) dealt with barrier options systematically and covered all possible variants of barrier options. A barrier option is cheaper than the corresponding plain-vanilla options because the possibility of knock-out or knock-in is priced explicitly as a rebate. This price rebate is the so-called knock-out (in) discount. Pricing the knock-out discount is essentially a problem of evaluating the probability that barrier option hits the pre-determined barrier. Obviously, this probability depends on the volatility of the underlying asset. Under the usual assumption of constant volatility, the volatility has a stable impact on these hitting probabilities. However, if the volatility itself follows a stochastic process, a sudden change of its value might considerably affect the hitting probabilities. For instance, in the case where stock prices and volatilities are negatively correlated, the left tail of the distribution of stock prices will be thinner than a lognormal distribution. Thus, the hitting probabilities of knock-out options will tend to be smaller compared to constant volatilities. Consequently, the knock-out discount will be smaller. At the same time, stochastic volatilities have a significant effect on the values of plain vanilla options. Therefore, with these two effects together, stochastic volatilities have a more complicated impact on prices of knock-out options than on prices of plain vanilla options. To our knowledge, an exact analytical pricing model for barrier options with stochastic volatilities is not available yet. This section tries to partially fill this gap.

We take out-of-the-money (OTM) knock-out options as an example to highlight the issue of evaluating the hitting probabilities. A general form of an OTM knock-out call option is

$$\begin{aligned} C_{KO} &= e^{-rT}\left[(S(T)-K)\Pr(X(T)\geqslant \ln K, m_T^X \geqslant \ln H)\right] \\ &= e^{-rT}\left[(S(T)-K)\Pr(X(T)\geqslant \ln K)\right] \\ &\quad - e^{-rT}\left[(S(T)-K)\Pr(X(T)\geqslant \ln K, m_T^X < \ln H)\right], \end{aligned} \tag{10.7}$$

where H is the downside barrier and smaller than K. $m_T^X = \min_{u\in[0,T]} X(u)$ stands for the minimum of $X(u)$ up to time T. Formula (10.7) implies that an OTM knock-out call option is composed of two parts: a standard European call and a call involving the probability $\Pr(X(T)\geqslant \ln K, m_T^X < \ln H)$. If the volatility $v(t)$ is constant, then the joint distribution of $(X(T), m_T^X)$ satisfies

$$\begin{aligned} \Pr(X(T)\geqslant x, m_T^X \geqslant y) = N\left(\frac{-x+rT-\frac{1}{2}v^2T}{v\sqrt{T}}\right) \\ -e^{(2rv^{-2}-1)y}N\left(\frac{2y-x+rT-\frac{1}{2}v^2T}{v\sqrt{T}}\right) \end{aligned} \tag{10.8}$$

for every x, y such that $y \leqslant x$. This probability is the key to derive a closed-form solution for OTM knock-out call options, which is given by[3]

$$\begin{aligned} C_{KO} &= C - C_{discount} \\ &= C - S_0 e^{(2rv^{-2}+1)y} N(d_1^*) + K e^{-rT} e^{(2rv^{-2}-1)y} N(d_2^*), \end{aligned} \tag{10.9}$$

where C denotes a standard European-style call option price and

$$d_j^* = \frac{2y - x + rT \pm \frac{1}{2}v^2 T}{v\sqrt{T}}, \quad y = \ln(H/S_0), \quad x = \ln(K/S_0), \quad j = 1, 2.$$

The formula (10.9) was given for the first time by Merton (1973). The probability (10.8) shows two different points which are specific to barrier options: Firstly, the probability is two-dimensional and involves two events: the first one is the terminal value $X(T)$ at time T, the other one is the minimum m_T^X over the period up to time T; Secondly, the minimum involves the stopping time. Usually, the reflecting principle of the standard Brownian motion is applied to solve this problem. But, in the presence of stochastic volatilities in a stock price process, calculating the probability $\Pr(X(T) \geqslant \ln K, m_T^X < \ln H)$ becomes explosively complicated. It is not theoretically assured that the heuristic argument of the reflecting principle is still valid in stochastic volatility models. Instead of generally dealing with barrier options in stochastic volatility models, we address only two special cases where we can indeed arrive at exact closed-form solutions.

The following subsections are organized as follows: Firstly, we handle two special cases and give two closed-form solutions for barrier options respectively. The pricing formulas for other types of barrier options can be obtained correspondingly. Next, some numerical examples are given to demonstrate the special features of the new solutions with stochastic volatilities.

10.2.2 Two Special Cases

The first case is an OTM knock-out option on futures or forwards. Although they are not commonplace in financial markets, barrier options on futures provide investors with a potentially efficient insurance tool in trading futures. If futures or forward prices fall under a certain level, investors will get a margin call and are required to fulfil a maintenance margin. In this case, one can hedge unexpected cash shortcomings by using a long position in barrier put options. Another direct application of this special case may be the valuation of Libor barrier caplets and floorlets since Libors as forwards are assumed to follow a driftless geometric Brownian motion.

---

[3] See Goldman, Sosin and Gatto (1979), or Merton (1973). Harrison (1985), Musiela and Rutkowski (2005) give a detailed derivation of the joint distribution of (X_T, m_T^X) with $m_T^X = \min_{u\in[0,T]} X(u)$.

Generally, plain-vanilla option on futures and forwards are priced with Black76 formula in markets. Therefore, if an option can valued with Black76 formula, its barrier counterparts may be valued with stochastic volatility by the formula given below.

Within the framework of the risk-neutral pricing, we have the following relation

$$F^T(t) = \mathbf{E}[S(T)] = e^{r(T-t)}S(t), \qquad t \leqslant T, \tag{10.10}$$

where $F^T(t)$ denotes the futures price of spot price $S(T)$ at time t. Under this assumption we obtain the dynamics of futures price

$$dF^T(t) = F^T(t)v(t)dW(t), \qquad t \leqslant T, \tag{10.11}$$

or

$$dX(t) = -\frac{1}{2}v(t)^2 dt + v(t)dW(t), \quad \text{with } X_0 = 0, \tag{10.12}$$

where $X(t) = \ln(F^T(t)/F_0^T)$. Obviously, $F^T(t)$ is a martingale. Black (1976b) derived the pricing formula for options on futures, which differs from the Black-Scholes formula only in the replacement of S_0 by $e^{-rT}F_0^T$. We specify here either the volatility $v(t)$ as a mean-reverting Ornstein-Uhlenbeck process or the squared volatility $v(t)^2$ as a mean-reverting (double) square root process. However, it is assumed that these volatilities are not correlated with stock returns. In this subsection, we take the Ornstein-Uhlenbeck process for demonstrating our results and have the following proposition:

Proposition 10.2.1. *If $X(t)$ follows the process (10.7) and the volatility $v(t)$ follows a mean-reverting Ornstein-Uhlenbeck process, and they are mutually not correlated, then the probabilities of $\Pr(m_T^X \leqslant z_1)$ and $\Pr(X(T) \geqslant x, m_T^X \geqslant z_2)$ are given respectively by*

$$\begin{aligned}\Pr(m_T^X \leqslant z_1) = &\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z_1)}{i\phi}\right)d\phi \\ &+ e^{-z_1}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z_1)}{i\phi}\right)d\phi\right].\end{aligned} \tag{10.13}$$

with $z_1 \leqslant 0$ and

$$\begin{aligned}\Pr(X(T) \geqslant z_1, m_T^X \geqslant z_2) = &\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z_1)}{i\phi}\right)d\phi \\ &- e^{-z_2}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(z_1 - 2z_2))}{i\phi}\right)d\phi\right]\end{aligned} \tag{10.14}$$

with $z_1 \geqslant z_2$ and $z_2 \leqslant 0$. The CF $f_2(\phi)$ is defined by $f_2(\phi) = \boldsymbol{E}[\exp(i\phi X(T))]$.

The proof is given in Appendix C. By applying the results in Proposition 10.2.1, we can derive a pricing formula for OTM knock-out options on futures, which is expressed by

$$
\begin{aligned}
C_{KO} &= e^{-rT}\mathbf{E}\left[(F^T(T)-K)\cdot \mathbf{I}_{(X(T)\geqslant x, m_T^X \geqslant y)}\right] \\
&= e^{-rT}F_0^T\mathbf{E}\left[\frac{F^T(T)}{F_0^T}\cdot \mathbf{I}_{(X(T)\geqslant x, m_T^X \geqslant y)}\right] - e^{-rT}K\,\mathrm{Pr}(X(T)\geqslant x, m_T^X \geqslant y)
\end{aligned}
\tag{10.15}
$$

with $x=\ln(K/F_0^T)$, $y=\ln(H/F_0^T)$. Note that $F^T(T)/F_0^T$ implies a change of measure by which the original process (10.12) is switched to be

$$
dX(t) = \frac{1}{2}v(t)^2dt + v(t)d\widehat{w}(t), \qquad \text{with } X_0 = 0.
$$

This fact immediately leads to the following pricing formula for barrier options

$$
C_{KO} = e^{-rT}F_0^T F_1 - e^{-rT}KF_2 \tag{10.16}
$$

with

$$
\begin{aligned}
F_1 =& \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(-i\phi x)}{i\phi}\right)d\phi \\
&-e^{y}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(-i\phi(x-2y))}{i\phi}\right)d\phi\right]
\end{aligned}
\tag{10.17}
$$

and

$$
\begin{aligned}
F_2 =& \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi x)}{i\phi}\right)d\phi \\
&-e^{-y}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(x-2y))}{i\phi}\right)d\phi\right],
\end{aligned}
\tag{10.18}
$$

where $f_1(\phi)=\mathbf{E}[\exp((1+i\phi)X(T))]$ and $f_2(\phi)=\mathbf{E}[\exp(i\phi X(T))]$. Apparently, the discounted factor e^{-rT} can be replaced by the price of a zero-coupon bond *à la* CIR (1985b) or Vasicek (1977). For in-the-money (ITM) knock-out options, we have $K\leqslant H$ and $H\leqslant F_0^T$. In this case, the call option has the following form,

$$
C_{KO} = F_0^T e^{-rT}F_1(m_T^X\geqslant y) - Ke^{-rT}F_2(m_T^X\geqslant y) \tag{10.19}
$$

where

$$
\begin{aligned}
F_1 =& \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(-i\phi y)}{i\phi}\right)d\phi \\
&- e^{-y}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(i\phi y)}{i\phi}\right)d\phi\right]
\end{aligned}
$$

and

$$F_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi y)}{i\phi}\right)d\phi$$
$$- e^{-y}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi y)}{i\phi}\right)d\phi\right].$$

The formulas (10.16) and (10.19) together give the pricing formula for knock-out barrier options on futures under the assumption that the volatilities and the stock returns are not correlated.

In the second case, we consider barrier options on spot instruments. For this purpose, we adopt the stock price process as in (6.16), that is

$$dX(t) = r(t)\left(1 - \frac{1}{2}v^2\right)dt + v\sqrt{r(t)}dW_1, \qquad X(t) = \ln(S(t)/S_0). \tag{10.20}$$

Here the volatility is the term $v\sqrt{r(t)}$ times a constant v, and is then accompanied by the risk of interest rates. In other words, volatilities vary through time because they bear the same risk as interest rates that are specified as a mean-reverting square root process

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dW_2, \tag{10.21}$$

where $dW_1 dW_2 = 0$. Under this specification, we have the following proposition on the hitting probabilities:

Proposition 10.2.2. *If $X(t)$ and $r(t)$ follow the processes (10.20) and (10.21) respectively, then the probabilities of $\Pr(m_T^X \leqslant z_1)$ and $\Pr(X(T) \geqslant z_1, m_T^X \geqslant y)$ are given by*

$$\Pr(m_T^X \leqslant z_1) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z_1)}{i\phi}\right)d\phi$$
$$+ e^{-z_1\left(1-2v^{-2}\right)}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z_1))}{i\phi}\right)d\phi\right] \tag{10.22}$$

with $z_1 \leqslant 0$, and

$$\Pr(X(T) \geqslant z_1, m_T^X \geqslant z_2) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z_1)}{i\phi}\right)d\phi$$
$$- e^{-z_2\left(1-2v^{-2}\right)}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(z_1 - 2z_2))}{i\phi}\right)d\phi\right] \tag{10.23}$$

with $z_1 \geqslant z_2$ and $z_2 \leqslant 0$. The CF $f_2(\phi)$ is defined by $f_2(\phi) = \boldsymbol{E}[\exp(i\phi X(T))]$.

The detailed proof is given in Appendix C. Following the same steps already applied above, we obtain a pricing formula for knock-out barrier options on a spot instrument,

$$
\begin{aligned}
C_{KO} &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\left((S(T)-K)\cdot\mathbf{I}_{(X(T)\geqslant x, m_T^X\geqslant y)}\right)\right] \\
&= S_0\cdot F_1 - B(0,T)K\cdot F_2
\end{aligned} \tag{10.24}
$$

with

$$
\begin{aligned}
F_1 &= \frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(-i\phi x)}{i\phi}\right)d\phi \\
&- e^{x\left(1+2v^{-2}\right)}\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(-i\phi(x-2y))}{i\phi}\right)d\phi\right]
\end{aligned} \tag{10.25}
$$

and

$$
\begin{aligned}
F_2 &= \frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi x)}{i\phi}\right)d\phi \\
&- e^{-x\left(1-2v^{-2}\right)}\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(x-2y))}{i\phi}\right)d\phi\right].
\end{aligned} \tag{10.26}
$$

The pricing formula for ITM knock-out options is correspondingly given by

$$
C_{KO} = S_0F_1(m_T^X\geqslant y) - B(0,T)KF_2(m_T^X\geqslant y) \tag{10.27}
$$

where

$$
\begin{aligned}
F_1 &= \frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(-i\phi y)}{i\phi}\right)d\phi \\
&- e^{x\left(1+2v^{-2}\right)}\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_1(\phi)\frac{\exp(i\phi y)}{i\phi}\right)d\phi\right]
\end{aligned} \tag{10.28}
$$

and

$$
\begin{aligned}
F_2 &= \frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi y)}{i\phi}\right)d\phi \\
&- e^{-x\left(1-2v^{-2}\right)}\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi y)}{i\phi}\right)d\phi\right].
\end{aligned} \tag{10.29}
$$

The two CFs are respectively defined by

$$f_1(\phi) = \mathbf{E}\left[\frac{\exp\left(-\int_0^T r(t)dt\right)S(T)}{S_0}\exp(i\phi X(T))\right] \tag{10.30}$$

and

$$f_2(\phi) = \mathbf{E}\left[\frac{\exp\left(-\int_0^T r(t)dt\right)}{B(0,T)}\exp(i\phi X(T))\right]. \tag{10.31}$$

All of the above closed-form solutions display the same structure as Merton's solution given in (10.9), and consists of two components: the first one is the same as the standard European call options; the second is the so-called knock-out discount or price rebate. The implementation of these formulas presents no special difficulty since all CFs are known and are given in the above chapter. The put-call parity for barrier options is different from that for plain vanilla options. Because of the following relation

$$\Pr(X(T) \geqslant x, m_T^X \geqslant y) + \Pr(X(T) < x, m_T^X < y) = 1,$$

we have the parity for knock-out calls and knock-in puts:

$$C_{KO} + Ke^{-rT} = P_{KI} + S_0, \tag{10.32}$$

which enables us to obtain the pricing formula for knock-in put options. Other variants of barrier options can be evaluated by applying the schemes given in Rubinstein and Reiner (1991), or Rich (1994).

10.2.3 Numerical Examples

In this subsection, we present some numerical examples to demonstrate the special features of the above derived pricing formulas and to show how the stochastic volatilities or stochastic interest rates affect the values of knock-out options. At first, we examine the knock-out options on futures with two different stochastic volatility models, namely, the volatility as a mean-reverting Ornstein-Uhlenbeck process and the squared volatility as a mean-reverting square-root process. The CFs for these two processes have already been given in Chapter 3, respectively. For the implementation of the formulas (10.16) and (10.19), we simply set ρ equal to zero in the corresponding CFs.

Table (10.1) reports the case where the stochastic volatilities follow an Ornstein-Uhlenbeck process, and includes three panels with different values of the long-run mean θ of volatility. In every panel, we calculate the values of the knock-out call options by combining the different strike price K and the barrier H. Thus, both in-the-money (ITM) knock-out options $(K < H)$ and out-of-the-money (OTM) knock-out options $(K > H)$ are considered. To perform a comparison with the case of constant volatility, we evaluate the benchmark values of the knock-out options using Merton's formula with the expected average variance as given in (3.37). The price differences between the model values and the benchmarks are denoted by PD. In order to understand the following tables better, we briefly discuss a special feature associated with knock-out options on futures: If $H = K \leqslant F_0$, we always have the call prices equal to $e^{-rT}(F_0 - K)$, regardless of how the volatilities are specified. This feature can be explained as follows: Since the barrier H is set to be K, the options can not gain a premium for the case $F^T(t) < K, 0 \leqslant t \leqslant T$. Furthermore, $F^T(t)$ is a martingale, therefore such particular options have no time value and their values are simply the discounted positive difference between the current futures price F_0 and the strike price K. Consequently, the values of PD on the diagonal from Panel A to Panel C are zero. Another feature of knock-out options is that the options are worthless if $H = F_0$. This is apparent regardless of whether the underlying assets are futures or equity. Thus, both call prices and the PD in the last column in panels A, B and C are equal to zero.

Some observations are summarized as follows: Firstly, all exact theoretical values of the OTM knock-out options ($H < K$) are greater than the corresponding benchmark values, and we have $PD > 0$. At the same time, the exact theoretical values and the corresponding benchmark for the ITM knock-out options perform an opposite relation with $PD < 0$. This finding is valid for all panels in Table 3.1 and independent of the values of θ. Thus, Merton's solution seems to undervalue (overvalue) the OTM (ITM) knock-out options on futures. Secondly, the more the spot volatility differs from its long-run mean θ, the larger the magnitude of the undervaluation and overvaluation becomes. In Panel B and Panel C, we can see that the mispricing due to constant volatility is significant. If we calculate the call prices by simply using spot volatility, the price biases are much more remarkable. To save space, we do not list these values here. Moreover, the price biases PD do not display a simple increasing or decreasing relationship with both H and K. The pattern of the mispricing seems to be hump-shaped. Finally, by comparing the data across panels, we reveal that a higher long-run mean θ leads to higher prices of the OTM knock-out options and lower prices of the ITM knock-out options. Similarly, a lower long-run mean θ leads to lower prices of the OTM knock-out options and higher prices of the ITM knock-out options.

In Table (10.2), we give the theoretical values of the knock-out options with the squared volatility $v(t)^2$ (variance) as a square root process. To calculate a similar benchmark as in Table (10.1), the expected average variance has to be evaluated and is given by

$$\begin{aligned} AV &= \mathbf{E}\left[\frac{1}{T}\int_0^T v(t)^2 dt\right] \\ &= \frac{1}{\kappa T}(v^2 - \theta)(1 - e^{-\kappa T}) + \theta. \end{aligned} \tag{10.33}$$

As expected, these two stochastic volatility models perform the almost identical features, as shown in Table (10.2). All findings in Table (10.1) are confirmed again by Table (10.2) except that the magnitude of the mispricing due to constant volatility is not so considerable as in the model with an Ornstein-Uhlenbeck process.

Table (10.3) lists the theoretical values of the knock-out options on spot calculated by using the formulas (10.24) and (10.27). As discussed above, this model is based on a modified process of stock prices and a specification of interest rates as a square root process. This raises a problem in finding out a suitable benchmark for the purpose of comparison. In order to make Merton's solution to match this model, we replace the constant interest rate and the squared volatility by the expected average interest rate and the expected average variance, respectively. These two expected values can be calculated by (10.33) and have the following form

$$r_{const} = \frac{1}{\kappa_r T}(r-\theta_r)(1-e^{-\kappa_r T})+\theta_r, \tag{10.34}$$

$$v_{const} = v\sqrt{r_{const}} = v\sqrt{\frac{1}{\kappa_r T}(r-\theta_r)(1-e^{-\kappa_r T})+\theta_r}. \tag{10.35}$$

The benchmark values evaluated in this way best fit the values of our model in the case where the spot interest rate r is equal to θ_r, as shown in Panel G. But if the current interest rate r diverges from the long-run mean θ_r, the considerable prices biases occur. Since the underlying asset of the options is not futures but equity, it is not expected that the option prices in Table (10.3) are zero for $K = H$. The pattern of the price biases does not have a clear structure as in the model for barrier options on futures. This might be due to the choice of the benchmark. Apart from this exception, knock-out options on spot share all the features of these on futures. Thus, the model specified by (10.20) and (10.21) could be regarded as a reliable model for barrier options.

10.3 Lookback Options

10.3.1 Introduction

Lookback options are financial derivatives which allow their owners to purchase (sell) a certain asset at expiry at the minimum (maximum) price during the option's lifetime. This type of option makes it possible for investors to buy at the lowest and sell at the highest. However, lookback options offer no-free-lunch for investors and, by the rule of thumb, are almost doubly expensive as the corresponding plain vanilla options. A possible application of lookback options might be related to some sophisticated portfolios. For example, one can minimize the regret of missing the best interest (exchange) rate if (currency linked) bonds are issued. Obviously, lookback options are path-dependent and involve evaluating the probability of the maximum or the minimum of the underlying asset again. In the previous section on barrier options, we have derived all the probabilities associated with a certain barrier, both

H K		80	85	90	95	100	105	110
80	Call	19.506	15.336	11.480	8.170	5.531	3.571	2.211
	PD	0.000	-0.012	-0.031	-0.052	-0.060	-0.049	-0.026
85	Call	18.181	14.630	11.146	8.019	5.464	3.542	2.199
	PD	0.041	0.000	-0.039	-0.066	-0.072	-0.058	-0.031
90	Call	14.854	12.304	9.753	7.261	5.072	3.347	2.104
	PD	0.114	0.057	0.000	-0.051	-0.077	-0.071	-0.045
95	Call	8.803	7.494	6.185	4.877	3.600	2.487	1.623
	PD	0.120	0.080	0.040	0.000	-0.036	-0.052	-0.045
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

A: $v = 0.2, \theta = 0.2, \kappa = 2, \sigma = 0.1, T = 0.5, F_0 = 100, r = 0.05$

H K		80	85	90	95	100	105	110
80	Call	19.506	15.510	11.810	8.614	6.027	4.056	2.637
	PD	0.000	-0.087	-0.166	-0.223	-0.248	-0.236	-0.198
85	Call	17.963	14.630	11.354	8.388	5.917	4.004	2.612
	PD	0.054	0.000	-0.051	-0.086	-0.098	-0.088	-0.062
90	Call	14.453	12.103	9.753	7.450	5.393	3.721	2.463
	PD	0.120	0.060	0.000	-0.055	-0.088	-0.091	-0.071
95	Call	8.449	7.258	6.067	4.877	3.710	2.672	1.834
	PD	0.210	0.140	0.070	0.000	-0.065	-0.107	-0.119
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

B: $v = 0.2, \theta = 0.25, \kappa = 2, \sigma = 0.1, T = 0.5, F_0 = 100, r = 0.05$

H K		80	85	90	95	100	105	110
80	Call	19.506	15.173	11.164	7.736	5.043	3.102	1.813
	PD	0.000	-0.083	-0.174	-0.253	-0.288	-0.267	-0.207
85	Call	18.404	14.630	10.932	7.642	5.006	3.088	1.807
	PD	0.145	0.000	-0.139	-0.243	-0.288	-0.269	-0.209
90	Call	15.298	12.525	9.753	7.054	4.728	2.962	1.751
	PD	0.335	0.168	0.000	-0.154	-0.245	-0.254	-0.206
95	Call	9.213	7.767	6.322	4.877	3.473	2.284	1.402
	PD	0.334	0.223	0.111	0.000	-0.101	-0.154	-0.150
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

C: $v = 0.2, \theta = 0.15, \kappa = 2, \sigma = 0.1, T = 0.5, F_0 = 100, r = 0.05$

Table 10.1 The values of knockout calls on futures. The stochastic volatility is specified as a mean-reverting Ornstein-Uhlenbeck process.

maximum and minimum for two particular processes. Stochastic interest rates can be embodied into these probabilities if we introduce a modified stock price process as suggested in (6.16). In this section, we will discuss how to price lookback options with such stochastic factors.

We briefly review the valuation of a lookback option in the Black-Scholes's framework. The earliest study on this topic was done by Goldman, Sosin and Gatto (1979) who gave detailed arguments about the replication strategy, hedgeability and valuation associated with lookback options. Based on reasonable replication strategies, the risk-neutral pricing is well guaranteed. The pricing formula with constant volatility is given by

H K		80	85	90	95	100	105	110
80	Call	19.506	15.310	11.439	8.123	5.480	3.515	2.150
	PD	0.000	-0.003	-0.008	-0.013	-0.015	-0.013	-0.007
85	Call	18.194	14.630	11.133	7.995	5.429	3.496	2.143
	PD	0.011	0.000	-0.010	-0.017	-0.019	-0.015	-0.008
90	Call	14.850	12.302	9.753	7.261	5.065	3.326	2.067
	PD	0.030	0.015	0.000	-0.013	-0.020	-0.019	-0.012
95	Call	8.784	7.482	6.179	4.877	3.604	2.489	1.614
	PD	0.031	0.021	0.010	0.000	-0.009	-0.013	-0.012
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

D: $v^2 = 0.04, \theta = 0.04, \kappa = 2, \sigma = 0.1, T = 0.5, F_0 = 100, r = 0.05$

H / K		80	85	90	95	100	105	110
80	Call	19.506	15.503	11.800	8.605	6.016	4.039	2.610
	PD	0.000	-0.004	-0.009	-0.012	-0.013	-0.011	-0.007
85	Call	17.955	14.630	11.361	8.397	5.922	3.998	2.593
	PD	0.010	0.000	-0.009	-0.015	-0.016	-0.014	-0.008
90	Call	14.419	12.086	9.753	7.465	5.413	3.733	2.460
	PD	0.022	0.011	0.000	-0.010	-0.016	-0.016	-0.012
95	Call	8.409	7.232	6.054	4.877	3.722	2.688	1.845
	PD	0.022	0.015	0.007	0.000	-0.007	-0.010	-0.010
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

E: $v^2 = 0.04, \theta = 0.0625, \kappa = 2, \sigma = 0.1, T = 0.5, F_0 = 100, r = 0.05$

H K		80	85	90	95	100	105	110
80	Call	19.506	15.153	11.136	7.711	5.019	3.072	1.773
	PD	0.000	-0.001	-0.007	-0.014	-0.017	-0.014	-0.006
85	Call	18.410	14.630	10.927	7.633	4.992	3.063	1.770
	PD	0.011	0.000	-0.011	-0.019	-0.021	-0.016	-0.007
90	Call	15.271	12.512	9.753	7.064	4.737	2.956	1.728
	PD	0.037	0.019	0.000	-0.017	-0.024	-0.021	-0.011
95	Call	9.165	7.736	6.306	4.877	3.486	2.296	1.402
	PD	0.042	0.028	0.014	0.000	-0.012	-0.017	-0.013
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

F: $v^2 = 0.04, \theta = 0.0225, \kappa = 2, \sigma = 0.1, T = 0.5, F_0 = 100, r = 0.05$

Table 10.2 The values of knockout calls on futures. The stochastic variance is specified as a mean-reverting square root process.

$$
\begin{aligned}
C_{LB} = {} & S_0 N(d) - m e^{-rT} N(d - v\sqrt{T}) - \frac{S_0 v^2}{2r} N(-d) \\
& + e^{-rT} \frac{S_0 v^2}{2r} \left(\frac{m}{S_0} \right)^{2rv^{-2}} N(-d + 2rv^{-1}\sqrt{T}),
\end{aligned}
\tag{10.36}
$$

where

$$
d = \frac{\ln(S_0/m) + (r + \frac{1}{2}v^2)T}{v\sqrt{T}}.
$$

m denotes the realized minimum price until the current time. The first two terms in (10.36) represent the value of a standard European-style option whereas the last two terms in (10.36) can be interpreted as the strike-bonus since there is always a pos-

H / K		80	85	90	95	100	105	110
80	Call	21.529	17.176	13.103	9.534	6.608	4.364	2.751
	PD	-0.003	-0.006	-0.010	-0.015	-0.016	-0.012	-0.005
85	Call	20.294	16.513	12.792	9.398	6.551	4.341	2.742
	PD	0.006	-0.004	-0.013	-0.019	-0.020	-0.015	-0.007
90	Call	16.933	14.137	11.342	8.599	6.140	4.141	2.650
	PD	0.024	0.010	-0.004	-0.017	-0.023	-0.020	-0.012
95	Call	10.357	8.868	7.379	5.890	4.431	3.130	2.085
	PD	0.028	0.017	0.007	-0.003	-0.012	-0.016	-0.013
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

G: $r = 0.04, \theta_r = 0.04, \kappa_r = 2, \sigma_r = 0.1, T = 0.5, S_0 = 100, v = 1$

H / K		80	85	90	95	100	105	110
80	Call	21.783	17.584	13.645	10.170	7.279	5.007	3.316
	PD	-0.288	-0.464	-0.619	-0.727	-0.766	-0.733	-0.640
85	Call	20.327	16.743	13.210	9.957	7.180	4.963	3.297
	PD	-0.034	-0.261	-0.479	-0.643	-0.722	-0.711	-0.630
90	Call	16.741	14.130	11.520	8.952	6.615	4.661	3.142
	PD	0.230	0.013	-0.204	-0.408	-0.549	-0.598	-0.564
95	Call	10.119	8.742	7.366	5.990	4.637	3.408	2.386
	PD	0.283	0.150	0.018	-0.114	-0.238	-0.322	-0.348
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H: $r = 0.04, \theta_r = 0.0625, \kappa_r = 2, \sigma_r = 0.1, T = 0.5, S_0 = 100, v = 1$

H /K		80	85	90	95	100	105	110
80	Call	21.299	16.805	12.602	8.938	5.977	3.767	2.243
	PD	0.227	0.367	0.492	0.580	0.613	0.583	0.504
85	Call	20.279	16.301	12.392	8.857	5.948	3.757	2.239
	PD	0.021	0.208	0.386	0.520	0.582	0.568	0.498
90	Call	17.166	14.170	11.175	8.244	5.663	3.633	2.188
	PD	-0.202	-0.020	0.163	0.334	0.450	0.486	0.451
95	Call	10.642	9.027	7.411	5.796	4.219	2.845	1.786
	PD	-0.247	-0.134	-0.021	0.091	0.197	0.266	0.285
100	Call/PD	0.000	0.000	0.000	0.000	0.000	0.000	0.000

I: $r = 0.04, \theta_r = 0.0225, \kappa_r = 2, \sigma_r = 0.1, T = 0.5, S_0 = 100, v = 1$

Table 10.3 The values of knockout calls on spot. The underlying spot is specified as a modified stock price process, and stochastic interest is specified as a mean-reverting square root process.

sibility that the actual strike price falls lower than the realized m. Hence, lookback options are always in-the-money and, thus, cost more than their standard counterparts.

10.3.2 Pricing Formulas with Stochastic Factors

We now attempt to value lookback options with stochastic volatilities or stochastic interest rates in the same economic setting as given in Chapter 3 and Chapter 6. We now consider a lookback call option on futures, whose terminal payoff is defined by

$$C_{LB}(T) = F_T^T - \min(m, m_T^F), \tag{10.37}$$

where $m_T^F = \min_{t\in[0,T]} F_t^T$ and m is the realized minimum price during the past option's lifetime. By the risk-neutral pricing, we have

$$\begin{aligned} C_{LB} &= e^{-rT}\mathbf{E}[F_T^T - \min(m, m_T^F)] \\ &= e^{-rT}F_0^T - e^{-rT}\mathbf{E}\left[\min(m, m_T^F)\right] \\ &= e^{-rT}F_0^T - e^{-rT}\mathbf{E}\left[\min(m, F_0^T\exp(m_T^X)\right], \end{aligned} \tag{10.38}$$

where $m_T^X = \min X(t)$ and $X(t) = \ln(F_t^T/F_0^T)$. Following the above procedure, we obtain

$$\begin{aligned} e^{rT}C_{LB} &= F_0^T - \mathbf{E}\left[\min(m, F_0^T\exp(m_T^X)\right] \\ &= F_0^T - \mathbf{E}\left[F_0^T\exp(m_T^X)\cdot \mathbf{I}_{(F_0^T\exp(m_T^X)\leqslant m)} + m\cdot \mathbf{I}_{(F_0^T\exp(m_T^X)>m)}\right] \\ &= F_0^T - F_0^T\mathbf{E}\left[\exp(m_T^X)\cdot \mathbf{I}_{(m_T^X\leqslant z)}\right] - m\mathbf{E}\left[\mathbf{I}_{(m_T^X>z)}\right] \\ &= F_0^T - F_0^T\mathbf{E}\left[\exp(m_T^X)\cdot \mathbf{I}_{(m_T^X\leqslant z)}\right] - m\Pr(m_T^X > z) \end{aligned} \tag{10.39}$$

with $z = \ln(m/F_0^T)$. According to Proposition 10.2.1 in the above Section the probability $\Pr(m_T^X > z)$ can be calculated as follows:

$$\begin{aligned} L_1 &= \Pr(m_T^X > z) \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z)}{i\phi}\right)d\phi \\ &- e^{-z}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z)}{i\phi}\right)d\phi\right]. \end{aligned} \tag{10.40}$$

Differentiating $\Pr(m_T^X \leqslant u)$ with respect to u yields the density function $pdf(u)$ of m_T^X,

$$\begin{aligned} pdf(u) &= \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\exp(-i\phi u)\right)d\phi \\ &- e^{-u}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi u)}{i\phi}\right)d\phi\right] \\ &+ e^{-u}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\exp(i\phi u)\right)d\phi. \end{aligned}$$

Rearranging it yields

$$pdf(u) = \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[f_2(\phi)e^{-i\phi u} + h(u;\phi)e^{i\phi u}\right] d\phi - \frac{1}{2}e^{-u} \tag{10.41}$$

with

$$h(u;\phi) = f_2(\phi)e^{-u}\frac{(i\phi-1)}{i\phi}.$$

We evaluate $\mathbf{E}\left[\exp(m_T^X)\cdot \mathbf{I}_{(m_T^X \leqslant z)}\right]$ by the Fourier transform,

$$\begin{aligned}
L_2 &= \mathbf{E}\left[\exp(m_T^X)\cdot \mathbf{I}_{(m_T^X \leqslant z)}\right] = \int_{-\infty}^{z} e^u pdf(u)du, \quad u = M_T^X, \\
&= \frac{1}{\pi}\int_{-\infty}^{z}\int_0^{\infty} \mathrm{Re}\{f_2(\phi)e^{u-i\phi u} + h(u;\phi)e^{u+i\phi u}\}d\phi du - \int_{-\infty}^{z}\frac{1}{2}du \\
&= \frac{1}{\pi}\int_0^{\infty}\int_{-\infty}^{z} \mathrm{Re}\left(f_2(\phi)e^{(1-i\phi)u}\right) dud\phi \\
&+ \frac{1}{\pi}\int_0^{\infty}\int_{-\infty}^{z} \mathrm{Re}\left(f_2(\phi)e^{i\phi u}\frac{(i\phi-1)}{i\phi}\right) dud\phi - \int_{-\infty}^{z}\frac{1}{2}du \\
&= \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_2(\phi)\frac{1}{1-i\phi}e^{(1-i\phi)z}\right) d\phi \\
&+ \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_2(\phi)\frac{e^{i\phi z}-e^{-i\phi\infty}}{i\phi}\right) d\phi \\
&- \frac{1}{2}\int_{-\infty}^{z} d\phi - \frac{1}{\pi}\int_0^{\infty}\int_{-\infty}^{z} \mathrm{Re}\left(f_2(\phi)\frac{e^{i\phi u}}{i\phi}\right) dud\phi.
\end{aligned} \tag{10.42}$$

Note the identity

$$\frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_2(\phi)\frac{e^{i\phi z}}{i\phi}\right) = \Pr(X(T) > -z),$$

we have

$$\frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_2(\phi)\frac{e^{-i\phi\infty}}{i\phi}\right) = \Pr(X(T) > \infty) - \frac{1}{2} = -\frac{1}{2}.$$

Additionally, in order to obtain a tractable expression for the third and fourth terms in (10.42), we use the well-known fact $\frac{2}{\pi}\int_0^{\infty}\frac{\sin(\phi u)}{\phi}d\phi = sign(u)$ to express $\frac{1}{2}\int_{-\infty}^{z} du$ as

$$-\frac{1}{2}\int_{-\infty}^{z} du = \frac{1}{\pi}\int_0^{\infty}\int_{-\infty}^{z}\frac{\sin(\phi u)}{\phi}dud\phi.$$

Using these manipulations, we obtain

$$
\begin{aligned}
L_2 &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{e^{(1-i\phi)z}}{1-i\phi} + \frac{e^{i\phi z}}{i\phi}\right)d\phi \\
&\quad - \frac{1}{\pi}\int_0^\infty\int_{-\infty}^{z} \mathrm{Re}\left(f_2(\phi)\frac{e^{i\phi u}}{i\phi} + \frac{\sin(\phi u)}{\phi}\right)dud\phi \\
&= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{e^{(1-i\phi)z}}{1-i\phi} + \frac{e^{i\phi z}}{i\phi}\right)d\phi \\
&\quad + \lim_{a\to-\infty}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{e^{i\phi z}-e^{i\phi a}}{\phi^2} - \frac{\cos(\phi z)-\cos(\phi a)}{\phi^2}\right)d\phi.
\end{aligned}
\tag{10.43}
$$

The numerical integrals in (10.43) display a well-behaved and fast convergence for a sufficiently small a, say -10.[4] The pricing formula for lookback options takes then a form of

$$
C_{LB} = e^{-rT}(F_0^T - mL_1 - F_0^T L_2). \tag{10.44}
$$

The interpretation of this pricing formula is straightforward: L_1 is the probability that the expected minimum up to maturity is smaller than the realized minimum while L_2 gives the opposite probability but under another measure. The technique applied in the above procedure is somehow different from anywhere else in this book. We do not change measure to evaluate $\mathbf{E}\left[\exp(m_T^X)\cdot \mathbf{I}_{(m_T^X\leqslant z)}\right]$. Instead we calculate it by using the density function.

If $X(t)$ and $r(t)$ follow the processes (10.20) and (10.21) respectively, we can obtain a pricing formula for lookback options on spot in the same way,

$$
\begin{aligned}
C_{LB} &= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\left[S(T) - \min(m, S_0\exp(m_T^X)\right]\right] \\
&= S_0 - B(0,T)(mL_1 + S_0L_2)
\end{aligned}
\tag{10.45}
$$

with

$$
\begin{aligned}
L_1 &= \Pr(m_T^X > z) \\
&= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z)}{i\phi}\right)d\phi \\
&= e^{-z\left(1-2v^{-2}\right)}\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z)}{i\phi}\right)d\phi\right]
\end{aligned}
\tag{10.46}
$$

and

[4] The number $a = -10$ for this intergral can be regarded to be sufficiently small because we have $m = 4.54\times 10^{-5}F_0^T$ ($\ln(m/F_0^T) = -10$) for this case and the probability $\Pr(m_T^X \leqslant -10)$ is actually zero. Thus, we only need to implement the integral over the interval $[a,z]$ instead of $[-\infty,z]$.

$$
\begin{aligned}
L_2 &= \frac{-1+2v^{-2}}{4v^{-2}} e^{2v^{-2}z} \\
&\quad + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\left[\frac{e^{(1-i\phi)z}}{1-i\phi} + \frac{1-2v^{-2}-i\phi}{\phi^2-2v^{-2}i\phi} e^{(2v^{-2}+i\phi)z}\right]\right) d\phi \\
&= e^{2v^{-2}z}[\Pr(X(T) > -z) - 0.25v^2] \\
&\quad + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\left[\frac{e^{(1-i\phi)z}}{1-i\phi} + \frac{e^{(2v^{-2}+i\phi)z}}{\phi^2-2v^{-2}i\phi}\right]\right) d\phi,
\end{aligned} \tag{10.47}
$$

where $f_2(\phi)$ is defined by

$$
f_2(\phi) = \mathbf{E}\left[\frac{\exp\left(-\int_0^T r(t)dt\right)}{B(0,T)} \exp(i\phi X(T))\right].
$$

A similar procedure can be applied to price lookback put options. However, one must know the probability distributions involving maximum. We have the following proposition:

Proposition 10.3.1. *(a): If $X(t)$ follows the process (10.12) and the volatility $v(t)$ follows a mean-reverting Ornstein-Uhlenbeck process, and they are mutually not correlated, then the probabilities of* $\Pr(X(T) \leqslant z_1, M_T^X \geqslant y)$ and $\Pr(M_T^X \leqslant z_2)$ *with $z_1 \leqslant z_2$ and $z_2 \geqslant 0$ are respectively given by*

$$
\begin{aligned}
&\Pr(X(T) \leqslant z_1, M_T^X \geqslant z_2) \\
&\qquad = e^{z_2}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(z_1-2z_2))}{i\phi}\right) d\phi\right]
\end{aligned} \tag{10.48}
$$

and

$$
\begin{aligned}
\Pr(M_T^X \leqslant z_2) &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z_2)}{i\phi}\right) d\phi \\
&\quad - e^{-z_2}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z_2)}{i\phi}\right) d\phi\right].
\end{aligned} \tag{10.49}
$$

(b): If $X(t)$ and $r(t)$ follow the processes (10.20) and (10.21) respectively, then the probabilities of $\Pr(X(T) \leqslant z_1, M_T^X \geqslant z_2)$ *and* $\Pr(M_T^X \leqslant z_2)$ *with $z_1 \leqslant z_2$ and $z_2 \geqslant 0$ are given by*

$$
\begin{aligned}
&\Pr(X(T) \leqslant z_1, M_T^X \geqslant z_2) \\
&\qquad = e^{-z_2\left(1-2v^{-2}\right)}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi(z_1-2z_2))}{i\phi}\right) d\phi\right]
\end{aligned} \tag{10.50}
$$

and

$$\Pr(M_T^X \leqslant z_2) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z_2)}{i\phi}\right)d\phi$$
$$- e^{-z_2\left(1-2v^{-2}\right)}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z_2)}{i\phi}\right)d\phi\right]. \quad (10.51)$$

A detailed proof is given in Appendix C. Since the terminal payoff of lookback put options on futures is equal to $\max_{t\in[0,T]} F^T(t) - F^T(T)$, the corresponding theoretical prices can be evaluated by

$$\begin{aligned} P_{LB} &= e^{-rT}\mathbf{E}\left[\max F^T(t) - F^T(T)\right] \\ &= e^{-rT}\mathbf{E}\left[M \Pr\left(M_T^X \leqslant z\right) + F_0^T \exp(M_T^X)\cdot \mathbf{I}_{(M_T^X > z)} - F^T(T)\right] \\ &= e^{-rT}[M\cdot L_1 + F_0^T \cdot L_2 - F_0^T], \end{aligned} \quad (10.52)$$

where M is the realized maximum up to now and z is equal to $\ln(M/F_0^T)$. L_1 and L_2 are respectively given by

$$L_1 = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z)}{i\phi}\right)d\phi$$
$$- e^{-z}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z)}{i\phi}\right)d\phi\right]$$

and

$$\begin{aligned} L_2 &= \mathbf{E}[\exp(M_T^X)\cdot \mathbf{I}_{(M_T^X > z)}] \\ &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z)}{i\phi}\right)d\phi \\ &+ \lim_{a\to\infty}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{e^{(1-i\phi)a} - e^{(1-i\phi)z}}{1-i\phi}\right)d\phi \\ &+ \lim_{a\to\infty}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{e^{i\phi a} - e^{i\phi z}}{\phi^2} - \frac{\cos(\phi a) - \cos(\phi z)}{\phi^2}\right)d\phi. \end{aligned}$$

In a practical implementation, setting the number a as 10 offers us very accurate results. The last integral in L_2 converges faster at a quadratic rate of ϕ. The derivation of the pricing formula for lookback put options on spot follows an analogous way,

$$P_{LB} = B(0,T)[M\cdot L_1 + S_0\cdot L_1] - S_0 \quad (10.53)$$

with

$$L_1 = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi z)}{i\phi}\right)d\phi$$
$$- e^{-z\left(1-2v^{-2}\right)}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z)}{i\phi}\right)d\phi\right]$$

and

$$L_2 = \lim_{a\to\infty}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{e^{(1-i\phi)a} - e^{(1-i\phi)z}}{1-i\phi}\right)d\phi$$
$$+ \frac{1}{1-b}\lim_{a\to\infty}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\left(e^{(1-b+i\phi)a} - e^{(1-b+i\phi)z}\right)}{1-b+i\phi}\right)d\phi$$
$$- \frac{b}{1-b}e^{(1-b)z}\left[\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi z)}{i\phi}\right)d\phi\right],$$

where $b = 1-2v^{-2}$. With these complex pricing formulas having been derived, we close this section.

10.4 Asian Options

10.4.1 Introduction

Asian options are referred to as a class of options whose payoff at maturity depends on a certain average of some prices of the underlying asset prior to the expiration. The average designed in a contract could be arithmetic or geometric. Asian options based on an arithmetic average dominate trading in the OTC markets. The exotic average feature of Asian options is especially designed for those thinly traded assets whose prices can be manipulated easily. Hence, the average component in Asian options allows their holders to control the risk of manipulation better. Another usage of Asian options, which is not emphasized in the literature, is the diversification of volatile market prices. In this section, we study the effect of stochastic volatilities on the pricing of Asian options and give closed-form pricing formulas for two cases: a mean-reverting Ornstein-Uhlenbeck process and a mean-reverting square-root process. This is a natural extension of our modular pricing by using the Fourier transform. An Asian option is usually cheaper than its standard counterpart. Thus, it is also an economic alternative for hedging and other risk management.

A still unsolved problem in financial economics is that one can not find a closed-form solution for arithmetic average Asian options even in the Black-Scholes world. However, the closed-form solution for Asian options with a geometric average is available since the geometric average of stock prices is still lognormally distributed. If the time to maturity of an option has the same length as the average period specified in the option contract, we have the following formula for Asian options with continuous geometric average:

$$C_{asian} = S_0 e^{-(\frac{1}{2}rT+\frac{1}{12}v^2T)}N(d+\sqrt{T/3}v) - Ke^{-rT}N(d), \tag{10.54}$$

where

$$d = \frac{\ln(S_0/K)+\frac{1}{2}T(r-\frac{1}{2}v^2)}{\sqrt{T/3}v}.$$

In comparison with the Black-Scholes formula for standard options, this formula has two additional correction terms: $1/2$ for $T(r-\frac{1}{2}v^2)$ and $\sqrt{1/3}$ for $\sqrt{T}v$. This is the reason why Asian options are cheaper than their standard counterparts. In the following subsections, at first, we explain how these two terms enter the valuation formula. Next, we incorporate stochastic factors including random jumps into the pricing formulas. The third subsection is about the approximation methods for arithmetic average Asian options. Finally, we deal with a pricing model for Asian interest rate options.

10.4.2 The Black-Scholes World

In this subsection, we give an alternative pricing formula to (10.54) by applying the Fourier inversion. This formula is equivalent to the existing one. The purpose of giving an alternative pricing formula, which, at first glance, is more complicated, is twofold. Firstly, we show why hedging with Asian options is different from hedging with plain vanilla options. In other words, why we need to modify the method of risk-neutral pricing used for plain vanilla options. Secondly, we make preparations for the valuation of Asian options with stochastic factors. Some notations are defined here:

- n_0 : the number of total observation points in the whole average period.
- n : the number of remaining observation points for calculating the average stock price. For out-of-the-period (OTP) Asian options we have $n_0 = n$.
- T_0 : time from the current time to the first observation point. For in-the-period (ITP) or at-the-period (ATP) Asian options $T_0 = 0$.
- h : observation frequency defined by the whole average period dividing n.

Instead of using the arithmetic average, we use the geometric average of the observed stock prices to price Asian options. The discrete geometric average is defined as follows:

$$GA(n) = \left(\prod_{j=1}^{n} S_j\right)^{1/n}. \tag{10.55}$$

where S_j is the stock price at the j-th observation point. The price of an Asian option is then generally given by

$$
\begin{aligned}
C_{asian} &= e^{-rT}\mathbf{E}\left[\{(nGA(n)+A)/n_0-K\}\cdot\mathbf{1}_{\{nGA(n)+A\geqslant K\cdot n_0\}}\right] \\
&= e^{-rT}\mathbf{E}\left[\{(nGA(n)+A)/n_0-K\}\cdot\mathbf{1}_{\{GA(n)\geqslant(K\cdot n_0-A)/n\}}\right],
\end{aligned}
\tag{10.56}
$$

where A is the sum of stock prices from the first day of the average period to the current time, and is equal to zero for the OTP- and ATP-Asian options.

Defining

$$
K_0=\frac{K\cdot n_0-A}{n} \quad \text{and} \quad K_1=K-\frac{A}{n_0}, \tag{10.57}
$$

we have

$$
\begin{aligned}
C_{asian} &= e^{-rT}\mathbf{E}\left[\left(\frac{n}{n_0}GA(n)-K_1\right)\cdot\mathbf{1}_{\{GA(n)\geqslant K_0\}}\right] \\
&= e^{-rT}\frac{n}{n_0}\mathbf{E}\left[GA(n)\cdot\mathbf{1}_{\{GA(n)\geqslant K_0\}}\right]-e^{-rT}K_1\mathbf{E}\left[\mathbf{1}_{\{GA(n)\geqslant K_0\}}\right].
\end{aligned}
\tag{10.58}
$$

What we need to do is calculating the CFs of $GA(n)$ under the risk-neutral measure. We consider the Black-Scholes world where the volatility is constant. The logarithm of $GA(n)$ can be expanded as follows:

$$
\begin{aligned}
\ln GA(n) &= \frac{1}{n}\ln\left(\prod_{j=1}^{n}S_j\right)=\frac{1}{n}\sum_{j=1}^{n}\ln S_j=\frac{1}{n}\sum_{j=1}^{n}x_j \\
&= x_0+(r-\frac{1}{2}v^2)\frac{1}{n}\sum_{j=1}^{n}T_j+\frac{v}{n}\sum_{j=1}^{n}W_j \\
&= x_0+(r-\frac{1}{2}v^2)\left(T_0+\frac{h(n+1)}{2}\right)+\frac{v}{n}\sum_{j=1}^{n}W_j.
\end{aligned}
\tag{10.59}
$$

So we have

$$
\begin{aligned}
&\mathbf{E}[\exp(i\phi\ln GA(n)] \\
&= \mathbf{E}\left[\exp\left(i\phi\left(x_0+(r-\frac{1}{2}v^2)\left(T_0+\frac{h(n+1)}{2}\right)+\frac{v}{n}\sum_{j=1}^{n}W_j\right)\right)\right] \\
&= \mathbf{E}\left[\exp\left(i\phi x_0+i\phi(r-\frac{1}{2}v^2)t_1+\frac{i\phi v}{n}\sum_{j=1}^{n}W_j\right)\right] \\
&= \exp\left(i\phi x_0+i\phi(r-\frac{1}{2}v^2)t_1\right)\mathbf{E}\left[\exp\left(\frac{i\phi v}{n}\sum_{j=1}^{n}W_j\right)\right] \\
&= \exp\left(i\phi x_0+i\phi(r-\frac{1}{2}v^2)t_1\right)\exp\left(\frac{-\phi^2}{2}Var\left(\frac{v}{n}\sum_{j=1}^{n}W_j\right)\right)
\end{aligned}
\tag{10.60}
$$

with
$$t_1 = T_0 + \frac{h(n+1)}{2}.$$

To calculate $Var\left(\frac{\nu}{n}\sum_{j=1}^{n} W_j\right)$, the following equation

$$\begin{aligned}
\frac{1}{n}\sum_{j=1}^{n} W_j &= \frac{1}{n}\sum_{j=1}^{n}(W(T_0) + j\Delta W_j) = W(T_0) + \sum_{j=1}^{n}\frac{j}{n}\Delta W_j \\
&= W(T_0) + \sum_{j=1}^{n}\frac{j}{n}W(h)
\end{aligned}$$

is the key to the following calculation. Due to the Markovian property of the Brownian motion we immediately obtain

$$\begin{aligned}
\exp\left(\frac{\phi^2}{2}Var\left(\frac{\nu}{n}\sum_{j=1}^{n} W_j\right)\right) &= \exp\left(\frac{\nu^2\phi^2}{2}\left(T_0 + \sum_{j=1}^{n}\frac{j^2}{n^2}h\right)\right) \\
&= \exp\left(\frac{\nu^2\phi^2}{2}\left(T_0 + \sum_{j=1}^{n}\frac{j^2}{n^2}h\right)\right) \\
&= \exp\left(\frac{\nu^2\phi^2}{2}\left(\frac{h(n+1)(2n+1)}{6n} + T_0\right)\right) \\
&= \exp\left(\frac{\nu^2\phi^2}{2}t_2\right),
\end{aligned}$$

where
$$t_2 = \frac{h(n+1)(2n+1)}{6n} + T_0.$$

Therefore we have

$$\begin{aligned}
f_2(\phi; \ln GA(n)) &= \mathbf{E}[\exp(i\phi \ln GA(n)] \\
&= \exp\left(i\phi x_0 + i\phi(r - \frac{1}{2}\nu^2)t_1 - \frac{\nu^2\phi^2}{2}t_2\right). \qquad (10.61)
\end{aligned}$$

Now we consider the following expected value,

$$\mathbf{E}\left[GA(n)\cdot \mathbf{1}_{\{GA(n)\geqslant K_0\}}\right].$$

In order to guarantee the risk-neutral pricing and a suitable hedging, we have to find a suitable process $M(t)$ such that $GA(n)/S_0M(t)$ is a likelihood process with an expected value of one. $M(t) = e^{rt}$, usually suggested for standard options, obviously can no longer satisfy the required condition since $\mathbf{E}[GA(n)] = \exp\left(x_0 + (r - \frac{1}{2}\nu^2)t_1 + \frac{\nu^2}{2}t_2\right)$. This means that the usual arbitrage argument for European options is no longer valid for Asian options. An appropriate $M(t)$ should be the term $\exp\left((r-\frac{1}{2}v^2)t_1+\frac{v^2}{2}t_2\right)$, with which we have $\mathbf{E}\left[GA(n)/S_0M(t)\right]=1$. Thus,

$$\mathbf{E}\left[GA(n)\cdot\mathbf{I}_{\{GA(n)\geqslant K_0\}}\right]=S_0M(t)\mathbf{E}\left[\frac{GA(n)}{S_0M(t)}\cdot\mathbf{I}_{\{GA(n)\geqslant K_0\}}\right] \tag{10.62}$$

with

$$M(t)=\exp\left((r-\frac{1}{2}v^2)t_1+\frac{v^2}{2}t_2\right).$$

The expectation in the above equality implies a probability measure due to the special term $S_0M(t)$. It is well worth discussing $M(t)$ in more detail because it includes some useful information for hedging. Due to the correlation among the sequence of stock prices, t_2 is somehow smaller than t_1 and $M(t)$ can not be reduced to e^{rt}. Hence, the volatility v plays an important role in discounting $GA(n)$ to S_0, which is not the case for plain vanilla options. The larger the volatility v is, the smaller the value of the "discounted" spot price $S_0\exp\left((r-\frac{1}{2}v^2)t_1+\frac{v^2}{2}t_2\right)$ is. This effect shows the usefulness and effectiveness of Asian options in eliminating the undesired impact of volatility on option prices. The CF of $\mathbf{E}\left[\frac{GA(n)}{S_0M(t)}\cdot\mathbf{1}_{\{GA(n)\geqslant K_0\}}\right]$ is given by

$$\begin{aligned} f_1(\phi;\ln GA(n)) &= \mathbf{E}\left[\frac{GA(n)}{S_0M(t)}\exp(i\phi\ln GA(n)\right] \\ &= \exp\left(i\phi x_0+i\phi r-\frac{i\phi}{2}v^2t_1+\frac{v^2(2i\phi-\phi^2)}{2}t_2\right). \end{aligned} \tag{10.63}$$

Hence, the pricing formula for Asian stock options with the discrete geometric average may be given by

$$C_{asian}=\exp\left(-rT+(r-\frac{1}{2}v^2)t_1+\frac{v^2}{2}t_2\right)\frac{n}{n_0}S_0F_1-e^{-rT}K_1F_2. \tag{10.64}$$

where

$$F_j=\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{f_j(\phi;\ln GA(n))\exp(-i\phi\ln K_0)}{i\phi}\right)d\phi,\quad j=1,2. \tag{10.65}$$

If n approaches infinity, we actually arrive at the pricing formula for options with the continuous geometric average. In the case of ATP-Asian options, we have $t_1=\frac{1}{2}T$ and $t_2=\frac{1}{3}T$. Consequently, the pricing formula for Asian options with a continuous geometric average is given by

$$C_{asian}=\exp\left(-\frac{rT}{2}-\frac{v^2T}{12}\right)\frac{n}{n_0}S_0F_1-e^{-rT}K_1F_2, \tag{10.66}$$

which is identical to formula (10.54).

10.4.3 Asian Options in a Stochastic World

In this subsection, we take stochastic volatilities and stochastic interest rates into account in pricing geometric average Asian options. Intuitively, $\ln GA(n)$ should follow a Brownian motion since $GA(n)$ is lognormally distributed. In order to obtain the concrete form of the process of $\ln GA(n)$, we consider $\mathbf{E}[GA(n)]$. From the results in the above subsection, we rewrite

$$\begin{aligned}
\mathbf{E}[GA(n)] &= \mathbf{E}[\exp(\ln GA(n)] \\
&= \exp\left(x_0 + (r - \frac{1}{2}v^2)t_1 + \frac{v^2}{2}t_2\right) \\
&= \mathbf{E}\left[\exp\left(x_0 + \frac{t_1}{T}(r - \frac{1}{2}v^2)T + (\sqrt{t_2/T}v)W_T\right)\right] \\
&= \mathbf{E}\left[\exp\left(x_0 + \int_0^T \frac{t_1}{T}(r - \frac{1}{2}v^2)dt + \int_0^T (\sqrt{t_2/T}v)dW\right)\right].
\end{aligned} \tag{10.67}$$

This expression with stochastic integral implies a "visual" stochastic process of $\ln GA(n)$, which takes the following form

$$\begin{aligned}
d\ln GA(t;n) &= \frac{t_1}{T}\left(r - \frac{1}{2}v^2\right)dt + \sqrt{\frac{t_2}{T}}vdW, \\
&= \varepsilon_1\left(r - \frac{1}{2}v^2\right)dt + (\sqrt{\varepsilon_2}v)dW, \qquad \ln GA(0;n) = x_0,
\end{aligned} \tag{10.68}$$

with

$$\varepsilon_1 = \frac{t_1}{T}, \qquad \text{and} \qquad \varepsilon_2 = \frac{t_2}{T}.$$

In the case of ATP-Asian options with infinity n, $t_1/T = 1/2$ and $t_1/T = 1/3$. The process can be rewritten as

$$d\ln GA(t) = \frac{1}{2}\left(r - \frac{1}{2}v^2\right)dt + \sqrt{\frac{1}{3}}vdW, \qquad \ln GA(0;n) = x_0. \tag{10.69}$$

We call such a process of $\ln GA(t)$ "visual" since this process can not be directly observed in the marketplaces and is derived from the stock price process. The derivation of the rigorous process of $GA(t;n)$ in the presence of stochastic volatilities and interest rates is mathematically very cumbersome and does not result in a tractable form. To incorporate stochastic factors into geometric average Asian options, we have the following assumption:

Assumption: The geometric average of n sequential stock prices of Asian options $GA(t;n)$ follows the process (10.68) even in the presence of stochastic volatilities and stochastic interest rates.

If the pricing formula for Asian options implies no arbitrage, as mentioned above, then there exists a process $M(t)$ such that the following equation holds,

$$\mathbf{E}\left[\frac{GA(T)}{S_0 M(T)}\right] = 1 \Longleftrightarrow M(T) = \mathbf{E}\left[GA(T)\right]/S_0. \tag{10.70}$$

Consequently, the CFs $f_1(\phi)$ and $f_2(\phi)$ are given by

$$\begin{aligned} f_1(\phi) &= \mathbf{E}\left[\frac{GA(n)}{S_0 M(T)}\exp(i\phi \ln GA(n)\right] = \mathbf{E}\left[\frac{\exp((1+i\phi)\ln GA(n))}{S_0 M(T)}\right] \\ &= \mathbf{E}\left[\frac{\exp((1+i\phi)\ln GA(n))}{\mathbf{E}\left[GA(T)\right]}\right] = \mathbf{E}\left[\frac{g(1+i\phi)}{\mathbf{E}\left[g(1)\right]}\right] \end{aligned} \tag{10.71}$$

and

$$f_2(\phi) = \mathbf{E}\left[g(i\phi)\right], \tag{10.72}$$

where the function $g(a) = \exp(a \ln GA(n))$. Again, in order to calculate the CFs, we classify the volatilities in two popular cases.

Case 1: A mean-reverting Ornstein-Uhlenbeck process.

Similar to the Schöbel and Zhu (1999) model, we may compute the following expected value,

$$\begin{aligned} &\mathbf{E}\left[g(a)\right] \\ &= \mathbf{E}\left[\exp(a\ln GA(n))\right] \\ &= \mathbf{E}\left[\exp\left(ax_0 + a\int_0^T \varepsilon_1(r - \frac{1}{2}v(t)^2)dt + a\int_0^T (\sqrt{\varepsilon_2}v(t))dW(t)\right)\right] \\ &= \mathbf{E}\left[\exp\left(ax_0 + a\varepsilon_1 rT - \left(\frac{a\varepsilon_1}{2} - \frac{a\sqrt{\varepsilon_2}\rho\kappa}{\sigma} - \frac{a^2\varepsilon_2(1-\rho^2)}{2}\right)\int_0^T v^2(t)dt\right.\right. \\ &\quad \left.\left. - \frac{a\sqrt{\varepsilon_2}\rho\kappa\theta}{\sigma}\int_0^T v(t)dt + \frac{a\sqrt{\varepsilon_2}\rho}{2\sigma}v^2(T) - \frac{a\sqrt{\varepsilon_2}\rho}{2\sigma}v_0^2 - \frac{a\sqrt{\varepsilon_2}\rho\sigma T}{2}\right)\right]. \end{aligned} \tag{10.73}$$

Case 2: A mean-reverting square-root process.

As in the Heston (1993) model, computing the following expected value is straightforward,

$$\begin{aligned} &\mathbf{E}\left[g(a)\right] \\ &= \mathbf{E}\left[\exp(a\ln GA(n))\right] \\ &= \mathbf{E}\left[\exp\left(ax_0 + a\int_0^T \varepsilon_1(r - \frac{1}{2}V(t)dt + a\int_0^T (\sqrt{\varepsilon_2 V(t)})dW(t)\right)\right] \\ &= \mathbf{E}\left[\exp\left(ax_0 + a\varepsilon_1 rT - \left(\frac{a\varepsilon_1}{2} - \frac{a\sqrt{\varepsilon_2}\rho\kappa}{\sigma} - \frac{a^2\varepsilon_2(1-\rho^2)}{2}\right)\int_0^T V(t)dt\right.\right. \\ &\quad \left.\left. + \frac{a\sqrt{\varepsilon_2}\rho}{\sigma}V(T) - \frac{a\sqrt{\varepsilon_2}\rho}{\sigma}V(0) - \frac{a\sqrt{\varepsilon_2}\rho\kappa\theta T}{\sigma}\right)\right]. \end{aligned} \tag{10.74}$$

The values of $f_1(\phi)$, $f_2(\phi)$ and $M(T)$ can be computed for a as $1+i\phi, i\phi$ and 1, respectively. The pricing formula for Asian options with stochastic volatilities is therefore given by the following form,

$$C_{asian} = e^{-rT}\left(S_0 M(T)\frac{n}{n_0}F_1 - K_1 F_2\right). \tag{10.75}$$

For these two cases, we have integrated stochastic volatility into the geometric Asian options. Integrating stochastic interest rates follows a similar line. We now consider the issue of incorporating jumps into Asian options and demonstrate the method by applying lognormal jumps. From (7.20) and (10.68), we obtain the following risk-neutral process

$$d\ln GA(t;n) = \varepsilon_1\left(r - \lambda\mu_J - \frac{1}{2}v^2\right)dt + (\sqrt{\varepsilon_2}v)dW(t) + \ln(1+J)\frac{1}{n}\sum_{j=1}^{n} dY_j \tag{10.76}$$

with

$$Y_j = Y(T_0 + hj), \qquad \ln GA(0;n) = x_0.$$

Note that the jump component is independent of the volatility, we only need to calculate the additional jump terms for CF $f_1(\phi)$:

$$\begin{aligned}
f_1^{PJ} &= \mathbf{E}\left[\exp\left(\ln(1+J)\frac{1+i\phi}{n}\sum_{j=1}^{n} Y_j\right) - (1+i\phi)\lambda\mu_J t_1\right] \\
&= \mathbf{E}\left[\exp\left(\ln(1+J)\frac{1+i\phi}{n}[nY_0 + \sum_{j=1}^{n}(n-j+1)\Delta Y_j]\right) - (1+i\phi)\lambda\mu_J t_1\right] \\
&= \exp\left(\lambda T_0[(1+\mu_J)^{(1+i\phi)}e^{\frac{1}{2}i\phi(1+i\phi)\sigma_J^2} - 1] - (1+i\phi)\lambda\mu_J t_1\right) \\
&\times \mathbf{E}\left[\exp\left(\ln(1+J)\frac{1+i\phi}{n}\sum_{j=1}^{n}(n-j+1)\Delta Y_j\right)\right],
\end{aligned}$$

Using the Markovian property of the Poisson process, we have

$$\begin{aligned}
f_1^{PJ} &= \exp\left(\lambda T_0[(1+\mu_J)^{(1+i\phi)}e^{\frac{1}{2}i\phi(1+i\phi)\sigma_J^2} - 1] - (1+i\phi)\lambda\mu_J t_1\right) \\
&\times \exp\left(\sum_{j=1}^{n}\left(\lambda h[(1+\mu_J)^{k_j}\exp(\frac{1}{2}k_j(k_j-1)\sigma_J^2) - 1]\right)\right) \\
&= \exp\left(\lambda T_0[(1+\mu_J)^{(1+i\phi)}e^{\frac{1}{2}i\phi(1+i\phi)\sigma_J^2} - 1] - (1+i\phi)\lambda\mu_J t_1\right) \\
&\times \exp\left(\lambda h\sum_{j=1}^{n}(1+\mu_J)^{k_j}\exp(\frac{1}{2}k_j(k_j-1)\sigma_J^2) - \lambda(T-T_0)\right)
\end{aligned}$$

$$f_1^{PJ} = \exp\left(\lambda T_0(1+\mu_J)^{(1+i\phi)} e^{\frac{1}{2}i\phi(1+i\phi)\sigma_J^2} - \lambda T - (1+i\phi)\lambda\mu_J t_1\right) \times \exp\left(\lambda h \sum_{j=1}^{n}(1+\mu_J)^{k_j}\exp(\frac{1}{2}k_j(k_j-1)\sigma_J^2)\right) \tag{10.77}$$

with

$$k_j = \frac{(1+i\phi)(n+j-1)}{n}.$$

For CF $f_2(\phi)$ we have a similar result but with another k_j,

$$\begin{aligned} f_2^{PJ} &= \mathbf{E}\left[\exp\left(\ln(1+J)\frac{i\phi}{n}\sum_{j=1}^{n}Y_j\right) - i\phi\lambda\mu_J t_1\right] \\ &= \exp\left(\lambda T_0(1+\mu_J)^{i\phi} e^{\frac{1}{2}i\phi(i\phi-1)\sigma_J^2} - \lambda T - i\phi\lambda\mu_J t_1\right) \times \\ &\quad \exp\left(\lambda h \sum_{j=1}^{n}(1+\mu_J)^{k_j}\exp(\frac{1}{2}k_j(k_j-1)\sigma_J^2)\right) \end{aligned} \tag{10.78}$$

with

$$k_j = \frac{i\phi(n+j-1)}{n}.$$

10.4.4 Approximations for Arithmetic Average Asian Options

Although Asian options based on the geometric average can be valued analytically by a closed-form formula, these options are not commonplace in the OTC markets. The most popular Asian options are based on the arithmetic average of the underlying asset, which, unfortunately, is no longer lognormally distributed. In other words, the arithmetic average does not follow a geometric Brownian motion. This raises a serious problem in pricing arithmetic average Asian options. In order to fill this gap, some approximation methods based on the pricing formula for geometric average Asian options are suggested. Boyle (1977) was the first to apply the Monte Carlo control-variate method to get the approximative values of arithmetic average Asian options. However, the Monte Carlo method is very time-consuming, inefficient and inaccurate for Asian options as reported in Fu, Madan and Wang (1995).

In this subsection, we briefly review the existing approximation methods that attempt to reduce the bias between the geometric average and arithmetic average, hoping that the previously derived formulas can be used. All these approximations are suggested in the framework of constant volatilities. Nevertheless, the methods presented below attempt to obtain better values for arithmetic average Asian options by correcting the moments of the geometric average, which can be calculated by the CFs. Therefore, the following corrections are also applicable for stochastic volatilities and stochastic interest rates.

(1). **Zhang's approach** (1997).

In his approach, the general mean is the key to associate arithmetic average with geometric average. The general mean is defined as follows:

$$GM(\gamma;a) = \left(\frac{1}{n}\sum_{i=1}^{n} a_i^{\gamma}\right)^{1/\gamma} \qquad \text{for} \quad \gamma > 0. \tag{10.79}$$

We need the following two special properties of $GM(\gamma;a)$ for approximation:

$$GM(1;a) = \left(\frac{1}{n}\sum_{i=1}^{n} a_i\right) = AA(a) \tag{10.80}$$

and

$$\lim_{\gamma\to 0} GM(\gamma;a) = \lim_{\gamma\to 0}\left(\frac{1}{n}\sum_{i=1}^{n} a_i^{\gamma}\right)^{1/\gamma} = \left(\prod_{i=1}^{n} a_i\right)^{1/n} = GA(a). \tag{10.81}$$

Thus, we can consider the arithmetic average and the geometric average as two special values of the function $GM(\gamma;a)$ at the point $\gamma = 1$ and $\gamma = 0$, respectively. Additionally, $GM(\gamma;a)$ is an increasing continuous function in γ. Using Taylor's expansion, we have

$$AA(a) = GM(1) \approx GM(0) + GM'(0) = GA(a)[1 + \frac{1}{2}Var(\ln a)], \tag{10.82}$$

where $GM'(\varphi)$ is the first derivation of GM. Zhang's main result can be summarized as follows:

$$AA(S) \cong kGA(S) \tag{10.83}$$

with

$$k = 1 + \frac{1}{2}EV + \frac{1}{4}(VV + EV)^2, \tag{10.84}$$

where

$$EV = \frac{(n^2-1)h}{6}\left(\frac{1}{2}(r - \frac{1}{2}v^2)^2 h + \frac{1}{n}v^2\right)$$

and

$$VV = \frac{(n^2-1)(3n^2-2)}{15n^3}\left(1 - \frac{1}{2}v^2\right)^2 v^2 h^3.$$

Obviously, when n approaches 1, the number k will degenerate to 1. When n approaches infinity, we obtain the limiting k which is given by

$$k = 1 + \frac{1}{24}\left((r - \frac{1}{2}v^2)nh\right)^2 + \frac{1}{576}\left((r - \frac{1}{2}v^2)nh\right)^4. \tag{10.85}$$

By using the corrected geometric average, we can calculate the approximative prices of the arithmetic average Asian options.

(2). **Vorst's approach** (1990).

Instead of correcting the geometric average, Vorst suggests a simple method to correct the strike price by using the difference between $\mathbf{E}[GA(n)]$ and $\mathbf{E}[AA(n)]$. The new effective strike price is therefore given by

$$K_{new} = K + \mathbf{E}[GA(n)] - \mathbf{E}[AA(n)].$$

Because $\mathbf{E}[GA(n)] \leqslant \mathbf{E}[AA(n)]$, we always have $K_{new} \leqslant K$. This correction uses only the first moment whereas Zhang's correction uses not only the first but also the second moments. The expectation value $\mathbf{E}[AA(n)]$ can be obtained analytically by the standard tool.

(3). **The modified geometric average** (MGA).

This method is more accurate than Vorst's approach as reported by Lévy and Turnbull (1992). The main idea of MGA is that if $\ln AA(n)$ is normally distributed according to $N(\mu, \sigma)$, then $\mathbf{E}[AA(n)] = \exp(\mu + \frac{1}{2}\sigma^2)$. It immediately follows that $\mu = \ln \mathbf{E}[AA(n)] - \frac{1}{2}\sigma^2$. In the framework of risk-neutral pricing, MGA implies the replacement of r by $\ln \mathbf{E}[AA(n)]$, namely,

$$r_{new} = \ln \mathbf{E}[AA(n)]. \tag{10.86}$$

Beside the above mentioned approximation approaches, a number of suggestions have been proposed to improve the performance of pricing arithmetic average options based on the available closed-form formula. These works include Lévy (1990), Turnbull and Wakeman (1991).

10.4.5 A Model for Asian Interest Rate Options

In this subsection, we deal with a special interest rate derivative whose payoff at maturity depends on the average of interest rates in a specified period prior to maturity. Such financial instruments are called Asian interest rate options and, then are a direct function of interest rates. This feature is different from a bond option whose payoff is indirectly affected by the interest rate. We can consider Asian interest rate options as an innovation of the interest rate cap or floor with average attribute. Therefore, Asian interest rate call (put) options are also called cap (floor) on the average interest rates.

Longstaff (1995) addressed the issue of pricing and hedging this particular option in the framework of the Vasicek model and found that it is an important alternative hedging instrument for managing borrowing costs. As Asian stock options do, this type of derivative is always less costly than the corresponding full cap and less sensitive to changes in interest rates. Ju (1997) studied Asian interest rate options by applying the Fourier transform. His idea is that if the density function of the underlying average is known, then the option based on this average can be given in an integral form. The unfavorable feature of his solution is that one does not have a solution form *à la* Black-Scholes and can not get the hedge ratios explicitly. Here we enhance Ju's idea and give a closed-form solution for Asian interest rate options by using the underlying CFs. The following results are valid both for the Vasicek (1977) model and the CIR (1985b) model.

Without loss of generality, we let the notional amount in the option contract be equal to one. Let $y(T) = \int_0^T r(u)du$, the average rate is then $y(T)/T$. We consider a case where the time to maturity is shorter than or equal to the length of the average period, i.e., ATP- or ITP-Asian options. Let A denote the sum of the interest rates from the first day of the average period to the current time. For the call option we have

$$\begin{aligned}
C_{AIR} &= \mathbf{E}\left[\exp\left(-\int_0^T r(u)du\right)\left(\frac{1}{T}\left(A+\int_0^T r(u)du\right)-K\right)^+\right] \\
&= \mathbf{E}\left[\exp\left(-y(T)\right)\left(\frac{y(T)}{T}-K_0\right)^+\right], \quad K_0 = K-\frac{A}{T} \\
&= \mathbf{E}\left[\exp\left(-y(T)\right)\left(\frac{y(T)}{T}-K_0\right)\cdot\mathbf{1}_{(y\geqslant TK_0)}\right] \\
&= T^{-1}\mathbf{E}\left[e^{-y(T)}y(T)\cdot\mathbf{1}_{(y\geqslant TK_0)}\right]-K_0\mathbf{E}\left[e^{-y(T)}\cdot\mathbf{1}_{(y\geqslant TK_0)}\right] \\
&= \frac{V(T)}{T}\mathbf{E}\left[\frac{e^{-y(T)}y(T)}{V(T)}\cdot\mathbf{1}_{(y\geqslant TK_0)}\right]-B(T)K_0\mathbf{E}\left[\frac{e^{-y(T)}}{B(0,T)}\cdot\mathbf{1}_{(y\geqslant TK_0)}\right] \\
&= \frac{V(T)}{T}F_1 - B(0,T)K_0F_2, \qquad (10.87)
\end{aligned}$$

where

$$B(0,T) = \mathbf{E}\left[\exp\left(-\int_0^T r(u)du\right)\right] = \mathbf{E}\left[e^{-y(T)}\right]$$

is a zero-coupon bond price maturing at time T and

$$V(T) = \mathbf{E}\left[\exp\left(-\int_0^T r(u)du\right)\left(-\int_0^T r(u)du\right)\right] = \mathbf{E}\left[e^{-y(T)}y(T)\right]$$

is a discounted price of the "underlying asset" $y(T)$. An interesting feature of Asian interest rate options is that they are discounted by using their own underlying. Obviously, $e^{-y(T)}/B(T)$ and $e^{-y(T)}y(T)/V(T)$ define two martingales, respectively, and guarantee no arbitrage over time. Hence, the characteristic function of F_1 can be given by

$$f_1(\phi) = \mathbf{E}\left[\frac{e^{-y(T)}y(T)}{V(T)}e^{i\phi y}\right] = \mathbf{E}\left[\frac{y(T)e^{(i\phi-1)y(T)}}{V(T)}\right]$$
$$= \mathbf{E}\left[\frac{y(T)}{V(T)}\exp\left((i\phi-1)\int_0^T r(u)du\right)\right]. \tag{10.88}$$

Similarly,

$$f_2(\phi) = \mathbf{E}\left[\frac{e^{-y(T)}}{B(T)}e^{i\phi y(T)}\right] = \mathbf{E}\left[\frac{e^{(i\phi-1)y(T)}}{B(T)}\right]$$
$$= \mathbf{E}\left[\frac{1}{B(T)}\exp\left((i\phi-1)\int_0^T r(u)du\right)\right]. \tag{10.89}$$

$V(T)$ and $f_1(\phi)$ can be calculated by the following feature of expectation,

$$V(T) = -\frac{\partial}{\partial\lambda}\mathbf{E}\left[e^{-\lambda y(T)}\right]|_{\lambda=1} \tag{10.90}$$

and

$$f_1(\phi) = -\frac{1}{V(T)}\frac{\partial}{\partial\lambda}\mathbf{E}\left[e^{-\lambda y(T)}\right]|_{\lambda=1-i\phi}\,. \tag{10.91}$$

Dependent on whether the interest rate is specified either as a mean-reverting square-root process or as a mean-reverting Ornstein-Uhlenbeck process, we can give the full formula of the expression $-\frac{\partial}{\partial\lambda}\mathbf{E}\left[e^{-\lambda y(T)}\right]$.

Case 1: A mean-reverting square root process (CIR, 1985b; Ju, 1997).

$$-\frac{\partial}{\partial\lambda}\mathbf{E}\left[e^{-\lambda y(T)}\right] = -B(0,T)$$
$$\times\left[\frac{2\kappa\theta\left[2\gamma e^{-\gamma T}+\kappa_1(1-e^{-\gamma T})\right]}{4\gamma^3 e^{(\kappa-\gamma)T}}\left[\gamma e^{-\gamma T}(\kappa_1-1)+\kappa_1(1-\gamma T)(1-e^{-\gamma T})+\gamma\right]\right.$$
$$\left.-\frac{r\left(1-e^{-\gamma T}\right)}{\gamma+0.5\kappa_2(1-e^{-\gamma T})}-\frac{\lambda r\sigma^2\left[\gamma T e^{-\gamma T}-0.5(1-e^{-2\gamma T})\right]}{\gamma[\gamma+0.5\kappa_2(1-e^{-\gamma T})]^2}\right] \tag{10.92}$$

with

$$\kappa_1 = \kappa+\gamma, \qquad \kappa_2 = \kappa-\gamma.$$

Case 2: A mean-reverting Ornstein-Uhlenbeck process (Vasicek, 1977).

$$-\frac{\partial}{\partial\lambda}\mathbf{E}\left[e^{-\lambda y(T)}\right]$$
$$= -B(0,T)\left[\left(\frac{\sigma^2}{\kappa^3}-\frac{\theta}{\kappa}\right)(e^{-\kappa T}-1)-\frac{\lambda\sigma^2}{2\kappa^3}(e^{-2\kappa T}-1)+\frac{\lambda^{-2}r}{\kappa}\right]. \tag{10.93}$$

By using these two results, we can finally compute $f_1(\phi)$. We obtain immediately

$$F_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_j(\phi)\frac{-i\phi K_0 T}{i\phi}\right) d\phi, \quad j = 1,2. \tag{10.94}$$

The put option price is given by using the put-call parity,

$$Put_{AIR} = B(0,T)K_0 F_1 - \frac{V(T)}{T}F_2 \tag{10.95}$$

where

$$F_j = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_j(\phi)\frac{-i\phi K_0 T}{i\phi}\right) d\phi, \quad j = 1,2. \tag{10.96}$$

The expressions of (10.87) and (10.95) possess a clear structure and the hedge-ratios can be derived analytically. In comparison with Ju's expression in which F_1 and F_2 are put together in a single integral form, this pricing formula can be understood more easily and interpreted more meaningfully. All of these features are of practical and theoretical importance. Some extensions can also be made, for example, for interest rates with stochastic volatility. (Longstaff and Schwarz, 1992). However, such an extension would be technically rather tedious and cause some difficulties in implementation.

10.5 Correlation Options

10.5.1 Introduction

Accompanied by a radical development of derivative markets worldwide in the past decades, more and more trading activities involving different markets or different products raise an increasing demand for effective new financial instruments for the purpose of hedging and arbitrage. To meet these needs, correlation options are generated as a generic term for a class of options such as exchange options, quotient options and spread options. Usually, the payoffs of correlation options are affected by at least two underlying assets. Hence, the correlation coefficients among these underlying assets play a crucial role in the valuation of correlation options and indicate the nomenclature of such particular options. Among correlation options, spread options have been introduced in several exchanges as their official financial products while quotient options and product options are mainly provided and traded in the OTC markets. The New York Mercantile Exchange (NYMEX) launched an option on a crack (fuel) futures spread on October 7, 1994. The New York Cotton Exchange (NYCE) and the Chicago Board of Trade (CBOT) proposed options on the cotton calendar spread and on the soybean complex spread, respectively.

While correlation options are drawing more and more attention in risk management, their valuation presents a great challenge in finance theory. Particularly, the valuation of spread options is a proxy for the complexity of this pricing issue. The earliest version of the pricing formula for spread options is a simple application of the Black-Scholes formula by assuming that the spread between two asset prices follows a geometric Brownian motion, and is then referred to as one-factor model. The limitations and problems of this extremely simplified model have been discussed by Garman (1992). The two-factor model where the underlying assets follow two distinct geometric Brownian motions, respectively, and the correlation between them is also permitted, requests calculating an integral over the cumulative normal distribution. Shimko (1994) developed a pricing formula for spread options incorporating stochastic convenience yield as an enhanced version of a two factor model. Wilcox (1990) applied an arithmetic Brownian motion to specify the spread movement. The resulting pricing formula, however, is inconsistent with the principle of no arbitrage. Recently, Poitras (1998) modeled the underlying asset as an arithmetic Brownian motion and constructed three partial differential equations (PDE) for two underlying assets and their spread respectively to avoid the arbitrage opportunities occurring in the Wilcox model. But the drawback of specifying asset price process as an arithmetic Brownian motion still remains in his model. Hence, we are falling in a dilemma: specifying asset prices as a geometric Brownian motion leads to double integration in the pricing formula, while modeling asset prices as an arithmetic Brownian motion makes it feasible to obtain a tractable pricing formula, but it is not coherent with standard models for options. Exchange options can be considered as a special case of spread options where the strike price is set to be zero. Due to this contractual simplification, a simple pricing formula *à la* Black-Scholes for exchange options can be derived (Margrabe, 1978). Product options and quotient options are two other examples of correlation options and their valuations in a two-factor model present no special difficulty by applying the bivariate normal distribution function. However, it has not yet been studied how one can price them in an environment of stochastic volatilities and stochastic interest rates.

In this section, we attempt to apply the Fourier transform technique to obtain alternative pricing formulas for exchange options, product options and quotient options. The favorite feature of the Fourier transform is that it presents not only an elegant pricing formula allowing for a single integration, but it also accommodates a general specification of state variables, for example, stochastic volatilities, stochastic interest rates and even jump components. By employing the martingale approach, we can construct some new types of equivalent martingale measures that ensure the absence of arbitrage in the risk-neutral valuation of quotient or product options. However, as in the Black-Scholes world, we can not give a simple tractable pricing formula for spread options incorporating stochastic volatility and stochastic interest rates.

To incorporate stochastic volatilities into correlation options, we specify the asset price process with two independent Brownian motions with stochastic volatilities simultaneously, that is, the asset price processes take the following form,

$$\frac{dS_j(t)}{S_j(t)} = r(t)dt + \sigma_{j1}\sqrt{V_1(t)}dW_1(t) + \sigma_{j2}\sqrt{V_2(t)}dW_2(t), \quad j = 1, 2. \tag{10.97}$$

Setting $X_1(t) = \ln[S_1(t)/S_1(0)]$ and $X_2(t) = \ln[S_2(t)/S_1(0)]$ yields

$$\begin{aligned} dX_j(t) = & \left[r(t) - \frac{1}{2}\sigma_{j1}^2 V_1(t) - \frac{1}{2}\sigma_{j2}^2 V_2(t)\right] dt \\ & + \sigma_{j1}\sqrt{V_1(t)}dW_1(t) + \sigma_{j2}\sqrt{V_2(t)}dW_2(t) \end{aligned} \tag{10.98}$$

with $X_1(0) = 0$ and $X_2(0) = 0$. For the purpose of demonstrating how to model the stochastic volatilities for exchange options, we adopt the Heston model and let the variance $V_j(t)$ follow a square root process,

$$dV_j(t) = \kappa_j[\theta_j - V_j(t)]dt + \sigma_j\sqrt{V_j(t)}dZ_j(t), \quad j = 1, 2. \tag{10.99}$$

$$dW_1(t)dW_2(t) = 0, \quad dW_1(t)dZ_1(t) = \rho_1 dt, \quad dW_2(t)dZ_2(t) = \rho_2 dt. \tag{10.100}$$

For an Ornstein-Uhlenbeck process we can proceed in a same manner. The instantaneous correlation coefficient between dS_1/S_1 and dS_2/S_2 is therefore stochastic and given by

$$\begin{aligned} \rho(t) &= Corr\left(\frac{dS_1(t)}{S_1(t)}, \frac{dS_2(t)}{S_2(t)}\right) \\ &= \frac{\sigma_{11}\sigma_{21}V_1(t) + \sigma_{12}\sigma_{22}V_2(t)}{\sqrt{\sigma_{11}^2 V_1(t) + \sigma_{12}^2 V_2(t)}\sqrt{\sigma_{21}^2 V_1(t) + \sigma_{22}^2 V_2(t)}}. \end{aligned} \tag{10.101}$$

The uncertainty of $\rho(t)$ is caused by the variances $V_j(t)$. Since $V_j(t) \geqslant 0$, the parameters σ_{ij} determine whether these two underlying assets are correlated positively or negatively. By setting $V_j(t)$ to be deterministic, i.e., the values of θ_j, σ_j and κ_j are nil, we obtain the usual two-factor model. In this case, the processes in (10.97) are simplified to be

$$\begin{aligned} \frac{dS_j(t)}{S_j(t)} &= r(t)dt + \sigma_{j1}dW_1(t) + \sigma_{j2}dW_2(t) \\ &= r(t)dt + \eta_j dW_j^*(t), \quad j = 1, 2, \end{aligned} \tag{10.102}$$

where $\eta_j = \sqrt{\sigma_{j1}^2 + \sigma_{j2}^2}$ and the new Wiener processes $W_j^*(t)$ are composed of $W_1(t)$ and $W_2(t)$ with the different weights such that $dW_1^*(t)dW_2^*(t) = \rho dt$. Therefore the model in (10.98) is more extensive than an usual two-factor model.

10.5.2 Exchange Options

A European-style exchange option entitles its holder to exchange one asset for another and can be considered as an extended case of a standard option. To value this type of options, we assume that both assets are risky and follow a geometric Brownian motion. Thus, the strike price is no longer deterministic. In fact, this is the only point which distinguishes an exchange option from a plain vanilla option. Margrabe (1978) developed a pricing formula for exchange options in the Black-Scholes framework. Although exchange options are not traded on organized exchanges, they are implied in many financial contracts, for example, performance incentive fee, margin account and dual-currency option bonds etc. The valuation of exchange options is therefore a basis for understanding and valuing these financial contracts. The payoff of exchange options at expiration is given by

$$C(T) = \begin{cases} S_1(T) - S_2(T) & \text{if} \quad S_1(T) > S_2(T) \\ 0 & \text{if} \quad S_1(T) > S_2(T) \end{cases} . \tag{10.103}$$

Obviously, exchange call options on asset 1 for asset 2 are identical to exchange put options on asset 2 for asset 1. If these two assets are stochastically dependent on each other, then valuation of exchange options is linked to a joint distribution of two assets. Additionally, with the price process of one asset, say S_2, being reduced to be deterministic, exchange options degenerate to the corresponding standard options. According to Margrabe (1978), the pricing formula *à la* Black-Scholes can be expressed as:

$$C = S_1 N(d) - S_2 N(d - v\sqrt{T}) \tag{10.104}$$

with

$$d = \frac{\ln(S_1/S_2) + \frac{1}{2}v^2 T}{v\sqrt{T}}, \qquad v = \sqrt{v_1^2 + v_2^2 - 2\rho v_1 v_2},$$

where $v_i, i = 1,2$, represents the volatility of the i-th asset and ρ is their correlation coefficient. The purpose of this subsection is to develop a model for exchange options capturing stochastic volatilities. On the analogy of the standard option pricing formulas, we can rewrite the pricing formula for exchange options as:

$$\begin{aligned} C(S_1, T; S_2) &= S_1 F_1(S_1(T) \geqslant S_2(T)) - S_2 F_2(S_1(T) \geqslant S_2(T)) \\ &= S_1 F_1\left(\ln\frac{S_1(T)}{S_2(T)} \geqslant 0\right) - S_2 F_2\left(\ln\frac{S_1(T)}{S_2(T)} \geqslant 0\right) \\ &= S_1 F_1\left(X_1(T) - X_2(T) \geqslant \ln\frac{S_2(0)}{S_1(0)}\right) \\ &\quad - S_2 F_2\left(X_1(T) - X_2(T) \geqslant \ln\frac{S_2(0)}{S_1(0)}\right). \end{aligned} \tag{10.105}$$

It is obvious that instead of using a T-forward measure to get a suitable CF for F_2, we need to take S_2 as numeraire to derive the CF for F_2. The way of proceeding is the same as for F_1. Consequently, we obtain two CFs

$$f_1(\phi) = \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\exp\left((i\phi+1)X_1(T) - i\phi X_2(T)\right)\right] \qquad (10.106)$$

and

$$f_2(\phi) = \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\exp\left(i\phi X_1(T) - (i\phi-1)X_2(T)\right)\right]. \qquad (10.107)$$

The exercise probabilities F_1 and F_2 are then

$$F_j = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{f_j(\phi)}{i\phi}\exp\left(-i\phi\ln\frac{S_2(0)}{S_1(0)}\right)\right)d\phi, \quad j=1,2. \qquad (10.108)$$

Under the specifications given in the above subsection, we can calculate $f_1(\phi)$ and $f_2(\phi)$ as follows:

$$\begin{aligned}
& f_1(\phi) \\
&= \mathbf{E}\left[\exp\left(-\int_0^T r(t)dt\right)\exp\left((i\phi+1)X_1(T) - i\phi X_2(T)\right)\right] \\
&= \mathbf{E}\Bigg[\exp\Bigg(-\int_0^T r(u)du + (i\phi+1)\Bigg(\int_0^T r(u)du - \frac{1}{2}\sigma_{11}^2\int_0^T V_1(u)du \\
&- \frac{1}{2}\sigma_{12}^2\int_0^T V_2(u)du + \sigma_{11}\int_0^T \sqrt{V_1(u)}dW_1(u) + \sigma_{12}\int_0^T \sqrt{V_2(u)}dW_2(u)\Bigg) \\
&- i\phi\Bigg(\int_0^T r(u)du - \frac{1}{2}\sigma_{21}^2\int_0^T V_1(u)du - \frac{1}{2}\sigma_{22}^2\int_0^T V_2(u)du \\
&+ \sigma_{21}\int_0^T \sqrt{V_1(u)}dW_1(u) + \sigma_{22}\int_0^T \sqrt{V_2(u)}dW_2(u)\Bigg)\Bigg)\Bigg].
\end{aligned}$$

Denoting

$$\Lambda_j = \int_0^T V_j(u)du \qquad j=1,2,$$

and decomposing $dW_j = \rho_j dZ_j + \sqrt{1-\rho_j^2}dy_j$ with $dZ_j dy_j = 0$, we may obtain

$$
\begin{aligned}
f_1(\phi) &= \exp[-A_1\frac{\rho_1}{\sigma_1}(v_1(0)+\kappa_1\theta_1 T)-A_2\frac{\rho_2}{\sigma_2}(V_2(0)+\kappa_2\theta_2 T)] \\
&\quad\times \mathbf{E}\left[\exp\left(\frac{1}{2}\{A_1^2(1-\rho_1^2)+2A_1\frac{\rho_1\kappa_1}{\sigma_1}-B_1\}\Lambda_1\right.\right. \\
&\quad+\ \frac{1}{2}\{A_2^2(1-\rho_2^2)+2A_2\frac{\rho_2\kappa_2}{\sigma_2}-B_2\}\Lambda_2 \\
&\quad+\ \left.\left.A_1\frac{\rho_1}{\sigma_1}V_1(T)+A_2\frac{\rho_2}{\sigma_2}V_2(T)\right)\right]
\end{aligned}
\tag{10.109}
$$

with

$$
A_j=\sigma_{1j}+i\phi(\sigma_{1j}-\sigma_{2j}) \qquad B_j=\sigma_{1j}^2+i\phi(\sigma_{1j}^2-\sigma_{2j}^2), \quad j=1,2.
$$

By using the usual technique of calculating expectations, we obtain the first CF for stochastic volatility following a mean-reverting square root process. Note that $f_1(\phi)$ in (10.106) and $f_2(\phi)$ in (10.107) are symmetric, we immediately have

$$
\begin{aligned}
f_2(\phi) &= \exp[-A_1^*\frac{\rho_2}{\sigma_2}(V_2(0)+\theta_2 T)-A_2^*\frac{\rho_1}{\sigma_1}(V_1(0)+\theta_1 T)] \\
&\quad\times \mathbf{E}\left[\exp\left(\frac{1}{2}\{A_1^{*2}(1-\rho_2^2)+2A_1^*\frac{\rho_2\kappa_2}{\sigma_2}-B_1^*\}\Lambda_2+A_1^*\frac{\rho_2}{\sigma_2}V_2(T)\right.\right. \\
&\quad+\ \left.\left.\frac{1}{2}\{A_2^{*2}(1-\rho_1^2)+2A_2^*\frac{\rho_1\kappa_1}{\sigma_1}-B_2^*\}\Lambda_1+A_2^*\frac{\rho_1}{\sigma_1}V_1(T)\right)\right]
\end{aligned}
\tag{10.110}
$$

with

$$
A_j^*=\sigma_{2j}+i\phi(\sigma_{2j}-\sigma_{1j}) \qquad B_j^*=\sigma_{2j}^2+i\phi(\sigma_{2j}^2-\sigma_{1j}^2), \quad j=1,2.
$$

With these two CFs being calculated, we have the closed-form pricing formula for exchange call options. It is easy to establish a model including random jumps and stochastic volatilities which follow a mean-reverting Ornstein-Uhlenbeck process by modular pricing. We omit these works here.

It is worth mentioning some additional properties of exchange options: First of all, the two characteristic functions are perfectly symmetric and independent on the specification of interest rates. This means that interest rates are not of any relevance for the values of exchange options. This property is also observable in formula (10.105). Secondly, if S_2 is deterministic, then the pricing formula is simplified to the standard one. Thirdly, the exchange call options on asset S_1 for asset S_2 is identical to the exchange put options on asset S_2 for S_1. Hence, by applying the put-call parity of exchange options $C(S_1;K=S_2)+S_2=P(S_1;K=S_2)+S_1$,[5] we can obtain the corresponding pricing formula for exchange put options.

[5] The put-call parity of exchange options differs from the normal one for European options. We derive this parity by directly applying definitions of exchange put and call options. Once again, interest rates do not play a role in this parity.

10.5.3 Quotient Options

Quotient options are options written on the ratio of two assets and, therefore, are also called ratio options. Assume that two underlying assets follow the processes given in (10.102), then the logarithms of their prices are binormally distributed. A straightforward calculation using the risk-neutral valuation approach yields a pricing formula for quotient options as follows,

$$Cq = e^{-rT+\eta_2(\eta_2-\rho\eta_1)T}\frac{S_1(0)}{S_2(0)}N(d_1) - e^{-rT}KN(d_2) \tag{10.111}$$

with

$$\begin{aligned} d_2 &= \frac{\ln[S_1(0)/KS_2(0)] - \frac{1}{2}T(\eta_1^2-\eta_2^2)}{\sqrt{(\eta_1^2-2\rho\eta_1\eta_2+\eta_2^2)T}}, \\ d_1 &= d_2 + \sqrt{(\eta_1^2-2\rho\eta_1\eta_2+\eta_2^2)T}. \end{aligned}$$

For simplicity, we assume the notional amount of quotient options to be one. Now we give an alternative valuation expression for quotient options. We try to incorporate some stochastic state variables into the pricing formula.

$$\begin{aligned} Cq &= e^{-rT}\mathbf{E}\left[\left(\frac{S_1(T)}{S_2(T)}-K\right)\mathbf{1}_{\left(\frac{S_1(T)}{S_2(T)}\geqslant K\right)}\right] \\ &= e^{-rT}\mathbf{E}\left[\left(\frac{S_1(T)}{S_2(T)}-K\right)\mathbf{1}_{(X_1(T)-X_2(T)\geqslant \ln K+\ln S_2(0)-S_1(0))}\right] \\ &= e^{-rT}\frac{S_1(0)}{S_2(0)}\mathbf{E}\left[\exp(X_1(T)-X_2(T))\mathbf{1}_{(X_1(T)-X_2(T)\geqslant \ln K^*)}\right] \\ &- e^{-rT}\mathbf{E}\left[K\mathbf{1}_{(X_1(T)-X_2(T)\geqslant \ln K^*)}\right] \end{aligned} \tag{10.112}$$

with $K^* = K\,S_2(0)/S_1(0)$. Let $q(t) = \mathbf{E}[\exp(X_1(t)-X_2(t))]$, we can immediately construct a martingale

$$M(t) = \exp(X_1(t)-X_2(t))/q(t),$$

which has an expected value of one and implies a change of measure. Therefore, we obtain the following CFs of F_1 and F_2,

$$\begin{aligned} f_1(\phi) &= \mathbf{E}\left[M(T)\exp(i\phi(X_1(T)-X_2(T))\right] \\ &= \mathbf{E}\left[\exp((1+i\phi)(X_1(T)-X_2(T))/q(T)\right] \end{aligned}$$

and

$$f_2(\phi) = \mathbf{E}\left[i\phi(X_1(T)-X_2(T))\right].$$

The pricing formula for call options takes the form of

$$Cq = e^{-rT} q(T) \frac{S_1(0)}{S_2(0)} F_1 - e^{-rT} K F_2. \tag{10.113}$$

Obviously, for the case of $K = 1$, the quotient options are reduced to exchange options whose value is equal to $S_2(0)Cq$. For the simple two-factor model given in (10.102), we can derive $f_j(\phi)$ as follows:

$$f_1(\phi) = \exp\left(-\frac{T}{2}(1+i\phi)(\eta_1^2 - \eta_2^2) + T(\eta_2^2 - \rho\eta_1\eta_2) + \frac{T}{2}(1+i\phi)^2(\eta_1^2 - 2\rho\eta_1\eta_2 + \eta_2^2)\right) \tag{10.114}$$

and

$$f_2(\phi) = \exp\left(-\frac{T}{2} i\phi(\eta_1^2 - \eta_2^2) - \frac{T}{2}\phi^2(\eta_1^2 - 2\rho\eta_1\eta_2 + \eta_2^2)\right) \tag{10.115}$$

with

$$q(T) = \exp((\eta_2^2 - \rho\eta_1\eta_2)T). \tag{10.116}$$

We can easily find that the exponent term $(\eta_2^2 - \rho\eta_1\eta_2)T$ in $q(T)$ is exactly the exponent term appearing in the pricing formula for quotient options in a two-factor model in (10.111). The CFs and $q(T)$ incorporating stochastic volatilities are much more complicated and are given in Appendix D.

10.5.4 Product Options

Product options are options whose payoff at expiration depends on the product of the prices of two underlying assets. A special example of this type of options is the foreign equity option with domestic strike price or the domestic equity option with foreign strike price. In these cases, two underlying assets are the foreign (domestic) equity and the exchange rate of domestic (foreign) currency against foreign (domestic) currency. For example, let $S_1(t)$ denote the price of a stock quoted in USD, and $S_2(t)$ denote the exchange rate of EUR against USD. Since the strike price K of a foreign domestic option, from the point of view of a German investor, is given in EUR, the terminal payoff of such an option is equal to $Max[\ S_1(T)S_2(T) - K, 0]$. Product options designed to link an asset and a foreign currency can provide international investors with the possibility to not only hedge the undesired variability of asset prices, but also protect them from bearing the exchange rate risks. The existing pricing formula for product options in a two-factor model is given by

$$Cp = e^{(r+\rho\eta_1\eta_2)T} S_1(0)S_2(0)N(d_1) - e^{-rT} KN(d_2) \tag{10.117}$$

with

$$d_2 = \frac{\ln[S_1(0)S_2(0)/K] + 2rT - \frac{1}{2}(\eta_1^2 + \eta_2^2)T}{\sqrt{(\eta_1^2 + 2\rho\eta_1\eta_2 + \eta_2^2)T}},$$
$$d_2 = d_1 + \sqrt{(\eta_1^2 + 2\rho\eta_1\eta_2 + \eta_2^2)T}.$$

Alternatively, according to the payoff function of product options, it is straightforward to have the following pricing equation:

$$\begin{aligned} Cp &= e^{-rT}\mathbf{E}\left[(S_1(T)S_2(T) - K)\,\mathbf{1}_{(S_1(T)S_2(T) \geqslant K)}\right] \\ &= e^{-rT}S_1(0)S_2(0)\mathbf{E}\left[\exp(X_1(T) + X_2(T))\mathbf{1}_{(X_1(T)+X_2(T) \geqslant \ln K^*)}\right] \\ &- e^{-rT}K\mathbf{E}\left[\mathbf{1}_{(X_1(T)+X_2(T) \geqslant \ln K^*)}\right] \end{aligned} \tag{10.118}$$

with $K^* = K/(S_1(0)S_2(0))$. Similarly to quotient options, we construct a martingale

$$M(t) = \frac{\exp(X_1(T) + X_2(T))}{p(t)}$$

with $p(t) = \mathbf{E}[\exp(X_1(T) + X_2(T))]$. The two CFs are then given by

$$\begin{aligned} f_1(\phi) &= \mathbf{E}\left[M(T)\exp(i\phi(X_1(T) + X_2(T))\right] \\ &= \mathbf{E}\left[\exp((1 + i\phi)(X_1(T) + X_2(T))/p(T)\right] \end{aligned}$$

and

$$f_2(\phi) = \mathbf{E}\left[i\phi(X_1(T) + X_2(T))\right].$$

Product options can be priced as follows:

$$Cp = e^{-rT}p(T)S_1(0)S_2(0)F_1 - e^{-rT}KF_2. \tag{10.119}$$

In a two-factor model, the CFs and $p(T)$ take the following concrete form,

$$\begin{aligned} f_1(\phi) = \exp\Bigg(&2i\phi rT - \eta_1\eta_2 T - \frac{T}{2}(1 + i\phi)(\eta_1^2 + \eta_2^2) \\ &+ \frac{T}{2}(1 + i\phi)^2(\eta_1^2 + 2\rho\eta_1\eta_2 + \eta_2^2)\Bigg) \end{aligned} \tag{10.120}$$

and

$$f_2(\phi) = \exp\left(2i\phi rT - \frac{T}{2}i\phi(\eta_1^2 + \eta_2^2) - \frac{T}{2}\phi^2(\eta_1^2 + 2\rho\eta_1\eta_2 + \eta_2^2)\right), \tag{10.121}$$

with

$$p(T) = \exp(2rT + \rho\eta_1\eta_2 T).$$

As in the case of quotient options, the term $e^{-rT}p(T)$ is exactly equal to the term appearing in the pricing formula for product options in a two-factor model in (10.117). The CFs and $p(T)$ with stochastic volatilities are given in Appendix D.

10.6 Other Exotic Options

Exotic options are generally classified into two classes: the first one has an exotic probability pattern, i.e., path-dependent probabilities to exercise options; the other has an unusual payoff pattern, i.e., complex payoff structure at maturity. Barrier options and Asian options are two typical path-dependent options and have been dealt with in previous sections whereas digital options, basket options and chooser options belongs to the second group. Options with both features are very rare and lack necessary intuition in practical hedging and risk management. Generally speaking, options with an exotic payoff pattern can be more easily evaluated than path-dependent options. In this section, we briefly discuss how to value digital options and chooser options in our extended framework.

Digital optionsare also called binary options and have a discontinuous payoff. This means that option holders have either a certain amount of cash or nothing at the time of exercise, dependently on whether the terminal stock price is greater than the strike price or not. Therefore, the prices of digital call options can be expressed by

$$C_{digital} = B(0,T) \cdot Z \cdot F(S(T) > K), \tag{10.122}$$

where Z is the contractual amount of cash. Obviously, the probability $F(S(T) > K)$ corresponds to the probability F_2 in the pricing formula for plain vanilla options, which can be derived by modular pricing according to the particular processes of $S(t)$, volatility $v(t)$ and interest rate $r(t)$ as well as random jumps.

However, due to the discontinuous payoff pattern, hedging with digital options is difficult, especially with near ATM-digital options. If volatility smile is present, the values of digital options are extremely sensitive to the slope of smile. To overcome this problem, there are different approximation methods. The first approach to value a digital call in practice is the so-called call spread, namely,

$$C_{digital}(K) = [C(K-\varepsilon) - C(K+\varepsilon)]/(2\varepsilon).$$

Similarly, put spread may be applied to approximate digital put,

$$P_{digital}(K) = [P(K+\varepsilon) - P(K-\varepsilon)]/(2\varepsilon).$$

The other approach approximates the discontinuous payoff step-function by a continuous function as follows:

$$C(T)_{digital} = \begin{cases} 0 & \text{if} \quad S(T) < K - \varepsilon \\ Z\frac{S(T)-K+\varepsilon}{2\varepsilon} & \text{if} \quad K - \varepsilon \leqslant S(T) \leqslant K + \varepsilon \\ Z & \text{if} \quad S(T) > K + \varepsilon \end{cases}, \tag{10.123}$$

where $\varepsilon > 0$. As $\varepsilon \to 0$, the new payoff will converge to the original digital payoff. The pricing formula for (10.123) can be obtained in the usual way. Hence, the hedge ratios are the limit cases of the hedge ratios of the new pricing formula.

Chooser options give option holders an opportunity to decide at a future time $T_0 < T$ whether options are calls or puts with a predetermined strike price K (identical for both call and put) in the remaining time to expiry $T - T_0$. Therefore, chooser options makes it possible for investors to hedge with ambiguous short or long positions up to time T_0. We have

$$C_{chooser} = B(0,T_0)\max\left[C(S_{T_0};T-T_0),P(S_{T_0};T-T_0)\right], \tag{10.124}$$

where $C(S_{T_0};T-T_0)$ denotes the call price at the time T_0 with a maturity of $T-T_0$. By using the put-call parity, we restate (10.124) as

$$\begin{aligned} C_{chooser} &= B(0,T_0) \\ &\times \max\left[C(S_{T_0};T-T_0), C(S_{T_0};T-T_0) - S_{T_0} + KB(0,T-T_0)\right] \\ &= B(0,T_0) \\ &\times \left[C(S_{T_0};T-T_0) + \max[KB(0,T-T_0) - S_{T_0}, 0]\right] \\ &= C(S_0;T) + P(S_0;T_0,K_0) \end{aligned} \tag{10.125}$$

with

$$K_0 = KB(0,T-T_0).$$

This result implies that a chooser option is composed of a call option with a maturity of T and a strike price of K, and a put option with a maturity of T_0 and a strike price of K_0. Because of the choice property, a chooser option is more expensive than a plain vanilla option. Thus, more choices are accompanied by more costs . Based on the formula (10.125), we can value chooser options with stochastic volatilities, stochastic interest rates and random jumps.

These two examples demonstrate that the valuation of exotic options with unusual payoff can be generally reduced to the valuation problem for normal European options. modular pricing is then applicable to most of the exotic options with unusual payoffs.

10.7 Appendices

C: Proofs for Probabilities Involving a Maximum or a Minimum

Proof of Proposition 10.2.1:

Since $X(t)$ is defined to be $X(t) = \ln S(t) - \ln S_0 = x(t) - x_0$, we have $P = \Pr(m_T^X \leqslant z_1) = \Pr(m_T^X \leqslant \ln K - x_0)$. It is sufficient to prove that the probability $\Pr(m_T^X \leqslant \ln K - x_0)$ given in (10.13) satisfies the Kolmogorov backward equation with respect to the backward variable x_0: (Here we express x and v as x_0 and v_0, respectively, to remind us that they are backward variables)

$$\frac{1}{2}v^2\frac{\partial^2 P}{\partial x_0^2} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial v_0^2} - \frac{1}{2}v^2\frac{\partial P}{\partial x_0} + \kappa(\theta - v)\frac{\partial P}{\partial v_0} = \frac{\partial P}{\partial \tau}$$

subject to the boundary condition

$$P = \Pr(m_T^X \leqslant 0) = 1.$$

We rewrite $P = \Pr(m_T^X \leqslant \ln K - x_0)$ as

$$P = \Pr(m_T^X \leqslant \ln K - x_0) = P_1 + e^{-z_1}P_2$$

with

$$P_1 = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(\ln K - x_0))}{i\phi}\right)d\phi,$$

$$P_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(i\phi(\ln K - x_0))}{i\phi}\right)d\phi.$$

It is clear that $P = \Pr(m_T^X \leqslant 0) = P_1 + P_2 = 1$ and the boundary condition is then fulfilled. Since $v(t)$ follows a mean-reverting Ornstein-Uhlenbeck process, as shown in Section 3.3, the CF $f_2(\phi)$ for the case of $\rho = 0$ can be suggested to have a solution of the form

$$f_2(\phi) = \exp\left(\frac{1}{2}H_3(\tau)v_0^2 + H_4(\tau)v_0 + H_5(\tau)\right).$$

It is easy obtain the derivatives of P_1 with respect to x_0 and v_0 as follows:[6]

[6] For convenience, an operator $\mathscr{G}$ is introduced to express

$$\mathscr{G}(z) = -\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(x(T) - x_0))}{i\phi}z\right)d\phi.$$

$$
\begin{aligned}
\frac{\partial P_1}{\partial x_0} &= -\frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left(f_2(\phi)\frac{\exp(-i\phi(\ln K - x_0))}{i\phi} i\phi \right) d\phi = \mathscr{G}(i\phi),\\
\frac{\partial^2 P_1}{\partial x_0^2} &= \mathscr{G}(-\phi^2), \qquad\qquad \frac{\partial P_1}{\partial v_0} = \mathscr{G}(H_3(\tau)v_0 + H_4(\tau)),\\
\frac{\partial^2 P_1}{\partial v_0^2} &= \mathscr{G}(H_3(\tau) + \{H_3(\tau)v_0 + H_4(\tau)\}^2),\\
\frac{\partial P}{\partial \tau} &= \mathscr{G}\left(\frac{1}{2}(H_3)_\tau^2 v_0 + (H_4)_\tau v_0 + (H_5)_\tau\right).
\end{aligned}
$$

Setting them into the above backward equation yields the following equation:

$$
\begin{aligned}
&-\frac{1}{2}v_0^2(\phi^2 + i\phi) + \frac{1}{2}\sigma^2(H_3(\tau) + \{H_3(\tau)v_0 + H_4(\tau)\}^2)\\
&+\kappa(\theta - v_0)(H_3(\tau)v_0 + H_4(\tau))\\
&= \frac{1}{2}(H_3)_\tau^2 v_0 + (H_4)_\tau v_0 + (H_5)_\tau.
\end{aligned}
$$

After rearranging the equation with respect to v_0^2 and v_0, we arrive at three ODEs:

$$
\begin{aligned}
\frac{1}{2}(H_3)_\tau &= -s_1^* - \kappa H_3 + \frac{1}{2}\sigma^2 H_3^2,\\
(H_4)_\tau &= -s_2^* + \kappa\theta H_3 - \kappa H_4 + \sigma^2 H_3 H_4,\\
(H_5)_\tau &= \kappa\theta H_4 + \frac{1}{2}\sigma^2 H_4^2 + \frac{1}{2}\sigma^2 H_3,
\end{aligned}
$$

which are identical to the ODEs in Appendix A and determine the suggested functions H_3, H_4 and H_5. Hence, P_1 satisfies the Kolmogorov backward equation. For the term $e^{-z_1}P_2$ we can proceed in the same way and finally obtain the same ODEs for $e^{-z_1}P_2$. Since both P_1 and $e^{-z_1}P_2$ fulfill the backward equation, $\Pr(m_T^X \leqslant 0)$ fulfills the same backward equation. The proof is completed.

The probability $\Pr(X(T) \geqslant \ln K - x_0, m_T^X \geqslant \ln H - y_0)$ is two-dimensional, and consequently, must satisfy the backward equation

$$
\frac{1}{2}v_0^2\frac{\partial^2 P}{\partial z^2} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial v_0^2} - \frac{1}{2}v_0^2\frac{\partial P}{\partial z} + \kappa(\theta - v_0)\frac{\partial P}{\partial v_0} = \frac{\partial P}{\partial \tau}
$$

for $z \in (x_0, y_0)$ subject to two boundary conditions

$$
\Pr(X(T) \geqslant \ln K - x_0, m_T^X \geqslant 0) = 0
$$

and

$$
\Pr(X(T) \geqslant \ln K - x_0, m_T^X \geqslant \ln K - x_0) = \Pr(m_T^X \geqslant \ln K - x_0).
$$

The proof follows the same way as done for $\Pr(m_T^X \leqslant 0)$ and is not shown in details.

Proof of Proposition 10.2.2:
The Kolmogorov backward equation is

$$\frac{1}{2}rv^2\frac{\partial^2 P}{\partial x_0^2}+\frac{1}{2}r\sigma^2\frac{\partial^2 P}{\partial r^2}+r(1-v^2)\frac{\partial P}{\partial x_0}+\kappa(\theta-r)\frac{\partial P}{\partial r}=\frac{\partial P}{\partial \tau}$$

subject to the boundary condition

$$P=\Pr(m_T^X\leqslant 0)=1.$$

The boundary condition is obviously satisfied by $\Pr(m_T^X\leqslant 0)$ given in (10.22). Suggesting a solution for $f_2(\phi)$ as in Section 3.3 with $\rho=0$, and carrying out the same procedure shown above, we can prove that $\Pr(m_T^X\leqslant 0)$ is actually equal to the equation (10.22).

Proof of Proposition 10.3.1:
The proof of this proposition parallels the proofs of propositions 10.2.1 and 10.2.2.

D: Derivation of the CFs for Correlation Options

Calculating the CFs for product options and quotient options can be generally reduced to calculating the expectation $\mathbf{E}[\exp(aX_1(T)+bX_2(T))]$ with a and b as arbitrary complex numbers. We expand this expectation in more details as follows (see also Subsection 10.5.2):

$$\begin{aligned}
&\mathbf{E}\left[\exp\left(aX_1(T)+bX_2(T)\right)\right]\\
=&\exp[(a+b)rT-A_1\frac{\rho_1}{\sigma_1}(V_1(0)+\kappa_1\theta_1 T)-A_2\frac{\rho_2}{\sigma_2}(V_2(0)+\kappa_2\theta_2 T)]\\
\times&\,\mathbf{E}\left[\exp\left(\frac{1}{2}\{A_1^2(1-\rho_1^2)+2A_1\frac{\rho_1\kappa_1}{\sigma_1}-B_1\}\int_0^T V_1(u)du+A_1\frac{\rho_1}{\sigma_1}V_1(T)\right.\right.\\
+&\left.\left.\frac{1}{2}\{A_2^2(1-\rho_2^2)+2A_2\frac{\rho_2\kappa_2}{\sigma_2}-B_2\}\int_0^T V_2(u)du+A_2\frac{\rho_2}{\sigma_2}V_2(T)\right)\right]
\end{aligned}$$

with

$$A_j=a\sigma_{1j}+b\sigma_{2j}\qquad B_j=a\sigma_{1j}^2+b\sigma_{2j}^2,\quad j=1,2.$$

By applying the formula (3.16) we can obtain the closed-form solution for the expected value. For $a=1$ and $b=1$, we can calculate $p(T)$ for product options and $q(T)$ for quotient options. Setting different values for a and b enables us to get the closed-form solutions for CFs.

Case 1: Quotient options.

$f_1(\phi)$ has the following coefficients,

$$a = 1 + i\phi, \qquad b = -(1 + i\phi).$$

$f_2(\phi)$ has the following coefficients,

$$a = i\phi, \qquad b = -i\phi.$$

Case 2: Product options.
$f_1(\phi)$ has the following coefficients,

$$a = 1 + i\phi, \qquad b = 1 + i\phi.$$

$f_2(\phi)$ has the following coefficients,

$$a = i\phi, \qquad b = i\phi.$$

Chapter 11
Libor Market Model with Stochastic Volatilities

In this chapter we address the smile modeling with stochastic volatility within the setup of Libor Market Model (LMM). Meanwhile, LMM is a standard pricing model for interest rate derivatives. Since LMM is based on a series of lognormal dynamics, the methods for building up the smile models, particularly with stochastic volatilities, can be adopted from the previous chapters. After a brief introduction to interest derivative markets in Section 11.1, we review in Section 11.2 a standard LMM with which we may gain the first impression of the high dimensionality and the large dependence structure of LMM. Next in Section 11.3, swaption valuation and Swap Market Model (SMM) are explained. The consistent valuation of swaptions and swap-rate linked products (CMS products) with LMM remains a challenge in quantitative finance. The key section in this chapter is Section 11.4 where five stochastic volatility LMMs are discussed. All of these five models apply characteristic functions or moment-generating functions for pricing caps and swaptions to different extent. While CFs do not find significant applications in the models of Andersen and Brotherton-Ratcliffe (2001), and Piterbarg (2003), the models of Zhang and Wu (2006), Zhu (2007) as well as Belomestny, Matthew and Schoenmakers (2007) have used CFs extensively and consequently to arrive at the closed-form pricing formulas for caplets and swaptions. The last section provides some conclusive remarks.

11.1 Introduction

Interest rate derivatives market is the most efficient and liquid one among all derivative markets. In terms of trading volume and exotic variants, interest rate derivatives, represented by caps/floors and swaptions, are more actively traded and innovative than equity and FX counterparts. The high trading volume of interest rate derivatives reflects the higher trading volume of the underlying markets: money, swap and bond markets. Therefore, interest rate derivatives play an important role for corporations and financial institutions in managing and hedging interest risks of treasury, credit and bond portfolios. Additionally, some interest rate structures are

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4_11,

used to enhance potential yields by sophisticated market participants. The basic reference indices for interest rate derivatives are Libors[1] and swap rates, that are the underlyings of caps/floors and swaptions, respectively. In early period of this market, the Black'76 formula is extensively applied to value and quote caps/floors and swaptions. Until now, the implied volatilities of caps and swaptions computed with the Black'76 formula are used as quotations for these plain-vanilla options. However, the unworried application of Black'76 formula for the valuation of interest rate options has been questionable from the point of view of consistent valuation.

Since more than one decade Libor Market Model (LMM), sometimes called the BGM model for the honor of a contribution of Brace, Gaterek and Musiela (1997), and also simultaneously developed by Jamshidian (1997), and Mitersen, Sandmann and Sondermann (1997), has established itself to be the mostly applied benchmark model for interest rate derivatives. LMM has some desired features allowing for an easy calibration to ATM cap and swaption market prices and providing rich flexibility to fit complicated ATM caplet volatility structure. LMM and its counterpart model for swap rates, Swap Market Model (SMM), justify the use of the Black'76 formula for caps/floors and swaptions, and impose financial meanings on the pricing measures associated with these formulas. The main assumption in LMM is that each forward Libor follows a driftless geometric Brownian motion under its own forward measure. The modeling similarity between LMM and the Black Scholes model for equity options becomes evident.

As in equity option market, if we value caps and swaptions with the Black'76 formula, we will encounter the same problem with volatility: the implied volatilities retrieved from the Black'76 prices are different from strike to strike, from maturity to maturity. The smile patterns of caps/floors and swaptions are even more evident than equity counterparts. As a result, capturing smile effect is important for a sound valuation of interest rate derivatives, and the smile modeling becomes a great challenge for interest rate models. Generally, there are two approaches to model smile effect. The first one is the so-called local volatility models that are are already mentioned in Chapter 3. The second approach is the stochastic volatility models. In this chapter, we will first review the standard Libor market model, and then address how to incorporate stochastic volatility into Libor market model. We will see that it is not an easy task to recover smile effects for all Libors and swap rates simultaneously.

[1] Libor becomes meanwhile a synonym for interest rates index fixed in inter-bank markets. For EUR market, however, Euribor sponsored by European Banking Federation in Brüssel is the benchmark rate for most interest rate derivatives since the introduction of EURO in 1999.

11.2 Standard Libor Market Model

11.2.1 Model Setup

Given a tenor structure $(T_0 = 0, T_1, T_2, \cdots, T_M, T_{M+1})$, we define a family of forward Libors

$$L_j(t) = \frac{1}{\tau_j}\Big(\frac{B(t,T_j)}{B(t,T_{j+1})} - 1\Big), \qquad t \leqslant T_j, \tag{11.1}$$

with $\tau_j = T_{j+1} - T_j$ and

$$B(t,T) = \mathbf{E}[\exp(-\int_t^T r(u)du)|\mathscr{F}_t]$$

as the time t price of a zero-coupon bond maturing at time T, and $r(u)$ is a short rate of interest in a given economy. T_{M+1} is the maximal time horizon of the yield curve we are considering. Obviously, all Libors $\{L_j(0), j > 0\}$ fixed at current time $T_0 = 0$ are deterministic, and recover the current yield curve, at least in the given tenor structure. By modeling the set of the Libors $\{L_j(t), j > 0\}$, we also model the dynamics of yield curve. Following the usual LMM set-up (see Brace-Gatarek-Musiela (1997), Jamshidian (1997)), we assume that under the forward measure Q^{j+1} associated with a zero-coupon bond $B(t,T_{j+1})$ as the corresponding numeraire, $L_j(t)$ are specified in LMM to follow a stochastic process

$$dL_j(t) = v_j(t)L_j(t)dW_j(t), \quad 1 \leqslant j \leqslant M. \tag{11.2}$$

Between different Wiener processes dW_j and dW_i exists a correlation $dW_i dW_j = \rho_{ij}dt$. Differently from the model for asset given in Section 2.1 where only one forward measure is used, a family of forward measures $Q^{j+1}, j = 1, 2, \cdots, M$, are defined for the corresponding forward Libors. This indicates the complexity of the M-dimensional model setup (11.2). The above model is also termed as Market Model since the market pricing and quoting formula Black'76 for caps and floors may be derived and justified easily if we further assume that $v_j(t)$ is constant.

Denote CPL_j as the j-th caplet in a cap with a notional one, its payoff at time T_{j+1} is given by

$$CPL_j(T_{j+1}) = \tau_j \max[L_j(T_j) - K, 0].$$

Note that Libor $L_j(T_j)$ is fixed at time T_j and paid at time T_{j+1}. The market standard caps and floors follow this fixing-payment schedule. According to the risk-neutral valuation principle, we have

$$\begin{aligned} CPL_j(0) &= \tau_j \mathbf{E}^Q\Big[\exp\big(-\int_0^{T_{j+1}} r(u)du\big)\max[L_j(T_j) - K, 0]\Big] \\ &= \tau_j \mathbf{E}^Q\Big[\frac{1}{H(0,T_{j+1})}\max[L_j(T_j) - K, 0]\Big], \end{aligned} \tag{11.3}$$

where $H(T_{j+1}) = \exp(\int_0^{T_{j+1}} r(u)du)$ is the money market account up time T_{j+1} and Q stands for the risk-neutral measure. Recalling the Radon-Nikodym derivative $g_2(T_{j+1})$ in (2.9) defined in Chapter 2,

$$g_2(T_{j+1}) = \frac{dQ_{j+1}}{dQ}(T_{j+1}) = \frac{1}{H(T_{j+1})B(0,T_{j+1})},$$

and applying it to the above expectation yields

$$\begin{aligned} CPL_j(0) &= \tau_j B(0,T_{j+1})\mathbf{E}^{Q_{j+1}}\left[\max[L_j(T_j)-K,0]\right] \\ &= \tau_j B(0,T_{j+1})[L_j(0)N(d_1)-KN(d_2)] \end{aligned} \tag{11.4}$$

with

$$d_1 = \frac{\ln(L_j(0)/K)+\frac{1}{2}v_j^2 T_j}{v_j\sqrt{T_j}},$$

$$d_2 = \frac{\ln(L_j(0)/K)-\frac{1}{2}v_j^2 T_j}{v_j\sqrt{T_j}},$$

where we have utilized the model assumption that $L_j(t)$ is a martingale under the forward measure Q^{j+1}. Similarly, the pricing formula for the corresponding floorlet $FLL_j(t,T_{j+1})$ is given by

$$FLL_j(0) = \tau_j B(0,T_{j+1})[KN(-d_2)-L_j(0)N(-d_1)].$$

Caps and floors are a portfolio of successive caplets and floorlets respectively, that are

$$CP_j(0) = \sum_{h=1}^{j} CPL_h(0), \quad j=1,2,\cdots,M,$$

and

$$FL_j(0) = \sum_{h=1}^{j} FLL_h(0), \quad j=1,2,\cdots,M.$$

It is obvious that each Libor in the above setup lives in its own stochastic world that is connected with each other via the correlation ρ_{ij}. In this sense, LMM is so far not finished yet, and is not appropriate for value exotic derivatives that usually involved more than one Libors. In fact, as all Libors are not specified under a single measure, we can not simulate all Libors firmly. Therefore, to establish LMM completely, we shall unify all Libor processes under a single measure. To this end, we define an arbitrage forward measure Q^{k+1} as the target measure for an unified world. Particularly, we have to change the measure Q^{j+1} to Q^{k+1} for the Libor $L_j(t)$, and consider the following three cases:

Case 1: $j > k$, backward change: Instead of arriving at Q^{k+1} by only one step, we change measures step by step, it means that we first change the measure from Q^{j+1} to Q^j, and then from Q^j to Q^{j-1}, and so on, until Q^{k+1}. To the end, we define the corresponding Radon-Nikodym derivative

$$\frac{dQ^{j+1}}{dQ^j}(t) = \frac{B(t,T_{j+1})B(0,T_j)}{B(t,T_j)B(0,T_{j+1})} = \frac{1+\tau_j L_j(0)}{1+\tau_j L_j(t)}. \tag{11.5}$$

Denote $g(t) = \frac{dQ^{j+1}}{dQ^j}(t)$, by applying Itô's lemma we can derive the process of $g(t)$ as follows:

$$dg(t) = \frac{-\tau_j L_j(t)(1+\tau_j L_j(0))}{(1+\tau_j L_j(t))^2} v_j(t) dW_j(t) = g(t)\frac{-\tau_j L_j(t)}{1+\tau_j L_j(t)} v_j(t) dW_j(t).$$

Obviously, $g(t)$ is a driftless geometric Brownian motion, and we have the following Doléanss exponential

$$\mathbf{E}^{j+1}\Big[\exp\Big(\int_0 \gamma_u dW_j(u)\Big)\Big] = 1$$

with

$$\gamma_t = \frac{-\tau_j L_j(t) v_j(t)}{1+\tau_j L_j(t)}.$$

Let Z_j denote the Wiener process W_j under the new measure Q^j. According to the Girsanov theorem it exists the following relation between Z_j and W_j,

$$dZ_j = dW_j - \gamma_t < dW_j, dW_j >= dW_j + \frac{\tau_j L_j(t) v_j(t)}{1+\tau_j L_j(t)} dt.$$

It follows the process of $L_j(t)$ under the measure Q^j,

$$\frac{dL_j(t)}{L_j(t)} = v_j dZ_j = \frac{\tau_j L_j(t) v_j^2(t)}{1+\tau_j L_j(t)} dt + v_j(t) dW_j. \tag{11.6}$$

It is known that the change of measure leads only to the additional drift term in the original process. Repeating the above procedure step by step will collect all $(k-j-1)$ drift terms and yield the following process for $L_j(t)$ under the measure Q^{k+1},

$$dL_j(t) = L_j(t)[\mu_j(t)dt + v_j(t)dW_j(t)] \tag{11.7}$$

with

$$\mu_j(t) = v_j(t)\sum_{h=k+1}^{j} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}.$$

Case 2: $j = k$**, no change**: Obviously, in this case we do not need to change any measure, and the process $L_j(t)$ keeps unchanged.

Case 3: $j < k$**, forward change**: We still change the measures step by step, and define the following Radon-Nikodym derivative,

$$\frac{dQ^{j+1}}{dQ^{j+2}}(t) = \frac{B(t,T_{j+1})B(0,T_{j+2})}{B(t,T_{j+2})B(0,T_{j+1})} = \frac{1+\tau_{j+1}L_{j+1}(t)}{1+\tau_{j+1}L_{j+1}(0)}. \tag{11.8}$$

By setting $g(t) = \frac{dQ^{i+1}}{dQ^{i+2}}(t)$, we can derive

$$dg(t) = g(t)\frac{\tau_{j+1}L_{j+1}(t)}{1+\tau_{j+1}L_{j+1}(t)}v_{j+1}(t)dW_{j+1}(t).$$

The corresponding Doléans's exponential is

$$\mathbf{E}^{j+1}\Big[\exp\Big(\int_0 \gamma_u dW_{j+1}(u)\Big)\Big] = 1$$

with

$$\gamma_t = \frac{\tau_{j+1}L_{j+1}(t)v_{j+1}(t)}{1+\tau_{j+1}L_{j+1}(t)}.$$

Let Z_j denote the Brownian motion W_j under the measure Q^{j+2}. Applying the Girsanov theorem leads to

$$dZ_j = dW_j - \gamma_t < dW_{j+1}, dW_j >= dW_j - \frac{\tau_{j+1}L_{j+1}(t)v_{j+1}(t)}{1+\tau_{j+1}L_{j+1}(t)}\rho_{j,j+1}dt.$$

The process $L_j(t)$ under the measure Q^{j+2} is then given by

$$\frac{dL_j(t)}{L_j(t)} = v_j(t)dZ_j = -\frac{\tau_{j+1}L_{j+1}(t)v_j(t)v_{j+1}(t)\rho_{j,j+1}}{1+\tau_{j+1}L_{j+1}(t)}dt + v_j(t)dW_j. \tag{11.9}$$

Similarly to Case 1, we will finally obtain the process for $L_j(t)$ under the measure Q^{k+1},

$$dL_j(t) = L_j(t)[\mu_j(t)dt + v_j(t)dW_j(t)] \tag{11.10}$$

with

$$\mu_j(t) = -v_j(t)\sum_{h=j+1}^{k}\frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}.$$

Summing up the above results, we have the following theorem that formulates the standard LMM completely.

Theorem 11.2.1. *Given the model set-up in (11.2), the processes of $L_j(t)$ under an arbitrary forward measure $Q^{k+1}, 0 \leqslant k \leqslant M$, are given by*

$$dL_j(t) = L_j(t)[\mu_j(t)dt + v_j(t)dW_j] \tag{11.11}$$

with

$$\begin{cases} \mu_j(t) = v_j(t)\sum_{h=k+1}^{j} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}, & j > k, \\ \mu_j(t) = 0, \quad j = k, \\ \mu_j(t) = -v_j(t)\sum_{h=j+1}^{k} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}, & j < k. \end{cases}$$

This theorem is the key to a consistent simulation of Libors and the correct valuation of interest rate derivatives. In practical applications, two special measures are drawn more attentions and often used, that are the spot measure and the terminal measure. The spot measure is namely the measure Q^1, and the terminal measure is the measure Q^{M+1}.

Spot measure:

$$\mu_j(t) = v_j(t)\sum_{h=1}^{j} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}.$$

Note in this case, all libors L_h with $T_h < t$ are not alive, therefore, it is more rigorous to introduce a time mapping function as done in most literature, $\beta(t) = inf\{h : t < T_h\}$. the drift $\mu_j(t)$ is then rewritten in a "clean" way,[2]

$$\mu_j(t) = v_j(t)\sum_{h=\beta(t)}^{j} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}.$$

Terminal measure:

$$\mu_j(t) = -v_j(t)\sum_{h=j+1}^{M} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}.$$

Intuitively, these two measures are namely the forward measures at both end points of a yield curve. According to the above theorem, all processes of $L_j(t)$ have non-negative drifts under the spot measure, and have non-positive drifts under the terminal measure. In this sense, the simulation of LMM under the terminal measure should be somehow more robust than the simulation under the spot measure.

We now consider an arbitrary interest rate derivative $G(T_n)$ that is fixed at time T_n and will be paid at time $T_p, p \geqslant n$. The payoff of $G(T_n)$ could be to do with the past fixings, and could also depends on the forward Libors. Therefore, we define

$$L(T_n) = (L_1(T_1), L_2(T_2), \cdots, L_{n-1}(T_{n-1}), L_n(T_n), L_{n+1}(T_n), \cdots, L_M(T_n)),$$

where Libors $L_1(T_1), L_2(T_2), \cdots, L_{n-1}(T_{n-1})$ are the historical Libor fixings before T_n, $L_n(T_n)$, and Libors $L_{n+1}(T_n)$, $\cdots$, $L_M(T_n)$ are the forward rates, and determine the uncertainty of $G(T_n), G(T_n) = G(L(T_n))$. According to the risk-neutral valuation,

---

[2] Spot measure Q^1 is the nearest measure to the risk-neutral measure Q among all forward measures. While the money market account for Q collects interests instantaneously up to time t, the money market account for Q^1 in LMM is discrete and collects money firstly from $T_{\beta(t)}$. This subtle difference results in a small difference of drift terms of a Libor process under these two measures. For a concrete form of the risk-neutral dynamics in LMM, please see Brigo and Mercurio (2006).

the price of $G(T_n)$ is given by

$$G(0) = B(0,T_{k+1})\mathbf{E}^{k+1}\big[\frac{G(T_n)}{B^*(T_p,T_{k+1})}\big], \tag{11.12}$$

where $B^*(T_p,T_{k+1})$ represents the corresponding numeraire and is not necessarily a zero-coupon bond. Similarly to the cases for the change of measures, it follows

$$\begin{aligned}
B^*(T_p,T_{k+1}) &= B(T_p,T_{k+1}) = \prod_{j=p}^{k} \frac{1}{1+\tau_j L_j(T_p)}, && p < k+1;\\
B^*(T_p,T_{k+1}) &= 1, && p = k+1;\\
B^*(T_p,T_{k+1}) &= B^{-1}(T_{k+1},T_p) = \prod_{j=k+1}^{p-1} [1+\tau_j L_j(T_j)], && p > k+1.
\end{aligned}$$

Under the spot measure, the pricing formula is simplified to

$$G(0) = B(0,T_1)\mathbf{E}^1\Big[\frac{G(T_n)}{B^*(T_p,T_1)}\Big].$$

Under the terminal measure, we have

$$G(0) = B(0,T_{M+1})\mathbf{E}^{M+1}\Big[\frac{G(T_n)}{B(T_p,T_{M+1})}\Big].$$

If $p = k+1$, according to the martingale property of $B^*(T_p,T_{k+1}) = 1$, we obtain a particular valuation formula

$$G(0) = B(0,T_{k+1})\mathbf{E}^{k+1}\Big[G(T_n)\Big].$$

If we further assume that n is equal to k, and $G(T_k) = \tau_k \max[L_k(T_k) - K, 0]$, we can easily verify that this special case reduces to a standard caplet, and the final closed-form pricing formula is just the Black76.

11.2.2 Term Structure and Smile of Volatility

From the point of view of modeling, LMM is similar to the Black-Scholes model. Both models are based on log-normal processes and result in the similar pricing formulas associated with standard normal distribution. Not surprisingly, as in equity derivatives market, we can observe significant smile/skew structure of implied cap and floor volatilities. Figure (11.1) shows a market flat volatility surface of caps, as of February 11, 2009, which displays a strong smile/skew patterns. The implied volatilities with maturity up to 4 years smile symmetrically, and the long-term implied volatilities decrease with strikes. Flat volatility is a volatility that is valid for all caplets (floorlets) in a cap (floor). In more details, we have the following relation,

$$CP_j(\bar{v}_j) = \sum_{h=1}^{j} CPL_h(\bar{v}_j),$$

that implies all Libors $L_h(t), h \leqslant j$, have the same flat volatility $\bar{v}_j$. An inconsistence arises if we want to retrieve, say, the volatility of the second caplet CPL_2 from two quoted flat volatilities, say, $\bar{v}_4$ and $\bar{v}_5$. Therefore, the quoted flat volatilities by market can not applied to LMM directly. However, a simple bootstrapping allows us to retrieve the caplet volatilities.

Denote $\hat{v}_h$ as the volatility of h-th caplet, we express a cap with caplets as follows

$$\begin{aligned} CP_j(\bar{v}_j; K_i) &= \sum_{h=1}^{j} CPL_h(\hat{v}_h; K_i) \\ &= CP_{j-1}(\bar{v}_{j-1}; K_i) + CPL_j(\hat{v}_j; K_i), \end{aligned} \tag{11.13}$$

from which the caplet volatility $\hat{v}_j$ can be found directly for each strike K_i. Note $\bar{v}_1 = \hat{v}_1$, we will obtain all $\hat{v}_j, 1 < j < M$, recursively from the cap flat volatilities for each strike K_i. A smile surface of caplet volatilities spanned by maturity T_j and strike K_i are shown in Figure (11.2). We can observe that the form and shape of a caplet smile surface is very similar to the counterparts of a cap smile surface.

There is a simple relationship between caplet volatility and instantaneous volatility, the latter is actually used to simulate dynamics of a Libor, and is important to describe the local behavior of a Libor. That is

$$\hat{v}_j^2 T_j = \int_0^{T_j} v_j^2(t)dt. \tag{11.14}$$

By assuming that $v(t)$ is piecewise constant, we rewrite it as

$$\hat{v}_j^2 T_j = \sum_{h=1}^{j} v_j^2(T_h)\tau_h.$$

The piecewise constant $v_j(T_h)$ is also called the forward volatility for the period $[T_{h-1}, T_h]$. In the above expression, it is clear that each caplet volatility $\hat{v}_j$ is decomposed into its own term structure of forward volatility. On the other hand, observing that Libor $L_j(t)$ will become Libor $L_{j-1}(t-\tau_j)$ one period later, it is reasonable to propose a time-homogenous term structure of instantaneous volatility, this means

$$v_j(t) = v_{j-1}(t-\tau_j), \quad 0 < j < M+1,$$

which in turn implies

$$\hat{v}_j^2 T_j = \sum_{h=1}^{j} v_h^2(T_h)\tau_h. \tag{11.15}$$

A simple and useful parametrization for a time-homogeneous volatility term structure is the following

$$v_j(t) = v_j(t,a,b,c,d) := [a(T_j - t) + b]e^{-c(T_j - t)} + d. \tag{11.16}$$

With this formulation, the caplet volatility $\hat{v}_j^2$ may be expressed by

$$\hat{v}_j^2 T_j = \int_0^{T_j} \left[[a(T_j - t) + b]e^{-c(T_j - t)} + d\right]^2 dt.$$

As $\hat{v}_j^2, 0 < j < M+1$, have been retrieved from the market data, the parameters $\{a,b,c,d\}$ may be calibrated to fit the caplet volatilities. A favorite feature of this ABCD parametrization is its ability to capture the different forms of the volatility term structure. For example, the ATM caplet volatility curve in Figure (11.1) displays an inverse humped form with higher short-term and long-term volatilities and lower medium-term volatilities, due to the credit crunch since 2007.[3] Brigo and Mercurio (2006) provided an excellent discussion how to describe $v_h(t)$ with various parametric formulations.

Applying the retrieved forward or instantaneous volatilities to the processes in (11.11), we are able to simulate LMM in a simple manner. The simulation of LMM is a multi-factor simulation that we have briefly examined in Chapter 5 in the context of a multi-asset model for assets. The key for such simulation is to generate correlated Gaussian random variables, perhaps with the help of the Cholesky decomposition. For more advanced topics on the simulation of LMM, the readers are referred to Glasserman (2004).

11.3 Swap Market Model

11.3.1 Model Setup

Interest rate swaps are very liquid instruments and allow two trade participants to exchange fixed cash-flows against floating cash-flows. A swaption allows its holder to enter a pre-determined swap in its maturity. Therefore, besides strike and option maturity, the length of underlying swap is also an important parameter for swaption contract. The payoff of a payer (receiver) swap is given by

$$SW(t;n,m) = \sum_{j=n}^{m-1} \tau_j \alpha [L_j(t) - K] B(t, T_{j+1})$$

where $\alpha = 1$ for payer swap and $\alpha = -1$ for receiver swap. In a reset date when a swap contract is signed, the strike K is chosen to set the present value of the swap

[3] In usual cases, caplet volatility curves are hump-shaped as a camel's back with with lower short-term and long-term volatilities and higher medium-term volatilities.

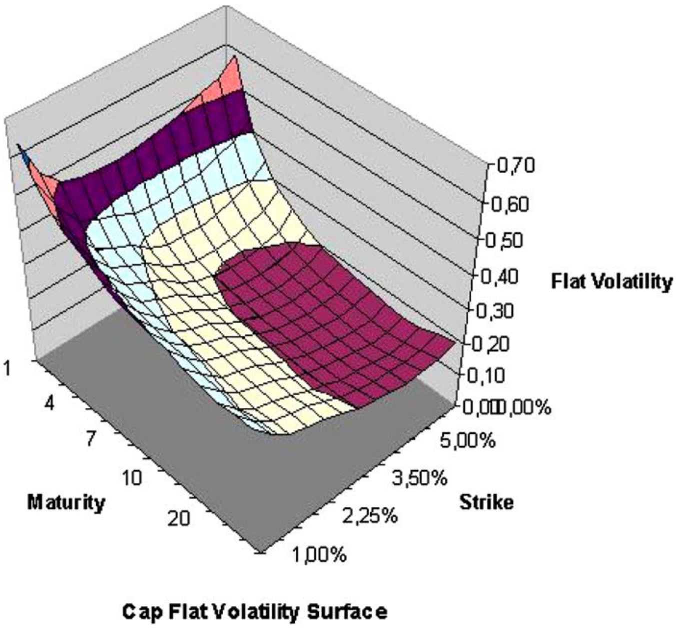

Fig. 11.1 The market flat volatility surface of caps, as of as of February 11, 2009.

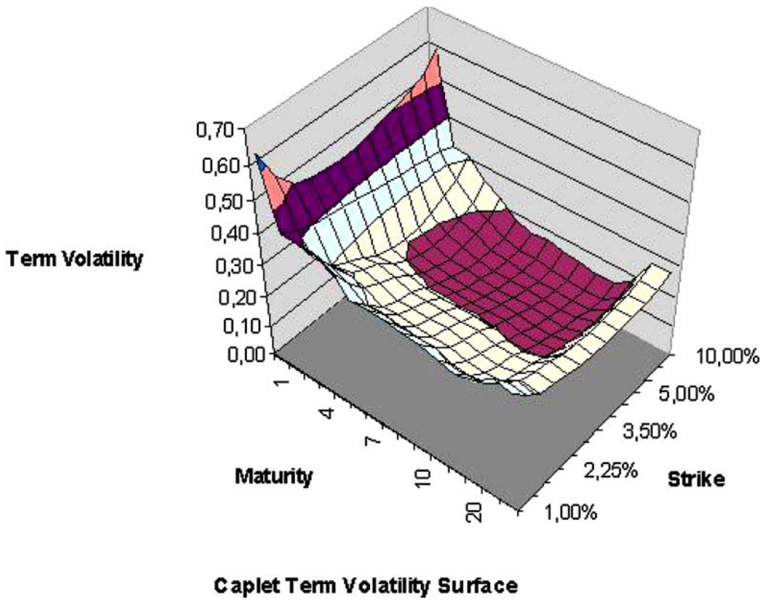

Fig. 11.2 The retrieved implied volatility surface of caplets, as of February 11, 2009.

equal to zero. Such a strike making the value of a swap equal to zero is referred to as swap rate, and may be easily expressed as

$$
\begin{aligned}
S_{n,m}(t) &= \frac{\sum_{j=n}^{m-1} \tau_j L_j(t) B(t,T_{j+1})}{\sum_{j=n}^{m-1} \tau_j B(t,T_{j+1})} \\
&= \frac{B(t,T_n) - B(t,T_m)}{A(t;n,m)},
\end{aligned} \tag{11.17}
$$

with

$$
A(t;n,m) = \sum_{j=n}^{m-1} \tau_j B(t,T_{j+1}),
$$

that is termed as swap annuity or as the present value of a basis point (PVBP), and may be interpreted as a weighted sum of $(m-n-1)$ discount factors during the swap period. Correspondingly, a payer (receiver) swaption has the following payoff

$$
\begin{aligned}
SWT(t;n,m) &= \max[SW(t;n,m),0] \\
&= \max\Big[\sum_{j=n}^{m-1} \tau_j \alpha [L_j(t) - K] B(t,T_{j+1}),0\Big] \\
&= \max\big[\alpha A(t;n,m)[S_{n,m}(t) - K],0\big].
\end{aligned} \tag{11.18}
$$

In the light of the above equation, swaption may be considered as an option with swap rate $S_{n,m}(t)$ as the underlying, and swap annuity as an effective discount factor. In fact, Swap Market Model (SMM) is based on swap measure with swap annuity as numeraire, and the corresponding Radon-Nikodym derivative between the risk-neutral measure and the swap measure is given by

$$
\frac{dQ}{dQ_S}(t) = \frac{H(t)A(0;n,m)}{H(0)A(t;n,m)} = \frac{H(t)A(0;n,m)}{A(t;n,m)}. \tag{11.19}
$$

Furthermore, we assume that the swap rate $S_{n,m}(t)$ under the swap measure follows a driftless geometric Brownian motion,

$$
\frac{dS_{n,m}(t)}{S_{n,m}(t)} = v_s dW(t). \tag{11.20}
$$

A straightforward application of the above numeraire change leads to a Black76-like pricing formula for payer swaption,

$$
\begin{aligned}
SWP(0;n,m) &= \mathbf{E}\Big[\frac{1}{H(T_n)} \max\big[A(T_n;n,m)[S_{n,m}(T_n) - K],0\big]\Big] \\
&= A(0;n,m)\mathbf{E}^S\big[\max[S_{n,m}(T_n) - K,0]\big] \\
&= A(0;n,m)\big[S_{n,m}(0)N(d_1) - KN(d_2)\big]
\end{aligned} \tag{11.21}
$$

with

$$d_1 = \frac{\ln(S_{n,m}(0)/K) + \frac{1}{2}v_s^2 T_n}{v_s\sqrt{T_n}},$$

$$d_2 = \frac{\ln(S_{n,m}(0)/K) - \frac{1}{2}v_s^2 T_n}{v_s\sqrt{T_n}}.$$

This is the broadly applied pricing formula for plain-vanilla swaptions in market, the corresponding implied volatilities are used as market quotations.

There are some important implications of SMM, which may help us better understand the essence of SMM and its relationship to LMM.

1. Firstly, swap measure Q_S associated with the particular numeraire $A(t;n,m)$ depends on n and m, this means that each swap measure is so particular that a SMM is only valid for a single swap rate. It is immensely tedious and involved to unify all swap rates in an unique swap measure.
2. Secondly, it is relatively easy to build up an unified market model for swap rates with a same tenor (see Musiela and Rutkowski (2006), Zhu (2007b)). Since Libor can be regarded as a swap rate with the tenor of one period, such a generalized SMM embraces LMM as a special case.
3. Next, swaptions can be classified into three subgroups according to n and m. As mentioned above, one group is the swaptions with constant tenor, it means that $(m-n)$ is equal to a constant, this group is called co-sliding swaptions or constant-tenor swaptions. Most CMS (Constant Maturity Swap) products use constant-tenor swap rates as underlying index. The second group includes all swaptions with the same terminal date T_m, and is called co-terminal swaptions. For examples, the underlyings of a Bermudan swaption are co-terminal swap rates. The so-called co-initial swaptions is a set of swaptions with the same start date T_n. A few interest rate products use co-initial swap rates as the underlying.
4. Finally, swap rate $S_{n,m}(t)$ is a function of $(m-n)$ Libors, and therefore, dependent on the correlations between these Libors. In this sense, swaptions are also the correlation products of Libors, and the market quotations of swaptions provide us with the information on the correlation structure of Libors. As a result, swaption variance may be approximated by the covariance of the weighted log returns of successive Libors (see Brigo and Mercurio (2006)).

Since swaps could last from one year to 30 years, the market quotations for swaptions have one dimension more than for caps, namely the length of underlying swap, or sometimes called the tenor of swap. Figure (11.3) shows the market ATM swaption volatilities as of February 11, 2009, that also shapes a surface of two dimensions of option maturity and swap tenor. For a certain swap tenor, we can draw an ATM swaption volatility curve from an ATM swaption surface. As shown in Figure (11.3), such a swaption volatility curve, as its caplet counterpart, is inversely humped in February, 2009.

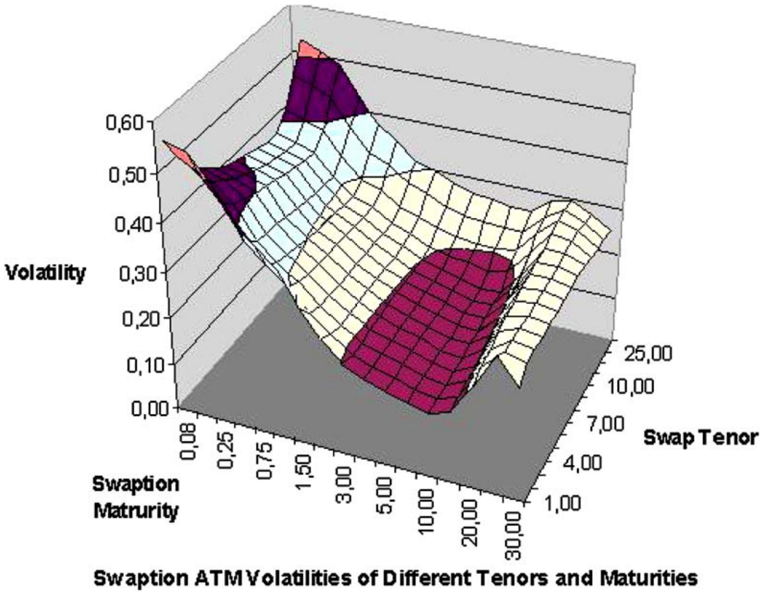

Fig. 11.3 Market ATM swaption volatilities, as of February 11, 2009.

11.3.2 Correlation Structure

We now briefly examine how to value a swaption in a standard Libor Market Model, and by this way, we may better expound how the correlation structure of Libors affects the pricing of swaptions. As seen in the pricing formula for caplet/floorlet in LMM, the values of caps and floors are independent of the correlation coefficients $\rho_{ij}, 0 < i, j < M+1$. Therefore, caps and floors are insensitive to correlations, and do not provide any information on correlations. However, many interest rate derivatives are sensitive to the correlations between Libors, and the simulation of LMM requires a well-defined correlation matrix. On the other hand, we know that swaptions imply such information. As the number of the correlation coefficients is $M(M-1)/2$, a parametrization of the correlation matrix is necessary, in order to make an efficient retrieving of correlations from swaptions feasible. Before we propose a parameterized correlation matrix, we may observe the following stylized facts:

1. Two closely neighbored Libors are more highly correlated than two distanced Libors. This implies that a change of an interest rate in a yield curve causes more changes in its near segment than in distanced segments.
2. The correlation of two Libors should be positive. As a result, a positive shift of an interest rate might leads to positive shifts of other interest rates.

To formulate these facts quantitatively, denote Σ as the matrix of instantaneous correlations,

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1M} \\ \rho_{21} & 1 & \cdots & \rho_{2M} \\ \cdots & \cdots & 1 & \cdots \\ \rho_{M1} & \rho_{M2} & \cdots & 1 \end{pmatrix}, \qquad \rho_{ij} = \rho_{ji} = W_i W_j. \tag{11.22}$$

The first parameterization of Σ is simple and based on a single parameter q,

$$\rho_{ij} = e^{-q|T_i - T_j|}, \qquad q > 0. \tag{11.23}$$

The resulting correlations ρ_{ij} satisfy the above stylized facts. Additionally, for a suitable q, the correlation matrix Σ may be positive semi-definite, and makes sense. The formulation (11.23) is then a simple and still reasonable parameterization.

A small extension of the formulation (11.23) may be the following,

$$\rho_{ij} = \rho + (1-\rho)e^{-q|T_i - T_j|}. \tag{11.24}$$

Obviously, when T_i is far away from T_j, their correlation ρ_{ij} approaches to ρ asymptotically, whereas ρ_{ij} in the formulation (11.23) goes to zero. Therefore, the parameter ρ is interpreted as the asymptotical instantaneous correlation. Rebonato (1999) proposed a similar formulation with three parameters,

$$\rho_{ij} = \rho + (1-\rho)e^{-|i-j|(p - q(\max(i,j)-1))}. \tag{11.25}$$

However, the correlation matrix of this parameterization is often not positive semi-definite. Additional repair of the correlation matrix Σ via correcting eigenvalues of Σ is then necessary.

Schoenmakers and Coffey (2000) suggested several full rank formulations for the correlation matrix. The favorite features of the resulting matrix are always positive semi-definite, and parsimonious in parameters. The following two formulations are of particular interest in practical applications.

1. Robust full-rank formulation:

$$\rho_{ij} = \exp\left[-\frac{|i-j|}{M-1}\left(-\ln\rho + q\frac{M-1-i-j}{M-2}\right)\right]. \tag{11.26}$$

2. Improved robust full-rank formulation:

$$\begin{aligned}\rho_{ij} = \exp\Bigg[-\frac{|i-j|}{M-1}\Bigg(&-\ln\rho \\ &+ q\frac{i^2+j^2+ij-3Mi+3Mj+3i+3j+2M^2-M-4}{(M-2)(M-3)}\Bigg)\Bigg].\end{aligned} \tag{11.27}$$

It should be emphasized that the correlation matrix Σ describes the correlations of the Brownian motions in Libor process $L_j(t)$, not the correlations between Libors themselves. The correlations of Libors involve not only Libors themselves, but also their volatilities. As shown in Brigo-Mercurio (2006) and Rebonato (1999), such Libor correlation could be approximated by

$$\rho(L_i(T_k), L_j(T_k)) \approx \rho_{ij} \frac{\int_0^{T_k} v_i(t) v_j(t) dt}{\sqrt{\int_0^{T_k} v_i^2(t) dt} \sqrt{\int_0^{T_k} v_j^2(t) dt}}, \qquad k \leqslant \min(i,j), \tag{11.28}$$

that is also called terminal correlation in financial literature. It is easy to observe that even if ρ_{ij} is equal to 1, $\rho(L_i(T_k), L_j(T_k))$ may not be 1. Additionally, the volatilities of $L_i(t)$ and $L_j(t)$ have a significant effect on $\rho(L_i(T_k), L_j(T_k))$. An extreme example may be the case where two volatility functions $v_i(t)$ and $v_j(t)$ are nearly orthogonal and makes the terminal correlation near to zero. Therefore, the parametrization of volatility functions has an immediate and direct impact on the values of terminal correlations.

Now we recall the definition of swap rate $S_{n,m}(t)$ in (11.17) and rewrite it as follows

$$\begin{aligned} S_{n,m}(t) &= \frac{\sum_{j=n}^{m-1} \tau_j L_j(t) B(t, T_{j+1})}{\sum_{j=n}^{m-1} \tau_j B(t, T_{j+1})} \\ &= \sum_{j=n}^{m-1} \omega_j(t) L_j(t) \end{aligned} \tag{11.29}$$

with

$$\omega_j(t) = \frac{\tau_j B(t, T_{j+1})}{A(t; n, m)}.$$

Obviously, $\omega_j(t)$ itself is stochastic. In order to express swap rate as a linear combination of Libors, we perform the following approximation

$$S_{n,m}(t) \approx \sum_{j=n}^{m-1} \omega_j(0) L_j(t),$$

with the weights frozen at the origin,

$$\omega_j(0) = \frac{\tau_j B(0, T_{j+1})}{A(0; n, m)}.$$

Note that $\omega_j(t)$ itself depends on zero-coupon bond and therefore also on $L(t)$, Hull and White (1999), Andersen and Andreasen (2000) give the more sophisticated formulations of the weights $\omega_j(t)$ by deriving the complete dynamics of $S_{n,m}(t)$ dependently on $L(t)$.[4] Without the detailed derivation of these two formulations for $\omega_j(t)$, we just give them for reference:

[4] Many numerical examples conducted by author have shown that the difference between all three versions of $w_h(t)$ is negligible in most cases.

Hull and White (1999):

$$\omega_j^{HW}(t) = \omega_j(t) + \sum_{i=n}^{m-1} L_i(t) \frac{\partial \omega_i(t)}{\partial L_j(t)}. \tag{11.30}$$

Andersen and Andreasen (2000):

$$\omega_j^{AA}(t) = \frac{\partial S_{n,m}(t) L_j(t)}{\partial L_j(t) S_{n,m}(t)}. \tag{11.31}$$

It is clear that the swaption annuity $A(t;n,m)$ is a weighted sum of different zero-coupon bonds $B(t,T_{j+1})$, which are used as numeraire for forward measure in LMM. By intuition, a swap measure should be also a weighted sum of different forward measures Q^{j+1}. This indicates that the process of a swap rate can be appropriately approximated by a weighted sum of different Libor processes under their own measures Q^{j+1} and does not involve some specific forward measures. As a consequence, if we approximate the swaption volatility v_s by caplet volatilities, v_s is uniquely determined by the volatilities $v_j(t)$ of Libors $L_j(t)$ and their correlations, and is independent of the drift adjustments $\mu_j(t)$. This feature distinguishes the pricing of a swaption from the pricing of a CMS derivatives. This subtle difference between swap rate and CMS rate leads to the different expression of their stochastic processes, and will be discussed later in more details. Consequently we have the following process for $S_{n,m}(t)$,

$$\begin{aligned} dS_{n,m}(t) &\approx \sum_{j=n}^{m-1} \omega_j(0) dL_j(t) \\ &= \sum_{j=n}^{m-1} \omega_j(0) L_j(t) v_j(t) dW_j. \end{aligned}$$

The next step is re-formulate $dS_{n,m}(t)$ in a log-normal process as follows:

$$\begin{aligned} \frac{dS_{n,m}(t)}{S_{n,m}(t)} &= S_{n,m}(0)^{-1} \sum_{j=n}^{m-1} \omega_j(0) L_j(0) v_j(t) dW_j \\ &= v_S(t) dW_S(t), \end{aligned} \tag{11.32}$$

where the volatility of swap rate takes the following form

$$v_S^2(t) = \sum_{i=n}^{m-1} \sum_{j=n}^{m-1} \frac{\omega_i(0)\omega_j(0) L_i(0) L_j(0)}{S_{n,m}(0)^2} \rho_{ij} v_i(t) v_j(t). \tag{11.33}$$

The approximated swap rate process in (11.32) has no drift and coincides with the one in SMM.

By freezing the dynamics $S_{n,m}(t)$ partially to $S_{n,m}(0)$, we can clearly see the effects of the correlation coefficient ρ_{ij} on the volatility of a swap rate, and at the same time, can conclude that the forward measure Q^k does not change the volatility $v_S(t)$. The formula (11.33) is referred to as the Rebonato formula, and establishes a close relationship between Libors and swaptions within LMM. Given the term structure of Libor volatility $v_j(t)$ and swaption volatilities v_s for different tenor and maturity pair (n,m), we are able to estimate the correlation parameters in the above various correlation formulations by considering the following optimization:

$$\boldsymbol{\Phi} = arg\min_{\boldsymbol{\Phi}} \sum |v_s^2(\boldsymbol{\Phi}) - (v_s^{Market})^2|, \tag{11.34}$$

where $\boldsymbol{\Phi}$ denotes the set of the parameters for correlation matrix. $v_s(\boldsymbol{\Phi})$ and v_s^{Market} are the theoretical and market swaption volatilities respectively.

Another immediate application of the Rebonato formula is to bootstrap the volatility $v_j(t)$ of Libor dynamics given an endogenous correlation matrix. To this end, we assume that $v_j(t)$ is piecewise constant, the Rebonato formula may be rewritten in the following discrete form,

$$v_S^2(T_n) = \frac{1}{T_n} \sum_{i=n}^{m-1} \sum_{j=n}^{m-1} \frac{\omega_i(0)\omega_j(0)L_i(0)L_j(0)}{S_{n,m}(0)^2} \rho_{ij} \sum_{h=0}^{T_{n-1}} v_i(T_h)v_j(T_h)\tau_h. \tag{11.35}$$

Based on this formulation, we have a forward bootstrapping procedure to retrieve the volatility $v_j(t), 0 < j < M+1$:

- Step 1: For the first Libor L_1, $v_1(T_0)$ is just the caplet volatility of L_1. No further calculation is required.
- Step 2: This step includes two small steps.
  1. Step 2a: The volatility $v_S^2(T_1;1,2)$ of swap rate $S(t;1,2)$ has to do with $v_1(T_0)$ and $v_2(T_0)$ only. Given $v_S^2(T_1;1,2)$ and $v_1(T_0)$, we can simply compute $v_2(T_0)$ by applying the formula (11.35).
  2. Step 2b: For the caplet volatility $\hat{v}_2$, we recall the following relationship

     $$\hat{v}_2^2(\tau_0 + \tau_1) = v_2^2(T_0)\tau_0 + v_2^2(T_1)\tau_1.$$

     which allows us to obtain $v_2(T_1)$ immediately. This step in fact fits the caplet volatility of a Libor.

- Step 3: This step includes three small steps.
  1. Step 3a: After $v_i(T_j), 1 \leqslant i \leqslant M-1, 1 \leqslant j \leqslant M-2$, are derived, we attempt to solve $v_M(T_0)$. To this end, we use the volatility $v_S^2(T_1;1,n)$ of $S(t;1,M)$. It is easy to see now that $v_M(T_0)$ is the only variable that is unknown in Rebonato's formula (11.35). The final equation for $v_M(T_0)$ is a quadratic equation. To ensure the solution to be positive, it is reasonable to take the bigger one of two solutions.
  2. Step 3b: We repeat step 3a and use the volatility $v_S^2(T_j;j,M), 0 < j < M-1$, of swap rates $S(t;j,M)$ to obtain $v_M(T_j)$.

3. Step 3c: Similar to Step 2b. We compute $v_M(T_{M-1})$ of L_M via the caplet volatility $\hat{v}_M$.

By denoting $v_i(T_{j-1})$ as v_{ij}, the above bootstrapping procedure delivers a time-inhomogeneous volatility term structure for each Libor, which can be illustrated by the following matrix:

	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots\cdots$	$(T_M, T_{M+1}]$
$L_1(t)$	v_{11}	No	$\cdots\cdots$	No
$L_2(t)$	v_{21}	v_{22}	$\cdots\cdots$	No
$\vdots$	$\vdots$	$\vdots$	$\cdots\cdots$	$\vdots$
$L_M(t)$	v_{M1}	v_{M2}	$\cdots\cdots$	$v_{M,M}$

The bootstrapping method is also called the cascade calibration, and performs a self-calibration to caps and swaption markets. Its advantages lie in the good fitting of LMM to the ATM swaption volatilities and the easy implementation without the application of extensive optimization methods. But the drawbacks of the cascade calibration are also obvious. Firstly, there is often no real solution in Step 3a and 3b. Secondly, the retrieved volatilities often oscillate strongly. As a result, the term structure of volatility is not smooth, and by nature, is not time-homogeneous.

11.3.3 Convexity Adjustments for CMS

Many interest derivatives use the constant-tenor swap rate $S_{n,m}(T_l)$ for a constant $c = m - n$ as the underlying index, and pay it at another time T_{k+1}. Generally, a swap rate discounted with a zero-coupon bond instead of a swap annuity is called a constant maturity swap (CMS) rate. In other words, a CMS rate is a swap rate under a forward measure. Applying Itô's lemma, we can derive the process of $S_{n,m}(t)$ under an arbitrary forward measure Q^{k+1},

$$\begin{aligned}
dS_{n,m}(t) &\approx \sum_{j=n}^{m-1} \omega_j(0) dL_j(t) \\
&\approx \sum_{j=n}^{m-1} \omega_j(0) L_j(t) \mu_j^k dt + \sum_{j=n}^{m-1} \omega_j(0) L_j(t) \sigma_j(t) dW_j \\
&\approx S_{n,m}(t)[\mu_c^k(0) dt + v_S(0) dW_S],
\end{aligned}$$

with

$$\mu_c^k(0) = S_{n,m}(0)^{-1} \sum_{j=n}^{m-1} \omega_j(0) L_j(0) \mu_j^k.$$

To avoid notational confusion, we denote $y(t)$ as CMS rate in this subsection, the process of CMS rate then reads

$$\frac{dy(t)}{y(t)} = \mu_c^k(0)dt + v_S(0)dW_S. \tag{11.36}$$

We can easily observe that the only difference between the dynamics of swap rate and CMS rate is the drift. The drift term μ_c^k in the process for CMS rate is caused by the change of a swap measure to a forward measure Q^{k+1}, and generates the co-called convexity adjustment for CMS rate. To compute convexity adjustment in LMM, we calculate the expected value $\mathbf{E}^{k+1}[y(T_l)], l \leqslant n$, at fixing time T_l under a forward measure Q^{k+1},

$$\begin{aligned} y^* &= \mathbf{E}^{k+1}[y(T_l)] = y(0)\exp(\int_0^{T_l} \mu_c^k(t)dt) \\ &\approx y(0)\exp(\mu_c^k(0)T_l). \end{aligned} \tag{11.37}$$

Note $y(0) = S_{n,m}(0)$, it follows immediately that the convexity adjustment is equal to

$$cv = y^* - y(0) = S_{n,m}(0)\left[\exp(\mu_c^k(0)T_l) - 1\right]. \tag{11.38}$$

According to the choice of the numeraire, we can gauge the convexity adjustment cv approximately. Again, three cases may be examined:

1. The payment time T_{k+1} is larger than the swap maturity time T_m. In this case, all drift terms μ_j^k are negative, and as a result, $\mu_c^k(0)$ is also negative. It follows then a negative convexity adjustment that is caused by delayed payments.
2. The payment time T_{k+1} is smaller than the swap start time T_n, but not smaller than the fixing time T_l. In this case, all drift terms μ_j^k are positive. Consequently, there is a positive $\mu_c^k(0)$ and a positive convexity adjustment. The positive convexity is due to earlier payments.
3. The third case is more interesting. The payment time T_{k+1} is the middle point of T_m and T_n. This leads to a drift term $\mu_c^k(0)$ near to zero. As a result, we will have an almost zero convexity adjustment due to a compensation of positive and negative drift terms μ_j^k.

Next we attempt to examine the convexity adjustment of CMS rate within SMM. A natural approach is to change a forward measure Q^{k+1} to a swap measure Q^S, and the corresponding Radon-Nikodym derivative is

$$\frac{dQ^{k+1}}{dQ^S} = \frac{A(0;n,m)B(t,T_{k+1})}{A(t;n,m)B(0,T_{k+1})}. \tag{11.39}$$

Pelsser (2000), Hunt-Kennedy-Pelsser (2000) proposed a Linear Swap Model (LSM) that delivers the following approximation,

$$\frac{B(t,T_{k+1})}{A(t;n,m)} \approx \alpha + \beta_k S_{n,m}(t), \tag{11.40}$$

with

$$\alpha = \frac{1}{\sum_{j=n}^{m-1} \tau_j},$$

$$\begin{aligned}
\beta_k &= \frac{B(0,T_{k+1})A^{-1}(0;n,m) - \alpha}{S_{n,m}(0)} \\
&= \frac{B(0,T_{k+1}) - \alpha A(0;n,m)}{B(0,T_n) - B(0,T_m)}.
\end{aligned}$$

Applying this result to the above Radon-Nikodym derivative yields

$$\frac{dQ^{k+1}}{dQ^S} \approx \frac{\alpha + \beta_k S_{n,m}(t)}{\alpha + \beta_k S_{n,m}(0)}.$$

Now we are able to evaluate the expected value $\mathbf{E}^{k+1}[y(T_l)], l \leqslant n$, at fixing time T_l under the swap measure Q^S. It follows

$$\begin{aligned}
y^* &= \mathbf{E}^{k+1}[y(T_l)] = \mathbf{E}^{k+1}[S_{n,m}(T_l)] \\
&= \mathbf{E}^S\Big[\frac{\alpha + \beta_k S_{n,m}(T_l)}{\alpha + \beta_k S_{n,m}(0)} S_{n,m}(T_l)\Big] \\
&= \frac{1}{\alpha + \beta_k S_{n,m}(0)} \mathbf{E}^S\big[\alpha S_{n,m}(T_l) + \beta_k S_{n,m}^2(T_l)\big] \\
&= \frac{\alpha S_{n,m}(0) + \beta_k S_{n,m}^2(0)\exp(v_S^2 T_l)}{\alpha + \beta_k S_{n,m}(0)}.
\end{aligned} \tag{11.41}$$

The resulting convexity adjustment for CMS rate in SMM is given by

$$cv = y^* - y(0) = \frac{\beta_k S_{n,m}^2(0)(\exp(v_S^2 T_l) - 1)}{\alpha + \beta_k S_{n,m}(0)}, \tag{11.42}$$

where we can observe that the sign of the convexity adjustment cv is determined by the sign of β_k since $\alpha + \beta_k S_{n,m}(0)$ is usually positive. Recall that β_k is defined by

$$\frac{B(0,T_{k+1}) - \alpha A(0;n,m)}{B(0,T_n) - B(0,T_m)}$$

and note that $\alpha A(0;n,m)$ is about equal to the discount factor up to the middle point between T_n and T_m, namely

$$\alpha A(0;n,m) \approx B(0,T^*), \quad T^* = (T_n + T_m)/2 + T_n,$$

it is easy to verify that the convexity adjustment cv in the Linear Swap Model exhibits the similar properties as in the LMM.

11.4 Incorporating Stochastic Volatility

As shown in Figure (11.2), the smile/skew effect of implied volatilities of caps and swaptions also displays a similar pattern as observed in stock option markets. It is evident for stock options that the leptokurtic feature of the distribution of stock returns mainly generates smile and skew, and as discussed in Chapter 3, stochastic volatility models such as in Heston (1993), or in Schöbel and Zhu (1999) can explain smile/skew effect successfully, and have became the main models for smile modeling. Accompanied by the success story for stock options, it is a natural step to extend LMM by adopting stochastic volatilities to capture volatility smile/skew observed in cap and swaption market.

Various extended LMMs with smile modeling are suggested and can mainly fall into two categories: The first is the local volatility approach including constant elasticity of variance (CEV) model (Andersen-Andreasen, 2000) and displaced diffusion model. Both CEV model and displaced diffusion model can generate a monotone down-slope skew of implied volatilities, but fail to create symmetric smile. The second approach is to describe volatility with a stochastic process. Most models so far with this approach use a mean-reverting square root process á la Heston for variance to specify a part of total variance of a Libor, see Andersen and Brotherton-Ratcliffe (ABR, 2001), Wu and Zhang (2002), Piterbarg (2003), Zhu (2007), as well as Belomestny, Mathew and Schoenmaker (BMS, 2007). The models of ABR and Piterbarg do not allow for a correlation between Libors and theirs stochastic volatilities, and therefore lack the capability to create skew for implied cap volatilities without adding additional model components. The zero correlation remarkably restricts the ability of the proposed stochastic volatility LMMs to capture different individual smile/skew patterns of caps and swaptions.

Hagan, Kumar, Lesniewski and Woodward (HKLW, 2002) applied a so-called SABR model to Libor process where the volatility is governed by a geometric Brownian motion which is correlated with Libor diffusion. Since SABR model admits an explicit expansion solution for implied volatilities of caplets and swaptions, it is very popular for brokers and traders to smooth the market implied volatilities. However, the stochastic volatility in SABR model displays no mean-reversion on the contrary to the well-documented empirical feature of a volatility process. Driven by a geometric Brownian motion, the stochastic volatility in SABR model is not guaranteed to be stationary in a long way. HKLW did not provide a comprehensive treatment to the application of SABR to LMM, and, for example, it is not clear which processes should Libors and stochastic volatilities follow under an unique forward measure. There is no link between cap pricing and swaption pricing in an usual SABR model. In practice, SABR model as a stochastic volatility model is often used to interpolate and extrapolate market volatilities. Recently, Mercurio and Morini (2007), Rebonato (2007), as well as Hagan and Lesniewski (2008) attempt to fill some theoretical gaps of an isolated SABR model and build up an extended LMM with SABR stochastic volatility allowing for non-zero correlation between the dynamics of Libors and SARR volatilities. However, this class of models do not involve the Fourier transform, and then are beyond the theoretical framework discussed in this book.

All of the mentioned problems in the financial literature on LMM indicate the open challenges in modeling smile effect by incorporating stochastic volatility into LMM. Firstly, every Libor and its volatility in LMM live under its own forward measure. If stochastic volatility is correlated with its Libor, its dynamics must also be changed with the change of forward measure and the associated numeraire. The volatility correlation is a desired feature but makes modeling much more complicated. Secondly, if each Libor has its own stochastic volatility process such as in the Heston model, we have at least 5 parameters more for each Libor. In a practical application to 20 Libors, we need 100 parameters only for expressing stochastic volatilities. No calibration procedure can deal with such a large number of parameters efficiently, and no trader and risk manager will feel comfort with so many parameters. Thirdly, even when we can build an extended LMM with stochastic volatilities, it is not straightforward to derive a reliable analytic or semi-analytic pricing formula for swaption. Without a closed-form solution for swaption on hand, it is impossible to perform a quick and efficient calibration to market swaption volatilities, and therefore also impossible to price exotic interest rate derivatives. As pointed out by Brigo and Mercurio (2006), analytical tractability is an important issue regarding the stochastic volatility model. In many existing models such as Andersen and Brotherton-Ratcliffe (2001), Piterbarg (2003), the cost for arriving at analytical or approximated analytical pricing formula for cap and swaption is to set a zero correlation between Libor and its volatility. Different techniques such as Taylor expansion, moment-matching, Fourier transform are intensively applied to obtain a tractable pricing formula.

In this section, we briefly address five extended LMMs with stochastic volatility component, they are Andersen and Brotherton-Ratcliffe model (2001), Piterbarg model (2003), Wu and Zhang model (2006), Zhu model (2007) as well as BMS model (2007). As shown below, the first two models do not consequently apply CF to search for option pricing formula, although a mean-reverting square root process is used to specify the dynamics of Libor volatility. Wu and Zhang model, Zhu model and BMS model not only incorporate stochastic volatility, but also admit arbitrary correlation between Libor and its volatility. All three models utilize CFs as a powerful and convenient tool to express the pricing formulas both for caplets and for swaptions.

11.4.1 The Andersen and Brotherton-Ratcliffe Model

Andersen and Brotherton-Ratcliffe (ABR, 2001) proposed the following dynamics for Libors $L_j(t)$ coupled with a stochastic volatility $V(t)$ under the forward measure Q^{j+1},

$$\begin{aligned} dL_j(t) &= q_j(t)g_1(L_j(t))\sqrt{V(t)}dW_j(t), \\ dV(t) &= \kappa(\theta - V(t))dt + \sigma g_2(V(t))dW(t), \end{aligned} \tag{11.43}$$

where the Wiener process $W_j(t), 0 < j < M+1$, are mutually correlated, $W_i(t)W_j(t) = \rho_{ij}t$, but not correlated with the Wiener process $W(t)$, $W_j(t)W(t) = 0$. The zero-correlation between Libor and volatility does not admit generating volatility skew, and therefore, would be a restriction of the model setup. To overcome this drawback, ABR introduced a function $g_1(x)$ to capture skew, and $g_1(x)$ may be called skew function. For example, a functional form $g_1(x) = x^a, 0 < a < 1$, maps the Libor process to a constant elasticity of variance (CEV) process that allows for a closed-form option pricing formula. An other alternative for the functional form $g_1(x)$ is the displaced diffusion, namely $g_1(x) = 1 - ax$, that is also known to generate volatility skew. The deterministic function $q_j(t)$ is used to capture the term structure of ATM caplet volatility.

The variance process $V(t)$ in the ABR model is also rather generally formulated with the diffusion function $g_2(L_j(t))$, and provides the diffusions of all Libors with the same dynamics. Obviously, setting $g_2(x) = x^{1/2}$ yields a mean-reverting square root process, and all known features of the Heston model discussed in Chapter 3 may be applied here. However, supported by their empirical work, ABR suggested to choose a function $g_2(x) = x^{3/4}$. The resulting variance process $V(t)$ is still stationary and the zero-value of variance is not accessible. Due to the volatility function $q_j(t)$ that controls the level of ATM volatility, the mean value θ can be scaled to be 1, and we have one parameter less to estimate. Generally, we can not immediately derive the analytical pricing formula for caplets if q_j, g_1 and g_2 are arbitrary functions.

Under the spot measure Q_1, the Libor processes read

$$\begin{aligned} dL_j(t) &= q_j(t)g_1(L_j(t))\sqrt{V(t)}\left[\mu_j dt + dW_j(t)\right], \\ dV(t) &= \kappa(\theta - V(t))dt + \sigma g_2(V(t))dW(t), \end{aligned} \tag{11.44}$$

with

$$\mu_j = \sqrt{V(t)} \sum_{h=\beta(t)}^{j} \frac{\tau_h \rho_{hj} q_h(t) g_1(L_h(t)) g_1(L_j(t))}{1 + \tau_h L_h(t)}.$$

Because the variance $V(t)$ is uncorrelated with $L_j(t)$, the change of measure does not affect the process of $V(t)$. As shown below, this zero-correlation allows ABR to treat the deterministic part and the stochastic part in the diffusion term separately. Instead of applying Fourier transform, ABR used an asymptotic expansion to obtain an efficient approximated pricing formula for caplets and swaptions.

Firstly, Andersen and Brotherton-Ratcliffe considered only the deterministic part in their model by assuming that $V(t)$ is constant and equal to 1. Therefore, the ABR model becomes a local volatility model that may be treated by an asymptotic expansion technique. Andersen and Brotherton-Ratcliffe finally arrived at the following pre-solution for caplet,

$$CPL_j(0, V = 1) = \tau_j B(0, T_{j+1})[L_j(0)N(d_1) - KN(d_2)] \tag{11.45}$$

with

$$
\begin{aligned}
d_{1,2} &= \frac{\ln(L_j(0)/K) \pm \frac{1}{2}C_j^2(L_j(0),x)}{C_j(L_j(0),x)} \\
x &= \frac{1}{T_j}\int_0^{T_j} q_j^2(t)dt, \\
C_j(L_j(0),x) &= \sum_{i=0}^{1}(xT_j)^{i+1/2}\Omega_i(L_j(0)) + \mathscr{O}(T_j^{5/2}), \\
\Omega_0(L_j(0)) &= \frac{\ln(L_j(0)/K)}{\int_K^{L_j(0)} g_1^{-1}(u)du}, \\
\Omega_1(L_j(0)) &= -\frac{\Omega_0(L_j(0))}{(\int_K^{L_j(0)} g_1^{-1}(u)du)^2}\ln\Big[\Omega_0(L_j(0))\big(\frac{L_j(0)K}{g_1(L_j(0))g_1(K)}\big)\Big].
\end{aligned}
$$

In spite of the approximation error of $\mathscr{O}(T_j^{5/2})$, ABR claimed that the above expansion formula for caplet values is often accurate enough for long maturities.

Based on the pre-solution in (11.45), ABR incorporated the stochastic volatility given in (11.44). Note that stochastic variance $V(t)$ is independent of the Libor $L_j(t)$, as indicated in Section 3.1, the integrated variance up to time T_j can be directly taken into account in a Black-Scholes-style pricing formula. In this case, the pricing formula (11.45) could be rewritten as

$$
\begin{aligned}
CPL_j(0,V) &= \tau_j B(0,T_{j+1})\mathbf{E}^{j+1}\big[[L_j(0)N(d_1) - KN(d_2)]\big] \\
&= \tau_j B(0,T_{j+1})\mathbf{E}^{j+1}\big[y(x,\cdots)\big] \qquad (11.46)
\end{aligned}
$$

with

$$
x = \frac{1}{T_j}\int_0^{T_j} q_j^2(t)V(t)dt.
$$

A fourth-order Taylor expansion leads to

$$
\begin{aligned}
CPL_j(0,V) &\approx \tau_j B(0,T_{j+1}) \\
&\quad \times \sum_{n=0}^{4}\frac{1}{n!}\frac{\partial^n y}{\partial x^n}(x^*)\mathbf{E}^{j+1}\big[(x-x^*)^n\big], \qquad (11.47)
\end{aligned}
$$

where the convention $\partial^0 y/\partial x^0 = y$ is used, and x^* is set to

$$
\begin{aligned}
x^* &= \mathbf{E}^{j+1}[x] = \frac{1}{T_j}\int_0^{T_j} q_j^2(t)\mathbf{E}^{j+1}[V(t)]dt \\
&= \frac{1}{T_j}\int_0^{T_j} q_j^2(t)[\theta + (V(t)-\theta)e^{-\kappa t}]dt. \qquad (11.48)
\end{aligned}
$$

The unknown quantities in the above Taylor expansion are $\mathbf{E}^{j+1}\big[(x-x^*)^n\big]$. To evaluate these moments of x, we apply the Fourier transform $\mathbf{E}^{j+1}[e^{-ix}]$ and then compute them according to the method given by (2.31) in Chapter 2. Thanks to the previous approximations, Andersen and Brotherton-Ratcliffe derived also an estimate for the implied volatility of a caplet on L_j, which is not given here.

The next step to do is swaption pricing. To this end, Andersen and Brotherton-Ratcliffe attempted to establish a similar process for swap rate to the Libor process given in (11.44), and perform the following approximation,

$$\begin{aligned} dS_{n,m}(t) &= \sum_{j=n}^{m-1} \frac{\partial S_{n,m}(t)}{\partial L_j(t)} g_1(L_j) q_j(t) \sqrt{V(t)} dW_j(t) \\ &\approx g_1(S_{n,m}(t)) \sum_{j=n}^{m-1} \omega_j(t) q_j(t) \sqrt{V(t)} dW_j(t), \end{aligned} \tag{11.49}$$

with

$$\omega_j(t) = \frac{\partial S_{n,m}(t) g_1(L_j(t))}{\partial L_j(t) g_1(S_{n,m}(t))}.$$

Expressing the diffusion term of the above swap rate process in a compact manner yields

$$dS_{n,m}(t) = g_1(S_{n,m}(t)) q_{n,m}(t) \sqrt{V(t)} dW^s(t), \tag{11.50}$$

where we have

$$q_{n,m}(t) dW^s(t) = \sum_{j=n}^{m-1} \omega_j(t) q_j(t) dW_j(t).$$

The process for the swap rate $dS_{n,m}(t)$ in (11.50) then takes the similar form as the Libor process in (11.44). As a result, all pricing formulas and techniques for caplets can be applied for swaptions with some appropriate replacements of $L_j(t)$ with $S_{n,m}(t)$, q_1 with $q_{n,m}$, and T_j with T_n, respectively.

11.4.2 The Piterbarg Model

Piterbarg (2003) proposed a local-stochastic LMM that is structurally similar to the ABR model. In details, the dynamics of Libors $L_j(t)$ and the related stochastic volatility process under the forward measure Q^{j+1} are given by

$$\begin{aligned} dL_j(t) &= [\beta_j(t) L_j(t) + (1-\beta_j(t)) L_j(0)] \sqrt{V(t)} \sum_{k=1}^{K} q_k(t) dW_k(t), \\ dV(t) &= \kappa(V_0 - V(t)) dt + \sigma \sqrt{V(t)} dW(t), \end{aligned} \tag{11.51}$$

with $dW_k dW = 0, 0 < k \leqslant K$. To compare the Piterbarg model with the ABR model, we need to set

$$g_1(L_j(t)) = [\beta_j(t) L_j(t) + (1-\beta_j(t)) L_j(0)]$$

and

$$g2(V(t)) = \sqrt{V(t)}.$$

However, the extended LMM of Piterbarg is a model with K factors that control the dynamics of all M Libors. By setting $K = M$, we may arrive at a full factor model for Libors, and with the help of some specific $q_k(t)$, may further reconcile the setups of Piterbarg and ABR. Particularly here, the stochastic variance $V(t)$ follows a special mean-reverting square root process in which the mean level θ is set to be spot variance V_0. As in the ABR model, $V(t)$ is uncorrelated with $L_j(t)$ and provides the identical variance dynamics for all Libors. The skews in implied caplet volatilities are generated here by a displaced diffusion, namely by the term $\beta_j(t)L_j(t) + (1-\beta_j(t))L_j(0)$ where $\beta_j(t)$ is a deterministic function of time and controls time-dependent skews.

Based on the model setup for Libors given in (11.51), we can follow the same procedure as described in the ABR model to approximate the process for swap rate $S_{n,m}(t)$ under the swap measure. It follows by some calculations,

$$\begin{aligned} dS_{n,m}(t) &= [\beta_{n,m}(t)S_{n,m}(t) + (1-\beta_{n,m}(t))S_{n,m}(0)]\sqrt{V(t)}\sum_{k=1}^{K} q_k(t;n,m)dW_k(t) \\ &= [\beta_{n,m}(t)S_{n,m}(t) + (1-\beta_{n,m}(t))S_{n,m}(0)]\sqrt{V(t)}q_{n,m}(t)dU(t) \end{aligned} \tag{11.52}$$

where

$$\begin{aligned} q_{n,m}^2(t) &= \sum_{k=1}^{K} q_k^2(t;n,m), \\ q_k(t;n,m) &= \sum_{i=n}^{m-1} \omega_i q_k(t), \\ \beta_{n,m}(t) &= \sum_{i=n}^{m-1} \gamma_i \beta_i(t), \\ \omega_i &= \frac{\partial S_{n,m}(0)L_j(0)}{\partial L_j(0)S_{n,m}(0)}, \\ \gamma_i &= \frac{\sum_{k=1}^{K} q_k(i)q_k(t;n,m)}{\sum_{k=1}^{K} q_k^2(t;n,m)}. \end{aligned}$$

The approximated processes for swap rates keep the same structure as the ones for Libors. Hence in the following, we consider only the valuation of swaptions and the calibration of this model to swaption market data. Piterbarg's contribution is that the swap rate process in (11.52) may be further approximated by a simple process via Markovian projection. The technique of Markovian projection has been briefly discussed in Chapter 4. The projected process has usually a simpler form with constant parameters, and still admits a consistent distribution as the original one, here, the process in (11.52), in the sense of first and second moments. Piterbarg proposed the following simple process for the Markovian projection,

$$dy(t) = \lambda_{n,m}[b_{n,m}y(t) + (1-b_{n,m})y(0)]\sqrt{V(t)}dU(t). \tag{11.53}$$

for each index pair (n,m). The parameters $\lambda_{n,m}$, $b_{n,m}$ and σ are regarded as effective fitting parameters in which the market implied distribution of $y(t)$ is encoded. Piterbarg provided the following estimates for effective skew $b_{n,m}$ and volatility $\lambda_{n,m}$.

Effective skew $b_{n,m}$: The skew is a weighted average of the original skew function $\beta_{n,m}(t)$.

$$\begin{aligned} b_{n,m} &= \int_0^{T_n} \beta_{n,m}(t)w(t)dt, \qquad (11.54)\\ w(t) &= \frac{D(t)|q_{n,m}(t)|^2}{\int_0^{T_n} D(t)|q_{n,m}(t)|^2 dt}, \\ D(t) &= V_0^2 \int_0^t |q_{n,m}(s)|^2 ds + V_0\sigma^2 e^{-\kappa t}\int_0^t |q_{n,m}(s)|^2 \frac{e^{\kappa s}-e^{-\kappa s}}{2\kappa} ds. \end{aligned}$$

Effective volatility $\lambda_{n,m}$: The effective volatility λ is found by fitting the ATM prices implied by the original process and the projected process in the following sense,

$$\mathbf{E}\Big[g\Big(\int_0^{T_n} |q_{n,m}(t)|^2 V(t)dt\Big)\Big] = \mathbf{E}\Big[g\Big(\int_0^{T_n} \lambda_{n,m}^2 V(t)dt\Big)\Big]$$

with

$$g(x) = \frac{S_{n,m}(0)}{b_{n,m}}(2N(b_{n,m}\sqrt{x}/2)-1),$$

where $N(\cdot)$ denotes the cumulative Gaussian distribution function. However, it is difficult to calculate the above two expected values directly. Piterbarg's approach to overcoming this problem is to approximate $g(x)$ with a functional form $g(x) \approx a_0 + a_1 e^{-a_2 x}$ where the parameters $\{a_0, a_1, a_2\}$ are determined by fitting the local second-order Taylor expansion around the mean of $\int_0^{T_n} |q_{n,m}(t)|^2$. This leads finally to the following equation for finding effective $\lambda_{n,m}$,

$$\mathbf{E}\Big[\exp\big(-a_2\int_0^{T_n} |q_{n,m}(t)|^2 V(t)dt\big)\Big] = \mathbf{E}\Big[\exp\big(-a_2\int_0^{T_n} \lambda_{n,m}^2 V(t)dt\big)\Big], \tag{11.55}$$

with

$$a_2 = -\frac{g''(x_0)}{g'(x_0)}, \quad x_0 = V_0\int_0^{T_n} |q_{n,m}(t)|^2 dt.$$

Calculating the first expected value

$$\mathbf{E}\Big[\exp\big(-a_2\int_0^{T_n} |q_{n,m}(t)|^2 V(t)dt\big)\Big]$$

is equivalent to calculating the following expected value

$$m_1(u) = \mathbf{E}\big[e^{-u\Lambda(T_n)}\big], \quad \Lambda(T_n) = \int_0^{T_n} |q_{n,m}(t)|^2 V(t)dt.$$

Generally $m_1(u)$ can be computed numerically. For the second expectation

$$\mathbf{E}\Big[\exp\big(-a_2\int_0^{T_n}\lambda_{n,m}^2 V(t)dt\big)\Big],$$

note $\lambda_{n,m}$ is constant, and define

$$m_2(u) = \mathbf{E}\big[\exp\big(-u\int_0^{T_n} V(t)dt\big)\big],$$

which admits an analytical solution and can be calculated with the formula (3.20). Therefore we can calculate the second expected value analytically. Finding the effective volatility $\lambda_{n,m}$ is then equivalent to solving

$$m_1\big(-\frac{g''(x_0)}{g'(x_0)}\big) = m_2\big(-\frac{g''(x_0)}{g'(x_0)}\lambda_{n,m}^2\big). \tag{11.56}$$

After the effective skew $b_{n,m}$ and the effective volatility $\lambda_{n,m}$ have been found, the relationship between the model parameters $\{\beta_{n,m}(t), q_{n,m}(t)\}$ and the effective market parameters $\{b_{n,m}, \lambda_{n,m}\}$ is established. The calibration to market swaption prices is then equivalent to solving the following two minimization problems:

$$\min\sum_{n,m}(b_{n,m} - b_{n,m}^{Market})^2$$

and

$$\min\sum_{n,m}(\lambda_{n,m} - \lambda_{n,m}^{Market})^2.$$

where the market implied parameters $\{b_{n,m}^{Market}, \lambda_{n,m}^{Market}, \sigma_{n,m}\}$ including the global parameter κ in the stochastic volatility process $V(t)$ are retrieved in an other calibration procedure where the projected model in (11.53) is matched to the actual market swaption prices. The effective parameters $\{b_{n,m}, \lambda_{n,m}\}$ sever then as the intermediates to search for the original model parameters $\{\beta_{n,m}(t), q_{n,m}(t)\}$ via the market implied ones $\{b_{n,m}^{Market}, \lambda_{n,m}^{Market}\}$. The whole calibration in the Piterbarg model includes two sub-calibrations. The advantages of Piterbarg's Markovian projection is to reformulate the calibration problem with the projected processes that admits possibly an easy calibration of market parameters or a simple derivation of the pricing formulas, instead of performing the calibration with the original processes directly.

11.4.3 The Wu and Zhang Model

As shown above, both the ABR model and the Piterbarg model incorporate the Heston's stochastic variance dynamics into to LMM, but are restricted to zero-correlation case. The nice feature of a correlated stochastic volatility model to generate down-sloping skews does not show to advantage in their models. The volatility skews in both models are generated via various local volatility formulations. Additionally, Fourier transform and CF do not play a key role in expressing the final option pricing formulas for caplets and swaptions. Wu and Zhang (2006) proposed an extended LMM with stochastic volatility, which fills these two gaps and allows for (1) a non-zero correlation between the dynamics of Libors and volatility, and (2) applies extensively the technique of Fourier transform to derive pricing formulas.

The stochastic processes for Libors $L_j(t), 1 \leqslant j \leqslant M$, and the corresponding variance process $V(t)$ in Wu and Zhang's model under the risk-neutral measure Q take the following forms,

$$\begin{aligned} \frac{dL_j(t)}{L_j(t)} &= \sqrt{V(t)}q_j(t) \cdot [\eta_j(t)\sqrt{V(t)}dt + dW(t)], \\ dV(t) &= \kappa(\theta - V(t))dt + \sigma\sqrt{(V(t)}dZ(t), \end{aligned} \tag{11.57}$$

where $W(t)$ is a multi-dimensional independent Brownian motion. $q_j(t)$ is a deterministic vector function with the same size as $W(t)$. The stochastic variance $V(t)$ follows a mean-reverting square root process where $Z(t)$ is a one-dimensional Brownian motion that is correlated with $W(t)$ in the following way

$$\frac{< q_j(t), dW(t) > dZ(t)}{|q_j(t)|} = \rho_j(t)dt. \tag{11.58}$$

It is clear that the correlation coefficient $\rho_j(t)$ is time-dependent, precisely, dependent on the deterministic volatility function $q_j(t)$. Hence, this correlation can not be given explicitly, or at least, is difficult to be used to control volatility skews when we note that $q_j(t)$ is initially used to control volatility term structure. $\eta_j(t)$ is also a deterministic vector function and has a special form

$$\eta_{j-1}(t) - \eta_j(t) = \frac{\tau_j L_j(t)}{1 + \tau_j L_j(t)} q_j(t).$$

It can be shown easily that such special construction of $\eta_j(t)$ eases the change of measure. Particularly, with the help of $\eta_j(t)$, the Libor processes and variance process under the respective forward measure Q^{j+1} are given in a simpler form,

$$\begin{aligned} \frac{dL_j(t)}{L_j(t)} &= \sqrt{V(t)}q_j(t) \cdot dW(t), \\ dV(t) &= [\kappa\theta - (\kappa + \sigma\xi_j(t))V(t)]dt + \sigma\sqrt{(V(t)}dZ(t), \end{aligned} \tag{11.59}$$

with

$$\begin{aligned}\xi_j(t) &= \sum_{h=1}^{j} \frac{\tau_j L_j(t)}{1+\tau_j L_j(t)} |q_j(t)| \rho_h(t) \\ &= \sum_{h=1}^{j} \frac{\tau_j L_j(t)}{1+\tau_j L_j(t)} < q_j(t), W(t) > Z(t). \end{aligned} \tag{11.60}$$

It is easy to verify that the disappearance of the term $\eta_j(t)$ in the process $L_j(t)$ and the appearance of the new term $\xi_j(t)$ in the process $V(t)$ are the results of the change of measure from Q to Q^{j+1}. As $\xi_j(t)$ is a time-dependent function, the usual methods for deriving CFs given in Section 3.2 are difficult to use. Wu and Zhang (2006) then used the freezing technique and approximated the $\xi_j(t)$ by freezing the forward Libors at the origin,

$$\xi_j(t) \approx \sum_{h=1}^{j} \frac{\tau_j L_j(0)}{1+\tau_j L_j(0)} |q_j(t)| \rho_h(t).$$

It follows immediately the reformulation for $dV(t)$,

$$dV(t) = \kappa[\theta - \xi^*(t)V(t)]dt + \sigma\sqrt{(V(t)}dZ(t)$$

with

$$\xi_j^*(t) = 1 + \frac{\sigma}{\kappa}\xi_j(t).$$

In spite of the freezing of $L_h(t)$ to their values at the origin, the time-dependent nature of the Wu and Zhang model does not change due to the terms $q_j(t)$. Denote $x(t) = \ln(L_j(t)/L_j(0))$, According to the Feynman-Kac theorem, the moment generating function $f(\phi,t) = \mathbf{E}[e^{\phi x(T_j)}|\mathscr{F}_t]$ of $x(t)$ satisfies the Kolmogorov's backward equation that we have used once in the Heston model,

$$\begin{aligned}&\frac{\partial f}{\partial t} + \kappa[\theta - \xi^*(t)V]\frac{\partial f}{\partial V} - \frac{1}{2}|q_j(t)|^2 V\Big(\frac{\partial f}{\partial x} - \frac{\partial^2 f}{\partial x^2}\Big) \\ &\quad + \frac{1}{2}\sigma^2 V \frac{\partial^2 f}{\partial V^2} + \sigma\rho_j(t)V|q_j(t)|\frac{\partial^2 f}{\partial x \partial V} = 0, \end{aligned} \tag{11.61}$$

subject to the boundary condition

$$f(\phi, t = 0) = e^{\phi x_0}.$$

This PDE can be solved explicitly as in the Heston model if it is assumed that the functions $q_j(t)$ and $\xi_j(t)$ are piecewise constant on the grid $\{t_0 = 0, t_1, t_2, \cdots, t_j = T_j\}$. The solution takes the following form

$$f(\phi, T_j, V_0, x_0) = \exp(H_1(T_j) + H_2(T_j)V_0 + \phi x_0), \tag{11.62}$$

where $H_1(T_j)$ and $H_2(T_j)$ are expressed by recursive formulas,

$$H_1(t_{h+1}) = H_1(t_h) + \frac{\kappa\theta}{\sigma^2}\left[(a_h + d_h)\Delta t_h - 2\ln\left(\frac{1-c_h e^{d_h \Delta t_h}}{1-c_h}\right)\right],$$

$$H_2(t_{h+1}) = H_2(t_h) + \frac{[a_h + d_h - \sigma^2 H_2(t_h)](1-e^{d_h \Delta t_h})}{\sigma^2(1-c_h e^{d_h \Delta t_h})}$$

with the following parameters

$$\begin{aligned} a_h &= \kappa\xi_j^*(t_h) - \sigma\rho_j(t_h)q_j(t_h)\phi, \\ d_h &= \sqrt{a_h^2 - \sigma^2 q_j^2(t_h)(\phi^2 - \phi)}, \\ c_h &= \frac{a_h + d_h - \sigma^2 H_2(t_h)}{a_h - d_h - \sigma^2 H_2(t_h)}. \end{aligned} \tag{11.63}$$

Having the closed-form solution for the moment generating function $f(\phi)$, we can construct two CFs under the Delta measure and the forward measure Q^{j+1}, under which two exercise probabilities are defined,

$$f_1(\phi) = f(-i+\phi) \quad \text{and} \quad f_2(\phi) = f(i\phi). \tag{11.64}$$

Zhang and Wu expressed the pricing formula for caplet in terms of the moment generating function $f(\phi)$, which reads

$$CPL_j(0) = \tau_j B(0, T_{j+1})[L_j(0)F_1 - KF_2], \tag{11.65}$$

where F_1 and F_2 are given by

$$F_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}[e^{i\phi \ln(L_j(0)/K)}]f(1+i\phi)}{\phi}d\phi$$

and

$$F_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}[e^{i\phi \ln(L_j(0)/K)}]f(i\phi)}{\phi}d\phi.$$

It can be verified easily that these two probabilities $F_k, k = 1,2$, may be rewritten in the usual forms throughout this book in terms of CFs, These are

$$F_k = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(f_k(\phi)\frac{\exp(i\phi\ln(L_j(0)/K))}{i\phi}\Big)d\phi, \quad k = 1,2.$$

To value swaptions, Wu and Zhang used the same approach as in the models of Andersen and Brotherton-Ratcliffe (2002) and Piterbarg (2003), and approximated $S_{n,m}(t)$ with a sum of weighted Libors, and kept the structure of the swap rate process identical to the structure of Libors. Finally, Wu and Zhang arrived at the following approximated dynamics for swap rate and the corresponding stochastic variance,

$$
\begin{aligned}
\frac{dS_{n,m}(t)}{S_{n,m}(t)} &= \sqrt{V(t)}\sum_{k=n}^{m-1}\omega_k(0)[q_k(t)\cdot dW^S(t)] \\
&= \sqrt{V(t)}q_{n,m}(t)dU^S(t), \qquad (11.66)\\
dV(t) &= \kappa(\theta-\xi_{n,m}(0)V(t))dt+\sigma\sqrt{V(t)}dZ^S(t),
\end{aligned}
$$

where dU^S and dZ^S are the Brownian motions under the swap measure, and are correlated with

$$\rho_{n,m} = \frac{\sum_{k=n}^{m-1}\omega_k(0)|q_k(t)|\rho_k}{|\sum_{k=n}^{m-1}\omega_k(0)q_k(t)|}.$$

The weights $\omega(0)$ and the drift adjustment $\xi_{n,m}(0)$ for $dV(t)$ are given by

$$\omega_k(0) = \frac{\partial S_{n,m}(0)L_k(0)}{\partial L_k(0)S_{n,m}}$$

and

$$\xi_{n,m}(0) = 1+\frac{\sigma}{\kappa S_{n,m}(0)}\sum_{k=n}^{m-1}\tau_k B(0,T_{k+1})\xi_k(0).$$

With these approximated processes for $S_{n,m}(t)$ and $V(t)$, all results including the pricing formula for caplet may be directly applied to swaptions with the corresponding replacements. Namely, we replace L_j with $S_{n,m}$, T_j with T_n, x_j with $x_{n,m}$, q_j with $q_{n,m}$, and ρ_j with $\rho_{n,m}$.

11.4.4 The Zhu Model

In this subsection we present an extended LMM of Zhu (2007), which incorporates stochastic the volatility under the market forward measure Q^{j+1} of each Libor $L_j(t)$. Unlike the above LMMs with stochastic variance, Zhu assumed that stochastic volatility $v_j(t)$ is governed by a mean-reverting Ornstein-Uhlenbeck process under the forward measure Q^{j+1}, as in Schöbel and Zhu (1999). In more details, the processes for Libors $L_j(t), 1 \leqslant j \leqslant M$, and the corresponding volatilities $v_j(t), 1 \leqslant j \leqslant M$, are given by

$$
\begin{aligned}
\frac{dL_j(t)}{L_j(t)} &= v_j(t)dW_j(t), \\
dv_j(t) &= \kappa_j(\theta_j - v_j(t))dt+\sigma_j(t)dW. \qquad (11.67)
\end{aligned}
$$

The Wiener process dW is correlated with the Wiener process dW_j of the Libor $L_j(t)$, i.e., $dW_j dW = \rho_j dt$, which means that stochastic volatility $v_j(t)$, unlike in Wu and Zhang's model (2006), has its own special correlation with $L_j(t)$. The correlation ρ_j is independent of any other deterministic and stochastic parameters and has the same financial meaning as in usual stochastic volatility models for equity derivatives. Furthermore, all stochastic volatilities $v_j(t), j \geqslant 1$, are driven by the same Wiener process $W(t)$. At first glance, this assumption seems to be a restriction. But if we look at the smile pattern of caplets in details, we can find that the smile structures of different caplets are mainly determined by the flexible volatility correlations ρ_j, and are less driven by the randomness of the volatility.

The second formulation of the Zhu model is to specify stochastic variance $V(t)$ to be a mean-reverting square-root process as in the Heston model,

$$\begin{aligned} \frac{dL_j(t)}{L_j(t)} &= \sqrt{V_j(t)}dW_j(t), \\ dV_j(t) &= \kappa_j(\theta_j - V_j(t))dt + \sigma_j(t)\sqrt{V_j(t)}dW. \end{aligned} \tag{11.68}$$

As indicated clearly by the specifications for the dynamics of Libors and the corresponding stochastic volatilities, the above specifications of the LMM have the following favorite features.

1. The Wiener processes of Libor and its stochastic volatility are individually correlated. This correlation is free and could be different for all Libors. The individual specification of the correlation between Libor and its volatility is important since it admits a fine control of smile patterns and could significantly improve the capability for a better fitting to market volatility data.
2. The Zhu model admits the different formulations for stochastic volatility. Although only mean-reverting Ornstein-Uhlenbeck process and mean-reverting square root process are proposed to model stochastic volatility and variance respectively. other interesting specifications discussed early in this book may be incorporated.
3. The pricing formula via the Fourier transform for caplets can be given in a straightforward way and takes the similar form to the well-known pricing solution for equity options without any approximation or the freezing technique.
4. A main contribution of the Zhu model is to derive the CFs of swaptions within an uncorrelated (orthogonal) framework where Libors are represented by some independent factors, namely the principal components. Due to the independence of factors, the CF of a swap rate can be simply approximated to be a product of the CFs of each factors. The pricing formula for swaption in the Zhu model coincides with the formulas in the existing stochastic models, and can be implemented in a very efficient way.
5. Finally, by introducing some deterministic nesting functions for the parameters of stochastic volatility process, the model becomes very parsimonious. As shown later, the same type of parameter of all stochastic volatility processes is nested by a deterministic function. Hence, this model is called a nested stochastic volatility LMM. To illustrate the idea of nesting, we consider all mean level parameters θ_j in the stochastic volatility processes. Obviously, in a pricing environment with 20 Libors there are 20 mean level parameters. A nesting function for θ_j is to parameterize these 20 parameters with a few parameters. As a parsimonious model, the calibration performs also efficiently and quickly.

To keep the number of parameters of stochastic volatility processes in a manageable range, Zhu introduced some deterministic functions to connect κ_j, θ_j, $v_j(0)$ and σ_j respectively. In particular, Zhu assumed the following term structure for κ_j, θ_j, $v_j(0)$ and σ_j,

$$\begin{aligned}
\kappa_j &= y_\kappa(T_j;\kappa,a) := \kappa e^{aT_j},\\
\theta_j &= y_\theta(T_j;\theta,b_1,b_2,b_3) := \theta + (b_1 T_j + b_2)e^{b_3 T_j},\\
v_j(0) &= y_v(T_j;v_0,c_1,c_2,c_3) := v + (c_1 T_j + c_2)e^{c_3 T_j},\\
\sigma_j &= y_\sigma(T_j;\sigma,d) := \sigma e^{dT_j},
\end{aligned} \tag{11.69}$$

where he attempted to capture different forms of the caplet volatility structure by specifying the mean levels θ_j and the spot volatilities $v_j(0)$ with a popular ABCD parametric function for the term volatility of caplets (Brigo and Mercurio (2006)). With these specifications we only need four functions with 11 parameters $\{\kappa,a,\theta,b_1,b_2,b_3,c_1,c_2,c_3,\sigma,d\}$ for κ_j, θ_j, $v_j(0)$ and σ_j, plus some parameters for the correlation ρ_j between $L_j(t)$ and $v_j(t)$, to specify the stochastic volatility processes for all Libors completely. The parameters of the stochastic processes for Libors are nested by certain deterministic functions that allow for a parsimonious but still reasonable formulation of stochastic volatility. Additionally, the correlation coefficients ρ_j may also be nested by a deterministic function. With a nesting function for ρ_j, the total number of parameters for this stochastic volatility LMM can be again reduced remarkably. Zhu suggested a nesting function for ρ_j with the following function

$$\rho_j = y_\rho(T_j;e_1,e_2) = e_1(1-e^{e_2 T_j}), \qquad -1 < e_1 < 1. \tag{11.70}$$

If e_1 lies in interval $(-1,0)$ and e_2 is slightly negative, the short-term Libors have a near zero correlation with their volatilities while the correlation between the long-term Libors with their volatilities will converge to e_1. With the given model setup and the nesting functions, the extended LMM embraces a well-known ABCD parametric LMM as special case. Setting

$$v_j(0) = \theta_j, \qquad \kappa \to 0, \qquad \sigma \to 0,$$

leads a degenerated volatility process $v_j(t) = v_j(0)$, and the Libor process converges to the LMM with a humped parametric volatility structure which is discussed extensively in Brigo and Mercurio (2006).

It is worth comparing Zhu's model with Wu and Zhang's model in terms of the relative volatility structure. As shown previously in Wu and Zhang's model, the relative volatility level of two Libors is given by

$$a(t) = \frac{q_i(t)\sqrt{V(t)}}{q_j(t)\sqrt{V(t)}} = \frac{q_i(t)}{q_j(t)} \tag{11.71}$$

with $q_j(t)$ as a deterministic function describing the term structure of volatility. Obviously the relative volatility level $a(t)$ is also a deterministic function, and therefore displays a fixed pattern which is certainly an unfavorable feature. In contrast, the relative volatility level in the extended LMM is defined by

$$a(t) = \frac{\sqrt{V_i(t)}}{\sqrt{V_j(t)}}, \tag{11.72}$$

which is a stochastic process. From the point of view of modeling, the specification of the stochastic volatility in the Zhu model is more flexible in describing the relative dynamics of the volatilities of two Libors.

In the following, we only take the mean-reverting Ornstain-Ulenbeck process to illustrate Zhu's model. The dynamics of Libor $L_j(t)$ and its volatility $v_j(t)$ given in (11.67) are associated only with the forward measure Q^{j+1}. For the purposes of simulation and pricing exotic interest rate products, we have to change the processes of $L_j(t)$ and $v_j(t)$ under an unique measure such as the spot measure or the terminal measure Q^{M+1}. Generally under an arbitrary forward measure Q^{k+1}, the stochastic processes for Libors and volatilities take the following forms

$$\begin{aligned} dL_j(t) &= L_j(t)[\mu_j(t) + v_j(t)dW_j], \\ dv_j(t) &= [\kappa_j\theta_j - \kappa_j v_i + \xi_j(t)]dt + \sigma_j dW, \end{aligned} \tag{11.73}$$

with

$$\begin{cases} \mu_j(t) = v_j(t)\sum_{h=k+1}^{j} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}, & j > k, \\ \xi_j(t) = \sigma_j \sum_{h=k+1}^{j} \frac{v_h(t)\rho_h\tau_h L_h(t)}{1+\tau_h L_h(t)}, & j > k, \end{cases}$$

$$\begin{cases} \mu_j(t) = 0, & j = k, \\ \xi_j(t) = 0, & j = k, \end{cases}$$

$$\begin{cases} \mu_j(t) = -v_j(t)\sum_{h=j+1}^{k} \frac{v_h(t)\rho_{jh}\tau_h L_h(t)}{1+\tau_h L_h(t)}, & j < k, \\ \xi_j(t) = -\sigma_j \sum_{h=j+1}^{k} \frac{v_h(t)\rho_h\tau_h L_h(t)}{1+\tau_h L_h(t)}, & j < k. \end{cases}$$

The proof follows the usual ways as used extensively in other LMMs. The valuation of caplets here is straightforward, and Zhu's extended LMM admits a closed-form pricing formula for caplet and floorlet via the Fourier transform under each forward measure as in the Wu and Zhang model, which reads

$$CPL_j(0) = \tau_j B(0, T_{j+1})[L_j(0)\cdot F_1 - K\cdot F_2] \tag{11.74}$$

with

$$F_k = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(f_k(\phi))\frac{\exp(-i\phi\ln K)}{i\phi}\Big)d\phi, \quad k = 1, 2.$$

The functions $f_k(\phi), k = 1, 2$, is the corresponding CFs under two martingale measures dQ_*^{j+1} and Q^{i+1}, which are related via the Radon-Nikodym derivative

$$\frac{dQ_*^{j+1}}{dQ^{j+1}} = \frac{L_j(t)}{L_j(0)},$$

where $\frac{L_j(t)}{L_j(0)}$ defines an exponential martingale according to the process of $L_j(t)$. For $v_j(t)$ as a mean-reverting Ornstein-Uhlenbeck process, the exact forms of $f_k(\phi)$ may be expressed by

$$f_k(\phi) = e^{i\phi \ln L_j(0)} f_k^{SZ}(\phi), \tag{11.75}$$

where $f_k^{SZ}(\phi), k = 1, 2$, are given in (3.32) and (3.36). Similarly, for $V_j(t)$ as a mean-reverting square root process, the exact forms of $f_k(\phi)$ may be written as by

$$f_k(\phi) = e^{i\phi \ln L_j(0)} f_k^{Heston}(\phi), \tag{11.76}$$

where $f_k^{Heston}(\phi), k = 1, 2$, are given in (3.16).

To prepare for the valuation of swaptions, we need a reformulation of LMM via the principle component analysis (PCA). Note that the correlation matrix $\Sigma = \{\rho_{ij}\}$ given in (11.22) is positive semi-definite. It is known that a positive semi-definite matrix has the following decomposition

$$\Sigma = U \Lambda U^{\top}, \tag{11.77}$$

where U is an eigenvector matrix with the eigenvectors of Σ as column vector. Λ is the eigenvalue matrix with the eigenvalues of Σ as diagonal element,

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_M \end{pmatrix}.$$

Assume that $\{\lambda_j\}_{1 \leqslant j \leqslant M}$ is a non-increasing sequence. The first few dominate components are often called principal components. By using the principal components, we can rewrite the Brownian motion W_i with D independent standard Brownian motions Z_j,

$$W_j \approx \sum_{l=1}^{D} \sqrt{\lambda_l} u_{jl} Z_l, \quad 1 \leqslant j \leqslant M, \quad 1 \leqslant l \leqslant D. \tag{11.78}$$

Setting W_j into the Libor process yields

$$dL_j(t) = L_j(t)\big[\mu_j(t)dt + v_j(t) \sum_{l=1}^{D} \sqrt{\lambda_l} u_{jl} dZ_l\big]. \tag{11.79}$$

Now we denote ρ_l^Z as the correlation between Z_l and dW. To preserve the correlation between W_j and dW, the following condition on Z_l and dW must be fulfilled,

$$Corr < \sum_{l=1}^{D} \sqrt{\lambda_l} u_{jl} dZ_l, dW > = \sum_{l=1}^{D} \sqrt{\lambda_l} u_{jl} \rho_l^Z dt \tag{11.80}$$

$$= \sum_{l=1}^{D} \eta_{jl} \rho_l^Z dt = \rho_j dt, \tag{11.81}$$

where $\eta_{jl} = \sqrt{\lambda_l} u_{jl}$. Setting $D = M$ leads to a linear equation system with M unknown variables ρ_l^Z,

$$\sum_{l=1}^{M} \eta_{jl} \rho_l^Z = \rho_j. \qquad 1 \leqslant j, l \leqslant M, \tag{11.82}$$

which can be solved for ρ_j^Z by a standard numerical procedure, for example, Gaussian elimination. For a factor representation of $L_j(t)$ we need only $\rho_l^Z, 1 \leqslant l \leqslant D$, and reformulate the dynamics of of the Libors and theirs stochastic volatilities,

$$\begin{aligned} dL_j(t) &= L_j(t) v_j(t) \sum_{l=1}^{D} \eta_{jl} dZ_l, \\ dv_j(t) &= [\kappa_j \theta_j - \kappa_j v_j(t)] dt + \sigma_j dW, \\ dW(t) dZ_l(t) &= \rho_l^Z dt. \qquad 1 \leqslant l \leqslant D. \end{aligned} \tag{11.83}$$

The factor representation (11.83) for the extended stochastic volatility LMM are based on the first D principal components, not whole components. Additionally, a correction of variance to the original value is important when we implement an exact Monte-Carlo simulation and the approximated swaption pricing formula.

As in other extended LMMs, the process for the swap rate is approximated in a lognormal form as a sum of weighted Libors, and takes the following form

$$\frac{dS_{n,m}(t)}{S_{n,m}(t)} = \sum_{k=n}^{m-1} \omega_k(0) v_k(t) dW_k(t), \tag{11.84}$$

where the weights $\omega_k(0)$ is identical to the one in the previous models. Setting the factor representation (11.83) of $L_h(t)$ into the above approximated swap rate process yields

$$\frac{dS_{n,m}(t)}{S_{n,m}(t)} = \sum_{l=1}^{D} \sum_{k=n}^{m-1} g_{kl} v_k(t) dZ_l(t)$$

with

$$g_{kl} = \omega_k(0) \eta_{kl}.$$

In order to simplify the expression of the swap rate process and makes its structure more transparent, we summarize the terms $\sum_{k=n}^{m-1} g_{kl} v_k(t)$ with a single process. By applying Itô's lemma, we can show that given the mean-reverting Ornstein-Uhlenbeck volatility process $v_k(t)$ in (11.67), the sum $v_s(t) = \sum_{k=n}^{m-1} q_k v_k(t)$ with the constant weights q_k should be governed by an other mean-reverting Ornstein-

Uhlenbeck process,

$$dv_s(t) = \kappa_s(\theta_s - v_s(t))dt + \sigma_s dW \tag{11.85}$$

with

$$v_s(0) = \sum_{k=n}^{m-1} q_k v_k(0), \qquad \kappa_s \approx \frac{\sum_{k=n}^{m-1} q_k \kappa_k v_k(0)}{v_s(0)},$$

$$\sigma_s = \sum_{k=n}^{m-1} q_k \sigma_k, \qquad \theta_s = \frac{\sum_{k=n}^{m-1} q_k \kappa_k \theta_k}{\kappa_s}.$$

Finally, we rewrite the swap rate process in SMM by using D independent Wiener processes with its individual stochastic volatility process $v_l^s(t)$,

$$\frac{dS_{n,m}(t)}{S_{n,m}(t)} = \sum_{l=1}^{D} v_l^s(t) dZ_l(t) \tag{11.86}$$

with

$$v_l^s(t) = \sum_{k=n}^{m-1} g_{kl} v_k(t).$$

Based on the simple process for swap rate $S_{n,m}(t)$ in (11.86), we are able to derive the pricing formula for swaptions by applying CFs. Let $X(t) = \ln S_{n,m}(t) - \ln S_{n,m}(0)$, we have

$$dX(t) = \sum_{l=1}^{D} \left[-\frac{1}{2}(v_l^s)^2(t)dt + v_l^s(t)dZ_l(t) \right]. \tag{11.87}$$

Although $v_l^s(t), 1 \leqslant l \leqslant D$, are correlated with each other, but $Z_l(t), 1 \leqslant l \leqslant D$, are not. This independence leads to

$$Corr < v_l^s(t)dZ_l(t), v_k^s(t)dZ_k(t) >= 0, \qquad 1 \leqslant l,k \leqslant D, \quad l \neq k.$$

Additionally, by Itô' rule the cross terms $(v_l^s)^2(t)dt \cdot (v_k^s)^2(t)dt$, $(v_l^s)^2(t)dt \cdot v_k^s(t)dZ_k(t)$ have at least an order of $(dt)^{\frac{3}{2}}$, and do not deliver a dominate effect for the dynamics of $X(t)$. The correlation between $v_l^s(t)$ and $Z_l(t)$ is given by $dW(t)dZ_l(t) = \rho_j^Z dt$ where ρ_j^Z has been calculated by the factorization procedure.

Using the independence of $v_l^s(t)dZ_l(t)$ and neglecting the second-order effect of cross terms, Zhu derived the approximated characteristic functions of $X(t)$ under the swap measure Q^S. For some complex number p, the moment-generating function for $X(t)$ is given by

$$\begin{aligned} f_X(p,T) &= \mathbf{E}^S \left[\exp\left(pX(T) \right) \right] \\ &= \mathbf{E}^S \left[\exp\left(-\frac{p}{2} \sum_{l=1}^{D} \int_0^T (v_l^s)^2(t)dt + \sum_{l=1}^{D} \int_0^T v_l^s(t)dZ_l(t) \right) \right] \end{aligned}$$

$$= \prod_{l=1}^{D} \mathbf{E}^S \Big[\exp\Big(-\frac{p}{2} \int_0^T (v_l^s)^2(t)dt + p \int_0^T v_l^s(t)dZ_l(t) \Big) \Big]$$
$$= \prod_{l=1}^{D} f_l^*(p,T) \tag{11.88}$$

with

$$f_l^*(p,T) = \mathbf{E}\Big[\exp\Big(-\frac{p}{2} \int_0^T (v_l^s)^2(t)dt + p \int_0^T v_l^s(t)dZ_l(t) \Big) \Big].$$

The moment-generating function $f_X(p,T)$ is then approximated by a product of the moment-generating function $f_l^*(\phi,p)$ of each factor $v_l^s(t)dZ_l(t)$. The concrete functions of $f_l^*(\phi,p)$ are given in the Schöbel-Zhu model (see Section 3.3). If we specify the stochastic variance $V_j(t)$ to be a mean-reverting square root process as in the Heston model, the function $f_l^*(\phi,p)$ is determined correspondingly (see Section 3.2). Applying the moment-generating function, we define two CFs with two different p,

$$f_1^s(\phi) = \exp\Big(i\phi \ln S_{n,m}(0) \Big) f_X(-i+\phi, T_n), \tag{11.89}$$

$$f_2^s(\phi) = \exp\Big(i\phi \ln S_{n,m}(0) \Big) f_X(i\phi, T_n). \tag{11.90}$$

The pricing formula for a payer swaption via CFs is given by

$$SWP(0) = A_{n,m}(0)[S_{n,m}(0) \cdot F_1 - K \cdot F_2] \tag{11.91}$$

with

$$F_k = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\Big(f_k^s(\phi)) \frac{\exp(i\phi \ln K)}{i\phi} \Big) d\phi, \quad k = 1,2.$$

The receiver swaption takes the similar form by applying the put-call parity,

$$SWR(0) = A_{n,m}(0)[K \cdot (1-F_1) - S_{n,m}(0) \cdot (1-F_2)]. \tag{11.92}$$

The numerical implementation is identical to other stochastic volatility models except that we have to integrate the CFs of each factor into a single CF. This pricing formula for swaption is numerically very efficient and admits a quickly pricing and calibration. For a practical application, five factors are usually sufficient to cover most dynamics of a yield curve reflected in swaptions.

11.4.5 The Belomestny, Matthew and Schoenmakers Model

Belomestny, Matthew and Schoenmakers(BMS, 2007) proposed an extended LMM with a high-dimensional specially structured system of mean-reverting square variance processes. Additionally, the model of BMS involves a deterministic volatility term structure $q_i(t)$ as in the ABR model and the Piterbarg model, and a proportion factor α that links the dynamics of standard LMM and the added dynamics of stochastic volatilities. In particular, the BMS model[5] has the following structure under some measure,

$$\begin{aligned}\frac{dL_j(t)}{L_j(t)} &= (\cdots)dt + \sqrt{1-\alpha^2}[q_j(t)\cdot dW(t)] + \alpha[\beta_j(t)\cdot dU(t)],\\ dU_k &= \sqrt{V_k(t)}dZ_k(t), \quad 1 \leqslant k \leqslant D, \qquad (11.93)\\ dV_k(t) &= \kappa_k(1-V_k(t))dt + \sigma_k\sqrt{V_k(t)}(\rho_k Z_k(t) + \sqrt{1-\rho_k^2}dZ_k^*(t), \qquad (11.94)\end{aligned}$$

where Z and Z^* are two independent D-dimensional uncorrelated standard Brownian motions, and both also independent of W that is a M-dimensional uncorrelated standard Brownian motions. Both $q_j(t) \in \mathbb{R}^N$ and $\beta_j \in \mathbb{R}^D$ are deterministic vector functions. The parameter α describes the leverage effect between Brownian motions W and U, and may also be interpreted as a perturbation to standard LMM. If α reduces to zero, this model degenerates to a standard LMM. In a more compact manner, the above system can be reformulated in a vector-valued stochastic process of dimension M driven by $(N+2D)$ standard Brownian motions $\mathscr{W}$, that takes the form

$$\frac{dL_j(t)}{L_j(t)} = (\cdots)dt + \Gamma_j \cdot d\mathscr{W} \qquad (11.95)$$

with

$$\Gamma_j = \begin{pmatrix} \sqrt{1-\alpha^2}q_{j1} \\ \sqrt{1-\alpha^2}q_{j2} \\ \cdot \\ \cdot \\ \sqrt{1-\alpha^2}q_{jN} \\ \alpha\beta_{j1\sqrt{V_1}} \\ \cdot \\ \cdot \\ \alpha\beta_{jD\sqrt{V_D}} \end{pmatrix}, \quad \mathscr{W} = \begin{pmatrix} dW_1 \\ dW_2 \\ \cdot \\ \cdot \\ dW_N \\ dZ_1 \\ \cdot \\ \cdot \\ dZ_D \end{pmatrix},$$

which shows the high dimensionality of the BMS model.

To value caplets and floorlets as well as swaptions, we need, however, the dynamics of Libors and volatilities under the respective forward measure, that can be obtained via the standard change of measures. The system of whole dynamics given in (11.93) under the terminal measure Q^{M+1} reads

[5] We consider here a slightly simplified BMS model with which Belomestny, Matthew and Schoenmakers actually worked.

$$\begin{aligned}\frac{dL_j(t)}{L_j(t)} &= \sqrt{1-\alpha^2}[q_j(t)\cdot dW(t)] + \alpha\sum_{k=1}^{D}\beta_{jk}\sqrt{V_k(t)}dZ_k(t),\\ dV_k(t) &= \kappa_k^*(\theta_k^* - V_k(t))dt + \sigma_k\sqrt{V_k(t)}\rho_k Z_k(t) + \sqrt{1-\rho_k^2}dZ_k^*(t)\end{aligned} \tag{11.96}$$

with

$$\kappa_k^* = \kappa_k - \alpha\sigma_k\rho_k\sum_{i=j+1}^{M}\frac{\tau_i L_i(0)}{1+\tau_i L_i(0)}\beta_{ik}$$

and

$$\theta_k^* = \kappa_k/\kappa_k^*.$$

In order to value swaptions, Belomestny, Matthew and Schoenmakers applied the same approach as used in other extended LMMs and obtain the following process for swap rate

$$\frac{dS_{n,m}(t)}{S_{n,m}(t)} = \Big[\sum_{i=n}^{m-1}\omega_i(0)\Gamma_i\Big]\cdot d\mathscr{W}^S \tag{11.97}$$

with

$$\omega_i(0) = \frac{\partial S_{n,m}(0)L_i(0)}{\partial L_i(0)S_{n,m}(0)}.$$

The process for variance $V_i(t)$ under swap measure is much more complicated. By freezing the values of Libors $L(t)$ at the values at the origin, we may obtain

$$dV_k^S(t) = \kappa_k^S(\theta_k^S - V_k(t))dt + \sigma_k\sqrt{V_k(t)}\rho_k Z_k^S(t) + \sqrt{1-\rho_k^2}dZ_k^{*,S}(t), \tag{11.98}$$

with

$$\kappa_k^S = \kappa_k - \alpha\sigma_k\rho_k\sum_{h=n}^{m-1}\frac{\tau_h B(0,T_h)}{A_{n,m}(0)}\sum_{i=h+1}^{M}\frac{\tau_i L_i(0)}{1+\tau_i L_i(0)}\beta_{ik}$$

and

$$\theta_k^S = \kappa_k/\kappa_k^S.$$

The calibration of this LMM includes two sub-calibrations. The first one is a pre-calibration to ATM caplet or ATM swaption market prices. This pre-calibration is performed with the standard LMM having no stochastic volatility, and its goal is to find $q_j(t), 1 \leqslant j \leqslant M$. The second calibration will fit the smile pattern of caplet or swaption volatilities by searching for appropriate parameter values of the volatility process $V_j(t)$ and the leverage parameter α. For the pre-calibration we refer the readers to Section 11.2, or in the connection with the BMS model here, to Schoenmakers (2005). For the second calibration, Belomestny, Matthew and Schoenmakers imposed the following assumptions and restrictions:

1. To reduce the number of parameters, there are some restrictions for covariance structure, from which a special equation for q_i and β_i holds

$$\int_0^t q_i^\top q_j ds = \int_0^t \beta_i^\top \beta_j ds.$$

Note that q_i is used in the BMS model to fit the market ATM volatilities of caps and swaptions, and has been determined in the pre-calibration, β_i will be estimated from q_i.

2. The number M of Libors is equal to the number D of the Brownian motions driving stochastic volatilities. $\beta \in \mathbb{R}^{M\times D}$ is a squared upper triangular matrix of a rank M. This means that the last Libor $L_M(t)$ depends only on the volatility dynamics U_M, and the Libor $L_{M-1}(t)$ only on U_M and U_{M-1}, and so forth.
3. The spot variance $V_k(0)$ is equal to θ_k, namely $V_k(0) = 1$. The parameters left to estimate are then κ, σ and ρ for each Libor.

Thanks to the above assumptions, and especially to the upper triangular structure of β, a recursive strategy for the estimation of κ, σ, ρ and α from L_M down to L_1 can be developed. For a more detailed calibration, please see Belomestny, Matthew and Schoenmakers (2007). With the respect to the assumption (2), the volatility dynamics of a Libor $L_j(t)$ are solely determined by the variance $V_k, j \leqslant k \leqslant M$. As a result, the Libor process can be rewritten as

$$\frac{dL_j(t)}{L_j(t)} = \sqrt{1-\alpha^2}[q_j(t)\cdot dW(t)] + \alpha \sum_{k=j}^{D} \beta_{jk}\sqrt{V_k(t)}dZ_k(t). \tag{11.99}$$

Now we consider the valuation of caplets via the Fourier transform in the BMS model. As indicated in the above Libor process, each Libor process includes $D-j$ independent stochastic volatility processes $V_k(t)$, that in turn have only the deterministic parameters, we face then a similar valuation problem as for swaptions in the Zhu model. Belomestny, Matthew and Schoenmakers derived the following CF for $\ln L_j(t)$ under the forward measure Q^{j+1},

$$f(\phi) = e^{i\phi \ln L_j(0)} f_0^*(\phi) \sum_{k=j}^{D} f_k^*(\phi) \tag{11.100}$$

where

$$f_0^*(\phi) = \exp\Big(-\frac{1}{2}(1-\alpha^2)(\phi^2+i\phi)\int_0^{T_j} |q_j|^2 dt\Big)$$

and

$$f_k^*(\phi) = \exp(H_{k,1}(T_j) + H_{k,2}(T_j)V_k(0))$$

with

$$H_{k,1}(T_j) = \frac{\kappa_k^*\theta_k^*}{\sigma_k^2}\Big[(a_k+d_k)T_j - 2\ln\Big(\frac{1-c_k e^{d_k T_j}}{1-c_k}\Big)\Big]$$

and

$$H_{k,2}(T_j) = \frac{[a_k+d_k](1-e^{d_k T_j})}{\sigma_k^2(1-c_k e^{d_k T_j})}.$$

The parameters a_k, b_k and c_k are given by[6]

$$
\begin{aligned}
a_k &= \kappa_k - i\sigma_k\rho_k\alpha\beta_{jk}\phi, \\
d_k &= \sqrt{a_k^2 - \alpha^2\beta_{jk}^2\sigma_k^2(\phi^2 + i\phi)}, \\
c_k &= \frac{a_k + d_k}{a_k - d_k}.
\end{aligned}
\tag{11.101}
$$

The deviation of the pricing formula for caplets/floorlets then follows the same way as discussed above in the Zhang and Wu model, or in the Zhu model.[7] Particularly, we have the two CFs for the exercise probabilities,

$$f_1(\phi) = f(-i + \phi), \quad f_2(\phi) = f(\phi).$$

The derivation of the pricing formula for swaptions follows the same way by applying the process (11.98) for swap rate.

11.5 Conclusive Remarks

In this chapter, we have briefly addressed the standard LMM and shown that the market pricing approach for caplets and floorlets via Black76 formula agrees with the standard LMM where each Libor lives in its own stochastic world with its own probability measure. This also highlights the nature of the high-dimensionality and the mutual correlation of LMM. Since a swaption may be regarded as a correlation product of Libors, the analytical valuation of swaption in LMM is never exact and requires some ad-hoc approximations. However, a swaption can be valued analytically with a Black-like formula in Swap Market Model (SMM) in which a swap annuity is used as numeraire. It can be shown that LMM and SMM do not agree with each other.

The smile/skew effects in cap markets and swaption markets are evident, and ever stronger than these ones in equity markets. To recover the implied volatility surface of caps and swaptions, LMM, as equity option pricing model, should accommodate stochastic volatility as the first choice for extension. However, adding stochastic volatility dynamics to LMM increases dramatically the technical complexity and the number of model parameters. In particular, the (semi-) analytical valuation of swaptions in extended LMMs will encounter more problems. We have examined five stochastic volatility LMMs. The first two models of Andersen and Brotherton-Ratcliffe (2000), and Piterbarg (2003) are in fact a mixture of local volatility and stochastic volatility. In their models, stochastic volatilities are uncorrelated with the

[6] The index j is suppressed in the following expressions for notation simplicity.

[7] The pricing formula for caplets given by BMS, however, takes a different form which is similar to Attari's formula in Chapter 4.

corresponding Libors, and local volatilities are responsible for generating down-slope skews. the Zhang and Wu model (2006), the Zhu model (2007) and the BMS model (2007) allow not only stochastic volatility in LMM, but also a non-zero correlation between stochastic volatility and its Libor. Consequently, these three models utilize intensively the Fourier transform as a powerful tool to derive caplet and swaption pricing formulas.

Recently, SABR model has drawn a strong attention in smile modeling and is combined with LMM. However, the technique to develop SABR model is beyond the scope of Fourier transform, hence SABR model is not addressed in this book. To capture the smile/skew effects, similarly to equity modeling, jump-diffusion model may be an alternative way. For example, Glasserman and Kou (2003), Belomestny and Schoenmakers (2006) proposed two special LMMs with jump components.

In spite of a number of vigorous efforts to search for a "better" LMM, not all valuation issues with respect to smile/skew effects are satisfactorily solved. In particular, recovering swaption smiles for all tenors and option maturities is still a challenge. In other words, a global consistent calibration to swaption markets is a stress test for all extended LMMs. From the point of view of practitioners, global consistent calibration to swaption markets is of great importance for reliable and efficient hedging of interest rate derivatives. For example, Bermudan swaptions, CMS spread options and range accruals are inhomogeneous in terms of the underlying rate and the structure. As a result, any trading book of interest rate derivatives in investment banks including such instruments may be exposed to the risks of these inhomogeneous underlying rates. To hedge these risks efficiently in a portfolio, a global consistent calibration is then necessary, and even crucial. A partial fitting of swaption smiles may result in a non-agreement with Vega risks in a trading book. Additionally, parameter parsimoniousness is often an underestimated aspect. Every additional model parameter increases model risk, and reduces the manageability of a model. Therefore, a reasonable tradeoff between model parsimoniousness and model capability of fitting markets should be achieved. Nevertheless, thanks to the Fourier transform and CFs, as shown above, remarkable achievements of the smile modeling within the LMM framework have been arrived.

References

1. M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulae, Graphs and Mathematical Tables*. Dover Publications, New York, 9th edition, 1965.
2. H. Albrechter, P. Mayer, W. Schoutens, and J. Tistaert. The little Heston trap. Radon Institute, Austrian Academy of Sciences, and Graz University of Technology, 2006.
3. L. Andersen. Efficient simulation of the Heston stochastic volatility model. Bank of America Securities, 2007.
4. L. Andersen and R. Brotherton-Ratcliffe. Extended Libor market models with stochastic volatility. *Journal of Computational Finance*, 9:1–40, 2005.
5. A. Antonov and T. Misirpashaev. Markovian projection onto a displaced diffusion: Generic formulas with applications. Available at SSRN: http://ssrn.com/abstract=937860, 2006.
6. M. Attari. Option pricing using Fourier transforms: A numerically efficient simplification. Charles River Associates, Inc., 2004.
7. G.S. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52:2003–2049, 1997.
8. G.S. Bakshi and Z. Chen. Equilibrium valuation of foreign exchange claims. *Journal of Finance*, 52:799–826, 1997.
9. G.S. Bakshi and D. Madan. Spanning and derivative security valuation. *Journal of Financial Economics*, 55:205–238, 2000.
10. C.A. Ball and A. Roma. Stochastic volatility option pricing. *Journal of Financial and Quantitative Analysis*, 29:581–607, 1994.
11. O.E. Barndorff-Nielsen and N. Shephard. 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.
12. O.E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Stochastics and Finance*, 2, 1998.
13. D. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies*, 6:69–107, 1996.
14. D. Beaglehole and M. Tenney. Corrections and additions to "a nonlinear equilibrium model of the term structure of interest rates". *Journal of Financial Economics*, 28:346–353, 1992.
15. D.R. Beaglehole, P.H. Dybvig, and G. Zhou. Going to extremes: correcting simulation bias in exotic option valuation. *Financial Analysts Journal*, 53:62–68, 1997.
16. S. Beckers. The constant elasticity of variance model and its implications for option pricing. *Journal of Finance*, 35:661–673, 1980.
17. S. Beckers. Variances of security price returns based on high, low and closing prices. *Journal of Business*, 56:97–112, 1983.
18. D. Belomestny, S. Matthew, and J. Schoenmakers. A stochastic volatility Libor model and its robust calibration. Discussion Paper, Humboldt University zu Berlin, 2007.

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4,

19. D. Belomestny and J. Schoenmakers. A jump-diffusion Libor model and its robust calibration. Discussion Paper, WIAS Berlin, 2006.
20. J. Bertoin. *Lévy Processes*. Cambridge University Press, Cambridge, 1996.
21. P. Billingsley. *Probability and Measure*. Wiley Inc, New York, 2th edition, 1986.
22. T. Björk. *Interest Rate Theory in Financial Mathematics*. Springer, Berlin, New York, 1996.
23. F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, 1976.
24. F. Black and J. Cox. Valuing corporate securities: some effects of bond indenture provisions. *Journal of Finance*, 31:351–367, 1976.
25. F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.
26. P.P. Boyle. Options: A Monte-Carlo approach. *Journal of Financial Economics*, 4:323–338, 1977.
27. D. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 6:265–296, 1979.
28. D. Brigo and F. Mercurio. *Interest Rate Models; Theory and Practice*. Springer, Heidelberg, Berlin, 2nd edition, 2006.
29. M. Broadie, P. Glasserman, and S. Kou. A continuity correction for discrete barrier options. *Mathematical Finance*, 7:325–349, 1997.
30. M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump-diffusion processes. *Operations Research*, 54:217–231, 2006.
31. S.J. Brown and P.H. Dybvig. The empirical implications of the Cox, Ingersoll and Ross theory of the term structure of interest rates. *Journal of Finance*, 41:617–632, 1986.
32. P. Carr and D. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.
33. P. Carr, H. Geman, D.H. Madan, and M. Yor. Stochastic volatility for Levy process. Working Paper 645, Unversités de Paris 6, 7, 2001.
34. P. Carr and D. Madan. Option valuation with the fast Fourier transform. Working Paper, Morgan Stanley and University of Maryland, 1998.
35. P. Carr and L. Wu. Finite moment log stable process and option pricing. *Journal of Finance*, 58:753–777, 2003a.
36. P. Carr and L. Wu. What type of process underlying options? *Journal of Finance*, 58:2581–2610, 2003b.
37. P. Carr and L. Wu. Time-changed Lévy process and option pricing. *Journal of Financial Economics*, 71:113–141, 2004.
38. R. Chen and L. Scott. Interest rate options in multifactor Cox-Ingersoll-Ross models of the term structure. *Journal of Derivatives*, 3:53–72, 1995.
39. M. Chernov, A. Gallant, E. Ghysels, and G. Tauchen. Notes on options pricing: Constant elasticity of variance diffusions. Working Paper, Stanford University, 1975.
40. M. Chernov, A. Gallant, E. Ghysels, and G. Tauchen. A new class of stochastic volatility models with jumps: Theory and estimation. Working paper, The Pennsylvania State University, 1999.
41. K. Chourdakis. Option pricing using the fractional FFT. *Journal of Computational Finance*, 8:1–18, 2005.
42. K.L. Chung. *A Course in Probability Theory*. Academic Press, New York, 2nd edition, 1974.
43. R. Cont and P. Tankov. *Financial Modeling with Jump Processes*. Chapman and Hall/CRC Press, London, New York, Singapore, 2003.
44. G. Coutadon. The pricing of options on default-free bonds. *Journal of Financial and Quantitative Analysis*, 17:75–100, 1982.
45. D.R. Cox and H.D. Miller. *The Theory of Stochastic Process*. Chapman and Hall Ltd, New York, 3th edition, 1972.
46. J.C. Cox, J.E. Ingersoll Jr., and S.A. Ross An intertemporal general equilibrium model of asset prices. *Econometrica*, 53:363–384, 1985a.

47. J.C. Cox and S.A. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166, 1976.
48. J.C. Cox, S.A. Ross, and M. Rubinstein. Option pricing: a simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.
49. J.C. Cox and M. Rubinstein. *Options Markets*. Prentice-Hall Inc, New Jersey, 1985.
50. Q. Dai and K. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55:1943–1978, 2000.
51. E. Derman. Regimes of volatility. *Risk*, 12(4):55–59, 1999.
52. E. Derman and I. Kani. Riding on the smile. *Risk*, 7:32–39, 1994.
53. U. Dothan. On the term structure of interest rates. *Journal of Financial Economics*, 6:59–69, 1978.
54. J.C. Duan. A unified theory of option pricing under stochastic volatility-from GRACH to diffusion. Working Paper, Department of Finance, Hong Kong University of Science and Technology, 1996.
55. D. Duffie. *Security Markets: Stochastic Models*. Academic Press, Boston, 1988.
56. D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey, 2nd edition, 1996.
57. D. Duffie and R. Kan. A yield factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.
58. D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376, 2000.
59. B. Dumas, J. Fleming, and R. Whaley. Implied volatility functions: Empirical tests. *Journal of Finance*, 53:111–127, 1998.
60. B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.
61. E. Eberlein, U. Keller, and K. Prause. New insights into smile mispricing, and value at risk: The hyperbolic model. *Journal of Business*, 71:371–406, 1998.
62. E.F. Fama. The behavior of stock market prices. *Journal of Business*, 38:34–105, 1965.
63. G.S. Fishman. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Verlag, New York, 1996.
64. J.-P. Fouque, G. Papanicolaou, and K.R. Sircar. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press, Cambridge, 2000.
65. M. Fu, D. Madan, and T. Wang. Pricing continuous time Asian options: A comparison of analytical and Monte-Carlo methods. Working paper, University of Maryland, 1995.
66. M. Garman. Spread the load. *Risk*, 5:68–84, 1992.
67. M. Garman and S. Kohlhagen. Foreign currency options values. *Journal of International Money and Finance*, 2:231–237, 1983.
68. J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley & Sons Inc., New York, 2006.
69. H. Geman, N. EL Karoui, and J.-C. Rochet. Changes of numeraire, changes of probability and option pricing. *Journal of Applied Probability*, 32:443–458, 1995.
70. R. Geske. The valuation of compound options. *Journal of Financial Economics*, 7:63–81, 1979.
71. P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, Heidelberg, New York, 2004.
72. P. Glasserman and S. Kou. The term structure of simple forward rates with jump risk. *Mathematical Finance*, 13:393–410, 1997.
73. M. Goldman, H. Sosin, and M. Gatto. Path dependent options: Buy at the low, sell at the high. *Journal of Finance*, 34:111–127, 1979.
74. I. Gyöngy. Mimicking the one-dimensional distributions of processes having an Ito differential. *Probability Theory and Related Fields*, 71:501–516, 1986.
75. J.M. Harrison. *Brownian Motion and Stochastic Flow Systems*. Wiley Press, New York, 1985.
76. J.M. Harrison and D. Kreps. Martingales and arbitrage in multiperiod security markets. *Journal of Economic Theory*, 20:381–408, 1979.

77. J.M. Harrison and S.R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.
78. D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation. *Econometrica*, 60:77–105, 1992.
79. S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.
80. S. Heston. A simple new formula for options with stochastic volatility. Working Paper, John M. Olin School of Business, Washington University in St. Louis, 1997.
81. S. Heston and S. Nandi. A closed-form GRACH option model. Working Paper, Federal Reserve Bank of Atlanta, 1997.
82. T.S. Ho and S.B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41:1011–1029, 1986.
83. J. Huang and L. Wu. Specification analysis of option pricing models based on time-changed Lévy processes. *Journal of Financial Economics*, 59:1405–1440, 2004.
84. J. Hull. *Options, Futures and Other Derivatives*. Prentice Hall Inc, New Jersey, 6th edition, 2006.
85. J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.
86. J. Hull and A. White. Pricing interest rate derivative securities. *Review of Financial Studies*, 3:573–592, 1990.
87. J. Hull and A. White. One-factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis*, 28:235–254, 1993.
88. J. Hull and A. White. Forward rate volatilities, swap rate volatilities, and the implementation of the Libor market model. Working Paper, University of Toronto, 1999.
89. P. Hunt and J. Kennedy. *Financial Derivatives in Theory and Practice*. John Wiley & Sons, 2004.
90. P. Hunt, J. Kennedy, and A. Pelsser. Markov-functional interest rate models. *Finance and Stochastics*, 4:391–408, 2000.
91. S.H. Hurst, E. Platen, and S.T. Rachev. Option pricing for a logstable asset price model. *Mathematical and Computer Modeling*, 29:105–119, 1999.
92. Jr. J.E. Ingersoll. *Theory of Financial Decision Making*. Rowman & Littelfield Publishers, Inc, Maryland, 1987.
93. F. Jamshidian Libor and swap market models and measures. *Finance and Stochastics*, 1:293–330, 1997.
94. J.M. Chambers, C.L. Mallows, and B.W. Stuck. A method for simulating stable random variables. *Journal of American Statistical Association*, 71:340–344, 1976.
95. H. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distribution, Vol 1*. John Wiley & Sons Inc, New York, 1987.
96. H. Johnson and D. Shanno. Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis*, 22:143–151, 1987.
97. P. Jorion. On jump processes in the foreign exchange and stock markets. *Review of Financial Studies*, 1:427–446, 1988.
98. N.J. Ju. Fourier transformation, martingale, and the pricing of average rate derivatives. Working Paper, Haas School of Business, University of California, Berkeley, 1997.
99. C. Kahl and P. Jaeckel. Fast strong approximation Monte-Carlo schemes for stochastic volatility models. ABN AMRO and University of Wuppertal, 2005.
100. C. Kahl and P. Jaeckel. Not-so-complex logarithms in the Heston model. ABN AMRO and University of Wuppertal, 2006.
101. S. Karlin and H. Taylor. *A First Course in Stochastic Process*. Academic Press, San Diego, 1975.
102. S. Karlin and H. Taylor. *A Second Course in Stochastic Process*. Academic Press, San Diego, 1975.
103. F. Kilin. Accelerating the calibration od stochastic volatility models. Working Paper, HfB, Frankfurt/M, May 2007.

104. S.G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.
105. N. Krause and U. Nögel. On the pricing of forward-starting option in Heston model on stochastic volatility. *Stochastics and Finance*, 9:233–250, 2005.
106. R.W. Lee. Option pricing by transform methods: Extensions, unification and error control. *Journal of Computational Finance*, 7:51–86, 2004.
107. E. Levy. Pricing European average-rate currency options. *Journal of International Money and Finance*, 11:474–491, 1992.
108. E. Levy and S. Turnbull. Average intelligence. *Risk*, 5:53–59, 1992.
109. A. Lewis. *Option Valuation Under Stochastic Volatility*. Finance Press, California, 2000.
110. A. Lewis. A simple option pricing formula for general jump-diffusion and other exponential Lévy processes. Working Paper, OptionCity.Net, 2001.
111. A. Lipton and A. Sepp. Stochastic volatility models and Kelvin waves. *Journal of Physics A: Mathematical and Theoretical*, 41:344012, 2008.
112. F.A. Longstaff. A nonlinear general equilibrium model of the term structure of interest rates. *Journal of Financial Economics*, 23:195–224, 1989.
113. F.A. Longstaff. Multiple equilibria and term structure models. *Journal of Financial Economics*, 32:333–344, 1992.
114. F.A. Longstaff. Hedging interest rate risk with options on average interest rates. *Journal of fixed Income*, 4:37–45, 1995.
115. F.A. Longstaff and E.S. Schwartz. Interest rate volatility and the term structure: A two factor general equilibrium model. *Journal of Finance*, 47:1259–1282, 1992.
116. F.A. Longstaff and E.S. Schwartz. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–819, 1995.
117. R. Koekkoek, R. Lord, and D. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *forthcoming in: Quantitative Finance*, 2008.
118. R. Lord and C. Kahl. Optimal Fourier inversion in semi-analytical option pricing. Rabo Bank International and University of Wuppertal, 2006.
119. J. MacBeth and L. Merville. An empirical examination of the Black-Scholes call option pricing model. *Journal of Finance*, 34:369–382, 1979.
120. D. Madan, P. Carr, and E.C. Chang. The variance Gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.
121. D. Madan and F. Milne. Option pricing with VG martingale components. *Mathematical Finance*, 1, 1991.
122. A.G. Malliaris and W.A. Brock. *Stochastic Methods in Economics and Finance*. North-Holland, New York, 6th edition, 1991.
123. B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.
124. W. Margrabe. The value of an option to exchange one asset for another. *Journal of Finance*, 33:177–186, 1978.
125. J.H. McCulloch. *Statistical Methods in Finance*, chapter Financial applications of stable distributions. Elsevier Science, Amsterdam, 1996.
126. Meeting of the American Statistical Association. *Studies of Stock Price Volatility Changes*, 1976.
127. R.C. Merton. *Continuous-time Finance*. Basil Blackwell, Oxford, 1990.
128. R.C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3:373–413, 1971.
129. R.C. Merton. Theory of rational option pricing. *Bell Journal of Economic and Management Science*, 4:141–183, 1973a.
130. R.C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41:867–888, 1973b.
131. R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
132. J.R. Michael, W. Schucany, and R. Haas. Generating random variates using transformations with multiple roots. *American Statistician*, 30:88–90, 1976.

133. K. Miltersen, K. Sandmann, and D. Sondermann. Closed form solutions for term structure derivatives with log-normal interest rates. *Journal of Finance*, 52:409–430, 1997.
134. M. Musiela and R. Rutkowski. *Martingale Methods in Financial Modelling*. Springer-Verlag, Berlin Heidelberg, 2nd edition, 2005.
135. H. Nagel. *Zur Optionsbewertung bei Stochastischer Volatilität am Beispiel der DTB: Eine Theoretische und Empirische Analyse*. PhD thesis, University of Tübingen, 1999.
136. S. Nandi. How important is the correlation between returns and volatility in a stochastic volatility model? empirical evidence from pricing and hedging in the s&p 500 index options market. *Journal of Banking and Finance*, 22:589–610, 1998.
137. D.B. Nelson and D.P. Foster. Asymptotic filtering theory for univariate ARCH models. *Econometrica*, 62:1–41, 1994.
138. U. Nögel and S. Mikhailov. Heston's stochastic volatility model: Implementation, calibration and some extensions. *WILMOTT Magazine*, July, 2003.
139. B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, Berlin Heidelberg New York, 6th edition, 2003.
140. A. Pelsser. *Efficient Methods for Valuing Interest Rate Derivatives*. Springer Verlag, Heidelberg, Berlin, 2000.
141. V. Piterbarg. A stochastic volatility forward Libor model with a term structure of volatility smile. Working Paper, Bank of America, 2003.
142. V. Piterbarg. Markovian projection method for volatility calibration. Working Paper, Barclays Capital, 2006.
143. G. Poitras. Spread options, exchange options, and arithmetic Brownian motion. *The Journal of Futures Markets*, 18:487–517, 1998.
144. I. Popova and P. Ritchken. On bounding option prices in paretian stable markets. *Journal of Derivatives*, 2:32–43, 1998.
145. S.A. Popovici. *An Analysis of Stochastic Volatility Models*. PhD thesis, University of Bonn, University Paris 6, 2003.
146. J.M. Poterba and L.H. Summers. Mean-reversion in stock prices: Evidences and implications. *Journal of Financial Economics*, 22:27–60, 1988.
147. R. Priestley. *Introduction to Complex Analysis*. Oxford University Press, 1990.
148. P. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin Heidelberg New York, 3 edition, 1995.
149. R. Rabinovitch. Pricing stock and bond options when the default-free rate is stochastic. *Journal of Financial and Quantitative Analysis*, 24:447–457, 1989.
150. R. Rebonato. *Interest Rate Option Models*. Wiley, Chichester, 2nd edition, 1998.
151. D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 1991.
152. D.R. Rich. *Advances in Futures and Options Research*, volume 7, chapter The Mathematical Foundations of Barrier Option-Pricing Theory, pages 267–311. JAI Press Inc, 1994.
153. M. Rubinstein. Nonparametric tests of alternative options pricing models using all reported trades and quotes on the 30 most active CBOE options from august 23, 1976 through august 31,1978. *Journal of Finance*, 40:455–480, 1985.
154. M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49:771–818, 1994.
155. M. Rubinstein and E. Reiner. Breaking down the barriers. *Risk*, 4:28–35, 1991.
156. P.A. Samuelson. Rational theory of warrant prices. *Industrial Management Review*, 6:13–31, 1965.
157. K.I. Sato. *Lévy Process and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.
158. R. Schmalensee and R. Trippi. Common stock volatility expectations implied by option premia. *Journal of Finance*, 33:129–147, 1978.
159. R. Schöbel. *Kapitalmarkt und zeitkontinuierliche Bewertung*. Physica-Verlag, Heidelberg, 1995.
160. R. Schöbel and J.W. Zhu. Stochastic volatility with an Ornstein-Uhlenbeck process: an extension. *European Finance Review*, 3:23–46, 1999.

161. R. Schöbel and J.W. Zhu. A Fourier analysis approach to the valuation of interest rate derivatives. Working Paper, University of Tübingen, 2000.
162. J. Schoenmakers. *Robust Libor Modeling and Pricing of Derivative Products*. Chapman and Hall/CRC Press, London, New York, Singapore, 2005.
163. J. Schoenmakers and C. Coffey. Stable implied calibration of a multi-factor Libor model via a semi-parametric correlation structure. Discussion Paper 611, Weierstrass Institute, 2000.
164. W. Schoutens. *Lévy Processes in Finance*. John Wiley & Sons, London, 2003.
165. W. Schoutens, E. Simons, and J. Tistaert. Model risk for exotic and moment derivatives. Working Paper, K.U. Leuven and ING SWE, 2004.
166. W. Schoutens, E. Simons, and J. Tistaert. Model risk for exotic and moment derivatives. K.U. Leuven and ING SWE Financial Modeling, 2004.
167. L.O. Scott. Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Financial Quantitative Analysis*, 22:419–438, 1987.
168. L.O. Scott. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates. *Mathematical Finance*, 7:413–426, 1997.
169. A. Sepp. Pricing European-style options under jump-diffusion processes with stochastic volatility: Application of Fourier transform. University of Tartu, Estonia, 2003.
170. N. Shephard. *Stochastic Volatility: Selected Readings*. Oxford University Press, Oxford, 2005.
171. D.C. Shimko. Options on futures spreads: Hedging, speculation, and valuation. *The Journal of Futures Markets*, 14:183–213, 1994.
172. K. Singleton. Estimation of affine diffusion models based on the empirical characteristic function. Working Paper, Stanford University, 1999.
173. E.M. Stein and J.C. Stein. Stock price distribution with stochastic volatility: An analytical approach. *Review of Financial Studies*, 4:727–752, 1991.
174. A. Stuard and O.J. Keith. *Kendall's Advanced Theory of Statistics, Vol. 1*. Edward Arnold, London, 5th edition, 1991.
175. E.C. Titchmarsh. *Introduction to the Theory of Fourier Integrals*. Chelsea Publication Co., New York, 1986.
176. S. Turnbull and L.M. Wakeman. A quick algorithm for pricing european average options. *Journal of Financial Quantitative Analysis*, 26:377–389, 1991.
177. O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.
178. J. Voit. *The Statistical Mechanics of Financial Markets*. Springer Verlag, New York, Heidelberg, 2003.
179. A.C.F. Vorst. Analytical boundaries and approximations of the prices and hedge ratios of average exchange rate options. Working Paper, Erasmus University Rotterdam, 1990.
180. R. Weron. On the Chambers-Mallows-Stuck method for simulating skewed random variables. *Statist. and Praob. Letter*, 28:165–171, 1996.
181. R. Weron. Correction to: On the Chambers-Mallows-Stuck method for simulating skewed random variables. Working Paper, Hugo Steinhaus Center, Wroclaw University of Technology, 1996a.
182. J.B. Wiggins. Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19:351–372, 1987.
183. D. Wilcox. Energy futures and options: Spread options in energy markets. Working Paper, Goldman Sachs & Co., New York, 1990.
184. L. Wu. Modeling financial security returns using Lévy processes. Working Paper, Zicklin School of Business, Baruch College, City University of New York, 2006.
185. L.X. Wu and F. Zhang. Libor market model with stochastic volatility. *Journal of Industrial and Management Optimization*, 2:199–227, 2006.
186. U. Wystup *FX Options and Structured Products*. Wiley, Chichester, England, 2006.
187. P.G. Zhang. *Exotic Options: A Guide to Second Generation Options*. World Scientific Publishing, Singapore, 1997.
188. C.S. Zhou. Jump risk, default rates, and credit spreads. Working paper, The Anderson Graduate School of Management, University of California, Riverside, 1998.

189. J.W. Zhu *Modular Pricing of Options*. Springer Verlag, Berlin, Heidelberg and New York, 2000.
190. J.W. Zhu. An extended Libor market model with nested stochastic volatility dynamics. Working Paper, Sal. Oppenheim, Frankfurt am Main, available: http://ssrn.com/abstract=955352, 2007.
191. J.W. Zhu. Generalized swap market model and the valuation of interest rate derivatives. Working Paper, LPA, Frankfurt am Main, available: http://ssrn.com/abstract=1028710, 2007b.
192. J.W. Zhu. A simple and exact simulation approach to the Heston model. Working Paper, LPA, Frankfurt am Main, available: http://ssrn.com/abstract=1153950, 2008.

Index

ABR model, 295
activity rate, 183
affine jump-diffusion, 168
alpha log-stable model, 194
alpha-stable distribution, 200
Andersen and Andreasen's weights, 289
approximation algorithm, 101
arbitrage-free, 10
arithmetic average Asian options, 244, 252
arithmetic Brownian motion, 258
Arrow-Debreu price, 32
ARV models, 47
Asian interest rate options, 254
Asian option, 244
at-the-period, 245
Attari's formula, 80
autoregressive random variance models, 47

backward change, 276
barrier option, 226
Bermudan swaption, 285
binary option, 266
Black-Scholes PDE, 6
Black-Scholes formula, 6
Black76 formula, 229
BMS model, 312
BNS model, 193
Bochner's procedure, 178
bootstrapping method, 291
Box-Mueller algorithm, 114
branch cut, 100
Broadie-Kaya Scheme, 122
Brownian motion, 2
butterfly volatility, 19

caching technique, 89
calendar spread, 257
call option, 5
call spread, 266
caplet, 275
caplet volatility, 281
caps, 275
Carr and Madan's formula, 78
cascade calibration, 291
characteristic function, 25
chi-square distribution, 49
Cholesky-decomposition, 131
chooser option, 267
CMS, 291
co-initial swaptions, 285
co-sliding swaptions, 285
co-terminal swaptions, 285
complex number, 99
complex power function, 100
compound jump process, 156
computation accuracy, 94
constant elasticity of variance, 46, 296
constant maturity swap, 291
convexity adjustment, 292
correlation matrix, 130, 287
correlation option, 257
Cox-Ingosoll-Ross model, 138
crack futures spread, 257
cross-correlation, 130
curvature of smile, 84

Deléans exponential, 24
Delta, 81
Delta measure, 24
Delta volatility, 19
density measure, 174
deterministic volatility function, 46
digital option, 266
discounted CF, 32, 170

J. Zhu, *Applications of Fourier Transform to Smile Modeling*, Springer Finance,
DOI 10.1007/978-3-642-01808-4,

discrete geometric average, 245
double square root, 63
double-exponential distribution, 166
drift, 3
drift compensator, 185
dynamic hedging, 7

eigenvalues, 132
equivalent martingale measure, 12
Euler scheme, 114
exchange option, 260
exotic options, 223
exponential distribution, 163

fast Fourier transform, 78
Feller's condition, 49, 116
Feynman-Kac theorem, 7
FFT, 78
filtration, 2
finite activity, 176
finite moment log-stable process, 196
finite variation, 176
first passage time, 181
flat volatility, 280
floorlet, 276
floors, 275
forward change, 278
forward measure, 24
forward volatility, 281
forward-starting option, 224
Fourier transform, 25
fractional FFT, 92
full rank formulations, 287
full truncation scheme, 124
futures price, 229
FX option, 17

Gamma, 81
gamma distribution, 193, 199
gamma function, 194
gamma Ornstein-Uhlenback process, 186
gamma process, 179
GARCH, 47
Garman-Kohlhagen formula, 18
Gaussian integration, 86
generalized Fourier transform, 188
generalized hyperbolic distribution, 179
geometric Brownian motion, 3
Girsanov theorem, 12
Gyöngy theorem, 108

Heston model, 48
hitting probability, 227
Hull and White's weights, 289

implied volatility, 15
in-the-money, 17
in-the-period, 245
infinite activity, 176
infinite variation, 176
infinitely divisible, 177
interest rate derivatives, 273
inverse Gaussian distribution, 181, 200
inverse transform method, 199
inverse-transform algorithm, 114
investment certificates, 223
Itô process, 4
Itô's lemma, 4
iteration algorithm, 101
ITM knock-out, 230

jump intensity, 158
jump size, 158
jump-diffusion process, 156

Kahl-Jäckel scheme, 124
Kolmogorov's backward equation, 39
Kou model, 165
kurtosis, 3, 31

Lévy exponent, 176
Lévy measure, 174
Lévy process, 174
Lévy triple, 176
Lévy-Khintchine theorem, 176
Laplace transform, 188
leptokurtic, 46
leverage effect, 188
leverage-neutral measure, 189
Libor, 275
Libor Market Model, 274
linear swap model, 292
Lipschitz condition, 4
LMM, 274
local volatility model, 46
local volatility process, 108
log-normal scheme, 117
lognormal jump, 160
Longstaff model, 144
lookback option, 235

main argument, 99
market completeness, 9
market price of risk, 8
Markovian projection, 108, 299
martingale, 12
mean Vega, 82
mean-reversion property, 46
Milstein scheme, 115

mimicking process, 110
model calibration, 104
model risk, 220
modified asset process, 140
modified geometric average, 254
modular pricing, 205
modulus, 99
moment-generating function, 31, 178
moment-matching, 115
money market account, 15
moneyness, 16
multi-asset model, 130
multi-domain integration, 88

nested stochastic volatility, 306
non-central chi-square distribution, 49
normal inverse Gaussian model, 181
numeraire, 14

ordinary differential equation, 3
Ornstein-Uhlenbeck process, 56
OTC markets, 223
OTM knock-out, 227
out and in parity, 233
out-of-the-money, 17
out-of-the-period , 245

Pareto jump, 163
partial integro-differential equation, 157
payer swap, 282
Piterbarg model, 298
Poisson distribution, 198
Poisson process, 154
portfolio, 5
present value, 7
principal components, 132, 309
principle component analysis, 309
probability measure, 2
probability space, 2
product option, 264
programming codes, 96
put option, 7
put spread, 266
put-call parity, 262
PVBP, 284

QE scheme, 120
quadratic variation, 176
quotient option, 263

Radon-Nikodym derivative, 12
Rayleigh distribution, 129
Rebonato formula, 290
receiver swap, 282
reflecting principle, 228
remaining uncertainty, 32
risk sensitivity, 80
risk-neutral valuation, 7
risk-reversal volatility, 19
rotation count algorithm, 102

SABR model, 294
Schöbel-Zhu model, 56
self-financing, 9
short rate process, 22
Simpson's rule, 92
simulation of maximum, 126
simulation of minimum, 126
skew, 16
skew function, 296
skewness, 3, 31
slope of smile, 84
smile, 16
smile surface, 16
sneer, 16
soybean complex spread, 257
spot measure, 279
spot Vega, 83
square root process, 48
stochastic clock model, 177
stochastic differential equation, 3
stochastic volatility model, 46
stop-loss strategy, 226
strike vector computation, 88
subordinator, 183
swap annuity, 284
swap contract, 282
swap market model, 284
swap rate, 284
swaption, 284
symmetric smile, 16

tenor structure, 275
term structure of interest rates, 137
terminal correlation, 288
terminal distribution, 109
terminal measure, 279
time-homogenous process, 4
trading strategy, 9
transformed volatility scheme, 118
Trapezoid rule, 92
truncation function, 175
truncation scheme, 117
two-factor model, 259

Vanna, 84
variance-gamma model, 179
Vasicek model, 142

Vega, 82
volatility, 3
volatility persistence, 46
volatility smile, 16
volatility term structure, 16, 282
Volga, 84

Wiener process, 2
Wu and Zhang model, 302

zero-coupon bond, 24, 136
Zhu model, 305