Managing Volatility Risk

Innovation of Financial Derivatives, Stochastic Models and Their Analytical Implementation

Chenxu Li

Submitted in partial fulfillment of the

Requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2010

ABSTRACT

Managing Volatility Risk

Innovation of Financial Derivatives, Stochastic Models and Their Analytical Implementation

Chenxu Li

This dissertation investigates two timely topics in mathematical finance. In particular, we study the valuation, hedging and implementation of actively traded volatility derivatives including the recently introduced timer option and the CBOE (the Chicago Board Options Exchange) option on VIX (the Chicago Board Options Exchange volatility index). In the first part of this dissertation, we investigate the pricing, hedging and implementation of timer options under Heston's (1993) stochastic volatility model. The valuation problem is formulated as a first-passage-time problem through a no-arbitrage argument. By employing stochastic analysis and various analytical tools, such as partial differential equation, Laplace and Fourier transforms, we derive a Black-Scholes-Merton type formula for pricing timer options. This work motivates some theoretical study of Bessel processes and Feller diffusions as well as their numerical implementation. In the second part, we analyze the valuation of options on VIX under Gatheral's double mean-reverting stochastic volatility model, which is able to consistently price options on S&P 500 (the Standard and Poor's 500 index), VIX and realized variance (also well known as historical variance calculated by the variance of the asset's daily return). We employ scaling, pathwise Taylor expansion and conditional Gaussian moments techniques to derive an explicit asymptotic expansion formula for pricing options on VIX. Our method is generally applicable for multidimensional diffusion models. The convergence of our expansion is justified via the theory of Malliavin-Watanabe-Yoshida. In numerical examples, we illustrate that the formula efficiently achieves desirable accuracy for relatively short maturity cases.

Contents

List of Tables

List of Figures

Part I

Bessel Process, Heston's Stochastic Volatility Model and Timer Option

Chapter 1

Introduction to Part I

1.1 A Brief Outline of Part I

The first part of this dissertation is motivated by the problems of pricing, hedging and implementation of timer options proposed in 2007 by Société Générale & Investment Banking as an innovative volatility derivative. Under Heston's (1993) stochastic volatility model, we rigorously formulate the perpetual timer call option valuation problem as a first-passage-time problem via a standard no-arbitrage argument and a stochastic representation of the solution to a boundary value problem. Motivated by this problem, we apply the time-change technique to find that the variance process modeled by Feller diffusion, running on a variance clock, is equivalent in distribution to a Bessel process with constant drift. We derive a joint density related to Bessel processes via Laplace transform techniques. Applying these results, we obtain a Black-Scholes-Merton type formula for pricing timer options. We also propose and compare several methods for implementation, including Laplace-Fourier transform inversion, Monte Carlo simulation and the alternating directional implicit scheme for partial differential equation with dimension reduction. At the theoretical level, we propose a method for dynamically hedging timer call options and discuss the computation of price sensitivities. As an extension, we consider the valuation of timer options under stochastic volatility with jump models.

1.2 Introduction: Feller Diffusion, Stochastic Variance Clock and Timer Option

The financial market exhibits hectic and calm periods. Prices exhibit large fluctuations when the market is hectic, and price fluctuations tend to be moderate when the market is mild. This uncertain fluctuation is defined as volatility, which has become one of the central features in financial modeling. A variety of volatility (or variance) derivatives, such as variance swaps and options on VIX (the Chicargo board of exchange volatility index), are now actively traded in the financial security markets.

A European call (put) option is a financial contract between two parties, the buyer and the seller of this type of option. It is the option to buy (sell) shares of stock at a specified time in the future for a specified price. The Black-Scholes-Merton (1973) model [10, 79] is a popular mathematical description of financial markets and derivative investment instruments. This model develops partial differential equations whose solution, the Black-Scholes-Merton formula, is widely used in the pricing of European-style options. However, the unrealistic assumption of constant volatility motivates that European options are usually quoted via Black-Scholes implied volatility, which is the volatility extracted from the market, i.e. the volatility implied by the market price of the option based on the Black-Scholes-Merton option pricing model. More explicitly, given the Black-Scholes-Merton model:

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where μ is the return and σ is volatility of the stock, the Black-Scholes-Merton formula for pricing a European call option with maturity T and strike K reads

$$C(\sigma) = BSM(S_0, K, T, \sigma, r) := S_0 N(d_1) - e^{-rT} K N(d_2),$$

where r is the interest rate assumed to be constant,

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)T\right], \\ d_2 &= \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_0}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)T\right], \end{aligned} \tag{1.1}$$

and

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du.$$

Given a market price C_M for the option, the Black-Scholes-Merton implied volatility is the volatility that equates model and market option prices, i.e., it is the value σ^* which solves equation

$$C(\sigma^*) = C_M.$$

Why and how timer options are developed? As reported in RISK [84],

> The price of a vanilla call option is determined by the level of implied volatility quoted in the market, as well as maturity and strike price. But the level of implied volatility is often higher than realised volatility, reflecting the uncertainty of future market direction. In simple terms, buyers of

> vanilla calls often overpay for their options. In fact, having analysed all stocks in the Euro Stoxx 50 index since 2000, SG CIB calculates that 80% of three-month calls that have matured in-the-money were overpriced.

In order to circumvent this problem and ensure that investors pay for the realized variance, Société Générale Corporate and Investment Banking (SG CIB) launched a new type of option (see report in Sawyer [84], Société Générale Asset Management [44] and Hawkins and Krol [59]), called "timer option". With a timer call option, the investor has the right to purchase the underlying asset at a pre-specified strike price at the first time when a pre-specified variance budget is consumed. Instead of fixing the maturity and letting the volatility float, we fix the volatility and let the maturity float. Thus, a timer Option can be viewed as a call option with random maturity. The maturity occurs at the first time the prescribed variance budget is exhausted.

There are several advantages of introducing timer options. According to Société Générale, a timer call option is cheaper than a traditional European call option with the same expected investment horizon, when realized volatility is less than implied volatility. With timer options, systematic market timing is optimized for the following reason. If the volatility increases, the timer call option terminates earlier. However, if the volatility decreases, the timer call option simply takes more time to reach its maturity. Moreover, financial institutions can use timer options to overcome the difficulty of pricing the call and put options whose implied volatility is difficult to quote. This situation usually happens in the markets where the implied volatility data does not exist or is limited. In consideration of applications to portfolio insurance, portfolio managers can use a timer put option on an index (or a well diversified portfolio) to limit their downside risk. They might be interested in hedging specifically against sudden market drops such as the crashes in 1987 and 2008. From the perspective of the financial institutions who offer timer options, if there is a market collapse, the sudden high volatility will cause the timer put options to be exercised rapidly, thus, protecting and hedging the fund's value. By contrast, European put options do not have this feature. With a timer put option, some uncertainty about the portfolio's outcome is represented by uncertainty about the variable time horizon (see Bick (1995) [8] for a similar discussion).

The Feller diffusion (see [41]), also called "square root diffusion", is widely used in mathematical finance due to its favorable properties and analytical tractability. The earliest application of this process in the literature of financial modeling can be found in Cox et al. [25] for the term structure of interest rates. Heston (1993) [61] employed Feller diffusions to model stochastic volatility. As one of the most popular and widely used stochastic volatility models, the variance process $\{V_t\}$ is assumed to follow the stochastic differential equation:

$$dV_t = \kappa(\theta - V_t)dt + \sigma_v \sqrt{V_t} dW_t^{(1)}, \tag{1.2}$$

where $\{W_t^{(1)}\}$ is a standard Brownian motion (in later chapters, we revisit this process). Geman and Yor [52] advocate using a stochastic variance clock, which runs fast if the volatility is high and runs slowly if the volatility is low, to model a non-constant volatility and measure financial time. Mathematically, the stochastic variance clock time can be defined as the first time the total realized variance achieves a certain level $b > 0$, i.e.

$$\tau_b = \inf \left\{ u \geq 0; \ \int_0^u V_s ds = b \right\}. \tag{1.3}$$

The distribution of the variance clock time τ_b plays an important role. Geman and Yor [52] give an explicit formula for this distribution under the Hull and White [62] model for stochastic volatility.

The problems of pricing, hedging and implementation of timer call options under Heston's [61] stochastic volatility model motivates our research in the following chapters. First of all, we formulate rigorously the timer call option valuation problem as a first-passage-time problem, via a standard no-arbitrage argument and the stochastic representation of the solution to a Dirichlet problem. Motivated by this problem, we apply the time-change technique to find that the variance process, which is modeled by Feller diffusion, running on a variance clock, is equivalent in distribution to a Bessel process with constant drift. In other words, we obtain a characterization of the distribution of (τ_b, V_{τ_b}). Further, we derive a joint density on Bessel processes and the integration of its reciprocal via Laplace transform techniques. Applying these results, we obtain a Black-Scholes-Merton type formula for pricing timer options. We also investigate and compare several methods for implementation, including Monte Carlo simulation, the alternating directional implicit scheme for partial differential equations with dimension reduction, and the analytical formula implementation via the Abate-Whitt (1992) algorithm (see [1]) on computing Laplace transform inversion via Fourier series expansion. In the analytical formula implementation, a rotation counting algorithm for correctly evaluating modified Bessel functions with complex argument is applied. We propose a method for dynamically hedging timer call options and discuss the computation of risk management parameters (price sensitivities). As an extension, we tentatively consider the valuation of timer option under stochastic volatility with jump models.

The organization of this part is as follows. In chapter 2, we formulate the perpetual timer call option valuation problem as a first-passage-time problem. In chapter 3, we investigate the connection between the Feller process and Bessel process with constant drift via variance clock time-change, and derive an explicit joint density on Bessel processes which is needed for characterizing the distribution of interest. In chapter 4, a Black-Scholes-Merton type formula for pricing timer options is derived as an application of the results in previous chapters. In chapter 5, various implementation techniques and numerical results are presented. In chapter 6, a dynamic hedging strategy and computation of price sensitivities are both discussed. A tentative consideration of jumps combined with stochastic volatility in the valuation of timer options is proposed in Appendix 10.

Chapter 2

Heston Model, Timer Option and a First-Passage-Time Problem

2.1 Realized Variance and Timer Option

First, we recall the definition of realized variance. Let $[0, T]$ $(T > 0)$ be an investment horizon. Let us define $\Delta t = T/n$ and suppose that the asset price is monitored at $t_i = i\Delta t,$ for $i = 0,\ 1,\ 2,\ ...,\ n$. According to the daily sampling convention, Δt is usually chosen as $1/252$ corresponding to the standard 252 trading days in a year. Let $\{S_t\}$ denote the price process of the underlying stock (or index). The realized variance for the period $[0, T]$ is defined as

$$\widehat{\sigma_T^2} := \frac{1}{(n-1)\Delta t} \sum_{i=0}^{n-1} \left(\log \frac{S_{t_{i+1}}}{S_{t_i}} \right)^2.$$

Next, we introduce the cumulative realized variance over time period $[0, T]$ as

$$RV_T = n\Delta t \cdot \widehat{\sigma_T^2} \approx \sum_{i=0}^{n-1} \left(\log \frac{S_{t_{i+1}}}{S_{t_i}} \right)^2 . \tag{2.1}$$

Upon purchasing a timer call option, the investor specifies a variance budget

$$B = \sigma_0^2 T_0,$$

where T_0 is an expected investment horizon, and σ_0 is the forecasted realized volatility during the investment period. A timer call option pays off $\max(S_{\mathcal{T}} - K, 0)$ at the first time $\mathcal{T}$ when the realized variance exceeds B, i.e. at the time

$$\mathcal{T} := \min \left\{ t_k, \sum_{i=1}^{k} \left(\log \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2 \geqslant B \right\} . \tag{2.2}$$

Similarly, a timer put option with strike K and variance budget B has a payoff $\max(K - S_{\mathcal{T}}, 0)$. Without loss of generality, for the problem of valuation, we focus on timer call options in this dissertation.

According to Hawkins and Krol (2008) [59], timer options are sometimes traded under a finite time-horizon constraint in practice, by slightly modifying the definition for the perpetual case explained in Sawyer (2008) [84]. This perpetuity can be regarded as the limiting case of a long-time horizon constraint. Therefore, even for a timer option with finite horizon, our investigation on the perpetual case may provide helpful information for the analysis. In addition, our study motivates some new research in stochastic analysis.

2.2 Heston's Stochastic Volatility Model

Suppose that the asset $\{S_t\}$ and its instantaneous variance $\{V_t\}$ follow Heston's stochastic volatility model (1993)[61]. In a filtered probability space $(\Omega, \mathbb{P}, \mathcal{G}, \{\mathcal{G}_t\})$, the joint dynamics of $\{S_t\}$ and $\{V_t\}$ are specified as

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t d\mathcal{B}_t^{(2)}, \\ dV_t &= \epsilon(\vartheta - V_t)dt + \sigma_v \sqrt{V_t} d\mathcal{B}_t^{(1)}, \end{aligned} \tag{2.3}$$

where $\{(\mathcal{B}_t^{(1)}, \mathcal{B}_t^{(2)})\}$ is a two-dimensional Brownian motion with instantaneous correlation ρ, i.e.,

$$d\mathcal{B}_t^{(1)} d\mathcal{B}_t^{(2)} = \rho dt.$$

Here μ represents the return of the asset; ϵ is the speed of mean-reversion of $\{V_t\}$; ϑ is the long-term mean-reversion level of $\{V_t\}$; σ_v is a parameter reflecting the volatility of $\{V_t\}$.

2.3 A First-Passage-Time Problem

We assume that the sampling is done continuously. Through quadratic variation calculation in the model of (2.3), it is straightforward to find that

$$\lim_{\Delta t \to 0} \sum_{i=1}^{m} \left(\log \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2 = \int_0^t V_s ds, \quad \text{a.s.,} \tag{2.4}$$

where $t = m\Delta t$ and $t_i = i\Delta t$. Thus, we define

$$I_t := \int_0^t V_s ds$$

as the continuous-time version of the cumulative realized variance over time period $[0, t]$. This continous-time setting motivates the definition of the first passage time:

$$\tau := \inf\left\{u \geq 0; \int_0^u V_s ds = B\right\}. \tag{2.5}$$

Thus, τ is obviously a continuous time approximation of $\mathcal{T}$ as defined in (2.2). In the continuous time setting, a timer call option is regarded as an option which pays off $\max(S_\tau - K, 0)$ at the random maturity time τ. In the following exposition, we represent the no-arbitrage price (in the sense which we explain momentarily) of a timer call option as the so called risk neutral expectation of the discounted payoff. Therefore, the valuation problem becomes a first-passage-time problem for the cumulative realized variance process $\{I_t\}$.

Timer options are traded in the over-the-counter market, where financial instruments such as stocks, bonds, commodities or derivatives are directly traded between parties. The original underlying asset on which timer options are written and the variance swap constitute a complete market. It follows that timer options can be priced according to the no-arbitrage rule as we explain momentarily. In other words, we use variance swaps as auxiliary hedging instruments. A variance swap is an actively traded over-the-counter financial derivative that allows one to speculate on or hedge risks associated with the magnitude of volatility of some underlying product, such as a stock index, exchange rate, or interest rate. One leg of the variance swap pays an amount based on the realized variance of the return of the underlying asset (as defined in (2.1)). The other leg of the swap pays a fixed amount, which is called the strike, quoted at the deal's origination. Thus the net payoff to the counterparties is the difference between the two legs and is settled in cash at the maturity of the swap. Mathematically, the payoff of a variance swap with maturity T and strike K_{var}

is given by $N_{var}(RV_T - K_{var})$, where constant N_{var} is called variance notional which converts the payoff into dollar terms. For the purpose of computation, we reasonably approximate the payoff by the continuous-time version of the realized variance as defined in (2.4). Therefore, we have that

$$N_{var}(RV_T - K_{var}) \approx N_{var}\left(\int_0^T V_s ds - K_{var}\right). \tag{2.6}$$

Demeterfi, et al (1999) [29] initiated the investigation of variance swaps. Broadie and Jain (2008) [16] thoroughly studied the pricing and hedging of variance swaps under Heston's (1993) stochastic volatility model.

To clarify our no-arbitrage pricing mechanism, we briefly go over the notion of the *market price of volatility risk* (also known as *volatility risk premium*) and Heston's (1993) [61] original modeling assumption on its particular functional form. Heston's (1993) [61] stochastic volatility model comes with the assumption that the market price of volatility risk takes a special form as a linear function of volatility $\sqrt{V_t}$. According to Heston (1993), this judicious choice was motivated by the consumption-based models proposed in Breeden (1979) [13] and the term structure model in Cox, et al. (1985) [26]. Though this choice is arbitrary, it becomes a standard for both academic and industrial research. A thorough comparison of the various specifications of the market price of volatility risk and a survey of their empirical estimation can be found in Lee (2001) [73].

Under a slightly broader framework, we recapitulate Heston's (1993) original idea as a necessary part of our current exposition. We follow the the presentation in Gatheral (2007) [48] (see page 5-7). Let us consider an arbitrary derivative security with payoff of the form $\mathcal{P}_1(S, V, I)$, where $\mathcal{P}_1$ is a functional of $\{(S_t, V_t, I_t)\}$. For example, a European call option with maturity T and strike K has payoff $\mathcal{P}_1(S, V, I) = \max(S_T - K, 0)$. Because of the Markov property of $\{(S_t, V_t, I_t)\}$, we assume that the price process of this security $\{C_t\}$ has the form

$$C_t = u(t, S_t, V_t, I_t), \tag{2.7}$$

for some function $u(t, s, v, x) : \ [0, \infty) \times \mathbf{R}_+^3 \to \mathbf{R}$, which is of class $C^{1,2}$. We similarly consider a volatility-dependent security with payoff of the form $\mathcal{P}_2(V, I)$, where $\mathcal{P}_2$ is a functional of $\{(V_t, I_t)\}$. For example, a variance swap with maturity T and strike K_{var} has payoff $\mathcal{P}_2(V, I) = N_{var}\,(I_T - K_{var})$ as we explained in (2.6). Because of the Markov property of $\{(V_t, I_t)\}$, we assume that the price process of this volatility-dependent security $\{F_t\}$ has the form

$$F_t = f(t, V_t, I_t), \tag{2.8}$$

for some function $f(t, v, x) : \ [0, \infty) \times \mathbf{R}_+^2 \to \mathbf{R}$, which is of class $C^{1,2}$.

The Cholesky decomposition on the correlated Brownian motion $\{(\mathcal{B}_t^{(1)}, \mathcal{B}_t^{(2)})\}$ allows us to rewrite Heston's model as

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t (\rho dZ_t^{(1)} + \sqrt{1 - \rho^2} dZ_t^{(2)}), \\ dV_t &= \epsilon(\vartheta - V_t) dt + \sigma_v \sqrt{V_t} dZ_t^{(1)}, \end{aligned} \tag{2.9}$$

where $\{(Z_t^{(1)}, Z_t^{(2)})\}$ is a standard two-dimensional Brownian motion.

Next, we resort to a standard no-arbitrage argument to recast the notions of a market price of volatility risk and risk-neutral; and, we return to the same argument momentarily to formulate the timer option valuation problem. First, we construct a self-financing portfolio with value process $\{P_t\}$ consisting of a share of $C_t = u(t, S_t, V_t, I_t)$, $-\Delta_t^{(1)}$ shares of the underlying asset with value S_t and $-\Delta_t^{(2)}$ shares of the variance swap with value $F_t = f(t, V_t, I_t)$. We also assume that $\{\Delta_t^{(1)}\}$ and $\{\Delta_t^{(2)}\}$ both satisfy technical conditions, such as adaptivity to the filtration $\{\mathcal{G}_t\}$ and integrability. Thus,

$$P_t = C_t - \Delta_t^{(1)} S_t - \Delta_t^{(2)} F_t. \tag{2.10}$$

Based on the self-financing assumption and an a priori assumption that both functions u and f are sufficiently smooth, we apply Ito's formula and collect dt, $dZ_t^{(1)}$ and $dZ_t^{(2)}$ terms to obtain that

$$\begin{aligned}
&dP_t = dC_t - \Delta_t^{(1)} dS_t - \Delta_t^{(2)} dF_t \\
=&\Bigg\{ \left[\frac{\partial u}{\partial t} + \epsilon(\vartheta - V_t)\frac{\partial u}{\partial v} + \mu S_t \frac{\partial u}{\partial s} + V_t \frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 V_t \frac{\partial^2 u}{\partial v^2} + \frac{1}{2} S_t^2 V_t \frac{\partial^2 u}{\partial s^2} + \rho \sigma_v S_t V_t \frac{\partial^2 u}{\partial s \partial v} \right] \\
&- \Delta_t^{(1)} \mu S_t - \Delta_t^{(2)} \left[\frac{\partial f}{\partial t} + \epsilon(\vartheta - V_t)\frac{\partial f}{\partial v} + V_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_v^2 V_t \frac{\partial^2 f}{\partial v^2} \right] \Bigg\} dt \\
&+ \left\{ \rho \sqrt{V_t} S_t \left(\frac{\partial u}{\partial s} - \Delta_t^{(1)} \right) + \sigma_v \sqrt{V_t} \left(\frac{\partial u}{\partial v} - \Delta_t^{(2)} \frac{\partial f}{\partial v} \right) \right\} dZ_t^{(1)} \\
&+ \sqrt{1-\rho^2}\sqrt{V_t} S_t \left(\frac{\partial u}{\partial s} - \Delta_t^{(1)} \right) dZ_t^{(2)}.
\end{aligned} \tag{2.11}$$

To make this portfolio instantaneously risk-free in the sense that the randomness induced by Brownian motion $\{Z_t^{(1)}, Z_t^{(2)}\}$ vanishes, we let

$$\frac{\partial u}{\partial s} - \Delta_t^{(1)} = 0 \tag{2.12}$$

in order to eliminate the $dZ_t^{(2)}$ risk, and let

$$\frac{\partial u}{\partial v} - \Delta_t^{(2)} \frac{\partial f}{\partial v} = 0 \tag{2.13}$$

in order to eliminate the $dZ_t^{(1)}$ risk. In order to rule out arbitrage, the expected return of this portfolio must be equal to the risk free rate r. Otherwise, investors would always be in favor of a risk-free portfolio with a higher deterministic rate of return; thus they can find arbitrage opportunities (see Bjork (1999) [9], p. 93). Thus,

$$\begin{aligned} dP_t = &\left\{ \frac{\partial u}{\partial t} + \epsilon(\vartheta - V_t)\frac{\partial u}{\partial v} + V_t \frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 V_t \frac{\partial^2 u}{\partial v^2} + \frac{1}{2} S_t^2 V_t \frac{\partial^2 u}{\partial s^2} + \rho \sigma_v S_t V_t \frac{\partial^2 u}{\partial s \partial v} \right\} dt \\ &- \Delta_t^{(2)} \left\{ \frac{\partial f}{\partial t} + \epsilon(\vartheta - V_t)\frac{\partial f}{\partial v} + V_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_v^2 V_t \frac{\partial^2 f}{\partial v^2} \right\} dt \\ =& r P_t dt = r(C_t - \Delta_t^{(1)} S_t - \Delta_t^{(2)} F_t) dt. \end{aligned} \tag{2.14}$$

We assume that $\frac{\partial u}{\partial v} \neq 0$ and $\frac{\partial f}{\partial v} \neq 0$. It follows that

$$\begin{aligned} &\frac{\partial u}{\partial t} + \epsilon(\vartheta - v)\frac{\partial u}{\partial v} + v \frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v \frac{\partial^2 u}{\partial v^2} + \frac{1}{2} s^2 v \frac{\partial^2 u}{\partial s^2} + \rho \sigma_v s v \frac{\partial^2 u}{\partial s \partial v} \\ =& \frac{\frac{\partial u}{\partial v}}{\frac{\partial f}{\partial v}} \left\{ \frac{\partial f}{\partial t} + \epsilon(\vartheta - v)\frac{\partial f}{\partial v} + v \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_v^2 v \frac{\partial^2 f}{\partial v^2} \right\} + r \left(u - \frac{\partial u}{\partial s} s - \frac{\frac{\partial u}{\partial v}}{\frac{\partial f}{\partial v}} f \right). \end{aligned} \tag{2.15}$$

Collecting all u-terms to the left-hand side and all f-terms to the right-hand side, we get

$$\frac{\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v \frac{\partial^2 u}{\partial v^2} + \frac{1}{2} s^2 v \frac{\partial^2 u}{\partial s^2} + \rho \sigma_v s v \frac{\partial^2 u}{\partial s \partial v} - ru + rs\frac{\partial u}{\partial s}}{\frac{\partial u}{\partial v}} = \frac{\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_v^2 v \frac{\partial^2 f}{\partial v^2} - rf}{\frac{\partial f}{\partial v}}. \tag{2.16}$$

This equation holds if and only if both sides equal a universal function h of independent variables v and t. We follow the exposition in Gatheral [48] to denote

$$h(t, v) = -[\epsilon(\vartheta - v) - \Lambda(t, v)\sqrt{v}],$$

where the function $\Lambda(t, v)$ is defined as the *market price of volatility risk.*

REMARK 1. Indeed, we may employ a volatility-dependent asset, e.g. variance swap, to illustrate the notion of the market price of volatility risk. Since the variance swap has no exposure to the risk induced by $\{Z_t^{(2)}\}$, we focus on the volatility risk exclusively. The infinitesimal excess growth satisfies that

$$\begin{aligned} dF_t - rF_t dt &= \left\{ \frac{\partial f}{\partial t} + \epsilon(\vartheta - V_t)\frac{\partial f}{\partial v} + V_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_v^2 V_t \frac{\partial^2 f}{\partial v^2} \right\} dt + \frac{\partial f}{\partial v}\sigma_v \sqrt{V_t} dZ_t^{(1)} \\ &= \sqrt{V_t}\frac{\partial f}{\partial v}\left\{ \Lambda(t, V_t)dt + \sigma_v dZ_t^{(1)} \right\}. \end{aligned} \tag{2.17}$$

According to Gatheral (2006) [48], $\Lambda(t, v)dt$ represents the extra return per unit of volatility risk $\sigma_v dZ_t^{(1)}$; and so, in analogy with the Capital Asset Pricing Model, Λ is known as the market price of volatility risk.

According to Heston (1993), Λ could be determined by one volatility dependent asset and then used to price all other securities. Motivated by the consumption-based capital asset pricing models (see Duffie (2001) [32] or Karatzas and Shreve (1998) [69]) proposed in Breeden (1979) [13] and the term structure model proposed in Cox, et al. (1985) [26], Heston's (1993) stochastic volatility model is equipped with a particular functional form of the market price of volatility risk:

$$\Lambda(t, V_t) = \eta\sqrt{V_t}. \tag{2.18}$$

In other words, the market price of volatility risk $\Lambda(t, V_t)$ is assumed to be proportional to volatility $\sqrt{V_t}$. This particular specification allows analytical tractability. This choice becomes standard when the model is used for pricing derivatives. In the following exposition, we denote $\kappa = \epsilon + \eta$ and $\theta = \frac{\epsilon\vartheta}{\varepsilon+\eta}$. Thus

$$h(t, v) = -\kappa(\theta - v).$$

Therefore, the PDE governing the price function of a European call option with payoff $\max\{S_T - K, 0\}$ reads

$$\frac{\partial u}{\partial t} + \kappa(\theta - v)\frac{\partial u}{\partial v} + rs\frac{\partial u}{\partial s} + v\frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 u}{\partial v^2} + \frac{1}{2}s^2 v\frac{\partial^2 u}{\partial s^2} + \rho\sigma_v sv\frac{\partial^2 u}{\partial s\partial v} - ru = 0.$$

By the Feynman-Kac theorem (see Karatzas and Shreve (1991) [68]), the price admits the following representation:

$$C_t = u(t, S_t, V_t) = \mathbb{E}^Q[e^{-r(T-t)}\max(S_T - K, 0)|\mathcal{G}_t]. \tag{2.19}$$

Here $\mathbb{Q}$ is a probability measure under which

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t}S_t(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)}), \quad S_0 = s, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_t^{(1)}, \quad V_0 = v, \end{aligned} \tag{2.20}$$

where $\{(W_t^{(1)}, W_t^{(2)})\}$ is a two-dimensional standard Brownian motion on the filtered probability space $(\Omega, \mathbb{Q}, \mathcal{G}, \{\mathcal{G}_t\})$; r is the instantaneous interest rate assumed to be constant; κ and θ are interpreted as the rate of mean-reversion and long-term reverting level respectively under probability measure $\mathbb{Q}$. According to Heston (1993), $\mathbb{Q}$ is interpreted as a risk-neutral probability measure. Under this measure all derivative securities, including timer options, written on $\{(S_t, V_t, I_t)\}$ are consistently priced based on the no-arbitrage principle.

REMARK 2. The change-of-measure from $\mathbb{P}$ to $\mathbb{Q}$ is specified as

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{G}_t} = \exp\left\{-\int_0^t \Theta_1(s)dZ_s^{(1)} - \int_0^t \Theta_2(s)dZ_s^{(2)} - \frac{1}{2}\int_0^t [\Theta_1(s)^2 + \Theta_1(s)^2]ds\right\}, \tag{2.21}$$

where

$$\Theta_1(t) = \frac{\eta}{\sigma_v}\sqrt{V_t}, \quad \Theta_2(t) = \frac{1}{\sqrt{1-\rho^2}}\left(\frac{\mu - r}{\sqrt{V_t}} - \rho\frac{\eta}{\sigma_v}\sqrt{V_t}\right). \tag{2.22}$$

By Girsanov's theorem,

$$\begin{aligned} W_t^{(1)} &= Z_t^{(1)} + \int_0^t \Theta_1(s)ds, \\ W_t^{(2)} &= Z_t^{(2)} + \int_0^t \Theta_2(s)ds, \end{aligned} \tag{2.23}$$

is a two-dimensional standard Brownian motion under probability measure $\mathbb{Q}$. A more general setting of change-of-measure for stochastic volatility models can be found in Lee (2001) [73].

In this article, since we focus on the no-arbitrage pricing and hedging of timer options under Heston's (1993) model equipped with the particular assumption on the functional form of market price of volatility risk as in (2.18), we do not intend to investigate the statistical measure and the specification of market price of volatility risk. We also assume that the model is calibrated to some bench-marked derivative securities such as European options (see Heston (1993) [61]); and, that variance swaps are priced as in Broadie and Jain (2008) [16]. Therefore, we perform no-arbitrage pricing for timer options, a redundant security, in a complete market where the risk-neutral measure $\mathbb{Q}$ is consistently fixed. An alternative perspective is to regard timer options as fundamental securities rather than path-dependent options in a complete market. Karatzas and Li (2009, 2010) [67] tentatively suggest a super-hedging approach to address the problem within an incomplete market environment. (see Karatzas and Shreve (1998) [69], Karatzas and Cvitanić (1993) [27], Karatzas and Kou (1996) [66], as well as Föllmer and Schied (2004) [43], etc.) This provides opportunities for some future research.

We are now in position to incorporate the random time τ and state the following proposition that formulates the timer option pricing problem as a first-passage-time problem.

PROPOSITION 1. *In the sense of no-arbitrage, the timer call option with strike K and variance budget B can be priced via the risk-neutral expectation of the discounted payoff, i.e.*

$$C_0 = \mathbb{E}^Q[e^{-r\tau}\max(S_\tau - K, 0)], \tag{2.24}$$

where

$$\tau = \inf\left\{u \geq 0, \int_0^u V_s ds = B\right\}.$$

Proof. Let us assume $C_t = u(t, S_t, V_t, I_t)$ introduced in (2.7) to be the price process of a timer call option for any $0 \leq t \leq \tau$ and some function $u(t, s, v, x) : [0, \infty) \times \mathbf{R}_+^3 \to \mathbf{R}$, which is of class $C^{1,2}$; and regard $F_t = f(t, V_t, I_t)$ introduced in (2.8) as the price process of a variance swap with maturity T_1 for some function $f(t, v, x) : [0, \infty) \times \mathbf{R}_+^2 \to \mathbf{R}$, which is of class $C^{1,2}$. Thus, the price of a timer call option at time $t \wedge \tau$ satisfies that

$$C_{t\wedge\tau} = u(t \wedge \tau; S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}).$$

Without loss of generality, we focus on the time interval $[0, \tau \wedge T_1]$. We form a self-financing portfolio with value process $\{P_t\}$ which consists of a share of $C_t = u(t, S_t, V_t, I_t)$, $-\Delta_t^{(1)}$ shares of the underlying asset with value S_t and $-\Delta_t^{(2)}$ shares of the variance swap with value $F_t = f(t, V_t, I_t)$. We also assume that $\{\Delta_t^{(1)}\}$ and $\{\Delta_t^{(2)}\}$ both satisfy the technical conditions, such as adaptivity to the filtration $\{\mathcal{G}_t\}$ and integrability. Thus,

$$P_t = C_t - \Delta_t^{(1)} S_t - \Delta_t^{(2)} F_t. \tag{2.25}$$

Based on Heston's assumption of the particular form of market price of volatility risk (see (2.18)), the same no-arbitrage argument on making the portfolio P_t risk-free (see

(2.25)) yields the PDE governing the timer option pricing function $u(t,s,v,x)$:

$$\frac{\partial u}{\partial t}+\kappa(\theta-v)\frac{\partial u}{\partial v}+rs\frac{\partial u}{\partial s}+v\frac{\partial u}{\partial x}+\frac{1}{2}\sigma_v^2 v\frac{\partial^2 u}{\partial v^2}+\frac{1}{2}s^2 v\frac{\partial^2 u}{\partial s^2}+\rho\sigma_v sv\frac{\partial^2 u}{\partial s\partial v}-ru=0, \quad (2.26)$$

for $(t,s,v,x)\in[0,+\infty)\times[0,+\infty)\times(0,+\infty)\times(0,B]$, with a boundary value condition:

$$u(t,s,v,B)=\max\{s-K\}.$$

The Feynman-Kac theorem (see Shreve (2004) [85] or a stronger version in Karatzas and Shreve (1991) [68]) suggests a candidate solution to PDE (2.26) as follows:

$$u(t\wedge\tau,s,v,x):=\mathbb{E}^Q[e^{-r(\tau-t\wedge\tau)}\max(S_\tau-K,0)|S_{t\wedge\tau}=s,V_{t\wedge\tau}=v,I_{t\wedge\tau}=x], \quad (2.27)$$

where the $\mathbb{Q}$-dynamics of Heston's model follows

$$\begin{aligned} dS_t &= rS_t dt+\sqrt{V_t}S_t(\rho dW_t^{(1)}+\sqrt{1-\rho^2}dW_t^{(2)}), \quad S_0=s;\\ dV_t &= \kappa(\theta-V_t)dt+\sigma_v\sqrt{V_t}dW_t^{(1)}, \quad V_0=v, \end{aligned} \quad (2.28)$$

where $(W_t^{(1)},W_t^{(2)})$ is a two-dimensional standard Brownian motion on the filtered probability space $(\Omega,\mathbb{Q},\mathcal{G},\{\mathcal{G}_t\})$.

Indeed, because the stochastic differential equation governing $\{(S_t,V_t,I_t)\}$ admits a unique weak solution in the sense of probability law, the theory of martingale problem (see Stroock and Varadhan (1969) [87, 88] or section 5.4 of Karatzas and Shreve (1991) [68]) guarantees that the time-homogenous diffusion $\{(S_t,V_t,I_t)\}$ enjoys the strong Markov property. Therefore, we have that

$$u(t\wedge\tau,S_{t\wedge\tau},V_{t\wedge\tau},I_{t\wedge\tau})=\mathbb{E}^Q[e^{-r(\tau-t\wedge\tau)}\max(S_\tau-K,0)|\mathcal{G}_{t\wedge\tau}]. \quad (2.29)$$

We notice that $\{e^{-rt\wedge\tau}u(t\wedge\tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau})\}$ is a martingale adapted to the filtration $\{\widehat{\mathcal{G}}_t\}$, where $\widehat{\mathcal{G}}_t = \mathcal{G}_{t\wedge\tau}$. A straightforward application of Ito's lemma suggests that

$$\begin{aligned}
&e^{-rt\wedge\tau}u(t\wedge\tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) \\
=&u(0, S_0, V_0, I_0) + \int_0^{t\wedge\tau} e^{-r\zeta}\sqrt{V_\zeta}\left(\rho S_\zeta \frac{\partial u}{\partial s} + \sigma_v \frac{\partial u}{\partial v}\right)(\zeta, S_\zeta, V_\zeta, I_\zeta)dW_\zeta^{(1)} \\
&+ \int_0^{t\wedge\tau} \sqrt{1-\rho^2}e^{-r\zeta}\sqrt{V_\zeta}S_\zeta \frac{\partial u}{\partial s}(\zeta, S_\zeta, V_\zeta, I_\zeta)dW_\zeta^{(2)} \\
&+ \int_0^{t\wedge\tau} e^{-r\zeta}\left[\frac{\partial u}{\partial t} + \kappa(\theta - V_\zeta)\frac{\partial u}{\partial v} + rS_\zeta \frac{\partial u}{\partial s} + V_\zeta \frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 V_\zeta \frac{\partial^2 u}{\partial v^2} + \frac{1}{2}S_\zeta^2 V_\zeta \frac{\partial^2 u}{\partial s^2}\right. \\
&\left. + \rho\sigma_v S_\zeta V_\zeta \frac{\partial^2 u}{\partial s \partial v} - ru\right](\zeta, S_\zeta, V_\zeta, I_\zeta)d\zeta
\end{aligned} \tag{2.30}$$

In order to make (2.30) a $\{\widehat{\mathcal{G}}_t\}$-martingale, the Lebesgue integral term must vanish. Therefore, the PDE (2.26) is satisfied. It is also obvious that $u(t, s, v, B) = \max\{s - K, 0\}$.

Thus, on $\{t < \tau\}$, the timer call option price process $\{C_t\}$ satisfies that

$$dC_t = rC_t dt + \sqrt{V_t}\left(\rho S_t \frac{\partial u}{\partial s} + \sigma_v \frac{\partial u}{\partial v}\right)dW_t^{(1)} + \sqrt{1-\rho^2}\frac{\partial u}{\partial s}\sqrt{V_t}S_t dW_t^{(2)}. \tag{2.31}$$

From (2.20), (2.17) and (2.31), we see that $\{e^{-rt\wedge\tau}C_{t\wedge\tau}\}$, $\{e^{-rt\wedge\tau}S_{t\wedge\tau}\}$ and $\{e^{-rt\wedge\tau}F_{t\wedge\tau}\}$ are all $\mathbb{Q}$-martingales. Thus, $\mathbb{Q}$ serves as a risk-neutral probability measure. Therefore, by the fundamental theorem of asset pricing (see Harrison and Pliska [57, 58] or section 5.4 of Shreve (2004) [85]), the market consisting of $\{(S_t, F_t, C_t)\}$ is free of arbitrage. With the initial capital $C_0 = u(0, S_0, V_0, I_0)$, the timer call option can be dynamically replicated via the following strategy (we return to this point in chapter 6):

$$\Delta_t^{(1)} = \frac{\partial u}{\partial s}(t, S_t, V_t, I_t), \quad \Delta_t^{(2)} = \frac{\partial u}{\partial v}\Big/\frac{\partial f}{\partial v}(t, S_t, V_t, I_t).$$

Hence, $u(t \wedge \tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau})$ reasonably prices the timer call option in the sense that the whole market is arbitrage free and the timer call option can be replicated dynamically using a self-financing portfolio.

Let us denote

$$\tau_x := \inf\left\{u \geq 0;\ \int_0^u V_s ds = B - x\right\}. \tag{2.32}$$

By the representation in (2.27) and the time homogeneity property of diffusion $\{S_t, V_t, I_t\}$, we obtain that

$$\begin{aligned} u(t\wedge\tau, s, v, x) =& \mathbb{E}^Q[e^{-r\tau_x}\max(S_{\tau_x+t\wedge\tau} - K, 0)|S_{t\wedge\tau} = s, V_{t\wedge\tau} = v, I_{t\wedge\tau} = x] \\ & \mathbb{E}^Q[e^{-r\tau_x}\max(S_{\tau_x} - K, 0)|S_0 = s, V_0 = v, I_0 = x]. \end{aligned} \tag{2.33}$$

We notice that the right-hand-side in the expression (2.33) is independent of t. Thus, we have that

$$\frac{\partial u}{\partial t} = 0,$$

which results in a Dirichlet problem for $u(s, v, x)$:

$$\begin{aligned} &\kappa(\theta - v)\frac{\partial u}{\partial v} + rs\frac{\partial u}{\partial s} + v\frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 u}{\partial v^2} + \frac{1}{2}s^2 v\frac{\partial^2 u}{\partial s^2} + \rho\sigma_v sv\frac{\partial^2 u}{\partial s\partial v} - ru = 0, \\ &u(s, v, B) = \max\{s - K, 0\}. \end{aligned} \tag{2.34}$$

Therefore, by letting $x = 0$ in (2.33), the initial timer call option price is represented as the risk-neutral expectation of the discounted payoff, i.e.

$$C_0 = \mathbb{E}^Q[e^{-r\tau}\max(S_\tau - K, 0)], \tag{2.35}$$

where

$$\tau = \inf\left\{u \geq 0, \int_0^u V_s ds = B\right\}. \tag{2.36}$$

□

Indeed, we have the following result regarding the uniqueness of the price. Similar to our discussion on timer call options, we denote $w(t, s, v, x)$ the the price function of a timer put option with payoff $\max(K' - S_\tau, 0)$ $(K' > 0)$.

PROPOSITION 2. *The price of a timer call option with payoff* $\max(S_\tau - K, 0)$ *can be uniquely represented by*

$$u(t \wedge \tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) = \mathbb{E}^Q[e^{-r(\tau - t\wedge\tau)} \max(S_\tau - K, 0)|\mathcal{G}_{t\wedge\tau}]; \tag{2.37}$$

the price of a timer put option with payoff $\max(K' - S_\tau, 0)$ *can be uniquely represented by*

$$w(t \wedge \tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) = \mathbb{E}^Q[e^{-r(\tau - t\wedge\tau)} \max(K' - S_\tau, 0)|\mathcal{G}_{t\wedge\tau}]. \tag{2.38}$$

We give a sketch of the proof here. we start from the timer put options whose payoff are bounded. By the same no arbitrage argument, we obtain the PDE governing the timer put option pricing function $w(t, s, v, x)$:

$$\frac{\partial w}{\partial t} + \kappa(\theta - v)\frac{\partial w}{\partial v} + rs\frac{\partial w}{\partial s} + v\frac{\partial w}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 w}{\partial v^2} + \frac{1}{2}s^2 v\frac{\partial^2 w}{\partial s^2} + \rho\sigma_v sv\frac{\partial^2 w}{\partial s \partial v} - rw = 0, \tag{2.39}$$

for $(t, s, v, x) \in [0, +\infty) \times [0, +\infty) \times (0, +\infty) \times (0, B]$, with a boundary value condition:

$$w(t, s, v, B) = \max\{K' - s\}.$$

Because of the time value, the price of the time put option must be bounded by K', i.e.

$$0 < w(t \wedge \tau, s, v, x) \leq K'.$$

Following the same reasoning as before, we obtain a candidate solution to the PDE boundary value problem 2.39:

$$\widehat{w}(t \wedge \tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) = \mathbb{E}^{Q}[e^{-r(\tau - t\wedge\tau)} \max(K' - S_\tau, 0)|\mathcal{G}_{t\wedge\tau}]. \tag{2.40}$$

Because of the upper bound of w, we have that

$$w(t\wedge\tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) \equiv \widehat{w}(t\wedge\tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) = \mathbb{E}^{Q}[e^{-r(\tau - t\wedge\tau)} \max(K' - S_\tau, 0)|\mathcal{G}_{t\wedge\tau}]. \tag{2.41}$$

This uniqueness can be justified using the Theorem 5.7.6 in Karatzas and Shreve (1988) [68]. Similarly, because of

$$\max\{S_\tau - K, 0\} = \max\{K - S_\tau, 0\} + S_\tau - K,$$

the timer call option can be replicated by a combination of a timer put option, the underlying stock and a contract paying fixed value K at τ. Therefore, the no-arbitrage price of the timer call option must admits the representation:

$$u(t \wedge \tau, S_{t\wedge\tau}, V_{t\wedge\tau}, I_{t\wedge\tau}) = \mathbb{E}^{Q}[e^{-r(\tau - t\wedge\tau)} \max(S_\tau - K, 0)|\mathcal{G}_{t\wedge\tau}]. \tag{2.42}$$

REMARK 3. The argument after the PDE (2.26) for proving (2.24) can be alternatively carried out as follows. We recall that the timer options considered in this dissertation are perpetual in the sense that its maturity depends only on the first time when the variance budget is exhausted. By the definition of timer options, given any arbitrary variance budget B, exhausted realized variance I and starting states of S and V, the timer option price function $u(t, s, v, x)$ is essentially independent of the initial time t. In other words, for any $t_1 > t_2 > 0$ and $0 < x < B$, we have that

$$u(t_1, s, v, x) = u(t_2, s, v, x). \tag{2.43}$$

Therefore, we have that

$$\frac{\partial u}{\partial t} = 0,$$

which simplifies the original parabolic PDE (2.26) for pricing timer option to an elliptic equation. Considering a boundary condition on the plane:

$$\Gamma = \{(\xi_1, \xi_2, B), \xi_1 \in \mathbb{R}, \xi_2 \in \mathbb{R}\},$$

we obtain the following Dirichlet problem:

$$\begin{aligned} &\kappa(\theta - v)\frac{\partial u}{\partial v} + rs\frac{\partial u}{\partial s} + v\frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 u}{\partial v^2} + \frac{1}{2}s^2 v\frac{\partial^2 u}{\partial s^2} + \rho\sigma_v sv\frac{\partial^2 u}{\partial s\partial v} - ru = 0, \\ &u(s, v, B) = \max\{s - K, 0\}. \end{aligned} \tag{2.44}$$

For $0 < x < B$, τ_x defined in (2.32) is the first time when the three dimensional diffusion process $\{\Xi_t^x\}$, where $\Xi_t^x = (S_t, V_t, I_t + x)$, exits the domain $\mathfrak{D} = \mathbb{R}^+ \times \mathbb{R}^+ \times [0, B]$, i.e.

$$\tau_x = \inf\{u \geq 0;\ \Xi_u^x \in \mathfrak{D}^c\} = \inf\left\{u \geq 0;\ \int_0^u V_s ds = B - x\right\}.$$

From the relation between the Dirichlet problem and the stochastic differential equations (see section 5.7 of Karatzas and Shreve [68]), we obtain a stochastic representation of the solution to (2.34):

$$u(t, s, v, x) = u(v, s, x) = \mathbb{E}^Q\left[e^{-r\tau_x}\max(S_{\tau_x} - K, 0)\right], \tag{2.45}$$

Hence, the initial timer call option price is represented as the risk-neutral expectation of the discounted payoff, i.e.

$$C_0 = \mathbb{E}^Q[e^{-r\tau} \max(S_\tau - K, 0)], \tag{2.46}$$

where

$$\tau = \inf \left\{ u \geq 0, \int_0^u V_s ds = B \right\}. \tag{2.47}$$

Chapter 3

Feller Diffusion, Bessel Process and Variance Clock

In this chapter, we present a characterization of the joint distribution of variance clock time (1.3) and variance via a Bessel process with constant drift. We also explicitly derive a joint density on Bessel processes. These theoretical results are applied in the problem of timer option valuation. In the following exposition, we make a modeling assumption that the Feller condition $2\kappa\theta - \sigma_v^2 \geq 0$ holds. According to Going-Jaeschke and Yor (1999) [54], zero is an unattainable point for the variance process $\{V_t\}$ under the Feller condition. This can be seen from the Feller's test (see Karatzas and Shreve (1991) [68], section 5.5). If the Heston model is calibrated to the timer option price data using our analytical results, this parameter assumption should be included.

3.1 Connect Feller Diffusion and Bessel Process by Variance Clock

Motivated by the problem of pricing timer options, it is natural to investigate the joint distribution of (V_τ, τ). It turns out that the Feller diffusion running on the variance clock is equivalent in distribution to a Bessel process with constant drift, and that the variance clock time is equivalent in distribution to an integration functional on this Bessel process.

THEOREM 1. *For any $B > 0$, under the risk neutral probability measure $\mathbb{Q}$, we have a distributional identity for the bivariate random variable (V_τ, τ):*

$$(V_\tau, \tau) =^{law} \left(\sigma_v X_B, \int_0^B \frac{ds}{\sigma_v X_s} \right), \tag{3.1}$$

where τ is defined in (2.5) and $\{V_t\}$ is defined in (2.20). Here $\{X_t\}$ is a Bessel process with index $\nu = \frac{\kappa\theta}{\sigma_v^2} - \frac{1}{2}$ (dimension $\delta = \frac{2\kappa\theta}{\sigma_v^2} + 1$) and constant drift $\mu = -\frac{\kappa}{\sigma_v}$, which is governed by SDE:

$$dX_t = \left(\frac{\kappa\theta}{\sigma_v^2 X_t} - \frac{\kappa}{\sigma_v} \right) dt + d\widetilde{\mathcal{B}}_t, \quad X_0 = \frac{V_0}{\sigma_v}, \tag{3.2}$$

where $\{\widetilde{\mathcal{B}}_t\}$ is a standard one dimensional Brownian motion.

REMARK 4. For any $\delta \geq 2$, δ-dimensional Bessel process BES^δ is a diffusion process $\{\mathcal{R}_t\}$ which serves as the unique strong solution to SDE:

$$d\mathcal{R}_t = \frac{\delta - 1}{2\mathcal{R}_t} dt + d\mathcal{W}(t), \quad \mathcal{R}_0 = r \geq 0, \tag{3.3}$$

where $\{\mathcal{W}(t)\}$ is a standard Brownian motion. Alternatively, we denote this Bessel process $BES^{(\nu)}$, where $\nu = \delta/2 - 1$ is defined as its index. For any $\mu \in \mathbf{R}$, we similarly define BES_{μ}^{δ}, the δ-dimensional Bessel process with drift μ, by a diffusion process $\{\mathcal{R}_t^{\mu}\}$ which serves as the unique strong solution to SDE:

$$d\mathcal{R}_t^{\mu} = \left(\frac{\delta - 1}{2\mathcal{R}_t^{\mu}} + \mu\right) dt + d\mathcal{W}(t), \quad \mathcal{R}_0^{\mu} = r' \geq 0. \tag{3.4}$$

Also, we denote this Bessel process with drift $BES_{\mu}^{(\nu)}$, where $\nu = \delta/2 - 1$ is defined as its index. For more detailed studies on Bessel process and Bessel process with drift, readers are referred to Revuz and Yor (1999) [83], Karatzas and Shreve (1991) [68] as well as Linetsky (2004) [74].

Because of our assumption about the Feller condition $2\kappa\theta - \sigma_v^2 \geq 0$ for $\{V_t\}$, the dimension and index of Bessel process with drift satisfies that

$$\nu = \frac{\kappa\theta}{\sigma_v^2} - \frac{1}{2} \geq 0 \text{ and } \delta = \frac{2\kappa\theta}{\sigma_v^2} + 1 \geq 2.$$

According to Linetsky (2004) [74], zero is unattainable for process $\{X_t\}$ in this case.

Proof. Let

$$\tau_t = \inf\left\{u \geq 0, \int_0^u V_s ds = t\right\},$$

For $M_t = \int_0^t \sqrt{V_s} dW_s^{(1)}$, we apply *Dubins-Dambis-Schwarz* theorem of local martingale representation via time changed Brownian motion (see Karatzas and Shreve [68]). We obtain that

$$M(\tau_t) = \int_0^{\tau_t} \sqrt{V_s} dW_s^{(1)} = \mathfrak{B}_t,$$

where $\{\mathfrak{B}_t\}$ is a standard one dimensional Brownian motion. Because $f(u) = \int_0^u V_s ds$ is an increasing C^1 function, it is easy to find that

$$\tau_t = \int_0^t \frac{1}{V_{\tau_s}} ds.$$

Thus

$$V_{\tau_t} = V_0 + \int_0^{\tau_t} \kappa(\theta - V_s)ds + \sigma_v \int_0^{\tau_t} \sqrt{V_s} dW_s^{(1)}.$$

Therefore, we see that

$$V_{\tau_t} = V_0 + \int_0^t \frac{\kappa(\theta - V_{\tau_s})}{V_{\tau_s}} ds + \sigma_v \mathfrak{B}_t.$$

By letting $\mathfrak{X}_t = \frac{V_{\tau_t}}{\sigma_v}$, we have that

$$\mathfrak{X}_t = \frac{V_0}{\sigma_v} + \int_0^t \left(\frac{\kappa\theta}{\sigma_v^2 \mathfrak{X}_u} - \frac{\kappa}{\sigma_v} \right) du + \mathfrak{B}_t. \tag{3.5}$$

Thus,

$$\begin{aligned} V_\tau &= V_{\tau_B} = \sigma_v \mathfrak{X}_B, \\ \tau &= \tau_B = \int_0^B \frac{ds}{V_{\tau_s}} = \int_0^B \frac{ds}{\sigma_v \mathfrak{X}_s}. \end{aligned} \tag{3.6}$$

Hence, the identity (3.1) is justified by the uniqueness of the solution to SDE (3.2). □

3.2 A Joint Density on Bessel Process

The Bessel process with constant drift can be closely related to a standard Bessel process via change-of-measure. In this section, we present and derive two equivalent expressions of a joint density on standard Bessel processes which are applied in the analytical valuation of timer option. Based on the transition density of Bessel processes with drift obtained in Linetsky (2004) [74], we employ the technique of Laplace transform inversion and change of measure to derive the first expression in Theorem 2. Based on a joint density on Bessel with exponential stopping in Borodin and Salminen (2001) [12], we present an alternative expression of our density in Theorem 3 via inverse Laplace transform on the time variable.

3.2.1 The First Expression of the Density

THEOREM 2. *For Bessel process $\{R_t\}$ with index $\nu \geq 0$ and any positive real number B , the joint density*

$$p(x,t)dxdt := \mathbb{P}_0\left(R_B \in dx, \int_0^B \frac{du}{R_u} \in dt\right), \tag{3.7}$$

admits the following analytical representation:

$$p(x,t) = \frac{2}{\pi}\int_0^\infty \cos(t\xi) Re\left\{\phi(-i\xi|x)\right\} d\xi, \tag{3.8}$$

where

$$\phi(\beta|x) = \exp\left\{-\frac{1}{2}\mu_2^2 B + \mu_2(R_0 - x)\right\}\frac{p_{\mu_2}(B;R_0,x)}{p_0(B;R_0,x)}, \tag{3.9}$$

with

$$\mu_2 = \frac{\beta}{(\nu+\frac{1}{2})}.$$

Here, $p_\mu(t;x,y)$ is the transition density of a Bessel process with drift μ and index ν, i.e.

$$\begin{aligned} p_\mu(t;x,y) =& \frac{1}{2\pi}\int_0^{+\infty} e^{-\frac{1}{2}(\mu^2+\rho^2)t}\left(\frac{y}{x}\right)^{\nu+\frac{1}{2}} e^{\mu(y-x)+\pi\frac{\beta}{\rho}} M_{i\frac{\beta}{\rho},\nu}(-2i\rho x) M_{-i\frac{\beta}{\rho},\nu}(2i\rho y) \\ &\cdot \left|\frac{\Gamma\left(\frac{1}{2}+\nu+i\frac{\beta}{\rho}\right)}{\Gamma(1+\nu)}\right|^2 d\rho. \end{aligned} \tag{3.10}$$

And, $p_0(t;x,y)$ is the transition density of the standard Bessel process with the index ν, i.e.

$$p_0(t;x,y) = \frac{1}{t}\left(\frac{y}{x}\right)^\nu y \exp\left\{\frac{-(x^2+y^2)}{2t}\right\} I_\nu\left(\frac{xy}{t}\right), \tag{3.11}$$

where $I_\nu(z)$ is the modified Bessel function of the first kind with index ν defined by

$$I_\nu(z) = \sum_{k=0}^{+\infty} \frac{(\frac{z}{2})^{\nu+k}}{k!\Gamma(\nu+k+1)}.$$

REMARK 5. ***(The Confluent Hypergeometric Functions)***

In Theorem 2, $\Gamma(\cdot)$ is the gamma function and $M_{\chi,\nu}(\cdot)$ is the Whittaker function related to Kummer confluent hypergeometric function (see Buchholz [18]). The Whittaker function can be defined as

$$M_{\chi,\nu}(z) = z^{\nu+\frac{1}{2}} e^{-\frac{z}{2}} M\left(\nu - \chi + \frac{1}{2}, 1+2\nu; z\right),$$

where

$$M(a,b,;z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}$$

is the Kummer confluent hypergeometric function and

$$(a)_n = a(a+1)\cdot \cdots \cdot (a+n-1), (a)_0 = 1$$

is the Pochhammer symbol which is also regarded as rising factorial. When $a = b$, ${}_1F_1(a,b,;z)$ is exactly e^z. As an entire function, it resembles the exponential function on the complex plane.

To prove Theorem (2), we start with an absolute continuity relation between Bessel processes with different constant drifts.

LEMMA 1. (**Absolute Continuity Between Bessel Process with Different Drifts**)

We suppose that the stochastic process $\{\rho_t\}$ follows the law of $BES_{\mu_1}^{(\nu)}$ ($\nu \geq 0$) on a *filtered probability space* $(\Omega, \mathbb{P}_{\mu_1}, \mathcal{F}, \{\mathcal{F}_t\})$. *Under probability measure* $\mathbb{P}_{\mu_2}$ *defined by*

$$d\mathbb{P}_{\mu_2}\Big|_{\mathcal{F}_t} = \exp\left\{(\mu_1 - \mu_2)(\rho_0 - \rho_t) + (\mu_1 - \mu_2)\int_0^t \frac{2\nu+1}{2\rho_s}ds + \frac{1}{2}(\mu_1^2 - \mu_2^2)t\right\} d\mathbb{P}_{\mu_1}\Big|_{\mathcal{F}_t},$$

$\{\rho_t\}$ *follows the law of* $BES^{(\nu)}_{\mu_2}$.

Proof. Let $\{\alpha_t\}$ be a standard Brownian motion on filtered probability space $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\})$. Thus, the unique strong solution of the stochastic integral equation

$$\rho_t = \rho_0 + \int_0^t \frac{2\nu+1}{2\rho_s}ds + \alpha_t$$

follows the law of a $BES^{(\nu)}$. Let $\mathbb{P}_{\mu_1}$ be a new probability measure defined by

$$d\mathbb{P}_{\mu_1}\Big|_{\mathcal{F}_t} = \exp\left\{\mu_1\alpha_t - \frac{1}{2}\mu_1^2 t\right\} d\mathbb{P}\Big|_{\mathcal{F}_t}.$$

By Girsanov's theorem,

$$\{\beta_t^{(1)} = \alpha_t - \mu_1 t\}$$

is a standard Brownian motion under probability measure $\mathbb{P}_{\mu_1}$. Thus, we obtain a Bessel process $\{\rho_t\}$ with index ν and drift μ_1 governed by the stochastic integral equation:

$$\rho_t = \rho_0 + \int_0^t \left(\frac{2\nu+1}{2\rho_s} + \mu_1\right) ds + \beta_t^{(1)}.$$

By algebraic computation, it follows that

$$d\mathbb{P}_{\mu_2}\Big|_{\mathcal{F}_t} = \exp\left\{(\mu_1 - \mu_2)(\rho_0 - \rho_t) + (\mu_1 - \mu_2)\int_0^t \frac{2\nu+1}{2\rho_s}ds + \frac{1}{2}(\mu_1^2 - \mu_2^2)t\right\} d\mathbb{P}_{\mu_1}\Big|_{\mathcal{F}_t},$$

which is equivalent to

$$d\mathbb{P}_{\mu_2}\Big|_{\mathcal{F}_t} = \exp\left\{-(\mu_1-\mu_2)\beta_t^1 - \frac{1}{2}(\mu_1-\mu_2)^2 t\right\} d\mathbb{P}_{\mu_1}\Big|_{\mathcal{F}_t}.$$

Therefore,

$$d\mathbb{P}_{\mu_2}\Big|_{\mathcal{F}_t} = \exp\left\{\mu_2\alpha_t - \frac{1}{2}\mu_2^2 t\right\} d\mathbb{P}\Big|_{\mathcal{F}_t}.$$

Again, by Girsanov's theorem,

$$\{\beta_t^{(2)} = \alpha_t - \mu_2 t\}$$

is a Brownian motion under probability measure $\mathbb{P}_{\mu_2}$. Thus, it follows that

$$\rho_t = \rho_0 + \int_0^t \left(\frac{2\nu+1}{2\rho_s} + \mu_2\right) ds + \beta_t^{(2)}.$$

Therefore, under probability measure $\mathbb{P}_{\mu_2}$, the process $\{\rho_t\}$ follows the law of $BES_{\mu_2}^{(\nu)}$.

□

Next, we derive a Laplace transform of an integral functional of Bessel bridge in order to characterize the conditional distribution of $\int_0^t \frac{du}{R_u}$ given $R_t = x$.

LEMMA 2. (**A Laplace Transform for an Integral Functional of the Bessel Bridge**)

$$\mathbb{E}_{R_0}^{\mu_1}\left[\exp\left\{-\beta\int_0^t \frac{du}{R_u}\right\}\Big| R_t = x\right] = \exp\left\{-\frac{1}{2}(\mu_1^2-\mu_2^2)t - (\mu_1-\mu_2)(R_0 - x)\right\}\frac{p_{\mu_2}(t;R_0,x)}{p_{\mu_1}(t;R_0,x)}, \tag{3.12}$$

where $p_{\mu_i}(t;R_0,x)$ is the transition density of the Bessel process with index $\nu \geq 0$ and *drift* μ_i*, for* $i = 1, 2$*. Here* μ_1 *and* μ_2 *are related by*

$$\mu_2 = \mu_1 + \frac{\beta}{\nu + \frac{1}{2}}, \quad \forall \beta > 0.$$

$\mathbb{E}^{\mu_1}_{R_0}$ *denotes the expectation associated with probability measure* $\mathbb{P}^{\mu_1}_{R_0}$ *under which* $\{R_t\}$ *is a Bessel process with constant drift* μ_1*.*

Proof. By applying Lemma 1 and conditioning, we deduce that

$$p_{\mu_2}(t; R_0, x) = \frac{d}{dy}\left[\mathbb{P}^{\mu_2}_{R_0}(R_t \le y)\right]$$

$$= \frac{d}{dy}\left[\mathbb{E}^{\mu_1}_{R_0} 1\{R_t \le y\} \exp\left\{(\mu_1 - \mu_2)(R_0 - R_t) + (\mu_1 - \mu_2)\int_0^t \frac{2\nu + 1}{2R_s} ds + \frac{1}{2}(\mu_1^2 - \mu_2^2)t\right\}\right]$$

$$= \frac{d}{dy}\Big[\int_0^y \mathbb{E}^{\mu_1}_{R_0}\left(\exp\left\{(\mu_1 - \mu_2)(R_0 - z) + (\mu_1 - \mu_2)\int_0^t \frac{2\nu + 1}{2R_s} ds + \frac{1}{2}(\mu_1^2 - \mu_2^2)t\right\} \Big| R_t = z\right) \mathbb{P}^{\mu_1}_{R_0}(R_t \in dz)\Big]$$

$$= \mathbb{E}^{\mu_1}_{R_0}\left[\exp\left\{(\mu_1 - \mu_2)(R_0 - y) + (\mu_1 - \mu_2)\int_0^t \frac{2\nu + 1}{2R_s} ds + \frac{1}{2}(\mu_1^2 - \mu_2^2)t\right\} \Big| R_t = x\right] p_{\mu_1}(t; R_0, x). \tag{3.13}$$

Let $\beta = -(\mu_1 - \mu_2)\frac{2\nu+1}{2\nu}$, i.e.

$$\mu_2 = \mu_1 + \frac{\beta}{\nu + \frac{1}{2}}, \quad \forall \beta > 0.$$

We obtain the conditional Laplace transform

$$\mathbb{E}^{\mu_1}_{R_0}\left[\exp\left\{-\beta \int_0^t \frac{du}{R_u}\right\} \Big| R_t = x\right] = \exp\left\{-\frac{1}{2}(\mu_1^2 - \mu_2^2)t - (\mu_1 - \mu_2)(R_0 - x)\right\} \frac{p_{\mu_2}(t; R_0, x)}{p_{\mu_1}(t; R_0, x)}. \tag{3.14}$$

$\square$

Based on the knowledge of transition density of Bessel process with constant drift (see Linetsky [74]), we invert the Laplace transform (3.12) to find the density $p(x,t)$. We state a useful lemma as follows.

LEMMA 3. (Inverting Moment Generating Function of A Nonnegative Random Variable)

Suppose Y is a nonnegative random variable with moment generating function $\phi(s) = \mathbb{E}\exp(-sY)$. Its probability density function can be represented by

$$f(x) = \frac{2}{\pi}\int_0^\infty \cos(x\theta) Re\left\{\phi(-i\theta)\right\} d\theta,$$

while its probability cumulative function is

$$F(x) = \frac{2}{\pi}\int_0^\infty \frac{\sin(x\theta)}{\theta} Re\left\{\phi(-i\theta)\right\} d\theta.$$

This Lemma and its proof can be found in Abate and Whitt (1992) [1]. We are now in position to prove Theorem 2.

Proof. We follow the setting and notations in Lemma 2. By letting $\mu_1 = 0$, $t = B$ and denoting the right-hand side of Laplace transform (3.12) $\phi(\beta|x)$, we deduce that

$$\phi(\beta|x) = \exp\left\{-\frac{1}{2}\mu_2^2 B + \mu_2(R_0 - x)\right\}\frac{p_{\mu_2}(B; R_0, x)}{p_0(B; R_0, x)}, \tag{3.15}$$

where

$$\mu_2 = \frac{\beta}{\nu + \frac{1}{2}}, \quad \text{for all } \beta > 0.$$

We obtain the conditional density

$$\mathbb{P}_0\left(\int_0^B \frac{ds}{R_s} \in dt \middle| R_B = x\right) = \frac{2}{\pi}\int_0^\infty \cos(t\xi) Re\left\{\phi(-i\xi|x)\right\} d\xi. \tag{3.16}$$

The joint density is therefore

$$\begin{aligned}
&\mathbb{P}_0\left(\int_0^B \frac{ds}{R_s} \in dt, R_B \in dx\right) \\
=&\mathbb{P}_0\left(\int_0^B \frac{ds}{R_s} \in dt \middle| R_B = x\right) \mathbb{P}_0\left(R_B \in dx\right) \\
=&\frac{2}{\pi}\int_0^\infty \cos(t\xi) Re\left\{\phi(-i\xi|x)\right\} d\xi dx dt.
\end{aligned} \tag{3.17}$$

Combining all the above steps and the knowledge of the transition density of Bessel processes with constant drift (see Linetsky [74]), we justify Theorem 2.

□

3.2.2 The Second Expression of the Density

Based on a joint density on Bessel process with exponential stopping in Borodin and Salminen (2001) [12], we present the second expression in Theorem 3 via inverse Laplace transform on the time variable as follows.

THEOREM 3. *For Bessel process $\{R_t\}$ with index $\nu \geq 0$ and any positive real number B, the joint density*

$$p(x,t)dxdt := \mathbb{P}_0\left(R_B \in dx, \int_0^B \frac{du}{R_u} \in dt\right), \tag{3.18}$$

admits the following analytical representation:

$$p(x,t) = \frac{2e^{\gamma B}}{\pi}\int_0^\infty \cos(By) Re\Bigg\{\Bigg\{\frac{\sqrt{2\lambda}x^{\nu+1}}{X_0^\nu \sinh\left(t\sqrt{\frac{\lambda}{2}}\right)}\exp\left\{-(X_0+x)\sqrt{2\lambda}\coth\left(t\sqrt{\frac{\lambda}{2}}\right)\right\}$$

$$\cdot I_{2\nu}\left(\frac{2\sqrt{2\lambda X_0 x}}{\sinh\left(t\sqrt{\frac{\lambda}{2}}\right)}\right)\Bigg\}\Bigg|_{\lambda=\gamma+iy}\Bigg\}dy, \quad \textit{for any } \gamma > 0. \tag{3.19}$$

We briefly justify this result. First of all, the following result (see Borodin and Salminen [12]) exhibits a joint distribution on Bessel process and the integration functional of its reciprocal stopped at an independent exponential time.

LEMMA 4. *Suppose that $\{X_t\}$ is a Bessel process with index $\nu \geq 0$ and T is an independent exponential time with intensity λ. We have that*

$$\mathbb{P}_0\left(X_T \in dx, \int_0^T \frac{du}{X_u} \in dt\right)$$

$$= \frac{\lambda\sqrt{2\lambda}x^{\nu+1}}{X_0^\nu \sinh\left(t\sqrt{\frac{\lambda}{2}}\right)}\exp\left\{-\frac{(X_0+x)\sqrt{2\lambda}\cosh\left(t\sqrt{\frac{\lambda}{2}}\right)}{\sinh\left(t\sqrt{\frac{\lambda}{2}}\right)}\right\} I_{2\nu}\left(\frac{2\sqrt{2\lambda X_0 x}}{\sinh\left(t\sqrt{\frac{\lambda}{2}}\right)}\right) dxdt. \tag{3.20}$$

We document the proof of this result in Appendix 10. The exponential stopping is equivalent to the Laplace transform on time in the following sense.

$$\mathbb{P}_0\left(X_T \in dx, \int_0^T \frac{du}{X_u} \in dt\right) = \lambda\int_0^{+\infty} e^{-\lambda s}\mathbb{P}_0\left(X_s \in dx, \int_0^s \frac{du}{X_u} \in dt\right) ds \tag{3.21}$$

Thus, we are in position to invert this Laplace transform to obtain the joint density at any fixed time. We need to employ a damping factor to ensure the integrability of the transformed function. According to Abate and Whitt (1992) [1], we spell out the following lemma.

LEMMA 5. *Let* $\phi(s) = \int_0^{+\infty} e^{-sx} f(x)dx$ *denote the single-sided Laplace transform of function* $f(x)$*, the Bromwich integral for inverting Laplace transform satisfies*

$$f(t) = \frac{1}{2\pi i}\int_{\gamma - i\infty}^{\gamma + i\infty} e^{st}\phi(s)ds = \frac{2e^{\gamma t}}{\pi}\int_0^{\infty} Re(\phi(\gamma + iy))\cos(ty)dy,$$

where $\gamma \in \mathbb{R}^+$ *is chosen so as to ensure that* $\phi(s)$ *has no singularities on or to the right of it.*

Hence, the joint density (3.19) follows directly. Thus, the proof of Theorem 3 is complete.

Chapter 4

A Black-Scholes-Merton Type Formula for Pricing Timer Option

As an application of the theoretical results on Feller diffusions and Bessel processes in previous chapters, a Black-Scholes-Merton type formula for pricing timer call option is presented in this chapter.

4.1 A Black-Scholes-Merton Type Formula for Pricing Timer Option

THEOREM 4. *Under Heston's (1993) stochastic volatility model* (2.20)*, the price of a timer call option (represented as* (2.24)*) with strike K and variance budget B admits the following analytical formula:*

$$C_0 = \mathcal{C}(S_0, K, r, \rho, V_0, \kappa, \theta, \sigma_v, B; d_0, d_1, d_2) = S_0\Pi_1 - K\Pi_2, \tag{4.1}$$

where

$$\Pi_i = \int_0^\infty \int_0^\infty \Omega_i\left(\sigma_v x, \frac{t}{\sigma_v}\right) p(x,t)dxdt, \quad \textit{for } i = 1, 2. \tag{4.2}$$

Here,

$$\begin{aligned}
\Omega_1(v,\xi) &= N(d_1(v,\xi)) \exp\left\{d_0(v,\xi) + \frac{\kappa}{\sigma_v^2}(V_0 - v) + \frac{\kappa^2\theta}{\sigma_v^2}\xi - \frac{\kappa^2}{2\sigma_v^2}B\right\}; \\
\Omega_2(v,\xi) &= N(d_2(v,\xi)) \exp\left\{-r\xi + \frac{\kappa}{\sigma_v^2}(V_0 - v) + \frac{\kappa^2\theta}{\sigma_v^2}\xi - \frac{\kappa^2}{2\sigma_v^2}B\right\},
\end{aligned} \tag{4.3}$$

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du,$$

and

$$\begin{aligned}
d_0(v,\xi) &= \frac{\rho}{\sigma_v}(v - V_0 - \kappa\theta\xi + \kappa B) - \frac{1}{2}\rho^2 B, \\
d_1(v,\xi) &= \frac{1}{\sqrt{(1-\rho^2)B}}\left[\log\left(\frac{S_0}{K}\right) + r\xi + \frac{1}{2}B(1-\rho^2) + d_0(v,\xi)\right], \\
d_2(v,\xi) &= \frac{1}{\sqrt{(1-\rho^2)B}}\left[\log\left(\frac{S_0}{K}\right) + r\xi - \frac{1}{2}B(1-\rho^2) + d_0(v,\xi)\right], \\
(d_2(v,\xi) &= d_1(v,\xi) - \sqrt{(1-\rho^2)B}).
\end{aligned} \tag{4.4}$$

Here, $p(x,t)$ is the explicit joint density in Theorem 2 or Theorem 3 with $\nu = \frac{\kappa\theta}{\sigma_v^2} - \frac{1}{2} \geq 0$.

In order to prove Theorem 4, we begin with a conditional Black-Scholes-Merton type formula. By conditioning on the variance path $\{V_t\}$, we obtain the following proposition.

PROPOSITION 3.

$$C_0 = \mathbb{E}^{\mathbb{Q}}[S_0 e^{d_0(V_\tau,\tau)} N(d_1(V_\tau,\tau)) - Ke^{-r\tau} N(d_2(V_\tau,\tau))], \tag{4.5}$$

where

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du,$$

and

$$\begin{aligned} d_0(v, \xi) &= \frac{\rho}{\sigma_v}(v - V_0 - \kappa\theta\xi + \kappa B) - \frac{1}{2}\rho^2 B, \\ d_1(v, \xi) &= \frac{1}{\sqrt{(1-\rho^2)B}} \left[\log\left(\frac{S_0}{K}\right) + r\xi + \frac{1}{2}B(1-\rho^2) + d_0(v,\xi) \right], \\ d_2(v, \xi) &= \frac{1}{\sqrt{(1-\rho^2)B}} \left[\log\left(\frac{S_0}{K}\right) + r\xi - \frac{1}{2}B(1-\rho^2) + d_0(v,\xi) \right]. \end{aligned} \tag{4.6}$$

REMARK 6. When $\rho = 0, \sigma_v = 0, \kappa = 0$, we have only one Brownian motion $\{W_t^{(2)}\}$, which drives the asset process. In this case, the variance $V_t = V_0$ is constant and

$$dS_t = rS_t dt + \sqrt{V_0} S_t dW_t^{(2)}.$$

For $B = V_0 T$, it is easy to see that $\tau = T$. Thus

$$S_\tau = S_T = S_0 \exp\left\{ rT - \frac{1}{2}B + \sqrt{B}Z \right\}.$$

It is obvious that $d_0 = 0$; d_1 and d_2 agree with the Black-Scholes-Merton [10] case, i.e.

$$\begin{aligned} d_1 &= \frac{1}{\sqrt{V_0 T}} \left[\log\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}V_0\right) T \right], \\ d_2 &= \frac{1}{\sqrt{V_0 T}} \left[\log\left(\frac{S_0}{K}\right) + \left(r - \frac{1}{2}V_0\right) T \right]. \end{aligned} \tag{4.7}$$

Therefore, the price of the timer call option with variance budget $B = V_0 T$ coincides with the Black-Scholes-Merton (see Black and Scholes [10] and Merton [79]) price of a call option with maturity T and strike K. That is

$$BSM(S_0, K, T, \sqrt{V_0}, r) = S_0 N(d_1) - Ke^{-rT} N(d_2).$$

We present a proof of Proposition 3 as follows.

Proof. ***(Proof of Proposition 3)***

We begin by representing the system (2.20) as

$$\begin{aligned} S_t &= S_0 \exp\left\{rt - \frac{1}{2}\int_0^t V_s ds + \rho \int_0^t \sqrt{V_s} dW_s^{(1)} + \sqrt{1-\rho^2}\int_0^t \sqrt{V_s} dW_s^{(2)}\right\}, \\ V_t &= V_0 + \kappa\theta t - \kappa \int_0^t V_s ds + \sigma_v \int_0^t \sqrt{V_s} dW_s^{(1)}. \end{aligned} \tag{4.8}$$

Conditioning on V_τ and τ is equivalent to fixing the whole path of the variance process as well as the driving Brownian motion $W^{(1)}$. On the conditioned probability space, within which the variance path $\{V_t\}$ is fixed, we can regard $\sqrt{V_s}$ as a deterministic function. Thus, we obtain, after some straightforward algebraic computations,

$$(S_\tau | \tau = t, V_\tau = v) =^{law} S_0 \exp\left\{N\left(rt - \frac{1}{2}B + \frac{\rho}{\sigma_v}(v - V_0 - \kappa\theta t + \kappa B), (1-\rho^2)B\right)\right\}, \tag{4.9}$$

where $N(\alpha, \beta^2)$ represents a normal variable with mean α and variance β^2. For simplicity, let us denote

$$p = r\tau - \frac{1}{2}B + \frac{\rho}{\sigma_v}(V_\tau - V_0 - \kappa\theta\tau + \kappa B), \quad q = \sqrt{(1-\rho^2)B}. \tag{4.10}$$

Thus,

$$C_0 = \mathbb{E}^Q\left\{\mathbb{E}^Q\left[e^{-r\tau}\max\{S_\tau - K, 0\} | V_\tau, \tau\right]\right\} = \mathbb{E}\left[e^{-r\tau}\max\{S_0 \exp\{p + qZ\} - K, 0\}\right], \tag{4.11}$$

where the previous expectation, and hereafter in this proof, is taken under the conditional probability measure $(\mathbb{Q}|V_\tau, \tau)$. It follows that

$$\begin{aligned}
&\mathbb{E}\left[e^{-r\tau}\max\{S_0\exp\{p+qZ\}-K,0\}\right]\\
=&e^{-r\tau}\mathbb{E}\left[(S_0\exp\{p+qZ\}-K)1\{S_0\exp\{p+qZ\}\geq K\}\right]\\
=&e^{-r\tau}S_0\mathbb{E}\left[\exp\{p+qZ\}1\left\{Z\geq\frac{1}{q}\left(\log\left(\frac{K}{S_0}\right)-p\right)\right\}\right]\\
&-e^{-r\tau}K\mathbb{E}1\left\{Z\geq\frac{1}{q}\left(\log\left(\frac{K}{S_0}\right)-p\right)\right\},
\end{aligned}\tag{4.12}$$

Thus, Proposition 3 follows the straightforward calculation of the above two terms based on the standard normal distribution. □

Let us recall that, for any $B>0$ which represents the variance budget,

$$\tau=\inf\left\{u\geq 0,\int_0^u V_s ds=B\right\}.$$

Theorem 1 states that, under the risk neutral probability measure $\mathbb{Q}$,

$$(V_\tau,\tau)=^{law}\left(\sigma_v X_B,\int_0^B\frac{ds}{\sigma_v X_s}\right),\tag{4.13}$$

where $\{X_t\}$ is a Bessel process with constant drift governed by SDE:

$$dX_t=\left(\frac{\kappa\theta}{\sigma_v^2 X_t}-\frac{\kappa}{\sigma_v}\right)dt+d\widetilde{\mathcal{B}}_t, X_0=\frac{V_0}{\sigma_v}.$$

In the following steps, we change the probability measure to identify a standard Bessel process, which is well studied (see Revuz and Yor (1999) [83]). Then, we apply the explicit joint density in Theorem 2 or Theorem 3 to obtain the analytical pricing formula in Theorem 4.

PROPOSITION 4. *Under the probability measure $\mathbb{P}_0$ where*

$$\frac{d\mathbb{Q}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_t} = \exp\left\{\frac{\kappa}{\sigma_v}\left(\frac{V_0}{\sigma_v} - X_t\right) + \frac{\kappa}{\sigma_v}\int_0^t \frac{\kappa\theta}{\sigma_v^2 X_s}ds - \frac{1}{2}\left(\frac{\kappa}{\sigma_v}\right)^2 t\right\}, \tag{4.14}$$

we have

$$(V_\tau, \tau) =^{law} \left(\sigma_v X_B, \int_0^B \frac{ds}{\sigma_v X_s}\right).$$

Here, $\{X_t\}$ is a standard Bessel process which is governed by SDE:

$$dX_t = \frac{\kappa\theta}{\sigma_v^2 X_t}dt + d\widehat{\mathcal{B}}_t, \quad X_0 = \frac{V_0}{\sigma_v}, \tag{4.15}$$

The timer call option price admits the following representation:

$$C_0 = \mathbb{E}^{P_0}\Psi\left(\sigma_v X_B, \int_0^B \frac{ds}{\sigma_v X_s}\right), \tag{4.16}$$

where

$$\Psi(v, \xi) = S_0\Omega_1(v, \xi) - K\Omega_2(v, \xi). \tag{4.17}$$

Here, functions Ω_1 and Ω_2 are given in Theorem 4.

Proof. Let

$$\widehat{\mathcal{B}}_t = \widetilde{\mathcal{B}}_t - \frac{\kappa}{\sigma_v}t.$$

Under a new probability measure $\mathbb{P}_0$, where

$$\frac{d\mathbb{P}_0}{d\mathbb{Q}}\Big|_{\mathcal{F}_t} = \exp\left\{\frac{\kappa}{\sigma_v}\widetilde{\mathcal{B}}_t - \frac{1}{2}\left(\frac{\kappa}{\sigma_v}\right)^2 t\right\},$$

$\{\widehat{B}_t\}$ is a standard Brownian motion. It obvious that

$$d\mathbb{Q}\Big|_{\mathcal{F}_t} = \exp\left\{\frac{\kappa}{\sigma_v}\left(\frac{V_0}{\sigma_v} - X_t\right) + \frac{\kappa}{\sigma_v}\int_0^t \frac{\kappa\theta}{\sigma_v^2 X_s}ds - \frac{1}{2}\left(\frac{\kappa}{\sigma_v}\right)^2 t\right\} d\mathbb{P}_0\Big|_{\mathcal{F}_t}.$$

Therefore, under measure $\mathbb{P}_0$, $\{X_t\}$ is a standard Bessel process satisfying

$$dX_t = \frac{\kappa\theta}{\sigma_v^2 X_t}dt + d\widehat{\mathcal{B}}_t, X_0 = \frac{V_0}{\sigma_v}.$$

$\square$

Combining with Theorem 2, we complete the proof of Theorem 4.

4.2 Reconcilement with the Black-Scholes-Merton (1973)

An idea similar to timer options can be traced back to Bick (1995) [8], which proposed a quadratic variation based and model-free portfolio insurance strategy to synthesize a put-like protection with payoff $\max\{K'e^{r\tau} - S_\tau, 0\}$ for some $K' > 0$. This strategy avoids the problem of volatility mis-specification in the traditional put-protection approach. Dupire (2005) [39] applies this similar idea to the "business time delta hedging" of volatility derivatives under the assumption that the interest rate is zero. Working under a general semi-martingale framework, Carr and Lee (2009) [21] investigate the hedging of options on realized variance. As an example, Carr and Lee (2009) derive a model-free strategy for replicating a class of claims on asset price when realized variance reaches a barrier. Using the method proposed in Carr and Lee (2009), we are able to price and replicate a payoff in the form of $\max(S_\tau - Ke^{r\tau}, 0)$. It deserves to notice that this payoff coincides with the Societe Generale's timer call options with payoff $\max(S_\tau - K, 0)$ considered in our paper, when the interest rate r is assumed to be zero and there is no finite maturity horizon.

When $r = 0\%$, we simply have that

$$S_t = S_0 \exp\left\{\int_0^t \sqrt{V_u} dW_u^s - \frac{1}{2}\int_0^t V_u du\right\}, \tag{4.18}$$

where

$$W_u^s = \rho W_t^{(1)} + \sqrt{1-\rho^2} W_t^{(2)}. \tag{4.19}$$

Recall that the variance budget is calculated as $B = \sigma_0^2 T_0$, where $[0, T_0]$ is the expected investment horizon and σ_0 is the forecasted annualized realized volatility. Based on the definition of τ in (2.5), we apply the Dubins-Dambis-Schwarz theorem (see Karatzas and Shreve (1991) [68]) to obtain that

$$\int_0^\tau \sqrt{V_u} dW_u^s = \mathcal{W}_B^s, \tag{4.20}$$

where $\{\mathcal{W}_t^s\}$ is a standard Brownian motion. So, we have that

$$S_\tau = S_0 \exp\left\{\mathcal{W}_B^s - \frac{1}{2}B\right\}. \tag{4.21}$$

Therefore, the price of the timer call option with strike K and variance budget $B = \sigma_0^2 T_0$ can be expressed by the Black-Scholes-Merton (1973) formula:

$$C_0 = \mathbb{E}^Q\left[\max\{S_\tau - K, 0\}\right] = BSM(S_0, K, T_0, \sigma_0, 0) = S_0 N(d_1) - KN(d_2), \tag{4.22}$$

where

$$d_1 = \frac{1}{\sqrt{B}} \left[\log\left(\frac{S_0}{K}\right) + \frac{1}{2}B \right],$$
$$d_2 = \frac{1}{\sqrt{B}} \left[\log\left(\frac{S_0}{K}\right) - \frac{1}{2}B \right]. \tag{4.23}$$

It is obvious that (4.22) is a special case of the formula in Theorem 4. To check this, we first recall that the formula in Theorem 4 is equivalent to (4.16), i.e.

$$C_0 = \mathbb{E}^{P_0} \Psi \left(\sigma_v X_B, \int_0^B \frac{ds}{\sigma_v X_s} \right). \tag{4.24}$$

Because of (4.15), we notice that

$$d_0 \left(\sigma_v X_B, \int_0^B \frac{ds}{\sigma_v X_s} \right) = \rho \left(\widehat{\mathcal{B}}_B + \frac{\kappa}{\sigma_v} B \right) - \frac{1}{2}\rho^2 B. \tag{4.25}$$

Similarly, we express d_1, d_2 and

$$\exp\left\{ \frac{\kappa}{\sigma_v} \left(\frac{V_0}{\sigma_v} - X_B \right) + \frac{\kappa}{\sigma_v} \int_0^B \frac{\kappa\theta}{\sigma_v^2 X_s} ds - \frac{1}{2} \left(\frac{\kappa}{\sigma_v} \right)^2 B \right\} = \exp\left\{ -\frac{\kappa}{\sigma_v} \widehat{\mathcal{B}}_B - \frac{1}{2} \left(\frac{\kappa}{\sigma_v} \right)^2 B \right\} \tag{4.26}$$

in terms of the Brownian motion $\widehat{\mathcal{B}}$. Straightforward computation on the normal distributions yields (4.22), which is independent of ρ, for the case of $r = 0\%$.

4.3 Comparison with European Options

According to RISK [84],

> "High implied volatility means call options are often overpriced. In the timer option, the investor only pays the real cost of the call and doesn't suffer from high implied volatility," says Stephane Mattatia, head of the hedge fund engineering team at SG CIB in Paris.

Under the case of $r = 0\%$, we provide a justification of this feature. More precisely, we have the following proposition.

PROPOSITION 5. *Assuming $r = 0\%$, the timer call option with strike K and expected investment horizon T_0 and forecasted realized volatility σ_0, i.e. variance budget $B = \sigma_0^2 T_0$, is less expensive than the European call option with strike K and maturity T_0, when the implied volatility $\sigma_{imp}(K, T_0)$ associated to the strike K and maturity T_0 is higher than the realized volatility σ_0, i.e.*

$$C_0^{European} \geq C_0^{Timer}. \tag{4.27}$$

Indeed, by (4.22) we have that

$$\begin{aligned} C_0^{Timer} =& \mathbb{E}^Q \left[\max\{S_\tau - K, 0\} \right] \\ =& BSM(S_0, K, T_0, \sigma_0, 0) \\ \leq& BSM(S_0, K, T_0, \sigma_{imp}(K, T_0), 0) = C_0^{European}. \end{aligned} \tag{4.28}$$

For the general case of $r > 0\%$, we illustrate the comparison between timer call options and European call options with the same expected investment horizon (maturity) by a numerical example in Chapter 5. We also note that timer options are less expensive than the corresponding American options, which have the different nature of randomness in maturity.

4.4 Timer Options Based Applications and Strategies

The comparison between timer options and European options in Proposition 5 heuristically motivates an option strategy for investors to capture the spread between the realized and implied volatility risk. For example, if an investor believes that the current implied volatility is higher than the subsequent realized volatility over a certain period, she would take a long position in a timer call option and a short position in a European option with the same strike and the expected investment horizon. By setting the variance budget below the implied variance level, she would expect to receive a net profit. In fact, one reason is that the timer call option is less expensive than the European counterpart in this strategy. Another reason is that the lower realized volatility would leverage the asset return, which results in a net profit at the maturity.

Though the timer option payoff ($\max\{K' - S_\tau, 0\}$, for some $K' > 0$) considered in this paper is different from the put protection ($\max\{K'e^{r\tau} - S_\tau, 0\}$) considered in Bick (1995) [8], timer put options may serve as effective tools for portfolio insurance. With a timer put option written on an index (a well diversified portfolio), the uncertainty about the index's outcome is replaced by the variability in time horizon. However, under the assumption of Heston's stochastic volatility model, the distribution of this variable time horizon, under the physical probability measure, can be easily adapted from Theorem 1.

Similar to the comparison in Proposition 5, timer put options are able to offer relatively cheaper cost of portfolio insurance and protection. If the realized variance is low, the timer put options take a long time to mature. Compared to the regularly rolled European put options for protecting the downside risk of a portfolio, the timer put options require less frequency of rollings, resulting in a reduction in the cost for implementing the protection.

4.5 Some Generalizations

We provide several straightforward extensions. Recall that $I_t = \int_0^t V_s ds$ denotes the exhausted variance over time period $[0, t]$. We can modify the above formula to obtain the price during the whole lifetime of the timer call option.

COROLLARY 1. The timer call option price at time $t \wedge \tau$ is

$$\begin{aligned} C_{t\wedge\tau} =& \mathbb{E}^Q\left[e^{-r(\tau - t\wedge\tau)} \max\{S_\tau - K, 0\} \middle| \mathcal{F}_{t\wedge\tau}\right] \\ =& C(S_{t\wedge\tau}, K, r, \rho, V_{t\wedge\tau}, \kappa, \theta, B - I_{t\wedge\tau}; d_0, d_1, d_2). \end{aligned} \tag{4.29}$$

Also, we obtain a put-call parity for timer options:

COROLLARY 2. (Timer Put-Call Parity) Put-Call parity holds for timer call and put options, i.e.

$$C_0 - P_0 = S_0 - K\mathbb{E}^Q e^{-r\tau},$$

where

$$\mathbb{E}^Q e^{-r\tau} = \int_0^{+\infty}\int_0^{+\infty} \exp\left\{\frac{\kappa}{\sigma_v}\left(\frac{V_0}{\sigma_v} - x\right) + \left(\frac{\kappa^2\theta}{\sigma_v^3} - \frac{r}{\sigma_v}\right)t - \frac{1}{2}\left(\frac{\kappa}{\sigma_v}\right)^2 B\right\} p(x,t)dxdt. \tag{4.30}$$

This corollary follows because

$$\max\{S_\tau - K, 0\} - \max\{K - S_\tau, 0\} = S_\tau - K.$$

Suppose the index/stock pays dividend according to a known dividend yield. Under the risk neutral probability measure, the dynamics of the underlying follows that

$$\begin{aligned} dS_t &= S_t[(r-d)dt + \sqrt{V_t}(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)})], \\ dV_t &= \kappa(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_t^{(1)}, \end{aligned} \tag{4.31}$$

where $(W_t^{(1)}, W_t^{(2)})$ is a two dimensional standard Brownian motion; r is the instantaneous interest rate; and d is the dividend yield.

COROLLARY 3. The price of a timer call option on a stock index paying dividend at rate d is

$$C_0^d = C(S_0, K, r, \rho, V_0, \kappa, \theta, \sigma_v, B; d_0', d_1', d_2'),$$

where

$$\begin{aligned} d_0'(v,\xi) &= \frac{\rho}{\sigma_v}(v - V_0 - \kappa\theta\xi + \kappa B) - \frac{1}{2}\rho^2 B - d\xi, \\ d_1'(v,\xi) &= \frac{1}{\sqrt{(1-\rho^2)B}}\left[\log\left(\frac{S_0}{K}\right) + r\xi + \frac{1}{2}B(1-\rho^2) + d_0(v,\xi)\right], \\ d_2'(v,\xi) &= \frac{1}{\sqrt{(1-\rho^2)B}}\left[\log\left(\frac{S_0}{K}\right) + r\xi - \frac{1}{2}B(1-\rho^2) + d_0(v,\xi)\right]. \end{aligned} \tag{4.32}$$

Chapter 5

Implementation and Numerical Examples

In this chapter, we propose the numerical implementation of the analytical formula in Theorem 4, the PDE boundary value problem (5.5), and two benchmarked Monte Carlo simulation schemes. As shown in the following numerical examples, these different methods are roughly competitive; and they produce implementation results which are in agreement. The analytical formula approach is bias-free. However, the challenge arises from the correct valuation of the special function as part of the Fourier transform and from the error control in the Laplace transform inversion procedure. The implementation of the dimension-reduced PDE (5.5) is fast but the error analysis is not transparent because of the non-attainable singularity when $v = 0$. With regard to the Monte Carlo simulation, which serves as a benchmark to verify the computing results from other methods, the efficiency is enhanced when an Euler scheme on Bessel process with predictor and corrector is employed. In the end of this chapter, we illustrate numerically the comparison between timer call options and European call options with the same expected investment horizon (maturity). We also plot a convex profile of the increase of the timer option price versus the raise of the variance budget. The optimal choices of implementation strategy, numerical stability and robustness of the algorithms are open research opportunities for further investigation. This section is limited to the proposal of each numerical method.

5.1 Analytical Implementation via Laplace Transform Inversion

The implementation of the analytical formula in Theorem 4 consists of several steps. To begin with, we map the infinite integration domain to a finite rectangular domain $[0,1] \times [0,1]$ via transform

$$x = -\log(u), \quad t = -\log(z).$$

Then, we implement the two dimensional integration on $[0,1] \times [0,1]$ via trapezoidal rule. The key here is to efficiently evaluate the joint density at each grid point on $[0,1] \times [0,1]$. We implement the Bessel joint density according to the expression in Theorem 3. In this regard, we use Abate and Whitt [1] algorithm about inverting Laplace transforms via Fourier series expansion. When evaluating the modified Bessel function, we perform analytical correction of the Bessel function in order to accomplish the continuity of the Fourier transform (see Broadie and Kaya [17] for a similar discussion). We briefly discuss the necessity of considering analytic continuation of the Bessel function and outline the algorithm.

Let $A(y)$ denote the argument of the Bessel function term in the integrand of (3.19), i.e.

$$A(y) = \frac{2\sqrt{2(\gamma + iy)X_0 x}}{\sinh\left(t\sqrt{\frac{\gamma+iy}{2}}\right)}. \tag{5.1}$$

In graph (5.1(a)), we observe the winding of $A(y)$ through the winding picture of scaled argument

$$\gamma(y) = A(y)\frac{\log\log|A(y)|}{|A(y)|}.$$

We recall that the modified Bessel function of the first kind is defined as

$$I_{2\nu}(z) = \sum_{k=0}^{+\infty} \frac{(\frac{z}{2})^{2(\nu+k)}}{k!\Gamma(2\nu+k+1)},$$

where the power functions are multi-valued. This multi-valued property can be seen from the definition of complex power functions. Because

$$z^\alpha := |z|^\alpha e^{i(arg(z)+2n\pi)\alpha}, \quad \text{for any integer } n,$$

there are different values for z^α when α is not an integer. Therefore, $I_{2\nu}(z)$ is multi-valued when 2ν is not an integer. When the inverse Laplace (Fourier) transform in (3.19) is implemented, the winding of $z = A(y)$ must be captured to ensure the continuity of the transformed function. However, most computation packages automatically map the complex numbers into the principal branch $(-\pi, \pi]$. This fact might cause the discontinuity of the Bessel function when its argument $A(y)$ goes across the negative real line. Therefore, we keep track of the winding number of the argument $A(y)$ by a rotation counting algorithm and employ the following analytical continuation formula (see Abramowitz and Stegun [2], Broadie and Kaya [17])

$$I_\nu(ze^{m\pi i}) = e^{m\nu\pi i}I_\nu(z).$$

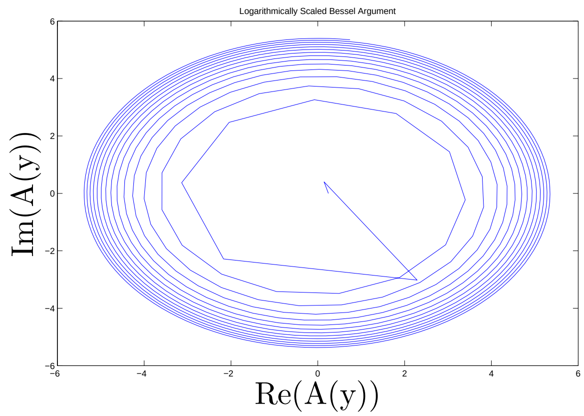
We simply calculate the function on the principal branch and multiply it by a factor $e^{m\nu\pi i}$. Also, we need to take into account the potential numerical overflow in the evaluation procedure. Figure 5.1(b) and Figure 5.1(c) show the effect of rotation counting on the phase angle of the Bessel argument.

We obtain the joint density in (3.19) by inverting the Laplace transform via the Abate and Whitt [1] algorithm. The integration with respect to dy is essentially a Fourier inversion of a damped function. Proper discretization, truncation and choice of damping parameter γ is essential. Trapezoidal rule, though primitive, works well here since the integrand is oscillatory. Errors tend to cancel. We choose step size $h = \frac{\pi}{2B}$ to discretize the integration. This step size, a quarter of a period of the trigonometric function in the integrand, enables us to capture the function oscillation and the tradeoff with computational expenses very well. We choose a significantly large number of steps in order to minimize the truncation error. According to Abate and Whitt [1], we choose a large damping factor γ to control the discretization error. In Figure 5.2, we plot the joint density:

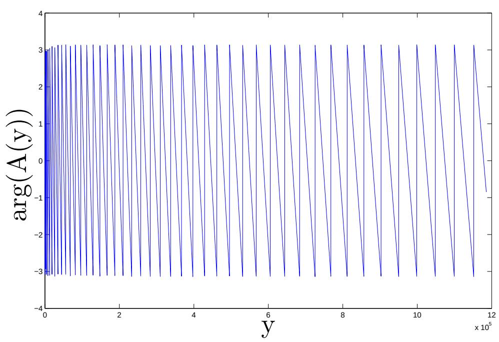
$$\mathbb{P}_0\left(e^{-R_B} \in du, e^{-\int_0^B \frac{du}{R_u}} \in dt\right) = f(x,t)dxdt.$$

Because the integrand decays very fast when t is large, we truncate the total integration domain to enhance the implementation efficiency.

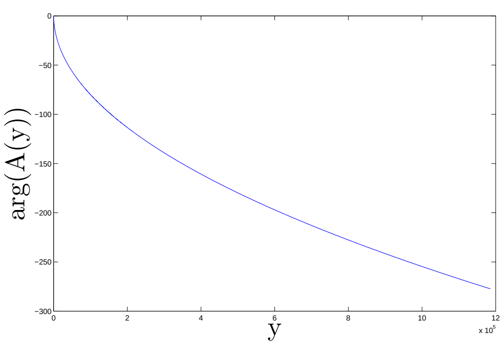
The implementation is performed on a laptop PC with a Intel(R) Pentium(R) M 1.73GHz processor and 1GB of RAM running Windows XP Professional. We code the algorithm using MATLAB. Throughout this section, we use an arbitrary parameter set in Table 5.1 only for the purpose of illustration. In this parameter set, the variance budget B is calculated from $B = \sigma_0^2 T_0$, where the forecasted volatility σ_0 is assumed



(a) Winding of the Bessel Argument



(b) Uncorrected Angle



(c) Corrected Angle

Figure 5.1: Rotation Counting Algorithm for the Bessel Argument

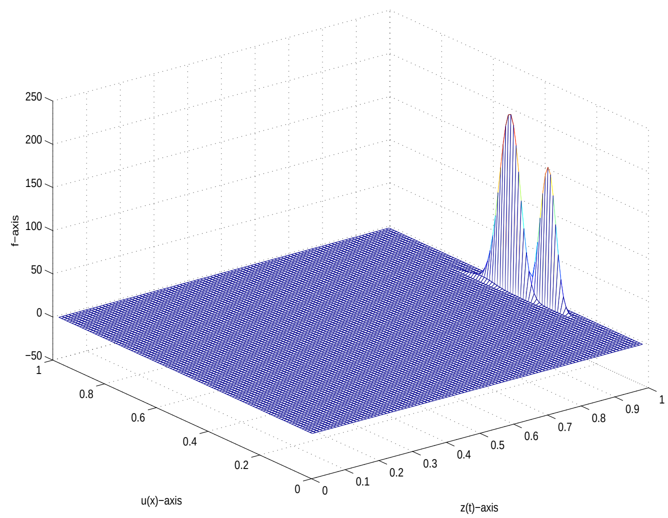


Figure 5.2: The joint density of $\left(e^{-R_B}, e^{-\int_0^B \frac{du}{R_u}}\right)$

to be 23% and the expected investment horizon T_0 is chosen to be 0.5 year. By implementing the analytical formula, we compute the bias-free timer option price as $C_0 = 7.5848$ in 75.6131 CPU seconds. Similarly, we compute the risk-neutral expected maturity as 0.5356 according to

$$\mathbb{E}^Q\tau = \frac{1}{\sigma_v}\int_0^{+\infty}\int_0^{+\infty} t\exp\left\{\frac{\kappa}{\sigma_v}\left(\frac{V_0}{\sigma_v} - x\right) + \frac{\kappa^2\theta}{\sigma_v^3}t - \frac{1}{2}\left(\frac{\kappa}{\sigma_v}\right)^2 B\right\} p(x,t)dxdt. \tag{5.2}$$

Model and Option Parameters	
Stock Price S_0	100
Option Strike K	100
Variance Budget B	0.0265
Correlation ρ	-0.5000
Initial Variance V_0	0.0625
Volatility of the Variance σ_v	0.1000
Rate of Mean Reverting κ	2.0000
Long Run Variance θ	0.0324
Instantaneous Interest Rate r	0.04

Table 5.1: Model and Option Parameters

5.2 ADI Implementation of the PDE with Dimension Reduction

To begin with, we recall that the original PDE governing the timer option pricing function $u(t, s, v, x)$ is

$$\frac{\partial u}{\partial t} + \kappa(\theta - v)\frac{\partial u}{\partial v} + rs\frac{\partial u}{\partial s} + v\frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 u}{\partial v^2} + \frac{1}{2}s^2 v\frac{\partial^2 u}{\partial s^2} + \rho\sigma_v sv\frac{\partial^2 u}{\partial s\partial v} - ru = 0, \quad (5.3)$$

for $(t, s, v, x) \in [0, +\infty) \times [0, +\infty) \times (0, +\infty) \times (0, B]$. Given an arbitrary variance budget, the timer option price $u(t, s, v, x)$ is essentially independent of the initial time t. In other words, timer options are perpetual. Therefore, we have that

$$\frac{\partial u}{\partial t} = 0.$$

This simplifies the original parabolic PDE (2.26) for pricing timer option to an elliptic equation. Considering a boundary condition on plane:

$$\Gamma = \{(\xi_1, \xi_2, B), \xi_1 \in \mathbb{R}, \xi_2 \in \mathbb{R}\},$$

we obtain the following Dirichlet problem:

$$\begin{aligned} &\kappa(\theta - v)\frac{\partial u}{\partial v} + rs\frac{\partial u}{\partial s} + v\frac{\partial u}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 u}{\partial v^2} + \frac{1}{2}s^2 v\frac{\partial^2 u}{\partial s^2} + \rho\sigma_v sv\frac{\partial^2 u}{\partial s\partial v} - ru = 0, \\ &u(s, v, B) = \max\{s - K, 0\}. \end{aligned} \tag{5.4}$$

Subsequently, we rewrite the equation (2.34) by dividing $v(>0)$ on both sides and regard x as a new time variable. After this dimension reduction procedure, a new parabolic PDE initial value problem can be formulated as follows.

PROPOSITION 6. *The timer call option price is governed by the following two dimensional parabolic initial value problem*

$$\begin{cases} \dfrac{\partial u}{\partial x} + \dfrac{1}{2}\sigma_v^2\dfrac{\partial^2 u}{\partial v^2} + \dfrac{1}{2}s^2\dfrac{\partial^2 u}{\partial s^2} + \rho\sigma_v s\dfrac{\partial^2 u}{\partial s\partial v} + \kappa\left(\dfrac{\theta}{v} - 1\right)\dfrac{\partial u}{\partial v} + \dfrac{rs}{v}\dfrac{\partial u}{\partial s} - \dfrac{r}{v}u = 0, \\ \\ u(s, v, B) = \max(s - K, 0), \end{cases} \tag{5.5}$$

where x serves as a temporal variable which corresponds to the stochastic total variance clock (see (1.3)*).*

We briefly discuss the implementation of the PDE (5.5). As the following algorithm and numerical result shows, the PDE implementation is fast because we regard the total variance as a clock through dimension reduction. The variance clock introduces a small but sensitive scale. The error control depends heavily on the specification of boundary conditions, computation domain, time steps and spatial discretization.

First, we set up the following boundary conditions of either Dirichlet, Neuman or Robin type:

$$
\begin{aligned}
&u(0,v,x)=0,\\
&\lim_{s\to+\infty}\frac{\partial u}{\partial s}(s,v,x)=1,\\
&\kappa\theta\frac{\partial u}{\partial v}(s,0,x)+rs\frac{\partial u}{\partial s}(s,0,x)-ru(s,0,x)=0,\\
&\lim_{v\to+\infty}u(s,v,x)=\max(s-K,0).
\end{aligned}
\tag{5.6}
$$

The first two conditions are straightforward. The third boundary condition on $v=0$ is obtained from plugging in zero to the original PDE (2.26) for pricing timer option. This is based on an a priori assumption that u is sufficiently smooth around $v=0$. The last boundary condition reconciles the fact that large initial volatility causes immediate exercise of the timer call.

We discretize the domain $[0,S]\times[0,V]$ such that there are $I+1$ nodes in s-direction and $J+1$ nodes in v-direction. i.e.

$$
S=I\Delta s,\quad V=J\Delta v.
$$

For all interior points $(1\leqslant i\leqslant I,\quad 1\leqslant j\leqslant J)$, we propose a $\lambda-$ scheme:

$$
(\mathbf{1}-\lambda\mathcal{A}_1-\lambda\mathcal{A}_2)U^{n+1}=[\mathbf{1}+\mathcal{A}_0+(\mathbf{1}-\lambda)\mathcal{A}_1+(1-\lambda)\mathcal{A}_2]U^n,\tag{5.7}
$$

where we categorize the derivatives such that $\mathcal{A}_0$ corresponds to the mixed derivative term, $\mathcal{A}_1$ corresponds to the spatial derivatives in the s direction and $\mathcal{A}_2$ corresponds to the spatial derivatives in the v direction. The ru term can be divided into $\mathcal{A}_1$ and $\mathcal{A}_2$. Therefore,

$$\begin{aligned}
\mathcal{A}_0 =& \frac{1}{4}\rho\sigma_v s R_{sv}\delta_{sv}, \\
\mathcal{A}_1 =& \frac{1}{2}s^2 R_{s2}\delta_{ss} + \frac{1}{2}\frac{rs}{v}R_s\delta_s - \frac{1}{2}\frac{r}{v}\Delta x, \\
\mathcal{A}_2 =& \frac{1}{2}\sigma_v^2 R_{v2}\delta_{vv} + \frac{1}{2}\kappa\left(\frac{\theta}{v} - 1\right)R_v\delta_v - \frac{1}{2}\frac{r}{v}\Delta x,
\end{aligned} \tag{5.8}$$

where

$$R_s = \frac{\Delta x}{\Delta s}, \quad R_v = \frac{\Delta x}{\Delta v}, \quad R_{s2} = \frac{\Delta x}{\Delta s^2}, \quad R_{v2} = \frac{\Delta x}{\Delta v^2}, \quad R_{sv} = \frac{\Delta x}{\Delta s \Delta v}.$$

The difference operators are defined as

$$\begin{aligned}
\delta_s U_{i,j} &= U_{i+1,j} - U_{i-1,j}, \\
\delta_{ss} U_{i,j} &= U_{i+1,j} - 2U_{i,j} + U_{i-1,j}, \\
\delta_v U_{i,j} &= U_{i,j+1} - U_{i,j-1}, \\
\delta_{vv} U_{i,j} &= U_{i,j+1} - 2U_{i,j} + U_{i,j-1}, \\
\delta_{sv} U_{i,j} &= U_{i+1,j+1} + U_{i-1,j-1} - U_{i-1,j+1} - U_{i+1,j-1},
\end{aligned} \tag{5.9}$$

By adding $\lambda^2\mathcal{A}_1\mathcal{A}_2U^{n+1}$ on both sides of (5.7) and ignoring higher order term, we design a variation of the previous scheme as

$$(\mathbf{1} - \lambda\mathcal{A}_1 - \lambda\mathcal{A}_2 + \lambda^2\mathcal{A}_1\mathcal{A}_2)U^{n+1} = [\mathbf{1} + \mathcal{A}_0 + (\mathbf{1} - \lambda)\mathcal{A}_1 + (1 - \lambda)\mathcal{A}_2 + \lambda^2\mathcal{A}_1\mathcal{A}_2]U^n. \tag{5.10}$$

Factorizing the difference operator on the left hand side and regrouping the terms, we obtain that

$$(1 - \lambda\mathcal{A}_1)(1 - \lambda\mathcal{A}_2)U^{n+1} = [\mathbf{1} + \mathcal{A}_0 + (\mathbf{1} - \lambda)\mathcal{A}_1 + \mathcal{A}_2]U^n - (1 - \lambda\mathcal{A}_1)\lambda\mathcal{A}_2U^n. \tag{5.11}$$

No.t-Steps N_t	No.V-Steps N_v	No. S-Steps N_s	Computed Price	CPU (Seconds)
10	40	80	8.2921	1.4319
20	80	160	7.3756	2.7091
40	160	320	7.4634	18.3646
80	320	640	7.5852	138.8510
160	640	1280	7.5849	879.6288

Table 5.2: A numerical example on pricing timer call option via ADI scheme

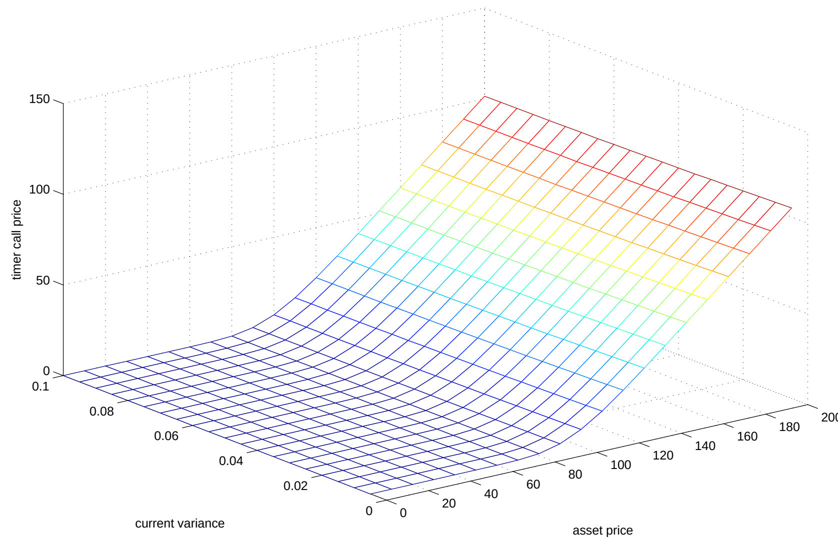
We propose an ADI scheme using Douglas-Rachford [31] method as follows:

$$\begin{aligned} (1-\lambda\mathcal{A}_1)U^{n+\frac{1}{2}} &= [\mathbf{1}+\mathcal{A}_0+(\mathbf{1}-\lambda)\mathcal{A}_1+\mathcal{A}_2]U^n, \\ (1-\lambda\mathcal{A}_2)U^{n+1} &= U^{n+\frac{1}{2}}-\lambda\mathcal{A}_2U^n \end{aligned} \tag{5.12}$$

Detailed formulation of this ADI scheme is documented in Appendix 10. We implement this algorithm by choosing S and V sufficiently large (for example, let $S = 800$, $V = 5$). Table 5.2 exhibits the numerical experiment using the parameter set in Table 5.1. As numbers of steps along different direction (N_t, N_s, N_v) increase proportionally, the values converge to a level closed to the bias-free value obtained from the formula implementation in the previous section. We observe that the complexity of this ADI scheme is approximately

$$\mathcal{O}(N_t)\times(\mathcal{O}(N_s)\times\mathcal{O}(N_v)+\mathcal{O}(N_v)\times\mathcal{O}(N_s)) = \mathcal{O}(N_t)\times\mathcal{O}(N_s)\times\mathcal{O}(N_v).$$

Figure 5.3 exhibits the surface of the timer call price. We note that it might not be optimal to have proportional step sizes and the choice of the proportional factor is a topic of further research. In this numerical experiment, we focus on the illustration of the validity of our AID scheme.



(a) Timer call option price Surface

Figure 5.3: Timer call surfaces implemented from PDE ADI scheme

5.3 Monte Carlo Simulation

In this section, we propose two Monte Carlo simulation schemes and compare their implementation with the PDE and the analytical formula. Algorithm 1 is a relatively brute force method. Compared to using the discounted payoff $e^{-r\tau}\max\{S_\tau - K, 0\}$ as the estimator directly, Proposition 3 offers a conditional Monte Carlo simulation estimator, which achieves the variance reduction. To further enhance the efficiency, we employ the discounted asset price $e^{-r\tau}S_\tau$ as a control variate. (τ, V_τ) is approximated via a "time-checking" algorithm based on the exact simulation of the variance path $\{V_t\}$. We notice that Algorithm 1 also works, when the Feller condition $2\kappa\theta - \sigma_v^2 \geq 0$ is violated. As we can see in Appendix 10, Algorithm 1 can be adapted to incorporate the consideration of jump risk. Algorithm 2 is an enhanced implementation based on Proposition 4. Instead of manipulating the original variance process $\{V_t\}$, we employ an Euler scheme on the Bessel process with predictor and corrector, which is applied to improve accuracy by averaging current and future levels of the drift coefficient. Also, as an alternative to Algorithm 2, an almost-exact simulation scheme based on the transition law of the underlying Bessel process is proposed in Algorithm 3. We briefly describe each algorithm as follows and exhibit the computing performance of Algorithm 1 and Algorithm 2.

Algorithm 1. For each replication, we perform the following steps:

1. Simulate the discretized variance process $\{V_t\}$ path exactly, according to its transition law, along the time grids $t_i = i\Delta t$ where i=1,2,3,...(see Remark 7 for details about sampling the square root process).

2. Evaluate the total variance $\int_0^t V_s ds$ by trapezoidal rule, i.e.

$$\int_0^{j\Delta t} V_s ds \approx \Delta t \left[\frac{V_0 + V(j\Delta t)}{2} + \sum_{k=1}^{j-1} V(k\Delta t) \right], \tag{5.13}$$

   and, check the first time when the variance budget is exhausted. We record the first j (denoted by j_{min}) when

$$\Delta t \left[\frac{V_0 + V(j\Delta t)}{2} + \sum_{k=1}^{j-1} V(k\Delta t) \right] \geq B. \tag{5.14}$$

   We regard $(V(j_{min}\Delta t), j_{min}\Delta t)$ as an approximation of (V_τ, τ).

3. Given τ, we sample S_τ via Euler scheme (see Glasserman (2004) [53]).

4. We employ the discounted asset price $e^{-r\tau}S_\tau$ as a control variate. The estimator is spelt out as

$$\widetilde{C_\tau} = S_0 e^{d_0(V_\tau,\tau)} N(d_1(V_\tau,\tau)) - Ke^{-r\tau} N(d_2(V_\tau,\tau)) + \beta^*(e^{-r\tau}S_\tau - S_0), \tag{5.15}$$

where the optimal coefficient β^* is the opposite number of the slope of the least-square regression line of the samples of $S_0 e^{d_0(V_\tau,\tau)}N(d_1(V_\tau,\tau))-Ke^{-r\tau}N(d_2(V_\tau,\tau))$ against that of $e^{-r\tau}S_\tau$. We use the sample version to estimate β^* in the implementation, i.e. use the ratio of the covariance of the two statistics and the variance of the control variable.

REMARK 7. By Cox et al. [25], the distribution of V_t given V_u where $u < t$, up to a scale factor, is a noncentral chi-squared distribution, i.e.

$$V_t = \frac{\sigma_v^2(1-e^{-\kappa(t-u)})}{4\kappa}\chi_d'^2\left(\frac{4\kappa e^{-\kappa(t-u)}}{\sigma_v^2(1-e^{-\kappa(t-u)})}V_u\right), \quad \text{for } t > u,$$

where $\chi_d'^2$ denotes the noncentral chi-squared random variable with d degrees of freedom, non-centrality parameter λ and $d = \frac{4\theta\kappa}{\sigma_v^2}$. See Broadie and Kaya [17] for detailed study about the exact simulation of process $\{V_t\}$.

Algorithm 2 relies on the representation of timer call option price via Bessel processes in Proposition 4. Here we propose an Euler scheme with Predictor-Corrector on the Bessel SDE. We apply the Brownian scaling property to derive the following property, in order to recast the variance budget B as a model parameter.

COROLLARY 4. Let $\{Y_t, Z_t\}$ be a bivariate diffusion process governed by stochastic differential equation:

$$\begin{aligned} dY_t &= B\kappa\cdot\frac{\theta}{Y_t}dt + \sigma_v\sqrt{B}d\beta_t;\ Y_0 = V_0, \\ dZ_t &= \frac{B}{Y_t}dt;\ Z_0 = 0, \end{aligned} \tag{5.16}$$

where $\{\beta_t\}$ is a standard Brownian motion. Under probability measure $\mathbb{P}_0$, the following distributional identity holds,

$$\left(\sigma_v X_B, \int_0^B \frac{ds}{\sigma_v X_s}\right) \equiv^{law} (Y_1, Z_1).$$

Thus, the algorithm can be described as follows.

Algorithm 2. For each replication, we perform an Euler iteration with predictor and corrector (see Jaeckel [63]) in order to compensate the error caused by the unattainable zero pole of Bessel processes. Numerical experiments shows that this algorithm works very well.

1. Predictor:

$$\hat{Y}(i+1) = \hat{Y}(i) + \frac{B\kappa\theta}{\hat{Y}(i)}\Delta t + \sigma_v\sqrt{B}\sqrt{\Delta t}\beta(i), \hat{Y}(0) = V_0,$$

2. Corrector:

$$Y(i+1) = Y(i) + \left[\frac{\alpha B\kappa\theta}{\hat{Y}(i+1)} + \frac{(1-\alpha)B\kappa\theta}{\hat{Y}(i)}\right]\Delta t + \sigma_v\sqrt{B}\sqrt{\Delta t}\beta(i),$$

   $Y(0) = V_0$, where $\beta(i)$'s are standard normal random variables; α is an adjusting coefficient which is usually chosen as $\frac{1}{2}$.

3. $$Z(i+1) = Z(i) + \frac{1}{2}\Delta t B\left[\frac{1}{Y(i)} + \frac{1}{Y(i+1)}\right], Z(0) = 0.$$

According to Proposition 4, the estimator is $\Psi(Y_1, Z_1)$.

As an alternative to Algorithm 2, an almost-exact simulation based on the transition law of the underlying Bessel process is proposed as follows. We start with the following lemma about the transition law of Bessel processes (see Delbaen and Shirakawa [28]).

LEMMA 6. (Delbaen and Shirakawa, Law of Squared Bessel Process) *For any $\delta > 0$, we have, for δ dimensional Squared Bessel process X_t^2,*

$$X_t^2 =^{law} t \cdot \chi_\delta'^2 \left(\frac{X_0^2}{t} \right),$$

where $\chi_\delta'^2 \left(\frac{X_0^2}{t} \right)$ means a noncentral chi-squared random variable with degree of freedom δ and non-centrality parameter $\frac{X_0^2}{t}$.

We note that the noncentral chi-square distribution $\chi_\delta'^2(\alpha)$ with degree of freedom δ and non-centrality parameter $\alpha > 0$ has the density

$$f_{\chi^2}(x; \delta, \alpha) = \frac{1}{2} e^{-\frac{(\alpha+x)}{2}} \left(\frac{x}{\alpha} \right)^{\frac{\nu}{2}} I_\nu(\sqrt{x\alpha}) 1\{x > 0\},$$

where $\nu = \frac{\delta}{2} - 1$. The Bessel transition density can be expressed in terms of the noncentral chi-square density as

$$p^{(\nu)}(t; x, y) = \left(\frac{2y}{t} \right) f_{\chi^2} \left(\frac{y^2}{t}; \delta, \frac{x^2}{t} \right).$$

Now, we briefly describe an almost-exact simulation algorithm.

Algorithm 3. (An Alternative Simulation Scheme in Algorithm 2) Let the length of time grid be $\Delta t = \frac{B}{m}$, We simulate exactly the whole path of X_t, i.e.

$$X_0, X_{\Delta t}, X_{2\Delta t}, \ \ldots, \ X_B,$$

by sampling noncentral chi-squared distribution iteratively according to

$$X_{i\Delta t} =^{law} \sqrt{\Delta t \cdot \chi'^2_{\frac{2\kappa\theta}{\sigma_v^2}+1}\left(\frac{X^2_{(i-1)\Delta t}}{\Delta t}\right)}. \tag{5.17}$$

The noncentral chi-squared random variable has degree of freedom $d = \frac{2\kappa\theta}{\sigma_v^2} + 1 > 1$. According to Johnson et al. [65],

$$\chi'^2_d(\lambda) = (Z + \sqrt{\lambda})^2 + \chi^2_{d-1}.$$

The chi-squared distribution has CDF

$$\mathbb{P}(\chi^2 \leqslant y) = \frac{1}{2^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)} \int_0^y e^{-\frac{z}{2}} z^{\frac{d}{2}-1} dz,$$

and the Gamma distribution $Gamma(a, \beta)$ with shape parameter a and scale parameter β has PDF

$$f_{a,\beta}(y) = \frac{1}{\Gamma(a)\beta^a} y^{a-1} e^{-\frac{y}{\beta}}, \forall y \geqslant 0.$$

Therefore, we find that

$$\chi^2_d = Gamma(\frac{d}{2}, 2),$$

by letting $a = \frac{d}{2}, \beta = 2$. The iteration equation (5.17) becomes

$$X_{i\Delta t} = \sqrt{\Delta t \cdot \left(\left(Z_i + \sqrt{\frac{X^2_{(i-1)\Delta t}}{\Delta t}}\right)^2 + \chi^2_{d-1}\right)}, X_0 = \frac{V_0}{\sigma_v}, \tag{5.18}$$

where $d = \frac{2\kappa\theta}{\sigma_v^2} + 1$ and

$$\chi^2_{d-1} = \chi^2_{\frac{2\kappa\theta}{\sigma_v^2}} = Gamma\left(\frac{\kappa\theta}{\sigma_v^2}, 2\right) = 2Gamma\left(\frac{\kappa\theta}{\sigma_v^2}, 1\right),$$

where the shape parameter of the Gamma variables is $\frac{\kappa\theta}{\sigma_v^2}$. To sample Gamma variables, we use Fishman's algorithm (see Fishman [42]).

Thus, we can sample V_τ through $\sigma_v X_B = \sigma_v X_{m\Delta t}$. We use trapezoidal rule to approximate τ as $\frac{1}{X_t}$

$$\tau = \int_0^B \frac{ds}{\sigma_v X(s)} \approx \frac{1}{2}\sum_{i=1}^m \frac{1}{\sigma_v}\left(\frac{1}{X_{\frac{i-1}{m}B}} + \frac{1}{X_{\frac{i}{m}B}}\right)\frac{B}{m}.$$

According to Duffie and Glynn [34], it is asymptotically optimal to increase the number of time steps proportional to the square root of the number of simulation trials for the first order Euler discretization. For Algorithm 1, we suppose the number of simulation trials is n and number of time steps for going from 0 to T_0 (the specified investment horizon) is k. Thus we apply the optimality principle in Duffie and Glynn [34] in the sense that $k = [\sqrt{n}]$ ($\Delta t = T_0/k$). Table 5.3 and Figure 5.4 shows our implementation results. It is obvious that Algorithm 2 significantly outperforms Algorithm 1.

5.4 Miscellaneous Features

Finally, we illustrate the comparison between timer call options and European call options with the same expected investment horizon (maturity) by a numerical example. Indeed, we specify an investment horizon, e.g. $T_0 = 0.25$. Using the model

(a) Simulation with Algorithm 1

No.Trials n	No.Steps k	Bias	Standard Error	RMS Error	Computing Time (CPU sec.)
10,000	100	0.1191	0.0617	0.1341	0.734
40,000	200	0.0343	0.0312	0.0464	5.718
160,000	400	0.0303	0.0156	0.0341	44.719
640,000	800	0.0250	0.0078	0.0262	354.719
2,560,000	1,600	0.0129	0.0070	0.0147	2863.83
10,240,000	3,200	0.0058	0.0033	0.0067	25596.63

(b) Simulation with Algorithm 2

No.Trials n	No.Steps k	Bias	Standard Error	RMS Error	Computing Time (CPU sec.)
10,000	100	0.0421	0.0521	0.0670	0.46
40,000	200	0.0149	0.0256	0.0299	1.95
160,000	400	0.0088	0.0129	0.0156	15.38
640,000	800	0.0073	0.0065	0.0098	118.12
2,560,000	1,600	0.0059	0.0032	0.0067	978.53
10,240,000	3,200	0.0049	0.0016	0.0051	9926.36

Table 5.3: We implement Algorithm 1 and Algorithm 2 using the parameters in Table 5.1. The bias-free timer option value computed from the analytical formula is $C_0 = 7.5848$. It turns out that Algorithm 2, employing a predictor-and-corrector correction with Bessel process simulation, performs much better than brute force Algorithm 1. We also note that the simulated values of the risk-neutral expectation of the maturity from Algorithm 1 and Algorithm 2 associated to the lowest RMS errors obtained in this numerical experiment are 0.5358 and 0.5355, respectively.

parameters in Table 9.1 and varying the value of S_0, we calculate the expected realized variance over period $[0, T_0]$ as

$$\sigma_0^2 = \frac{1}{T_0}\mathbb{E}^Q\left(\int_0^{T_0} V_s ds\right) = \theta + \frac{V_0 - \theta}{\kappa T_0}\left(1 - e^{-\kappa T_0}\right).$$

We observe, from Figure 5.5, that the timer call option with variance budget $B = \sigma_0^2 T_0$ is less expensive than the European call option with maturity T_0. This numerical experiment suggests that, as the implied volatility tends to be mostly higher than the realized volatility, timer options can be applied to suppress the extra cost of the high

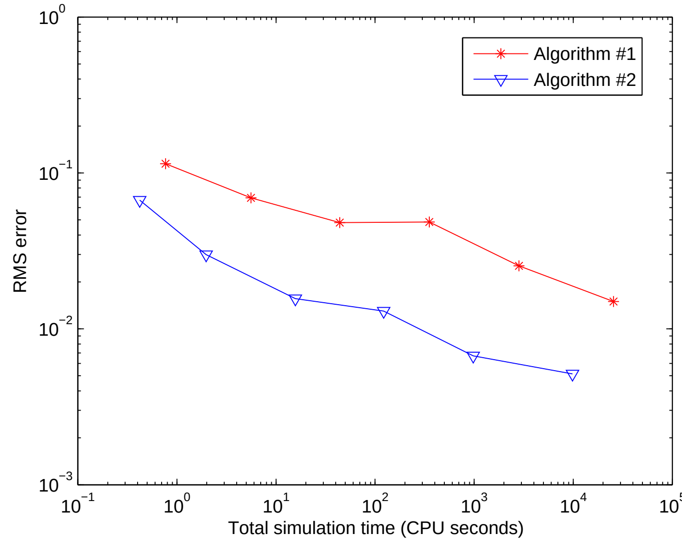


Figure 5.4: Convergence of the RMS errors of Monte Carlo simulation Algorithm 1 and Algorithm 2

implied volatility of European options.

As we can see in Figure 5.6, the price of a timer call option increases as the variance budget increases and manifests tiny convexity.

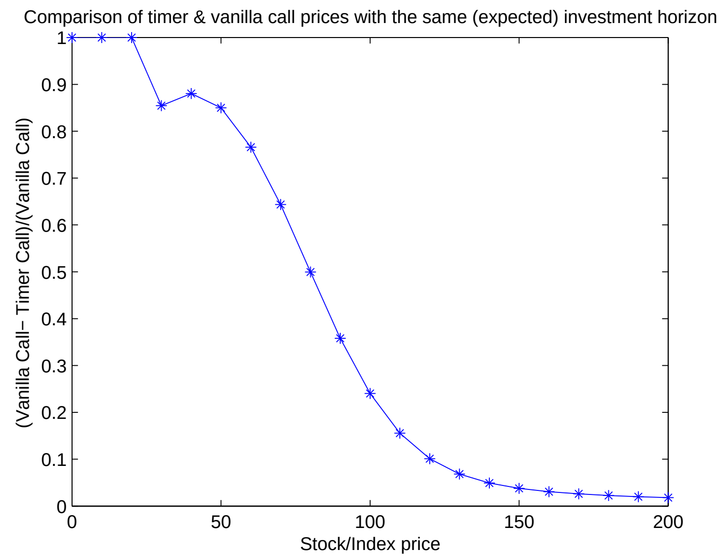


Figure 5.5: Comparison between the prices of timer call options and European call options

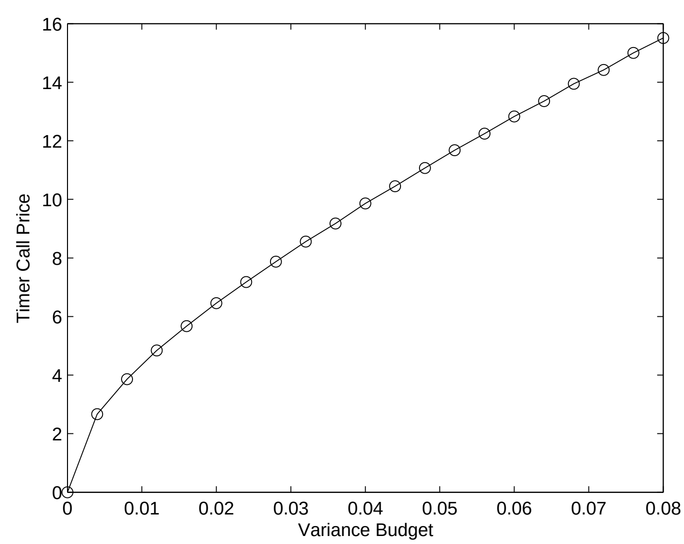


Figure 5.6: Timer call option price increases as variance budget increases

Chapter 6

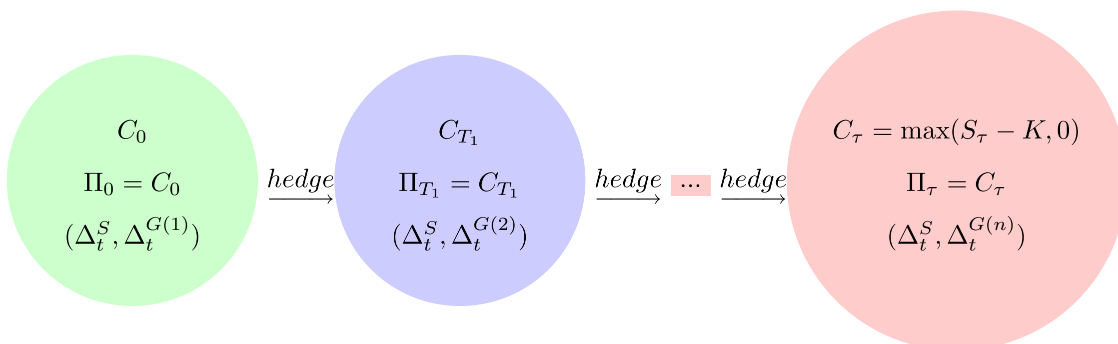
Dynamic Hedging Strategies

In this chapter, we discuss dynamic hedging strategies for timer options at the theoretical level. We also briefly illustrate the issue of the computation of price sensitivities used for hedging and risk-management purposes.

6.1 Dynamic Hedging Strategies

The underlying asset and the riskless money market account constitute an incomplete market due to the uncertain volatility risk. Thanks to the market completion via volatility-dependent assets (e.g. variance swap), we formulate the pricing of timer option as a first-passage-time problem via a standard no-arbitrage argument in Chapter 2. In this chapter, we focus on the dynamic hedging (replication) of timer option based on market completion.

We hedge the timer call option by the original underlying asset and a sequence of contiguous auxiliary volatility-dependent securities. In the following exposition, we take variance swap as an example of the auxiliary hedging instruments. Because we need to hedge the timer call option until its random maturity τ, we may need to replace matured variance swaps with new ones during the whole life of the timer call option. We regard $[0,\tau]$ as a composition of hedging periods $\mathcal{D}_i = [\sum_{j=0}^{i-1} T_j \wedge \tau, \sum_{j=0}^{i} T_j \wedge \tau],\ i = 1, 2, 3, ...$ (we define $T_0 = 0$). On $\mathcal{D}_i$ a variance swap with maturity T_i and price process $\{G_i(t)\}$ is employed as an auxiliary hedging instrument. In fact, in order to hedge a timer option we need not only to rebalance the replicating portfolio locally within each period $\mathcal{D}_i$, but we also need to replace the expired variance swap with a new one at each time T_i, if $T_i < \tau$, until time τ when the total variance budget is consumed. The following chart illustrates this hedging mechanism.



Without loss of generality, we focus on the first hedging period $[0, T_1 \wedge \tau]$. Let $\{\Pi_t\}$ denote the value process of the self-financing hedging portfolio. Suppose that at time t the portfolio consists of Δ_t^s shares of asset with price S_t and Δ_t^G shares of variance swap with price G_t (note that we drop the subscription 1 of G_1 without introducing any confusion), and the rest of the money fully invested in the riskless money market account. Therefore,

$$d\Pi_t = \Delta_t^S dS_t + \Delta_t^G dG_t + r(\Pi_t - \Delta_t^S S_t - \Delta_t^G G_t)dt.$$

It is equivalent to

$$d(e^{-rt}\Pi_t) = e^{-rt}\left[\Delta_t^S(dS_t - rS_t dt) + \Delta_t^G(dG_t - rG_t dt)\right].$$

Let $G_t = \mathcal{G}(t, V_t, I_t)$ for some function $\mathcal{G}$. According to Broadie and Jain (2008) [16], we have that

$$d(e^{-rt}G_t) = e^{-rt}\frac{\partial \mathcal{G}}{\partial v}\sigma_v\sqrt{V_t}dW_t^{(1)}.$$

Thus the replicating portfolio satisfies

$$d(e^{-rt}\Pi_t) = e^{-rt}\Big[\Delta_t^G\frac{\partial \mathcal{G}}{\partial v}\sigma_v\sqrt{V_t}dW_t^{(1)} + \Delta_t^S\sqrt{V_t}S_t(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)})\Big]. \tag{6.1}$$

On the other hand, for the timer option price we deduce that

$$d(e^{-rt}C_t) = e^{-rt}\Big[\frac{\partial u}{\partial v}\sigma_v\sqrt{V_t}dW_t^{(1)} + \frac{\partial u}{\partial s}\sqrt{V_t}S_t(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)})\Big]. \tag{6.2}$$

Finally, we let

$$d(e^{-rt}C_t) = d(e^{-rt}\Pi_t).$$

Therefore, the dynamic hedging strategy localized within a certain hedging period $\mathcal{D}_i$ reads

$$\begin{cases} \Delta_t^S = \dfrac{\partial u}{\partial s} \\ \Delta_t^G = \dfrac{\frac{\partial u}{\partial v}}{\frac{\partial \mathcal{G}}{\partial v}}. \end{cases} \tag{6.3}$$

REMARK 8. The hedging procedure can be also done with fixed maturity European options. Employing a European option with maturity T and price process $\{H_t\}$ ($H_t = \mathcal{H}(t, S_t, V_t, I_t)$ for some function $\mathcal{H}(t, s, v, I)$) as the auxiliary security, the strategy for dynamically hedging a timer call option reads

$$\begin{cases} \Delta_t^S = \dfrac{\partial u}{\partial s} - \dfrac{\dfrac{\partial u}{\partial v} \cdot \dfrac{\partial \mathcal{H}}{\partial s}}{\dfrac{\partial \mathcal{H}}{\partial v}} \\ \Delta_t^H = \dfrac{\dfrac{\partial u}{\partial v}}{\dfrac{\partial \mathcal{H}}{\partial v}}. \end{cases} \tag{6.4}$$

6.2 Computation of Price Sensitivities

From Section 6.1, we see that both timer call Delta (the sensitivity with respect to S_0) and Vega (the sensitivity with respect to V_0) are important for dynamic hedging of timer options. Price sensitivities on other model parameters might be useful for risk management purpose. In this section, we discuss the computation of sensitivities based on various computational methods we discussed previously.

Upon implementing the analytical formula in Theorem 4, we can use the corresponding finite difference to approximate a sensitivity. For a general parameter Θ (could be S_0,V_0, etc)

$$\frac{\partial C_0}{\partial \Theta} \approx \frac{C_0(\Theta + \delta\Theta) - C_0(\Theta)}{\delta\Theta}.$$

It is straightforward to implement it by choosing small $\delta\Theta$ so as to achieve a pre-specified error tolerance level. We can also perform path-wise simulation (see Broadie and Glasserman [15]) to compute price sensitivities. We can obtain the sensitivities almost for free, once the simulation of underlying process is done for valuation. However, we need to tolerate the bias in this approach. Based on Proposition 4 and Corollary (5.16), we can apply pathwise approach combined with Euler discretization on the Bessel process in order to compute the price sensitivities with respect to $\Theta \in \{B, \kappa, \theta, \sigma_v, V_0\}$. By taking the derivatives inside the expectation, we obtain the following proposition.

PROPOSITION 7. *The pathwise estimator for the price sensitivity with respect to $\Theta \in \{B, \kappa, \theta, \sigma_v, V_0\}$ admits the following form.*

$$\widehat{\frac{\partial C_0}{\partial \Theta}} = \Bigg[S_0 e^{d_0(Y_1,Z_1)} N(d_1(Y_1,Z_1))D_0 + S_0 e^{d_0(Y_1,Z_1)}\phi(d_1(Y_1,Z_1))D_1 - Ke^{-rZ_1}\phi(d_2(Y_1,Z_1))D_2 + Kre^{-rZ_1}N(d_2(Y_1,Z_1))\left(\frac{\partial Z_1}{\partial\Theta} + \frac{\kappa}{\sigma_v^2}\frac{\partial Y_1}{\partial \Theta} - \frac{\kappa^2\theta}{\sigma_v^2}\frac{\partial Z_1}{\partial\Theta}\right)\Bigg]\exp\left\{\frac{\kappa}{\sigma_v^2}(V_0 - Y_1) + \frac{\kappa^2\theta}{\sigma_v^2}Z_1 - \frac{\kappa^2 B}{2\sigma_v^2}\right\}, \tag{6.5}$$

where

$$\begin{aligned} D_0 =& \frac{\partial d_0}{\partial v}(Y_1,Z_1)\frac{\partial Y_1}{\partial\Theta} + \frac{\partial d_0}{\partial \xi}(Y_1,Z_1)\frac{\partial Z_1}{\partial\Theta} + \frac{\partial d_0}{\partial \Theta}(Y_1,Z_1) - \frac{\kappa}{\sigma_v^2}\frac{\partial Y_1}{\partial\Theta} + \frac{\kappa^2\theta}{\sigma_v^2}\frac{\partial Z_1}{\partial\Theta},\\ D_1 =& \frac{\partial d_1}{\partial v}(Y_1,Z_1)\frac{\partial Y_1}{\partial\Theta} + \frac{\partial d_1}{\partial \xi}(Y_1,Z_1)\frac{\partial Z_1}{\partial\Theta} + \frac{\partial d_1}{\partial \Theta}(Y_1,Z_1),\\ D_2 =& \frac{\partial d_2}{\partial v}(Y_1,Z_1)\frac{\partial Y_1}{\partial\Theta} + \frac{\partial d_2}{\partial \xi}(Y_1,Z_1)\frac{\partial Z_1}{\partial\Theta} + \frac{\partial d_2}{\partial \Theta}(Y_1,Z_1), \end{aligned} \tag{6.6}$$

and

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

and

$$\begin{aligned} \frac{\partial d_0}{\partial v}(v,\xi) &= \frac{\rho}{\sigma_v}, & \frac{\partial d_0}{\partial \xi}(v,\xi) &= -\frac{\rho}{\sigma_v}\kappa\theta,\\ \frac{\partial d_1}{\partial v}(v,\xi) &= \frac{\rho}{\sigma_v\sqrt{(1-\rho^2)B}}, & \frac{\partial d_1}{\partial \xi}(v,\xi) &= \frac{r - \frac{\rho}{\sigma_v}\kappa\theta}{\sqrt{(1-\rho^2)B}},\\ \frac{\partial d_2}{\partial v}(v,\xi) &= \frac{\rho}{\sigma_v\sqrt{(1-\rho^2)B}}, & \frac{\partial d_2}{\partial \xi}(v,\xi) &= \frac{r - \frac{\rho}{\sigma_v}\kappa\theta}{\sqrt{(1-\rho^2)B}}. \end{aligned} \tag{6.7}$$

In order to implement Proposition 7, we just need to simulate $(\frac{\partial Y_1}{\partial\Theta}, \frac{\partial Z_1}{\partial\Theta})$. This can be obtained directly, once we simulate the trajectories of (Y_t, Z_t). We demonstrate the method through the following example of computing Vega, i.e. $\frac{\partial C_0}{\partial V_0}$.

We differentiate $\{Y_t, Z_t\}$ with respect to V_0. By the dynamics in (5.16), we obtain that,

$$
\begin{aligned}
d\left(\frac{\partial Y_t}{\partial V_0}\right) &= -\frac{B\kappa\theta}{Y_t^2}\left(\frac{\partial Y_t}{\partial V_0}\right)dt; \quad \frac{\partial Y_0}{\partial V_0} = 1,\\
d\left(\frac{\partial Z_t}{\partial V_0}\right) &= -\frac{B}{Y_t^2}\left(\frac{\partial Y_t}{\partial V_0}\right)dt; \quad \frac{\partial Z_0}{\partial V_0} = 0.
\end{aligned}
\tag{6.8}
$$

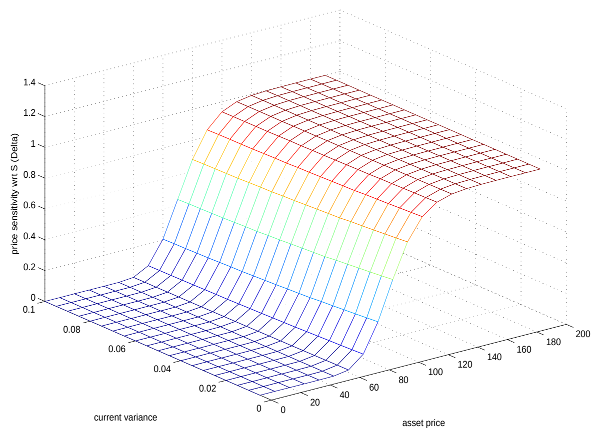
Therefore, by discretization, we design the following Euler scheme to simulate $(\frac{\partial Y_1}{\partial V_0}, \frac{\partial Z_1}{\partial V_0})$.

$$
\begin{aligned}
\left(\frac{\partial Y}{\partial V_0}\right)(i+1) &= \left(\frac{\partial Y}{\partial V_0}\right)(i) - \frac{B\kappa\theta}{Y^2(i)}\left(\frac{\partial Y}{\partial V_0}\right)(i)\Delta t,\\
\left(\frac{\partial Z}{\partial V_0}\right)(i+1) &= \left(\frac{\partial Z}{\partial V_0}\right)(i) - \frac{B}{Y^2(i)}\left(\frac{\partial Y}{\partial V_0}\right)(i)\Delta t.
\end{aligned}
\tag{6.9}
$$

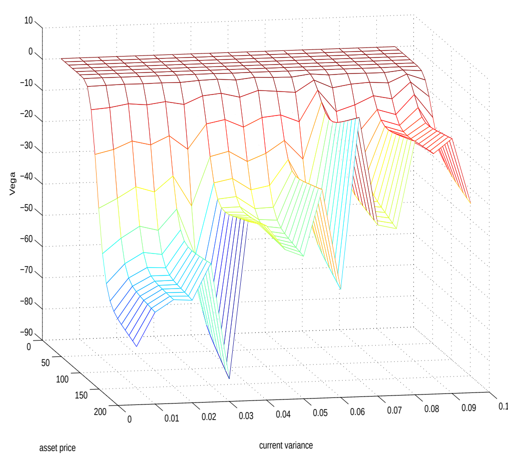
We take the terminal value from these iterations to evaluate the estimators for the price sensitivities.

As to the computation of sensitivities via PDE approach, we use finite difference to approximate the corresponding derivatives. It is straightforward to obtain the price sensitivity with respect to S_0 (the so called Delta) and the price sensitivity with respect to V_0 (the so called Vega) once the price surface 6.1 is obtained.

For the purpose of illustration, we plot the surface of timer call Delta and Vega, which are both important for hedging, in Figure 6.1. We observe that the Delta of the timer call option is bounded between zero and one, while the Vega keeps negative with considerable fluctuation.



(a) Timer call Delta $\frac{\partial C}{\partial S_0}$



(b) Timer call Vega $\frac{\partial C}{\partial V_0}$

Figure 6.1: Numerical examples of time call option price sensitivities: Delta and Vega

Part II

Efficient Valuation of VIX Options under Gatheral's Double Log-normal Stochastic Volatility Model

Chapter 7

Introduction to Part II

7.1 A Brief Outline of Part II

Consistent modeling and pricing of options on S&P 500 (the Standard and Poor's 500 index), VIX (the CBOE volatility index) and realized variance are an important issues for the financial derivatives market. Gatheral's (2007, 2008) double mean-reverting stochastic volatility model achieves such a goal to reduce the model risk. However, the non-affine structure of this model leads to the analytical intractability in the sense that the characteristic function of the underlying variables might not exist in closed-form. At the time when the model was proposed, its calibration involved massive Monte Carlo simulation. In this part, we begin with some analysis of the model. Then, we develop a generally applicable asymptotic expansion method for efficient valuation of derivatives under non-affine multidimensional diffusion models; and demonstrate it in the valuation of options on VIX under Gatheral's double log-normal model. Our probabilistic method is a combination of scaling, pathwise Taylor expansion, inductive computation of correction terms, calculation of conditional mixed Brownian moments, etc. The non-linear payoff function of an option on VIX is complicated; nevertheless convergence of our expansion is rigorously justified via the theory of Malliavin-Watanabe-Yoshida. In numerical examples, we demonstrate that the formula efficiently achieves desirable accuracy for relatively short maturity cases. Following our probabilistic computation approach and theoretical basis, better accuracy for longer maturity options can be obtained by adding more correction terms.

7.2 Modeling VIX

VIX, the Chicago Board Options Exchange (CBOE) Volatility Index (see [23]), is the premier benchmark for U.S. stock market volatility. It provides investors a direct and effective way to understand volatility. As an implied volatility index, VIX measures market expectations of near term (next 30 calendar days) volatility conveyed by stock index option prices. Since volatility often signifies financial turmoil, VIX is often referred to as the "investor fear gauge".

According to CBOE [23], independent of model specification, an index σ^2 is calculated by averaging the weighted prices of out-of-the-money puts and calls on S&P 500 (the Standard & Poor's 500 index), i.e.

$$\sigma^2 = \frac{2}{\Delta T} \sum_i \frac{\Delta K_i}{K_i^2} e^{r\Delta T} Q(K_i) - \frac{1}{\Delta T} \left[\frac{F}{K_0} - 1 \right]^2, \tag{7.1}$$

where the various parameters are defined as follows.

ΔT :	Time to expiration,
F :	Forward index level derived from index option prices,
K_i :	Strike price of i-th out-of-the-money option,
ΔK_i :	Interval between strike prices,
K_0 :	First strike below the forward index level F,
r :	Risk-free interest rate to expiration,
$Q(K_i)$:	The midpoint of the bid-ask spread for each option with strike K_i.

Conventionally, VIX is calculated as

$$VIX = \sigma \times 100. \tag{7.2}$$

The squared VIX satisfies that

$$VIX^2 = \frac{2 \times 10^4}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) - \frac{10^4}{T} \left[\frac{F}{K_0} - 1 \right]^2. \tag{7.3}$$

Following Carr and Wu (2006) [22], we introduce the theoretical proxy of VIX and demonstrate the necessity of employing a model, which is able to consistently price options on VIX and S&P 500. To begin with, we recall the definition of theoretical VIX, which is used in the modeling and valuation of derivative securities on VIX. On the filtered risk neutral probability space $(\Omega, \mathbb{Q}, \mathcal{F}, \{\mathcal{F}_t\})$, we assume the asset dynamics to be

$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t} dW_t,$$

where $\{V_t\}$ denotes the stochastic variance process. Let us define the realized variance over the time interval $[t, t+\Delta T]$ to be

$$RV_{t,t+\Delta T} := \frac{1}{\Delta T}\int_t^{t+\Delta T} V_s ds.$$

Further, $\widehat{VIX}_t^2$, the theoretical squared VIX (see Carr and Wu [22]) is defined as

$$\widehat{VIX}_t^2 = \mathbf{E}^Q[RV_{t,t+\Delta T}|\mathcal{F}_t] = \mathbf{E}^Q\left[\frac{1}{\Delta T}\int_t^{t+\Delta T} V_s ds \middle| \mathcal{F}_t\right]. \tag{7.4}$$

For now, we exclude the presence of jump in the asset dynamics. Otherwise, jumps will contribute to the realized variance as well. The theoretical squared VIX (7.4) can be statically replicated via forward-starting S&P 500 put and call options, i.e.

$$\widehat{VIX}_t^2 = \frac{2}{\Delta T}\left[\int_{F_t}^{\infty} \frac{e^{r\Delta T}}{K^2} C^{Model}(t, t+\Delta T, K)dK + \int_0^{F_t} \frac{e^{r\Delta T}}{K^2} P^{Model}(t, t+\Delta T, K)dK\right], \tag{7.5}$$

where $F_t = S_t e^{r(T-t)}$ denotes the forward price of S_t contracted for a later time $T > t$. $C^{Model}(t, t+\Delta T, K)$ and $P^{Model}(t, t+\Delta T, K)$ denote the model price of forward started call and put options with maturity ΔT, respectively. This static replication idea has been successfully applied to the study of variance swaps (see Derman et al. [29]). Thus, the theoretical squared VIX (7.4) is represented by a strip of model-generating out-of-the-money S&P 500 option prices over a wide range of strikes.

Suppose that we have a stochastic volatility model which fits the entire S&P 500 options smile. In other words, for any strike K,

$$\begin{aligned}C^{Model}(t,t+\Delta T,K) =& e^{-r\Delta T}\mathbb{E}_t^Q(S_{t+\Delta T}-K)^+\\ =& C_{BS}(\sigma_{imp}(K,\Delta T),S_t) = C^{Market}(t,t+\Delta T,K)\\ P^{Model}(t,t+\Delta T,K) =& e^{-r\Delta T}\mathbb{E}_t^Q(K-S_{t+\Delta T})^+\\ =& P_{BS}(\sigma_{imp}(K,\Delta T),S_t) = P^{Market}(t,t+\Delta T,K).\end{aligned} \tag{7.6}$$

where $C^{Market}(t,t+\Delta T,K)$ and $P^{Market}(t,t+\Delta T,K)$ denote the market price of forward started call and put options with maturity T. Here C_{BS} and P_{BS} denote the Black-Scholes call and put option pricing formulae, respectively; and $\sigma_{imp}(K,\Delta T)$ denotes the implied volatility associated with strike K and maturity ΔT. Therefore, it follows that

$$\widehat{VIX}_t^2 = \frac{2}{\Delta T}\left[\int_{F_t}^{\infty}\frac{e^{r\Delta T}}{K^2}C^{Market}(t,t+\Delta T,K)dK + \int_0^{F_t}\frac{e^{r\Delta T}}{K^2}P^{Market}(t,t+\Delta T,K)dK\right]. \tag{7.7}$$

Hence, we can interpret σ (defined in (7.3)) as a discretization of its theoretical counterpart defined in (7.4). i.e.

$$\widehat{VIX}^2 \approx \sigma^2. \tag{7.8}$$

Therefore, we have that

$$\widehat{VIX}\times 100 \approx VIX. \tag{7.9}$$

An excellent stochastic volatility model, under which the approximate relation (7.8) holds, allows us to use theoretical VIX (7.4) as a proxy to study the pricing of VIX options and futures, etc. within a continuous time framework. Therefore, we need a model which matches S&P 500 options market prices across all strikes and also produces correct VIX option market prices. Under the assumption of absence of jumps in the asset dynamics, the error resulted from using theoretical VIX is solely induced by the discretization of the integration in (7.7) on strikes.

Similar to the Black-Scholes-Merton model (1973) [10, 79] for pricing equity options, Whaley's (1993) [91] model directly regarded VIX as a geometric Brownian motion with constant volatility. Following the modeling approach in Cox et al. (1985) [26] and Heston (1993) [61], Grunbichler and Longstaff (1996) [55] specified the dynamics of VIX as a mean-reverting square root process. Detemple and Osakwe (2000) [30] employed a logarithmic mean-reverting process for pricing options on volatility. Carr and Lee (2007) [20] (also see Carr and Wu (2006) [22]) proposed a model-free approach by using the associated variance and volatility swap rates as the model inputs. Following Eraker et al. (2003, 2004) [40, 64], Lin and Chang (2009) modeled the S&P 500 index and the S&P 500 stochastic volatility by correlated jump-diffusion processes. Following the approach proposed in Bergomi (2005, 2009) [6, 7], Cont and Kokholm (2009) studied a modeling framework for the joint dynamics of an index and a set of forward varaince swap rates written on this index.

7.3 Historical Developments on Modeling Multi-factor Stochastic Volatility

The Black-Scholes-Merton (1973) model [10, 79] assumes that the underlying volatility is constant over the life of the derivative, and is unaffected by the changes in the underlying asset price. However, one of the most prominent shortcomings of this model is that it cannot explain features of the implied volatility surface such as volatility smile and skew, which indicate that implied volatility does tend to vary with respect to strike price and expiration. Stochastic volatility is a popular approach to resolve such deficiency of the Black-Scholes-Merton (1973) model. In particular, the early attempts of two-factor stochastic volatility models, such as Hull and White (1987) [62] and Heston (1993) [61], assume that the volatility of the underlying price is a stochastic process rather than a constant, which makes it possible to model derivatives more accurately. Two-factor stochastic volatility models can fit the volatility smiles. However, these models are restrictive in the modeling of the relationship between the S&P 500 volatility level and the slope of the smile. The consistent modeling of options on VIX and realized variance is confronted with even more challenge.

In the literature of equity/index option valuation, the deficiencies of the two-factor stochastic volatility model have traditionally been addressed by adding jump components to the dynamics. See for example Bates (1996) [4], Duffie, Pan and Singleton (2000) [36], Broadie, Chernov and Johannes (2007) [14] and Lipton (2002) [75]. We do not doubt the usefulness of this approach. Instead, we claim that adding additional factors to the volatility process can alternatively circumvent the model deficiencies. Here, we do not investigate whether multi-factor models or jump processes are more appropriate for modeling option data.

Early considerations of using multi-factor stochastic volatility models to improve the goodness of fit for S&P 500 returns can be found in Bates (1997) [5] and Gallant et. al (1999) [47]. Duffie, Pan and Singleton (2000) [36] propose an affine class * of three-factor stochastic volatility models. Based on an assumption on the specific correlation structure, an intermediate stochastic trend component is added to volatility, which generalizes Heston's (1993) [61] two-factor stochastic volatility model. Also relying on the affine class, Heston et al. (2009) [24] propose a three-factor stochastic volatility

*For the affine class models, see Duffie and Kan (1996) [35], Duffie, Filipovic and Schachermayer (2003) [33].

model by adding a second CIR process of variance, which is independent of the one in Heston (1993) [61]. Due to the special requirement on the correlation structure, this model is analytically tractable in the sense that the characteristic function of the underlying admits closed-form formula.

Gatheral's double mean-reverting stochastic volatility model (2007, 2008) [50, 49, 51] generalizes the affine three-factor model in Duffie, Pan and Singleton (2000) [36] by allowing more flexibility in the correlation structure and the specification of the diffusion components. Supported by some empirical study, the variance is more properly modeled by a double-CEV type mean-reverting process instead of using square root processes. Thus, it falls into the non-affine class, which is quite computationally challenging.

7.4 Gatheral's Double Mean-Reverting Stochastic Volatility Model

Under the risk-neutral probability measure, Gatheral's double mean-reverting stochastic volatility model (hereafter DMR-SV) is specified as follows:

$$\begin{aligned}\frac{dS_t}{S_t} &= r dt + \sqrt{V_t} dW(t), \quad S_0 = s_0 > 0, \\ dV_t &= \kappa(V_t' - V_t)dt + \xi_1 V_t^{\alpha} dZ_1(t), \quad V_0 = v_0 > 0, \\ dV_t' &= c(z_3 - V_t')dt + \xi_2 V_t'^{\beta} dZ_2(t), \quad V_0' = v_0' > 0,\end{aligned} \tag{7.10}$$

where the Brownian motions are correlated as

$$dZ_1(t)dZ_2(t) = \rho dt, \quad dW(t)dZ_1(t) = \rho_1 dt, \quad dW(t)dZ_2(t) = \rho_2 dt. \tag{7.11}$$

This model is presented in Gatheral (2007) [49], its further development in Gatheral (2008a) [51] and Gatheral (2008b) [50]. In this model, $\{S_t\}$ represents the price process of a stock or index. The instantaneous variance level V_t reverts to the moving intermediate level V_t' at rate k, while V_t' reverts to the long-term level z_3 at the slower rate $c < k$. In the dynamics (7.10), r is the instantaneous interest rate, which is assumed to be a constant; ξ_1 and ξ_2 are the volatility parameter for $\{V_t\}$ and $\{V_t'\}$, respectively; α and β are the constant elasticity parameters for $\{V_t\}$ and $\{V_t'\}$, respectively. By adding such a moving intermediate level, an additional freedom is given compared to traditional two-factor stochastic volatility models, e.g. Heston (1993) [61], etc. It provides a new perspective, from financial economics, for us to capture more about the dynamics of stochastic volatility. In this dissertation, we focus on the study of the double mean-reverting volatility process $\{(V_t, V_t')\}$.

The CEV (constant elasticity of volatility) type parameters are chosen as $\alpha, \beta \in [1/2, 1]$. According to Gatheral [50], the models are classified as:

- Double Heston for the case of $\alpha = \beta = 1/2$,
- Double Log-normal for the case $\alpha = \beta = 1$,
- Double CEV for any other cases.

We note that both V_t and V_t' are always finite and positive. For the case of double log-normal ($\alpha = \beta = 1$), this property of (V_t, V_t') can be seen from its explicit solution.

PROPOSITION 8. *When $\alpha = \beta = 1$, the linear stochastic differential equation* *governing the bivariate diffusion* $\{(V_t, V_t')\}$ *can be solved explicitly as*

$$\begin{aligned} V'(t) =& \exp\left\{-ct + \xi_2 B_2(t) - \frac{1}{2}\xi_2^2 t\right\} \\ &\cdot \left(V_0' + cz_3 \int_0^t \exp\left\{cu - \xi_2 B_2(u) + \frac{1}{2}\xi_2^2 u\right\} du\right), \\ V(t) =& \exp\left\{-\kappa t + \xi_1 B_1(t) - \frac{1}{2}\xi_1^2 t\right\} \\ &\cdot \left(V_0 + \kappa \int_0^t \exp\left\{\kappa u - \xi_1 B_1(u) + \frac{1}{2}\xi_1^2 u\right\} V'(u) du\right). \end{aligned} \tag{7.12}$$

For the case $\alpha = \beta = 1$, the probability distribution of (V_t, V_t') is not strictly log-normal, however, according to the local behavior of the dynamics we rationally call it double log-normal, which is consistent with Gatheral (2008) [50]. We note that with $\xi_2 = 0$ and $c = 0$, the double log-normal model becomes the λ-SABR model of Henry-Labordere (2008) [60]. In this case, the dynamics of the variance process $\{V(t)\}$ follows a continuous-time version of the well-known GARCH$(1, 1)$ model. It assumes that the randomness of the variance process varies with the variance, as opposed to the square root of the variance in the Heston (1993) [61] model. It is a special case of the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model (see Bollerslev (1986) [11]), which is popular for estimating stochastic volatility. We note that Gatheral's model is compatible with the models on forward variance such as Dupire (1993) [38], Bergomi (2005) [6] and Buehler (2006) [19]. The dynamics of Gatheral's model (7.10) is implied by properly selecting a variance curve functional (see Gatheral (2008) [50]).

It makes sense to add at least one more factor to the traditional two-factor stochastic volatility models, because we know from the principle component analysis (PCA) of volatility surface time series that there are at least three important modes of fluctuation, namely level, term structure and skew. By employing three factors, Gatheral's

DMR-SV model is able to price options on S&P 500, VIX and realized variance consistently. It generates a correct S&P 500 volatility skew for short expirations with no need for jumps. Moreover, the model fits the option market with time homogeneous parameters.

Gatheral (2007, 2008) [50, 49, 51] demonstrates that the DMR-SV model is able to fit consistently to the market of options on VIX, S&P 500 and realized variance very well. Gatheral's [50] step-by-step calibration strategy begins with estimating some parameters via the analysis of time series data, and then identifies others via fitting options on S&P 500 and VIX. In the procedure of fitting option values, his method heavily relies on Monte Carlo simulation, which is very time-consuming when a large amount of computation is needed. Gatheral (2007, 2008) [50, 49, 51] suggested that the local log-normal structure (the case of $\alpha = \beta = 1$) works well in terms of fitting the market; nevertheless the optimal parameter set, which is freely calibrated to the market, indicates that $\alpha = \beta = 0.94 \approx 1$. In this paper, based on the tradeoff between the performance and the relative simplicity of the case $\alpha = \beta = 1$, we are motivated to provide and implement an efficient asymptotic expansion formula for pricing VIX options under the double log-normal model. We explicitly derive the first three expansion terms and prove the validity of our approach rigorously. Following our probabilistic computation approach and theoretical basis, better accuracy for longer maturity options can be obtained by adding more correction terms.

Though the double log-normal model works well, a minor limitation one should keep in mind, as commented in Staunton (2009) [86], is that the underlying variable (V_t, V_t') is conditionally stable.

PROPOSITION 9. *(V_t, V_t') has finite second moments if and only if*

$$2\kappa > \xi_1^2 \text{ and } 2c > \xi_2^2. \tag{7.13}$$

Proof. Indeed, this can be seen from straightforward computation on the moments through Ito's lemma. Let us take $V'(t)$ as an example. We can compute that

$$\mathbb{E}V'(t) = e^{-ct}V_0' + z_3(1 - e^{-ct}). \tag{7.14}$$

Thus, by applying Ito's lemma and solving an ODE, we find that

$$\begin{aligned}\mathbb{E}V'(t)^2 =& e^{-(2c-\xi_2^2)t}V_0'^2 + \frac{2cz_3(V_0' - z_3)}{c - \xi_2^2}(e^{-ct} - e^{-(2c-\xi_2^2)t}) + \frac{2cz_3^2}{2c - \xi_2^2}(1 - e^{-(2c-\xi_2^2)t}) \\ \to& \frac{2c}{2c - \xi_2^2}z_3^2 \quad \text{as } t \to \infty, \quad \text{only if } 2c > \xi_2^2.\end{aligned} \tag{7.15}$$

Similarly, we find the stability condition $2\kappa > \xi_1^2$ for $\{V(t)\}$. □

If the stability condition (7.13) is violated, one of the practical consequences of this is that the Monte Carlo simulation would generate unreasonably large values of $V'(t)$, which might lead to the divergence of the numerical computation results. In addition, the growth of the second moment of V_t as time t increasing is inconsistent with the historical time series of VIX and the price of options on VIX. Therefore, we advocate the consideration of the parameters restriction (7.13) when the double log-normal model is calibrated to the market data.

The rest of this part is organized as follows. In chapter 8, we set up the valuation problem for VIX option under the double log-normal model. Then, we present an asymptotic expansion formula for pricing VIX options. By deriving this formula, we essentially provide a generally applicable asymptotic expansion method for efficient valuation of derivatives under non-affine multidimensional diffusion models. In chapter 9, we explore the computing performance by conducting a set of numerical experiments. In chapter 10, the validity of our expansion is justified via the theory of Malliavin-Watanabe-Yoshida.

Chapter 8

Valuation of Options on VIX under Gatheral's Double Log-normal Stochastic Volatility Model

There are mainly two approaches to derive asymptotic expansions for derivatives valuation. One is the PDE based singular (regular) perturbation method, for example, Hagan et al. (2002) [56] and Andersen and Brotherton-Ratcliffe (2005) [3]. Another one is based on probabilistic techniques. For example, Kunitomo and Takahashi (2001) [70] and Osajima (2007) [82] employed the computation of the conditional expectation of Wiener-Ito chaos (see Nualart et al. (1988) [81]) to derive asymptotic expansions for the valuation of interest rate derivatives. In this section, we derive an efficient asymptotic expansion formula for pricing options on VIX under Gatheral's model probabilistically. Beyond the valuation of options on VIX under Gatheral's model, our method provides a generally applicable platform for a wide variety of valuation problems. We begin with the pathwise Taylor expansion of the underlying diffusion and write the expansion terms in the form of Lebesgue integration. Subsequently, we inductively compute each correction terms. The computation is based on the conditional mixed Brownian moments, which is easier to implement compared with the Wiener-Ito chaos. The convergence of our asymptotic expansion is rigorously justified via the Theory of Malliavin-Watanabe-Yoshida.

8.1 Basic Setup

In the following sections, we use VIX to denote $\widehat{VIX}$ for notational convenience, since we work under a consistent model and ignore the discretization error in approximation (7.8). First, the squared VIX can be represented by a linear combination of instantaneous variance, intermediate level and long-term level.

PROPOSITION 10. *Under Gatheral's DMR-SV models,*

$$VIX_t^2 = \mathbb{E}^Q\left[\frac{1}{\Delta T}\int_t^{t+\Delta T} V_s ds \middle| \mathcal{F}_t\right] = a_1 V_t + a_2 V_t' + a_3 z_3, \tag{8.1}$$

where

$$\begin{aligned} a_1 &= \frac{1}{\Delta T}\cdot\frac{1-e^{-\kappa\Delta T}}{\kappa}, \\ a_2 &= \frac{1}{\Delta T}\cdot\frac{\kappa}{\kappa - c}\left[\frac{1-e^{-c\Delta T}}{c} - \frac{1-e^{-\kappa\Delta T}}{\kappa}\right], \\ a_3 &= 1 - \frac{1}{\Delta T}\cdot\frac{1-e^{-\kappa\Delta T}}{\kappa}. \end{aligned} \tag{8.2}$$

Based on the fact that the function $f(x) = (1-e^{-x})/x$ is decreasing and bounded by 1 as well as the modeling requirement of $c < k$, we have that $a_1 > 0, a_2 > 0, a_3 > 0$. Because of the mean-reverting feature in V_t and V_t', the mean-reversion of VIX is modeled explicitly from (8.1).

Proof. By taking expectations on the bivariate diffusion of (V_t, V_t'), we arrive at an ODE system

$$\begin{aligned} \mathbb{E}_t V_{t+\Delta T} - V_t &= \kappa \int_t^{t+\Delta T} (\mathbb{E}_t V_s' - \mathbb{E}_t V_s) ds, \\ \mathbb{E}_t V_{t+\Delta T}' - V_t &= c z_3 \Delta T - c \int_t^{t+\Delta T} \mathbb{E}_t V_s' ds. \end{aligned} \tag{8.3}$$

It is straightforward to solve it to obtain $\mathbb{E}_t V_s$ for $s > t$. Thus, by integration, we find the coefficients a_1, a_2 and a_3 as defined in (8.2) for representing squared VIX. This yields the statement in (8.1). □

Following a standard no-arbitrage argument (see the treatment of stochastic volatility in Gatheral (2006) [48]), the price of a VIX call option with maturity T and strike K can be represented as the risk-neutral expectation of the discounted payoff, i.e.

$$C_t = e^{-r(T-t)} \mathbb{E}^Q \left[\left(100 \times \sqrt{a_1 V_T + a_2 V_T' + a_3 z_3} - K \right)^+ \middle| \mathcal{F}_t \right]. \tag{8.4}$$

In this expression, we multiply 100 with the theoretical VIX defined in (7.4), according to the convention and definition of VIX (see (7.3) and (7.2)).

REMARK 9. We start from a physical specification of the double mean-reverting stochastic volatility model. Empirically, we add two auxiliary securities, within which volatility risk are embedded, to complete the market. For instance, we may choose S&P500 options or variance swaps to accomplish this market completion procedure. Then, we construct a risk-free self-financing portfolio consisting of Δ_t shares of VIX_t^2, which can be replicated by a stripe of out-of-money options (see (7.7) and (7.3)), $\Delta_t^{(1)}$ shares of the auxiliary security with value process $\{A_1(t)\}$ and $\Delta_t^{(2)}$ shares of the auxiliary security with value process $\{A_2(t)\}$. By taking a long position in the VIX call option (with value C_t^{VIX}) and by shorting all others, the portfolio value at time

t is given by

$$P(t) = C_t^{VIX} - \Delta_t VIX_t^2 - \Delta_t^{(1)} A_1(t) - \Delta_t^{(2)} A_2(t).$$

Therefore, by a no-arbitrage argument based on $dP(t) = rP(t)dt$ and a straightforward calculation using Ito's lemma, we find a PDE governing the VIX option price given a judicious choice of the mathematical form of the market price of volatility risk. By the Feynman-Kac theorem (see Karatzas and Shreve (1991)), the VIX option price admits the risk-neutral expectation form as in (8.23), while the corresponding risk-neutral specification of the double mean-reverting model is exhibited as in (7.10). Meanwhile, a theoretical delta-hedging strategy is suggested as

$$\Delta_t = \frac{\partial C_t^{VIX}}{\partial VIX_t^2}, \quad \Delta_t^{(1)} = \frac{\partial C_t^{VIX}}{\partial A_1(t)}, \quad \Delta_t^{(2)} = \frac{\partial C_t^{VIX}}{\partial A_2(t)}. \tag{8.5}$$

By straightforward computation, we find that

$$\begin{aligned}
\frac{\partial C_t^{VIX}}{\partial VIX_t^2} &= \frac{\partial C_t^{VIX}}{\partial V_t}\frac{\partial V_t}{\partial VIX_t^2} + \frac{\partial C_t^{VIX}}{\partial V_t'}\frac{\partial V_t'}{\partial VIX_t^2}, \\
\frac{\partial C_t^{VIX}}{\partial V_t} &= \frac{\partial C_t^{VIX}}{\partial A_1(t)}\frac{\partial A_1(t)}{\partial V_t} + \frac{\partial C_t^{VIX}}{\partial A_2(t)}\frac{\partial A_2(t)}{\partial V_t}, \\
\frac{\partial C_t^{VIX}}{\partial V_t'} &= \frac{\partial C_t^{VIX}}{\partial A_1(t)}\frac{\partial A_1(t)}{\partial V_t'} + \frac{\partial C_t^{VIX}}{\partial A_2(t)}\frac{\partial A_2(t)}{\partial V_t'}.
\end{aligned} \tag{8.6}$$

Thus, $(\Delta_t, \Delta_t^{(1)}, \Delta_t^{(2)})$ can be solved from (8.6) as

$$\begin{aligned}
\Delta_t &= \frac{1}{a_1}\frac{\partial C_t^{VIX}}{\partial V_t} + \frac{1}{a_2}\frac{\partial C_t^{VIX}}{\partial V_t'}, \\
\Delta_t^{(1)} &= \frac{\frac{\partial A_2(t)}{\partial V_t'}\frac{\partial C_t^{VIX}}{\partial V_t} - \frac{\partial A_2(t)}{\partial V_t}\frac{\partial C_t^{VIX}}{\partial V_t'}}{\frac{\partial A_2(t)}{\partial V_t'}\frac{\partial A_1(t)}{\partial V_t} - \frac{\partial A_2(t)}{\partial V_t}\frac{\partial A_1(t)}{\partial V_t'}}, \\
\Delta_t^{(2)} &= \frac{\frac{\partial A_1(t)}{\partial V_t'}\frac{\partial C_t^{VIX}}{\partial V_t} - \frac{\partial A_1(t)}{\partial V_t}\frac{\partial C_t^{VIX}}{\partial V_t'}}{\frac{\partial A_2(t)}{\partial V_t}\frac{\partial A_1(t)}{\partial V_t'} - \frac{\partial A_2(t)}{\partial V_t'}\frac{\partial A_1(t)}{\partial V_t}}.
\end{aligned} \tag{8.7}$$

The dynamic hedging portfolio might be expensive, but it can be constructed by a certain amount of shares of the two auxiliary securities and a stripe of out-the-money (OTM) options which replicates VIX^2. More specifically, in the hedging portfolio, we need to hold

$$\Delta_i(t) = \frac{2 \times 10^4}{T} \frac{\Delta K_i}{K_i^2} e^{rT} \left(\frac{1}{a_1} \frac{\partial C_t^{VIX}}{\partial V_t} + \frac{1}{a_2} \frac{\partial C_t^{VIX}}{\partial V_t'} \right)$$

number of shares of an European option with strike K_i.

8.2 An Asymptotic Expansion Formula for VIX Option Valuation

Despite the simplicity in (8.1), the payoff of European call options on VIX is highly nonlinear. In addition, the double log-normal model is not analytically tractable in the sense that the joint characteristic function of (V_t, V_t') is not easy to obtain in closed form. Therefore, we employ an asymptotic expansion technique to find an explicit formula for approximating the option price:

$$C_0^{VIX} = e^{-rT} \mathbb{E}^Q \left[\left(100 \times \sqrt{a_1 V_T + a_2 V_T' + a_3 z_3} - K \right)^+ \right]. \tag{8.8}$$

We let $\epsilon = \xi_2 \sqrt{T}$, the total volatility of the intermediate level, be the quantity based on which asymptotic expansion is performed. This choice is inspired by the calibration result in Gatheral (2008) [50]. As of April 2007, Gatheral (2008) [50] finds that $\xi_1 = 2.6$ and $\xi_2 = 0.45$ when a double CEV model is calibrated to the VIX option market; $\xi_1 = 7$ and $\xi_2 = 0.94$ when a double log-normal model is calibrated. Both of these cases suggest that ξ_2 should be a relatively small quantity. In addition to the requirement that $c < k$, we may expect that the intermediate level reverts at a lower speed and is more stable. However, we note that these modeling assumptions are still subject to further empirical tests if possible. Under certain market condition where the intermediate level of the volatility is not so volatile (say, $\xi_2 < 1$), the choice $\epsilon = \xi_2\sqrt{T}$ may accelerate the convergence comparing to expanding according to $\sqrt{T}$. We also note that this choice of expansion parameter resembles that of the SABR expansion based on the "small vol. of vol." in Hagan (2002) et al. [56]. Even for the case when ϵ is not small enough, we can still conduct an expansion to obtain a satisfactory approximation once enough correction terms are added. An analog to support this reasoning is that the Taylor expansion for exponential function e^x can be performed, i.e.

$$e^x = \sum_{n=0}^{N} \frac{1}{n!}x^n + \mathcal{O}\left(\frac{|x|^{N+1}}{(N+1)!}\right), \quad \text{as } n \to \infty, \text{ for any arbitrary } x \in \mathbb{R}. \tag{8.9}$$

In Theorem 5, we present the key result about an asymptotic expansion representation of the VIX option price. In order to obtain significant accuracy of the approximation for actively traded short maturity options, for example one-month options, we explicitly expand the price up to the first four correction terms, each of which is a combination of polynomials, normal density and cumulative distribution functions. Higher order terms can be similarly computed by using the method we present in the next section. In Remark 10, we give an alternative asymptotic expansion formula, for the first four correction terms, based on the choice of parameter as $\epsilon = \sqrt{T}$, which exclusively depends on the option maturity. In section 8.3, detailed derivation of the formula in Theorem 5 is explicitly carried out via scaling, pathwise Taylor expansion around $\epsilon = 0$, computation of conditional mixed Brownian moments, etc.

THEOREM 5. *The price of an option on VIX with strike K and maturity T admits the following asymptotic expansion:*

$$C_0^{VIX} = 100 \times \xi e^{-r\frac{\epsilon^2}{\xi_2^2}} \left\{ \epsilon\Theta_1(y) + \epsilon^2\Theta_2(y) + \epsilon^3\Theta_3(y) + \epsilon^4\Theta_4(y) + \mathcal{O}(\epsilon^5) \right\}. \tag{8.10}$$

Here

$$y = \frac{K/100 - X_0}{\epsilon\xi},$$

where $\epsilon = \xi_2\sqrt{T}$ and

$$X_0 = \sqrt{a_1V_0 + a_2V_0' + a_3z_3}, \quad \xi = \frac{\sqrt{a_1^2\xi_1^2V_0^2 + a_2^2\xi_2^2V_0'^2 + 2\rho a_1a_2\xi_1\xi_2V_0V_0'}}{2X_0\xi_2}.$$

In the expansion (8.10)*, we have that*

$$\begin{aligned}
\Theta_1(y) =& \phi(y) - y(1 - N(y)), \\
\Theta_2(y) =& (\Lambda_0 + \Lambda_1)(1 - N(y)) + \Lambda_1 y\phi(y), \\
\Theta_3(y) =& \left[\frac{1}{2}\Lambda_2 + \Lambda_3 + (2 + y^2)\Lambda_4\right]\phi(y), \\
\Theta_4(y) =& \left(\Lambda_5y^7 + \Lambda_6y^5 + \Lambda_7y^3 + \Lambda_8y\right)\phi(y) + 2\Lambda_9(1 - N(y)),
\end{aligned} \tag{8.11}$$

where Λ_i's are either constants or polynomials of y:

$$
\begin{aligned}
\Lambda_0 =& \alpha\overline{\sigma}_1^2 + \beta\overline{\sigma}_2^2 + \gamma, \\
\Lambda_1 =& \alpha\Omega_1^2 + \beta\Omega_2^2 + \delta, \\
\Lambda_2 =& 3\omega_2^2(y) + \omega_1^2(y) + \omega_0^2(y) + 2\omega_0(y)\omega_2(y), \\
\Lambda_3 =& \theta_7 + \frac{1}{12\xi X_0}\left(\theta_1\Omega_1 + \theta_2\Omega_2 + 3\theta_3\Omega_1\overline{\sigma}_1^2 + 3\theta_4\Omega_2\overline{\sigma}_2^2 + \frac{\theta_5\Omega_1}{2\xi_2^2} + \frac{\theta_6\Omega_2}{2\xi_2^2}\right), \\
\Lambda_4 =& \theta_8 + \frac{1}{12\xi X_0}\left(\theta_3\Omega_1^3 + \theta_4\Omega_2^3\right), \\
\Lambda_5 =& \frac{\Lambda_1^3}{6}, \\
\Lambda_6 =& \frac{1}{48\xi^2 X_0^2}[a_1^2V_0^2\xi_1^5\Omega_1^5 + a_1a_2V_0'V_0\xi_1^2\xi_2^2\Omega_2^2(\xi_1\Omega_1 + \xi_2\Omega_2)\Omega_1^2 + a_2^2V_0'^2\xi_2^5\Omega_2^5 + 24\xi^3X_0(-2\Lambda_1^2 \\
&+ \theta_8 - \Lambda_4) + 24\xi^2X_0^2\Lambda_1(-2\Lambda_1^2 + (\alpha\bar{\sigma}_1^2 + \beta\bar{\sigma}_2^2 + \gamma)\Lambda_1 + 4(\alpha\Omega_1\bar{\sigma}_1 - \beta\Omega_2\bar{\sigma}_2)^2)],
\end{aligned}
$$

$\Lambda_7 = \frac{1}{48\xi^2X_0^2}[a_2^2V_0'\xi_2^3(\frac{5cz_3}{\xi_2^2} - 2V_0'((\frac{2}{\xi_2^2} - 5\bar{\sigma}_2^2)\xi_2^2 + \frac{4c}{\xi_2^2}))\Omega_2^3 + a_1^2V_0\xi_1^2\Omega_1^2(\frac{\kappa V_0'(5\xi_1\Omega_1+3\xi_2\Omega_2)}{\xi_2^2} - 2V_0\xi_1\Omega_1((\frac{2}{\xi_2^2} - 5\bar{\sigma}_1^2)\xi_1^2 + \frac{4\kappa}{\xi_2^2})) - \xi^2X_0(-a_1V_0\xi_1^4\Omega_1^4 + 192\alpha^2\xi\bar{\sigma}_1^2\Omega_1^2 - 384\alpha\beta\xi\Omega_2\bar{\sigma}_1\bar{\sigma}_2\Omega_1 - a_2V_0'\xi_2^4\Omega_2^4 + 24\xi^2\Lambda_1^2 + 192\beta^2\xi\Omega_2^2\bar{\sigma}_2^2 - 24\xi\theta_7 + 24\xi\Lambda_3 + 48\xi^2\Lambda_4 + 96\xi\Lambda_1(\alpha\bar{\sigma}_1^2 + \beta\bar{\sigma}_2^2 + \gamma)) + 24\xi^2X_0^2(-4(\alpha\bar{\sigma}_1^2 + \beta\bar{\sigma}_2^2 + \gamma)\Lambda_1^2 + (3\alpha^2\bar{\sigma}_1^4 + 2\alpha(-8\alpha\Omega_1^2 + 3\beta\bar{\sigma}_2^2 + \gamma)\bar{\sigma}_1^2 + 32\alpha\beta\Omega_1\Omega_2\bar{\sigma}_2\bar{\sigma}_1 + 3\beta^2\bar{\sigma}_2^4 + \gamma^2 + 2\beta(\gamma - 8\beta\Omega_2^2)\bar{\sigma}_2^2)\Lambda_1 + 4\gamma(\alpha\Omega_1\bar{\sigma}_1 - \beta\Omega_2\bar{\sigma}_2)^2) - a_1a_2(-\frac{cV_0z_3\xi_1^2(2\xi_1\Omega_1+3\xi_2\Omega_2)\Omega_1^2}{\xi_2^2} - \kappa V_0'^2\Omega_2^2(3\xi_1\Omega_1 + 5\xi_2\Omega_2) + V_0'V_0(\Omega_1(3\Omega_2^2(\frac{1}{\xi_2^2} - \bar{\sigma}_1^2)\xi_2^2 + 6\Omega_1\Omega_2\bar{\sigma}_1\bar{\sigma}_2\xi_2^2 + \Omega_1^2((\frac{1}{\xi_2^2} - \bar{\sigma}_2^2)\xi_2^2 + \frac{2c}{\xi_2^2}))\xi_1^3 + \xi_2\Omega_2(\Omega_2^2(\frac{1}{\xi_2^2} - \bar{\sigma}_1^2)\xi_2^2 + 6\Omega_1\Omega_2\bar{\sigma}_1\bar{\sigma}_2\xi_2^2 + 3\Omega_1^2((\frac{1}{\xi_2^2} - \bar{\sigma}_2^2)\xi_2^2 + \frac{2c}{\xi_2^2}))\xi_1^2 + 6\kappa\Omega_1\Omega_2^2\xi_1 + 2\kappa\xi_2\Omega_2^3))],$

$\Lambda_8 = \frac{1}{48}(120\alpha^3\bar{\sigma}_1^6 + \frac{3(5a_1^2V_0^2\Omega_1\xi_1^5 - 48\alpha^2\xi^3X_0 + 24\alpha^2\xi^2X_0^2(5\beta\bar{\sigma}_2^2 + \gamma - 2\Lambda_1))\bar{\sigma}_1^4}{\xi^2X_0^2} - \frac{1}{\xi^2X_0^2}[3(2X_0(-a_1V_0\Omega_1^2\xi_1^4 + 16\alpha^2\xi^2\Omega_1^2 + 48\alpha\beta\xi\bar{\sigma}_2^2 + 16\alpha\gamma\xi + 8\alpha\xi^2\Lambda_1)\xi^2 + 8\alpha X_0^2(-15\beta^2\bar{\sigma}_2^4 - 6\beta(\gamma - 2\Lambda_1)\bar{\sigma}_2^2 + \gamma(8\alpha\Omega_1^2 - \gamma + 4\Lambda_1))\xi^2 + a_1V_0\xi_1^2(\frac{4a_1\xi_1(V_0(\xi_1^2+2\kappa)-2\kappa V_0')\Omega_1}{\xi_2^2} + a_2(\xi_1\Omega_1 + \xi_2\Omega_2)(V_0'((\frac{1}{\xi_2^2} - 3\bar{\sigma}_2^2)\xi_2^2 + \frac{2c}{\xi_2^2}) - \frac{2cz_3}{\xi_2^2})))\bar{\sigma}_1^2] + \frac{1}{\xi^2X_0^2}[6\bar{\sigma}_2(32\alpha\beta X_0(\xi^2 + 2\gamma X_0)\Omega_1\Omega_2\xi^2 + a_1a_2\xi_1\xi_2(-\frac{2\kappa\Omega_2V_0'^2}{\xi_2} + V_0(\frac{\Omega_2\xi_1^2}{\xi_2} + \Omega_1((\frac{1}{\xi_2^2} - \bar{\sigma}_2^2)\xi_2^2 + \frac{2c}{\xi_2^2})\xi_1 + \frac{2\kappa\Omega_2}{\xi_2})V_0' - \frac{2cV_0z_3\xi_1\Omega_1}{\xi_2^2}))\bar{\sigma}_1] + 120\beta^3\bar{\sigma}_2^6 + \frac{3a_1V_0\xi_1^4\Omega_1^4}{X_0} + \frac{3a_2V_0'\xi_2^4\Omega_2^4}{X_0} +$

$$3(\frac{5a_2^2V_0'^2\Omega_2\xi_2^5}{\xi^2X_0^2}+24\beta^2(\gamma-2\Lambda_1)-\frac{48\beta^2\xi}{X_0})\bar{\sigma}_2^4+8\gamma^3-\frac{72\xi^2\Lambda_1^2}{X_0}-\frac{8\kappa a_1V_0'\xi_1^2\rho_{13}^2}{(\zeta^2+2\eta\rho\zeta+\eta^2)X_0\xi_2^2}-\frac{8ca_2z_3\rho_{23}^2}{(\zeta^2+2\eta\rho\zeta+\eta^2)X_0}+$$

$$\frac{4\kappa a_1V_0'\rho_{23}^2}{(\zeta^2+2\eta\rho\zeta+\eta^2)X_0}-\frac{6a_1V_0\xi_1^4\Omega_1^2}{X_0\xi_2^2}+\frac{12\kappa a_1V_0'\xi_1^2\Omega_1^2}{X_0\xi_2^2}-\frac{12\kappa a_1V_0\xi_1^2\Omega_1^2}{X_0\xi_2^2}+\frac{9\eta\kappa\rho a_1^2V_0'V_0\xi_1^3\Omega_1^2}{\xi^2X_0^2\xi_2^4}+\frac{9\zeta\kappa a_1^2V_0'V_0\xi_1^3\Omega_1^2}{\xi^2X_0^2\xi_2^4}-$$

$$\frac{6a_2V_0'\xi_2^2\Omega_2^2}{X_0}+\frac{12ca_2z_3\Omega_2^2}{X_0}-\frac{12ca_2V_0'\Omega_2^2}{X_0}+\frac{9c\zeta\rho a_2^2V_0'z_3\Omega_2^2}{\xi^2X_0^2\xi_2}+\frac{9c\eta a_2^2V_0'z_3\Omega_2^2}{\xi^2X_0^2\xi_2}-\frac{9\zeta\kappa\rho a_1a_2V_0'^2\Omega_2^2}{\xi^2X_0^2\xi_2}-\frac{9\eta\kappa a_1a_2V_0'^2\Omega_2^2}{\xi^2X_0^2\xi_2}-$$

$$\frac{1}{\xi^2X_0^2}[3(8\beta\gamma X_0^2(8\beta\Omega_2^2-\gamma+4\Lambda_1)\xi^2+2X_0(8\beta\Lambda_1\xi^2+16\beta\gamma\xi+(16\beta^2\xi^2-a_2V_0'\xi_2^4)\Omega_2^2)\xi^2+$$

$$a_2V_0'(4a_2\xi_2(V_0'(\xi_2^2+2c)-2cz_3)\Omega_2-a_1(2\kappa V_0'-V_0(\xi_1^2+2\kappa))(\xi_1\Omega_1+\xi_2\Omega_2)))\bar{\sigma}_2^2]-48\gamma^2\Lambda_1-$$

$$\frac{48\gamma\xi^2\Lambda_1}{X_0}-\frac{48\xi^2\Lambda_3}{X_0}-\frac{144\xi^2\Lambda_4}{X_0}+\frac{3\kappa a_1a_2V_0'^2\xi_1\rho_{13}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^2}+\frac{3\kappa a_1^2V_0'V_0\xi_1^3\rho_{13}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^4}-\frac{6c\kappa a_1a_2V_0'z_3\xi_1\rho_{13}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^4}-$$

$$\frac{6\kappa^2a_1^2V_0'^2\xi_1\rho_{13}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^4}+\frac{6c\kappa a_1a_2V_0'^2\xi_1\rho_{13}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^4}+\frac{6\kappa^2a_1^2V_0'V_0\xi_1\rho_{13}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^4}+\frac{3ca_2^2V_0'z_3\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2}-\frac{3\kappa a_1a_2V_0'^2\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2}$$

$$-\frac{6c^2a_2^2z_3^2\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}+\frac{3ca_1a_2V_0z_3\xi_1^2\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}-\frac{3\kappa a_1^2V_0'V_0\xi_1^2\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}+\frac{6c^2a_2^2V_0'z_3\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}+\frac{6c\kappa a_1a_2V_0z_3\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}+$$

$$\frac{6\kappa^2a_1^2V_0'^2\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}-\frac{6c\kappa a_1a_2V_0'^2\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}-\frac{6\kappa^2a_1^2V_0'V_0\rho_{23}}{\xi^2\sqrt{\zeta^2+2\eta\rho\zeta+\eta^2}X_0^2\xi_2^3}+\frac{3a_1a_2V_0'V_0\xi_1^3\Omega_1}{\xi^2X_0^2\xi_2^2}-\frac{9\kappa a_1a_2V_0'^2\xi_1\Omega_1}{\xi^2X_0^2\xi_2^2}+$$

$$\frac{6\kappa a_1a_2V_0'V_0\xi_1\Omega_1}{\xi^2X_0^2\xi_2^2}-\frac{6c\rho a_1a_2V_0z_3\xi_1^2\Omega_1}{\xi^2X_0^2\xi_2^3}+\frac{6\kappa\rho a_1^2V_0'V_0\xi_1^2\Omega_1}{\xi^2X_0^2\xi_2^3}+\frac{3a_1^2V_0^2\xi_1^5\Omega_1}{\xi^2X_0^2\xi_2^4}-\frac{6ca_1a_2V_0z_3\xi_1^3\Omega_1}{\xi^2X_0^2\xi_2^4}+\frac{12\kappa a_1^2V_0^2\xi_1^3\Omega_1}{\xi^2X_0^2\xi_2^4}-$$

$$\frac{21\kappa a_1^2V_0'V_0\xi_1^3\Omega_1}{\xi^2X_0^2\xi_2^4}+\frac{6ca_1a_2V_0'V_0\xi_1^3\Omega_1}{\xi^2X_0^2\xi_2^4}+\frac{12c\kappa a_1a_2V_0'z_3\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}-\frac{12c\kappa a_1a_2V_0z_3\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}+\frac{12\kappa^2a_1^2V_0'^2\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}-\frac{12c\kappa a_1a_2V_0'^2\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}+$$

$$\frac{12\kappa^2a_1^2V_0^2\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}-\frac{24\kappa^2a_1^2V_0'V_0\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}+\frac{12c\kappa a_1a_2V_0'V_0\xi_1\Omega_1}{\xi^2X_0^2\xi_2^4}+\frac{3a_2^2V_0'^2\xi_2\Omega_2}{\xi^2X_0^2}+\frac{4\kappa a_1V_0'\xi_1\Omega_1\Omega_2}{X_0\xi_2}+\frac{9\zeta\kappa\rho a_1a_2V_0'^2\xi_1\Omega_1\Omega_2}{\xi^2X_0^2\xi_2^2}+$$

$$\frac{9\eta\kappa a_1a_2V_0'^2\xi_1\Omega_1\Omega_2}{\xi^2X_0^2\xi_2^2}+\frac{9c\eta\rho a_1a_2V_0z_3\xi_1^2\Omega_1\Omega_2}{\xi^2X_0^2\xi_2^3}+\frac{9c\zeta a_1a_2V_0z_3\xi_1^2\Omega_1\Omega_2}{\xi^2X_0^2\xi_2^3}-\frac{9\eta\kappa\rho a_1^2V_0'V_0\xi_1^2\Omega_1\Omega_2}{\xi^2X_0^2\xi_2^3}-\frac{9\zeta\kappa a_1^2V_0'V_0\xi_1^2\Omega_1\Omega_2}{\xi^2X_0^2\xi_2^3}+$$

$$\frac{3a_1a_2V_0'V_0\xi_1^2\Omega_2}{\xi^2X_0^2\xi_2}-\frac{21ca_2^2V_0'z_3\Omega_2}{\xi^2X_0^2\xi_2}+\frac{12ca_2^2V_0'^2\Omega_2}{\xi^2X_0^2\xi_2}+\frac{3\kappa a_1a_2V_0'^2\Omega_2}{\xi^2X_0^2\xi_2}+\frac{6\kappa a_1a_2V_0'V_0\Omega_2}{\xi^2X_0^2\xi_2}-\frac{6\kappa\rho a_1a_2V_0'^2\xi_1\Omega_2}{\xi^2X_0^2\xi_2^2}+$$

$$\frac{12c^2a_2^2z_3^2\Omega_2}{\xi^2X_0^2\xi_2^3}-\frac{9ca_1a_2V_0z_3\xi_1^2\Omega_2}{\xi^2X_0^2\xi_2^3}+\frac{3\kappa a_1^2V_0'V_0\xi_1^2\Omega_2}{\xi^2X_0^2\xi_2^3}+\frac{6ca_1a_2V_0'V_0\xi_1^2\Omega_2}{\xi^2X_0^2\xi_2^3}-\frac{24c^2a_2^2V_0'z_3\Omega_2}{\xi^2X_0^2\xi_2^3}+\frac{12c\kappa a_1a_2V_0'z_3\Omega_2}{\xi^2X_0^2\xi_2^3}-$$

$$\frac{12c\kappa a_1a_2V_0z_3\Omega_2}{\xi^2X_0^2\xi_2^3}+\frac{12c^2a_2^2V_0'^2\Omega_2}{\xi^2X_0^2\xi_2^3}-\frac{12c\kappa a_1a_2V_0'^2\Omega_2}{\xi^2X_0^2\xi_2^3}+\frac{12c\kappa a_1a_2V_0'V_0\Omega_2}{\xi^2X_0^2\xi_2^3}-\frac{6a_1a_2V_0'V_0\xi_1^3\xi_2^2\Omega_2\bar{\sigma}_1^3\bar{\sigma}_2}{\xi^2X_0^2}-\frac{48\gamma^2\xi}{X_0}),$$

$$\Lambda_9=\frac{1}{96X_0}[3a_1V_0\Omega_1^4\xi_1^4-\frac{6a_1V_0\Omega_1^2\xi_1^4}{\xi_2^2}+\frac{3a_1V_0\xi_1^4}{\xi_2^4}-\frac{8\kappa a_1V_0'\rho_{13}^2\xi_1^2}{(\zeta^2+2\eta\rho\zeta+\eta^2)\xi_2^2}+\frac{8\kappa a_1V_0'\rho_{13}^2\xi_1^2}{\xi_2^4}+\frac{12\kappa a_1V_0'\Omega_1^2\xi_1^2}{\xi_2^2}-$$

$$\frac{12\kappa a_1V_0\Omega_1^2\xi_1^2}{\xi_2^2}-\frac{12\kappa a_1V_0'\xi_1^2}{\xi_2^4}+\frac{12\kappa a_1V_0\xi_1^2}{\xi_2^4}-\frac{4\eta\kappa a_1V_0'\Omega_1\xi_1}{\xi_2^3}-\frac{4\zeta\kappa\rho a_1V_0'\Omega_1\xi_1}{\xi_2^3}+\frac{4\kappa a_1V_0'\Omega_1\Omega_2\xi_1}{\xi_2}+3a_2V_0'\xi_2^4\Omega_2^4+$$

$$(3a_1V_0\xi_1^4-72\alpha^2\xi^2)\bar{\sigma}_1^4+(3a_2V_0'\xi_2^4-72\beta^2\xi^2)\bar{\sigma}_2^4-24\gamma^2\xi^2-72\xi^2\Lambda_1^2+\frac{4\kappa a_1V_0'\rho_{23}^2}{\zeta^2+2\eta\rho\zeta+\eta^2}-\frac{8ca_2z_3\rho_{23}^2}{\zeta^2+2\eta\rho\zeta+\eta^2}-$$

$$\frac{4\kappa a_1V_0'\rho_{23}^2}{\xi_2^2}+\frac{8ca_2z_3\rho_{23}^2}{\xi_2^2}-6a_2V_0'\xi_2^2\Omega_2^2-12ca_2V_0'\Omega_2^2+12ca_2z_3\Omega_2^2-6(8\beta\Lambda_1\xi^2+8\beta(2\beta\Omega_2^2+$$

$$\gamma)\xi^2+a_2\xi_2^2(V_0'((\frac{1}{\xi_2^2}-\Omega_2^2)\xi_2^2+\frac{2c}{\xi_2^2})-\frac{2cz_3}{\xi_2^2}))\bar{\sigma}_2^2+3a_2V_0'-48\gamma\xi^2\Lambda_1-48\xi^2\Lambda_3-144\xi^2\Lambda_4+$$

$$192\alpha\beta\xi^2\Omega_1\Omega_2\bar{\sigma}_1\bar{\sigma}_2-6\bar{\sigma}_1^2(-a_1V_0\Omega_1^2\xi_1^4+\frac{a_1V_0\xi_1^4}{\xi_2^2}-\frac{2\kappa a_1V_0'\xi_1^2}{\xi_2^2}+\frac{2\kappa a_1V_0\xi_1^2}{\xi_2^2}+8\alpha\gamma\xi^2+16\alpha^2\xi^2\Omega_1^2+$$

$$24\alpha\beta\xi^2\bar{\sigma}_2^2+8\alpha\xi^2\Lambda_1)+\frac{12ca_2V_0'}{\xi_2^2}-\frac{12ca_2z_3}{\xi_2^2}-\frac{12\kappa^2a_1V_0'}{\xi_2^4}-\frac{12c\kappa a_1V_0'}{\xi_2^4}+\frac{12c^2a_2V_0'}{\xi_2^4}+\frac{12\kappa^2a_1V_0}{\xi_2^4}+\frac{12c\kappa a_1z_3}{\xi_2^4}-$$

$$\frac{12c^2a_2z_3}{\xi_2^4}].$$

$N(y)$ and $\phi(y)$ are standard normal CDF and PDF respectively, i.e.,

$$N(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad \phi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

The various coefficients are defined as

$$\begin{aligned} &\alpha = \frac{a_1\xi_1^2 V_0}{4\xi X_0},\ \beta = \frac{a_2\xi_2^2 V_0'}{4\xi X_0},\ \delta = -\frac{\xi}{2X_0},\ \zeta = \frac{a_1\xi_1 V_0}{2\sqrt{F(0)}\xi},\ \eta = \frac{a_2\xi_2 V_0'}{2\sqrt{F(0)}\xi}, \\ &\gamma = \frac{1}{4\xi\sqrt{F(0)}\xi_2^2}[2a_1\kappa(V_0' - V_0) - a_1\xi_1^2 V_0 + 2a_2 c(z_3 - V_0') - a_2\xi_2^2 V_0'], \end{aligned} \tag{8.12}$$

and

$$\omega_0(y) = \Lambda_1 y^2 + \gamma, \quad \omega_1(y) = (2\alpha\overline{\sigma}_1\Omega_1 - 2\beta\overline{\sigma}_2\Omega_2)y, \quad \omega_2(y) = \Lambda_0 - \gamma, \tag{8.13}$$

where

$$\begin{aligned} &\Omega_1 = \frac{\zeta + \rho\eta}{\zeta^2 + \eta^2 + 2\rho\zeta\eta}, \qquad \Omega_2 = \frac{\eta + \rho\zeta}{\zeta^2 + \eta^2 + 2\rho\zeta\eta}, \qquad \overline{\sigma}_1^2 = \frac{(1-\rho^2)\eta^2}{\xi_2^2(\zeta^2 + \eta^2 + 2\rho\zeta\eta)}, \\ &\overline{\sigma}_2^2 = \frac{(1-\rho^2)\zeta^2}{\xi_2^2(\zeta^2 + \eta^2 + 2\rho\zeta\eta)}, \quad \rho_{13} = \frac{\zeta + \rho\eta}{\sqrt{\zeta^2 + \eta^2 + 2\rho\zeta\eta}}, \quad \rho_{23} = \frac{\eta + \rho\zeta}{\sqrt{\zeta^2 + \eta^2 + 2\rho\zeta\eta}}, \end{aligned} \tag{8.14}$$

as well as

$$\begin{aligned} &\theta_1 = (-6a_1\kappa\xi_1 V_0 - 3a_1 V_0 \xi_1^3)\frac{1}{\xi_2^2}, \quad \theta_2 = \left[6a_1\kappa\xi_2 V_0' + a_2(-6c\xi_2 V_0' - 3\xi_2^3 V_0')\right]\frac{1}{\xi_2^2}, \\ &\theta_3 = a_1\xi_1^3 V_0, \quad \theta_4 = a_2\xi_2^3 V_0', \quad \theta_5 = 6a_1\kappa\xi_1 V_0', \quad \theta_6 = 6a_2\xi_2 c z_3 - 6a_1\kappa\xi_2 V_0', \\ &\theta_7 = -\frac{\xi\Lambda_0}{X_0}, \quad \theta_8 = -\frac{\xi\Lambda_1}{X_0}. \end{aligned} \tag{8.15}$$

REMARK 10. If choosing $\epsilon = \sqrt{T}$ as the expansion parameter, we keep most of the expressions in Theorem 5. For example, we make the following modifications of our $\mathcal{O}(\epsilon^4)$-expansion for $\epsilon = \xi_2\sqrt{T}$. First, the expansion is expressed as

$$C_0^{VIX} = 100 \times \xi e^{-rT}\left\{\epsilon\Theta_1(y) + \epsilon^2\Theta_2(y) + \epsilon^3\Theta_3(y) + \mathcal{O}(\epsilon^4)\right\}. \tag{8.16}$$

The modifications in the coefficients are

$$\xi = \frac{\sqrt{a_1^2\xi_1^2V_0^2 + a_2^2\xi_2^2V_0'^2 + 2\rho a_1a_2\xi_1\xi_2V_0V_0'}}{2X(0)}, \tag{8.17}$$

and

$$\Lambda_3 = \theta_7 + \frac{1}{12\xi X_0}\left(\theta_1\Omega_1 + \theta_2\Omega_2 + 3\theta_3\Omega_1\overline{\sigma}_1^2 + 3\theta_4\Omega_2\overline{\sigma}_2^2 + \frac{\theta_5\Omega_1}{2} + \frac{\theta_6\Omega_2}{2}\right), \tag{8.18}$$

and

$$\gamma = \frac{1}{4\xi\sqrt{F(0)}}[2a_1\kappa(V_0' - V_0) - a_1\xi_1^2V_0 + 2a_2c(z_3 - V_0') - a_2\xi_2^2V_0'], \tag{8.19}$$

and

$$\overline{\sigma}_1^2 = \frac{(1-\rho^2)\eta^2}{\zeta^2+\eta^2+2\rho\zeta\eta}, \quad \overline{\sigma}_2^2 = \frac{(1-\rho^2)\zeta^2}{\zeta^2+\eta^2+2\rho\zeta\eta}, \tag{8.20}$$

as well as

$$\begin{aligned} \theta_1 =& -6a_1\kappa\xi_1V_0 - 3a_1V_0\xi_1^3, \\ \theta_2 =& 6a_1\kappa\xi_2V_0' + a_2(-6c\xi_2V_0' - 3\xi_2^3V_0'). \end{aligned} \tag{8.21}$$

Due to the similarity, in the subsequent sections, we derive and implement the asymptotic expansion formula in Theorem 5 only.

8.3 Derivation of the Asymptotic Expansion

In this section, we start with scaling the model and then perform a pathwise Taylor expansion around $\epsilon = 0$. Further, we express the expansion of the payoff using generalized Wiener functionals. Finally, we provide a clear road map about explicit computation of each correction term. We begin with scaling the model and then expand it into a pathwise Taylor expansion form. The explicit computation follows from conditional Gaussian mixed moments.

8.3.1 Scaling of the Model

We begin with scaling the model to bring forth the finer local behavior of the diffusion process. We let

$$V^{\epsilon}(t) = V_{\epsilon^2 t}, \quad V'^{\epsilon}(t) = V'_{\epsilon^2 t}, \tag{8.22}$$

for any $\epsilon > 0$. In this dissertation, we choose $\epsilon = \xi_2 \sqrt{T}$.

REMARK 11. A good analog to make sense of this scaling procedure is as follows. Let us recall the construction of Brownian motion $\{W_t\}$ (see Karatzas and Shreve [68]). For a simple random walk M_n, we have

$$\frac{1}{n} M_{[n^2 t]} \longrightarrow W_t.$$

It is intuitive that we obtain a Brownian motion via looking at longer horizon and scaling the magnitude by $\frac{1}{n}$. Now, we do the reverse of this construction procedure. Starting from Brownian motion $\{W_t\}$, we look at a short horizon $\frac{t}{n^2}$ and amplify the magnitude of each moving step by multiplying n times, in this way we recognize much finer local behavior of the Brownian motion. Therefore, a random walk is obtained, i.e.

$$nW_{\frac{t}{n^2}} \sim M_t.$$

Here $\frac{1}{n}$ plays the same role as the scaling parameter ϵ does for our valuation problem. It is natural to expect better analytical tractability for the scaled diffusion (8.22) because we magnify the local behavior of the original diffusion (V_t, V_t').

Thus, the VIX option value can be represented as follows.

$$\begin{aligned} C_0^{VIX} =& e^{-rT}\mathbb{E}^Q\left[\left(\sqrt{a_1 V_T + a_2 V_T' + a_3 z_3} - K\right)^+\right] \\ =& e^{-r\frac{\epsilon^2}{\xi_2^2}}\mathbb{E}^Q\left[\left(\sqrt{a_1 V^\epsilon\left(\frac{1}{\xi_2^2}\right) + a_2 V'^\epsilon\left(\frac{1}{\xi_2^2}\right) + a_3 z_3} - K\right)^+\right]. \end{aligned} \tag{8.23}$$

We derive the dynamics of $(V^\epsilon(t), V'^\epsilon(t))$ by the Brownian scaling property. We also notice that the correlation between Brownian motions is invariant under scaling.

PROPOSITION 11. *Let*

$$B_i(t) = \frac{1}{\epsilon} Z_i(\epsilon^2 t). \tag{8.24}$$

Then,

$$dB_1(t)dB_2(t) = \rho dt.$$

i.e. the correlation between Brownian motions is invariant under scaling.

Proof. By computing the quadratic variation, we find that

$$\begin{aligned} dB_1(t)dB_2(t) =& d\langle B_1, B_2\rangle(t) = d\langle \frac{1}{\epsilon} Z_i(\epsilon^2 \cdot), \frac{1}{\epsilon} Z_i(\epsilon^2 \cdot)\rangle_t \\ =& \left(\frac{1}{\epsilon}\right)^2 d\langle \int_0^{\epsilon^2 \cdot} dZ_1(s), \int_0^{\epsilon^2 \cdot} dZ_2(s)\rangle_t = \rho dt. \end{aligned} \tag{8.25}$$

Alternatively, we arrive at the conclusion from the definition of quadratic variation.

$$\begin{aligned}\langle B_1, B_2\rangle_t =& \lim_{\Pi}\sum_i (B_1(t_{i+1}) - B_1(t_i))(B_2(t_{i+1}) - B_2(t_i)) \\ =& \left(\frac{1}{\epsilon}\right)^2 \lim_{\Pi}\sum_i (Z_1(\epsilon^2 t_{i+1}) - Z_1(\epsilon^2 t_i))(Z_2(\epsilon^2 t_{i+1}) - Z_2(\epsilon^2 t_i)) = \rho t\end{aligned} \tag{8.26}$$

□

Through integral variable substitution, we arrive at the following proposition.

PROPOSITION 12. *The scaled diffusion $\{(V^\epsilon(t), V'^\epsilon(t))\}$ is governed by the following dynamics.*

$$\begin{aligned} dV^\epsilon(t) &= \epsilon^2\kappa(V'^\epsilon(t) - V^\epsilon(t))dt + \epsilon\xi_1 V^\epsilon(t)dB_1(t), \quad V^\epsilon(0) = V_0, \\ dV'^\epsilon(t) &= \epsilon^2 c(z_3 - V'^\epsilon(t))dt + \epsilon\xi_2 V'^\epsilon(t)dB_2(t), \quad V'^\epsilon(0) = V_0', \end{aligned} \tag{8.27}$$

where $\{B_1(t)\}$ and $\{B_2(t)\}$ are two standard Brownian motions with instantaneous correlation ρ.

8.3.2 Derivation of the Asymptotic Expansion Formula

The proof of Theorem 5 is carried out in several steps as follows. We begin with setting up the framework of the asymptotic expansion. Then, a clear road map for computing explicit expansion terms via probabilistic approach follows.

Framework of the Asymptotic Expansion

We first find the Taylor expansion of the bivariate scaled diffusion $(V^\epsilon(t), V'^\epsilon(t))$ around $\epsilon = 0$. In order to facilitate the computation, we write the expansion coefficients as polynomials of the underlying Brownian motion and Lebesgue integrals with respect to time.

PROPOSITION 13. *The bivariate scaled diffusion* $(V^\epsilon(t), V'^\epsilon(t))$ *admits the following Taylor expansion around* $\epsilon = 0$*. For any* $n \in \mathbf{N}$*, we have*

$$
\begin{aligned}
V^\epsilon(t) &= \sum_{n=0}^{N} \frac{1}{n!} \frac{\partial^n V^\epsilon(t)}{\partial \epsilon^n}\bigg|_{\epsilon=0} \epsilon^n + \mathcal{O}(\epsilon^{N+1}), \\
V'^\epsilon(t) &= \sum_{n=0}^{N} \frac{1}{n!} \frac{\partial^n V'^\epsilon(t)}{\partial \epsilon^n}\bigg|_{\epsilon=0} \epsilon^n + \mathcal{O}(\epsilon^{N+1}),
\end{aligned}
\tag{8.28}
$$

where the derivatives of $(V^\epsilon(t), V'^\epsilon(t))$ *w.r.t* ϵ *satisfy that*

$$
\begin{aligned}
V^\epsilon(t)\bigg|_{\epsilon=0} =& V_0, \\
\frac{\partial V^\epsilon(t)}{\partial \epsilon}\bigg|_{\epsilon=0} =& \xi_1 V_0 B_1(t), \\
\frac{\partial^2 V^\epsilon(t)}{\partial \epsilon^2}\bigg|_{\epsilon=0} =& 2\kappa(V_0' - V_0)t + \xi_1^2 V_0 (B_1^2(t) - t), \\
\frac{\partial^3 V^\epsilon(t)}{\partial \epsilon^3}\bigg|_{\epsilon=0} =& 6\kappa V_0' \int_0^t [\xi_2 B_2(u) - \xi_1 B_1(u)] du + t[6\kappa\xi_1 (V_0' - V_0) - 3V_0\xi_1^3] B_1(t) + V_0 \xi_1^3 B_1(t)^3, \\
\frac{\partial^4 V^\epsilon(t)}{\partial \epsilon^4}\bigg|_{\epsilon=0} =& 12\kappa V_0' \int_0^t [\xi_1 B_1(u) - \xi_2 B_2(u)]^2 du + 24\kappa\xi_1 V_0' B_1(t) \int_0^t [\xi_2 B_2(u) - \xi_1 B_1(u)] du \\
& + t[12\kappa\xi_1^2 (V_0' - V_0) - 6V_0\xi_1^4] B_1(t)^2 + V_0\xi_1^4 B_1(t)^4 + t^2[12\kappa^2 (V_0 - V_0') \\
& + 12\kappa\xi_1^2 (V_0 - V_0') + 3V_0\xi_1^4 + 12\kappa c(z_3 - V_0') + 6\kappa V_0'(\xi_1^2 - \xi_2^2)],
\end{aligned}
\tag{8.29}
$$

as well as

$$
\begin{aligned}
V'^{\epsilon}(t)\Big|_{\epsilon=0} =& V'_0,\\
\frac{\partial V'^{\epsilon}(t)}{\partial \epsilon}\Big|_{\epsilon=0} =& \xi_2 V'_0 B_2(t),\\
\frac{\partial^2 V'^{\epsilon}(t)}{\partial \epsilon^2}\Big|_{\epsilon=0} =& 2c(z_3 - V'_0)t + \xi_2^2 V'_0 (B_2^2(t) - t)\\
\frac{\partial^3 V'^{\epsilon}(t)}{\partial \epsilon^3}\Big|_{\epsilon=0} =& -6cz_3\xi_2 \int_0^t B_2(u)du + t[6c\xi_2\left(z_3 - V'_0\right) - 3V'_0\xi_2^3]B_2(t) + V'_0\xi_2^3 B_2(t)^3,\\
\frac{\partial^4 V'^{\epsilon}(t)}{\partial \epsilon^4}\Big|_{\epsilon=0} =& -24cz_3\xi_2^2 B_2(t)\int_0^t B_2(u)du + 12cz_3\xi_2^2\int_0^t B_2(u)^2du + V'_0\xi_2^4 B_2(t)^4\\
& -6t\xi_2^2 B_2(t)^2[2c(V'_0 - z_3) + V'_0\xi_2^2] + 3t^2\left(2c + \xi_2^2\right)[2c(V'_0 - z_3) + V'_0\xi_2^2].
\end{aligned}
\tag{8.30}
$$

Proof. Without loss of generality, we perform the Taylor expansion for $V^{\epsilon}(t)$. We start with

$$
V_t^{\epsilon} = V_0^{\epsilon} + \int_0^t \epsilon^2\kappa(V'^{\epsilon}(u) - V^{\epsilon}(u))du + \int_0^t \epsilon\xi_1 V^{\epsilon}(u)dB_u^{(1)}, \quad V_0^{\epsilon} = V_0. \tag{8.31}
$$

and perform a Taylor expansion on ϵ around $\epsilon = 0$. It is obvious that $V^{\epsilon}(t)\Big|_{\epsilon=0} = V_0$. By differentiating (8.31) on both sides with respect to ϵ, we find that

$$
\begin{aligned}
\frac{\partial V^{\epsilon}(t)}{\partial \epsilon}\Big|_{\epsilon=0} =& \int_0^t 2\epsilon\kappa(V'^{\epsilon}(u) - V^{\epsilon}(u))du + \int_0^t \epsilon^2\kappa\left(\frac{\partial V'^{\epsilon}(u)}{\partial \epsilon} - \frac{\partial V^{\epsilon}(u)}{\partial \epsilon}\right)du\\
&+ \int_0^t \left(\xi_1 V^{\epsilon}(u) + \epsilon\xi_1\frac{\partial V^{\epsilon}(u)}{\partial \epsilon}\right)dB_1(u)\Big|_{\epsilon=0} = \xi_1 V_0 B_1(t).
\end{aligned}
\tag{8.32}
$$

Further differentiation yields that

$$
\frac{\partial^2 V^{\epsilon}(t)}{\partial \epsilon^2}\Big|_{\epsilon=0} = 2\kappa(V'_0 - V_0)t + \xi_1^2 V_0(B_1^2(t) - t)
$$

and

$$\frac{\partial^3 V^\epsilon(t)}{\partial\epsilon^3}\bigg|_{\epsilon=0} = (-6\kappa\xi_1 V_0 t - 3\xi_1^3 V_0 t) B_1(t) + 6\kappa\xi_2 V_0' t B_2(t) + \xi_1^3 V_0 B_1^3(t) \\ + 6\kappa\xi_1 V_0' \int_0^t u dB_1(u) - 6\kappa\xi_2 V_0' \int_0^t u dB_2(u). \tag{8.33}$$

By Ito's Lemma, we obtain the expression in Proposition 13 by rewriting the stochastic integrals (Wiener-Ito chao terms) in (8.33) as Lebesgue integrals. Similarly, we derive the higher order terms. □

REMARK 12. Since the system (8.27) can be solved explicitly as

$$\begin{aligned} V'^\epsilon(t) =& \exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\} \\ & \cdot \left(V_0' + \epsilon^2 c z_3 \int_0^t \exp\left\{\epsilon^2 cu - \epsilon\xi_2 B_2(u) + \frac{1}{2}\epsilon^2\xi_2^2 u\right\} du\right); \\ V^\epsilon(t) =& \exp\left\{-\epsilon^2\kappa t + \epsilon\xi_1 B_1(t) - \frac{1}{2}\epsilon^2\xi_1^2 t\right\} \\ & \cdot \left(V_0 + \epsilon^2\kappa \int_0^t \exp\left\{\epsilon^2\kappa u - \epsilon\xi_1 B_1(u) + \frac{1}{2}\epsilon^2\xi_1^2 u\right\} V'^\epsilon(u) du\right), \end{aligned} \tag{8.34}$$

the Taylor expansion around $\epsilon = 0$ can be obtained via direct differentiation of the above expression.

Next, let us denote

$$F(\epsilon) = a_1 V^\epsilon\left(\frac{1}{\xi_2^2}\right) + a_2 V'^\epsilon\left(\frac{1}{\xi_2^2}\right) + a_3 z_3. \tag{8.35}$$

By Taylor expansion, the underlying random variable in (8.23) reads

$$
\begin{aligned}
X^\epsilon = \sqrt{F(\epsilon)} &= \sqrt{F(0)} + \frac{F'(0)}{2\sqrt{F(0)}}\epsilon + \frac{1}{2}\left[-\frac{F'(0)^2}{4F(0)^{\frac{3}{2}}} + \frac{F''(0)}{2\sqrt{F(0)}}\right]\epsilon^2 \\
&+ \frac{1}{6}\left[\frac{3F'(0)^3}{8F(0)^{\frac{5}{2}}} - \frac{3F'(0)F''(0)}{4F(0)^{\frac{3}{2}}} + \frac{F'''(0)}{2\sqrt{F(0)}}\right]\epsilon^3 \\
&+ \frac{1}{24}\left[\frac{F^{(4)}(0)}{2\sqrt{F(0)}} - \frac{3F''(0)^2}{4F(0)^{3/2}} - \frac{15F'(0)^4}{16F(0)^{7/2}} - \frac{F^{(3)}(0)F'(0)}{F(0)^{3/2}} + \frac{9F'(0)^2F''(0)}{4F(0)^{5/2}}\right]\epsilon^4 + \mathcal{O}(\epsilon^5),
\end{aligned}
\tag{8.36}
$$

where

$$F(0) = a_1V_0 + a_2V_0' + a_3z_3,$$

and for $n \in \mathbf{N}$,

$$F^{(n)}(0) = \left[a_1\frac{\partial^n V^\epsilon}{\partial\epsilon^n}\left(\frac{1}{\xi_2^2}\right) + a_2\frac{\partial^n V'^\epsilon}{\partial\epsilon^n}\left(\frac{1}{\xi_2^2}\right)\right]\Bigg|_{\epsilon=0}. \tag{8.37}$$

Employing a deterministic quantity ξ to be determined for computational purpose, we rewrite that

$$X^\epsilon = \sqrt{F(\epsilon)} = X_0 + \xi(\epsilon X_1 + \epsilon^2 X_2 + \epsilon^3 X_3 + \epsilon^4 X_4 + \mathcal{O}(\epsilon^5)),$$

where

$$
\begin{aligned}
X_0 =& \sqrt{F(0)}, \\
X_1 =& \frac{1}{\xi}\frac{F'(0)}{2\sqrt{F(0)}}, \\
X_2 =& \frac{1}{2\xi}\left[-\frac{F'(0)^2}{4F(0)^{\frac{3}{2}}}+\frac{F''(0)}{2\sqrt{F(0)}}\right], \\
X_3 =& \frac{1}{6\xi}\left[\frac{3F'(0)^3}{8F(0)^{\frac{5}{2}}}-\frac{3F'(0)F''(0)}{4F(0)^{\frac{3}{2}}}+\frac{F'''(0)}{2\sqrt{F(0)}}\right], \\
X_4 =& \frac{1}{24\xi}\left[\frac{F^{(4)}(0)}{2\sqrt{F(0)}}-\frac{3F''(0)^2}{4F(0)^{3/2}}-\frac{15F'(0)^4}{16F(0)^{7/2}}-\frac{F^{(3)}(0)F'(0)}{F(0)^{3/2}}+\frac{9F'(0)^2F''(0)}{4F(0)^{5/2}}\right].
\end{aligned}
\tag{8.38}
$$

We emphasize following inductive algebraic relations, which are important for simplifying the calculation of each correction term based on the ones obtained in previous steps.

$$
\begin{aligned}
X_1 =& \frac{1}{\xi}\frac{F'(0)}{2X_0}, \quad X_2 = \frac{F''(0)}{4\xi X_0}-\frac{\xi X_1^2}{2X_0}, \quad X_3 = \frac{F'''(0)}{12\xi X_0}-\frac{\xi X_1 X_2}{X_0}, \\
X_4 =& \frac{F^{(4)}(0)}{48X_0}-\frac{\xi^2 X_2^2}{2X_0}-\frac{\xi^2 X_1 X_3}{X_0}.
\end{aligned}
\tag{8.39}
$$

We select ξ such that X_1 is a standard normal variable, i.e. $X_1 \sim N(0,1)$. Indeed, we have

$$X_1 = \zeta B\left(\frac{1}{\xi_2^2}\right)+\eta B_2\left(\frac{1}{\xi_2^2}\right),$$

where

$$\zeta = \frac{a_1\xi_1 V_0}{2\xi\sqrt{F(0)}}, \quad \eta = \frac{a_2\xi_2 V_0'}{2\xi\sqrt{F(0)}}.$$

It is easy to find that

$$VarX_1 = \frac{a_1^2\xi_1^2V_0^2+a_2^2\xi_2^2V_0'^2+2\rho a_1a_2\xi_1\xi_2V_0V_0'}{4\xi^2F(0)\xi_2^2}.$$

Thus, ξ is selected according to

$$\xi = \frac{1}{\xi_2}\sqrt{\frac{a_1^2\xi_1^2V_0^2 + a_2^2\xi_2^2V_0'^2 + 2\rho a_1 a_2 \xi_1\xi_2 V_0 V_0'}{4F(0)}}.$$

We rewrite the strike as $K = 100 \times (X_0 + \xi\epsilon y)$. Meanwhile, we define a translated underlying variable

$$Y^\epsilon = \frac{X^\epsilon - X_0}{\epsilon\xi} = X_1 + \epsilon X_2 + \epsilon^2 X_3 + \epsilon^3 X_4 + \mathcal{O}(\epsilon^4). \tag{8.40}$$

REMARK 13. The reason why we define the translated variable Y^ϵ is as follows. We note that $Y^\epsilon \to X_1 \sim N(0,1), \quad \text{as } \epsilon \to 0$ pathwise. The Malliavin non-degeneracy of X_1 allows us to obtain the uniform non-degeneracy of Y^ϵ, which verifies the validity of our asymptotic expansion. (See chapter 10 for detailed discussion.)

Let us denote $f(x) = x^+$. Thus, we have that

$$C_0^{VIX} = 100 \times \xi\epsilon e^{-r\frac{\epsilon^2}{\xi_2^2}}\mathbb{E}^Q f(Y^\epsilon - y).$$

In an appropriate weak sense,

$$\frac{\partial}{\partial x}f(x-y) = 1_{(y,+\infty)}(x), \quad \frac{\partial^2}{\partial x^2}f(x-y) = \delta_y(x), \quad \frac{\partial^3}{\partial x^3}f(x-y) = \delta_y'(x),$$

where δ_y denotes the Dirac function centered at y. By Taylor expansion, we find that

$$f(Y^\epsilon - y) = \sum_{n=0}^{N} \frac{1}{n!}\frac{\partial^{(n)}}{\partial x^n} f(Y^\epsilon - y)\bigg|_{\epsilon=0} \epsilon^n + O(\epsilon^{N+1}).$$

We denote $\Phi_n(y) = \frac{1}{n!}\frac{\partial^{(n)}}{\partial x^n} f(Y^\epsilon - y)\big|_{\epsilon=0}$ and recall that

$$Y^\epsilon = X_1 + \epsilon X_2 + \epsilon^2 X_3 + \epsilon^3 X_4 + O(\epsilon^4),$$

where

$$X_1 = Y^0, \quad X_2 = \left.\frac{\partial Y^\epsilon}{\partial \epsilon}\right|_{\epsilon=0}, \quad X_3 = \frac{1}{2}\left.\frac{\partial^2 Y^\epsilon}{\partial \epsilon^2}\right|_{\epsilon=0}, \quad X_4 = \frac{1}{6}\left.\frac{\partial^3 Y^\epsilon}{\partial \epsilon^3}\right|_{\epsilon=0}.$$

Thus, we find that

$$\begin{aligned}
\Phi_0(y) &= (X_1 - y)^+, \\
\Phi_1(y) &= 1_{(y,+\infty)}(X_1)X_2, \\
\Phi_2(y) &= \frac{1}{2}\left[\delta_y(X_1)X_2^2 + 1_{(y,+\infty)}(X_1)(2X_3)\right], \\
\Phi_3(y) &= \delta_y(X_1)X_2X_3 + \frac{1}{6}\delta_y'(X_1)X_2^3 + 1_{(y,+\infty)}(X_1)X_4.
\end{aligned} \tag{8.41}$$

Next, we calculate the correction terms $\mathbb{E}^Q[\Phi_0(y)]$, $\mathbb{E}^Q[\Phi_1(y)]$, $\mathbb{E}^Q[\Phi_2(y)]$ and $\mathbb{E}^Q[\Phi_3(y)]$ explicitly.

Preliminary Results on Brownian Moments

Our computation relies on the explicit knowledge of moments of the underlying Brownian motion. In this section, we collect several auxiliary results which are useful for the computation of each correction term.

PROPOSITION 14. *For the two-dimensional Brownian motion* $\{B_1(t), B_2(t)\}$ *and* *any $0 < u \le t$, we have the following conditional normal distributions.*

$$\begin{aligned}&\left(\begin{pmatrix} B_1(u) \\ B_2(t) \end{pmatrix} \middle| \zeta B_1(t) + \eta B_2(t) = y\right) \\ \sim N &\left(\begin{pmatrix} \frac{u}{t}\Omega_1 y \\ \Omega_2 y \end{pmatrix}, \begin{pmatrix} u - \frac{u^2}{t}\Omega_1(\zeta + \rho\eta) & \rho u - u\Omega_1(\eta + \rho\zeta) \\ \rho u - u\Omega_1(\eta + \rho\zeta) & t - t\Omega_2(\eta + \rho\zeta) \end{pmatrix}\right),\end{aligned} \tag{8.42}$$

and

$$\begin{aligned}&\left(\begin{pmatrix} B_1(t) \\ B_2(u) \end{pmatrix} \middle| \zeta B_1(t) + \eta B_2(t) = y\right) \\ \sim N &\left(\begin{pmatrix} \Omega_1 y \\ \frac{u}{t}\Omega_2 y \end{pmatrix}, \begin{pmatrix} t - t\Omega_1(\zeta + \rho\eta) & \rho u - u\Omega_1(\eta + \rho\zeta) \\ \rho u - u\Omega_1(\eta + \rho\zeta) & u - \frac{u^2}{t}\Omega_1(\eta + \rho\zeta) \end{pmatrix}\right),\end{aligned} \tag{8.43}$$

and

$$\begin{aligned}&\left(\begin{pmatrix} B_2(u) \\ B_2(t) \end{pmatrix} \middle| \zeta B_1(t) + \eta B_2(t) = y\right) \\ \sim N &\left(\begin{pmatrix} \frac{u}{t}\Omega_2 y \\ \Omega_2 y \end{pmatrix}, \begin{pmatrix} u - \frac{u^2}{t}\Omega_2(\eta + \rho\zeta) & u - u\Omega_2(\eta + \rho\zeta) \\ u - u\Omega_2(\eta + \rho\zeta) & t - t\Omega_2(\eta + \rho\zeta) \end{pmatrix}\right),\end{aligned} \tag{8.44}$$

and

$$\begin{aligned}&\left(\begin{pmatrix} B_1(u) \\ B_1(t) \end{pmatrix} \middle| \zeta B_1(t) + \eta B_2(t) = y\right) \\ \sim N &\left(\begin{pmatrix} \frac{u}{t}\Omega_1 y \\ \Omega_1 y \end{pmatrix}, \begin{pmatrix} u - \frac{u^2}{t}\Omega_1(\zeta + \rho\eta) & u - u\Omega_1(\zeta + \rho\eta) \\ u - u\Omega_1(\zeta + \rho\eta) & t - t\Omega_1(\zeta + \rho\eta) \end{pmatrix}\right),\end{aligned} \tag{8.45}$$

where the various coefficients are defined in Theorem 5.

Without loss of generality, we verify the distributional property (8.42) for the case $u = t = \frac{1}{\xi_2^2}$, which is the key for deriving the first three correction terms. Let us denote

$$Z_1 = \begin{pmatrix} B_1\left(\frac{1}{\xi_2^2}\right) \\ B_2\left(\frac{1}{\xi_2^2}\right) \end{pmatrix}, \quad Z_2 = \zeta B_1\left(\frac{1}{\xi_2^2}\right) + \eta B_2\left(\frac{1}{\xi_2^2}\right).$$

We note that $(Z_1, Z_2)^T$ is a three-dimensional normal variable $N(\mu, \Sigma)$, where

$$\mu = (\mu_1, \mu_2)^T = ((0,0)^T, 0)^T$$

and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \frac{1}{\xi_2^2} \begin{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} & \begin{pmatrix} \zeta + \rho\eta \\ \eta + \rho\zeta \end{pmatrix} \\ \begin{pmatrix} \zeta + \rho\eta & \eta + \rho\zeta \end{pmatrix} & \zeta^2 + \eta^2 + 2\rho\zeta\eta \end{pmatrix}.$$

Thus, we have the following lemma, which appears as a special case of Proposition 14.

LEMMA 7.

$$(Z_1 | Z_2 = y) \sim N(\overline{\mu}(y), \overline{\Sigma}),$$

where

$$\overline{\mu} = \begin{pmatrix} \overline{\mu}_1(y) \\ \overline{\mu}_2(y) \end{pmatrix} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y - \mu_2) = \frac{y}{\zeta^2 + \eta^2 + 2\rho\zeta\eta} \begin{pmatrix} \zeta + \rho\eta \\ \eta + \rho\zeta \end{pmatrix} \tag{8.46}$$

and

$$\overline{\Sigma} = (\overline{\Sigma}_{ij})_{2\times 2} = \begin{pmatrix} \overline{\sigma}_1^2 & \overline{\rho\sigma}_1\overline{\sigma}_2 \\ \overline{\rho\sigma}_1\overline{\sigma}_2 & \overline{\sigma}_2^2 \end{pmatrix} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \frac{1}{\xi_2^2}\frac{1-\rho^2}{\zeta^2+\eta^2+2\rho\zeta\eta}\begin{pmatrix} \eta^2 & -\zeta\eta \\ -\zeta\eta & \zeta^2 \end{pmatrix}, \tag{8.47}$$

with $\overline{\rho} = -1$.

Based on Proposition 14 and Lemma 7, we are able to compute explicitly some Gaussian moments, which are important to the calculation of the correction terms. For example, we consider the case of $u = t = 1/\xi_2^2$. The conditional moment generating function of $(B_1(t), B_2(t))$ given $X_1 = \zeta B_1(t) + \eta B_2(t) = y$ is

$$M(\vartheta_1, \vartheta_2) := \exp\left\{\overline{\mu}_1(y)\vartheta_1 + \overline{\mu}_2(y)\vartheta_2 + \frac{1}{2}\vartheta_1^2\overline{\Sigma}_{11} + \vartheta_1\vartheta_2\overline{\Sigma}_{12} + \frac{1}{2}\vartheta_2^2\overline{\Sigma}_{22}\right\}, \quad \vartheta_1,\ \vartheta_2 \in \mathbf{R}.$$

Thus, the conditional moment satisfies

$$M_{ij} := E[B_1(t)^i B_2(t)^j | \zeta B_1(t) + \eta B_2(t) = y] = \frac{\partial^{i+j}}{\partial^i\vartheta_1 \partial^j\vartheta_2} M(\vartheta_1, \vartheta_2)\Big|_{\vartheta_1=\vartheta_2=0}.$$

Using this idea, we explicitly compute the following conditional Gaussian moments, which are useful for the derivation of each correction term.

COROLLARY 5. For the two-dimensional Brownian motion $\{B_1(t), B_2(t)\}$, the following identities on conditional moments hold.

$$
\begin{aligned}
\mathbb{E}\left(B_1\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x\right)=&\Omega_1 x,\\
\mathbb{E}(B_1^2\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_1^2 x^2+\overline{\sigma}_1^2,\\
\mathbb{E}(B_1^3\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_1^3 x^3+3\overline{\sigma}_1^2\Omega_1 x,\\
\mathbb{E}(B_1^4\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_1^4 x^4+6\overline{\sigma}_1^2\Omega_1^2 x^2+3\overline{\sigma}_1^4,\\
\mathbb{E}(B_1^5\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_1^5 x^5+10\overline{\sigma}_1^2\Omega_1^3 x^3+15\Omega_1\overline{\sigma}_1^4 x,\\
\mathbb{E}\left(B_2\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x\right)=&\Omega_2 x,\\
\mathbb{E}(B_2^2\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_2^2 x^2+\overline{\sigma}_2^2,\\
\mathbb{E}(B_2^3\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_2^3 x^3+3\overline{\sigma}_2^2\Omega_2 x,\\
\mathbb{E}(B_2^4\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_2^4 x^4+6\overline{\sigma}_2^2\Omega_2^2 x^2+3\overline{\sigma}_2^4,\\
\mathbb{E}(B_2^5\left(t\right)|\zeta B_1(t)+\eta B_2(t)=x)=&\Omega_2^5 x^5+10\overline{\sigma}_2^2\Omega_2^3 x^3+15\Omega_2\overline{\sigma}_2^4 x,
\end{aligned}
\tag{8.48}
$$

and

$$
\begin{aligned}
\mathbb{E}(B_1(u)|\zeta B_1(t)+\eta B_2(t)=x)=&\frac{u}{t}\frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\rho_{13},\\
\mathbb{E}(B_1^2(u)|\zeta B_1(t)+\eta B_2(t)=x)=&\rho_{13}^2[\frac{u(t-u)}{t}+(\frac{u}{t})^2\frac{x^2}{\zeta^2+\eta^2+2\rho\zeta\eta}]+(1-\rho_{13}^2)u,\\
\mathbb{E}(B_2(u)|\zeta B_1(t)+\eta B_2(t)=x)=&\frac{u}{t}\frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\rho_{23},\\
\mathbb{E}(B_2^2(u)|\zeta B_1(t)+\eta B_2(t)=x)=&\rho_{23}^2[\frac{u(t-u)}{t}+(\frac{u}{t})^2\frac{x^2}{\zeta^2+\eta^2+2\rho\zeta\eta}]+(1-\rho_{23}^2)u,
\end{aligned}
\tag{8.49}
$$

and

$$
\begin{aligned}
\mathbb{E}(B_1(t)^2B_2(t)|\zeta B_1(t)+\eta B_2(t)=x) =&\Omega_1^2\Omega_2x^3-2\overline{\sigma}_1\overline{\sigma}_2\Omega_1x+\overline{\sigma}_1^2\Omega_2x,\\
\mathbb{E}(B_1(t)B_2(t)^2|\zeta B_1(t)+\eta B_2(t)=x) =&\Omega_1\Omega_2^2x^3-2\overline{\sigma}_1\overline{\sigma}_2\Omega_2x+\overline{\sigma}_2^2\Omega_1x,\\
\mathbb{E}(B_1(t)^3B_2(t)^2|\zeta B_1(t)+\eta B_2(t)=x) =&-6\overline{\sigma}_1\overline{\sigma}_2\Omega_1^2\Omega_2x^3+(\overline{\sigma}_2^2+x^2\Omega_2^2)\Omega_1^3x^3+6\overline{\sigma}_1^2\overline{\sigma}_2^2\Omega_1x\\
&-6\overline{\sigma}_1^3\overline{\sigma}_2\Omega_2x+3\overline{\sigma}_1^2\Omega_1x(\overline{\sigma}_2^2+x^2\Omega_2^2),\\
\mathbb{E}(B_1(t)^2B_2(t)^3|\zeta B_1(t)+\eta B_2(t)=x) =&-6\overline{\sigma}_1\overline{\sigma}_2\Omega_1\Omega_2^2x^3+(\overline{\sigma}_1^2+x^2\Omega_1^2)\Omega_2^3x^3+6\overline{\sigma}_1^2\overline{\sigma}_2^2\Omega_2x\\
&-6\overline{\sigma}_1\overline{\sigma}_2^3\Omega_1x+3\overline{\sigma}_2^2\Omega_2x(\overline{\sigma}_1^2+x^2\Omega_1^2),
\end{aligned}
\tag{8.50}
$$

and

$$
\begin{aligned}
\mathbb{E}(B_1(u)B_2(u)|\zeta B_1(t)+\eta B_2(t)=x) =&\rho u-\frac{(\eta+\zeta\rho)\Omega_1u^2}{t}+\frac{x^2\Omega_1\Omega_2u^2}{t^2},\\
\mathbb{E}(B_1(t)B_2(u)|\zeta B_1(t)+\eta B_2(t)=x) =&u\rho-u(\eta+\zeta\rho)\Omega_1+\frac{u\Omega_1\Omega_2x^2}{t},\\
\mathbb{E}(B_1(t)B_1(u)|\zeta B_1(t)+\eta B_2(t)=x) =&\frac{u\rho_{13}^2}{t}\frac{x^2}{\zeta^2+\eta^2+2\rho\zeta\eta}+u(1-\rho_{13}^2),\\
\mathbb{E}(B_2(t)B_2(u)|\zeta B_1(t)+\eta B_2(t)=x) =&\frac{u\rho_{23}^2}{t}\frac{x^2}{\zeta^2+\eta^2+2\rho\zeta\eta}+u(1-\rho_{23}^2),
\end{aligned}
\tag{8.51}
$$

as well as

$$
\begin{aligned}
\mathbb{E}(B_1(t)^2 B_1(u)|\zeta B_1(t)+\eta B_2(t)=x) =& 2x\Omega_1\left(u-u(\zeta+\eta\rho)\Omega_1\right) \\
&+\frac{ux\Omega_1\left(x^2\Omega_1^2-t(\zeta+\eta\rho)\Omega_1+t\right)}{t}, \\
\mathbb{E}(B_1(t)^2 B_2(u)|\zeta B_1(t)+\eta B_2(t)=x) =& 2x\Omega_1\left(u\rho-u(\zeta+\eta\rho)\Omega_2\right) \\
&+\frac{ux\Omega_2\left(x^2\Omega_1^2-t(\zeta+\eta\rho)\Omega_1+t\right)}{t}, \\
\mathbb{E}(B_1(u)B_2(t)^2|\zeta B_1(t)+\eta B_2(t)=x) =& 2x\Omega_2\left(u\rho-u(\eta+\zeta\rho)\Omega_1\right) \\
&+\frac{ux\Omega_1\left(x^2\Omega_2^2-t(\eta+\zeta\rho)\Omega_2+t\right)}{t}, \\
\mathbb{E}(B_2(u)B_2(t)^2|\zeta B_1(t)+\eta B_2(t)=x) =& 2x\Omega_2\left(u-u(\eta+\zeta\rho)\Omega_2\right) \\
&+\frac{ux\Omega_2\left(x^2\Omega_2^2-t(\eta+\zeta\rho)\Omega_2+t\right)}{t},
\end{aligned}
\tag{8.52}
$$

where the various coefficients are defined in Theorem 5.

Subsequently, we calculate each correction term explicitly. We exhibit the framework and techniques employed in the computation as follows.

Calculation of the Leading Order Term: $\mathbb{E}^Q[\Phi_0(y)]$

It is easy to find that

$$
\mathbb{E}^Q[\Phi_0(y)] = \mathbb{E}^Q(X_1-y)^+ = \int_{\mathbb{R}} (x-y)^+\phi(x)dx = yN(y)+\phi(y)-y. \tag{8.53}
$$

Calculation of the Second Correction Term: $\mathbb{E}^Q[\Phi_1(y)]$

To begin with, we observe that

$$
\begin{aligned}
\mathbb{E}^Q[\Phi_1(y)] =& \mathbb{E}^Q[1_{(y,+\infty)}(X_1)X_2] = \int_{\mathbb{R}} \mathbb{E}^Q[1_{(y,+\infty)}(x)X_2|X_1=x]\phi(x)dx \\
=& \int_y^\infty \mathbb{E}^Q[X_2|X_1=x]\phi(x)dx.
\end{aligned}
\tag{8.54}
$$

Thus, we focus on the calculation of conditional expectation $\mathbb{E}^Q[X_2|X_1 = x]$ as follows. We recall that

$$X_1 = \zeta B_1 \left(\frac{1}{\xi_2^2}\right) + \eta B_2 \left(\frac{1}{\xi_2^2}\right).$$

By algebraic computation, we find that

$$X_2 = -\frac{\xi^2 X_1^2}{2\xi X_0} + \frac{F'''(0)}{4\xi X_0} = \alpha \left(B_1\left(\frac{1}{\xi_2^2}\right)\right)^2 + \beta \left(B_2\left(\frac{1}{\xi_2^2}\right)\right)^2 + \gamma + \delta X_1^2,$$

where $\alpha, \beta, \gamma, \delta, \zeta, \eta$ are defined in Theorem 5. Hence, it follows from the computation of moments based on Lemma 7 that

$$\mathbb{E}^Q[X_2|X_1 = x] = \alpha(\overline{\sigma}_1^2 + \mu_1^2) + \beta(\overline{\sigma}_2^2 + \mu_2^2) + \gamma + \delta x^2 = \Lambda_0 + \Lambda_1 x^2,$$

where

$$\Lambda_0 = \alpha\overline{\sigma}_1^2 + \beta\overline{\sigma}_2^2 + \gamma$$

and

$$\Lambda_1 = \alpha\Omega_1^2 + \beta\Omega_2^2 + \delta.$$

Finally, we compute explicitly that

$$\mathbb{E}^Q\Phi_1(y) = \int_y^\infty (\Lambda_0 + \Lambda_1 x^2)\phi(x)dx = (\Lambda_0 + \Lambda_1)(1 - N(y)) + \Lambda_1 y\phi(y).$$

Calculation of the Third Correction Term: $\mathbb{E}^Q[\Phi_2(y)]$

By conditioning, we have that

$$
\begin{aligned}
\mathbb{E}^Q\Phi_2(y) =& \frac{1}{2}\int_{\mathbb{R}} \mathbb{E}\left[\delta_y(X_1)X_2^2 + 1_{(y,+\infty)}(X_1)(2X_3)|X_1 = x\right]\phi(x)dx \\
=& \frac{1}{2}\mathbb{E}(X_2^2|X_1 = y)\phi(y) + \int_y^{\infty} \mathbb{E}(X_3|X_1 = x)\phi(x)dx.
\end{aligned}
\tag{8.55}
$$

Denote the two-dimensional normal variable $(\overline{B}_1, \overline{B}_2) \sim N(\overline{\mu}, \overline{\Sigma})$. Thus, it is easy to see that

$$
\mathbb{E}(X_2^2|X_1 = y) = \mathbb{E}(\alpha\overline{B}_1^2 + \beta\overline{B}_2^2 + \gamma + \delta y^2)^2.
$$

By Lemma 7, we decompose the correlated normal variables as:

$$
\begin{aligned}
\overline{B}_1 =& \overline{\sigma}_1 W_1 + \overline{\mu}_1, \\
\overline{B}_2 =& \overline{\sigma}_2(\overline{\rho}W_1 + \sqrt{1-\overline{\rho}^2}W_2) + \overline{\mu}_2 = -\overline{\sigma}_2 W_1 + \overline{\mu}_2,
\end{aligned}
\tag{8.56}
$$

where $W_1 \sim N(0,1)$. Using the fundamental fact

$$
\mathbb{E}W_i = 0,\ \mathbb{E}W_i^2 = 1,\ \mathbb{E}W_i^3 = 0,\ \mathbb{E}W_i^4 = 3, \quad \text{for i=1, 2,}
$$

we find that

$$
\Lambda_2 := \mathbb{E}(X_2^2|X_1 = y) = 3\omega_2^2(y) + \omega_1^2(y) + \omega_0^2(y) + 2\omega_0(y)\omega_2(y), \tag{8.57}
$$

where

$$
\omega_0(y) = (\alpha\Omega_1^2 + \beta\Omega_2^2 + \delta)y^2 + \gamma, \quad \omega_1(y) = (2\alpha\overline{\sigma}_1\Omega_1 - 2\beta\overline{\sigma}_2\Omega_2)y, \quad \omega_2(y) = \alpha\overline{\sigma}_1^2 + \beta\overline{\sigma}_2^2. \tag{8.58}
$$

Now, we compute the conditional expectation $\mathbb{E}(X_3|X_1 = x)$. It is straightforward to find that

$$
\mathbb{E}(X_3|X_1 = x) = \theta_7 x + \theta_8 x^3 + \frac{1}{12\xi X_0}\mathbb{E}(F'''(0)|X_1 = x)
$$

where

$$F'''(0) = \theta_1 B_1\left(\frac{1}{\xi_2^2}\right)+\theta_2 B_2\left(\frac{1}{\xi_2^2}\right)+\theta_3 B_1^3\left(\frac{1}{\xi_2^2}\right)+\theta_4 B_2^3\left(\frac{1}{\xi_2^2}\right)+\theta_5\int_0^{\frac{1}{\xi_2^2}} udB_1(u)+\theta_6\int_0^{\frac{1}{\xi_2^2}} udB_2(u),$$

with the coefficients defined in the Theorem 5. By Corollary 5, we have that

$$\begin{aligned}
&\mathbb{E}\left(B_1\left(\frac{1}{\xi_2^2}\right)\middle|X_1=x\right)=\Omega_1 x,\\
&\mathbb{E}\left(B_2\left(\frac{1}{\xi_2^2}\right)\middle|X_1=x\right)=\Omega_2 x,\\
&\mathbb{E}\left(B_1^3\left(\frac{1}{\xi_2^2}\right)\middle|X_1=x\right)=\overline{\mu}_1^3+3\overline{\mu}_1\overline{\sigma}_1^2=\Omega_1^3x^3+3\overline{\sigma}_1^2\Omega_1 x,\\
&\mathbb{E}\left(B_2^3\left(\frac{1}{\xi_2^2}\right)\middle|X_1=x\right)=\overline{\mu}_2^3+3\overline{\mu}_2\overline{\sigma}_2^2=\Omega_2^3x^3+3\overline{\sigma}_2^2\Omega_2 x,\\
&\mathbb{E}\left(\int_0^{\frac{1}{\xi_2^2}} udB_1(u)\middle|X_1=x\right)=\frac{\rho_{13}}{2\xi_2^2}\frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}=\frac{\Omega_1}{2\xi_2^2}x,\\
&\mathbb{E}\left(\int_0^{\frac{1}{\xi_2^2}} udB_2(u)\middle|X_1=x\right)=\frac{\rho_{23}}{2\xi_2^2}\frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}=\frac{\Omega_2}{2\xi_2^2}x,
\end{aligned}\tag{8.59}$$

where

$$\Omega_1=\frac{\zeta+\rho\eta}{\zeta^2+\eta^2+2\rho\zeta\eta},\quad \Omega_2=\frac{\eta+\rho\zeta}{\zeta^2+\eta^2+2\rho\zeta\eta},$$

and

$$\rho_{13}=\frac{\zeta+\rho\eta}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}},\quad \rho_{23}=\frac{\eta+\rho\zeta}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}.$$

REMARK 14. (*A Brownian-Bridge-Based Calculation*) The above identities can be alternatively obtained from the Cholesky decomposition of correlated Brownian motions and the properties of Brownian bridge. We let

$$B_3(w)=\frac{\zeta B_1(w)+\eta B_2(w)}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}},$$

which forms a standard Brownian motion $\{B_3(t)\}$ satisfying

$$Corr(dB_1, dB_3) = \rho_{13}dt, \quad Corr(dB_2, dB_3) = \rho_{23}dt.$$

Thus,

$$\mathbb{E}\left(\int_0^{\frac{1}{\xi_2^2}} udB_i(u)\middle|X_1 = x\right) = \mathbb{E}\left(\int_0^w udB_i(u)\middle|B_3(w) = \frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\right), \quad (8.60)$$

where $w = \frac{1}{\xi_2^2}$. To compute this expectation, we first observe that

$$\int_0^w udB_i(u) = wB_i(w) - \int_0^w B_i(u)du. \quad (8.61)$$

It is easy to obtain that, for $i = 1, 2$,

$$B_i(t) = \rho_{i3}B_3(t) + \sqrt{1-\rho_{i3}^2}B_i'(t), \quad (8.62)$$

where $\{B_i'(t)\}$ is a Brownian motion independent of $\{B_3(t)\}$. So, equation (8.60) can be further deduced from interchanging the order of conditional expectation and the Lebesgue integration on time, i.e.

$$\begin{aligned}
&\mathbb{E}\left(\int_0^{\frac{1}{\xi_2^2}} udB_i(u)\middle|X_1 = x\right)\\
=&\mathbb{E}\left(wB_i(w)\middle|B_3(w) = \frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\right) - \mathbb{E}\left(\int_0^w B_i(u)du\middle|B_3(w) = \frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\right)\\
=&w\mathbb{E}\left(B_i(w)\middle|B_3(w) = \frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\right) - \int_0^w \mathbb{E}\left(B_i(u)\middle|B_3(w) = \frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\right)du.
\end{aligned} \quad (8.63)$$

Based on the property of Brownian bridge, we observe that

$$\mathbb{E}\left(B_3(u)\middle|B_3(w)=\frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}\right)=\frac{u}{w}\frac{x}{\sqrt{\zeta^2+\eta^2+2\rho\zeta\eta}}. \tag{8.64}$$

This, we complete the computation of (8.60).

Finally, it follows that

$$E(X_3|X_1=x)=\Lambda_3 x+\Lambda_4 x^3,$$

where

$$\begin{aligned}\Lambda_3&=\theta_7+\frac{1}{12\xi X_0}\left(\theta_1\Omega_1+\theta_2\Omega_2+3\theta_3\Omega_1\overline{\sigma}_1^2+3\theta_4\Omega_2\overline{\sigma}_2^2+\frac{\theta_5\Omega_1}{2\xi_2^2}+\frac{\theta_6\Omega_2}{2\xi_2^2}\right),\\ \Lambda_4&=\theta_8+\frac{1}{12\xi X_0}\left(\theta_3\Omega_1^3+\theta_4\Omega_2^3\right).\end{aligned} \tag{8.65}$$

Hence, by explicit integration we find that

$$\int_y^{+\infty}E(X_3|X_1=x)\phi(x)dx=\int_y^{+\infty}(\Lambda_3 x+\Lambda_4 x^3)\phi(x)dx=\phi(y)(\Lambda_3+(2+y^2)\Lambda_4).$$

Calculation of the Fourth Correction Term: $\mathbb{E}^Q[\Phi_3(y)]$

We briefly outline the framework and techniques for computing the fourth order correction term in this section. First, by conditioning, we have that

$$\mathbb{E}\Phi_3(y)=\int_{-\infty}^{+\infty}E(\Phi_3(y)|X_1=x)\phi(x)dx, \tag{8.66}$$

where

$$\begin{aligned}\mathbb{E}(\Phi_3(y)|X_1=x)=&\mathbb{E}(\delta_y(X_1)X_2X_3|X_1=x)\\ &+\frac{1}{6}\mathbb{E}(\delta_y'(X_1)X_2^3|X_1=x)+\mathbb{E}(1_{(y,+\infty)}(X_1)X_4|X_1=x).\end{aligned} \tag{8.67}$$

Therefore, we find that

$$\mathbb{E}\Phi_3(y) = A(y) + B(y) + C(y),$$

where

$$\begin{aligned} A(y) &= \mathbb{E}(X_2X_3|X_1 = y)\phi(y), \\ B(y) &= -\frac{1}{6}\frac{\partial}{\partial y}\{\mathbb{E}(X_2^3|X_1 = y)\phi(y)\}, \\ C(y) &= \int_y^{+\infty} \mathbb{E}(X_4|X_1 = x)\phi(x)dx. \end{aligned} \tag{8.68}$$

Here, we have used the following integration by parts formula for Dirac Delta function, i.e. for any function g and $n \in \mathbb{N}$,

$$\int_{-\infty}^{+\infty} g(x)\delta_y^{(n)}(x)dx = -\int_{-\infty}^{+\infty} \frac{\partial g}{\partial x}(x)\delta_y^{(n-1)}(x)dx.$$

Next, we compute $A(y), B(y)$ and $C(y)$, respectively.

We begin with the following inductive algebraic relations, which allow us to make use of the previous computation results to derive new ones. These observations are helpful for reducing the computational load.

$$X_2 = \frac{F''(0)}{4\xi X_0} - \frac{\xi X_1^2}{2X_0}, \quad X_3 = \frac{F'''(0)}{12\xi X_0} - \frac{\xi X_1 X_2}{X_0}, \quad X_4 = \frac{F^{(4)}(0)}{48X_0} - \frac{\xi^2 X_2^2}{2X_0} - \frac{\xi^2 X_1 X_3}{X_0}. \tag{8.69}$$

For term $A(y)$, we have that

$$\begin{aligned} A(y) =& \Bigg\{ \frac{1}{12\xi X_0}\left[\frac{1}{4\xi X_0}\mathbb{E}(F''(0)F'''(0)|X_1 = y) - \frac{\xi y^2}{2X_0}\mathbb{E}(F'''(0)|X_1 = y)\right] \\ & - \frac{\xi y}{X_0}\mathbb{E}(X_2^2|X_1 = y)\Bigg\}\phi(y), \end{aligned} \tag{8.70}$$

where

$$\mathbb{E}(X_2^2|X_1=y)=3\omega_2^2(y)+\omega_1^2(y)+\omega_0^2(y)+2\omega_0(y)\omega_2(y),$$

and

$$\mathbb{E}(F'''(0)|X_1=y)=12\xi X_0[(\Lambda_3-\theta_7)y+(\Lambda_4-\theta_8)y^3],$$

and $\mathbb{E}(F''(0)F'''(0)|X_1=y)$ is a new quantity to compute. Based on (8.37), $F''(0)F'''(0)$ can be expanded explicitly as a combination of polynomials and Lebesgue integrals, with respect to time, of the underlying Brownian motions. Thus, we accomplish the explicit computation of conditional moments by using the identities in Corollary 5. Without loss of generality, we demonstrate the computation on one of such terms. For example, we need to compute

$$\mathbb{E}\left[B_1(t)^2\int_0^t B_2(u)du\middle|\zeta B_1(t)+\eta B_2(t)=x\right].$$

Indeed, this conditional expectation is equal to

$$\begin{aligned}
&\int_0^t \mathbb{E}(B_1(t)^2B_2(u)|\zeta B_1(t)+\eta B_2(t)=x)du\\
=&\int_0^t\left[2x\Omega_1\left(u\rho-u(\zeta+\eta\rho)\Omega_2\right)+\frac{ux\Omega_2\left(x^2\Omega_1^2-t(\zeta+\eta\rho)\Omega_1+t\right)}{t}\right]du \qquad (8.71)\\
=&\frac{1}{2}t\Omega_1^2\Omega_2y^3+\left(t^2\rho\Omega_1+\frac{1}{2}t^2\Omega_2-\frac{3}{2}t^2\zeta\Omega_1\Omega_2-\frac{3}{2}t^2\eta\rho\Omega_1\Omega_2\right)y.
\end{aligned}$$

For term $B(y)$, we observe that

$$
\begin{aligned}
&\mathbb{E}^Q(X_2^3|X_1=y)\\
=&\mathbb{E}(\alpha\overline{B}_1^2+\beta\overline{B}_2^2+\gamma+\delta y^2)^3\\
=&\mathbb{E}(\alpha(\overline{\sigma}_1W_1+\overline{\mu}_1)^2+\beta(-\overline{\sigma}_2W_1+\overline{\mu}_2)^2+\gamma+\delta y^2)^3\\
=&\mathbb{E}(\omega_2(y)W_1^2+\omega_1(y)W_1+\omega_0(y))^3\\
=&15\omega_2(y)^3+9\omega_0(y)\omega_2(y)^2+3(\omega_0(y)^2\omega_2(y)+\omega_0(y)\omega_1(y)^2)+\omega_0(y)^3.
\end{aligned}
\tag{8.72}
$$

For term $C(y)$, we observe that

$$
\mathbb{E}(X_4|X_1=x)=\frac{1}{48X_0}\mathbb{E}(F^{(4)}(0)|X_1=x)-\frac{\xi^2}{2X_0}\mathbb{E}(X_2^2|X_1=x)-\frac{\xi^2x}{X_0}\mathbb{E}(X_3|X_1=x),
$$

where it is known from the previous computation that

$$
\mathbb{E}(X_2^2|X_1=x)=3\omega_2^2(x)+\omega_1^2(x)+\omega_0^2(x)+2\omega_0(x)\omega_2(x),
$$

and that

$$
\mathbb{E}(X_3|X_1=x)=\Lambda_3x+\Lambda_4x^3.
$$

For any constants $c_i,\ i=1,...,5$, we have that

$$
\begin{aligned}
&\int_y^{+\infty}(c_0+c_1x+c_2x^2+c_3x^3+c_4x^4)\phi(x)dx\\
=&(c_0+c_2+3c_4)(1-N(y))+\phi(y)[(c_1+2c_3)+(c_2+3c_4)y+c_3y^2+c_4y^3].
\end{aligned}
\tag{8.73}
$$

This fact renders an explicit calculation of term $C(y)$.

Finally, the price of VIX option C_0^{VIX} (see (8.23)) admits the following asymptotic expansion:

$$C_0^{VIX} = 100 \times \xi e^{-r\frac{\epsilon^2}{\xi_2^2}} \left\{ \epsilon\Theta_1(y) + \epsilon^2\Theta_2(y) + \epsilon^3\Theta_3(y) + \epsilon^4\Theta_4(y) + \mathcal{O}(\epsilon^5) \right\}, \tag{8.74}$$

where

$$\begin{aligned}
\Theta_1(y) =& \phi(y) - y(1 - N(y)), \\
\Theta_2(y) =& (\Lambda_0 + \Lambda_1)(1 - N(y)) + \Lambda_1 y\phi(y), \\
\Theta_3(y) =& \frac{1}{2}\Lambda_2\phi(y) + \phi(y)(\Lambda_3 + (2 + y^2)\Lambda_4), \\
\Theta_4(y) =& \left(\Lambda_5 y^7 + \Lambda_6 y^5 + \Lambda_7 y^3 + \Lambda_8 y\right)\phi(y) + 2\Lambda_9(1 - N(y)).
\end{aligned} \tag{8.75}$$

Chapter 9

Implementation and Numerical Examples

In this chapter, we demonstrate the efficiency of our asymptotic expansion formula in Theorem 5 with some numerical experiments. First, we employ a bias-corrected Monte Carlo simulation scheme as a benchmark, to which the implementation of the formula is compared. We recall that

$$C_0^{VIX} = 100 \times e^{-rT}\mathbb{E}\left(\sqrt{a_1 V_T + a_2 V_T' + a_3 z_3} - \frac{K}{100}\right)^+.$$

9.1 Benchmark from Monte Carlo Simulation: Euler Scheme with Partial Truncation

Due to the positivity of the diffusion processes (V_t, V_t'), we adopt a biased-corrected Euler discretization scheme of partial truncation (see Lord et al. (2008)[76]). By the

Cholesky decomposition of Brownian motion, we rewrite the dynamics of (V_t, V_t') as follows

$$\begin{aligned} dV_t =& \kappa(V_t' - V_t)dt + \xi_1 V_t d\beta_1(t), \\ dV_t' =& c(z_3 - V_t')dt + \xi_2 V_t' d[\rho d\beta_1(t) + \sqrt{1-\rho^2} d\beta_2(t)], \end{aligned} \tag{9.1}$$

where $\{(\beta_1(t), \beta_2(t))\}$ is a standard two-dimensional Brownian motion. Thus the partial truncation scheme can be designed as

$$\begin{aligned} V_{i+1} =& V_i + \kappa(V_i' - V_i)\Delta t + \xi_1(V_i \vee 0)\sqrt{\Delta t} Z_i^{(1)}, \\ V_{i+1}' =& V_i' + c(z_3 - V_i')\Delta t + \xi_2(V_i' \vee 0)[\rho\sqrt{\Delta t} Z_i^{(1)} + \sqrt{1-\rho^2}\sqrt{\Delta t} Z_i^{(2)}], \end{aligned} \tag{9.2}$$

where $\{Z_i^{(1)}, Z_i^{(2)}\}$ is an independent two-dimensional standard normal sequences. In the implementation, we choose a reasonable length of the time step and number of simulation trials to minimize the mean square error.

9.2 Implementation of our Asymptotic Expansion Formula

To demonstrate the numerical performance of our asymptotic expansion formula, we use model parameters similar to those estimated in Gatheral (2008) [50, 51], which were found by calibrating the model to the market VIX option prices data in April 2007. We adjust the value of parameter ξ_2 slightly so that the stability condition (7.13) is satisfied. We assume a risk free rate of 4.00%. Table 9.1 gives this parameter set. Accordingly, the initial value for VIX is calculated as $\text{VIX}_0 = 17.47$. We conduct numerical experiments to compute the prices of options on VIX for different strikes and maturities. Because of the large trading volume and the high liquidity of options on VIX with strikes around the initial value of the underlying VIX, we select the range of strikes used for our numerical illustration to be from 14 to 23. For each case, we compare the computing performance of the formulae employing the expansion up to indicated order of convergence, including $\mathcal{O}(\epsilon^5), \mathcal{O}(\epsilon^4), \mathcal{O}(\epsilon^3)$ and $\mathcal{O}(\epsilon^2)$. We regard the Monte Carlo simulation results as benchmark. In Table 9.2, all numerical results are exhibited. It takes about 60 milliseconds to compute one value from our asymptotic expansion formula while the simulation takes much longer time, on average about 230 CPU seconds for each by sampling 10^3 time steps with 10^6 simulation trials, in order to achieve satisfactory level of standard error (in the magnitude of 10^{-5} in our examples shown in Table 9.2). We can observe a significant saving in computing time using our asymptotic expansion formula as compared to Monte Carlo simulation.

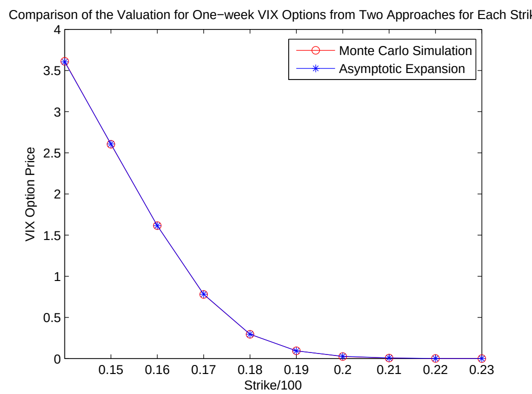
INPUT PARAMETERS	Values
Risk Free Rate r	0.04
Time Horizon of VIX ΔT	0.082
Correlation ρ	0.57
Initial Instantaneous Variance V_0	0.0137
Initial Intermediate Level V_0'	0.0208
Long-Term Intermediate Level z_3	0.078
Rate of Mean Reversion κ	5.5
Rate of Mean Reversion c	0.1
Volatility of the Instantaneous Variance ξ_1	2.6
Volatility of the Intermediate Level ξ_2	0.44

Table 9.1: Input Parameters

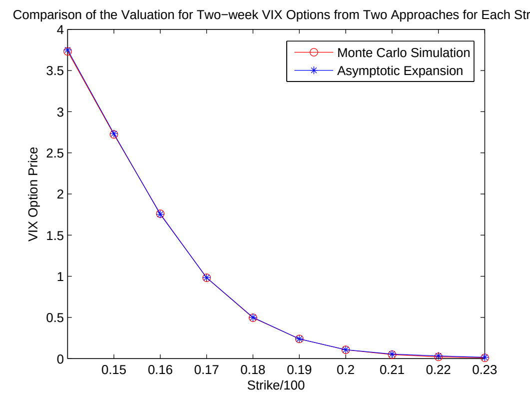
Also, we plot the "value matching" between asymptotic expansion valuation and Monte Carlo simulation. We observe that the satisfactory fitting of the VIX option prices is obtained for actively traded short maturity (for example, up to one-month) options as shown in Figure 9.1. Motivated by the geometric Brownian motion model for VIX in Whaley (1993) [91], in the numerical comparison between our asymptotic expansion formula and Monte Carlo simulation, we also convert the VIX option prices computed from both of the two approaches into the Black-Scholes implied volatility. From Figure 9.2, we observe the fitting of the Black-Scholes implied volatility for the relatively short maturities, one-month at least, and the deviation magnified at the out-of-money side. We also plot and illustrate the computing errors from applying the formulae with various convergence orders, including $\mathcal{O}(\epsilon^5), \mathcal{O}(\epsilon^4), \mathcal{O}(\epsilon^3)$ and $\mathcal{O}(\epsilon^2)$ in Figure 9.3(a). For the sake of completeness, we plot the absolute errors in Figure 9.3(b). We observe the significant improvement on the numerical performance by adding more correction terms in our asymptotic expansion. All implementations algorithms are programmed in C++ and executed on a laptop PC with a Intel(R) Pentium(R) M 1.73GHz processor and 1GB of RAM running Windows XP Professional.

Strike $K/100$	14%	15%	16%	17%	18%	19%	20%	21%	22%	23%
One-Week Options ($T = 0.020$)										
Simulation	3.61737	2.61029	1.62319	0.791609	0.307392	0.102895	0.0288509	0.008203	0.0019101	0.000525892
$\mathcal{O}(\epsilon^5)$	3.61195	2.61458	1.62181	0.794438	0.309904	0.10202	0.0314812	0.0105287	0.00246737	0.000268122
$\mathcal{O}(\epsilon^4)$	3.61526	2.6117	1.61779	0.794843	0.31096	0.10765	0.0353047	0.00842071	0.00106723	6.38707e-005
$\mathcal{O}(\epsilon^3)$	3.60992	2.60225	1.62238	0.819213	0.334626	0.110925	0.0257443	0.00345634	0.000241747	8.45948e-006
$\mathcal{O}(\epsilon^2)$	3.46876	2.47755	1.53341	0.749812	0.259619	0.0575766	0.00757776	0.000561627	2.26518e-005	4.74539e-007
Two-Week Options ($T = 0.040$)										
Simulation	3.74353	2.73528	1.77447	0.999683	0.514822	0.253688	0.115983	0.0546813	0.0241671	0.0112798
$\mathcal{O}(\epsilon^5)$	3.76485	2.744	1.76633	1.001	0.518915	0.253306	0.117863	0.0602953	0.0371945	0.0199837
$\mathcal{O}(\epsilon^4)$	3.75286	2.72414	1.75957	1.00332	0.522148	0.266348	0.143391	0.0765393	0.0335712	0.0105687
$\mathcal{O}(\epsilon^3)$	3.7263	2.72042	1.80068	1.07601	0.593829	0.304875	0.13746	0.0496842	0.0134282	0.00261447
$\mathcal{O}(\epsilon^2)$	3.47802	2.5217	1.64916	0.936648	0.444189	0.169694	0.0506926	0.0115692	0.0019821	0.000251593
Three-Week Options ($T = 0.060$)										
Simulation	3.85696	2.84889	1.90744	1.15636	0.66977	0.38257	0.207609	0.116959	0.0639251	0.0356879
$\mathcal{O}(\epsilon^5)$	3.9048	2.85061	1.88643	1.15137	0.671838	0.380767	0.206683	0.114592	0.0798014	0.0649857
$\mathcal{O}(\epsilon^4)$	3.85972	2.81643	1.87967	1.15701	0.678537	0.401012	0.254474	0.171235	0.108433	0.0573259
$\mathcal{O}(\epsilon^3)$	3.83055	2.84557	1.97562	1.29278	0.813076	0.493557	0.279588	0.139707	0.0586066	0.0200014
$\mathcal{O}(\epsilon^2)$	3.50327	2.58384	1.75956	1.08295	0.589673	0.278553	0.112243	0.0380465	0.0107281	0.00249429
One-Month Options ($T = 0.083$)										
Simulation	3.97402	2.9668	2.04005	1.30086	0.81126	0.506535	0.304207	0.189602	0.116331	0.072158
$\mathcal{O}(\epsilon^5)$	4.03703	2.95017	1.9965	1.28147	0.804713	0.498159	0.296003	0.170324	0.114329	0.104945
$\mathcal{O}(\epsilon^4)$	3.9533	2.90465	1.99223	1.2926	0.817018	0.527392	0.367472	0.278368	0.211932	0.146608
$\mathcal{O}(\epsilon^3)$	3.94944	2.9942	2.16762	1.51541	1.03838	0.698638	0.451386	0.269705	0.143754	0.066617
$\mathcal{O}(\epsilon^2)$	3.54476	2.66137	1.87558	1.22434	0.730706	0.393925	0.189795	0.0809836	0.0303704	0.00994781
One-and-One-Half-Month Options ($T = 0.125$)										
Simulation	4.15701	3.15166	2.24055	1.50635	1.00931	0.685393	0.451759	0.307855	0.20924	0.142487
$\mathcal{O}(\epsilon^5)$	4.2109	3.07856	2.13401	1.43862	0.966496	0.646782	0.414967	0.243632	0.139806	0.112256
$\mathcal{O}(\epsilon^4)$	4.07047	3.02339	2.1401	1.46437	0.993537	0.695583	0.527361	0.441224	0.38937	0.33528
$\mathcal{O}(\epsilon^3)$	4.17172	3.26713	2.49596	1.87892	1.4063	1.04625	0.763316	0.534019	0.350053	0.211046
$\mathcal{O}(\epsilon^2)$	3.63616	2.8024	2.06262	1.43938	0.945713	0.581203	0.332128	0.175556	0.0854423	0.03814
Two-Month Options ($T = 0.167$)										
Simulation	4.30802	3.30456	2.40238	1.66554	1.1602	0.823556	0.569775	0.406407	0.289565	0.207371
$\mathcal{O}(\epsilon^5)$	4.31628	3.14835	2.20539	1.52009	1.0522	0.727549	0.47945	0.276282	0.126356	0.0563953
$\mathcal{O}(\epsilon^4)$	4.13725	3.09444	2.22862	1.56638	1.09979	0.800743	0.633207	0.556767	0.52741	0.50376
$\mathcal{O}(\epsilon^3)$	4.39851	3.53538	2.80233	2.20809	1.73945	1.3684	1.06471	0.806773	0.585569	0.401381
$\mathcal{O}(\epsilon^2)$	3.73395	2.93583	2.22609	1.62023	1.12688	0.745763	0.467793	0.277145	0.154597	0.0809722

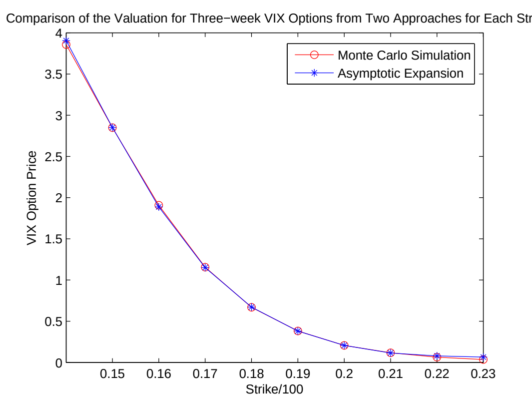
Table 9.2: The table above shows the prices of options on VIX for different strikes and maturities, computed from different methods: Monte Carlo simulation and the asymptotic expansion formulae. We compare the computing performance of the formulae computation employing the expansion up to each indicated order of convergence, including $\mathcal{O}(\epsilon^5), \mathcal{O}(\epsilon^4), \mathcal{O}(\epsilon^3)$ and $\mathcal{O}(\epsilon^2)$.



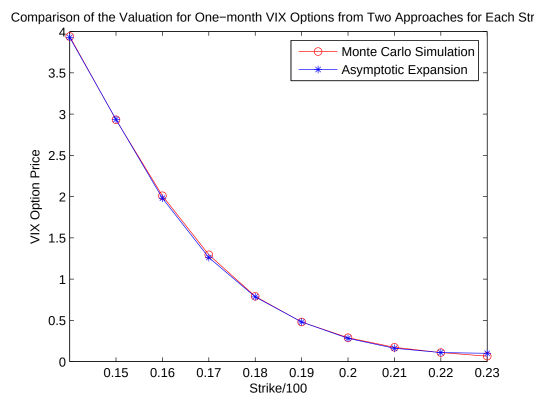
(a) One-week



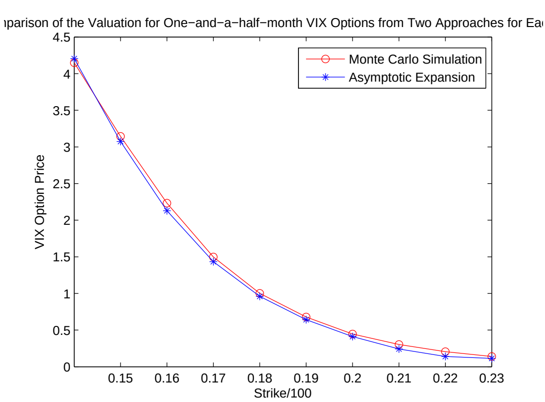
(b) Two-week



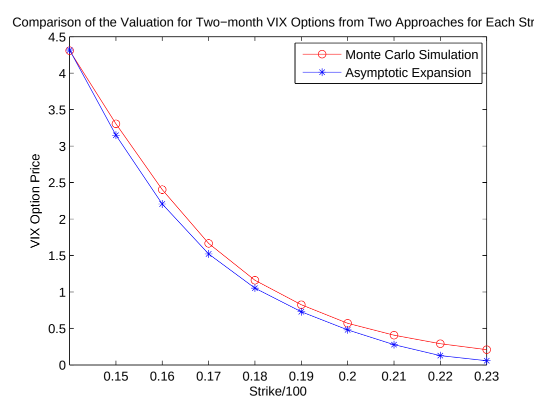
(c) Three-week



(d) One-month

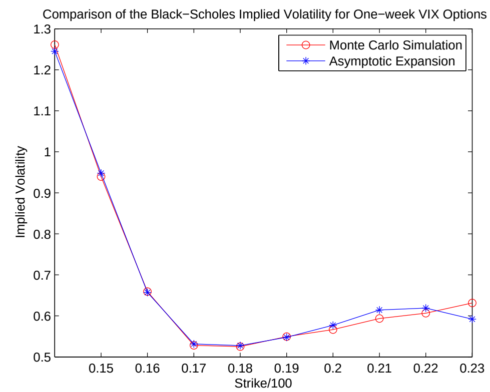


(e) One-and-a-half-month

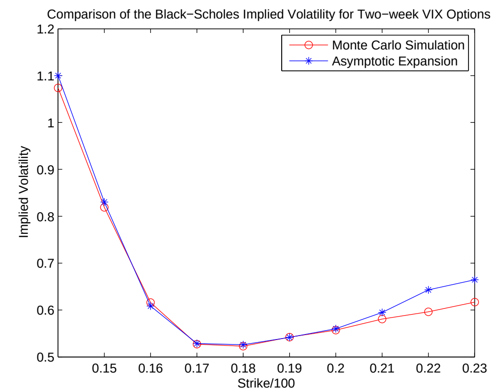


(f) Two-month

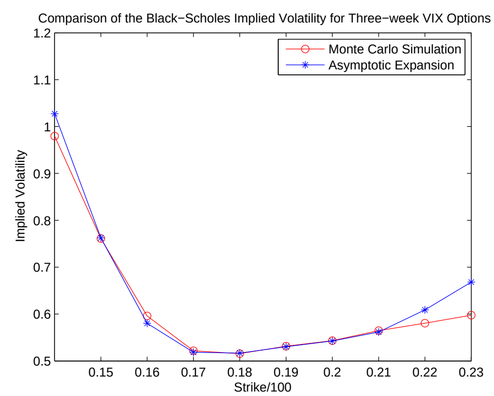
Figure 9.1: This set of graphs illustrates the comparison of the VIX option prices computed from the Monte Carlo simulation and our $\mathcal{O}(\epsilon^5)$ asymptotic expansion. The maturities range from one week to two month.



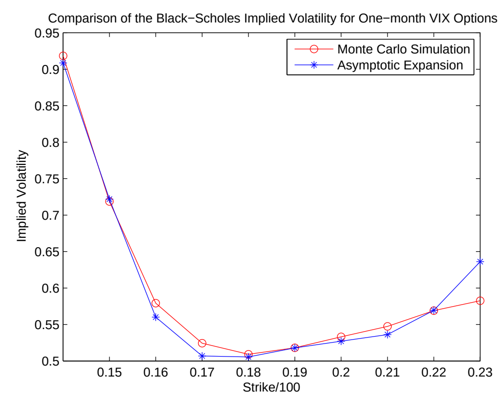
(a) One-week



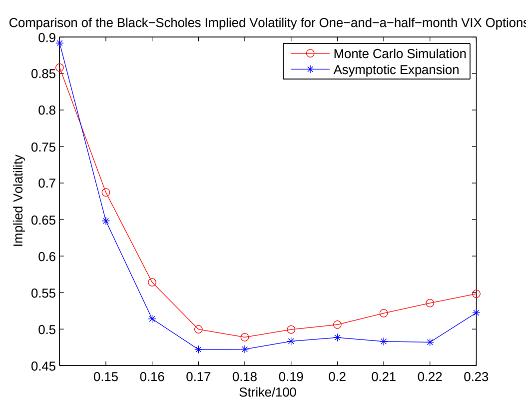
(b) Two-week



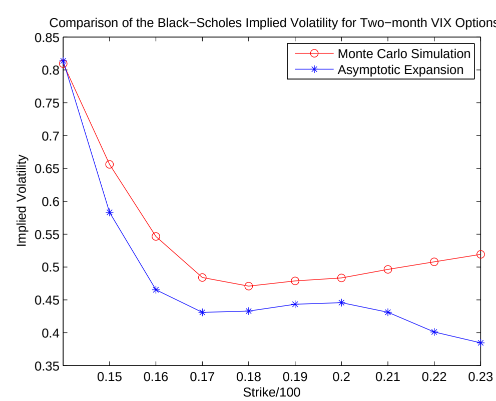
(c) Three-week



(d) One-month

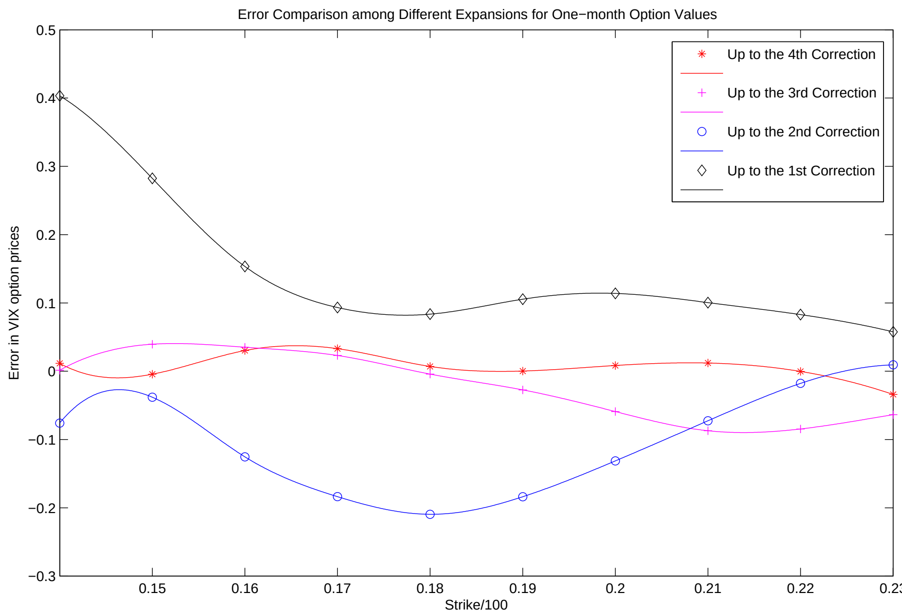


(e) One-and-a-half-month

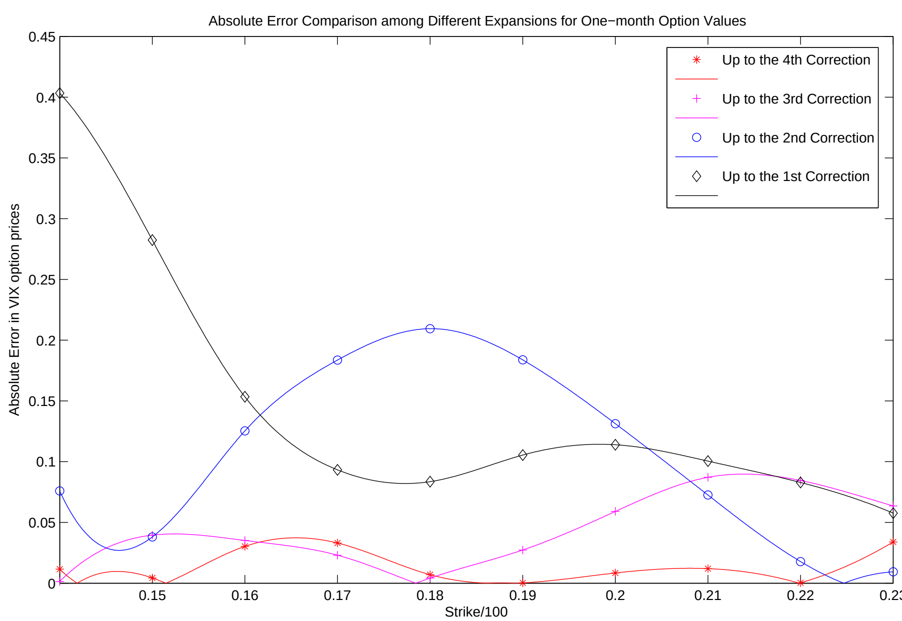


(f) Two-month

Figure 9.2: This set of graphs illustrates the comparison of the Black-Scholes implied volatilities computed from the Monte Carlo simulation and our $\mathcal{O}(\epsilon^5)$ asymptotic expansion. The maturities range from one week to two month.



(a) Error



(b) Absolute Error

Figure 9.3: This graph illustrates the comparison of errors and absolute errors, respectively, resulting from employing the asymptotic expansion formulae up to different convergence orders, including $\mathcal{O}(\epsilon^5), \mathcal{O}(\epsilon^4), \mathcal{O}(\epsilon^3)$ and $\mathcal{O}(\epsilon^2)$. For instance, we compute the prices of options on VIX with maturity one month. The error is interpreted by the difference between the formulae value and Monte Carlo simulation value. The absolute error is the absolute value of the error.

Chapter 10

On the Validity of the Asymptotic Expansion

In this chapter, we justify the validity of our asymptotic expansion in Theorem 5. The following theorem characterizes the magnitude of the asymptotic error explicitly.

THEOREM 6. *Let us denote $C_0^{VIX}(k)$ the price of VIX option with strike k. There exists $R > 0$, such that, for any $N \in \mathbb{N}$, we have*

$$\left| C_0^{VIX}(100 \times (X_0 + \epsilon \xi y)) - 100 \times \xi e^{-r\frac{\epsilon^2}{\xi_2^2}} \sum_{k=1}^{N} \epsilon^k \Theta_k(y) \right| \leq R\epsilon^{N+1},$$

where the various quantities are defined in Theorem 5.

We employ the Malliavin-Watanabe-Yoshida theory on asymptotic expansion for generalized Wiener functional to justify Theorem 6. For the reader's convenience, we document a brief introduction to the theory of Malliavin calculus and the Malliavin-Watanabe-Yoshida theory on asymptotic expansion, where the Watanabe theory (1987) [90] and its further refinement by Yoshida [93, 92, 94] are briefly surveyed.

First of all, we establish the following proposition which leads to Theorem 6.

PROPOSITION 15. *For any arbitrary $y \in \mathbf{R}$ and $f(x) = \max(x, 0)$, we have that*

$$|\mathbb{E}^Q f(Y^\epsilon - y) - \mathbb{E}^Q[\Phi_0(y) + \epsilon\Phi_1(y) + \epsilon^2\Phi_2(y) + \epsilon^3\Phi_3(y)]| \leq R\epsilon^4, \tag{10.1}$$

for some constant $R > 0$, where $\Phi_i(y),\ i = 1, 2, 3, 4$, are defined in (8.41).

The proof of Proposition 15 is carried out in several steps as follows. We sketch the proof and omit some tedious routine details. In the proof, the notations related to the Malliavin calculus are explained in Appendix 10.

First of all, by the Cholesky decomposition, we identify a two-dimensional Wiener process $\{W_1(t), W_2(t)\}$ such that

$$\begin{aligned} B_1(t) &= W_1(t), \\ B_2(t) &= \rho W_1(t) + \sqrt{1-\rho^2}W_2(t). \end{aligned} \tag{10.2}$$

In the following exposition, our justification is carried out on the two-dimensional Wiener space associated to $\{W_1(t), W_2(t)\}$.

Because the SDE system is linear, by Theorem 2.2.2 of Nualart (2006), the bivariate random variable $(V^\epsilon(t), V'^\epsilon(t))$ is in D^∞. Based on Theorem 7.1 in Malliavin and Thalmaier (2004) [78] and Theorem 3.3 in Watanabe (1987), we find that the coefficients for the asymptotic expansion of $(V^\epsilon(t), V'^\epsilon(t))$ in Proposition 13 are in D^∞; the asymptotic expansion (8.28) lies in D^∞. Let us regard X^ϵ as a function $G(v, v') = \sqrt{a_1 v + a_2 v' + a_3 z_3}$ acting on the positive diffusion $\{(V^\epsilon(t), V'^\epsilon(t))\}$. Employing the chain rule and Proposition 1.5.1 in Nualart (2006) [80], we obtain that $X^\epsilon \in D^\infty$. Noticing the expression of the coefficients in the expansion of X^ϵ, we claim that $X_i \in D^\infty$, for $i = 0, 1, 2, ...$ Following an argument using elementary Taylor expansion as in Chapter. 7 of Malliavin and Thalmaier (2004) [78], we obtain that, for any $n \in \mathbb{N}$,

$$\left\| X^\epsilon - \sum_{k=0}^{n} \frac{1}{k!} \frac{\partial^k X^\epsilon}{\partial \epsilon^k}\bigg|_{\epsilon=0} \epsilon^k \right\|_{D_p^s} = \mathcal{O}(\epsilon^{n+1}), \quad \text{for any} \quad s > 0, p \in \mathbb{N}. \tag{10.3}$$

i.e.

$$\left\| X^\epsilon - \left(X_0 + \xi \sum_{k=1}^{n} \epsilon^k X_k \right) \right\|_{D_p^s} = \mathcal{O}(\epsilon^{n+1}), \quad \text{for any} \quad s > 0, p \in \mathbb{N}. \tag{10.4}$$

Without loss of generality, we sketch the proof of (10.4) for the case of $n = 1$:

$$\| X^\epsilon - (X_0 + \epsilon \xi X_1) \|_{D_p^s} = \mathcal{O}(\epsilon^2). \tag{10.5}$$

Indeed, by the integral form of Taylor expansion and integration variable substitution, we deduce that

$$\begin{aligned} X^\epsilon - (X_0 + \epsilon \xi X_1) =& X^\epsilon - X_0 - \frac{\partial X}{\partial \epsilon}\bigg|_{\epsilon=0} \epsilon \\ =& \int_0^\epsilon \int_0^u \frac{\partial^2 X}{\partial \epsilon^2}\bigg|_{\epsilon=s} ds du \\ =& \epsilon^2 \int_0^1 \int_0^w \frac{2F''(r\epsilon)F(r\epsilon) - F'(r\epsilon)^2}{4F(r\epsilon)^{\frac{3}{2}}} dr dw. \end{aligned} \tag{10.6}$$

Because

$$\| X^\epsilon - (X_0 + \epsilon \xi X_1) \|_{D_p^s} \le \epsilon^2 \int_0^1 \int_0^w \left\| \frac{2F''(r\epsilon)F(r\epsilon) - F'(r\epsilon)^2}{4F(r\epsilon)^{\frac{3}{2}}} \right\|_{D_p^s} dr dw, \tag{10.7}$$

(10.5) follows from the fact that

$$\left\| \frac{2F''(r\epsilon)F(r\epsilon) - F'(r\epsilon)^2}{4F(r\epsilon)^{\frac{3}{2}}} \right\|_{D_p^s} < +\infty, \tag{10.8}$$

uniformly for $0 \le r, \epsilon \le 1$. Indeed, (10.8) can be routinely proved via some explicit computation based on (8.34) and (8.35). Without loss of generality, we focus on the L^p-norm. Because of the triangle inequality and the fact $a_1, a_2, a_3 > 0$, we obtain that

$$\begin{aligned} \left\| \frac{2F''(r\epsilon)F(r\epsilon) - F'(r\epsilon)^2}{4F(r\epsilon)^{\frac{3}{2}}} \right\|_{L^p} \le & \frac{1}{(\sqrt{a_3 z_3})^3} \| 2F''(r\epsilon)F(r\epsilon) - F'(r\epsilon)^2 \|_{L^p} \\ \le & \frac{1}{(\sqrt{a_3 z_3})^3} (2\|F''(r\epsilon)F(r\epsilon)\|_{L^p} + \|F'(r\epsilon)^2\|_{L^p}). \end{aligned} \tag{10.9}$$

We further express F, F' and F'' in terms of the explicit solution of V and V' in (8.34). The application of the triangle inequality and the Hölder inequality as well as the explicit computation of moments of the geometric Brownian motion components in (8.34) yields (10.8) uniformly for $0 \le r, \epsilon \le 1$. We omit the tedious details, as a very similar argument is given momentarily in the subsequent verification procedure.

It is straightforward to find that the case of $n = 4$ of (10.4) implies the asymptotic expansion of D^∞ random variable Y^ϵ in D^∞ up to the third order, i.e.

$$\|Y^\epsilon - (X_1 + \epsilon X_2 + \epsilon^2 X_3 + \epsilon^3 X_4)\|_{D_p^s} = \mathcal{O}(\epsilon^4), \quad \text{for any} \quad s > 0, p \in \mathbb{N}.$$

Next, we only need to verify that the underlying variable Y^ϵ is uniformly non-degenerate in the sense of Malliavin. We follow the approach proposed in Yoshida (1992) [93] to verify a truncated version of the uniform non-degenerate condition on the Malliavin covariance of Y^ϵ. Denote the Malliavin covariance matrix

$$\Sigma(\epsilon) := \langle DY^\epsilon, DY^\epsilon\rangle_{L^2([0,T];\mathbb{R}^2)} = \sum_{k=1}^{2}\langle D^k Y^\epsilon, D^k Y^\epsilon\rangle_{L^2([0,T])} = \sum_{k=1}^{2}\int_0^t (D_s^k Y^\epsilon)^2 ds. \tag{10.10}$$

By the definition in (8.40) and the chain rule of Malliavin differentiation,

$$D^i Y^\epsilon = \frac{a_1 D^i V^\epsilon(t) + a_2 D^i V'^\epsilon(t)}{2\epsilon\xi\sqrt{F(\epsilon)}}. \tag{10.11}$$

The limiting distribution of Y^ϵ satisfies that

$$Y^\epsilon \to X_1 \sim N(0,1), \quad \text{as } \epsilon \to 0, \quad \text{a.e. } \Omega.$$

We thus define the limiting Malliavin covariance as

$$\Sigma(0) := \langle DY^0, DY^0\rangle_H = \sum_{k=1}^{2}\int_0^t (D_s^k X_1)^2 ds. \tag{10.12}$$

It is obvious that

$$\begin{aligned} D_s^1 X_1 &= \frac{a_1\xi_1 V_0 + a_2\xi_2 V_0'\rho}{2\xi\sqrt{F(0)}} 1_{[0,t]}(s), \\ D_s^2 X_1 &= \frac{a_2\xi_2 V_0'\sqrt{1-\rho^2}}{2\xi\sqrt{F(0)}} 1_{[0,t]}(s). \end{aligned} \tag{10.13}$$

It follows from a direct computation that $\Sigma(0) > 0$.

Let us define

$$\eta_c^\epsilon = c\int_0^t |D_s Y^\epsilon - D_s Y^0|^2 ds = c\int_0^t [(D_s^1 Y^\epsilon - D_s^1 Y^0)^2 + (D_s^2 Y^\epsilon - D_s^2 Y^0)^2] ds, \tag{10.14}$$

where $c > 0$. In the following proposition, we justify a truncated version of the Malliavin uniform non-degeneracy condition (69) as proposed in Yoshida (1992) [93].

LEMMA 8. *The Malliavin covariance* $\Sigma(\epsilon)$ *defined in* (10.10) *is uniformly non-degenerated under truncation, i.e. there exists* $c_0 > 0$ *such that for any* $c > c_0$ *and any* $p > 1$,

$$\sup_{\epsilon \in [0,1]} \mathbb{E}[1\{\eta_c^\epsilon \leq 1\}(\det(\Sigma(\epsilon)))^{-p}] < +\infty. \tag{10.15}$$

Proof. By the triangle and Cauchy-Schwarz inequality,

$$|D_s Y^\epsilon (D_s Y^\epsilon)^T - D_s Y^0 (D_s Y^0)^T| \leq |D_s Y^\epsilon - D_s Y^0|^2 + 2|D_s Y^0||D_s Y^\epsilon - D_s Y^0|.$$

Noticing that $\eta_c^\epsilon \leq 1$ is equivalent to

$$\int_0^t |D_s Y^\epsilon - D_s Y^0|^2 ds \leq \frac{1}{c},$$

we thus have

$$\begin{aligned} |\Sigma(\epsilon) - \Sigma(0)| \leq & \int_0^t |D_s Y^\epsilon - D_s Y^0|^2 ds + \int_0^t 2|D_s Y^0||D_s Y^\epsilon - D_s Y^0| ds \\ \leq & \int_0^t |D_s Y^\epsilon - D_s Y^0|^2 ds + 2\left(\int_0^t |D_s Y^0|^2\right)^{\frac{1}{2}} \left(\int_0^t |D_s Y^\epsilon - D_s Y^0|^2 ds\right)^{\frac{1}{2}} \\ \leq & \frac{1}{c} + 2\sqrt{\frac{\Sigma(0)}{c}}. \end{aligned} \tag{10.16}$$

Hence, there exists c_0 such that, for any $c > c_0 > 0$,

$$|\Sigma(\epsilon)| \geq \Sigma(0) - |\Sigma(\epsilon) - \Sigma(0)| > \Sigma(0) - \left(\frac{1}{c_0} + 2\sqrt{\frac{\Sigma(0)}{c_0}}\right).$$

□

In order to apply the theory of Yoshida [93], we need to justify that the truncation in Lemma 8 is negligible in probability by verifying condition (70).

LEMMA 9. *Following the definition of η_c^ϵ in* (10.14), *we have*

$$\mathbb{P}\left(|\eta_c^\epsilon| > \frac{1}{2}\right) = \mathcal{O}(\epsilon^n), \tag{10.17}$$

for any $n = 1, 2, 3, ...$

Proof. We prove that, for any $n \in \mathbf{N}$ there exist a constant c_n such that

$$\mathbb{P}\left(|\eta_c^\epsilon| > \frac{1}{2}\right) < c_n \epsilon^{2n}. \tag{10.18}$$

Noticing the fact that

$$\left\{|\eta_c^\epsilon| > \frac{1}{2}\right\} \subseteqq \left\{c\int_0^t (D_s^1 Y^\epsilon - D_s^1 Y^0)^2 ds > \frac{1}{4}\right\} \cup \left\{c\int_0^t (D_s^2 Y^\epsilon - D_s^2 Y^0)^2 ds > \frac{1}{4}\right\}.$$

we have,

$$\mathbb{P}\left(|\eta_c^\epsilon| > \frac{1}{2}\right) \leq \mathbb{P}\left(\int_0^t (D_s^1 Y^\epsilon - D_s^1 Y^0)^2 ds > \frac{1}{4c}\right) + \mathbb{P}\left(\int_0^t (D_s^2 Y^\epsilon - D_s^2 Y^0)^2 ds > \frac{1}{4c}\right). \tag{10.19}$$

It is sufficient to prove that,

$$\mathbb{P}\left(\int_0^t (D_s^i Y^\epsilon - D_s^i Y^0)^2 ds > \frac{1}{4c}\right) = \mathcal{O}(\epsilon^{2n}), \tag{10.20}$$

for $i = 1, 2$ and any $n \in \mathbb{N}$. Without loss of generality, we provide the proof for the case of $i = 1$.

By the Chebyshev-Markov inequality, we have

$$\begin{aligned}
&\mathbb{P}\left(\int_0^t (D_s^1 Y^\epsilon - D_s^1 Y^0)^2 ds > \frac{1}{4c}\right) \\
\leq &(4c)^n \mathbb{E}\left(\int_0^t (D_s^1 Y^\epsilon - D_s^1 Y^0)^2 ds\right)^n \\
= &(4c)^n \mathbb{E}\left(\int_0^t \left[d_1(s) + d_2(s)\right]^2 ds\right)^n,
\end{aligned} \tag{10.21}$$

where

$$d_1(s) := \frac{a_2}{2\xi}\left(\frac{D_s^1 V'^\epsilon(t)}{\epsilon\sqrt{F(\epsilon)}} - \frac{\xi_2 V_0'\rho}{\sqrt{F(0)}}\right), \quad d_2(s) := \frac{a_1}{2\xi}\left(\frac{D_s^1 V^\epsilon(t)}{\epsilon\sqrt{F(\epsilon)}} - \frac{\xi_1 V_0}{\sqrt{F(0)}}\right). \tag{10.22}$$

By the Holder inequality and the convexity property of power functions, we deduce that

$$\begin{aligned}
\mathbb{E}\left(\int_0^t \left[d_1(s) + d_2(s)\right]^2 ds\right)^n \leq &\mathbb{E}\left(\int_0^t \left[d_1(s) + d_2(s)\right]^{2n} ds\right) \cdot t^{\frac{n}{n'}} \\
\leq &2^{2n-1} t^{\frac{n}{n'}} \int_0^t [\mathbb{E}(d_1(s))^{2n} + \mathbb{E}(d_2(s))^{2n}] ds
\end{aligned} \tag{10.23}$$

where $\frac{1}{n} + \frac{1}{n'} = 1$.

Without loss of generality, we justify that

$$\int_0^t \mathbb{E}(d_1(s))^{2n} ds = \mathcal{O}(\epsilon^{2n}). \tag{10.24}$$

By the Malliavin differentiation chain rule and the fact $D_s^i W_j(t) = \delta_{ij} 1\{s \leq t\}$, a direct computation based on the explicit solution in (8.34) yields that

$$\begin{aligned}
D_s^1 V'^\epsilon(t) = &\epsilon\xi_2\rho 1_{[0,t]}(s) V'^\epsilon(t) - \epsilon^3 c z_3 \xi_2 \rho \exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2 \xi_2^2 t\right\} \\
&\cdot \int_0^t \exp\left\{\epsilon^2 cu - \epsilon\xi_2 B_2(u) + \frac{1}{2}\epsilon^2\xi_2^2 u\right\} 1_{[0,u](s)} du.
\end{aligned} \tag{10.25}$$

Thus,

$$\mathbb{E}\int_0^t (d_1(s))^{2n} ds = \mathbb{E}\int_0^t \left[I_1(s) + I_2(s)\right]^{2n} ds, \tag{10.26}$$

where

$$I_1(s) := \xi_2 \rho \left(\frac{V'^{\epsilon}(t)}{\sqrt{F(\epsilon)}} - \frac{V_0'}{\sqrt{F(0)}} \right) \tag{10.27}$$

and

$$\begin{aligned} &I_2(s) \\ &:= -\frac{\epsilon^2 \xi_2 \rho c z_3 \exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\} \int_0^t \exp\left\{\epsilon^2 cu - \epsilon\xi_2 B_2(u) + \frac{1}{2}\epsilon^2 \xi_2^2 u\right\} 1_{[0,u](s)} du}{\sqrt{F(\epsilon)}}. \end{aligned} \tag{10.28}$$

Following a similar inequality estimate as in (10.23), it suffices to show that

$$\mathbb{E}\int_0^t \left(I_k(s)\right)^{2n} ds = O(\epsilon^{2n}), \quad \text{for } k = 1, 2. \tag{10.29}$$

Indeed, based on the fact of $a_1 > 0, a_2 > 0, a_3 > 0$, we deduce that

$$\begin{aligned} \mathbb{E}\int_0^t \left(I_1(s)\right)^{2n} ds =& t \cdot \mathbb{E}\left(\frac{V'^{\epsilon}(t) - V_0'}{\sqrt{F(\epsilon)}} + \left(\frac{1}{\sqrt{F(\epsilon)}} - \frac{1}{\sqrt{F(0)}} \right) V_0' \right)^{2n} \\ \le& C_1 \mathbb{E}\left(\frac{V'^{\epsilon}(t) - V_0'}{\sqrt{F(\epsilon)}} \right)^{2n} + C_2 \mathbb{E}\left(\frac{1}{\sqrt{F(\epsilon)}} - \frac{1}{\sqrt{F(0)}} \right)^{2n} \\ \le& \frac{C_1}{F(0)^n} \mathbb{E}(V'^{\epsilon}(t) - V_0')^{2n} + C_2 \mathbb{E}\left(\frac{F(\epsilon) - F(0)}{(\sqrt{F(\epsilon)} + \sqrt{F(0)})\sqrt{F(\epsilon)F(0)}} \right)^{2n} \\ \le& \frac{C_1}{F(0)^n} \mathbb{E}(V'^{\epsilon}(t) - V_0')^{2n} + C_2 \mathbb{E}\left(\frac{F(\epsilon) - F(0)}{2F(0)\sqrt{F(0)}} \right)^{2n} \\ \le& \widehat{C}_1 \mathbb{E}(V'^{\epsilon}(t) - V_0')^{2n} + \widehat{C}_2 \mathbb{E}(V^{\epsilon}(t) - V_0)^{2n}. \end{aligned} \tag{10.30}$$

Thus, it is sufficient to show that

$$\mathbb{E}(V'^{\epsilon}(t) - V'_0)^{2n} = O(\epsilon^{2n}) \tag{10.31}$$

and

$$\mathbb{E}(V^{\epsilon}(t) - V_0)^{2n} = O(\epsilon^{2n}). \tag{10.32}$$

Without loss of generality, we briefly justify (10.31) here. Indeed, we have that

$$\begin{aligned} V'^{\epsilon}(t) - V'_0 = & \left(\exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\} - 1\right) V'_0 \\ & + \epsilon^2 c z_3 \exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\} \int_0^t \exp\left\{\epsilon^2 cu - \epsilon\xi_2 B_2(u) + \frac{1}{2}\epsilon^2\xi_2^2 u\right\} du. \end{aligned} \tag{10.33}$$

It is straightforward to obtain that

$$\begin{aligned} & \lim_{\epsilon\to 0} \frac{1}{\epsilon^{2n}} \mathbb{E}\left(\exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\} - 1\right)^{2n} \\ = & \lim_{\epsilon\to 0} \frac{1}{\epsilon^{2n}} \mathbb{E} \sum_{k=0}^{2n} \binom{2n}{k} \exp\left\{\epsilon k\xi_2 B_2(t) - \epsilon^2 kct - \frac{1}{2}\epsilon^2 k\xi_2^2 t\right\} (-1)^{2n-k} \\ = & \lim_{\epsilon\to 0} \frac{1}{\epsilon^{2n}} \sum_{k=0}^{2n} \binom{2n}{k} \exp\left\{\left[\frac{1}{2}\xi_2^2 k(k-1) - kc\right]\epsilon^2 t\right\} (-1)^{2n-k} = \xi_2^{2n} \frac{(2n)! t^n}{2^n n!} \end{aligned} \tag{10.34}$$

and

$$\begin{aligned} & \mathbb{E}\left(\epsilon^2 c z_3 \exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\} \int_0^t \exp\left\{\epsilon^2 cu - \epsilon\xi_2 B_2(u) + \frac{1}{2}\epsilon^2\xi_2^2 u\right\} du\right)^{2n} \\ = & \mathcal{O}(\epsilon^4). \end{aligned} \tag{10.35}$$

Hence (10.31) is proved and (10.32) follows quite similarly.

Similarly, we find that

$$
\begin{aligned}
&\mathbb{E}\int_0^t (I_2(s))^{2n}\, ds\\
\leq &C_3(t)\epsilon^{4n}\mathbb{E}\left(\exp\left\{-\epsilon^2 ct + \epsilon\xi_2 B_2(t) - \frac{1}{2}\epsilon^2\xi_2^2 t\right\}\int_0^t \exp\left\{\epsilon^2 cu - \epsilon\xi_2 B_2(u) + \frac{1}{2}\epsilon^2\xi_2^2 u\right\} du\right)\\
= &\mathcal{O}(\epsilon^{4n}).
\end{aligned}
\tag{10.36}
$$

Therefore, (10.18) is verified, which is equivalent to (10.17).

$\square$

Hence, for the tempered distribution $T_y \in \mathcal{S}'(\mathbb{R})$ defined as

$$T_y(x) = f(x-y) = (x-y)^+,$$

we apply the Malliavin-Watanabe-Yoshida theory surveyed in Section .2 to conclude the validity of the asymptotic expansion in Proposition 15. Therefore, Theorem 6 is proved accordingly.

Glossary

arbitrage — the practice of taking advantage of a price differential between two or more markets: striking a combination of matching deals that capitalize upon the imbalance, the profit being the difference between the market prices, i

CBOE — Chicago Board Options Exchange, located at 400 South LaSalle Street in Chicago, the largest U.S. options exchange with annual trading volume that hovered around one billion contracts at the end of 2007, i

diversified portfolio — a risk management technique, related to hedging, that mixes a wide variety of investments within a portfolio, 5

European option a financial contract between two parties, the buyer and the seller of this type of option. It is the option to buy (sell) shares of stock at a specified time in the future for a specified price, 3

hedge a position established in one market in an attempt to offset exposure to price fluctuations in some opposite position in another market with the goal of minimizing one's exposure to unwanted risk, 156

implied volatility the volatility implied by the market price of the option based on an option pricing model (i.e. the volatility that, when used in a particular pricing model, yields a theoretical value for the option equal to the current market price of that option), 3

out-of-the-money For a call, when an option's strike price is higher than the market price of the underlying asset; For a put, when the strike price is below the market price of the underlying asset, 74

over-the-counter A decentralized market of securities not listed on an exchange where market participants trade over the telephone, facsimile or electronic network instead of a physical trading floor, 12

portfolio insurance a method of hedging a portfolio of stocks against the market risk by short selling stock index futures, 5

price sensitivities the quantities representing the sensitivities of derivatives such as options to a change in underlying parameters on which the value of an instrument or portfolio of financial instruments is dependent, 2

realized variance the variance of the return of a financial instrument over a specified period, 5

realized volatility the volatility of a financial instrument over a specified period, 5

S&P 500 a free-float capitalization-weighted index published since 1957 of the prices of 500 large-cap common stocks actively traded in the United States, i

Societe Generale one of the main European financial services companies and also maintains extensive activities in others parts of the world, 2

timer option an Exotic option, that allows buyers to specify the level of volatility used to price the instrument, 2

variance swap an over-the-counter financial derivative that allows one to speculate on or hedge risks associated with the magnitude of movement, i.e. volatility, of some underlying product, like an exchange rate, interest rate, or stock index, 3

VIX the ticker symbol for the Chicago Board Options Exchange Volatility Index, a popular measure of the implied volatility of S&P 500 index options, i

volatility most frequently refers to the standard deviation of the continuously compounded returns of a financial instrument within a specific time horizon, i

Bibliography

[1] J. Abate and W. Whitt. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems Theory and Applications*, 10(1-2):5–87, 1992.

[2] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* A Wiley-Interscience Publication. John Wiley and Sons Inc., New York, 1984. Reprint of the 1972 edition, Selected Government Publications.

[3] L. B. Andersen and R. Brotherton-Ratcliffe. Extended libor market models with stochastic volatility. *Journal of Computational Finance*, 9(1), 2005.

[4] D. S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9(1):69–107, 1996.

[5] D. S. Bates. Post-'87 crash fears in S&P 500 futures options. Jan. 1997.

[6] L. Bergomi. Smile dynamics II. *Risk*, October 2005.

[7] L. Bergomi. Smile dynamics III. *Risk*, March 2009.

[8] A. Bick. Quadratic-variation-based dynamic strategies. *Management Science*, 41(4), 1995.

[9] T. Bjork. *Arbitrage Theory in Continuous Time*. Oxford University Press, 1999.

[10] F. Black and M. S. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973.

[11] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, April 1986.

[12] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion—Facts and Formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, second edition, 2002.

[13] D. T. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296, September 1979.

[14] M. Broadie, M. Chernov, and M. Johannes. Model specification and risk premia: Evidence from futures options. *Journal of Finance*, 62(3):1453–1490, 06 2007.

[15] M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management Science*, 42(2):269–285, 1996.

[16] M. Broadie and A. Jain. Pricing and hedging volatility derivatives. *The Journal of Derivatives*, 15(3):7–24, spring 2008.

[17] M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2):217–231, 2006.

[18] H. Buchholz. *The Confluent Hypergeometric Function with Special Emphasis on its Applications*. Translated from the German by H. Lichtblau and K. Wetzel. Springer Tracts in Natural Philosophy, Vol. 15. Springer-Verlag New York Inc., New York, 1969.

[19] H. Buehler. Consistent variance curve models. *Finance and Stochastics*, 10(2):178–203, April 2006.

[20] P. Carr and R. Lee. Realized volatility and variance: Options via swaps. *RISK*, pages 76–83, May 2007.

[21] P. Carr and R. Lee. Hedging variance options on continuous semimartingales. *Finance and Stochastics*, 14(2):179–207, April 2010.

[22] P. P. Carr and L. Wu. A tale of two indices. *Journal of Derivatives*, Spring 2006.

[23] CBOE. The CBOE volatility index - VIX. *Chicago Board Options Exchange*, 2009.

[24] P. Christoffersen, S. L. Heston, and K. Jacobs. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *SSRN eLibrary*, 2009.

[25] J. C. Cox, J. Ingersoll, Jonathan E, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[26] J. C. Cox, J. Ingersoll, Jonathan E, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[27] J. Cvitanić and I. Karatzas. Hedging contingent claims with constrained portfolios. 3(3):652–681, 1993.

[28] F. Delbaen and H. Shirakawa. A note on option pricing for the constant elasticity of variance model. *Financial Engineering and the Japanese Markets*, 9(2):85–99(15), 2002.

[29] K. Demeterfi, E. Derman, M. Kamal, and J. Zou. More than you ever wanted to know about volatility swaps. *Quantitative Strategies Research Notes, Goldman Sachs*, 1999.

[30] J. Detemple and C. Osakwe. The valuation of volatility options. *European Finance Review*, 4(1), 2000.

[31] J. Douglas, Jr. and H. H. Rachford, Jr. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82:421–439, 1956.

[32] D. Duffie. *Dynamic Asset Pricing Theory, 3rd edition*. Princeton University Press, Princeton, NJ, 2001.

[33] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. 13(3):984–1053, 2003.

[34] D. Duffie and P. Glynn. Efficient Monte Carlo simulation of security prices. *The Annals of Applied Probability*, 5(4):897–905, 1995.

[35] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 1996.

[36] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376, 2000.

[37] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[38] B. Dupire. Model art. *RISK*, 6(9):118–124, September 1993.

[39] B. Dupire. Exploring volatility derivatives: New advances in modelling. *Global Derivatives, Paris*, May 2005.

[40] B. Eraker. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance*, 59:1367–1403, 2004.

[41] W. Feller. Two singular diffusion problems. *The Annals of Mathematics*, 54:173–182, 1951.

[42] G. S. Fishman. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Series in Operations Research. Springer-Verlag, New York, 1996.

[43] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter & Co., Berlin, 2004.

[44] E. d. Fontenay. Hedge fund replication and structured products. *Societe Generale Asset Management*, 2007.

[45] E. Fourni, J.-M. Lasry, P.-L. Lions, J. Lebuchoux, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3(4):391–412, 1999.

[46] P. K. Friz. An Introduction to Malliavin Calculus. *lecture notes*, 2005.

[47] A. R. Gallant, C.-T. Hsu, and G. Tauchen. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *The Review of Economics and Statistics*, 81(4):617–631, November 1999.

[48] J. Gatheral. *The Volatility Surface, a Practitioner's Guide*. John Wiley and Sons, Hoboken, NJ, 2006.

[49] J. Gatheral. Developments in volatility derivatives pricing. *Discussion Paper at Global Derivatives, Paris*, May 2007. Available online at: `http://www.math.nyu.edu/fellows_fin_math/gatheral/VolDerivatives2007.pdf`.

[50] J. Gatheral. Consistent modeling of SPX and VIX options. *Discussion Paper at The Fifth World Congress of the Bachelier Congress Society, London*, July 2008. Avaliable online at: `http://www.math.nyu.edu/fellows_fin_math/gatheral/Bachelier2008.pdf`.

[51] J. Gatheral. Further developments in volatility derivatives pricing. *Discussion Paper at Global Derivatives Trading & Risk Management, Paris*, May 2008. Available online at: `http://www.math.nyu.edu/fellows_fin_math/gatheral/FurtherVolDerivatives2008.pd`

[52] H. Geman and M. Yor. Bessel processes, Asian options and perpetuities. *Mathematical Finance*, 3:349–375, 1993.

[53] P. Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2004. Stochastic Modelling and Applied Probability.

[54] A. Göing-Jaeschke and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[55] A. Grunbichler and F. A. Longstaff. Valuing futures and options on volatility. *Journal of Banking and Finance*, 20(6):985–1001, July 1996.

[56] P. D. Hagan, A.S.Lesniewski, and D.E.Woodward. Managing smile risk. *Wilmott Magazine Separtember*, pages 84–108, 2002.

[57] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. 11(3):215–260, 1981.

[58] J. M. Harrison and S. R. Pliska. A stochastic calculus model of continuous trading: complete markets. 15(3):313–316, 1983.

[59] D. Hawkins and S. Krol. Product overview: Timer options. *Lehman Brothers Equity Derivative Strategy*, January 2008. Available online at `http://www.scribd.com/doc/19601860/Lehman-Brothers-Product-Overview-Timer-Opti`

[60] P. Henry-Labordere. *Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing*. Chapman and Hall/CRC, 2008.

[61] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–43, 1993.

[62] J. C. Hull and A. D. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, June 1987.

[63] P. Jaeckel. *Monte Carlo Methods in Finance*. John Wiley and Sons Inc., New York, 2002.

[64] M. S. Johannes, B. Eraker, and N. Polson. The impact of jumps in volatility and returns. *Journal of Finance*, 53:1269–1300, 2003.

[65] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions. Vol. 2*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley and Sons Inc., New York, second edition, 1995. A Wiley-Interscience Publication.

[66] I. Karatzas and S. G. Kou. On the pricing of contingent claims under constraints. 6(2):321–369, 1996.

[67] I. Karatzas and C. Li. Personal communication and notes. 2009, 2010.

[68] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[69] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, New York, 1998.

[70] N. Kunitomo and A. Takahashi. The asymptotic expansion approach to the valuation of interest rate contingent claims. *Mathematical Finance*, 11(1):117–151, 2001.

[71] N. Kunitomo and A. Takahashi. On validity of the asymptotic expansion approach in contingent claim analysis. *The Annals of Applied Probability*, 13(3):914–952, 2003.

[72] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *Journal of the Faculty of Science, the University of Tokyo, Section IA, Mathematics*, 32:1–76, 1985.

[73] R. W. Lee. Implied and local volatilities under stochastic volatility. *International Journal of Theoretical and Applied Finance*, 4(1):45–89, 2001.

[74] V. Linetsky. The spectral representation of Bessel processes with constant drift: applications in queueing and finance. *Journal of Applied Probability*, 41(2):327–344, 2004.

[75] A. Lipton. The vol smile problem. *Risk Magazine*, 15:61–65, 2002.

[76] R. Lord, R. Koekkoek, and D. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. 2008.

[77] P. Malliavin. *Stochastic Analysis.* Grundlehren der Mathematischen Wissenschaften. 313. Berlin: Springer. xi, 342 p., 1997.

[78] P. Malliavin and A. Thalmaier. *Stochastic Calculus of Variations in Mathematical Finance.* Springer Finance. Berlin: Springer, 2006.

[79] R. C. Merton. Theory of rational option pricing. *Bell Journal of Economics*, 4(1):141–183, Spring 1973.

[80] D. Nualart. *The Malliavin Calculus and Related Topics. 2nd ed.* Probability and Its Applications. Berlin: Springer, 2006.

[81] D. Nualart, A. S. Ustunel, and M. Zakai. On the moments of a multiple wiener-ito integral and the space induced by the polynomials of the integral. *Stochastics*, 25(4):233–340, 1988.

[82] Y. Osajima. The asymptotic expansion formula of implied volatility for dynamic SABR model and FX hybrid model. *SSRN eLibrary*, 2007.

[83] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293. Springer-Verlag, Berlin, third edition, 1999.

[84] N. Sawyer. SG CIB launches timer options. *Risk*, 20(7), July 2007.

[85] S. E. Shreve. *Stochastic Calculus for Finance. II.* Springer Finance. Springer-Verlag, New York, 2004.

[86] M. Staunton. Exploring for VIX. *Wilmott Magazine*, pages 66–67, November 2009.

[87] D. W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients. I. 22:345–400, 1969.

[88] D. W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients. II. 22:479–530, 1969.

[89] M. Uchida and N. Yoshida. Asymptotic expansion for small diffusions applied to option pricing. *Statistical Inference for Stochastic Processes*, 7(3):189–223, October 2004.

[90] S. Watanabe. Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Annals of Probability*, 15:1–39, 1987.

[91] R. E. Whaley. Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives*, 1:71–84, 1993.

[92] N. Yoshida. Asymptotic expansions for statistics related to small diffusions. *Journal of Japan Statistical Society*, 22:139–159, 1992.

[93] N. Yoshida. Asymptotic expansions of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe . *Probability Theory and Related Fields*, 92(3):275–311, September 1992.

[94] N. Yoshida. Asymptotic expansion of Bayes estimators for small diffusions. *Probability Theory and Related Fields*, 95(4):429–450, December 1993.

Appendix

A Joint Density of Bessel Process at Exponential Stopping

In this appendix, we give detailed derivation of the joint density in equation 3.20. We begin with a useful result on a Laplace transform involving Bessel process stopped at an exponential time.

LEMMA 10.

$$\mathbb{E}^{P_0}\left[\exp\left\{-\gamma\int_0^T \frac{du}{X_u}\right\}\middle| X_T = x\right]\mathbb{P}_0(X_T \in dx)$$

$$=\frac{\sqrt{\lambda}\Gamma(\gamma+\frac{1}{2}+\frac{\gamma}{\sqrt{2\lambda}})}{\sqrt{2}\Gamma(2\nu+1)}\cdot\left(\frac{x}{X_0}\right)^{\nu+\frac{1}{2}}\cdot M_{-\frac{\gamma}{\sqrt{2\lambda}},\nu}(2(X_0\wedge x)\sqrt{2\lambda})\cdot W_{-\frac{\gamma}{\sqrt{2\lambda}},\nu}(2(X_0\vee x)\sqrt{2\lambda})dx, \tag{37}$$

where M and W are the Whittaker functions defined as follows:

$$\begin{aligned}
M(a,b,x) &= 1+\sum_{k=1}^{+\infty}\frac{a(a+1)\cdots(a+k-1)x^k}{b(b+1)\cdots(b+k-1)k!},\\
U(a,b,x) &= \frac{\pi}{\sin(\pi b)}\left[\frac{M(a,b,x)}{\Gamma(1+a-b)\Gamma(b)}-x^{1-b}\frac{M(1+a-b,2-b,x)}{\Gamma(a)\Gamma(2-b)}\right],\\
M_{n,m}(x) &= x^{m+\frac{1}{2}}e^{-\frac{x}{2}}M(m-n+\frac{1}{2},2m+1,x),\\
W_{n,m}(x) &= x^{m+\frac{1}{2}}e^{-\frac{x}{2}}U(m-n+\frac{1}{2},2m+1,x).
\end{aligned} \tag{38}$$

Proof. Let us assume that X_t is a Bessel process with drift μ under the probability measure $\mathbb{P}^\mu$ and S is an exponential random variable with parameter η. Thus

$$\mathbb{P}^\mu(X_S \leqslant y) = \int_0^{+\infty}\mathbb{P}^\mu(X_t \leqslant y|S=t)\mathbb{P}(S\in t) = \eta\int_0^{+\infty}e^{-\eta t}\mathbb{P}^\mu(X_t\leqslant y)dt$$

$$=\eta\int_0^{+\infty}e^{-\eta t}\mathbb{E}^{P_0}\left[1\{X_t\leqslant y\}\exp\left\{-\mu(X_0-X_t)-\mu\int_0^t\frac{2\nu+1}{2X_u}du-\frac{1}{2}\mu^2 t\right\}\right]dt$$

$$=\eta\int_0^{+\infty}e^{-\eta t}\left[\int_0^y\mathbb{E}^{P_0}\left[\exp\left\{-\mu(X_0-X_t)-\mu\int_0^t\frac{2\nu+1}{2X_u}du-\frac{1}{2}\mu^2 t\right\}\middle|X_t=z\right]\right.$$
$$\left.\mathbb{P}_0(X_t\in dz)\right]dt$$

$$=\eta\int_0^{+\infty}e^{-\left(\eta+\frac{\mu^2}{2}\right)t}\left[\int_0^y e^{-\mu(X_0-z)}\mathbb{E}^{P_0}\left[\exp\left\{-\mu\left(\nu+\frac{1}{2}\right)\int_0^t\frac{1}{X_u}du\right\};X_t\in dz\right]\right]dt$$

$$=\frac{\eta}{\eta+\frac{\mu^2}{2}}\int_0^y e^{-\mu(X_0-z)}\mathbb{E}^{P_0}\left[\exp\left\{-\mu\left(\nu+\frac{1}{2}\right)\int_0^T\frac{1}{X_u}du\right\};X_T\in dz\right], \tag{39}$$

where T is an exponential random variable with parameter $\eta+\frac{\mu^2}{2}$, which is independent of the underlying Bessel process with drift. By differentiating both sides of the above equation, we obtain that

$$\mathbb{P}^{\mu}(X_S \in dy) = \frac{\eta}{\eta + \frac{\mu^2}{2}} e^{-\mu(X_0 - y)} \mathbb{E}^{P_0}\left[\exp\left\{-\mu\left(\nu + \frac{1}{2}\right)\int_0^T \frac{1}{X_u} du\right\}; X_T \in dy\right].$$

Now let us identify

$$\gamma = \mu(\nu + \frac{1}{2}), \quad \eta + \frac{\mu^2}{2} = \lambda. \tag{40}$$

It follows that $\mu = \frac{\gamma}{\nu + \frac{1}{2}}$ and $\eta = \lambda - \frac{2\gamma^2}{(2\nu+1)^2}$. Therefore

$$\mathbb{E}^{P_0}\left[\exp\left\{-\gamma \int_0^T \frac{1}{X_u} du\right\}; X_T \in dy\right]$$

$$= \mathbb{P}^{\mu}(X_S \in dy) \cdot \frac{\eta + \frac{\mu^2}{2}}{\eta} e^{-\mu(X_0 - y)} \tag{41}$$

$$= (\eta + \frac{\mu^2}{2}) e^{-\mu(X_0 - y)} \int_0^{+\infty} e^{-\eta t} p_{\mu}(t; X_0, y) dt.$$

Thus, by Lemma 1 (Resolvent Kernel) in Linetsky (2004) [74], we identify s with η as well as

$$\kappa(s) = \kappa(\eta) = -\frac{\gamma}{\sqrt{2\lambda}}.$$

Thus, the result in Lemma 10 follows immediately. $\square$

Next, it deserves to notice the following explicit Laplace transform inversion.

LEMMA 11. (A Closed-form Inversion of a Laplace transform involving Whittaker functions) *For $\nu > -\frac{1}{2}$ and $0 \leqslant x \leqslant z$ the following explicit Laplace transform inversion holds*

$$\mathcal{L}_{\gamma}^{-1}\left(\Gamma\left(\frac{1}{2}+\nu+\gamma\right)M_{-\gamma,\nu}(x^2)\cdot W_{-\gamma,\nu}(z^2)\right)(y)$$
$$=\frac{\Gamma(2\nu+1)xz}{2\sinh\left(\frac{y}{2}\right)}\exp\left\{-\frac{(x^2+z^2)\cosh\left(\frac{y}{2}\right)}{\sinh\left(\frac{y}{2}\right)}\right\}I_{2\nu}\left(\frac{xz}{\sinh\left(\frac{y}{2}\right)}\right), \tag{42}$$

where the Hyperbolic functions are defined as

$$\sinh(x)=\frac{1}{2}(e^x-e^{-x}), \quad \cosh(x)=\frac{1}{2}(e^x+e^{-x}). \tag{43}$$

Here

$$I_{2\nu}(x)=\sum_{k=0}^{+\infty}\frac{(\frac{x}{2})^{2(\nu+k)}}{k!\Gamma(2\nu+k+1)}$$

is a modified Bessel function of the first kind with index 2ν

Finally, the joint density 3.20 can be obtained by inverting the Laplace transform in Lemma 10.

Proof.

$$\mathbb{P}_0\left(X_T\in dx, \int_0^T\frac{du}{X_u}\in dt\right)$$

$$=\mathcal{L}_{\gamma}^{-1}\left\{\mathbb{E}^{P_0}\left[\exp\left\{-\gamma\int_0^T\frac{1}{X_u}du\right\}; X_T\in dx\right]\right\}$$

$$=\mathcal{L}_{\gamma}^{-1}\left\{\frac{\sqrt{\lambda}\Gamma(\gamma+\frac{1}{2}+\frac{\gamma}{\sqrt{2\lambda}})}{\sqrt{2}\Gamma(2\nu+1)}\cdot\left(\frac{x}{X_0}\right)^{\nu+\frac{1}{2}}\cdot M_{-\frac{\gamma}{\sqrt{2\lambda}},\nu}(2(X_0\wedge x)\sqrt{2\lambda})\cdot W_{-\frac{\gamma}{\sqrt{2\lambda}},\nu}(2(X_0\vee x)\sqrt{2\lambda})\right\}. \tag{44}$$

By using the scaling property of Laplace transform:

$$\mathcal{L}\{f(at)\}(s) = \frac{1}{|a|} F\left(\frac{s}{a}\right),$$

where $\mathcal{L}\{f(t)\}(s) = F(s)$. By letting $a = \sqrt{2\lambda}$, we find the joint density expression in equation 3.20. $\square$

Detailed ADI Scheme

In this appendix, we document the detailed formulation of the ADI scheme in Section 5.2. In the implementation of (5.12), for each step n, we use a matrix $U^n = (U^n_{i,j})_{(I+1)\times(J+1)}$ to store the approximated values. In the first iteration of the ADI scheme (5.12), for each fixed $j \in \{1, 2, ..., J-1\}$ we set up an $I+1$ dimensional linear system to solve for $U^{n+\frac{1}{2}}_{\cdot,j}$. The system consists of two boundary conditions at $i = 0$ and $i = I$ respectively as well as $I-1$ equations for inner points. The specification of equations of inner points is straightforward. Here we need to figure out the boundary conditions for the ADI middle step points $U^{n+\frac{1}{2}}$. For $i = 0$, it follows from the the Dirichlet boundary condition that

$$U^{n+\frac{1}{2}}_{0,j} = (1 - \lambda\mathcal{A}_2)U^{n+1}_{0,j} + \lambda\mathcal{A}_2 U^n_{0,j} = 0.$$

However, at $i = I$ we operate $1 - \lambda\mathcal{A}_2$ on both sides of the Neumann boundary condition

$$U^{n+1}_{I,j} - U^{n+1}_{I-1,j} = \Delta s.$$

Thus, we find the implied boundary condition for the ADI middle step as

$$U^{n+\frac{1}{2}}_{I,j} - U^{n+\frac{1}{2}}_{I-1,j} = \lambda\mathcal{A}_2 U^n_{I,j} - \lambda\mathcal{A}_2 U^n_{I-1,j} + \left(1 + \frac{r\Delta t}{2j\Delta v}\right)\Delta s.$$

Hence the first iteration of the ADI scheme (5.12) is translated to a tri-diagonal linear system which can be solved efficiently by the Thomas algorithm of linear complexity.

$$\begin{cases} U_{0,j}^{n+\frac{1}{2}} = 0, \\ (1-\lambda\mathcal{A}_1)U_{1,j}^{n+\frac{1}{2}} = [\mathbf{1} + A_0 + (\mathbf{1}-\lambda)\mathcal{A}_1 + \mathcal{A}_2]U_{1,j}^n, \\ \vdots \\ (1-\lambda\mathcal{A}_1)U_{I-1,j}^{n+\frac{1}{2}} = [\mathbf{1} + A_0 + (\mathbf{1}-\lambda)\mathcal{A}_1 + \mathcal{A}_2]U_{I-1,j}^n, \\ U_{I,j}^{n+\frac{1}{2}} - U_{I-1,j}^{n+\frac{1}{2}} = \lambda\mathcal{A}_2 U_{I,j}^n - \lambda\mathcal{A}_2 U_{I-1,j}^n + \left(1+\frac{r\Delta t}{2j\Delta v}\right)\Delta s, \end{cases} \tag{45}$$

i.e.

$$\begin{pmatrix} 1 & & & & \\ a_{1,j} & b_{1,j} & c_{1,j} & & \\ & \ddots & \ddots & \ddots & \\ & & a_{I-1,j} & b_{I-1,j} & c_{I-1,j} \\ & & & -1 & 1 \end{pmatrix} \begin{pmatrix} U_{0,j}^{n+\frac{1}{2}} \\ U_{1,j}^{n+\frac{1}{2}} \\ \vdots \\ U_{I-1,j}^{n+\frac{1}{2}} \\ U_{I,j}^{n+\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ [\mathbf{1} + A_0 + (\mathbf{1}-\lambda)\mathcal{A}_1 + \mathcal{A}_2]U_{1,j}^n \\ \vdots \\ [\mathbf{1} + A_0 + (\mathbf{1}-\lambda)\mathcal{A}_1 + \mathcal{A}_2]U_{I-1,j}^n \\ \lambda\mathcal{A}_2 U_{I,j}^n - \lambda\mathcal{A}_2 U_{I-1,j}^n + \left(1+\frac{r\Delta t}{2j\Delta v}\right)\Delta s \end{pmatrix}.$$

where

$$a_{i,j} = \frac{\lambda r i \Delta t}{2j\Delta v} - \frac{\lambda i^2 \Delta t}{2},\ b_{i,j} = 1 + \frac{\lambda r \Delta t}{2j\Delta v} + \lambda i^2 \Delta t,\ c_{i,j} = -\frac{\lambda r i \Delta t}{2j\Delta v} - \frac{\lambda i^2 \Delta t}{2}.$$

Similarly we handle the second iteration of the ADI scheme (5.12) as follows. We fix $i \in \{0, 1, 2, ..., I\}$ and set up a linear equation system for solving $U_{i,\cdot}^{n+1}$. We also impose the boundary conditions at $v = 0$ and $v = \infty$. It follows that

$$\begin{cases} \frac{\kappa\theta}{\Delta v}(U_{i,1}^{n+1} - U_{i,0}^{n+1}) + ri(U_{i,0}^{n+1} - U_{i-1,0}^{n+1}) - rU_{i,0}^{n+1} = 0, \\ (1-\lambda\mathcal{A}_2)U_{i,1}^{n+1} = U^{n+\frac{1}{2}} - \lambda\mathcal{A}_2 U_{i,1}^{n}, \\ \vdots \\ (1-\lambda\mathcal{A}_2)U_{i,J-1}^{n+1} = U^{n+\frac{1}{2}} - \lambda\mathcal{A}_2 U_{i,J-1}^{n}, \\ U_{i,J}^{n+1} = \max(i\Delta s - K, 0), \end{cases} \tag{46}$$

i.e.

$$\begin{pmatrix} -r - \frac{\kappa\theta}{\Delta v} & \frac{\kappa\theta}{\Delta v} & & & \\ e_{1,j} & f_{1,j} & g_{1,j} & & \\ & \ddots & \ddots & \ddots & \\ & & e_{I-1,j} & f_{I-1,j} & g_{I-1,j} \\ & & & & 1 \end{pmatrix} \begin{pmatrix} U_{i,0}^{n+1} \\ U_{i,1}^{n+1} \\ \vdots \\ U_{i,J-1}^{n+1} \\ U_{i,J}^{n+1} \end{pmatrix} = \begin{pmatrix} -ri(U_{i,0}^{n+1} - U_{i-1,0}^{n+1}) \\ U^{n+\frac{1}{2}} - \lambda\mathcal{A}_2 U_{i,1}^{n} \\ \vdots \\ U^{n+\frac{1}{2}} - \lambda\mathcal{A}_2 U_{i,J-1}^{n} \\ \max(i\Delta s - K, 0) \end{pmatrix},$$

where

$$e_{i,j} = -\frac{\lambda\sigma_v^2\Delta t}{2\Delta v^2} + \frac{\lambda\kappa}{2}(\frac{\theta}{j\Delta v} - 1)\frac{\Delta t}{\Delta v},\ f_{i,j} = 1 + \frac{\lambda\sigma_v^2\Delta t}{\Delta v^2} + \frac{\lambda r\Delta t}{2j\Delta v},\ g_{i,j} = -\frac{\lambda\sigma_v^2\Delta t}{2\Delta v^2} - \frac{\lambda\kappa}{2}(\frac{\theta}{j\Delta v} - 1)\frac{\Delta t}{\Delta v}.$$

Consideration of Jump Combined with Stochastic Volatility

In this appendix, we outline some ideas on pricing timer call options under stochastic volatility models with jumps (see Bates [4] and Duffie, Pan and Singleton [37]). Unlike the case of Heston's model, the stochastic volatility with jump models might not render semi-closed form formulae for the valuation of timer options. However, we may still propose a conditional Black-Scholes-Merton formula and implement it via Monte Carlo simulation.

Under the risk neutral probability measure, Bates' [4] stochastic volatility with jump model is specified as

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= (r - \lambda\overline{\mu})dt + \sqrt{V_t}(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)}) + d\left[\sum_{i=1}^{N_t}(\xi_i^s - 1)\right], \\ dV_t &= \kappa(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_t^{(1)}, \end{aligned} \tag{47}$$

where $\{(W_t^{(1)}, W_t^{(2)})\}$ is a two dimensional standard Brownian motion; r is assumed to be the constant instantaneous interest rate; $\{N_t\}$ is a Poisson process with constant intensity λ; ξ_i^s is the relative jump size in the stock price. At time t_i, when jump occurs, we have that

$$S_{t_i} = S_{t_i-}\xi_i^s.$$

Let us assume that ξ^s follows a lognormal distribution with mean μ_s and variance σ_s^2, i.e.

$$\xi^s \sim \mathcal{LN}(\mu_s, \sigma_s^2).$$

Because

$$\bar{\mu} = \mathbb{E}\xi_i^s - 1 = \mathbb{E}e^{\mu_s+\sigma_s Z} - 1,$$

where $Z \sim \mathcal{N}(0,1)$, the parameter μ_s and $\bar{\mu}$ are related to each other by the equation

$$\mu_s = \log(1+\bar{\mu}) - \frac{1}{2}\sigma_s^2.$$

Under Bates' model (47), the theoretical realized variance consists of not only the diffusion part but also the compounded Poisson part (see Carr and Wu (2006) [22]).

PROPOSITION 16. *Under Bates' model* (47)*, we have that, for* $t = m\Delta t$*,*

$$\lim_{\Delta t \to 0} \sum_{j=1}^{m} \left(\log \frac{S_{t_j}}{S_{t_{j-1}}}\right)^2 = I_t^J := \int_0^t V_s ds + \sum_{i=1}^{N_t} [\log(\xi_i^s)]^2, \quad a.s. \tag{48}$$

In the continuous-time setting, we consider the first-passage-time:

$$\tau^J = \left\{ u \geq 0, \int_0^u V_s ds + \sum_{i=1}^{N_t} [\log(\xi_i^s)]^2 \geqslant B \right\}. \tag{49}$$

Let us denote

$$T_o = I_\tau^J - B = \int_0^{\tau^J} V_s ds + \sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2 - B \tag{50}$$

the size of overshoot over the variance budget B by the total variance. Similar to Proposition 3, we have the following conditional Black-Scholes-Merton formula for pricing timer options under Bates' stochastic volatility with jump model.

PROPOSITION 17. *The timer call option with strike K and variance budget B can be priced by the following conditional Black-Scholes-Merton formula:*

$$\begin{aligned} C_0 =&\mathbb{E}^{\mathbb{Q}}\Bigg[S_0 \prod_{i=1}^{N_{\tau^J}} \xi_i^s \cdot e^{d_0(V_{\tau^J},\tau^J,\sum_{i=1}^{N_{\tau^J}}[\log(\xi_i^s)]^2,T_o)} N(d_1(V_{\tau^J},\tau^J,\sum_{i=1}^{N_{\tau^J}}[\log(\xi_i^s)]^2,T_o)) \\ &- Ke^{-r\tau^J} N(d_2(V_{\tau^J},\tau^J,\sum_{i=1}^{N_{\tau^J}}[\log(\xi_i^s)]^2,T_o))\Bigg], \end{aligned} \tag{51}$$

where

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du,$$

and

$$\begin{aligned} d_0(v,\xi,p,q) &= \frac{\rho}{\sigma_v}(v - V_0 - \kappa\theta\xi + \kappa(B-p+q)) - \frac{1}{2}\rho^2(B-p+q), \\ d_1(v,\xi,p,q) &= \frac{1}{\sqrt{(1-\rho^2)(B-p+q)}}\left[\log\left(\frac{S_0}{K}\right) + r\xi + \frac{1}{2}(B-p+q)(1-\rho^2) + d_0(v,\xi,p,q)\right], \\ d_2(v,\xi,p,q) &= \frac{1}{\sqrt{(1-\rho^2)(B-p+q)}}\left[\log\left(\frac{S_0}{K}\right) + r\xi - \frac{1}{2}(B-p+q)(1-\rho^2) + d_0(v,\xi,p,q)\right]. \end{aligned} \tag{52}$$

Proof. First, we express (S_t, V_t) as

$$\begin{aligned} S_t &= S_0 \exp\left\{rt - \frac{1}{2}\int_0^t V_s ds + \rho\int_0^t \sqrt{V_s} dW_s^{(1)} + \sqrt{1-\rho^2}\int_0^t \sqrt{V_s} dW_s^{(2)}\right\} \prod_{i=1}^{N_t} \xi_i^s, \\ V_t &= V_0 + \kappa\theta t - \kappa\int_0^t V_s ds + \sigma_v \int_0^t \sqrt{V_s} dW_s^{(1)}. \end{aligned} \tag{53}$$

By conditioning on the variance budget consumption time τ^J, the variance V_{τ^J}, the total number of Poisson jumps N_{τ^J}, the jump sizes $(\xi_i^s)_{i=1}^{N_{\tau^J}}$, the total quadratic variation contributed by jumps $\sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2$ and the overshoot size of the total variation T_o, we find the conditional distribution of S_{τ^J}. We note that all the conditions are based on the variance process and the jump part of the asset dynamics, which are both independent of the Brownian motion $W_t^{(1)}$. Similar argument to the proof of Proposition 3 yields that

$$
\begin{aligned}
&\left(S_{\tau^J}\middle|\tau^J = t, V_{\tau^J} = v, N_{\tau^J} = n, (\xi_i^s)_{i=1}^{N_{\tau^J}} = (\xi_i)_{i=1}^n, \sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2 = p, T_o = q\right)\\
=^{law}&\Bigg(S_0 \prod_{i=1}^{N_t} \xi_i^s \exp\Bigg\{r\tau^J - \frac{1}{2}(B-p+q) + \frac{\rho}{\sigma_v}(V_{\tau^J} - V_0 - \kappa\theta\tau^J + \kappa(B-p+q))\\
&+ \sqrt{1-\rho^2}\int_0^{\tau^J} \sqrt{V_s}dW_s^{(2)}\Bigg\}\Bigg|\tau^J = t, V_{\tau^J} = v, N_{\tau^J} = n, (\xi_i^s)_{i=1}^{N_{\tau^J}} = (\xi_i)_{i=1}^n,\\
&\sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2 = p, T_o = q\Bigg)\\
=^{law}&S_0 \prod_{i=1}^{n} \xi_i \exp\Bigg\{N\Bigg(rt - \frac{1}{2}(B-p+q) + \frac{\rho}{\sigma_v}(v - V_0 - \kappa\theta t + \kappa(B-p+q)),\\
&(1-\rho^2)(B-p+q)\Bigg)\Bigg\}.
\end{aligned}
\tag{54}
$$

Therefore, by conditioning and further computation based on the standard normal distribution, we obtain that

$$
\begin{aligned}
C_0 =&\mathbb{E}^Q\left\{\mathbb{E}^Q\left[e^{-r\tau^J}\max(S_{\tau^J} - K, 0)\middle|(\xi_i^s)_{i=1}^{N_{\tau^J}}, N_{\tau^J}, \tau^J, V_{\tau^J}, \sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2, T_o\right]\right\}\\
=&\mathbb{E}^{\mathbb{Q}}\Bigg[S_0 \prod_{i=1}^{N_{\tau^J}} \xi_i^s \cdot e^{d_0(V_{\tau^J}, \tau^J, \sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2, T_o)} N(d_1(V_{\tau^J}, \tau^J, \sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2, T_o))\\
&- Ke^{-r\tau^J} N(d_2(V_{\tau^J}, \tau^J, \sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2, T_o))\Bigg].
\end{aligned}
\tag{55}
$$

$\square$

A conditional Monte Carlo simulation scheme can be implemented based on Proposition 17. Similar to Algorithm 1, we perform the exact simulation of the path of variance process $\{V_t\}$ and the jump components in (48). By the definition of τ^J in (49), we simulate it by a "time-checking" algorithm (see 1). Meanwhile, we record the size of overshoot T_o and the quadratic variation contributed by jumps, i.e. $\sum_{i=1}^{N_{\tau^J}} [\log(\xi_i^s)]^2$. Based on all these steps, we calculate the estimator:

$$S_0 \prod_{i=1}^{N_{\tau^J}} \xi_i^s \cdot e^{d_0} N(d_1) - K e^{-r\tau^J} N(d_2), \tag{56}$$

where $d_0,\ d_1,\ d_2$ are evaluated as in Proposition 17.

The Malliavin-Watanabe-Yoshida Theory: A Primer

In recent years, various researches on the applications of Malliavin calculus in quantitative finance are vividly carried out. For example, Fournie et al. (1999) [45] proposed to use Malliavin integration by parts formula to derive Monte Carlo simulation estimator for computing price sensitivities. Among many other applications, Malliavin calculus provides researchers a powerful tool to justify asymptotic expansion of derivative security payoffs and thus its no-arbitrage price approximation. Some examples applied to interest rates modeling can be found in Kunitomo and Takahashi (2001) [70] and Osajima (2007) [82], etc. More topics on the applications of Malliavin calculus in finance can be found in Malliavin and Thalmaier (2006) [78].

.1 Basic Setup of the Malliavin Calculus Theory

The Malliavin calculus, also known as the stochastic calculus of variations, is an infinite-dimensional differential calculus on the Wiener space. It was originally tailored to investigate regularity properties of the law of Wiener functionals such as solutions of stochastic differential equations. The theory was initiated by Malliavin and further developed by Stroock, Bismut, Watanabe, et al. Roughly speaking, the Malliavin calculus theory consists of two parts. First is the theory of the differential operators defined on Sobolev spaces of Wiener functionals. The second part of the theory establishes general criteria for a given random vector to possess a smooth density.

Instead of making an exhaustive survey of the Malliavin calculus theory, we briefly list some terminologies related to the asymptotic expansion theory of Malliavin-Watanabe-Yoshida. To learn more about the fundamentals of Malliavin calculus, readers are referred to Nualart (2006) [80], Malliavin (1997) [77], Kusuoka and Stroock (1991) [72] and Fritz (2005) [46], etc.

Let us denote $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\})$ the d-dimensional filtered Wiener space, where $\Omega = C_0([0,T], \mathbf{R}^d)$. The coordinate process $\{w(t)\}$ is a d-dimensional Brownian motion under the Wiener measure $\mathbb{P}$. Let H be the Cameron-Martin subspace of Ω, i.e.

$$H = \left\{ h = \left(\int_0^{\cdot} \dot{h}^1(s)ds, ..., \int_0^{\cdot} \dot{h}^d(s)ds \right); \quad \dot{h} \in L^2[0,T] \right\}.$$

The inner product of Hilbert space H is defined as

$$\langle h_1, h_2 \rangle_H = \sum_{k=1}^{d} \int_0^T \dot{h}_1^k(s) \dot{h}_2^k(s) ds,$$

for any $h_1, h_2 \in H$. The norm is equipped with

$$||h||_H = \left(\sum_{k=1}^{d} \int_0^T |\dot{h}^k(t)|^2 dt \right)^{\frac{1}{2}},$$

for $h \in H$. Let $F : \Omega \to \mathbb{R}$ be an $\mathcal{F}_T$ measurable random variable, which is also called Wiener functional. We further assume that $F \in L^p(\Omega)$, where $p > 1$. Given $h \in H$, we define the directional derivative of F as

$$D_h F(w) := \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} F(w+\epsilon h) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[F(w+\epsilon h)-F(w)\right]. \tag{57}$$

Thus, $D.F(w)$ is defined as a linear functional on Hilbert space H. By Riesz representation theory, there exists an element $D_sF(w) := (D^1F(w), D^2F(w), ..., D^dF(w)) \in H$ such that

$$D_hF(w) = \langle \dot{h}, DF(w)\rangle_H = \sum_{k=1}^{d}\int_0^T \dot{h}^k(s)D_s^kF(w)ds. \tag{58}$$

Denote $\mathbf{L}^p(\Omega : H)$ the collection of measurable map f from Ω to H such that $||f||_H \in \mathbf{L}^p(\Omega)$. If $DF \in \mathbf{L}^p(\Omega : H)$, DF is defined as the Malliavin derivative of F. It can be regarded as an H-valued random variable or a d-dimensional stochastic process. Consequently we are able to define higher order Malliavin derivatives. Let $\mathbf{P}$ denote the collection of polynomials on the Wiener space Ω. Let us define the s-times Malliavin norm $||\cdot||_{s,p}$ as

$$\|F\|_{s,p} = \left[\mathbb{E}\|F\|^p + \sum_{j=1}^{s}\mathbb{E}\|D^{(j)}F\|^p_{H^{\otimes j}}\right]^{\frac{1}{p}}. \tag{59}$$

By completing $\mathbf{P}$ (in $L^p(\Omega)$) according to norm $||\cdot||_{s,p}$, we construct a Banach space denoted by D_p^s, which collects all s-times Malliavin differentiable variables.

It is well known that, for $s \in \mathbb{N}$, norm (59) is equivalent to the one defined as follows. For any Wiener functional g, let

$$|||g|||_{s,p} = \| (I-\mathcal{L})^{\frac{s}{2}} g \|_{L^p(\Omega)}, \tag{60}$$

where $\mathcal{L}$ is the Ornstein-Uhlenbeck operator. However, using norm (60), we can obtain space D_p^s for any arbitrary real number $s \in \mathbb{R}$ by completing the Wiener polynomial space $\mathbf{P}$. For any $s \in \mathbb{N}$, the equivalence ensures that the completion using norm $|||\cdot|||_{s,p}$ also generates the space D_p^s. It is well known that the dual space of D_p^s is D_q^{-s}, i.e. $(D_p^s)' = D_q^{-s}$, where

$$\frac{1}{p} + \frac{1}{q} = 1.$$

If $s < 0$ the elements in D_p^s may not be ordinary random variables and they are usually interpreted as distributions on the Wiener space. If $F \in D_p^s$ and $G \in D_q^{-s}$, we denote the paring $\langle F, G\rangle$ by $\mathbb{E}(FG)$. Thus, we interpret norm $||\cdot||_{-s,q}$ as

$$||G||_{-s,q} = \sup \mathbb{E}(FG),$$

where the sup is taken over set

$$\{F \in D_p^s : ||F||_{s,p} < 1\}.$$

Consider the intersection

$$D^{\infty} := \bigcap_{p>1} \bigcap_{k>1} D_p^k, \quad D^{-\infty} := \bigcap_{p>1} \bigcup_{k>1} D_p^{-k}.$$

Here, D^{∞} is the set of Wiener functionals. $D^{-\infty}$ is the set of generalized Wiener functionals.

The original motivation of the Malliavin calculus theory has been to give a probabilistic proof of Hörmander's "sum of squares" theorem. The key task in this regard is to establish general criteria in terms of the Malliavin covariance matrix for a given random vector to possess a smooth density. In the applications of Malliavin calculus to specific examples, one usually tries to find sufficient conditions for these general criteria to be satisfied. To apply the Malliavin-Watanabe-Yoshida thoery of asymptotic expansion, it is crucial for us to justify the validity through the Malliavin covariance of the underlying variables.

Let G be a m-dimensional random variable $G = (G^1, G^2, ..., G^m)$ and denote

$$DG = (D^j G^i)_{m \times d},$$

which appears as a $m \times d$-matrix of H-valued random variable. The Malliavin covariance matrix is defined as

$$\Sigma(G) = (\Sigma(G)_{ij})_{m \times m}, \tag{61}$$

where

$$\Sigma(G)_{ij} = \langle DG^i, DG^j \rangle_H = \sum_{k=1}^{d} \langle D^k G^i, D^k G^j \rangle_H.$$

REMARK 15. The analog of Malliavin derivative operator in finite dimensional space $\mathbb{R}^n$ is the gradient operator. Given a $\mathcal{C}^1$-function $f : \mathbb{R}^n \to \mathbb{R}$, the derivative along direction $\overrightarrow{n}$ can be computed as

$$\frac{\partial f}{\partial \overrightarrow{n}} = \nabla f \cdot \overrightarrow{n}.$$

The resemblance to (58) convinces us that the Malliavin calculus generalizes the ordinary finite dimensional calculus to infinite dimensional settings. For more such analogies between the Malliavin calculus and the ordinary finite dimensional calculus, see Friz (2002) [46].

.2 The Malliavin-Watanabe-Yoshida Theory of Asymptotic Expansion

Suppose the random variable $G(\epsilon)$ admits the asymptotic expansion in D^∞, i.e.

$$G(\epsilon) = G_0 + \epsilon G_1 + \epsilon^2 G_2 + \cdots + \epsilon^n G_n + \mathcal{O}(\epsilon^{n+1}), \quad \text{in } D^\infty. \tag{62}$$

In other words,

$$\limsup_{\epsilon \to 0} \frac{1}{\epsilon^{n+1}} \left\| G(\epsilon) - [G_0 + \epsilon G_1 + \epsilon^2 G_2 + \cdots + \epsilon^n G_n] \right\|_{D_p^s} < +\infty, \quad \text{for any } p > 1, s > 0, \tag{63}$$

where $G_i \in D^\infty$ for $i = 0, 1, 2, 3, ...$

Let $\mathcal{S}$ be the real Schwartz space of rapidly decreasing C^∞-functions. Denote $\mathcal{S}'$ the space of Schwartz tempered distribution, which is the dual space of $\mathcal{S}$. Now, for $T \in \mathcal{S}'$, the question is under what condition we have that

$$\mathbb{E}[T(G(\epsilon))] = \mathbb{E}\Phi_0 + \epsilon \mathbb{E}\Phi_1 + \epsilon^2 \mathbb{E}\Phi_2 + \cdots + \epsilon^n \mathbb{E}\Phi_n + \mathcal{O}(\epsilon^{n+1}), \tag{64}$$

where

$$\Phi_0 = T(G_0), \quad \Phi_1 = \frac{\partial T}{\partial x}(G_0) G_1, \quad \Phi_2 = \frac{1}{2}\frac{\partial^2 T}{\partial x^2}(G_0) G_1^2 + \frac{\partial T}{\partial x}(G_0) G_2, \quad ... . \tag{65}$$

It is sufficient to answer under what condition we have the asymptotic expansion for generalized Wiener functional $T(G(\epsilon))$, i.e.

$$T(G(\epsilon)) = \Phi_0 + \epsilon \Phi_1 + \epsilon^2 \Phi_2 + \cdots + \epsilon^n \Phi_n + \mathcal{O}(\epsilon^{n+1}), \quad \text{in } D^{-\infty}. \tag{66}$$

In other words, there exists $s > 0$ such that, for all $p > 1$

$$\limsup_{\epsilon \to 0} \frac{1}{\epsilon^{n+1}} \left\| T(G(\epsilon)) - [\Phi_0 + \epsilon\Phi_1 + \epsilon^2\Phi_2 + \cdots + \epsilon^n\Phi_n] \right\|_{D_p^{-s}} < +\infty, \tag{67}$$

where $\Phi_i \in D_p^{-s}, \quad \text{for } i = 0, 1, 2, 3, \ldots$.

Watanabe (1987) successfully interpreted the composition functional $T(G(\epsilon))$ as generalized Wiener functionals (i.e. the Schwartz distribution on the probability space.) He justified that the asymptotic expansion in (66) and (64) is valid if the Malliavin covariance of $G(\epsilon)$ is uniformly non-degenerated in the sense that

$$\limsup_{\epsilon \to 0} \mathbb{E}[\det(\Sigma(G(\epsilon)))^{-p}] < \infty, \quad \text{for all } p \in (0, +\infty). \tag{68}$$

A crucial step to apply this successful theory is to verify the non-degeneracy of the Malliavin covariance. This is even not easy to do for some simple cases where the underlying Malliavin covariance is expressed by an integration of some adaptive processes. Yoshida (1992, 1993) [93, 92, 94] not only successfully applies the theory to statistical inference, but also gives a truncated version of the asymptotic expansion theory, in which verification of the uniform non-degeneracy of Malliavin covariance becomes easier. According to Yoshida (1992, 1993) [93, 92, 94], the validity of the expansion can be obtained if there exists a random sequence $\{\eta^\epsilon\} \subseteqq D^\infty$ such that the following two conditions are verified.

Condition.1: Uniform Non-degeneracy under Truncation

$$\sup_{\epsilon \in [0,1]} \mathbb{E}[1\{\eta^\epsilon \leq 1\}(\det(\Sigma(G(\epsilon))))^{-p}] < +\infty; \tag{69}$$

Condition.2: Negligible Probability of Truncation

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon^n} \mathbb{P}\left(|\eta^\epsilon| > \frac{1}{2} \right) = 0, \quad \text{for any } n = 1, 2, 3, ... \tag{70}$$

By reasonably modifying Watanabe (1987), Yoshida (1992, 1993) [93, 92, 94] establishes that the generalized Wiener functional $\Psi(\eta^\epsilon)T(G(\epsilon))$ admits the following expansion

$$\Psi(\eta^\epsilon)T(G(\epsilon)) = \Phi_0 + \epsilon\Phi_1 + \epsilon^2\Phi_2 + \cdots, \quad \text{in } D^{-\infty}, \tag{71}$$

where $\Psi : \mathbf{R} \to \mathbf{R}$ is a smooth function such that $0 \leq \Psi(x) \leq 1$ for $x \in \mathbf{R}$, $\Psi(x) = 1$, for $|x| < 1/2$ and $\Psi(x) = 0$, for $|x| \geq 1$. Thus, (64) follows directly. Interested readers are also referred to Uchida and Yoshida (2004) [89], Kunitomo and Takahashi (2001) [70, 71] for the development and application of the Malliavin-Watanabe-Yoshida theory in the valuation of contingent claims.