Regime-Switching Models with Mean Reversion and Applications in Option Pricing

by

Jie Yu

(Under the direction of Qing Zhang)

Abstract

In option pricing the underlying stock price is traditionally assumed to follow a geometric Brownian motion. However it is observed that the stock prices often switch between geometric Brownian motion and mean reversion behaviors. The contributions in this dissertation include constructing a regime-switching model incorporating both geometric Brownian motion and mean reversion models in which the switching is determined by a finite state Markov chain. This model leads to an effective mathematical framework for studying the valuation of the corresponding financial derivatives. Then we use a PDE method as well as a viscosity solution method to solve the non-smooth boundary value problem and to characterize the pricing of European options. In the second part of this dissertation we obtain a closed-form pricing formula for European call options using a successive approximation approach. A stochastic approximation method is used to estimate parameters under this model. Numerical experiments are carried out to compare our results with that of Monte Carlo simulation. Our effort is also devoted to providing applications involving model calibration and prediction of stock market trends using option market data. Finally, the pricing of perpetual American put options under the mean reversion model is studied. We obtain a closed-form solution in this case.

INDEX WORDS: Regime Switching, Mean Reversion, European Option Pricing, PDE method, Successive Approximation, Viscosity Solution, Stochastic Approximation

Regime-Switching Models with Mean Reversion and Applications in Option Pricing

by

Jie Yu

B.S., Fudan University, 2004

M.A., University of Georgia, 2006

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy,

Athens, Georgia

2009

Regime-Switching Models
with Mean Reversion
and Applications
in Option Pricing

by

Jie Yu

Approved:

Major Professor: Qing Zhang

Committee: Dhandapani Kannan
Edward A. Azoff
Jingzhi Tie
Robert Varley

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2009

Dedication

To my parents Airong He and Chunsheng Yu.

To my husband Zhaozhi Li.

Acknowledgments

My sincere gratitude goes to my advisor, Professor Qing Zhang, for inspiring and encouraging me to pursue a career in mathematical finance. I am very grateful to have his guidance and advices during my studies at the University of Georgia. He always gives me valuable suggestions when I pursue my career. I deeply appreciate his kindness and constant supports.

I would like to thank my committee members, Professor Dhandapani Kannan and Professor Edward A. Azoff for proofreading and commenting on my dissertation. I have learned a lot from them. I would also like to thank my committee members, Professor Robert Varley and Professor Jingzhi Tie for their helpful comments and serving on my committee.

I would like to thank Professor George Yin who shared his knowledge with me. I would also like to thank everyone who helped me during these five years I spent in this lovely town, Athens of Georgia.

Finally, I would like to thank my parents and my husband, who have provided huge support during all these years.

Table of Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1.1 Financial Derivatives

The development of various financial derivatives is one of the most noticeable events in the financial world during the past two decades. The fundamental changes in financial markets including changes in the foreign exchange markets, the credit markets and the capital markets as well as innovations in financial theory and increasing computerization over this period, have contributed to the growth of these financial derivatives.

Financial derivatives are financial contracts whose values are derived from the value of the underlying assets which can be commodities, stocks, indexes, weather conditions or other items. The main types of derivatives are forwards, futures, options, and swaps. They can be used to manage the risk of economic loss incurred at changes in the value of the underlying assets as well as increase the profit arising when the value of the underlying assets moves in a favorable way. In this dissertation we study stock option valuation problems under several model setups. A call option gives the holder the right to buy the underlying asset by a specified date(expiration date) for a pre-determined price(strike price). A put option gives the holder the right to sell the underlying asset by a specified date for a pre-determined price. European style option can only be exercised on the expiration date. On the other hand, American option can be exercised any time up to the expiration date.

Valuation of stock options has become a very popular research topic which has applications in both hedging and speculation. There has been a great deal of interest in using mathematical models in pricing options. The Black-Scholes model and various subsequent modifications have provided a reasonably good description of the market. And the Black-Scholes formula provides a closed-form solution to the European option pricing problem. On the other hand, pricing American put options is challenging due to the possibility of early exercise. The traditional American option valuation methods such as lattice and tree-based techniques are not very practical, because the computational cost increases rapidly as the number of underlying securities and other payoff-related variables increase. Instead of the traditional methods, Monte Carlo simulation is more practical and widely used.

1.2 Review of the Black-Scholes Model

One of the most famous results in the financial literature is the widely-used Black-Scholes model. The Black-Scholes model is based on some theoretical assumptions including that stock price behavior follows geometric Brownian motion. This model can be used to derive European option prices under the assumption that the dividends will not be paid during the life of the option, by using the critical determinants of the stock price, strike price, time to expiration, volatility and short-term risk free interest rate.

In the Black-Scholes model, the price of the underlying instrument $S(t)$ follows a geometric Brownian motion with constant drift μ and volatility σ given by the following stochastic differential equation,

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t), \tag{1.2.1}$$

where σ is the volatility and $W(t)$ is a Wiener process.

We assume that there are no arbitrage opportunities(i.e. the opportunity to buy an asset at a low price and then immediately sell it on a different market for a higher price); the stock pays no dividends during the option's life ; trading in the stock is continuous; there is no transaction costs or taxes; all securities are perfectly divisible (e.g. it is possible to buy 1/100th of a share); it is possible to borrow and lend cash at a constant risk-free interest rate.

Under these assumptions, Black and Scholes obtained the Black-Scholes formula in " The Pricing of Options and Corporate Liabilities" [4] in 1973. This formula became the foundation of modern mathematical finance. It gives the price of a European option by taking into account the price of the underlying stock, the volatility of the underlying, the time remaining before the option expires and the time value of money.

By Ito's Lemma, the Black-Scholes PDE is given as follows:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

where V represents the value of an option. The Black-Scholes PDE leads to the following formula for the price of a European call option with exercise price K on a stock currently trading at price $S(0)$ (i.e. the right to buy a share of the stock at price K at maturity T). The constant interest rate is r, and the constant stock volatility is σ.

At time $t = 0$ the call option price is given by the famous Black-Scholes formula

$$c = e^{-rT} E^x (S(T) - K)^+ = x\Phi(d_1) - Ke^{-rT}\Phi(d_2), \tag{1.2.2}$$

where Φ is the standard normal cumulative distribution function, and

$$d_1 = \frac{\log \frac{x}{K} + T(r + \frac{\sigma^2}{2})}{\sigma\sqrt{T}},$$

$$d_2 = d_1 - \sigma\sqrt{T}.$$

Similarly the put option price is

$$p = Ke^{-rT}\Phi(-d_2) - x\Phi(-d_1). \tag{1.2.3}$$

The Black-Scholes model and its corresponding results are widely used in industry. However, this model has some major limitations: It doesn't capture the volatility smile(i.e.

implied volatility of options based on the same underlying asset and same expiration date displays a U-shape across the various strike prices). It does not include the possibility of early exercise of an American option, unless the American call is on a non-dividend paying asset in which the call is always worth the same as its European equivalent as there is never any advantage in exercising early. It also has serious discrepancies due to its insensibility to random parameter changes such as changes in market trends.

Chapter 2

A Regime-Switching Model with Mean Reversion

2.1 Introdution

Mean reversion models are widely used in studies of energy and commodity markets. They capture movements of prices which tend to get back to some "equilibrium" level. In the financial literature, there is substantial empirical evidence that supports mean-reversion modeling. The predictability of stock returns in connection with mean reversion is also well researched topic in the empirical literature on financial economics.

In particular, studies that support mean reversion stock returns can be traced back to the 1930's; see Cowles and Jones [10]. Among many other studies, Fama and French [16] reported their findings of occurrence of mean reverting U.S. stock prices so that the predictable price variation accounts for large fractions of return variances over return horizons of several years; Poterba and Summers [27] used statistical tools to study the mean reversion of U.S. stock prices; in the paper of Gallagber and Taylor [17], the size and significance of the mean-revering component in U.S. stock prices is investigated, for the period from January 1949 to December 1997; Balvers, Wu and Gilliland [1] provided international evidence to support mean-reverting stock prices in 18 countries during the period 1969 to 1996.

Mean-reversion models are also popular with other applications. In Hafner and Herwartz's paper [21], mean reversion model is used to characterize stochastic volatility. In Blanco and Soronow's paper [5] it is used to price assets in energy markets; Bos, Ware and

Pavlov [7]'s studies are concerned with option pricing with a mean-reversion asset. Recently, Zhang and Zhang [38] developed an optimal trading rule involving buy-low and sell-high strategy under a mean reversion model.

In order to generate useful results in pricing financial derivatives and hedging applications it is necessary to incorporate mean-reversion of stock price in traditional modeling. There is much research development in stock market models. For example, studies of regime-switching were first made by Hamilton in 1989 [23] to describe a regime-switching time series; Marsi, Kabanov and Runggaldier researched mean variance hedging of regime-switching European options on stocks; Bauwens, Preminger and Rombouts [2] developed univariate regime-switching GARCH models wherein the conditional variance switches in time from one GARCH process to another; Bollen [6] employed lattice method and simulation to value options in regime-switching models; Guo [18] provided an explicit solution to an optimal stopping problem with regime switching; Zhang's paper [37] proposed a hybrid switching GBM model(Black-Scholes model) involving a number of GBMs modulated by a finite-state Markov chain with application in optimal selling rules, such switching processes can be used to represent market trends or the trends of an individual stock; Yao, Zhang and Zhou [32] also studied the European option prices under a regime switching model in which the rate of return and the volatility of the underlying asset depend on the market modes: "bullish" or "bearish".

The use of a regime-switching model helps to improve the traditional Black-Scholes model by allowing the parameters of stocks to have different structure over time. Our contributions in this chapter are constructing a mean-reversion model and the related regime-switching model as well as deriving a European option pricing formula under the mean-reversion model. The regime-switching framework incorporating both Black-Scholes and mean-reversion models into the system better reflects price movement in the real market place. In this model, parameters' change over time is determined by a finite-state Markov chain which is a discrete stochastic process.

2.2 The Mean-Reversion Model

This section is concerned with the description of the mean reversion model. The solution of stock prices in this model is also provided. The idea of the mean reversion model is that the stock prices fluctuate around some "equilibrium" level. When the price is moving away from that "equilibrium" level it will be eventually pulled back to it. And the further it is away the stronger the force is to pull it back.

Suppose the log-price $X(t)$ defined by $X(t) = \log S(t)$ is a mean-reverting process satisfying the following stochastic differential equation(SDE):

$$dX(t) = a(L - X(t))dt + \sigma dW(t), \tag{2.2.1}$$

with initial value $X(s) = x$, where L represents the "equilibrium" level, a determines how fast the log price gets back to the "equilibrium" level, σ is the volatility and $W(t)$ is a Wiener process.

Implied by the mean-reverting process (2.2.1), $X(t) > L$ results in a negative drift term $a(L - X(t))dt$, pulling $X(t)$ back down toward the equilibrium level. Similarly $X(t) < L$ results in a positive drift term, pulling $X(t)$ back up to the equilibrium level. The rate of reversion a determines the reversion speed of $X(t)$. The greater a is, the faster $X(t)$ converges to the equilibrium value. In addition, due to the stochastic term $dW(t)$, the value of $X(t)$ tends to oscillate around the equilibrium level.

To simplify the above stochastic differential equation, we define

$$V(t) = e^{at}(X(t) - L).$$

Using Ito's formula, we have

$$dV(t) = \sigma e^{at} dW(t).$$

Lemma 2.2.1. *If a stochastic process $V(t)$ satisfies $dV(t) = \sigma e^{at} dW(t)$, for $0 \le t \le T$, then*

$$V(t) = V_0 + W(\phi_t^{-1}). \tag{2.2.2}$$

where $\phi_t = \log(2at/\sigma^2 + 1)/(2a)$ and $\phi_t^{-1} = \sigma^2(e^{2at} - 1)/(2a)$.

Proof. We use change-of-time method in [30, Theorem 8.5.2]. Let $b(t, w) = b(t) = \sigma e^{at}$. This is bounded over $[0, T]$, $\sigma \le b(t) \le e^{aT}\sigma$. Define

$$\phi_t = \int_0^t \frac{e^{-2a\phi_s}}{\sigma^2} ds$$

with initial value $\phi_0 = 0$. Then by the change of time method, we have

$$V(t) = V_0 + W(\phi_t^{-1}).$$

Taking the derivative of ϕ_t, we get

$$\phi_t' = \frac{e^{-2a\phi_t}}{\sigma^2} = f(\phi_t).$$

Then the inverse function can be obtained by

$$\begin{aligned} t &= \int \frac{d\phi_t}{f(\phi_t)} + C_1 \\ &= \sigma^2 \int e^{2a\phi_t} d\phi_t + C_1 \\ &= \frac{\sigma^2}{2a} e^{2a\phi_t} + C. \end{aligned}$$

Since $\phi_0 = 0$, the constant must be $C = -\sigma^2/(2a)$.

Therefore we have $\phi_t^{-1} = e^{2a\phi_t}\sigma^2/(2a) - \sigma^2/(2a) = \sigma^2(e^{2at} - 1)/(2a)$.

It follows that $\phi_t = \log(2at/\sigma^2 + 1)/(2a)$. □

Using (2.2.2), we can write $X(t)$ in terms of $W(\phi_t^{-1})$ as follows:

$$X(t) = e^{-at}V(t) + L = e^{-at}V(0) + e^{-at}W(\phi_t^{-1}) + L.$$

The corresponding stock price $S(t)$ is

$$S(t) = e^{X(t)} = \exp\,[e^{-at}V(0) + L + e^{-at}W(\phi_t^{-1})],$$

where

$$\phi_t^{-1} = \sigma^2(e^{2at} - 1)/(2a).$$

2.3 European Call Option with Mean-Reversion

In this section, we focus on the valuation of European call options under the mean reversion model. We derive a closed-form formula and then study the dependence on various parameter values. Examples with different parameter values are given.

Theorem 2.3.1. *Suppose the stock log-price follows a mean reversion trend. Then the European call option price* $c = E^x[e^{-rT}(S_T - K)^+]$ *can be obtained by the following formula*

$$c = \mathrm{C}_0 + Ke^{-rT}\Phi(\tilde{d}_1) - b\Phi(\tilde{d}_2), \tag{2.3.3}$$

where Φ *is a standard Gaussian distribution function and*

$$\begin{aligned}
\mathrm{C}_0 &= e^{-rT}[\exp(L + e^{-aT}(X_0 - L) + \tfrac{\sigma^2}{4a}(1 - e^{-2aT})) - K],\\
\tilde{d}_1 &= \tfrac{1}{\sqrt{\phi_T^{-1}}}(e^{aT}(\log K - L) - (X_0 - L)),\\
\tilde{d}_2 &= \tilde{d}_1 - e^{-aT}\sqrt{\phi_T^{-1}},\\
\phi_T^{-1} &= \tfrac{\sigma^2}{2a}(e^{2aT} - 1),\\
b &= \exp[-rT + e^{-aT}(X_0 - L) + L + \tfrac{e^{-2aT}}{2}\phi_T^{-1}].
\end{aligned} \tag{2.3.4}$$

Proof. To get the closed-form solution of call option's price, we calculate the expectation by computing the integrals with respect to the Gaussian process $W(\phi_t^{-1})$ as follows.

$$\begin{aligned}
c &= E^x[e^{-rT}(S(T) - K)^+]\\
&= E^x e^{-rT}(\exp\,[e^{-at}V(0) + L + e^{-at}W(\phi_t^{-1})] - K)^+
\end{aligned}$$

$$= e^{-rT}\int_{(W(\phi^{-1})>e^{aT}(\log K-L)-V(0))}(\exp(e^{-aT}(V(0)+u)+L)-K)\frac{1}{\sqrt{2\pi\phi_t^{-1}}}e^{-\frac{u^2}{2\phi_T^{-1}}}du$$
$$= \frac{e^{-rT}}{\sqrt{2\pi\phi_t^{-1}}}\int_{(u>e^{aT}(\log K-L)-V(0))}\exp\left(-\frac{u^2}{2\phi_T^{-1}}\right)(\exp(e^{-aT}(V(0)+u)+L)-K)du.$$

To simplify the above integral, define,

$$P = e^{aT}(\log K - L) - V(0),$$
$$V(0) = \log S(0) - L,$$
$$\phi_T^{-1} = \frac{\sigma^2}{2a}(e^{2aT}-1),$$
$$A = -\frac{1}{2\phi_T^{-1}} < 0,$$
$$B = e^{-aT},$$
$$C = e^{-aT}V(0) + L.$$

Using these symbols, we have

$$c = \frac{e^{-rT}}{\sqrt{2\pi\phi_t^{-1}}}\left[\int_{(u>P)}\exp\left[-\frac{u^2}{2\phi_T^{-1}}+e^{-aT}(u+V(0))+L\right]du - K\int_{(u>P)}\exp(-\frac{u^2}{2\phi_T^{-1}})du\right]$$
$$= \frac{e^{-rT}}{\sqrt{2\pi\phi_t^{-1}}}\left[\int_{(u>P)}e^{Au^2+Bu+C}du - K\int_{(u>P)}e^{Au^2}du\right].$$

Consider the second integral first. Using change of variables $v = \sqrt{-2A}u$, we have

$$\int_{(u>P)}e^{Au^2}dx = \int_P^\infty e^{Au^2}du$$
$$= \frac{\sqrt{2\pi}}{\sqrt{-2A}}\int_{\sqrt{-2A}P}^\infty \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}}dv$$
$$= \sqrt{\frac{\pi}{-A}}\left[1-\int_{-\infty}^{\sqrt{-2A}P}\frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}dv\right]$$
$$= \sqrt{\frac{\pi}{-A}}[1-\Phi(\sqrt{-2A}P)].$$

Similarly we change variables $y = u + B/(2A)$ and compute the first integral as follows:

$$\begin{aligned}
&\int_{(u>P)} \exp(Au^2 + Bu + C)du \\
&= \int_{(u>P)} \exp\left[A\left(u + \frac{B}{2A}\right)^2 + \left(C - \frac{B^2}{4A}\right)\right]du \\
&= e^{(C-\frac{B^2}{4A})} \int_{(y>P+\frac{B}{2A})} e^{Ay^2} dy \\
&= e^{(C-\frac{B^2}{4A})} \sqrt{\frac{\pi}{-A}} \left[1 - \Phi\left(\sqrt{-2A}(P + \frac{B}{2A})\right)\right].
\end{aligned}$$

Substituting the above integrals and we get,

$$c = \frac{e^{-rT}}{\sqrt{2\pi\phi_t^{-1}}} \left[e^{(C-\frac{B^2}{4A})} \sqrt{\frac{\pi}{-A}} \left[1 - \Phi\left(\sqrt{-2A}(P + \frac{B}{2A})\right)\right] - K\sqrt{\frac{\pi}{-A}} \left[1 - \Phi(\sqrt{-2A}P)\right]\right].$$

This leads to the closed form call price:

$$c = \mathrm{C}_0 + Ke^{-rT}\Phi(\tilde{d}_1) - b\Phi(\tilde{d}_2).$$

where

$$\begin{aligned}
\mathrm{C}_0 &= e^{-rT}[\exp(L + e^{-aT}(X_0 - L) + \frac{\sigma^2}{4a}(1 - e^{-2aT})) - K], \\
\tilde{d}_1 &= \frac{1}{\sqrt{\phi_T^{-1}}}(e^{aT}(\log K - L) - (X_0 - L)), \\
\tilde{d}_2 &= \tilde{d}_1 - e^{-aT}\sqrt{\phi_T^{-1}}, \\
\phi_T^{-1} &= \frac{\sigma^2}{2a}(e^{2aT} - 1), \\
b &= \exp[-rT + e^{-aT}(X_0 - L) + L + \frac{e^{-2aT}}{2}\phi_T^{-1}].
\end{aligned}$$

□

Now we apply the above results in pricing European call options under mean reversion model. We also compare the results with the well-known Black-Scholes formula. Assume we have risk-free interest rate $r = 0.04$, volatility $\sigma = 0.4$, the expiration time of options is T=0.5, the speed of mean reversion is a=0.1.

We calculate the European call prices under the mean reversion model using the formula provided in this section and we compute the European call prices under GBM(Black-Scholes) model using Black-Scholes formula. We compare the differences between those two methods when the initial stock price and strike price are $S = 10, K = 11$, $S = 50, K = 53$ and $S = 100, K = 120$ respectively.

L	*Call price* (MR)	*Call price* (GBM)	*Relative error*
2.1	0.828167	0.8184	0.011934
2.4	0.897803	0.8184	0.097022

Table 2.1: Call prices under the MR and Black-scholes models when $S = 10, K = 11$

L	*Call price* (MR)	*Call price* (GBM)	*Relative error*
3.7	4.86736	4.8096	0.012009
3.93	5.14927	4.8096	0.070623

Table 2.2: Call prices under the MR and Black-scholes models when $S = 50, K = 53$

L	*Call price* (MR)	*Call price* (GBM)	*Relative error*
4.4	5.38629	5.3687	0.003275
4.7	5.90127	5.3687	0.099199

Table 2.3: Call prices under the MR and Black-scholes models when $S = 100, K = 120$

We can see that different equilibrium levels lead to different call prices under the mean reversion model. And the higher the level is, the higher the call price is.

Recall that L is the equilibrium level of stock log-price. This can be explained because higher stock equilibrium levels lead to greater benefit for stock holders and therefore more expensive call prices. In all three cases both methods give similar results. The error in last columns of the tables have two sources: model errors and errors from determining the stock mode.

2.4 A Regime-Switching Model with Mean Reversion

In this section we construct a framework which allows regime switching. Our goal here is to develop a regime-switching model in which the log stock price switches between two models: geometric Brownian motion and mean reversion. In addition, the switching in this model is determined by a two-state Markov chain. In practice it is easier to characterize this two-state framework as well as estimate the parameters in this model.

In particular we consider the regime-switching model,

$$dX(t) = \tilde{A}(X(t), \alpha(t))dt + \tilde{\sigma}(\alpha(t))dW(t), \tag{2.4.5}$$

where

$$\begin{aligned}
\tilde{A}(X(t), 1) &= \mu - \frac{\sigma_1^2}{2}, \\
\tilde{A}(X(t), 2) &= a(L - X(t)), \\
\tilde{\sigma}(1) &= \sigma_1, \\
\tilde{\sigma}(2) &= \sigma_2.
\end{aligned}$$

Suppose that $\alpha(t)$ is the system mode process. Let $Q = (q_{ij})$ denote the generator of $\alpha(t)$. When $\alpha(t) = 1$ the system follows a geometric Browian motion, i.e. the process $X(t)$ is governed by

$$dX(t) = (\mu - \frac{\sigma_1^2}{2})dt + \sigma_1 dW(t),$$

where μ represents the return rate and σ_1 is the corresponding volatility.

When $\alpha(t) = 2$ the log price is subject to mean reversion, i.e.,

$$dX(t) = a(L - X(t))dt + \sigma_2 dW(t),$$

where $a > 0$ is the rate of reversion, L is the equilibrium level and σ_2 is the volatility.

By Ito's formula, we can have a different version of above model in terms of the stock price $S(t)$ as follows.

$$dS(t) = A(S(t), \alpha(t))dt + \sigma(S(t), \alpha(t))dW(t), \tag{2.4.6}$$

where

$$\begin{aligned}
&A(S(t), 1) = \mu S(t), \\
&A(S(t), 2) = [a(L - \log S(t)) + \frac{1}{2}\sigma_2^2]S(t), \\
&\sigma(1) = \sigma_1 S(t), \\
&\sigma(2) = \sigma_2 S(t).
\end{aligned}$$

Valuation of European options and other applications based on above models are provided in the following chapters.

Chapter 3

Valuation of European Options under the Regime-Switching Model

3.1 Introduction

This chapter is concerned with valuation of European call options under the regime-switching model. The same problem for European put options can be derived similarly. The regime switching model has the clear advantage of capturing random change of the market environment. On the other hand, it is difficult to obtain a closed-form solution to the pricing of European options. Due to the associated non-differentiable boundary value problem, there is no guarantee that the associated PDEs that govern the stock prices in the regime-switching model have smooth solutions in addition to the non-uniqueness of the solution. These problems in practice can lead to deriving incorrect option prices and errors in hedging applications. To overcome these difficulties we devote our effort to two approaches which are convenient for treating possible non differentiable solutions. One is to use a system of PDEs with smoothed boundary conditions. The other one is the viscosity solution approach with which European option price can be characterized as the unique viscosity solution of a system of linear partial differential equations with variable coefficients.

3.2 The PDE Method

In order to study the possibility of existence and uniqueness of a solution to the European call option pricing problem, we define a system of PDE's with smooth boundary condition where the solution is slightly different from the solution associated with the original pricing problem.

Recall the regime-switching model,

$$dS(t) = A(S(t), \alpha(t))dt + \sigma(S(t), \alpha(t))dW(t)$$

where

$$A(S(t), 1) = \mu S(t),$$
$$A(S(t), 2) = [a(L - \log S(t)) + \frac{1}{2}\sigma_2^2]S(t),$$
$$\sigma(1) = \sigma_1 S(t),$$
$$\sigma(2) = \sigma_2 S(t).$$

Applying Ito's formula to $c(s, x, i)$, we have,

$$dc(s,x,i) = [\frac{\partial c}{\partial s}(s,x,i) + A(x,i)\frac{\partial c}{\partial x}(s,x,i) + \frac{1}{2}\sigma(i)^2\frac{\partial^2 c}{\partial x^2}(s,x,i)]dt$$
$$+\sigma(i)\frac{\partial c}{\partial x}(s,x,i)dW(t) + Qc(s,x,\cdot)(i) + d(\text{martingale})$$

where $Qc(s,x,\cdot)(i) = \sum_{j\neq i} q_{ij}(c(s,x,j) - c(s,x,i))$.

Then $c(s, x, i)$ satisfies the following PDE's,

$$\frac{\partial c}{\partial s}(s,x,i) + A(x,i)\frac{\partial c}{\partial x}(s,x,i) + \frac{1}{2}\sigma(i)^2\frac{\partial^2 c}{\partial x^2}(s,x,i) - rc(s,x,i) + Qc(s,x,\cdot)(i) = 0, \quad (3.2.1)$$

for $i = 1, 2$, with the boundary condition

$$c(T, x, i) = h(x) = (x - K)^+. \quad (3.2.2)$$

Note that the boundary function $h(x)$ is not differentiable at $x = K$. Smooth boundary conditions are typically required in existence problems. One way to overcome these difficulties is to construct an associated smooth boundary value problem via modification. Define $k_\delta(x)$ and $h_\delta(x)$ as follows:

$$k_\delta(x) = \begin{cases} \dfrac{a}{\delta}\exp(\dfrac{\delta^2}{x^2-\delta^2}), & \text{if } |x| < \delta, \\ 0, & \text{if } |x| \geq \delta, \end{cases}$$

$$h_\delta(x) = \int_{-\infty}^{\infty} h(y)k_\delta(x-y)dy,$$

where a is a constant such that $\int_{-\infty}^{\infty} k_\delta(x)dx = 1$. It can be verified that $h_\delta(x) = 0$ when $x \leq K - \delta$, $h_\delta(x) = x - K$ when $x > K - \delta$ and otherwise $|h_\delta(x) - h(x)| \leq \delta$. Moreover the nth derivative of $h_\delta(x)$ is bounded for $n = 1, 2, 3, 4$.

Define

$$c_\delta(s, x, i) = E[e^{-r(T-s)}h_\delta(S(T))|S(s) = x, \alpha(s) = i].$$

Note that

$$|c_\delta(s,x,i) - c(s,x,i)| \leq e^{-r(T-s)}E[|h_\delta(S(T)) - h(S(T))||S(s) = x, \alpha(s) = i]$$

$$\leq e^{-r(T-s)}\delta \leq \delta.$$

This implies that

$$\lim_{\delta\to 0} c_\delta(s,x,i) = c(s,x,i).$$

To study $c_\delta(s,x,i)$, use log-price and then define

$$\phi_\delta(s,y,i) = E[e^{-r(T-s)}h_\delta(e^{X(t)})|X(s) = y, \alpha(s) = i]. \tag{3.2.3}$$

Lemma 3.2.1. *For any $t \in [s,T]$, we have*

$$c(s,x,i) = E_{s,x,i}[e^{-r(t-s)}c(t, S(t), \alpha(t))],$$

and

$$\phi_\delta(s,y,i) = E_{s,y,i}[e^{-r(t-s)}\phi_\delta(t, X(t), \alpha(t))].$$

Proof.

Given any $t \in [s,T]$, conditioning on $(S(t),\alpha(t))$, note that

$$\begin{aligned} c(s,x,i) &= E[e^{-r(T-s)}h(S(T))|S(s)=x,\alpha(s)=i] \\ &= E[e^{-r(T-s)}E\left(h(S(T))|S(t),\alpha(t)\right)|S(s)=x,\alpha(s)=i] \\ &= E[e^{-r(t-s)}E\left(e^{-r(T-t)}h(S(T))|S(t),\alpha(t)\right)|S(s)=x,\alpha(s)=i] \\ &= E[e^{-r(t-s)}c(t,S(t),\alpha(t))|S(s)=x,\alpha(s)=i] \\ &= E_{s,x,i}[e^{-r(t-s)}c(t,S(t),\alpha(t))]. \end{aligned}$$

The second result follows in a similar way. □

Lemma 3.2.2. *Both $\phi_\delta(s,y,1)$ and $\phi_\delta(s,y,2)$ are in $C^{1,2}([s,T],R)$, and the following holds for some constant $C>0$,*

$$\phi_\delta(s,y,i) + \left|\frac{\partial}{\partial y}\phi_\delta(s,y,i)\right| + \left|\frac{\partial^2}{\partial y^2}\phi_\delta(s,y,i)\right| \le Ce^{2y}.$$

Proof.

Given the equation (2.4.5), define

$$G(s,t) = \int_s^t \tilde{A}(X(u),\alpha(u)))du + \tilde{\sigma}(\alpha(u))dW(u),$$

and rewrite $\phi_\delta(s,y,i)$ as

$$\phi_\delta(s,y,i) = E[e^{-r(T-s)}h_\delta(e^{y+G(s,t)})|X(s)=y,\alpha(s)=i] = E_{s,i}[e^{-r(T-s)}h_\delta(e^{y+G(s,t)}))].$$

First we compute the first and second derivatives of $\phi_\delta(s,y,i)$ with respect to y. The average rate of change of $h_\delta(e^{y+G(s,t)})$ is

$$\begin{aligned} &\frac{h_\delta(e^{y+\Delta y+G(s,t)}) - h_\delta(e^{y+G(s,t)})}{\Delta y} \\ &= \frac{h_\delta(e^{y+G(s,t)}) + (\Delta y)h'_\delta(e^{y+G(s,t)})e^{y+G(s,t)} + \frac{(\Delta y)^2}{2}[h'_\delta(e^{\xi+G(s,t)})e^{\xi+G(s,t)}]' - h_\delta(e^{y+G(s,t)})}{\Delta y} \end{aligned}$$

$$= h'_\delta(e^{y+G(s,t)})e^{y+G(s,t)} + \frac{\Delta y}{2}[h'_\delta(e^{\xi+G(s,t)})e^{\xi+G(s,t)} + h''_\delta(e^{\xi+G(s,t)})e^{2\xi+2G(s,t)}],$$

for some $\xi \in (y, y+\Delta y)$.

Next, we prove that

$$E_{s,i}[h'_\delta(e^{\xi+G(s,t)})e^{\xi+G(s,t)} + h''_\delta(e^{\xi+G(s,t)})e^{2\xi+2G(s,t)}]$$

is bounded. Since dh_δ/dx and d^2h_δ/dx^2 are bounded, e^ξ and $e^{2\xi}$ are bounded too for ξ between y and $y+\Delta y$, so it suffices to show that $E_{s,i}e^{G(s,t)}$ is bounded.

$$\begin{aligned}
E_{s,i}e^{G(s,t)} &= E_{s,i}\exp\left(\int_s^t \tilde{A}(X(u),\alpha(u))du + \int_s^t \tilde{\sigma}(\alpha(u))du\right) \\
&= E_{s,i}\exp\left(\int_s^t \tilde{A}(X(u),\alpha(u))du\right)\exp\left(\int_s^t \tilde{\sigma}(\alpha(u))du\right) \\
&\le \sqrt{E_{s,i}\exp(2\int_s^t \tilde{A}(X(u),\alpha(u))du)}\sqrt{E_{s,i}\exp(2\int_s^t \tilde{\sigma}(\alpha(u))du)} \\
&\le \sqrt{E_{s,i}\exp(2\int_s^t \tilde{A}(X(u),\alpha(u))du)}\sqrt{E_{s,i}\exp(2(\sigma_1+\sigma_2)(W(t)-W(s)))} \\
&= \sqrt{E_{s,i}\exp(2\int_s^t \tilde{A}(X(u),\alpha(u))du)}\left(\exp((\sigma_1+\sigma_2)^2(t-s))\right).
\end{aligned}$$

The second factor in the above product is obtained using the Gaussian property of $W(t) - W(s)$ which is bounded due to finite maturity. Therefore we only need to verify the boundedness of $E_{s,i}[\exp(2\int_s^t \tilde{A}(X(u),\alpha(u))du)]$. It is easy to see that

$$\begin{aligned}
\tilde{A}(X(t),\alpha(t)) &\le \max(\tilde{A}(X(t),1), \tilde{A}(X(t),2)) \\
&= \max(\mu - \frac{\sigma_1^2}{2}, a(L - X(t)) \\
&\le C_0 + C_1|X(t)|.
\end{aligned}$$

It follows that

$$\tilde{A}(X(t), \alpha(t)) \leq C_0 + C_1|X(t)|, \tag{3.2.4}$$

for some positive constants C_0, C_1. We have

$$E_{s,i}[\exp(2\int_s^t \tilde{A}(X(u), \alpha(u))du)] \leq E_{s,i}[\exp(2C_0(t-s) + 2C_1 \int_s^t |X(u)|du)].$$

By Ito's formula, we have

$$de^{X(t)} = e^{X(t)}(\tilde{A}(X(t), \alpha(t))dt + \tilde{\sigma}(\alpha(t))dW(t)) + \frac{\tilde{\sigma}(\alpha(t))^2}{2} e^{X(t)}dt.$$

Therefore,

$$Ee^{X(t)} = e^{X(s)} + E\int_s^t e^{X(u)}\left(\tilde{A}(X(u), \alpha(u)) + \frac{\tilde{\sigma}(\alpha(u))^2}{2}\right)du.$$

In the case of the mean reversion,

$$\begin{aligned} e^x\left(\tilde{A}(x, 2) + \frac{\tilde{\sigma}(2)^2}{2}\right) &= e^x(a(L-x) + \frac{\sigma_2^2}{2}) \\ &< a\exp(L + \frac{\sigma_2^2}{2a} - 1). \end{aligned}$$

In the case of the GBM,

$$e^x\left(\tilde{A}(x, 1) + \frac{\tilde{\sigma}(1)^2}{2}\right) = e^x(\mu + \frac{\sigma_1^2}{2}).$$

Let $C_2 = e^{X(s)} + a\exp(L + \sigma_2^2/(2a) - 1)(t-s)$ and $C_3 = \mu + \sigma_1^2/2$.

Then, we have

$$\begin{aligned} Ee^{X(t)} &\leq e^{X(s)} + E\int_s^t (a\exp(L + \frac{\sigma_2^2}{2a} - 1) + C_3 e^{X(u)}du \\ &= C_2 + C_3\int_s^t (Ee^{X(u)})du. \end{aligned}$$

By Gronwall's inequality, it follows that

$$Ee^{X(t)} \leq C_2 e^{C_3(t-s)}.$$

Since e^x is convex, we have

$$E_{s,i}[\exp(\int_s^t |X_u|du)] \leq \int_s^t E_{s,i}\exp(|X(u)|)du.$$

It follows that

$$E_{s,i}[\exp(2\int_s^t \tilde{A}(X(u),\alpha(u))du)]$$

is bounded and so is $E_{s,i}e^{G(s,t)}$. Then we have as $\Delta y \to 0$,

$$E_{s,i}[\frac{\Delta y}{2}[h_\delta'(e^{\xi+G(s,t)})e^{\xi+G(s,t)} + h_\delta''(e^{\xi+G(s,t)})e^{2\xi+2G(s,t)}] \to 0.$$

Therefore,

$$\begin{aligned}\frac{\partial \phi_\delta}{\partial y} &= \lim_{\Delta y \to 0} \frac{\phi_\delta(s, y+\Delta y, i) - \phi_\delta(s,y,i)}{\Delta y} \\ &= \lim_{\Delta y \to 0} \frac{E_{s,i}[e^{-r(T-s)}\left(h_\delta(e^{y+\Delta y+G(s,t)}) - h_\delta(e^{y+G(s,t)})\right)]}{\Delta y} \\ &= E_{s,i}[e^{-r(T-s)}h_\delta'(e^{y+G(s,t)})e^{y+G(s,t)}].\end{aligned}$$

The existence of $\partial^2\phi_\delta/\partial y^2$ can be verified in the same way.

$$\frac{\partial^2 \phi_\delta}{\partial y^2} = E_{s,i}[e^{-r(T-s)}\left(h_\delta'(e^{y+G(s,t)})e^{y+G(s,t)} + h_\delta''(e^{y+G(s,t)})e^{2y+2G(s,t)}\right)].$$

For δ small enough, note that

$$|h_\delta(x)| \leq |h(x)| + \delta \leq |x| + \delta \leq |x| + 1.$$

Then using the boundedness of $E_{s,i}e^{G(s,t)}$,

$$\begin{aligned}\phi_\delta(s,y,i) &= E_{s,i}[e^{-r(T-s)}h_\delta(e^{y+G(s,t)})] \\ &\leq e^{-r(T-s)}E_{s,i}[|h_\delta(e^{y+G(s,t)})|]\end{aligned}$$

$$\leq e^y e^{-r(T-s)} E_{s,i}[|e^{G(s,t)}| + 1]$$

$$\leq C_4 e^y,$$

for some constant $C_4 > 0$.

Similarly it can be verified that $|\partial\phi_\delta/\partial y| \leq C_5 e^y$ and $|\partial^2\phi_\delta/\partial y^2| \leq C_6 e^{2y}$ for some positive constants C_5 and C_6. This leads to, for some constant $C > 0$,

$$\phi_\delta(s,y,i) + \left|\frac{\partial}{\partial y}\phi_\delta(s,y,i)\right| + \left|\frac{\partial^2}{\partial y^2}\phi_\delta(s,y,i)\right| \leq Ce^{2y}.$$

In the following part we show the existence of $\partial\phi_\delta/\partial s$. By Lemma (3.2.1)

$$\phi_\delta(s,y,i) = E_{s,y,i}[e^{-r\Delta s}\phi_\delta(s+\Delta s, X(s+\Delta s), \alpha(s+\Delta s))]. \tag{3.2.5}$$

Similarly we have

$$\frac{\partial\phi_\delta}{\partial y}(s,y,i) = E_{s,y,i}[e^{-r(t-s)}\frac{\partial\phi_\delta}{\partial y}(t, X(t), \alpha(t))], \tag{3.2.6}$$

and

$$\frac{\partial^2\phi_\delta}{\partial y^2}(s,y,i) = E_{s,y,i}[e^{-r(t-s)}\frac{\partial^2\phi_\delta}{\partial y^2}(t, X(t), \alpha(t))]. \tag{3.2.7}$$

By the mean value theorem, for some $\xi \in (y, X(s+\Delta s))$, we have

$$\phi_\delta(s+\Delta s, X(s+\Delta s), \alpha(s+\Delta s))$$

$$= \phi_\delta(s+\Delta s, y, \alpha(s+\Delta s)) + (X(s+\Delta s) - y)\frac{\partial\phi_\delta}{\partial y}(s+\Delta s, \xi, \alpha(s+\Delta s)).$$

Therefore,

$$E_{s,y,i}[\phi_\delta(s+\Delta s, X(s+\Delta s), \alpha(s+\Delta s))]$$

$$= E_{s,y,i}[\phi_\delta(s+\Delta s, y, \alpha(s+\Delta s))] + O(\sqrt{\Delta s})$$

$$= \sum_{j=1}^{2}\phi_\delta(s+\Delta s, y, j)P(\alpha(s+\Delta s) = j|\alpha(s) = i) + O(\sqrt{\Delta s})$$

$$= \sum_{j=1}^{2} \phi_\delta(s+\Delta s, y, j)(\delta_{ij} + O(\Delta s)) + O(\sqrt{\Delta s})$$
$$= \phi_\delta(s+\Delta s, y, i) + O(\sqrt{\Delta s}).$$

From this and (3.2.5), we have, as $\Delta s \to 0$,

$$\phi_\delta(s+\Delta s, y, i) - \phi_\delta(s, y, i) = (1 - e^{-r\Delta s})\phi_\delta(s+\Delta s, y, i)O(\sqrt{\Delta s}) \to 0.$$

Using (3.2.6) and (3.2.7), we have similar results as $\Delta s \to 0$,

$$\frac{\partial \phi_\delta}{\partial y}(s+\Delta s, y, i) - \frac{\partial \phi_\delta}{\partial y}(s, y, i) \to 0,$$

and

$$\frac{\partial^2 \phi_\delta}{\partial y^2}(s+\Delta s, y, i) - \frac{\partial^2 \phi_\delta}{\partial y^2}(s, y, i) \to 0.$$

Then according to Dynkin's formula, for fixed $s+\Delta s$, we have

$$\begin{aligned}
&E_{s,y,i}[\phi_\delta(s+\Delta s, X(s+\Delta s), \alpha(s+\Delta s))] = \phi_\delta(s+\Delta s, y, i) \\
&+ E_{s,y,i}\int_s^{s+\Delta s} [\tilde{A}(X(u), \alpha(u))\frac{\partial \phi_\delta}{\partial y}(s+\Delta s, X(u), \alpha(u)) \\
&+ \frac{1}{2}\tilde{\sigma}(\alpha(u))^2 \frac{\partial^2 \phi_\delta}{\partial y^2}(s+\Delta s, X(u), \alpha(u)) + Q\phi_\delta(s+\Delta s, X(u), \cdot)(\alpha(u))]du.
\end{aligned}$$

Define

$$\begin{aligned}
I(s, s+\Delta s) = E_{s,y,i}&\int_s^{s+\Delta s} [\tilde{A}(X(u), \alpha(u))\frac{\partial \phi_\delta}{\partial y}(s+\Delta s, X(u), \alpha(u)) \\
&+ \frac{1}{2}\tilde{\sigma}(\alpha(u))^2 \frac{\partial^2 \phi_\delta}{\partial y^2}(s+\Delta s, X(u), \alpha(u)) + Q\phi_\delta(s+\Delta s, X(u), \cdot)(\alpha(u))]du.
\end{aligned}$$

Note that

$$\phi_\delta(s+\Delta s, y, i) = E_{s,y,i}[\phi_\delta(s+\Delta s, X(s+\Delta s), \alpha(s+\Delta s))] - I(s, s+\Delta s).$$

$$\phi_\delta(s,y,i) = E_{s,y,i}[e^{-r\Delta s}\phi_\delta(s+\Delta s, X(s+\Delta s), \alpha(s+\Delta s))].$$

It follows that

$$\frac{1}{\Delta s}[\phi_\delta(s+\Delta s,y,i)-\phi_\delta(s,y,i)] = \left(\frac{e^{r\Delta s}-1}{\Delta s}\right)\phi_\delta(s,y,i) - \frac{I(s,s+\Delta s)}{\Delta s}.$$

Therefore,

$$\frac{\partial\phi_\delta}{\partial s}(s,y,i) = \lim_{\Delta s\to 0}\frac{1}{\Delta s}[\phi_\delta(s+\Delta s,y,i)-\phi_\delta(s,y,i)]$$

exists because $I(s,s+\Delta s)/\Delta s$ has a limit as $\Delta s \to 0$. □

Theorem 3.2.3. *$c_\delta(s,x,i)$ is in $C^{1,2}([s,T],R)$ for some small positive δ, and it is the unique solution to (3.2.1) with boundary condition $c_\delta(T,x,i) = h_\delta(x)$.*

Proof.

We have $c_\delta(s,x,i) = \phi_\delta(s,y,i)$ and Lemma (3.2.2). Thus, it suffices to show that $\phi_\delta(s,y,i)$ is the unique solution to

$$\frac{\partial\phi_\delta}{\partial s}(s,y,i)+\tilde{A}(y,i)\frac{\partial\phi_\delta}{\partial y}(s,y,i)+\frac{1}{2}\tilde{\sigma}(i)^2\frac{\partial^2\phi_\delta}{\partial y^2}(s,y,i)-r\phi_\delta(s,y,i)+Q\phi_\delta(s,y,\cdot)(i) = 0, \quad (3.2.8)$$

with boundary

$$\phi_\delta(T,y,i) = h_\delta(e^y),$$

$$i = 1,2.$$

By lemma (3.2.1), for $t_1 < t_2$ in $[s,T]$, we have

$$\phi_\delta(t_1,y,i) = E_{t_1,y,i}[e^{-r(t_2-t_1)}\phi_\delta(t_2,X(t_2),\alpha(t_2))].$$

Then

$$e^{-rt_1}\phi_\delta(t_1,y,i) = E_{t_1,y,i}[e^{-rt_2}\phi_\delta(t_2,X(t_2),\alpha(t_2))].$$

Moreover, apply Dynkin's formula to $e^{-rt}\phi_\delta(t, X(t), \alpha(t))$, we have

$$E_{t_1,y,i}\int_{t_1}^{t_2}[\frac{\partial \phi_\delta}{\partial s}(u, X(u), \alpha(u)) + \tilde{A}(X(u), \alpha(u))\frac{\partial \phi_\delta}{\partial y}(u, X(u), \alpha(u))$$

$$+\frac{1}{2}\tilde{\sigma}(i)^2\frac{\partial^2 \phi_\delta}{\partial y^2}(u, X(u), \alpha(u)) - r\phi_\delta(u, X(u), \alpha(u)) + Q\phi_\delta(u, X(u), \cdot)(i)]du = 0.$$

For $u \in [t_1, t_2]$, dividing both sides of above equation by $t_2 - t_1$ and sending $t_2 \to t_1$, we obtain (3.2.8).

Assume $\phi_\delta(s, y, i)$ is a solution to the above system of PDEs (3.2.8). Note that

$$d(e^{-rt}\phi_\delta(t, X(t), \alpha(t))) = e^{-rt}\left[\frac{\partial \phi_\delta}{\partial s}(t, X(t), \alpha(t)) + \tilde{A}(X(t), \alpha(t))\frac{\partial \phi_\delta}{\partial y}(t, X(t), \alpha(t))\right.$$

$$\left.+\frac{1}{2}\tilde{\sigma}(\alpha(t))^2\frac{\partial^2 \phi_\delta}{\partial y^2}(t, X(t), \alpha(t)) - r\phi_\delta(t, X(t), \alpha(t)) + Q\phi_\delta(t, X(t), \cdot)(\alpha(t))\right]dt$$

$$+\tilde{\sigma}(\alpha(t))\frac{\partial \phi_\delta}{\partial y}(t, X(t), \alpha(t))dW(t) + d(\text{martingale}).$$

Integrating and taking conditional expectation of the above equation, we have

$$\phi_\delta(s, y, i) = E_{s,y,i}[e^{-r(T-s)}h_\delta(e^{X(T)})].$$

Therefore, $\phi_\delta(s, y, i)$ is the unique solution. $\square$

3.3 Viscosity Solution

The idea of viscosity solution was introduced two decades ago by Crandall and Lions [11]. The framework of viscosity solution is extremely effective to treat uniqueness of PDE in the absence of classical solutions. In this section, we show that the call option price is the only viscosity solution to the associated PDE.

Recall that

$$\begin{aligned}\phi(s, y, i) &= E[e^{-r(T-s)}h(e^{X(T)})|X(s) = y, \alpha(s) = i]\\ &= E^{s,y,i}[e^{-r(T-s)}h(e^{X(T)})].\end{aligned}$$

By Ito's formula, $\phi(s, y, i)$ satisfies the following PDE's

$$\frac{\partial \phi}{\partial s}(s,y,i) + \tilde{A}(y,i)\frac{\partial \phi}{\partial y}(s,y,i) + \frac{1}{2}\tilde{\sigma}(i)^2\frac{\partial^2 \phi}{\partial y^2}(s,y,i) - r\phi(s,y,i) + Q\phi(s,y,\cdot)(i) = 0, \quad (3.3.9)$$

$$i = 1, 2,$$

with the boundary conditions

$$\phi(T, y, i) = h(e^y) = (e^y - K)^+. \quad (3.3.10)$$

To solve the non-differentiable boundary problem developed in Section 3.2, we use the viscosity solution, a weak solution to a system of linear partial differential equations. We first study the existence of solutions of (3.3.9) and (3.3.10). Then the uniqueness of the viscosity solution will be addressed.

Definition 3.3.1 *Assume that $f(s, y, i)$ is continuous in (s, y), and there exists positive constant K and integer n, such that $|f| \leq K(1 + |y|^n)$, for $i = 1, 2$.*

1. *If $\frac{\partial \varphi}{\partial t}(s_0, y_0) + \tilde{A}(y_0, i)\frac{\partial \varphi}{\partial y}(s_0, y_0) + \frac{1}{2}\tilde{\sigma}(i)^2\frac{\partial^2 \varphi}{\partial y^2}(s_0, y_0) - rf(s_0, y_0, i)$ $+Qf(s_0, y_0, \cdot)(i) \geq 0$ for any $\varphi(s, y) \in C^2$ such that $f(s, y, i) - \varphi(s, y)$ has a local maximum at any such (s_0, y_0) for $i = 1, 2$, then $f(s, y, i)$ is a viscosity subsolution.*
2. *If $\frac{\partial \varphi}{\partial t}(s_0, y_0) + \tilde{A}(y_0, i)\frac{\partial \varphi}{\partial y}(s_0, y_0) + \frac{1}{2}\tilde{\sigma}(i)^2\frac{\partial^2 \varphi}{\partial y^2}(s_0, y_0) - rf(s_0, y_0, i)$ $+Qf(s_0, y_0, \cdot)(i) \leq 0$ for any $\varphi(s, y) \in C^2$ such that $f(s, y, i) - \varphi(s, y)$ has a local minimum at any such (s_0, y_0) for $i = 1, 2$, then $f(s, y, i)$ is a viscosity supersolution.*

$f(s, y, i)$ is called a viscosity solution if it is the viscosity subsolution and supersolution simultaneously.

Lemma 3.3.2 $\phi(s, y, \alpha(s)) = E^{s,y,\alpha(s)}[e^{-r(t-s)}\phi(t, X(t), \alpha(t))]$, *for any $t \in [s, T]$.*

The above result follows from lemma (3.2.1).

Lemma 3.3.3 *$\phi(s,y,i)$ is continuous in (s,y) for $i=1,2$. Moreover, for some positive constant C, we have $|\phi(s,y,i)|\le C(1+|y|)$.*

Proof. Since

$$\phi(s,y,i)=E^{s,y,i}[e^{-r(T-s)}(e^{X(T)}-K)^+].$$

Consider

$$|e^{-r(T-s_1)}(e^{X_1}-K)^+-e^{-r(T-s_2)}(e^{X_2}-K)^+|.$$

Applying the mean value theorem to $e^{-r(T-s_2)}$, since $|e^{-r(T-s_i)}|\le 1$ for $s_i\in[s,T]$, it follows that

$$\begin{aligned}&|e^{-r(T-s_1)}(e^{X_1}-K)^+-e^{-r(T-s_2)}(e^{X_2}-K)^+|\\ &\le |e^{-r(T-s_1)}(e^{X_1}-K)^+-e^{-r(T-s_1)}(e^{X_2}-K)^+|\\ &+|e^{-r(T-s_1)}(e^{X_2}-K)^+-e^{-r(T-s_2)}(e^{X_2}-K)^+|\\ &\le |e^{X_1}-e^{X_2}|+C_1|s_1-s_2||e^{X_2}-K|,\end{aligned}$$

for some positive constant C_1.

First let us show that $\phi(s,y,i)$ is continuous in y.

Assume $X_1(t)$ and $X_2(t)$ to be two solutions to (2.4.5). Their initial values are $X_1(s)=y_1$ and $X_2(s)=y_2$. Then we have

$$\begin{aligned}(X_1(t)-X_2(t))^2=[(y_1-y_2)+\int_s^t(\tilde{A}(X_1(\theta),\alpha(\theta))-\tilde{A}(X_2(\theta),\alpha(\theta)))d\theta\\ +\int_s^t(\tilde{\sigma}(\alpha(\theta))-\tilde{\sigma}(\alpha(\theta)))dW(\theta)]^2.\end{aligned}$$

Taking expectations of both sides of above equation conditioning on

$$X_1(s)=y_1,X_2(s)=y_2,\alpha(s)=i.$$

We have

$$
\begin{aligned}
E^{s,(y_1,y_2),i}(X_1(t)-X_2(t))^2 \leq{}& 3E^{s,(y_1,y_2),i}(y_1-y_2)^2 \\
&+3E^{s,(y_1,y_2),i}\Big(\int_s^t (\tilde{A}(X_1(\theta),\alpha(\theta))-\tilde{A}(X_2(\theta),\alpha(\theta)))d\theta\Big)^2 \\
&+3E^{s,(y_1,y_2),i}\Big(\int_s^t (\tilde{\sigma}(\alpha(\theta))-\tilde{\sigma}(\alpha(\theta)))dW(\theta)\Big)^2.
\end{aligned}
$$

By Ito's isometry,

$$
\begin{aligned}
E^{s,(y_1,y_2),i}(X_1(t)-X_2(t))^2 \leq{}& 3(y_1-y_2)^2 \\
&+3E^{s,(y_1,y_2),i}\Big(\int_s^t (\tilde{A}(X_1(\theta),\alpha(\theta))-\tilde{A}(X_2(\theta),\alpha(\theta)))d\theta\Big)^2 \\
&+3E^{s,(y_1,y_2),i}\int_s^t (\tilde{\sigma}(\alpha(\theta))-\tilde{\sigma}(\alpha(\theta)))^2 d\theta \\
\leq{}& 3(y_1-y_2)^2 \\
&+3E^{s,(y_1,y_2),i}\Big(\int_s^t (\tilde{A}(X_1(\theta),\alpha(\theta))-\tilde{A}(X_2(\theta),\alpha(\theta)))d\theta\Big)^2.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&(\tilde{A}(X_1(\theta),\alpha(\theta))-\tilde{A}(X_2(\theta),\alpha(\theta)))^2 \\
&\leq Max[((\mu-\frac{\sigma_1}{2})-(\mu-\frac{\sigma_1}{2}))^2,(a(L-X_1(\theta))-a(L-X_2(\theta)))^2] \\
&\leq Max[0,a^2(X_1(\theta)-X_2(\theta))^2] \\
&\leq a^2(X_1(\theta)-X_2(\theta))^2.
\end{aligned}
$$

It follows that

$$
E^{s,(y_1,y_2),i}(X_1(t)-X_2(t))^2 \leq 3(y_1-y_2)^2+3a^2\int_s^t E^{s,(y_1,y_2),i}(X_1(\theta)-X_2(\theta))^2 d\theta.
$$

By Gronwall's inequality, we have

$$
E^{s,(y_1,y_2),i}(X_1(T)-X_2(T))^2 \leq 3(y_1-y_2)^2\exp(3a^2(T-s)).
$$

Then,

$$
\begin{aligned}
E^{s,(y_1,y_2),i}|X_1(T) - X_2(T)| &\leq \sqrt{E^{s,(y_1,y_2),i}(X_1(T) - X_2(T))^2} \\
&\leq \sqrt{3(y_1 - y_2)^2 \exp(3a^2(T-s))} \\
&= \sqrt{3} \exp(\frac{3a^2(T-s)}{2})|y_1 - y_2|.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|\phi(s, y_1, i) - \phi(s, y_2, i)| &\leq E^{s,(y_1,y_2),i}[e^{-r(T-s)}|(e^{X_1(T)} - K)^+ - (e^{X_2(T)} - K)^+|] \\
&\leq E^{s,(y_1,y_2),i}[|e^{X_1(T)} - e^{X_2(T)}| + C_1|s-s||e^{X_2} - K|] \\
&\leq E^{s,(y_1,y_2),i}|e^{X_1(T)} - e^{X_2(T)}| \\
&\leq C_2 E^{s,(y_1,y_2),i}|X_1(T) - X_2(T)| \\
&\leq \sqrt{3} C_2 \exp(\frac{3a^2(T-s)}{2})|y_1 - y_2|.
\end{aligned}
$$

This implies continuity of $\phi(s, y, i)$ with respect to y.

Namely, given $\varepsilon > 0$, take

$$
\delta = \varepsilon \exp(-\frac{3a^2(T-s)}{2})/(\sqrt{3}C_2).
$$

Then for any $|y_1 - y_2| \leq \delta$, we have

$$
|\phi(s, y_1, i) - \phi(s, y_2, i)| \leq \varepsilon.
$$

Next we show that $\phi(s, y, i)$ is continuous with respect to s.

Define

$$
X_1(t) = X(t - (s_2 - s_1)),
$$

$$
\alpha_1(t) = \alpha(t - (s_2 - s_1)),
$$

where $s \leq s_1 \leq s_2 \leq T$ and $X(s_1) = y$.

Note that $dt = d(t-(s_2-s_1))$ and $dW(t) = dW(t-(s_2-s_1))$. Then,

$$X(t) = y + \int_{s_1}^{t} \tilde{A}(X(\theta), \alpha(\theta))d\theta + \int_{s_1}^{t} \tilde{\sigma}(\alpha(\theta))dW(\theta),$$

and

$$X_1(t) = y + \int_{s_2}^{t} \tilde{A}(X_1(\theta), \alpha_1(\theta))d\theta + \int_{s_2}^{t} \tilde{\sigma}(\alpha_1(\theta))dW(\theta).$$

It follows that

$$\begin{aligned}
X(t) - X_1(t) &= \int_{s_1}^{t} \tilde{A}(X(\theta), \alpha(\theta))d\theta + \int_{s_1}^{t} \tilde{\sigma}(\alpha(\theta))dW(\theta) - \int_{s_2}^{t} \tilde{A}(X_1(\theta), \alpha_1(\theta))d\theta \\
&\quad - \int_{s_2}^{t} \tilde{\sigma}(\alpha_1(\theta))dW(\theta) \\
&= \int_{s_1}^{t} \tilde{A}(X(\theta), \alpha(\theta))d\theta + \int_{s_1}^{t} \tilde{\sigma}(\alpha(\theta))dW(\theta) - \int_{s_1}^{t-(s_2-s_1)} \tilde{A}(X(\theta), \alpha(\theta))d\theta \\
&\quad - \int_{s_1}^{t-(s_2-s_1)} \tilde{\sigma}(\alpha(\theta))dW(\theta) \\
&= \int_{t-(s_2-s_1)}^{t} \tilde{A}(X(\theta), \alpha(\theta))d\theta + \int_{t-(s_2-s_1)}^{t} \tilde{\sigma}(\alpha(\theta))dW(\theta).
\end{aligned}$$

Taking expectation of both sides with conditioning on $X(s_1) = X_1(s_2) = y$ and $\alpha(s_1) = \alpha_1(s_2) = i$. It follows that

$$\begin{aligned}
E^{(s_1,s_2),y,i}(X(t) - X_1(t))^2 \leq 2E^{(s_1,s_2),y,i}&\left(\int_{t-(s_2-s_1)}^{t} \tilde{A}(X(\theta), \alpha(\theta))d\theta\right)^2 \\
&+2E^{(s_1,s_2),y,i}\left(\int_{t-(s_2-s_1)}^{t} \tilde{\sigma}(\alpha(\theta))dW(\theta)\right)^2.
\end{aligned}$$

Note that from (3.2.4),

$$\tilde{A}(X(t), \alpha(t)) \leq C_0 + C_1|X(t)|.$$

Then,

$$\begin{aligned}
E^{(s_1,s_2),y,i}(X(t) - X_1(t))^2 \leq 2E^{(s_1,s_2),y,i} &\int_{t-(s_2-s_1)}^{t} \tilde{A}^2(X(\theta), \alpha(\theta))d\theta(s_2 - s_1) \\
&+ 2E^{(s_1,s_2),y,i} \int_{t-(s_2-s_1)}^{t} \tilde{\sigma}^2(\alpha(\theta))d\theta
\end{aligned}$$

$$\le 2E^{(s_1,s_2),y,i}(C_0 + C_1|X(t)|)^2(s_2 - s_1)^2$$
$$+ 2(\sigma_1^2 + \sigma_2^2)(s_2 - s_1)$$
$$\le 4(C_0^2 + C_1^2 E^{(s_1,s_2),y,i}|X(t)|^2)(s_2 - s_1)^2$$
$$+ 2(\sigma_1^2 + \sigma_2^2)(s_2 - s_1).$$

Applying the Fubini theorem, we have

$$\int E^{(s_1,s_2),y,i}|X(t)|^2 dt = E^{(s_1,s_2),y,i} \int |X(t)|^2 dt < \infty.$$

This indicates that $E^{(s_1,s_2),y,i}|X(t)|^2$ is finite.

Thus for some positive constants C_2, C_3,

$$E^{(s_1,s_2),y,i}(X(t) - X_1(t))^2 \le C_2(s_2 - s_1)^2 + 2(\sigma_1^2 + \sigma_2^2)(s_2 - s_1),$$

and

$$E^{(s_1,s_2),y,i}|X(t) - X_1(t)| \le C_3\sqrt{(s_2 - s_1)^2 + (s_2 - s_1)}$$
$$\le C_3(\sqrt{(s_2 - s_1)^2} + \sqrt{s_2 - s_1})$$
$$= C_3((s_2 - s_1) + \sqrt{s_2 - s_1}).$$

$$|\phi(s_1, y, i) - \phi(s_2, y, i)| \le E^{(s_1,s_2),y,i}[|e^{X(T)} - e^{X_1(T)}| + C_1|s_1 - s_2||e^{X(T)} - K|]$$
$$\le C_4 E^{(s_1,s_2),y,i}|X(T) - X_1(T)|$$
$$+ C_1|s_1 - s_2|E^{(s_1,s_2),y,i}|e^{X(T)} - K|$$
$$\le C_5((s_2 - s_1) + \sqrt{s_2 - s_1})$$
$$+ C_1|s_1 - s_2|E^{(s_1,s_2),y,i}|e^{X(T)} - K|$$
$$\le C_5((s_2 - s_1) + \sqrt{s_2 - s_1}) + C_6(s_2 - s_1),$$

because of the boundedness of $E^{(s_1,s_2),y,i}|e^{X(T)} - K|$.

Therefore for some positive constants C_7 and C_8,

$$\begin{aligned}|\phi(s_1,y_1,i)-\phi(s,y,i)| &\le |\phi(s_1,y_1,i)-\phi(s,y_1,i)|+|\phi(s,y_1,i)-\phi(s,y,i)|\\ &\le C_7((s_1-s)+\sqrt{s_1-s})+C_8\exp(\frac{3a^2(T-s)}{2})|y_1-y|.\end{aligned}$$

As $(s_1,y_1)\to(s,y)$,

$$|\phi(s_1,y_1,i)-\phi(s,y,i)|\to 0;$$

this implies the continuity of $\phi(s,y,i)$ with respect to (s,y).

Now from above results, we have

$$|\phi(s,y_1,i)-\phi(s,y_2,i)|\le C_9|y_1-y_2|+C_{10}|s-s|E^{s,(y_1,y_2),i}|e^{X_2(T)}-K|.$$

In particular, take $y_2=-10$. Then,

$$\begin{aligned}|\phi(s,y_2,i)| &= E^{s,-10,i}|e^{-r(T-s)}(e^{X(T)}-K)^+|\\ &\le E^{s,-10,i}|(\frac{\exp(\int_s^t\tilde{A}(X(\theta),\alpha(\theta))d\theta+\int_s^t\tilde{\sigma}(\alpha(\theta))dW(\theta))}{e^{10}}-K)^+|\\ &\le \frac{E^{s,-10,i}\exp(\int_s^t\tilde{A}(X(\theta),\alpha(\theta))d\theta+\int_s^t\tilde{\sigma}(\alpha(\theta))dW(\theta))}{e^{10}}+K\\ &\le M.\end{aligned}$$

Since

$$|\phi(s,y_1,i)-\phi(s,y_2,i)|\le C_9|y_1|+10C_9,$$

then,

$$|\phi(s,y_1,i)|\le|\phi(s,y_2,i)|+C_9|y_1|+10C_9\le M+C_9|y_1|+10C_9.$$

Therefore, there exists some positive constant C, so that $|\phi(s,y,i)|\le C(1+|y|)$. □

Lemma 3.3.4 *Define*

$$\Lambda^{2,+}f(s,x,i) = \{(a,b,c) : f(t,y,i) \leq f(s,x,i) + a(t-s)b(y-x) + \frac{1}{2}(y-x)^2c + o(|y-x|^2),$$

as $(t,y) \to (s,x)\}$*, for* $f(s,x,i) : [0,T] \times R \times \{1,2\} \to R$*. And,*

$$\Lambda^{2,-}f(s,x,i) = -\Lambda^{2,+}(-f)(s,x,i).$$

Then for any $\varphi \in C^2([0,T] \times R)$ *so that* $f - \varphi$ *has a global max at* (s,x)*,* $\Lambda^{2,+}f(s,x,i)$ *consists of the set of* $\{\frac{\partial\varphi(s,x)}{\partial s}, \frac{\partial\varphi(s,x)}{\partial x}, \frac{\partial^2\varphi(s,x)}{\partial x^2}\}$*;*
Similarly, for any $\varphi \in C^2([0,T] \times R)$ *so that* $f - \varphi$ *has a global min at* (s,x)*,* $\Lambda^{2,-}f(s,x,i)$ *consists of the set of* $\{\frac{\partial\varphi(s,x)}{\partial s}, \frac{\partial\varphi(s,x)}{\partial x}, \frac{\partial^2\varphi(s,x)}{\partial x^2}\}$*.*

Remark:

1. Condition (1) in definition (3.3.1) can be replaced by

$$\forall i \in \{1,2\}, (a,b,c) \in \Lambda^{2,+}f(s,x,i),$$

$$a + \tilde{A}(y_0,i)b + \frac{1}{2}\tilde{\sigma}(i)^2c - rf(s_0,y_0,i) + Qf(s_0,y_0,\cdot)(i) \geq 0.$$

Then $f(s,x,i)$ is a viscosity subsolution.

2. Condition (2) in definition (3.3.1) can be replaced by

$$\forall i \in \{1,2\}, (a,b,c) \in \Lambda^{2,-}f(s,x,i),$$

$$a + \tilde{A}(y_0,i)b + \frac{1}{2}\tilde{\sigma}(i)^2c - rf(s_0,y_0,i) + Qf(s_0,y_0,\cdot)(i) \leq 0.$$

Then $f(s,x,i)$ is a viscosity supersolution.

Theorem 3.3.5 *The European call option price* $\phi(s,y,i)$ *is a viscosity solution to (3.3.9) with boundary condition (3.3.10).*

Proof.

Recall in Lemma (3.3.3) it is proved that $\phi(s,y,i)$ is continuous in (s,y) and $|\phi(s,y,i)| \leq C(1+|y|)$ for some constant $C > 0$. Then it suffices to show $\phi(s,y,i)$ is a viscosity subsolution and viscosity supersolution to (3.3.9) and (3.3.10).

First, we show that

$$\frac{\partial \varphi}{\partial t}(s,X(s)) + \tilde{A}(X(s),\alpha(s))\frac{\partial \varphi}{\partial y}(s,X(s)) + \frac{1}{2}\tilde{\sigma}(\alpha(s))^2\frac{\partial^2 \varphi}{\partial y^2}(s,X(s))$$

$$-r\phi(s,X(s),\alpha(s)) + Q\phi(s,X(s),\cdot)(\alpha(s)) \leq 0,$$

for any $\varphi(t,X(t)) \in C^2([s,T]\times R)$ such that $\phi(t,X(t),\alpha(t)) - \varphi(t,X(t))$ has a local minimum at $(s,X(s))$, for $i = 1,2$.

Assume that $\phi(t,X(t),\alpha(t)) - \varphi(t,X(t))$ has a local minimum at $(s,X(s))$ in a neighborhood $B(s,X(s))$.

Define

$$\bar{\psi}(t,y,i) = \begin{cases} \phi(t,y,i) & \text{if } i \neq \alpha(s), \\ \varphi(t,y) + \phi(s,X(s),\alpha(s)) - \varphi(s,X(s)) & \text{if } \ i = \alpha(s). \end{cases}$$

Let τ be the first time that $\alpha(\cdot)$ jumps away from $\alpha(s)$. And for $\theta \in [s,\tau]$, $(t,X(t))$ travels from $(s,X(s))$ and remains in $B(s,X(s))$. This implies that $\alpha(t) = \alpha(s)$ for all $t \in [s,\theta]$.

Applying Dynkin's formula, we get

$$E^{s,X(s),\alpha(s)}[e^{-r(\theta-s)}\bar{\psi}(\theta,X(\theta),\alpha(\theta))] - \bar{\psi}(s,X(s),\alpha(s))$$
$$= \bar{\psi}(s,X(s),\alpha(s)) + E^{s,X(s),\alpha(s)}\int_s^\theta e^{-r(t-s)}[\frac{\partial \bar{\psi}}{\partial t}(t,X(t),\alpha(t))$$
$$+ \tilde{A}(X(t),\alpha(t))\frac{\partial \bar{\psi}}{\partial y}(t,X(t),\alpha(t)) + \frac{1}{2}\tilde{\sigma}(\alpha(t))^2\frac{\partial^2 \bar{\psi}}{\partial y^2}(t,X(t),\alpha(t)) - r\bar{\psi}(t,X(t),\alpha(t))$$

$$+ Q\bar{\psi}(t, X(t), \cdot)(\alpha(s))]dt - \bar{\psi}(s, X(s), \alpha(s))$$
$$= E^{s,X(s),\alpha(s)} \int_s^\theta e^{-r(t-s)}[\frac{\partial \bar{\psi}}{\partial t}(t, X(t), \alpha(t)) + \tilde{A}(X(t), \alpha(t))\frac{\partial \bar{\psi}}{\partial y}(t, X(t), \alpha(t))$$
$$+ \frac{1}{2}\tilde{\sigma}(\alpha(t))^2 \frac{\partial^2 \bar{\psi}}{\partial y^2}(t, X(t), \alpha(t)) - r\bar{\psi}(t, X(t), \alpha(t)) + Q\bar{\psi}(t, X(t), \cdot)(\alpha(s))]dt.$$

Note that, for $s \le t \le \theta$,

$$\phi(t, X(t), \alpha(t)) - \varphi(t, X(t)) \ge \phi(s, X(s), \alpha(s)) - \varphi(s, X(s)).$$

Then, we have

$$\phi(t, X(t), \alpha(t)) \ge \varphi(t, X(t)) + \phi(s, X(s), \alpha(s)) - \varphi(s, X(s)). \tag{3.3.11}$$

By definition of $\bar{\psi}(t, X(t), \alpha(s))$, for any $t \in [s, \theta]$,

$$\phi(t, X(t), \alpha(t)) \ge \bar{\psi}(t, X(t), \alpha(s)),$$

and

$$\phi(s, X(s), \alpha(s)) = \bar{\psi}(s, X(s), \alpha(s)).$$

It follows that,

$$\frac{\partial \bar{\psi}}{\partial t}(t, X(t), \alpha(s)) = \frac{\partial \varphi}{\partial t}(t, X(t)),$$
$$\frac{\partial \bar{\psi}}{\partial y}(t, X(t), \alpha(s)) = \frac{\partial \varphi}{\partial y}(t, X(t)),$$
$$\frac{\partial \bar{\psi}^2}{\partial^2 y}(t, X(t), \alpha(s)) = \frac{\partial^2 \varphi}{\partial y^2}(t, X(t)).$$

Therefore, we have the following inequalities

$$E^{s,X(s),\alpha(s)}[e^{-r(\theta-s)}\phi(\theta, X(\theta), \alpha(\theta))] - \phi(s, X(s), \alpha(s))$$
$$= E^{s,X(s),\alpha(s)}[e^{-r(\theta-s)}\phi(\theta, X(\theta), \alpha(\theta))] - \bar{\psi}(s, X(s), \alpha(s))$$
$$\ge E^{s,X(s),\alpha(s)}[e^{-r(\theta-s)}\bar{\psi}(\theta, X(\theta), \alpha(\theta))] - \bar{\psi}(s, X(s), \alpha(s))$$

$$
\begin{aligned}
&= E^{s,X(s),\alpha(s)} \int_s^\theta e^{-r(t-s)} [\frac{\partial \bar{\psi}}{\partial t}(t, X(t), \alpha(t)) + \tilde{A}(X(t), \alpha(t)) \frac{\partial \bar{\psi}}{\partial y}(t, X(t), \alpha(t)) \\
&+ \frac{1}{2} \tilde{\sigma}(\alpha(t))^2 \frac{\partial^2 \bar{\psi}}{\partial y^2}(t, X(t), \alpha(t)) - r\bar{\psi}(t, X(t), \alpha(t)) + Q\bar{\psi}(t, X(t), \cdot)(\alpha(s))]dt \\
&= E^{s,X(s),\alpha(s)} \int_s^\theta e^{-r(t-s)} [\frac{\partial \varphi}{\partial t}(t, X(t)) + \tilde{A}(X(t), \alpha(t)) \frac{\partial \varphi}{\partial y}(t, X(t)) \\
&+ \frac{1}{2} \tilde{\sigma}(\alpha(t))^2 \frac{\partial^2 \varphi}{\partial y^2}(t, X(t)) - r\bar{\psi}(t, X(t), \alpha(t)) + Q\bar{\psi}(t, X(t), \cdot)(\alpha(s))]dt \\
&\geq E^{s,X(s),\alpha(s)} \int_s^\theta e^{-r(t-s)} [\frac{\partial \varphi}{\partial t}(t, X(t)) + \tilde{A}(X(t), \alpha(t)) \frac{\partial \varphi}{\partial y}(t, X(t)) \\
&+ \frac{1}{2} \tilde{\sigma}(\alpha(t))^2 \frac{\partial^2 \varphi}{\partial y^2}(t, X(t)) - r\phi(t, X(t), \alpha(t)) + Q\bar{\psi}(t, X(t), \cdot)(\alpha(s))]dt.
\end{aligned}
$$

Recall that

$$Q\phi(t, X(t), \cdot)(\alpha(s)) = \sum_{\beta \neq \alpha(s)} q_{\alpha(s),\beta} \left(\phi(t, X(t), \beta) - \phi(t, X(t), \alpha(s))\right).$$

We have

$$
\begin{aligned}
Q\bar{\psi}(t, X(t), \cdot)(\alpha(s)) = \sum_{\beta \neq \alpha(s)} & q_{\alpha(s),\beta}(\phi(t, X(t), \beta) \\
&-[\varphi(t, X(t)) + \phi(s, X(s), \alpha(s)) - \varphi(s, X(s))]).
\end{aligned}
$$

Using (3.3.11), we have

$$
\begin{aligned}
Q\bar{\psi}(t, X(t), \cdot)(\alpha(s)) &= \sum_{\beta \neq \alpha(s)} q_{\alpha(s),\beta}(\phi(t, X(t), \beta) - \phi(s, X(s), \alpha(s)) \\
&+ \varphi(s, X(s)) - \varphi(t, X(t))) \\
&\geq \sum_{\beta \neq \alpha(s)} q_{\alpha(s),\beta}(\phi(t, X(t), \beta) - \phi(s, X(s), \alpha(s)) + \varphi(s, X(s)) \\
&- \phi(t, X(t), \alpha(t)) + \phi(s, X(s), \alpha(s)) - \varphi(s, X(s))) \\
&= \sum_{\beta \neq \alpha(s)} q_{\alpha(s),\beta}(\phi(t, X(t), \beta) - \phi(t, X(t), \alpha(t))) \\
&= \sum_{\beta \neq \alpha(s)} q_{\alpha(s),\beta}(\phi(t, X(t), \beta) - \phi(t, X(t), \alpha(s))) \\
&= Q\phi(t, X(t), \cdot)(\alpha(s)).
\end{aligned}
$$

It follows that

$$\begin{aligned}
&E^{s,X(s),\alpha(s)}[e^{-r(\theta-s)}\phi(\theta,X(\theta),\alpha(\theta))]-\phi(s,X(s),\alpha(s))\\
&\geq E^{s,X(s),\alpha(s)}\int_s^\theta e^{-r(t-s)}[\frac{\partial\varphi}{\partial t}(t,X(t))+\tilde{A}(X(t),\alpha(t))\frac{\partial\varphi}{\partial y}(t,X(t))\\
&+\frac{1}{2}\tilde{\sigma}(\alpha(t))^2\frac{\partial^2\varphi}{\partial y^2}(t,X(t))-r\phi(t,X(t),\alpha(t))+Q\bar{\psi}(t,X(t),\cdot)(\alpha(s))]dt\\
&\geq E^{s,X(s),\alpha(s)}\int_s^\theta e^{-r(t-s)}[\frac{\partial\varphi}{\partial t}(t,X(t))+\tilde{A}(X(t),\alpha(t))\frac{\partial\varphi}{\partial y}(t,X(t))\\
&+\frac{1}{2}\tilde{\sigma}(\alpha(t))^2\frac{\partial^2\varphi}{\partial y^2}(t,X(t))-r\phi(t,X(t),\alpha(t))+Q\phi(t,X(t),\cdot)(\alpha(s))]dt.
\end{aligned}$$

Recall $\phi(s,X(s),\alpha(s))=E^{s,y,\alpha(s)}[e^{-r(\theta-s)}\phi(\theta,X(\theta),\alpha(\theta))]$ from Lemma (3.2.1). We get

$$\begin{aligned}
0\geq E^{s,X(s),\alpha(s)}&\int_s^\theta e^{-r(t-s)}[\frac{\partial\varphi}{\partial t}(t,X(t))+\tilde{A}(X(t),\alpha(t))\frac{\partial\varphi}{\partial y}(t,X(t))\\
+\frac{1}{2}\tilde{\sigma}(\alpha(t))^2&\frac{\partial^2\varphi}{\partial y^2}(t,X(t))-r\phi(t,X(t),\alpha(t))+Q\phi(t,X(t),\cdot)(\alpha(s))]dt.
\end{aligned}$$

Dividing both sides of above inequality by $(\theta-s)>0$, and sending $\theta\to s$, we have

$$\begin{aligned}
\frac{\partial\varphi}{\partial t}(s,X(s))+\tilde{A}(X(s),\alpha(s))\frac{\partial\varphi}{\partial y}(s,X(s))+\frac{1}{2}\tilde{\sigma}(\alpha(s))^2\frac{\partial^2\varphi}{\partial y^2}(s,X(s))\\
-r\phi(s,X(s),\alpha(s))+Q\phi(s,X(s),\cdot)(\alpha(s))\leq 0.
\end{aligned}$$

This implies that $\phi(s,X(s),\alpha(s))$ is a viscosity supersolution to (3.3.9). Similarly it can be proven that $\phi(s,X(s),\alpha(s))$ is also a viscosity subsolution to (3.3.9). And consequently it is a viscosity solution to (3.3.9) with boundary (3.3.10). □

Next, we verify the uniqueness of above viscosity solution. This is done through the Comparison Principle theorem.

Theorem 3.3.6 *(Comparison Principle) Assume* $f_1(s, y, i)$ *and* $f_2(s, y, i)$ *are continuous with respect to* (s, y)*, and there exist positive constants* K_1 *and* K_2 *so that*

$$f_n(s, y, i) \le K_n(1 + |y|), n = 1, 2.$$

If $f_1(s, y, i)$ *is a viscosity subsolution and* $f_2(s, y, i)$ *is a viscosity supersolution of (3.3.9) and (3.3.10), then*

$$f_1(s, y, i) \le f_2(s, y, i),$$

for all $(s, y, i) \in [s, T] \times R^+ \times \{1, 2\}$.

Proof. Define

$$\Psi(s, x, y, i) = f_1(s, y, i) - f_2(s, y, i) - \Phi(s, x, y),$$

where for $0 < \delta < 1$ and $0 < \varepsilon < 1$,

$$\Phi(s, x, y) = \frac{1}{\delta}(x - y)^2 + \varepsilon e^{(T-s)}(x^2 + y^2).$$

Note that

$$f_n(s, y, i) \le K_n(1 + |y|), n = 1, 2.$$

We have

$$\lim_{|x|+|y| \to \infty} \Psi(s, x, y, i) = -\infty, i = 1, 2. \tag{3.3.12}$$

Moreover, note that each $f_n(s, y, i)$ is continuous with respect to (s, y); so is $\Psi(s, x, y, i)$ with respect to (s, x, y). Thus $\Psi(s, x, y, i)$ has a global maximum. There exists $(s_\delta, x_\delta, y_\delta, \alpha')$ so that $\Psi(s_\delta, x_\delta, y_\delta, \alpha')$ is the global maximum.

It is easy to see that

$$\Psi(s_\delta, x_\delta, x_\delta, \alpha') + \Psi(s_\delta, y_\delta, y_\delta, \alpha') \le 2\Psi(s_\delta, x_\delta, y_\delta, \alpha').$$

This leads to

$$\frac{2}{\delta}(x_\delta - y_\delta)^2 \leq f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, y_\delta, \alpha') + f_2(s_\delta, x_\delta, \alpha') - f_1(s_\delta, y_\delta, \alpha').$$

Using the condition that $f_n(s, y, i) \leq K_n(1 + |y|), n = 1, 2$. There exists $K > 0$ such that

$$(x_\delta - y_\delta)^2 \leq \frac{\delta K}{2}(1 + |x_\delta| + |y_\delta|). \tag{3.3.13}$$

In the special case where $x = y = 0$,

$$\begin{aligned} |\Psi(s, 0, 0, \alpha_\delta)| &= |f_1(s, 0, \alpha') - f_2(s, 0, \alpha')| \\ &\leq K(1 + |x_\delta| + |y_\delta|), \end{aligned}$$

and

$$\Psi(s, 0, 0, \alpha') \leq \Psi(s_\delta, x_\delta, y_\delta, \alpha').$$

So for some positive K,

$$\begin{aligned} \varepsilon e^{(T-s)}(x_\delta^2 + y_\delta^2) &\leq |f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, y_\delta, \alpha')| + |\frac{1}{\delta}(x_\delta - y_\delta)^2| + |\Psi(s, 0, 0, \alpha')| \\ &\leq 3K(1 + |x_\delta| + |y_\delta|). \end{aligned}$$

For $s_\delta \in [s, T]$, there exists K_ε which depends on ε but not on δ such that,

$$|x_\delta| + |y_\delta| \leq K_\varepsilon.$$

Therefore, there is a convergent subsequence $(x_\delta), (y_\delta), (s_\delta)$ so that $\lim_{\delta \to 0} s_\delta = s_0$ and $\lim_{\delta \to 0} x_\delta = \lim_{\delta \to 0} y_\delta = x_0$.

It follows that

$$\lim_{\delta \to 0} \frac{2}{\delta}(x_\delta - y_\delta)^2 = 0.$$

From Crandall, Lions and Ishii's work ([11]), it can be verified that for each $\varepsilon' > 0$, there exists $a_{1\delta}$, $a_{2\delta}$, $C_{1\delta}$ and $C_{2\delta}$ such that

$$(a_{1\delta}, \frac{2}{\delta}(x_\delta - y_\delta) + 2\varepsilon e^{(T-s)} x_\delta, C_{1\delta}) \in \Lambda^{2,+} f_1(s_\delta, x_\delta, \alpha_\delta).$$

$$(-a_{2\delta}, -\frac{2}{\delta}(x_\delta - y_\delta) + 2\varepsilon e^{(T-s)}y_\delta, -C_{2\delta}) \in \Lambda^{2,+}(-f_2(s_\delta, y_\delta, \alpha')).$$

Equivalently we have

$$(a_{2\delta}, \frac{2}{\delta}(x_\delta - y_\delta) - 2\varepsilon e^{(T-s)}y_\delta, C_{2\delta}) \in \Lambda^{2,-}(f_2(s_\delta, y_\delta, \alpha')).$$

Since $f_1(s_\delta, x_\delta, \alpha')$ is a viscosity subsolution, it follows that

$$a_{1\delta} + \tilde{A}(x_\delta, \alpha')(\frac{2}{\delta}(x_\delta - y_\delta) + 2\varepsilon e^{(T-s)}x_\delta) + \frac{1}{2}\tilde{\sigma}(\alpha')^2 C_{1\delta} - rf_1(s_\delta, x_\delta, \alpha_\delta)$$

$$+Qf_1(s_\delta, x_\delta, \cdot)(\alpha') \geq 0. \tag{3.3.14}$$

Similarly since $f_2(s_\delta, y_\delta, \alpha')$ is a viscosity supersolution, we have

$$a_{2\delta} + \tilde{A}(y_\delta, \alpha')(\frac{2}{\delta}(x_\delta - y_\delta) + 2\varepsilon e^{(T-s)}y_\delta) + \frac{1}{2}\tilde{\sigma}(\alpha')^2 C_{2\delta} - rf_2(s_\delta, y_\delta, \alpha')$$

$$+Qf_2(s_\delta, y_\delta, \cdot)(\alpha') \leq 0. \tag{3.3.15}$$

Using these inequalities,

$$\frac{\partial \Phi}{\partial t}(s_\delta, y_\delta, \alpha') = \varepsilon e^{(T-s_\delta)}[x_\delta^2 + y_\delta^2],$$

and the maximum principle, we obtain

$$r(f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, \alpha')) \leq \frac{1}{2}\tilde{\sigma}(\alpha')^2[C_{1\delta} - C_{2\delta}] + [\tilde{A}(x_\delta, \alpha') - \tilde{A}(y_\delta, \alpha')]\frac{2}{\delta}(x_\delta - y_\delta)$$

$$+2\varepsilon e^{(T-s_\delta)}[x_\delta \tilde{A}(x_\delta, \alpha') + y_\delta \tilde{A}(y_\delta, \alpha')]$$

$$+Qf_1(s_\delta, x_\delta, \cdot)(\alpha') - Qf_2(s_\delta, y_\delta, \cdot)(\alpha') + \varepsilon e^{(T-s_\delta)}[x_\delta^2 + y_\delta^2], \tag{3.3.16}$$

and

$$-\left(\frac{1}{\varepsilon} + \|D^2_{(x,y)}\Phi(s_\delta, x_\delta, \alpha')\|\right) I \leq \begin{pmatrix} C_{1\delta} & 0 \\ 0 & -C_{2\delta} \end{pmatrix} \leq D^2_{(x,y)}\Phi(s_\delta, x_\delta, \alpha')$$

$$+\varepsilon(D^2_{(x,y)}\Phi(s_\delta, x_\delta, \alpha'))^2.$$

By definition, we have

$$D^2_{(x,y)}\Phi(s_\delta, x_\delta, \alpha') = \frac{2}{\delta}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + 2\varepsilon e^{(T-s_\delta)}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} C_{1\delta} & 0 \\ 0 & -C_{2\delta} \end{pmatrix} \le \frac{2}{\delta}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + [2\varepsilon e^{(T-s_\delta)} + 4\varepsilon'\varepsilon e^{2(T-s_\delta)}]\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$+\frac{8\varepsilon'(1+\varepsilon\delta e^{(T-s_\delta)})}{\delta^2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Sending $\varepsilon \to 0$ and setting $\varepsilon' = \delta/4$, we have

$$\begin{pmatrix} C_{1\delta} & 0 \\ 0 & -C_{2\delta} \end{pmatrix} \le \frac{4}{\delta}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

and then

$$C_{1\delta} - C_{2\delta} = (1,1)\begin{pmatrix} C_{1\delta} & 0 \\ 0 & -C_{2\delta} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \le 0.$$

As $\varepsilon \to 0$, we have

$$r(f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, \alpha')) \le a|x_\delta - y_\delta|\frac{2}{\delta}(x_\delta - y_\delta) + Qf_1(s_\delta, x_\delta, \cdot)(\alpha') - Qf_2(s_\delta, y_\delta, \cdot)(\alpha'). \tag{3.3.17}$$

That leads to

$$\lim_{\delta\to 0}\frac{2}{\delta}(x_\delta - y_\delta)^2 = 0.$$

So

$$r(f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, \alpha')) \le Qf_1(s_\delta, x_\delta, \cdot)(\alpha') - Qf_2(s_\delta, y_\delta, \cdot)(\alpha').$$

Recall that Ψ attains a maximum at $(s_\delta, x_\delta, y_\delta, \alpha')$. Then for all states and $x \in R$,

$$\Psi(s, x, x, i) \le \Psi(s_\delta, x_\delta, y_\delta, \alpha').$$

Setting $\delta \to 0$ and replacing x and s by x_δ and s_δ, respectively, we get

$$f_1(s,x,i) - f_2(s,x,i) - 2\varepsilon e^{(T-s)}x^2 \le f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, \alpha') - 2\varepsilon e^{(T-s_\delta)}x_\delta^2. \tag{3.3.18}$$

$$f_1(s_\delta, x_\delta, i) - f_2(s_\delta, x_\delta, i) \le f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, \alpha').$$

Since

$$Qf_1(s_\delta, x_\delta, \cdot)(\alpha') - Qf_2(s_\delta, y_\delta, \cdot)(\alpha') = \sum_{i \ne \alpha'} q_{\alpha',i}[f_1(s_\delta, x_\delta, i) - f_1(s_\delta, x_\delta, \alpha')$$

$$+f_2(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, i)] \le 0.$$

Then

$$f_1(s_\delta, x_\delta, \alpha') - f_2(s_\delta, x_\delta, \alpha') \le 0.$$

Letting $\varepsilon \to 0$ in (3.3.18), we obtain

$$f_1(s, x, i) \le f_2(s, x, i).$$

This completes the proof. $\square$

Recall that the definition of viscosity solution implies that it is a viscosity subsolution and a viscosity supersolution simultaneously. If $f_1(s, x, i)$ and $f_2(s, x, i)$ are both viscosity solutions to (3.3.9) with boundary (3.3.10), then according to Theorem (3.3.6) we get

$$f_1(s, x, i) \le f_2(s, x, i) \le f_1(s, x, i).$$

Thus,

$$f_1(s, x, i) = f_2(s, x, i).$$

This establishes uniqueness of the viscosity solution to (3.3.9) with boundary (3.3.10).

Chapter 4

Numerical Methods

4.1 Introduction

In this chapter we consider the successive approximation method for the valuation of European options with switching regimes as well as the stochastic approximation method for estimation of the parameters in the regime-switching model.

The idea of successive approximation is based on risk-neutral valuation theory and the fixed-point of a certain integral operator with a Gaussian kernel. Using the successive approximation method, one can calculate approximate European option prices without solving the system of differential equations given in the last chapter. To implement the results in practice, it is necessary to develop an approach to estimate the parameters which are not known a priori. In the second part of this chapter, we use the stochastic approximation method to provide an algorithm to accomplish this based on a large number of real market option prices. In practice the time needed for successfully estimating the parameter is a major concern. Usually more observations bring better estimates, whereas, computation on a large amount of data takes more time to reach the desired estimation accuracy. The rate of convergence is also discussed there.

4.2 Successive Approximation Method

Recall the regime-switching model in terms of log price $X(t)$,

$$dX(t) = \tilde{A}(X(t), \alpha(t))dt + \tilde{\sigma}(\alpha(t))dW(t), \tag{4.2.1}$$

where

$$\begin{aligned} \tilde{A}(X(t), 1) &= \mu - \frac{\sigma_1^2}{2}, \\ \tilde{A}(X(t), 2) &= a(L - X(t)), \\ \tilde{\sigma}(1) &= \sigma_1, \\ \tilde{\sigma}(2) &= \sigma_2. \end{aligned}$$

The state process $\alpha(t)$ is a Markov chain that takes values in $\{1, 2\}$ where 1 represents the geometric Brownian motion state and 2 represents the mean reversion state. $Q = (q_{ij})$ is the generator of $\alpha(t)$. In the case of at most one switching $q_{11} = -\lambda_1, q_{12} = \lambda_1, q_{21} = \lambda_2, q_{22} = -\lambda_2$, and $P(\tau > u|\alpha(s) = i) = e^{q_{ii}u}$, where τ is the stopping time defined next.

Define the stopping time process

$$\tau = \inf\,(t \geq s, \alpha(t) \neq \alpha(s)).$$

There are two possible situations. First of all, when $\tau > T$, no jump happens during the life of the option, and the stock price stays its same initial state. We can see that the Black-Scholes formula will give the price of the European options with initial state 1. On the other hand, the European call option prices can be obtained by the formula in Theorem (2.3.1) with initial state 2. The other situation is when $\tau < T$, there is at least one jump which occurs before the expiration of the option.

Define

$$\psi(s, y, i) = \frac{c(s, x, i)}{x}.$$

$y = \log(x)$ is the initial log price, s is initial time and i is the initial state. First we consider the case of no switching and the discounted payoff is given by

$$\psi^0(s, y, i) = e^{(-y)} E[e^{(-r(T-s))}(e^{X(T)} - K)^+ | X(s) = y, \alpha(u) = i, 0 \le u \le T].$$

Next, we consider the discounted payoff with the possibility of switching and the discounted payoff conditioned on the first jump time before expiration.

Theorem 4.2.1. *$\psi(s, y, i)$ is given as follows:*

$$\psi(s, y, i) = \mathcal{T}\psi(s, y, i) + e^{-\lambda_i (T-s)} \psi^0(s, y, i).$$

1. *if the initial state is geometric Brownian motion, then*

$$\mathcal{T}\psi(s, y, 1) = \int_s^T e^{-r(t-s)} \int_{-\infty}^{\infty} e^u \psi(t, y+u, 2) N\left((\mu - \frac{\sigma_1^2}{2})(t-s), \sigma_1^2 (t-s) \right) du$$
$$\lambda_1 e^{-\lambda_1 (t-s)} dt;$$

2. *if the initial state is mean reversion and subject to the condition $(\sigma_2^2/2 + aL - r) < 0$, then*

$$\mathcal{T}\psi(s, y, 2) = \int_s^T e^{-r(t-s)-y} \int_{-\infty}^{\infty} \exp(e^{-a(t-s)}(y - L + v) + L)$$
$$\psi(t, e^{-a(t-s)}(y - L + v) + L, 1) N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)} - 1) \right) dv \lambda_2 e^{-\lambda_2 (t-s)} dt.$$

Proof. Note that

$$c(s, x, i) = E[e^{(-r(T-s))}(S(T) - K)^+ | S(s) = x, \alpha(s) = i],$$

and $\psi(s, y, i)e^y = c(s, x, i)$. It follows that

$$\psi(s, y, i) = e^{-y} \left(E[e^{(-r(T-s))}(e^{X(T)} - K)^+ I_{(\tau \le T)} | X(s) = y, \alpha(s) = i] \right.$$
$$\left. + E[e^{(-r(T-s))}(e^{X(T)} - K)^+ I_{(\tau \ge T)} | X(s) = y, \alpha(s) = i] \right).$$

This equation can be written as

$$\psi(s,y,i) = e^{-y}E[e^{(-r(T-s))}(e^{X(T)}-K)^{+}I_{(\tau\le T)}|X(s)=y,\alpha(s)=i] + e^{-\lambda_i(T-s)}\psi^0(s,y,i).$$

The first term, by conditioning on $(\tau = t)$, is defined by $\mathcal{T}\psi(s,y,i)$ as follows:

$$\begin{aligned}
\mathcal{T}\psi(s,y,i) &= \int_s^T e^{-r(T-s)-y}E[(e^{X(T)}-K)^{+}|\tau=t,X(s)=y,\alpha(s)=i]\lambda_i e^{-\lambda_i(t-s)}dt \\
&= e^{-r(T-s)-y}\int_s^T E\left(E[(e^{X(T)}-K)^{+}|X(t),\alpha(t)]|\tau=t,X(s)=y,\alpha(s)=i\right) \\
&\qquad \lambda_i e^{-\lambda_i(t-s)}dt \\
&= e^{-r(T-s)-y}\int_s^T E\left(e^{X(t)}e^{r(T-t)}\psi(t,X(t),\alpha(t))|\tau=t,X(s)=y,\alpha(s)=i\right) \\
&\qquad \lambda_i e^{-\lambda_i(t-s)}dt.
\end{aligned}$$

When the initial state is geometric Brownian motion,

$$\begin{aligned}
\mathcal{T}\psi(s,y,1) &= \int_s^T e^{-r(T-s)-y}\left(\int_{-\infty}^{\infty} e^{y+u+r(T-t)}\psi(t,y+u,2)\right. \\
&\qquad \left. N((\mu-\frac{\sigma_1^2}{2})(t-s),\sigma_1^2(t-s))du\right)\lambda_1 e^{-\lambda_1(t-s)}dt \\
&= \int_s^T e^{-r(t-s)}\left(\int_{-\infty}^{\infty} e^{u}\psi(t,y+u,2)N\left((\mu-\frac{\sigma_1^2}{2})(t-s),\sigma_1^2(t-s)\right)du\right) \\
&\qquad \lambda_1 e^{-\lambda_1(t-s)}dt.
\end{aligned}$$

This is the same as the first result as in Theorem (4.2.1). We next prove the uniqueness of the solution. Let

$$\rho_1 = \int_s^T e^{-r(t-s)}\left(\int_{-\infty}^{\infty} e^{u}N\left((\mu-\frac{\sigma_1^2}{2})(t-s),\sigma_1^2(t-s)\right)du\right)\lambda_1 e^{-\lambda_1(t-s)}dt.$$

It is easily verified that

$$\int_{-\infty}^{\infty} e^{u}N\left((\mu-\frac{\sigma_1^2}{2})(t-s),\sigma_1^2(t-s)\right)du = e^{r(t-s)}.$$

Then

$$\rho_1 = \int_s^T e^{-r(t-s)}(e^{r(t-s)})\lambda_1 e^{-\lambda_1(t-s)}dt$$
$$= 1 - e^{-\lambda_1(T-s)}$$
$$< 1.$$

This leads to the uniqueness of above solution to the European call pricing problem with initial state geometric Brownian motion by the contraction mapping fixed point theorem.

Similarly, when the initial state is mean reversion, define $V(t) = e^{at}(X(t)-L)$ with initial value $V(s) = e^{as}(y - L)$. The associated differential equation for $V(t)$ is

$$dV(t) = \sigma_2 e^{at} dW(t).$$

Using the change-of-time method, the solution to the above SDE is

$$V(t) = V(s) + e^{as}W(\phi_{t-s}^{-1}),$$

where $\phi_{t-s}^{-1} = \sigma_2^2(e^{2a(t-s)} - 1)/(2a)$.

Then, $V(t) - V(s)$ is a Gaussian process with mean 0 and variance $\sigma_2^2(e^{2at} - e^{2as})/(2a)$. Using $X(t) = e^{-at}V(t) + L$ to calculate the above expectation, we have

$$\begin{aligned}
\mathcal{T}\psi(s, y, 2) &= \int_s^T e^{-r(T-s)-y} E[(e^{X(T)} - K)^+ | \tau = t, X(s) = y, \alpha(s) = 2]\lambda_2 e^{-\lambda_2(t-s)}dt \\
&= e^{-r(T-s)-y} \int_s^T E\left(E[(e^{X(T)} - K)^+ | X(t), \alpha(t)] | \tau = t, X(s) = y, \alpha(s) = 2\right) \\
&\qquad \lambda_2 e^{-\lambda_2(t-s)}dt \\
&= e^{-r(T-s)-y} \int_s^T E\left(e^{X(t)} e^{r(T-t)} \psi(t, X(t), \alpha(t)) | \tau = t, X(s) = y, \alpha(s) = 2\right) \\
&\qquad \lambda_2 e^{-\lambda_2(t-s)}dt \\
&= e^{-r(T-s)-y} \int_s^T E(\exp(e^{-at}V(t) + L)e^{r(T-t)} \\
&\qquad \psi(t, (e^{-at}V(t) + L), \alpha(t)) | \tau = t, X(s) = y, \alpha(s) = 2)\lambda_2 e^{-\lambda_2(t-s)}dt
\end{aligned}$$

$$= \int_s^T e^{-r(T-s)-y}\left[\int_{-\infty}^{\infty} e^{r(T-t)} \exp[e^{-at}(e^{as}(y-L)+e^{as}v)+L]\right.$$
$$\psi(t, e^{-at}(e^{as}(y-L)+e^{as}v)+L, 1)$$
$$\left. N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right) dv\right] \lambda_2 e^{-\lambda_2(t-s)}dt$$
$$= \int_s^T e^{-r(t-s)-y}\left[\int_{-\infty}^{\infty} \exp[e^{-a(t-s)}(y-L+v)+L]\psi(t, e^{-a(t-s)}(y-L+v)+L, 1)\right.$$
$$\left. N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right) dv\right] \lambda_2 e^{-\lambda_2(t-s)}dt.$$

It remains to prove uniqueness of the above solution. Let

$$\rho_2 = \int_s^T e^{-r(t-s)-y}\int_{-\infty}^{\infty} \exp[e^{-a(t-s)}(y-L+v))+L]$$
$$N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right) dv \lambda_2 e^{-\lambda_2(t-s)}dt.$$

We want to show that $0 \leq \rho_2 < 1$. Note that,

$$e^{-r(t-s)-y}\int_{-\infty}^{\infty} \exp[e^{-a(t-s)}(y-L+v))+L]N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right) dv$$
$$= e^{-r(t-s)}\exp(e^{-a(t-s)}y-y)\exp(L(1-e^{-a(t-s)}))\int_{-\infty}^{\infty}\exp(e^{-a(t-s)}v)\left(\frac{\exp\left(\frac{-v^2}{2\phi_{t-s}^{-1}}\right)}{\sqrt{2\pi\phi_{t-s}^{-1}}}\right)dv$$
$$\leq e^{-r(t-s)}\exp(L(1-e^{-a(t-s)}))\int_{-\infty}^{\infty}\exp(e^{-a(t-s)}v)\left(\frac{\exp\left(\frac{-v^2}{2\phi_{t-s}^{-1}}\right)}{\sqrt{2\pi\phi_{t-s}^{-1}}}\right)dv$$
$$= e^{-r(t-s)}\exp(L(1-e^{-a(t-s)}))\exp\left(\frac{\phi_{t-s}^{-1}e^{-2a(t-s)}}{2}\right)\int_{-\infty}^{\infty}\frac{\exp\left(-(v-e^{-a(t-s)}\phi_{t-s}^{-1})^2\right)}{2\phi_{t-s}^{-1}\sqrt{2\pi\phi_{t-s}^{-1}}}dv$$
$$= \exp[-r(t-s)+L(1-e^{-a(t-s)})+\frac{\phi_{t-s}^{-1}e^{-2a(t-s)}}{2}]$$
$$= \exp[-r(t-s)+L(1-e^{-a(t-s)})+\frac{\sigma_2^2(e^{2a(t-s)}-1)}{4a}e^{-2a(t-s)}].$$

Then,

$$\rho_2 \leq \int_s^T \exp[-r(t-s)+L(1-e^{-a(t-s)})+\frac{\sigma_2^2(e^{2a(t-s)}-1)}{4a}e^{-2a(t-s)}]\lambda_2 e^{-\lambda_2(t-s)}dt$$
$$= \int_0^{T-s}\exp[-ru+L(1-e^{-au})+\frac{\sigma_2^2(1-e^{-2au})}{4a}]\lambda_2 e^{-\lambda_2 u}du.$$

Note that $\exp[-ru + L(1-e^{-au}) + \sigma_2^2(1-e^{-2au})/(4a)] \leq 1$, we have

$$\rho_2 \leq \int_0^{T-s} \lambda_2 e^{-\lambda_2 u} du$$
$$= 1 - e^{-\lambda_2(T-s)} < 1.$$

The above assumption is equivalent to

$$-ru + L(1-e^{-au}) + \sigma_2^2(1-e^{-2au})/(4a) \leq 0,$$

for u in $[0, T-s]$.

Define $f(u) = -ru + L(1-e^{-au}) + \sigma_2^2(1-e^{-2au})/(4a)$, with $f(0) = 0$.

We can see that

$$f'(u) = -r + aLe^{-au} + \frac{\sigma_2^2 e^{-2au}}{2} < -r + aL + \frac{\sigma_2^2}{2} = f'(0).$$

Under the sufficient condition $aL + \sigma_2^2/2 - r < 0$, the inequalities

$$\| \tau\psi(s,y,2) \| \leq \rho_2 \| \psi(s,y,2) \|,$$

$$0 \leq \rho_2 < 1,$$

hold.

Therefore, by the contraction mapping principle,

$$\psi(s,y,2) = \int_s^T e^{-r(t-s)-y} \int_{-\infty}^{\infty} (\exp(e^{-a(t-s)}(y-L+v)+L))$$
$$\psi(t, e^{-a(t-s)}(y-L+v)+L, 1) N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right) dv \lambda_2 e^{-\lambda_2(t-s)} dt$$
$$+ e^{-\lambda_2(T-s)}\psi^0(s,y,2),$$

gives the unique solution to the European call pricing problem.

□

Finally, European call option prices can be calculated by

$$c(s,x,1) = x\psi(s,\log x,1),$$

and

$$c(s,x,2) = x\psi(s,\log x,2).$$

4.3 Comparison with the Monte Carlo Approach in the Single Jump Case

We consider a simpler case when the non-initial state is absorbing. So the second row of Q is zero. This means at most one jump happens during the life of the option. Based on the successive approximation method considered in the above section, the pricing formula for the single jump case is illustrated in the next theorem. All the following numerical examples are based on this single jump assumption.

Theorem 4.3.1. *In the case of geometric Brownian motion as initial state,*

$$\psi(s,y,1) = \int_s^T e^{-r(t-s)}\left(\int_{-\infty}^{\infty} e^u\psi^0(t,y+u,2)N\left((\mu-\frac{\sigma_1^2}{2})(t-s),\sigma_1^2(t-s)\right)du\right) \lambda_1 e^{-\lambda_1(t-s)}dt + e^{-\lambda_1(T-s)}\psi^0(s,y,1).$$

In the case of mean reversion as initial state,

$$\psi(s,y,2) = \int_s^T e^{-r(t-s)-y}\left(\int_{-\infty}^{\infty} \exp(e^{-a(t-s)}(y-L+v)+L) \psi^0(t,e^{-a(t-s)}(y-L+v)+L,1)N\left(0,\frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right)dv\right) \lambda_2 e^{-\lambda_2(t-s)}dt + e^{-\lambda_2(T-s)}\psi^0(s,y,2).$$

Now we use the Monte Carlo method to test the effectiveness of the analytical solution to European option price obtained above. We compare the option prices calculated by the analytical solution with that generated by the Monte Carlo method. And we graph option prices with respect to different initial stock prices.

In order to apply the Monte Carlo method to a recursive equation, we approximate equation (2.4.6) in a discrete form as follows:

$$\frac{S_{(k+1)\delta} - S_{k\delta}}{S_{k\delta}} = A(S_{k\delta}, \alpha_{k\delta})\delta + \sigma(\alpha_{k\delta})(W_{(k+1)\delta} - W_{k\delta}). \tag{4.3.2}$$

Because $W_{k\delta}$ is normally distributed, we can see that $W_{(k+1)\delta} - W_{k\delta}/\sqrt{\delta}$ is a Gaussian process with mean 0 and variance 1. Therefore we can simplify the equation (4.3.2) as

$$S_{(k+1)\delta} = S_{k\delta} + S_{k\delta}\left(A(S_{k\delta}, \alpha_{k\delta})\delta + \sigma\sqrt{\delta}u_{k\delta}\right). \tag{4.3.3}$$

where $u_{k\delta}$ is the standard Gaussian random variable.

We use the computer to repeatedly generate 1000 random numbers according to the standard normal distribution and generate a random number representing the switching time. In the recursive equation (4.3.3), $\alpha_{k\delta}$ stays the same when the iteration is before the switching time, and changes to the other state after the switching time. Then the repeated computation gives us a number which is our future stock price by Monte Carlo simulation. Repeating this procedure 1000 times with 10 different switching time, we can get 10000 possible future stock prices $S(T)_i$.

European call option price $Ee^{-rT}(S_T - K)^+$ can be translated into a discrete format as

$$c = e^{-rT}\frac{\sum_{i=1}^{10000}(S_T(i) - K)^+}{N} = \frac{e^{-rT}}{2N}\sum_{i=1}^{10000}[|S_T(i) - K| + (S_T(i) - K)]. \tag{4.3.4}$$

Applying equation (4.3.4) to those $S_T(i)$ to get the European call option prices. Below we compare the European call option prices obtained by analytical solution and those simulated by the Monte Carlo method.

First we consider the situation when the initial state is GBM. Take r=0.0467, σ_1=σ_2=0.2, T=0.24, λ=10, i=1. The following charts show the European call prices obtained by analytical solution (AS) and by the Monte Carlo method (MC) and the relative error of them, as stock price, strike price, a and L change respectively.

Stock price	40	42	44	46	48	50
AS	0.373813	0.83652	1.58522	2.62835	3.92282	5.39838
MC	0.348533	0.794791	1.52858	2.56488	3.86457	5.35767
Rel. Error	0.07254	0.05250	0.03705	0.02475	0.01507	0.0076

Table 4.1: Calls obtained by AS and MC when $K = 45, a = 1, L = 3.9, i = 1$

Stike price	37	40	42	45	47	50
AS	9.00712	6.12589	4.35873	2.21531	1.24105	0.42377
MC	8.86374	5.9971	4.2513	2.15318	1.2073	0.415329
Rel. Error	0.01618	0.02147	0.02527	0.02886	0.02795	0.02031

Table 4.2: Calls obtained by AS and MC when $S = 45.27, a = 1, L = 3.9, i = 1$

a	0.01	0.1	0.5	1	2	5
AS	2.06384	2.07719	2.13762	2.21531	2.37530	2.84506
MC	2.01618	2.02830	2.08304	2.15318	2.29673	2.70960
Rel. Error	0.02364	0.02411	0.02620	0.02885	0.03421	0.04999

Table 4.3: Calls obtained by AS and MC when $S = 45.27, K = 45, L = 3.9, i = 1$

L	3.75	3.78	3.83	3.87	3.91	3.95
AS	1.68452	1.78135	1.95309	2.09983	2.25484	2.41805
MC	1.71394	1.79327	1.93491	2.05679	2.18625	2.32327
Rel. Error	-0.01716	-0.00665	0.0094	0.02093	0.03137	0.0408

Table 4.4: Calls obtained by AS and MC when $S = 45.27, K = 45, a = 1, i = 1$

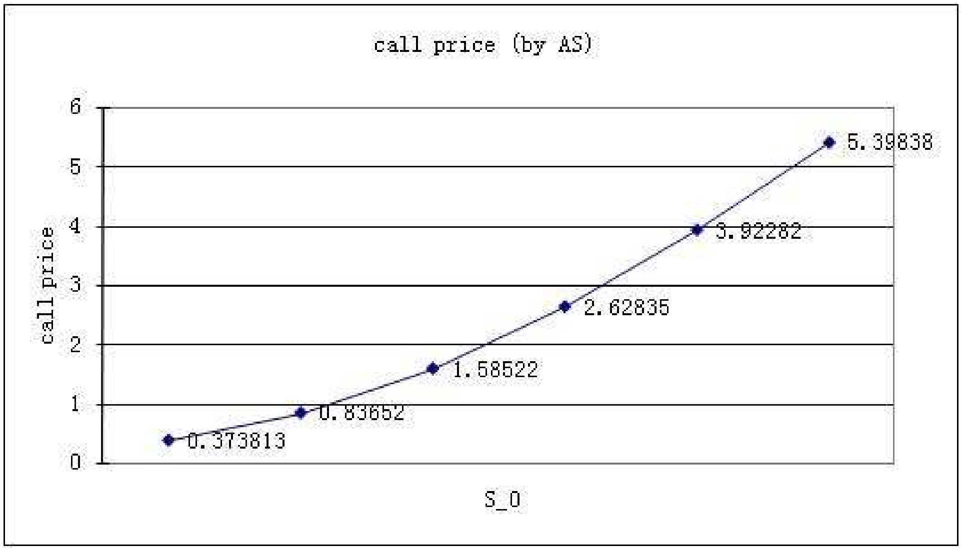

Figure 4.1: Graph of call price obtained by AS with initial state GBM

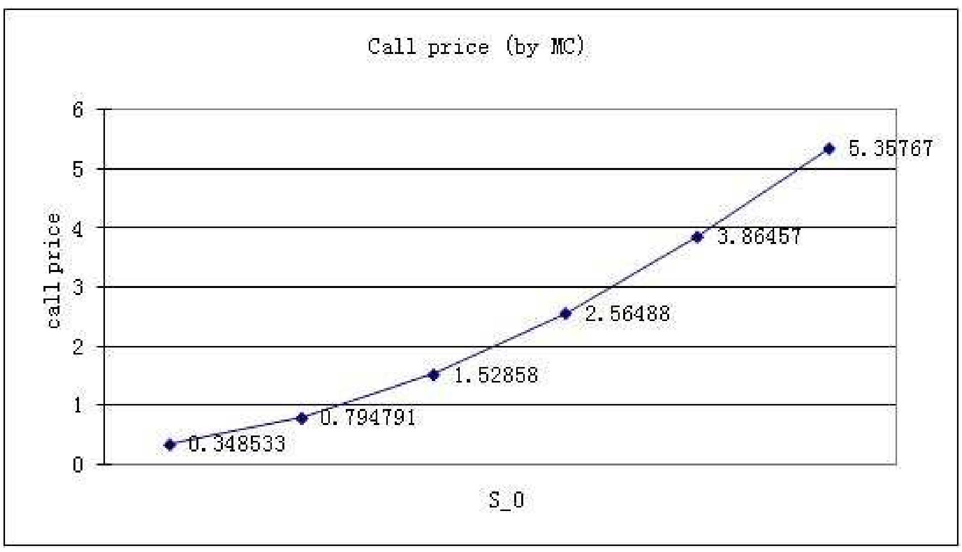

Figure 4.2: Graph of call price obtained by MC with initial state GBM

As we can see in these charts, the call prices using AS and MC are similar, and the relative errors are mostly below 0.03. As initial stock prices increases, so are those two set of call prices, that is also shown in the graphs.

The other case is when the initial state is MR. Take r=0.0467, σ_1=σ_2=0.2, T=0.24, λ=10, i=2. We have the following charts showing the European call prices computed using the analytical and Monte Carlo methods, as stock price, strike price, a and L change respectively.

Stock price	40	42	44	46	48	50
AS	0.36982	0.83634	1.60137	2.67797	4.02766	5.57836
MC	0.36223	0.81749	1.56213	2.61097	3.92655	5.43963
Rel. Error	0.02096	0.02305	0.02511	0.02566	0.02574	0.0255

Table 4.5: Calls obtained by AS and MC when $K = 45, a = 1, L = 3.9, i = 2$

Stike price	37	40	42	45	47	50
AS	8.91174	6.05986	4.33160	2.25043	1.29800	0.47521
MC	8.85833	6.00366	4.27284	2.19446	1.25089	0.44659
Rel. Error	0.00603	0.00936	0.01376	0.0255	0.03766	0.06408

Table 4.6: Calls obtained by AS and MC when $S = 45.27, a = 1, L = 3.9, i = 2$

a	0.01	0.1	0.5	1	2	5
AS	2.10456	2.11786	2.17694	2.25043	2.39512	2.79090
MC	2.03609	2.05042	2.11435	2.19446	2.35308	2.78265
Rel. Error	0.03362	0.03289	0.0296	0.0255	0.01786	0.00297

Table 4.7: Calls obtained by AS and MC when $S = 45.27, K = 45, L = 3.9, i = 2$

L	3.75	3.78	3.83	3.87	3.91	3.95
AS	1.91924	1.98096	2.08887	2.1797	2.2745	2.3733
MC	1.80936	1.87929	2.00372	2.11036	2.22326	2.34234
$Rel.\ Error$	0.06073	0.0541	0.0425	0.03286	0.02305	0.01322

Table 4.8: Calls obtained by AS and MC when $S = 45.27, K = 45, a = 1, i = 2$

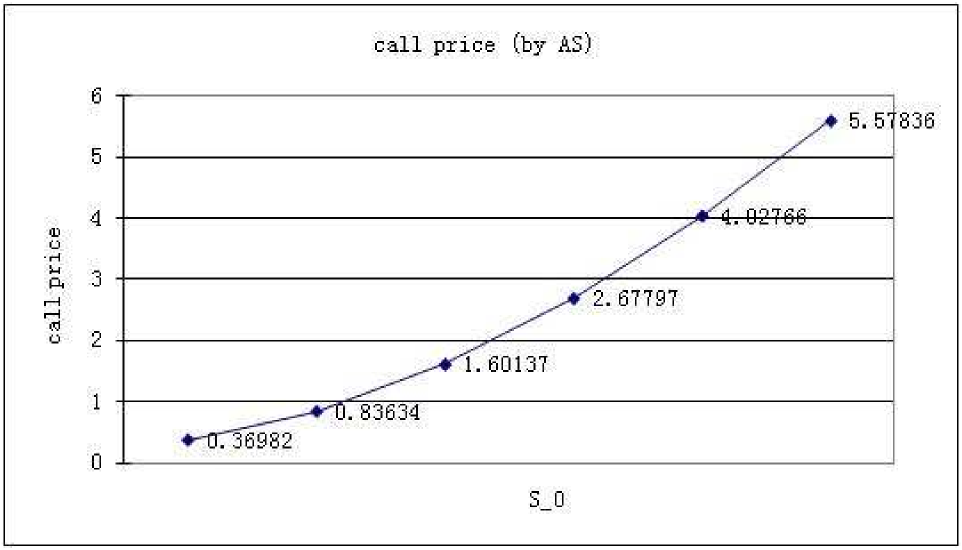

Figure 4.3: Graph of call price obtained by AS with initial state MR

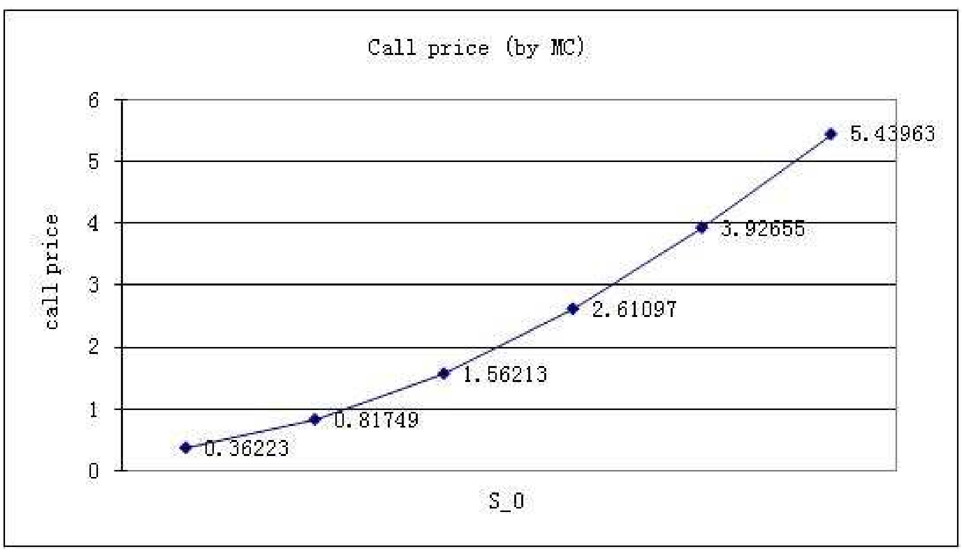

Figure 4.4: Graph of call price obtained by MC with initial state MR

In the case of initial state MR, we have similar results. Both methods give us similar solutions and most of the relative errors are below 0.04. The call prices also increase together with the initial stock prices.

These examples have verified that the analytical solutions given in Theorem (4.3.1) are consistent with the Monte Carlo method in pricing European call options.

4.4 Stochastic Approximation

The idea of stochastic approximation is to solve an optimization problem to help to estimate parameters of the model. The recursive procedure introduced in paper [36] is defined by

$$\sigma_{n+1} = \Pi\big[\sigma_n - \varepsilon_n[c(\sigma_n) - c_n]c_\sigma(\sigma_n)\big], \tag{4.4.5}$$

where $\{\varepsilon_n\}$ is a sequence of nonnegative decreasing step sizes satisfying $\varepsilon_n \to 0$ as $n \to \infty$, and $\sum_n \varepsilon_n = \infty$. We assume the step size $\{\varepsilon_n\}$ to be of the form $\varepsilon_n = O(1/n)$. $c_\sigma(\cdot)$ denotes the derivative of $c(\cdot)$ with respect to σ. For some $M > 0$ Π is defined by

$$\Pi[\sigma] = \begin{cases} 0, & \text{if } \sigma < 0, \\ M, & \text{if } \sigma > M, \\ \sigma, & \text{otherwise.} \end{cases}$$

To make equation (4.4.5) easier to work with, we rewrite it as

$$\sigma_{n+1} = \sigma_n - \varepsilon_n[c(\sigma_n) - c_n]c_\sigma(\sigma_n) + \varepsilon_n F_n, \tag{4.4.6}$$

where $\varepsilon_n F_n$ is to make sure that σ_{n+1} is bounded between 0 and M with the shortest distance needed to bring $\sigma_n - \varepsilon_n[c(\sigma_n) - c_n]c_\sigma(\sigma_n)$ back to the interval $[0, M]$ if it ever escapes from there.

Assume

(C1) $c(\cdot)$ is twice continuously differentiable.

(C2) There exists a $\overline{c}$ such that

$$\varepsilon_n \sum_{k=0}^{n-1} c_k \to \overline{c} \quad \text{w.p.1 as} \quad n \to \infty. \tag{4.4.7}$$

The functions obtained in Theorem (4.2.1)

$$\psi(s,y,1) = \int_s^T e^{-r(t-s)} \left(\int_{-\infty}^{\infty} e^u \psi(t, y+u, 2) N\left((\mu - \frac{\sigma_1^2}{2})(t-s), \sigma_1^2(t-s) \right) du \right)$$
$$\lambda_1 e^{-\lambda_1(t-s)} dt + e^{-\lambda_1(T-s)} \psi^0(s,y,1),$$

and

$$\psi(s,y,2) = \int_s^T e^{-r(t-s)-y} \left[\int_{-\infty}^{\infty} (\exp(e^{-a(t-s)}(y-L+v)+L)) \right.$$
$$\left. \psi(t, e^{-a(t-s)}(y-L+v)+L, 1) N\left(0, \frac{\sigma_2^2}{2a}(e^{2a(t-s)}-1)\right) dv \right]$$
$$\lambda_2 e^{-\lambda_2(t-s)} dt + e^{-\lambda_2(T-s)} \psi^0(s,y,2),$$

both have Gaussian density function in terms of σ_i inside the double integrals, so the condition (C1) obviously holds.

Since we assumed that $\{\varepsilon_n\}$ is of the form $\varepsilon_n = O(1/n)$. It follows from the law of large numbers that $\displaystyle\frac{1}{n}\sum_{k=0}^{n-1} c_k \to \overline{c}$ w.p.1, as $n \to \infty$, for some number $\overline{c}$. So the condition (C2) holds too.

In order to establish the convergence of the algorithm, we define a piecewise constant interpolation and then analyze a sequence of functions instead of the discrete iterates. Define

$$t_0 = 0, \quad t_n = \sum_{k=0}^{n-1} \varepsilon_k,$$
$$a(t) = \begin{cases} n, & t_n \le t < t_{n+1}, \quad \text{for} \quad t \ge 0, \\ 0, & \text{for} \quad t < 0. \end{cases}$$

Define the continuous-time interpolation $\sigma^0(\cdot)$ on $(-\infty, \infty)$ and its shift process as follows:

$$\sigma^0(t) = \begin{cases} \sigma_0, & \text{for } \ t < 0, \\ \sigma_n & \text{for } \ t \geq 0 \ \text{ and } \ t_n \leq t < t_{n+1}. \end{cases}$$

$$\sigma^n(t) = \sigma^0(t_n + t), \quad t \in (-\infty, \infty).$$

And to work with error term in the initial iteration equation as a function, define

$$\begin{aligned} F_n &= 0 \ \text{ and } \ [c(\sigma_n) - c_n]c_\sigma(\sigma_n) = 0, \quad \text{for } \ n < 0, \\ F^0(t) &= \sum_{k=0}^{a(t)-1} \varepsilon_k F_k \ \text{ for } \ t \geq 0, \\ F^n(t) &= \begin{cases} F^0(t_n + t) - F^0(t_n) = \displaystyle\sum_{k=n}^{a(t_n+t)-1} \varepsilon_k F_k, & t \geq 0, \\ -\displaystyle\sum_{k=a(t_n+t)}^{n-1} \varepsilon_k F_k, & t < 0. \end{cases} \end{aligned}$$

As in Kushner and Yin [29, Section 4.3], define a set $C(\sigma)$ as follows. For $\sigma \in (0, M)$, $C(\sigma)$ contains only the zero element; for $\sigma = 0$ or $\sigma = M$, $C(\sigma)$ is the infinite cone (interval $(-\infty, 0)$ or (M, ∞)) pointing in the direction away from $[0, M]$.

Theorem 4.4.1. *Assume conditions* (C1) *and* (C2) *hold, then* $(\sigma^n(\cdot), F^n(\cdot))$ *is equicontinuous w.p.1 in the extended sense* (see Kushner and Yin [29, p. 102]). *Let* $(\sigma(\cdot), F(\cdot))$ *be the limit of a convergent subsequence of* $(\sigma^n(\cdot), F^n(\cdot))$. *Then this limit satisfies the projected ordinary differential equation*

$$\dot{\sigma}(t) = -[c(\sigma(t)) - \overline{c}]c_\sigma(\sigma(t)) + f(t), \quad f(t) \in -C(\sigma(t)), \tag{4.4.8}$$

where f *is the minimal number needed to keep the solution in* $[0, M]$ *with*

$$F(t) = \int_0^t f(s)ds.$$

Proof. Define

$$\begin{aligned} \widetilde{c}_n &= -[c(\sigma_n) - \overline{c}]c_\sigma(\sigma_n), \quad \text{for } \ n \geq 0, \quad \text{and } \ \widetilde{c}_n = 0 \ \text{ for } \ n < 0, \\ \widehat{c}_n &= [c_n - \overline{c}]c_\sigma(\sigma_n) \ \text{ for } \ n \geq 0, \quad \text{and } \ \widehat{c}_n = 0 \ \text{ for } \ n < 0. \end{aligned}$$

Since (4.4.6) can be written as

$$\sigma_{n+1} = \sigma_n - \varepsilon_n[c(\sigma_n)c_\sigma(\sigma_n) - \bar{c}c_\sigma(\sigma_n)] + \varepsilon_n(c_n - \bar{c})c_\sigma(\sigma_n) + \varepsilon_n F_n, \tag{4.4.9}$$

We also have

$$\sigma_{n+1} = \sigma_n + \varepsilon_n \widetilde{c}_n + \varepsilon_n \widehat{c}_n + \varepsilon_n F_n.$$

Define new interpolations $\widetilde{c}^0(\cdot)$, $\widetilde{c}^n(\cdot)$, $\widehat{c}^0(\cdot)$, and $\widehat{c}^n(\cdot)$ of $\widetilde{c}_n$ and $\widehat{c}_n$ in the same way as we defined $F^0(\cdot)$ and $F^n(\cdot)$. Then

$$\sigma^n(t) = \sigma_n + \widetilde{c}^n(t) + \widehat{c}^n(t) + F^n(\cdot), \quad \text{for} \quad t \in (-\infty, \infty).$$

Here we verify an asymptotic rate of change condition holds. Using Kushner and Yin [29, Theorem 6.1.1] we can prove our theorem. Recall our definition of $\widehat{c}^0(\cdot)$,

$$\widehat{c}^0(t) \stackrel{\text{def}}{=} \sum_{k=0}^{a(t)-1} \varepsilon_k(c_k - \bar{c})c_\sigma(\sigma).$$

Recall that the asymptotic rate of change of $\widehat{c}^0(t)$ is said to go to 0 with probability 1 if for some $T > 0$,

$$\lim_n \sup_{j \geq n} \max_{0 \leq t \leq T} |\widehat{c}^0(jT + t) - \widehat{c}^0(jT)| = 0, \quad \text{w.p.1.}$$

Using

$$D_n \stackrel{\text{def}}{=} \sum_{k=0}^{n} [c_k - \bar{c}],$$

we get by means of a partial summation,

$$\sum_{k=m}^{n} \varepsilon_k[c_k - \bar{c}]c_\sigma(\sigma) = \varepsilon_n[D_{n+1} - D_m]c_\sigma(\sigma) + \sum_{k=m}^{n-1}[D_{k+1} - D_m][\varepsilon_k - \varepsilon_{k+1}]c_\sigma(\sigma).$$

Letting $m = 0$, $n = a(t) - 1$, we obtain

$$\widehat{c}^0(t) = \varepsilon_{a(t)-1} D_{a(t)} c_\sigma(\sigma) + \sum_{k=0}^{a(t)-2} D_{k+1} \frac{\varepsilon_k - \varepsilon_{k+1}}{\varepsilon_k} \varepsilon_k c_\sigma(\sigma).$$

Note that (C2) implies that $\varepsilon_{a(t)-1} D_{m(t)} c_\sigma(\sigma) \to 0$ as $a(t) \to \infty$ (or $n \to \infty$) and also

$$\sum_{k=0}^{a(t)-2} D_{k+1} \frac{\varepsilon_k - \varepsilon_{k+1}}{\varepsilon_k} \varepsilon_k c_\sigma(\sigma) = \sum_{k=0}^{a(t)-2} D_{k+1} O(\varepsilon_k^2) c_\sigma(\sigma) \to 0.$$

Thus the asymptotic rate of change of $\widehat{c}^0(t)$ goes to 0 w.p.1. Using Kushner and Yin [29, Theorem 6.1.1], the conclusion in the theorem follows. $\square$

Corollary 4.4.2. *In addition to the conditions in Theorem* 4.4.1*, suppose that* $c_\sigma(\sigma) \neq 0$ *for all* σ*, and that* σ_* *is the unique solution of* $c(\sigma) - \bar{c} = 0$ *with* $\sigma_* \in (0, M)$ *such that* σ_* *is in the set of locally asymptotic stable points of the projected ODE. Then* $\sigma_n \to \sigma_*$ *w.p.1.*

We have shown the convergence of algorithm (4.4.5). The next part of the section concerns the rate of convergence. We assume that $\varepsilon_n = 1/(n+1)$ and all the conditions of Corollary 4.4.2 hold. Since σ_* is strictly in the constrained set $(0, M)$, without loss of generality, we will drop the term $\varepsilon_n F_n$ in the rest of this section, and assume that the sequence of iterates $\{\sigma_n\}$ is nonnegative and uniformly bounded by M.
We want to exploit the asymptotic properties of the scaled sequence $u_n = \sqrt{n+1}(\sigma_n - \sigma_*)$. We use linearization and local analysis to show that the interpolation of u_n converges to a diffusion limit. We rewrite (4.4.6) with $\varepsilon_n F_n$ dropped as

$$\begin{aligned} u_{n+1} = \sqrt{\frac{n+2}{n+1}} u_n & - \frac{1}{n+1}\sqrt{\frac{n+2}{n+1}} c_\sigma^2(\sigma_*) u_n + \sqrt{\frac{n+2}{n+1}} \frac{1}{\sqrt{n+1}} (c_n - \bar{c}) c_\sigma(\sigma_*) \\ & + \sqrt{\frac{n+2}{n+1}} \frac{1}{n+1} (c_n - \bar{c}) g(u_n) + \frac{1}{n+1} o(|u_n|), \end{aligned} \tag{4.4.10}$$

where $g(\cdot)$ is a continuous function and $g(u) = O(|u|)$.

Define an auxiliary process v_n by

$$\begin{aligned} v_0 = u_0, v_{n+1} = v_n & - \frac{1}{n+1}\left(c_\sigma^2(\sigma_*) - \frac{1}{2}\right) v_n + \frac{1}{\sqrt{n+1}} (c_n - \bar{c}) c_\sigma(\sigma_*) \\ & + \frac{1}{n+1} (c_n - \bar{c}) g(v_n) + \frac{1}{n+1} o(|v_n|). \end{aligned} \tag{4.4.11}$$

Since

$$\sqrt{\frac{n+2}{n+1}} = 1 + \frac{1}{2(n+1)} + O\left(\frac{1}{(n+1)^2}\right), \tag{4.4.12}$$

we only need to study v_n for the asymptotics of u_n.

(C3) The $\{c_n - \bar{c}\}$ is a stationary ϕ-mixing sequence with 0 mean and mixing rate ϕ_k satisfying $\sum_{k=0}^{\infty} \phi_k^{1/2} < \infty$. In addition, suppose $c_\sigma^2(\sigma_*) > 1/2$.

Remark 4.4.3. Under (C3), it can be shown (see Kushner and Yin [29, Chapter 7]) that $\sum_{k=0}^{a(t)-1} \frac{1}{\sqrt{k+1}}(c_k - \overline{c})$ converges weakly to a real-valued standard Brownian motion with variance $\varsigma^2 t$, where

$$\varsigma^2 = E(c_0 - \overline{c})^2 + 2\sum_{k=1}^{\infty} E(c_k - \overline{c})(c_0 - \overline{c}).$$

Moreover, by using the well-known mixing inequality (see [3, p. 166]), we obtain

$$E\left|\sum_{k=0}^{n-1} \frac{1}{\sqrt{k+1}}(c_k - \overline{c})\right|^2 \le K$$

for some $K > 0$.

Lemma 4.4.4. *In addition to the conditions of Corollary* 4.4.2*, assume* (C3) *holds. Then* $\{v_n\}$ *is tight.*

Proof. We claim that $\sup_n E|v_n| < \infty$. To this end, define

$$A_{nk} = \begin{cases} \displaystyle\prod_{j=k+1}^{n} \left(1 - \frac{c_\sigma^2(\sigma_*) - \frac{1}{2}}{j+1}\right), & \text{if } \ k < n, \\ 1, & \text{if } \ k = n. \end{cases}$$

Then

$$\begin{aligned} v_{n+1} = A_{n0} v_0 &+ \sum_{k=0}^{n} \frac{1}{\sqrt{k+1}} A_{nk}(c_k - \overline{c}) c_\sigma(\sigma_*) \\ &+ \sum_{k=0}^{n} \frac{1}{k+1} A_{nk}(c_k - \overline{c}) g(v_k) + \sum_{k=0}^{n} \frac{1}{k+1} A_{nk} o(|v_k|). \end{aligned}$$

Note that in view of (C3), $E|c_k - \overline{c}||g(v_k)| \le K E|v_k|$. Thus, we obtain

$$E|v_{n+1}| \le |A_{n0}|E|v_0| + E\left|\sum_{k=0}^{n} \frac{1}{\sqrt{k+1}} A_{nk}(c_k - \overline{c}) c_\sigma(\sigma_*)\right| + K\sum_{k=0}^{n} \frac{1}{k+1}|A_{nk}|E|v_k|. \quad (4.4.13)$$

It is easily verified that by (C3),

$$\sum_{k=0}^{n} \frac{1}{k+1}|A_{nk}| = \sum_{k=0}^{n} \frac{1}{k+1} A_{nk} < \infty.$$

Using the mixing inequality in [3], we have

$$
\begin{aligned}
&E\left|\sum_{k=0}^{n}\frac{1}{\sqrt{k+1}}A_{nk}(c_k-\overline{c})c_\sigma(\sigma_*)\right| \\
&\leq E^{\frac{1}{2}}\left|\sum_{k=0}^{n}\frac{1}{\sqrt{k+1}}A_{nk}(c_k-\overline{c})c_\sigma(\sigma_*)\right|^2 \\
&=\left(\sum_{k=0}^{n}\sum_{j=0}^{n}\frac{1}{\sqrt{k+1}}\frac{1}{\sqrt{j+1}}A_{nk}A_{nj}(c_k-\overline{c})(c_j-\overline{c})c_\sigma^2(\sigma_*)\right)^{\frac{1}{2}} \\
&\leq K\left(\sum_{j=0}^{n}\frac{1}{j+1}A_{nj}^2\sum_{k>j}^{n}E(c_k-\overline{c})(c_j-\overline{c})c_\sigma^2(\sigma_*)\right)^{\frac{1}{2}} \\
&\leq K\left(\sum_{j=0}^{n}\frac{1}{j+1}A_{nj}^2\sum_{k>j}^{n}|E(c_k-\overline{c})(c_j-\overline{c})-E(c_k-\overline{c})E(c_j-\overline{c})|\right)^{\frac{1}{2}} \\
&\leq K\left(\sum_{j=0}^{n}\frac{1}{j+1}A_{nj}^2\sum_{k<j}^{n}\phi_{k-j}\right)^{\frac{1}{2}} \\
&\leq K<\infty.
\end{aligned}
$$

We use K as a generic positive constant, its value may change at different appearances. Adding the above estimates, we have

$$
E|v_{n+1}|\leq K+K\sum_{k=0}^{n}\frac{1}{k+1}|A_{nk}|E|v_k|. \tag{4.4.14}
$$

An application of the Gronwall's inequality leads to

$$
E|v_{n+1}|\leq K<\infty \quad\text{and}\quad \sup_n E|v_n|<\infty.
$$

The desired tightness then follows from the well-known Markov inequality

$$
P\left(|v_n|\geq\widetilde{K}\right)\leq\frac{\sup_n E|v_n|}{\widetilde{K}}.
$$

The lemma is proved. □

Lemma 4.4.5. *Under the conditions of Lemma* 4.4.4*,* $\lim_n E|v_n-u_n|=0$.

Proof. We merely use the expansions in (4.4.12), and the definitions of u_n and v_n in (4.4.10) and (4.4.11), respectively. Detailed calculation yields the desired result. □

Define a piecewise constant interpolation $v^0(t)$ and its shift $v^n(t)$ as in the first part of this section. Then we have

$$\begin{aligned} v^n(t+s) - v^n(t) = &\sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{j+1}\left(c_\sigma^2(\sigma_*) - \frac{1}{2}\right) v_j \\ &+ \sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{\sqrt{j+1}}(c_j - \overline{c})c_\sigma(\sigma_*) \\ &+ \sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{j+1}(c_j - \overline{c})g(v_j) \\ &+ \sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{j+1} h_j v_j. \end{aligned} \tag{4.4.15}$$

Theorem 4.4.6. *The sequence of interpolated estimation errors $\{v^n(\cdot)\}$ converges weakly to $v(\cdot)$, which is a solution of the stochastic differential equation*

$$dv = \left(c_\sigma^2(\sigma_*) - \frac{1}{2}\right) v dt + \varsigma c_\sigma(\sigma_*) dw, \tag{4.4.16}$$

where $w(\cdot)$ is a real-valued standard Brownian motion.

Proof. The theorem is proven in two steps. The first step establishes tightness, and the second step characterizes the limit process.

Step 1): Show that the sequence $\{v^n(\cdot)\}$ is tight in the space $D[0,\infty)$ of functions that are right continuous, have left limits and endowed with the Skorohod topology. To this end, we apply the tightness criterion in Kushner and Yin [29]. Without loss of generality and for notational simplicity, assume that $\{v_n\}$ is bounded (otherwise, we can use a truncation device as in [29]).

Then we obtain that for any $t > 0$, $\eta > 0$, and any $0 < s \le \eta$,

$$E|v^n(t+s) - v^n(t)|^2 \le I_1 + I_2 + I_3 + I_4, \tag{4.4.17}$$

where I_i for $i = 1, 2, 3, 4$ are four terms on the right-hand side of (4.4.15). By virtue of the boundedness of $\{v_k\}$,

$$\begin{aligned} I_1 &= E\left|\sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{j+1}\left(c_\sigma^2(\sigma_*) - \frac{1}{2}\right) v_j\right|^2 \\ &\le K \sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \sum_{k=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{j+1}\frac{1}{k+1} \\ &\le Ks^2 \le K\eta^2. \end{aligned} \tag{4.4.18}$$

Thus, taking $\limsup_n$ followed by $\lim_{\eta\to 0}$, the limit is 0.

The mixing inequality implies that

$$\begin{aligned} I_2 &= E\left|\sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{\sqrt{j+1}}(c_j - \overline{c})c_\sigma(\sigma_*)\right|^2 \\ &\le Ks \le K\eta. \end{aligned} \tag{4.4.19}$$

Thus, the double limits of this term also goes to 0. Likewise,

$$I_3 + I_4 \le E\left|\sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{\sqrt{j+1}}(c_j - \overline{c})g(v_j)\right|^2 + E\left|\sum_{j=a(t_n+t)}^{a(t_n+t+s)-1} \frac{1}{\sqrt{j+1}} h_j v_j\right|^2 \le K\eta. \tag{4.4.20}$$

Combining the estimates above, we arrive at

$$\lim_{\eta\to 0}\limsup_n E|v^n(t+s) - v^n(t)|^2 = 0. \tag{4.4.21}$$

Therefore, $\{v^n(\cdot)\}$ is tight.

Step 2) Characterization of the limit process. By Prohorov's theorem, we can extract a convergent subsequence and still denote it by $\{v^n(\cdot)\}$ for notational simplicity. Denote the limit by $v(\cdot)$. By the Skorohod representation, without changing notation, we may assume that the sequence $v^n(\cdot)$ converges to $v(\cdot)$ w.p.1 and the convergence is uniform in any bounded time interval. We proceed to establish that the limit is nothing but the desired diffusion process.

We shall show that $v(\cdot)$ is a solution of the martingale problem with operator

$$\mathcal{L}f(v) = \frac{1}{2}\varsigma^2 c_\sigma^2(\sigma_*)\frac{d^2 f(v)}{dv^2} + \left(c_\sigma^2(\sigma_*) - \frac{1}{2}\right) v\frac{df(v)}{dv}, \tag{4.4.22}$$

where $f(\cdot)$ is a C^2 function with compact support. To this end, we show that

$$f(v(t+s)) - f(v(t)) - \int_t^{t+s} \mathcal{L}f(v(\tau))d\tau \text{ is a martingale.}$$

To do so, for any bounded and continuous function $\rho(\cdot)$, any $t, s > 0$, any positive integer κ, and $0 \le t_1 \le t_2 \le \cdots \le t_\kappa \le t$, we will show

$$E\rho(v(t_i) : i \le \kappa)\Big[f(v(t+s)) - f(v(t)) - \int_t^{t+s} \mathcal{L}f(v(\tau))d\tau\Big] = 0.$$

First, by the weak convergence and the Skorohod representation, it is readily seen that

$$\begin{aligned} &E\rho(v^n(t_i) : i \le \kappa)[f(v^n(t+s)) - f(v^n(t))] \\ &\quad \to \rho(v(t_i) : i \le \kappa)[f(v(t+s)) - f(v(t))] \text{ as } n \to \infty. \end{aligned} \tag{4.4.23}$$

Let δ_n be a sequence of positive real numbers satisfying $\delta_n \to 0$ and select an increasing sequence $\{a_\ell(n)\}$ such that $a(t_n + t) = a_1(n) < a_2(n) < \cdots \le a(t_n + t + s) - 1$, and that for $a(t_n + t) \le a_\ell \le a_{\ell+1} \le a(t_n + t + s) - 1$,

$$\frac{1}{\delta_n} \sum_{j=a_\ell(n)}^{a_{\ell+1}(n)-1} \frac{1}{j+1} \to 1 \text{ as } n \to \infty.$$

In what follows, for notational simplicity, we suppress the n dependence in $a_\ell(n)$ and write it as a_ℓ instead. Denote by I_a the index set satisfying $a(t_n+t) \le a_\ell \le a_{\ell+1} \le a(t_n+t+s)-1$.

Using the notation defined above, we have

$$\begin{aligned} f(v^n(t+s)) - f(v^n(t)) &= \sum_{\ell\in I_m} [f(v_{a_{\ell+1}}) - f(v_{a_\ell})] \\ &= \sum_{\ell\in I_m} \frac{df(v_{a_\ell})}{dv}\Big[\sum_{j=a_\ell}^{a_{\ell+1}-1}(v_{j+1}-v_j)\Big] \\ &\quad + \sum_{\ell\in I_m}\frac{1}{2}\frac{d^2f(v_{a_\ell})}{dv^2}\Big[\sum_{j=a_\ell}^{a_{\ell+1}-1}(v_{j+1}-v_j)\Big]^2 + o(1), \end{aligned} \tag{4.4.24}$$

where $o(1)\to 0$ in probability uniformly in t.

It follows that

$$\begin{aligned} &E\rho(v^n(t_i) : i\le\kappa)\sum_{\ell\in I_m}\frac{df(v_{a_\ell})}{dv}\Big[\sum_{j=a_\ell}^{a_{\ell+1}-1}(v_{j+1}-v_j)\Big] \\ &= E\rho(v^n(t_i) : i\le\kappa)\sum_{\ell\in I_m}\frac{df(v_{a_\ell})}{dv}E_{a_\ell}\Big[-\sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{j+1}(c_\sigma^2(\sigma_*)-\frac{1}{2})v_j \\ &\qquad + \sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{\sqrt{j+1}}(c_j-\overline{c})c_\sigma(\sigma_*) \\ &\qquad + \sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{j+1}(c_j-\overline{c})g(v_j) + \sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{j+1}h_jv_j\Big]. \end{aligned} \tag{4.4.25}$$

Then

$$\begin{aligned} &\lim_n E\rho(v^n(t_i) : i\le\kappa)\sum_{\ell\in I_m}\frac{df(v_{a_\ell})}{dv}\Big[-\sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{j+1}(c_\sigma^2(\sigma_*)-\frac{1}{2})v_j\Big] \\ &= \lim_n E\rho(v^n(t_i) : i\le\kappa)\sum_{\ell\in I_m}\frac{df(v_{a_\ell})}{dv}\Big[-\sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{j+1}(c_\sigma^2(\sigma_*)-\frac{1}{2})v_{a_\ell}\Big] \\ &= E\rho(v(t_i) : i\le\kappa)\Big[-\int_t^{t+s}\frac{df(v(\tau))}{d\tau}(c_\sigma^2(\sigma_*)-\frac{1}{2})v(\tau)d\tau\Big]. \end{aligned} \tag{4.4.26}$$

As for the next term, using the mixing property,

$$\begin{aligned} &\lim_n E\rho(v^n(t_i) : i\le\kappa)\sum_{\ell\in I_m}\frac{df(v_{a_\ell})}{dv}\Big[-\sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{\sqrt{j+1}}(c_j-\overline{c})c_\sigma(\sigma_*)\Big] \\ &= \lim_n E\rho(v^n(t_i) : i\le\kappa)\sum_{\ell\in I_m}\frac{df(v_{a_\ell})}{dv}\Big[-\sum_{j=a_\ell}^{a_{\ell+1}-1}\frac{1}{\sqrt{j+1}}E_{a_\ell}(c_j-\overline{c})c_\sigma(\sigma_*)\Big] = 0. \end{aligned} \tag{4.4.27}$$

Using the continuity of $g(\cdot)$, detailed calculation also shows that

$$\begin{aligned}
&\lim_n E\rho(v^n(t_i) : i \le \kappa) \sum_{\ell\in I_m} \frac{df(v_{a_\ell})}{dv}\Big[\sum_{j=a_\ell}^{a_{\ell+1}-1} \frac{1}{j+1}(c_j - \overline{c})g(v_j)\Big] = 0,\\
&\lim_n E\rho(v^n(t_i) : i \le \kappa) \sum_{\ell\in I_m} \frac{df(v_{a_\ell})}{dv}\Big[\sum_{j=a_\ell}^{a_{\ell+1}-1} \frac{1}{j+1}h_j v_j\Big] = 0.
\end{aligned} \tag{4.4.28}$$

Likewise, similar estimates lead to

$$\begin{aligned}
&\lim_n E\rho(v^n(t_i) : i \le \kappa) \sum_{\ell\in I_m} \frac{1}{2}\frac{d^2 f(v_{a_\ell})}{dv^2}\Big[\sum_{j=a_\ell}^{a_{\ell+1}-1} (v_{j+1} - v_j)\Big]^2\\
&\quad = E\rho(v(t_i) : i \le \kappa)\Big[\frac{1}{2}\int_t^{t+s} \frac{d^2 f(v(\tau))}{dv^2}\varsigma^2 c_\sigma^2(\sigma_*)d\tau\Big];
\end{aligned} \tag{4.4.29}$$

we omit the details for brevity. Thus the desired result follows. □

The above shows that the scaled sequence $v^n(\cdot)$ converges weakly to $v(\cdot)$, a diffusion process.

Remark 4.4.7. Similarly, we can also construct a stochastic approximation algorithm using constant step size

$$\sigma_{n+1} = \Pi\big[\sigma_n - \varepsilon[c(\sigma_n) - c_n]c_\sigma(\sigma_n)\big]. \tag{4.4.30}$$

Convergence of corresponding algorithm can also be obtained.

Chapter 5

Applications

5.1 Model Calibration

To use our results in practice, one needs to know the values for the parameters σ_1, σ_2, a, L, and $Q = (\lambda_{ij})$, respectively. In fact, the parameter involved in Q in our setup, is only one dimensional. Thus, for simplicity, we write it as λ without subscripts. One choice for σ_1 is the implied volatility using the Black-Scholes formula. For other parameters, it is a traditional method to use least squares. The drawback of this method is the heavy load on the corresponding computation. In this paper, we have developed stochastic approximation algorithms to estimate these parameters. Compared to the least squares estimation approach, the stochastic approximation method takes much less computational time and is more efficient.

First, we consider the least squares regression approach to estimate parameters in the mean reversion model. Differential equation (2.2.1) can be translated into the following equation in a discrete form:

$$X_{n+1} = L(1 - e^{-a\delta}) + X_n e^{-a\delta} + \sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}}u. \tag{5.1.1}$$

where δ is some fixed step in the simulation procedure, and u is a standard Gaussian random variable.

To simplify equation (5.1.1), define

$$\beta_0 = L(1 - e^{-a\delta}), \beta_1 = e^{-a\delta}, \varepsilon = \sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}}u. \tag{5.1.2}$$

Then, we get

$$X_{n+1} = \beta_0 + \beta_1 X_n + \varepsilon,$$

where the standard deviation of ε is

$$\varepsilon = \sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}}.$$

We regress the above linear model using the least squares method on the time series of stock prices to estimate β_0, β_1 and $sd(\varepsilon)$. Then based on the relationship between β_0, β_1 and $sd(\varepsilon)$ and a, L, σ_2 that we derived in (5.1.2), we can get the estimation of a, L, σ_2.

Next, we consider the stochastic approximation algorithms in Case I (from GBM to MR) and Case II (from MR to GBM) separately. Since typically, the call price is not sensitive to the parameter a, we shall only apply the stochastic approximation algorithms to estimates for σ_2, L, λ in Case I and σ_1, λ in Case II.

Case I, (GBM $\rightarrow$ MR). We have the following steps searching for σ_2, L, and λ.

(a) Given values of parameters with the initial state being in GBM, create a set of prices C_{ij}^{real} with strike prices K_i and maturity dates T_j for $1 \leq i \leq m$ and $1 \leq j \leq n$.

(b) Assuming values of σ_1, a, and λ are known, estimate σ_2 using stochastic approximation implied by a single-jump regime-switching model based on different choices of L.

(c) Compute the sum of squared errors $\sum_{i,j}(C_{ij} - C_{ij}^{\text{real}})^2$ and the mean squared errors $\frac{1}{mn}\sum_{i,j}(C_{ij} - C_{ij}^{\text{real}})^2$ between original prices and prices C_{ij} obtained using each pair of σ_2 and λ with m different strike prices K_i and n different maturity dates T_j, where $1 \leq i \leq m$ and $1 \leq j \leq n$.

Values of parameters used for creating the initial set of call prices are as follows:

$$r = 0.025, \sigma_1 = 0.35, \sigma_2 = 0.33, S = 44.5, a = 10, i = 1, L = 3.9, \lambda = 25,$$

$$K = \{37.5, 40, 42.5, 45, 47.5, 50, 52.5, 55\} \text{ and } T = \{15/252, 35/252, 95/252\}.$$

For the stochastic approximation method, start with $\sigma_2 = 0.35$.

L	2	3	3.8	3.85	3.89	3.9	3.91	4
σ_2	0.43594	0.5	0.5	0.48848	0.38165	0.3304	0.25771	0.05126
SSE	1.49854	4.24594	4.24594	3.62989	0.32564	0.000018	0.49936	3.82984
MSE	0.06244	0.17691	0.17691	0.15125	0.01357	0	0.02081	0.15958

Table 5.1: Sums of squared errors, mean squared errors and optimal estimates of σ_2 with different values of L

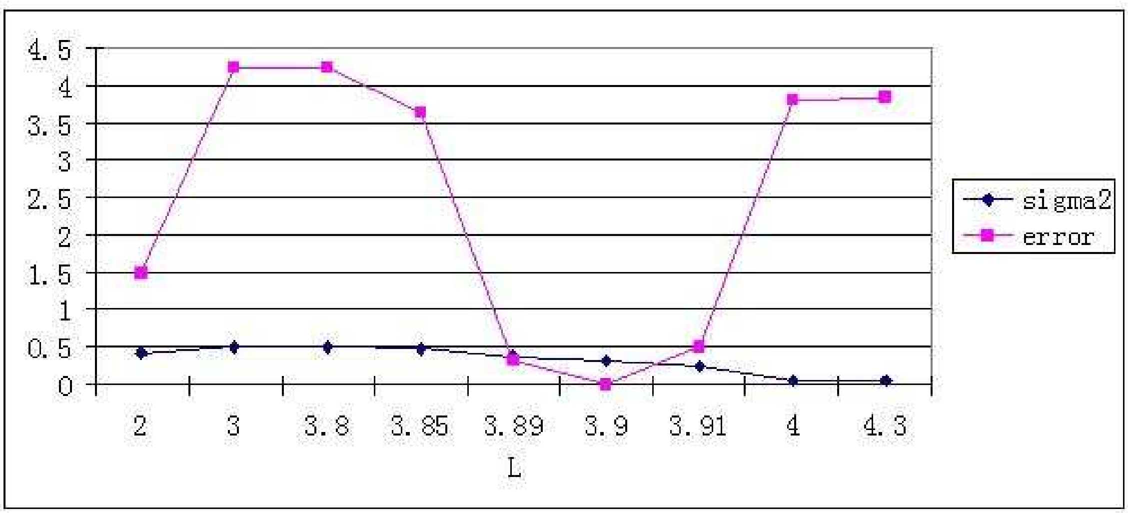

Figure 5.1: Sums of squared errors and optimal estimates of σ_2 with different L

Assuming values of σ_1, a, and L are known, estimate σ_2 using stochastic approximation implied by a single-jump regime-switching model based on different choices of λ, compute the sum of squared errors $\sum_{i,j}(C_{ij}-C_{ij}^{\text{real}})^2$ and the mean squared errors $\frac{1}{mn}\sum_{i,j}(C_{ij}-C_{ij}^{\text{real}})^2$ between original prices C_{ij}^{real} and prices C_{ij} obtained using each pair of σ_2 and λ with m different strike prices K_i and n different maturity dates T_j, where $1 \le i \le m$ and $1 \le j \le n$.

λ	1	2	5	10	25	50	100	252
σ_2	0.4296	0.39561	0.36625	0.35382	0.3304	0.31453	0.2978	0.27798
SSE	1.31145	0.53826	0.15616	0.06593	0.000018	0.02581	0.10805	0.27048
MSE	0.05464	0.02243	0.00651	0.00275	0	0.00108	0.00450	0.01127

Table 5.2: Sums of squared errors, mean squared errors and optimal estimates of σ_2 with different values of λ

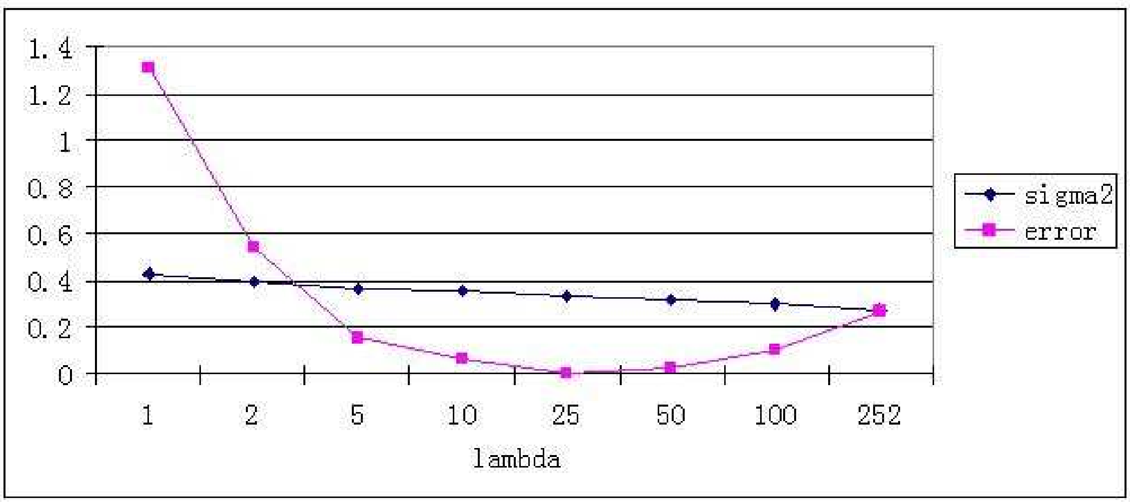

Figure 5.2: Sums of squared errors and optimal estimates of σ_2 with different λ

Case II, (MR $\rightarrow$ GBM).

(a) Given values of parameters with initial state in MR, create a set of prices C_{ij}^{real} with m different strike prices K_i and n different maturity dates T_j, where $1 \leq i \leq m$ and $1 \leq j \leq n$.

(b) Assuming values of σ_2, a, and L are known, estimate σ_1 using stochastic approximation algorithms with a single-jump regime-switching model based on different choices of λ.

(c) Compute the sum of squared errors $\sum_{i,j}(C_{ij} - C_{ij}^{\text{real}})^2$ and the mean squared errors $\frac{1}{mn}\sum_{i,j}(C_{ij} - C_{ij}^{\text{real}})^2$ between original prices C_{ij}^{real} and prices C_{ij} obtained using each pair of σ_1 and λ with m different strike prices K_i and n different maturity dates T_j.

Values of parameters used for creating the initial set of call prices are as follows:

$$r = 0.025, \sigma_1 = 0.35, \sigma_2 = 0.33, S = 44.5, a = 10, i = 2, L = 3.81, \lambda = 25,$$

$$K = \{37.5, 40, 42.5, 45, 47.5, 50, 52.5, 55\} \text{ and } T = \{15/252, 35/252, 95/252\}.$$

To apply the stochastic approximation method, start with the initial estimate $\sigma_1 = 0.33$.

λ	1	2	5	10	25	50	100	252
σ_1	0.5	0.5	0.42716	0.46665	0.35	0.05	0.05	0.05
SSE	17.6593	17.6593	4.45807	10.4684	1.07E-10	32.7053	32.7053	32.7053
MSE	0.73580	0.73580	0.18575	0.43618	0	1.36272	1.36272	1.36272

Table 5.3: Sums of squared errors, mean squared errors and optimal estimates of σ_1 with different values of λ

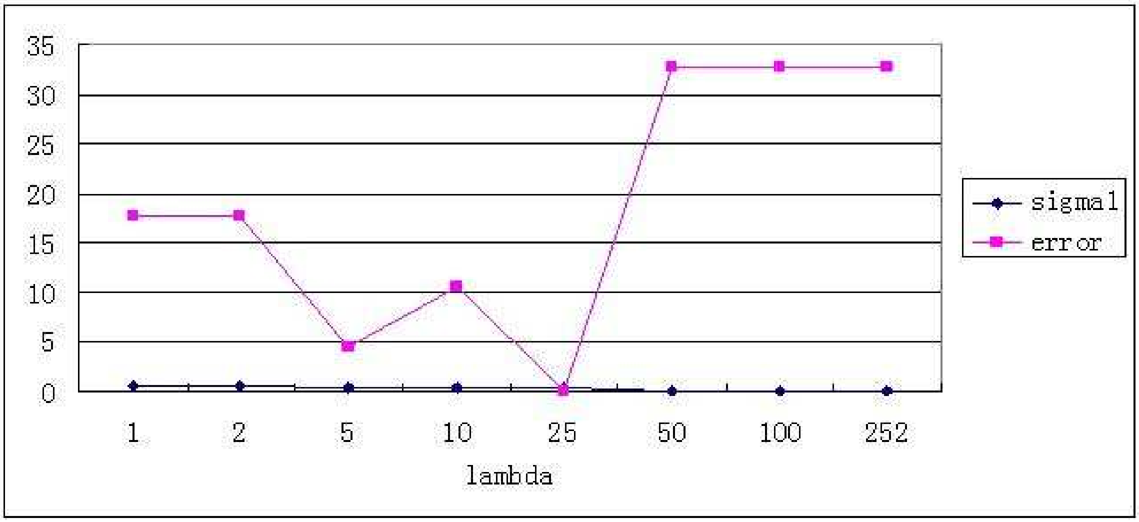

Figure 5.3: Sums of squared errors and optimal estimates of σ_1 with different λ

As can be seen that when we use different values of one of the parameters, the results of squared errors in all above tables show that the one that gives the smallest error is that with the value closest to the actual initial value of that parameter. This suggests stochastic approximation leads to the value which best approximates the actual value. The same conclusion applies for all parameters in the regime-switching model.

5.2 Prediction of Stock Trend

Typically, an option market is more sensitive than its underlying market. Because of the high leverage of the option market, it reacts to major news faster than the underlying market.

In this section, we use market option data to explore how the option market predicts the future underlying market mode (i.e., GBM vs. MR). In particular, we use option and stock market data to estimate the jump rate λ whose reciprocal gives the mode switching time.

One way to do this is to minimize the sum of squared differences $\sum_{i,j}(C_{ij} - C_{ij}^{\text{real}})^2$ between real call prices C_{ij}^{real} and derived call prices C_{ij} over different choices of λ with m different strike prices K_i and n different maturity dates T_j, where $1 \leq i \leq m$ and $1 \leq j \leq n$. Next, we work with two examples using the stocks of Gilead Sciences Inc. (GILD) and Intel Corp. (INTC).

For the GILD example, we use the stock prices from January 28, 2008 to March 5, 2008 to estimate σ_1, L, and a. The estimates for these parameters are obtained as follows: $\sigma_1 = 0.358393$, $L = 3.9$ and $a = 50$. The GILD stock price (daily close) during January 28 to March 5, 2008 and several following months, to May 30, 2008, are given in Figure (5.4).

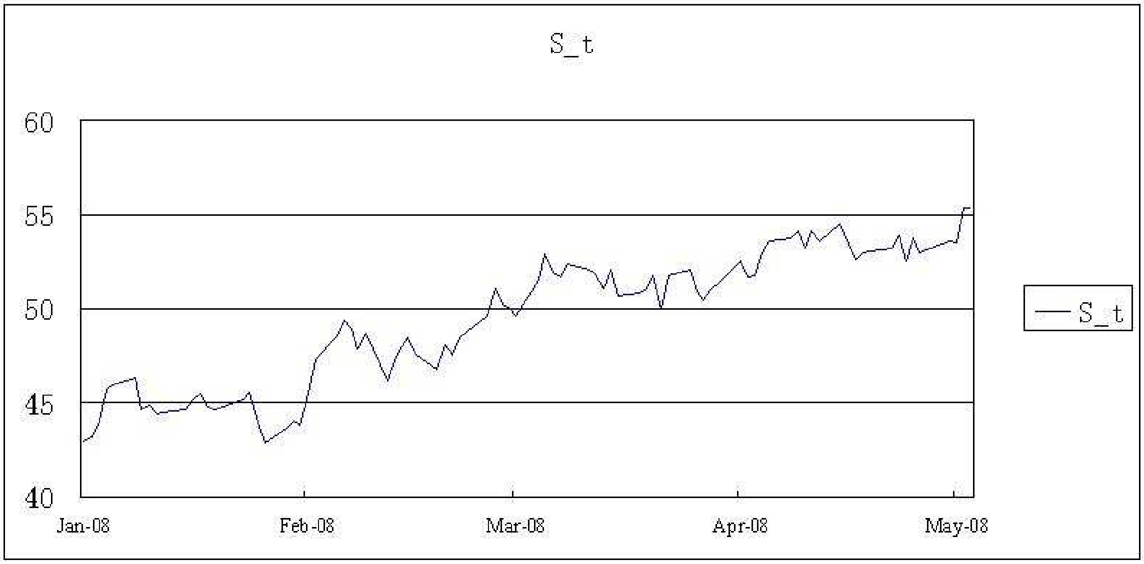

Figure 5.4: GILD: Jan. 28, 2008 to May 30, 2008

The stock price given in Figure (5.4) is divided into two parts: Jan 28 – Mar 5 and Mar 6 – May 30, 2008. It is convincing that the first part follows a GBM. We are to use the option price to predict the market mode (i.e., GBM vs. MR) after the middle point. We follow the steps used in Case I for estimating (λ, σ_2) and obtain the following squared errors:

λ	1	5	10	50
σ_2	0.31041	0.1521	0.1013	0.21127
SSE	25.2268	217.069	541.067	109.342
MSE	1.05112	9.04454	22.54446	4.55592

Table 5.4: Sums of squared errors, mean squared errors and optimal estimates of σ_2 with different values of λ

The results given in Table (5.4) indicate that the minimum is reached at $\lambda = 1$. This means that the mode switch will take on average of a year to occur. Therefore, one should not expect immediate mode switch over to a GBM. This is confirmed by the second part (Mar 6 – May 30) of the stock price, which also follows a GBM.

To test this result, we pick 3 dates after March 5, 2008: March 11, March 25 and March 26, 2008 and collect the real call prices. We use the Black-Scholes formula and the mean reversion formula separately to calculate call prices C_1 and C_2, then we compare those values with the real call prices. The squared errors are given in the following table.

$date$	$March11$	$March25$	$March26$
$err(C_1, C_{real})$	10.5351	2.53407	1.82575
$err(C_2, C_{real})$	27.457	56.7593	22.1717

Table 5.5: Comparison between GBM and mean reversion of GILD

The result in Table (5.5) supports our conclusion that the call prices obtained from Black-Scholes formula are closer to real prices than those under the mean reversion model. This means the state of this stock is very likely to remain in GBM in the next month.

Finally, we consider the second example with the daily close prices of INTC stock from January 28, 2008 to May 30, 2008 (see Figure (5.5)). In this case, the stock prices can be divided into two parts: Jan 28 – Mar 5 and Mar 6 – May 30, 2008. It appears that the first part follows a mean reversion model and the second part a GBM. We use the stock prices of the first part and linear regression to obtain $\sigma_2 = 0.358947964$, $L = 3.007900326$, and $a = 178.4244988$.

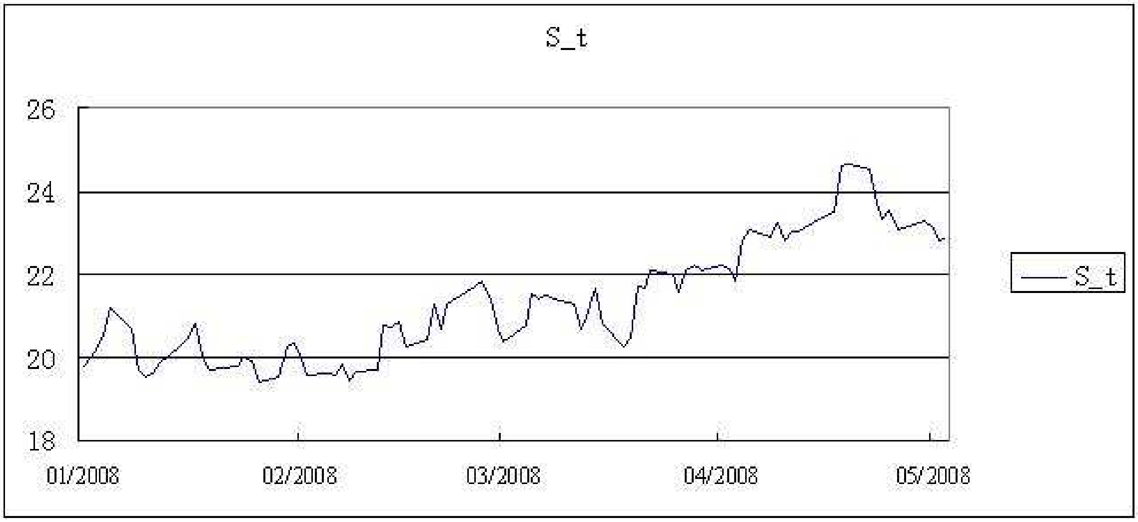

Figure 5.5: INTC: Jan. 28, 2008 to May 30, 2008

We follow the steps used in Case II for estimating (λ, σ_1) and obtain the following squared errors:

λ	10	20	100	200
σ_1	0.49976	0.41669	0.3459	0.34714
SSE	3.32802	2.43916	1.74524	1.40951
MSE	0.13867	0.10163	0.07272	0.05873

Table 5.6: Sums of squared errors, mean squared errors and optimal estimates of σ_1 with different values of λ

The results in Table (5.6) suggest that the best (associated with the smallest error) $\lambda = 200$. This means the average time needed to switch to the GBM mode is 1/200$\sim$ one day. This can be confirmed by Figure (5.5).

To test this result, we pick 2 dates: March 7, 2008 and March 10, 2008 and collect the real call prices. We use the Black-Scholes formula and the mean reversion formula separately to calculate the call prices C_1 and C_2, then we calculate the sum errors with the real call prices as follows.

$date$	$March7$	$March10$
$err(C_1, C_{real})$	0.197826	0.30096
$err(C_2, C_{real})$	1.74052	2.25789

Table 5.7: Comparison between GBM and mean reversion of INTC

This result also supports our conclusion that the smaller squared error with call prices obtained from the Black-Scholes formula suggests that the stock market is very likely to switch from MR to GBM right after March 6, 2007.

Chapter 6

Perpetual American Put Option Pricing

European call options are equivalent to American call options with no dividends during the life of the option. However, early exercise of American put options is sometimes profitable in the absence of dividends. Therefore, it is necessary to consider the pricing problem of American put options. There are some studies about perpetual American options. For example: Guo and Zhang's paper ([20]) provides an explicit optimal stopping rule and the corresponding value function in a closed form using the modified smooth fit technique. Levendorskii's paper ([25]) provides calculation of optimal exercise boundaries and rational prices for perpetual American call and put options. Boyarchenko and Levendorskii's paper ([8]) provides the solution to the pricing problem for perpetual American options in Markov-modulated Levy models.

In this chapter we focus on the American put option pricing problem under the mean reversion model, especially the perpetual put option for which we suppose the expiration date goes to infinity. The problem is to maximize

$$E^x[e^{-\rho t}(e^{X_t} - K)^+],$$

over possible stopping time τ, where ρ is the discount factor, K is the strike price, and x is the initial log-value of X_t.

Let $v(t, \xi)$ be the value function defined by

$$v(t, \xi) = e^{-\rho t}(e^{\xi} - K)^+.$$

Then the characteristic operator is.

$$\hat{A}v(s,x) = \frac{\partial v}{\partial t} + a(L-x)\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2}.$$

The problem turns out to be solving the value problem:

$$\frac{\partial v}{\partial t} + a(L-x)\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2} = 0. \tag{6.0.1}$$

We start with the form $e^{-\rho s}\phi(x)$ for $v(s,x)$; then the free boundary problem can be interpreted by

$$\frac{1}{2}\sigma^2\phi^{''}(x) + a(L-x)\phi^{'}(x) - \rho\phi(x) = 0, \tag{6.0.2}$$

with

$$\phi(x^*) = (K - e^{x^*})^+,$$

$$\phi'(x^*) = -e^{x^*}.$$

Define $y = L - x$ and $\psi(y) = \phi(L-y) = \phi(x)$, we have

$$\psi^{'}(y) = -\phi^{'}(L-y) = \phi^{'}(x),$$

$$\psi^{''}(y) = \phi^{''}(L-y) = \phi^{''}(x).$$

Equation (6.0.2) can be converted into

$$\frac{1}{2}\sigma^2\psi''(y) - ay\psi^{'}(y) - \rho\psi(y) = 0. \tag{6.0.3}$$

In order to solve equation (6.0.3), we set $z = \frac{\sqrt{2a}}{\sigma}y = \kappa y$; then the equation can be simplified as

$$f^{''}(z) - zf^{'}(z) - \lambda f(z) = 0,$$

where $\lambda = \frac{\rho}{a}, \kappa = \frac{\sqrt{2a}}{\sigma}$.

We use the transform $f(z) = e^{\frac{z^2}{4}}D(z)$. Therefore, $D(z)$ satisfies

$$D_{zz}(z) + [\frac{1}{2} - \frac{z^2}{4} - \lambda]D(z) = 0.$$

Two independent solutions to the above differential equation are

$$D(z) = \frac{1}{\Gamma(\lambda)} e^{\frac{-z^2}{4}} \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}-zt)} dt,$$

and

$$D(-z) = \frac{1}{\Gamma(\lambda)} e^{\frac{-z^2}{4}} \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+zt)} dt.$$

Then, we substitute these results to get the solution of $f(z)$:

$$f(z) = C_1 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}-zt)} dt + C_2 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+zt)} dt.$$

Then, we transform the variables backwards to get

$$\psi(y) = C_1 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}-\kappa yt)} dt + C_2 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+\kappa yt)} dt,$$

and

$$\phi(x) = C_1 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}-\kappa(L-x)t)} dt + C_2 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+\kappa(L-x)t)} dt.$$

Next, we show above solution for $\phi(x)$ indeed satisfies

$$-\rho\phi(x) + a(L-x)\phi^{'}(x) + \frac{1}{2}\sigma^2\phi^{''}(x) = 0.$$

Since $\phi(x)$ is bounded, then

$$e^{(-\frac{t^2}{2}-\kappa(L-x)t)} \to \infty, \text{ as } x \to \infty.$$

C_1 must be 0. Then We are left with

$$\phi(x) = C_2 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+\kappa(L-x)t)} dt.$$

Taking the derivative, we have

$$\phi^{'}(x) = C_2 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+\kappa(L-x)t)} (-\kappa t) dt,$$

$$\phi^{''}(x) = C_2 \int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2}+\kappa(L-x)t)} (\kappa^2 t^2) dt.$$

Reversing the sign in (6.0.2), we obtain

$$\begin{aligned}
&\rho\phi(x) - a(L-x)\phi^{'}(x) - \frac{1}{2}\sigma^2\phi^{''}(x)\\
&= \rho C_2 \int_0^\infty t^{\lambda-1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt - a(L-x)C_2\int_0^\infty t^{\lambda-1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}(-\kappa t)dt\\
&- \frac{1}{2}\sigma^2 C_2 \int_0^\infty t^{\lambda-1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}(\kappa^2 t^2)dt\\
&= C_2 a[\int_0^\infty e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt^\lambda + \int_0^\infty \kappa(L-x)t^\lambda e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt\\
&- \int_0^\infty t^{\lambda+1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt]\\
&= C_2 a[0 - \int_0^\infty t^\lambda de^{(-\frac{t^2}{2}+\kappa(L-x)t)} + \int_0^\infty \kappa(L-x)t^\lambda e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt\\
&- \int_0^\infty t^{\lambda+1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt]\\
&= C_2 a[\int_0^\infty t^{\lambda+1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt - \int_0^\infty \kappa(L-x)t^\lambda e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt\\
&+ \int_0^\infty \kappa(L-x)t^\lambda e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt - \int_0^\infty t^{\lambda+1}e^{(-\frac{t^2}{2}+\kappa(L-x)t)}dt] = 0.
\end{aligned}$$

This completes the proof. □

We substitute the solution for $\phi(x)$ into

$$\phi(x^*) = (K - e^{x^*})^+,$$

and

$$\phi'(x^*) = -e^{x^*}.$$

We have

$$C_2 \int_0^\infty t^{\lambda-1}e^{(-\frac{t^2}{2}+\kappa(L-x^*)t)}dt = (K - e^{x^*})^+ = e^{x^*} - K,$$

$$C_2 \int_0^\infty t^{\lambda-1}e^{(-\frac{t^2}{2}+\kappa(L-x^*)t)}(-\kappa t)dt = -e^{x^*}.$$

Dividing the first equation by the second one, we get

$$e^{x^*}\int_0^\infty t^{\lambda-1}e^{(-\frac{t^2}{2}+\kappa(L-x^*)t)}dt = \kappa(e^{x^*} - K)\int_0^\infty t^\lambda e^{(-\frac{t^2}{2}+\kappa(L-x^*)t)}dt.$$

This is the equation for the threshold x^*, and

$$C_2 = \frac{e^{x^*} - K}{\int_0^\infty t^{\lambda-1} e^{(-\frac{t^2}{2} + \kappa(L - x^*)t)} dt}.$$

In Levendorskii's working paper ([25]), the same problems have been studied. We found this working paper when we finished this section in Fall 2006. There are some minor differences between that study and ours. First of all, the rational value function in that paper is assumed indifferent to time. Secondly, the representations of the parabolic cylinder functions are either a series or integral, the integral representation is provided in this chapter, and the series representation is provided in that paper.

Bibliography

[1] R. Balvers, Y. Wu and E. Gilliland, *Mean Reversion across National Stock Markets and Parametric Constrain Investment Strategies*, Journal of Finance, 55(2) (2000), 745-772.

[2] L. Bauwens, A. Preminger and J. Rombouts., *Regime Switching Garch Models*, CORE Discussion Papers 2006011, Universit catholique de Louvain, Center for Operations Research and Econometrics (CORE) (2006).

[3] P. Billingsley, *Convergence of Probability Measures*, J. Wiley, New York, (1968).

[4] F. Black, M. Scholes, *The Pricing of Options and Corporate Liabilities*, The Journal of Political Economy, Vol. 81, No. 3, pp. 637-654, (1973).

[5] C. Blanco and D. Soronow, *Mean reverting processes – Energy price processes used for derivatives pricing and risk management*, Commodities Now, pp. 68-72, (June 2001).

[6] N.P.B. Bollen, *Valuing options in regime-switching models*, Journal of Derivatives, vol. 6, pp. 38-49, (1998).

[7] L.P. Bos, A.F. Ware and B.S. Pavlov, *On a semi-spectral method for pricing an option on a mean-reverting asset*, Quantitative Finance, Vol. 2, pp. 337-345, (2002).

[8] S. Boyarchenko, S. Levendorskii *Perpetual American options in regime-switching models* working paper series, (2007).

[9] J. Buffington and R.J. Elliott, *American options with regime switching*, International Journal of Theoretical and Applied Finance, vol. 5, pp. 497-514, (2002).

[10] A. Cowles and H. Jones, *Some posteriori probabilities in stock market action*, Econometrica, Vol. 5, pp. 280-294, (1937).

[11] M.G. Crandall, H. Ishii, and P.L. Lions, *Users guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., 27, (1992), 1-67.

[12] J.-C. Duan, I. Popova and P. Ritchken, *Option pricing under regime switching*, Quantitative Finance, vol. 2, pp.116-132, (2002).

[13] C. de Jong and R. Huisman, *Option formulas for mean-reverting power prices with spikes*, preprint.

[14] R.J. Elliott and P. E. Kopp, *Mathematics of Financial Markets*, Springer-Verlag, New York, (1998).

[15] S.N. Ethier and T.G. Kurtz, *Markov Process: Characterization and Convergence*, John Wiely, New York, (1986).

[16] E. Fama and K.R. French, *Permanent and temporary components of stock prices*, Journal of Political Economy, Vol. 96, pp. 246-273, (1988).

[17] L.A. Gallagher and M.P. Taylor, *Permanent and temporary components of stock prices: Evidence from assessing macroeconomic shocks*, Southern Economic Journal, Vol 69, pp. 345-362, (2002).

[18] X. Guo, *An explicit solution to an optimal stopping problem with regime switching*, J. App. Prob, vol. 38, pp. 464–481, (2001).

[19] X. Guo and Q. Zhang, *Optimal selling rules in a regime switching model*, IEEE Transactions on Automatic Control, vol. 50, pp. 1450-1455, (2005).

[20] X. Guo and Q. Zhang, *Closed-form solutions for perpetual American put options with regime switching*, SIAM journal on applied mathematics, Vol. 64, pp. 2034-2049 (2004).

[21] C.M. Hafner and H. Herwartz, *Option pricing under linear autoregressive dynamics, heteroskedasticity, and conditional leptokurtosis*, Journal of Empirical Finance, Vol. 8, pp. 1-34, (2001).

[22] M.R. Hardy, *A regime-switching model of long-term stock returns*, North American Actuarial Journal, vol. 5, pp. 41-53, (2001).

[23] J.D. Hamilton, *A new approach to the economic analysis of non-stationary time series*, Econometrica, vol.57, pp. 357-384, (1989).

[24] J.C. Hull, *Options, Futures, and Other Derivatives*, 3rd Ed., Prentice-Hall, Upper Saddle River, NJ, (1997).

[25] S. Levendorskii, *Perpetual American options and real options under mean-reverting processes*, working paper series.

[26] Lo, Andrew W., and A. Craig MacKinlay, *Stock market prices do not follow random walks: Evidence from a simple specification test*, Review of Financial Studies 1,(1988) 4166.

[27] J. Poterba and L. Summers, *Mean Reversion in Stock Prices: Evidence and Implications*, Journal of Financial Economics, 22 (1988), 27-59.

[28] H.J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, (1984).

[29] H.J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd Ed., Springer, New York, (2003).

[30] B. OKsendal, *Stochastic Differential Equations*, Springer-Verlag, New York, (2000).

[31] Q.S. Song, G. Yin, and Q. Zhang, *Stochastic optimization methods for buying-low-and-selling-high strategies*, Stochastic Analysis and Applications, (in press).

[32] D.D. Yao, Q. Zhang and X.Y. Zhou, *A regime-switching model for European options*, Stochastic Processes, Optimization, and Control Theory Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems, H.M. Yan, G. Yin, and Q. Zhang Eds., Vol. 94, Springer, New York, (2006), 281-300.

[33] G. Yin, *On extensions of Polyak's averaging approach to stochastic approximation*, Stochastics Stochastics Rep., **36** (1991), 245-264.

[34] G. Yin, R.H. Liu, and Q. Zhang, *Recursive algorithms for stock liquidation: A stochastic optimization approach*, SIAM J. Optim., **13**, (2002), 240-263.

[35] G. Yin, J.W. Wang, Q. Zhang, and Y.J. Liu, *Stochastic optimization algorithms for pricing American put options under regime-switching models*, J. Optim. Theory Appl., **131** (2006), 37-52.

[36] G. Yin, Q. Zhang, and C. Zhuang, *Stochastic approximation algorithms for parameter estimation in option pricing with regime switching*, Stochastic Analysis and Applications, **25** (2007), 1243-1261.

[37] Q. Zhang, *Stock trading: An optimal selling rule*, SIAM J. Control. Optim., **40**, (2001), 64-87.

[38] H. Zhang and Q. Zhang, *Trading a mean-reverting asset: Buy low and sell high*, Automatica, Vol. 44, pp. 1511-1518, (2008).

Appendix A

change of time method: Theorem 8.5.2 of [30]

Let $dY_t = v(t,\omega)dB_t$, $v \in R^{n\times m}$, $B_t \in R^m$ be an Ito integral in R^n, $Y_0 = 0$ and assume that

$$vv^T(t,\omega) = c(t,\omega)I_n,$$

for some process $c(t,\omega) \geq 0$. Let α_t, β_t be as in (8.5.1), (8.5.2). Then Y_{α_t} is an n-dimensional Brownian motion.

(8.5.1): $\beta_t = \beta(t,\omega) = \int_0^t c(s,\omega)ds.$

(8.5.2): $\alpha_t = inf\{s; \beta_s > t\}.$

Appendix B

equicontinuous in the extended sense: p. 102 of [29]

Suppose that for each n, $f_n(\cdot)$ is an R^r-valued measurable function on $(-\infty, \infty)$ and $\{f_n(0)\}$ is bounded. Also suppose that for each T and $\epsilon > 0$, there is a $\delta > 0$ such that

$$\limsup_{n} \sup_{0 \leq t-s \leq \delta, |t| \leq T} |f_n(t) - f_n(s)| \leq \epsilon.$$

Then we say that $\{f_n(\cdot)\}$ is equicontinuous in the extended sense.

Appendix C

Theorem 6.1.1 of [29]

Assume (5.1.1), algorithm (1.1), and the conditions (A1.1)-(A1.7), with H stisfying any one of the conditions (A4.3.1), (A4.3.2), or (A4.3.3). If the ϵ_n are random, assume (A1.8) in lieu of (5.1.1). Then there is a null set N such that for ω not in N, the set of functions $\{\theta^n(\omega,\cdot), Z^n(\omega,\cdot), n < \infty\}$ is equicontinuous. Let $(\theta(\omega,\cdot), Z(\omega,\cdot))$ denote the limit of some convergent subsequence. Then this pair satisfies the projected ODE (5.2.1), and $\{\theta_n(\omega)\}$ converges to some limit set of the ODE in H. If the constraint set is dropped, but $\{\theta_n\}$ is bounded with probability one, then for almost all ω, the limits $\theta(\omega,\cdot)$ of convergent subsequences of $\{\theta^n(\omega,\cdot)\}$ are trajectories of

$$\dot{\theta} = \bar{g}(\theta)$$

in some bounded invariant set and $\{\theta_n(\omega)\}$ converges to this invariant set. Let p_n be integer-valued functions of ω, not necessarily being stopping times or even measurable, but that go to infinity with probability one. Then the conclusions concerning the limits of $\{\theta^n(\cdot)\}$ hold with p_n replacing n. If $A \subset H$ is locally asymptotically stable in the sense of Liapunov for (5.2.1) and θ_n is in some compact set in the domain of attraction of A infinitely often with probability $\geq \rho$, then $\theta_n \to A$ with at least probability ρ. Suppose that (A5.2.6) holds. Then, for almost all ω, $\{\theta_n(\omega)\}$ converges to a unique S_i. Under the additional conditions of Theorem (5.2.5) on the mean ODE, for almost all ω, $\theta^n(\omega,\cdot)$ and $\theta_n(\omega)$ converge to a set of chain recurrent points in the limit or invariant set.

(5.1.1): $\sum_{n=0}^{\infty} \epsilon_n = \infty, \epsilon_n \geq 0, \epsilon_n \to 0$, for $n \geq 0$; $\epsilon_n = 0$, for $n < 0$.

(5.2.1): $\dot{\theta} = \bar{g}(\theta) + z$, $z \in -C(\theta)$.

Algorithm (1.1): $\theta_{n+1} = \prod_H[\theta_n + \epsilon_n Y_n]$.

(A1.1): $sup_n E|Y_n| < \infty$.

(A1.2): $g_n(\theta,\xi)$ is continuous in θ for each ξ and n.

(A1.3): There is a continuous function $\bar{g}(\cdot)$ such that for each θ, $\mu > 0$ and some $T > 0$,

$$\lim_{n\to\infty} P\left\{\sup_{j\geq n}\max_{0\leq t\leq T}\left|\sum_{i=m(jT)}^{m(jT+t)-1}\epsilon_i[g_i(\theta,\xi_i)-\bar{g}(\theta)]\right|\geq\mu\right\}=0.$$

(A1.4):

$$\lim_{n\to\infty} P\left\{\sup_{j\geq n}\max_{0\leq t\leq T}\left|\sum_{i=m(jT)}^{m(jT+t)-1}\epsilon_i\delta M_i\right|\geq\mu\right\}=0.$$

(A1.5):

$$\lim_{n\to\infty} P\left\{\sup_{j\geq n}\max_{0\leq t\leq T}\left|\sum_{i=m(jT)}^{m(jT+t)-1}\epsilon_i\beta_i\right|\geq\mu\right\}=0.$$

(A1.6): There are non-negative measurable functions $\rho_3(\cdot)$ and $\rho_{n4}(\cdot)$ of θ and ξ, respectively, such that $|g_n(\theta,\xi)|\leq\rho_3(\theta)\rho_{n4}(\xi)$, where $\rho_3(\cdot)$ is bounded on each bounded θ-set, and for each $\mu > 0$,

$$\lim_{\tau\to 0}\lim_{n\to\infty} P\left\{\sup_{j\geq n}\sum_{i=m(j\tau)}^{m(j\tau+\tau)-1}\epsilon_i\rho_{i4}(\xi_i)\geq\mu\right\}=0.$$

(A1.7): There are non-negative measurable functions $\rho_1(\cdot)$ and $\rho_{n2}(\cdot)$ of θ and ξ, respectively, such that $\rho_1(\cdot)$ is bounded on each bounded θ-set and $|g_n(\theta,\xi)-g_n(y,\xi)|\leq\rho_1(\theta-y)\rho_{n2}(\xi)$, where $\rho_1(\theta)\to 0$, as $\theta\to 0$ and for some $\tau > 0$,

$$P\left\{\limsup_{j}\sum_{i=j}^{m(t_j+\tau)}\epsilon_i\rho_{i2}(\xi_i)<\infty\right\}=1.$$

(A1.8): If ϵ_n is random, let it be F_n-measurable, with $\epsilon_n\geq 0$ and $\Sigma_n\epsilon_n=\infty$ with probability one.

(A4.3.1): H is a hyperrectangle. In other words, there are real numbers $a_i < b_i$, $i = 1, ..., r$, such that $H = \{x : a_i \leq x \leq b_i\}$.

(A4.3.2): Let $q_i(\cdot)$, $i = 1, ..., p$, be continuously differentiable real-valued functions on R^r, with gradients $q_{i,x}(\cdot)$. Without loss of generality, let $q_{i,x}(x) \neq 0$ if $q_i(x) = 0$. Define $H = \{x : q_i(x) \leq 0, i = 1, ..., p\}$. Then H is connected, compact and nonempty.

(A4.3.3): H is an R^{r-1} dimensional connected compact surface with a continuously differentiable outer normal. In this case, define $C(x), x \in H$, to be the linear span of the outer normal at x.

(A5.2.6): $\bar{g}(\cdot) = -f_\theta(\cdot)$ for continuously differentiable real-valued $f(\cdot)$ and $f(\cdot)$ is constant on each S_i.