2016-06-01

American Spread Option Pricing with Stochastic Interest Rate

An Jiang
Brigham Young University

American Spread Option Pricing with Stochastic Interest Rates

An Jiang

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Kening Lu, Chair
John Dallon
Kenneth Kuttler
Lennard Bakker
Tiancheng Ouyang

Department of Mathematics

Brigham Young University

June 2016

ABSTRACT

American Spread Option Pricing with Stochastic Interest Rates

An Jiang
Department of Mathematics, BYU
Doctor of Philosophy

In financial markets, spread option is a derivative security with two underlying assets and the payoff of the spread option depends on the difference of these assets. We consider American style spread option which allows the owners to exercise it at any time before the maturity. The complexity of pricing American spread option is that the boundary of the corresponding partial differential equation which determines the option price is unknown and the model for the underlying assets is two-dimensional.

In this dissertation, we incorporate the stochasticity to the interest rate and assume that it satisfies the Vasicek model or the CIR model. We derive the partial differential equations with terminal and boundary conditions which determine the American spread option with stochastic interest rate and formulate the associated free boundary problem. We convert the free boundary problem to the linear complimentarity conditions for the American spread option, so that we can go around the free boundary and compute the option price numerically. Alternatively, we approximate the option price using methods based on the Monte Carlo simulation, including the regression-based method, the Lonstaff and Schwartz method and the dual method. We make the comparisons among the option prices derived by the partial differential equation method and Monte Carlo methods to show the accuracy of the result.

Keywords: American Spread Option, Stochastic Interest Rate, Free Boundary Problem, Linear Complementarity Conditions, CIR Model, Monte Carlo Simulation

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Dr. Kening Lu. I could not have completed this dissertation without his support and guidance. Thanks to Dr. Lu, I had the opportunity to study such an interesting problem in financial mathematics. I also would like to thank my graduate committee members, friends and colleagues who encouraged me and gave me generous advice about my dissertation. I want to thank my parents. Although I studied abroad and rarely went home during my PhD program at Brigham Young University, my parents always supported and encouraged me to focus on my studies and let me have no worries about my family in China. Lastly, the person I would like to thank most is my dear wife. With her company and love, I get through the most difficult times.

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Option pricing can be originated from the beginning of the 20th century when the Brownian motion was introduced to describe financial behaviors. This drew people's attention from deterministic functions to random functions. The biggest question is to seek a counterpart of the classical calculus for random functions. The situation was resolved by Kiyosi Itô who introduced the stochastic integral and the celebrated Itô's formula. After that, Itô calculus embraced its most successful application. In 1973, Fischer Black, Myron Scholes [3] and Robert Merton [33] derived the notable Black-Scholes-Merton formula from their model. Their key idea is the so called delta hedging: buying and selling the underlying assets of the option in order to eliminate the chance of arbitrage. Their outstanding outcome inspired mathematicians to look at other financial derivative securities with more complicated structures, one of them is spread option.

1.1 Spread Option

An option is a contract which gives its owner the right but not the obligation to buy (call option) or sell (put option) the underlying assets at a specified price (strike price), on a specified date (European option) or any time before that date (American option).

An spread option is an option with two underlying assets which profits from their price difference. It is a very popular type of option in the currency markets, fixed income markets, foreign exchange markets, agricultural futures markets, commodity markets and energy markets etc. For example, the NOB spread measures the difference between municipal bonds and treasury bonds; the TED spread measures the difference between treasury bills and treasury bonds in the U.S. fixed income market. Although most of the spread options are traded over the counter, crush spread is traded on the Chicago Board of Trade (CBOT). It has three underlying indexes: future contracts of soybean, soybean oil and soybean meal. The payoff is the weighted difference between the soybean and its products. Spread option is widely used in the energy markets. For a single commodity, there are temporal spread and locational spread as a result of the time factor or transportation factor which makes the commodity price differ. In the case of two or more commodities, spread options measure the difference between the raw material and the refined product it produces, like crack spread for crude oil and refined petroleum and spark spread for natural gas and electricity.

Due to the diversity of the assets or indexes, various methods of analysis and numerical algorithms are applied in mathematical literature to study spread options. For example, Deng, Johnson and Sogomonian [14] studied two types of spread options (spark spread option and locational spread option) in the energy market. Johnson, Zulauf, Irwin and Gerlow [26] studied the crush spread in agricultural markets. Jones [27] studied TED spread in the fixed income markets. Carmona and Durrleman [8] provided a complete overview of the spread option category and early research history.

In 2000, Dempster and Hong [13] applied the Fast Fourier Transform(FFT) to study the spread option with a one-factor stochastic volatility model. Later in 2001, Hong [21] used the same FFT technique to study the spread option if the two underlying assets have stochastic correlation. Since then FFT has been employed to numerical computation.

1.2 American Spread Option

American option is a main type of option in the market, unlike the European option, it gives the owner the right to exercise the option at any time before the expiration date/maturity. It makes itself more flexible for investment than the European option, but more difficult for valuation. It requires us to solve an optimal stopping time problem, i.e.,

$$V(t,x) = \max_{\tau \in \mathcal{T}_{[t,T]}} \widetilde{E}[e^{-\int_t^\tau r(u)du}(K - S(\tau))|S(t) = x].$$

Here $V(t,x)$ is the price of the American option at time t and the underlying asset price is $S(t) = x$, r is the interest rate, τ is a stopping time, K is the strike price and the expectation is taken under the risk-neutral measure.

For a non-dividend paying underlying asset, it can be shown that the American call option price is the same as the European call option price. So we need only to consider the American put option. To understand its characteristic, it is better to consider the American perpetual put option first where there is no maturity ($T = \infty$), then consider the finite-expiration American put option. In [36], Shreve gives a thorough explanation.

For American option with multiple underlying assets, especially the American spread option, the main obstacle is the unknown early exercise boundary. Most of the techniques applied on the European spread option can't be transferred to the American spread option to obtain closed form solutions. Therefore, extensive amount of research has been done to numerically estimate the American spread option price.

In 1997, Broadie and Glasserman [7] introduced a stochastic mesh method based on Monte Carlo simulation. Later, Tsitsiklis and Van Roy [37, 38] proposed a dynamical programming algorithm. The idea is to use a regression formulation employing the least square method and basis functions to approximate the continuous value of the American option. Longstaff and Schwartz [31] revised their method using their interleaving value to give a low-biased estimator for the option price. During 2001 and 2002, Haugh, Kogan [20] and Rogers [35] independently and almost simultaneously developed dual methods which can give a tighter upper bound for the American option price. Together with the lower bound given by Longstaff and Schwartz method, it can provide an interval for the true option price value. All these methods are based on Monte Carlo simulation, they work well on multidimensional underlying assets, although they can only give a rough approximation.

In 2008, Jackson, Jaimungal and Surkov [25] applied the Fast Fourier Transform algorithm to price the American spread option with the underlying assets influenced by Levy processes. In 2011, Chiarella and Ziveyi [11] applied the numerical integration technique under the Black-Scholes framework. Later that year, they [10] applied the method of lines to price American spread option under the Heston stochastic volatility model.

1.3 The Objectives of the Dissertation

In this dissertation, we will incorporate stochasticity into the interest rate for the American spread option. The most basic stochastic interest model is the well-known Vasicek model. Luo used it in pricing the European spread option in [32]. However, there is a major drawback of the Vasicek model that the interest rate $r(t)$ may have a chance to go negative, which generally will not happen in the financial market. To compensate for this weakness, we employ the CIR model as well.

The option pricing strategy will be divided into two parts: partial differential equation approach and Monte Carlo method. In the first part, we will review the fundamental option pricing theory including the Itô calculus, risk-neutral pricing method and delta-hedging strategy, then we will apply them to the partial differential equation for the American spread option under Vasicek and CIR interest rate models. We will deal with the free boundary problem and the linear complementarity conditions and use finite difference method to get numerical approximation. In the second part, we will use different numerical methods based on the Monte Carlo simulation to approximate the option price. The algorithms are applied under both Vasicek and CIR models. Comparisons will be made among different interest rate models and also between the partial differential equation approach and the Monte Carlo method. Finally, we will conclude this dissertation with directions for future research.

Chapter 2. Stochastic Calculus

In this chapter, we will review some basic knowledge in stochastic calculus. Firstly, we will survey the basic probability concepts and Itô integral. Then we will review the Itô's formula. We will also discuss a conditional expectation identity for different measures. Moreover, we will present Girsanov Theorem and Martingale Representation Theorem. Finally, we conclude this chapter by discussing the Markov property for the solutions to stochastic differential equations.

The presentation of this chapter is inspired by [30], [36] and [34].

2.1 Itô Integral

Firstly, we give the definition for probability space, random variable and stochastic process.

Definition 2.1. A probability space is a triple $(\Omega, \mathcal{F}, P)$ consisting a nonempty sample space Ω, a σ-algebra $\mathcal{F}$ and a probability measure P.

Definition 2.2. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and Σ be a topological space with Borel σ-algebra $\mathcal{G}$. A random variable $X : \Omega \to \Sigma$ is a $\mathcal{F}$-measurable function, i.e., for any Borel subset $B \in \mathcal{G}$,

$$X^{-1}(B) = \{\omega : X(\omega) \in B,\ B \in \mathcal{G}\} \tag{2.1}$$

is in the σ-algebra $\mathcal{F}$.

For our purposes, we let the topological space Σ to be the Euclidean space $\mathbb{R}^n$, where n is a positive integer.

Definition 2.3. A stochastic process is a collection of t-parameterized random variables

$$X(t, \omega) : [0, \infty) \times \Omega \to \mathbb{R}^n.$$

For the sake of simple notation, we omit the variable ω when we write stochastic processes, i.e., $X(t) = X(t, \omega)$.

An important type of stochastic process is Brownian motion.

Definition 2.4. A stochastic process $W(t, \omega)$ is called a Brownian motion if it satisfies the following conditions:

(i) $P\{\omega : W(0, \omega) = 0\} = 1$.

(ii) For any $0 \leq s < t$, the random variable $W(t) - W(s) \sim N(0, t - s)$, i.e., normally distributed with mean 0 and variance $t - s$.

(iii) $W(t, \omega)$ has independent increments, i.e., for any $0 \leq t_1 < t_2 < \cdots < t_n$, the random variables

$$W(t_1), W(t_2) - W(t_1), ..., W(t_n) - W(t_{n-1})$$

are independent.

(iv) Almost all sample paths of $W(t, \omega)$ are continuous functions, i.e.,

$$P\{\omega : W(\cdot, \omega) \text{ is continuous}\} = 1.$$

Definition 2.5. A multidimensional stochastic process $W(t, \omega) = (W_1(t, \omega), ..., W_n(t, \omega))$ is a multidimensional Brownian motion if $W_1(t, \omega), ..., W_n(t, \omega)$ are independent and each $W_i(t, \omega),\ i = 1, ..., n$ is a Brownian motion.

Secondly, we review filtration and stopping time.

Definition 2.6. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A filtration on $(\Omega, \mathcal{F}, P)$ is a family of σ-algebras $\{\mathcal{F}(t),\ t \geq 0\}$, such that for any fixed t, $\mathcal{F}(t) \subset \mathcal{F}$ and if $0 \leq s \leq t$, $\mathcal{F}(s) \subset \mathcal{F}(t)$.

Definition 2.7. A stochastic process $X(t)$ is said to be adapted to a filtration $\{\mathcal{F}(t),\ t \geq 0\}$ if for any t, $X(t)$ is $\mathcal{F}(t)$-measurable.

Definition 2.8. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $W(t), t \geq 0$ be a Brownian motion. A filtration for the Brownian motion is a filtration $\{\mathcal{F}(t), t \geq 0\}$, on the probability space, satisfying:

(1) The Brownian motion $W(t), t \geq 0$ is adapted to the filtration $\{\mathcal{F}(t), t \geq 0\}$.

(2) For $0 \leq t < u$, the increment $W(u) - W(t)$ is independent of $\mathcal{F}(t)$.

A filtration for a Brownian motion contains two possibilities. In one case, the filtration is generated by a Brownian motion. In the other case, the filtration is generated by a Brownian motion and one or more other processes.

Definition 2.9. Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\{\mathcal{F}(t),\ t \geq 0\}$. A random variable $\tau : \Omega \to I$ is a stopping time if for any $t \in I$,

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}(t),$$

We denote the set of the stopping times by $\mathcal{T}_I$.

Next we review the Itô integral. Three steps are needed to define the Itô integral.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $W(t), a \leq t \leq b$, be a Brownian motion and $\{\mathcal{F}(t), a \leq t \leq b\}$ be a filtration for this Brownian motion. Let $L^2_{ad}([a,b] \times \Omega)$ denote the space of all adapted stochastic processes $f(t,\omega), a \leq t \leq b, \omega \in \Omega$, satisfying $\int_a^b E(|f(t)|^2)dt < \infty$.

Step 1: Firstly, we define the Itô integral for step stochastic process. Let $\{a = t_0, t_1, ..., t_n = b\}$ be a partition of the interval $[a,b]$. A step stochastic process in $L^2_{ad}([a,b] \times \Omega)$ is given by

$$f_n(t,\omega) = \sum_{i=1}^{n} \xi_{i-1}(\omega) 1_{[t_{i-1},t_i]}(t), \tag{2.2}$$

where ξ_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable and $E(\xi_{i-1}^2) < \infty$. The Itô integral for the step stochastic process in $L^2_{ad}([a,b] \times \Omega)$ is defined by

$$\int_a^b f_n(t)dW(t) := \sum_{i=1}^{n} \xi_{i-1}(W(t_i) - W(t_{i-1})). \tag{2.3}$$

Step 2: Let $f \in L^2_{ad}([a,b] \times \Omega)$. Then there exists a sequence $\{f_n, n \geq 1\}$ of step stochastic processes in $L^2_{ad}([a,b] \times \Omega)$ such that

$$\lim_{n\to\infty} \int_a^b E\{|f(t) - f_n(t)|^2\} dt = 0. \tag{2.4}$$

Step 3: Finally, we can define the Itô integral for $f \in L^2_{ad}([a,b] \times \Omega)$. Let $\{f_n, n \geq 1\}$ be the step stochastic processes in $L^2_{ad}([a,b] \times \Omega)$ in the previous step. Then

$$\int_a^b f(t) dW(t) := \lim_{n\to\infty} \int_a^b f_n(t) dW(t), \text{ in } L^2(\Omega). \tag{2.5}$$

We will see an important property of the Itô integral.

Definition 2.10. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{\mathcal{F}(t),\ 0 \leq t \leq T\}$, be a filtration on the probability space. An adapted stochastic process $M(t),\ 0 \leq t \leq T$, is called a *martingale* associated with the filtration $\{\mathcal{F}(t), 0 \leq t \leq T\}$, if

$$E[M(t)|\mathcal{F}(s)] = M(s) \tag{2.6}$$

for all $0 \leq s \leq t \leq T$.

Theorem 2.11. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $W(t), 0 \leq t \leq T$ be a Brownian motion and let $\{\mathcal{F}(t),\ 0 \leq t \leq T\}$, be a filtration on the probability space. Let $f \in L^2_{ad}([0,T] \times \Omega)$. The stochastic process defined by the Itô integral*

$$\int_a^t f(s) dW(s), \quad 0 \leq t \leq T, \tag{2.7}$$

is a martingale.

2.2 Itô's Formula

We are ready to show the celebrated Itô's formula which is used for computing the derivatives of composed functions involving stochastic processes.

Definition 2.12. Let $(\Omega, \mathcal{F}, P)$ be a probability space associated with a filtration $\{\mathcal{F}(t), a \leq t \leq b\}$. An Itô process is a stochastic process of the form

$$X(t) = X(a) + \int_a^t g(s)ds + \int_a^t f(s)dW(s), \quad a \leq t \leq b, \tag{2.8}$$

where $X(a)$ is $\mathcal{F}$-measurable, $f \in L^2_{ad}([a,b] \times \Omega)$ and g is an adapted process and $\int_a^b |g(t)|dt < \infty$ almost surely.

Theorem 2.13 (One-dimensional Itô's formula)**.** *Let $X(t)$ be an Itô process given by* (2.8)*. Suppose that $\theta(t,x)$ is a continuous function with continuous partial derivatives $\frac{\partial \theta}{\partial t}$, $\frac{\partial \theta}{\partial x}$ and $\frac{\partial^2 \theta}{\partial x^2}$. Then $\theta(t, X(t))$ is also an Itô process and*

$$\theta(t, X(t)) = \theta(a, X(a)) + \int_a^t \left[\frac{\partial \theta}{\partial t}(s, X(s)) + \frac{\partial \theta}{\partial x}(s, X(s))g(s) + \frac{1}{2}\frac{\partial^2 \theta}{\partial x^2}(s, X(s))f^2(s)\right] ds + \int_a^t \frac{\partial \theta}{\partial x}(s, X(s))f(s)dW(s). \tag{2.9}$$

Equivalently, the differential form is

$$d\theta(t, X(t)) = \frac{\partial \theta}{\partial t}(t, X(t))dt + \frac{\partial \theta}{\partial x}(t, X(t))g(t)dt + \frac{1}{2}\frac{\partial^2 \theta}{\partial x^2}(t, X(t))f^2(t)dt + \frac{\partial \theta}{\partial x}(t, X(t))f(t)dW(t). \tag{2.10}$$

In particular, if we let $f(t) \equiv 1$, $g(t) \equiv 0$ and $X(a) = 0$ then the Itô process $X(t)$ is just the Brownian motion $W(t)$ and we have the following two simple forms of Itô's formula.

Corollary 2.14 (The simplest form of Itô's formula)**.** *Let $f(x)$ be a C^2-function. Then*

$$f(W(t)) = f(W(a)) + \frac{1}{2}\int_a^t f''(W(s))ds + \int_a^t f'(W(s))dW(s). \tag{2.11}$$

Equivalently, the differential form is

$$df(W(t)) = \frac{1}{2}f''(W(t))dt + f'(W(t))dW(t). \tag{2.12}$$

Corollary 2.15 (Slightly gereralized Itô's formula). *Let $f(x)$ be a continuous function with continuous partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$. Then*

$$f(t, W(t)) = f(t, W(a)) + \int_a^t \left(\frac{\partial f}{\partial t}(s, W(s)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, W(s))\right) ds + \int_a^t \frac{\partial f}{\partial x}(s, W(s))dW(s). \tag{2.13}$$

Equivalently, the differential form is

$$df(t, W(t)) = \frac{\partial f}{\partial t}(t, W(t))dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, W(t))dt + \frac{\partial f}{\partial x}(t, W(t))dW(t). \tag{2.14}$$

Definition 2.16 (Multi-dimensional Itô process). Let $(\Omega, \mathcal{F}, P)$ be a probability space associated with a filtration $\{\mathcal{F}(t), a \leq t \leq b\}$, and let $W_1(t), ..., W_m(t)$ be m independent Brownian motions. Consider n Itô processes $X_t^{(1)}, ..., X_t^{(n)}$ given by

$$X_t^{(i)} = X_a^{(i)} + \int_a^t g_i(s)ds + \sum_{j=1}^{m}\int_a^t f_{ij}(s)dW_j(s), \quad 1 \leq i \leq n, \tag{2.15}$$

where $f_{ij} \in L^2_{ad}([a, b] \times \Omega)$ and g_i is an adapted process with $\int_a^b |g(t)|dt < \infty$ almost surely for all $1 \leq i \leq n$ and $1 \leq j \leq m$. If we introduce the matrices

$$W(t) = \begin{bmatrix} W_1(t) \\ \vdots \\ W_m(t) \end{bmatrix}, \quad X_t = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix},$$

$$f(t) = \begin{bmatrix} f_{11}(t) & \dots & f_{1m}(t) \\ \vdots & \ddots & \vdots \\ f_{n1}(t) & \dots & f_{nm}(t) \end{bmatrix}, \quad g(t) = \begin{bmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{bmatrix},$$

then the Itô processes can be written as a matrix equation:

$$X_t = X_a + \int_a^t g(s) + \int_a^t f(s)dW(s). \tag{2.16}$$

Theorem 2.17 (Multi-dimensional Itô's formula)**.** *Suppose that $\theta(t, x_1, ..., x_n)$ is a continuous function on $[a, b] \times \mathbb{R}^n$ with continuous partial derivatives $\frac{\partial \theta}{\partial t}$, $\frac{\partial \theta}{\partial x_i}$ and $\frac{\partial^2 \theta}{\partial x_i \partial x_j}$ for $1 \leq i, j \leq n$. Then the stochastic differential of $\theta(t, X_t^{(1)}, ..., X_t^{(n)})$ is given by*

$$\begin{aligned} d\theta(t, X_t^{(1)}, ..., X_t^{(n)}) = \frac{\partial \theta}{\partial t}(t, X_t^{(1)}, ..., X_t^{(n)})dt + \sum_{i=1}^{n} \frac{\partial \theta}{\partial x_i}(t, X_t^{(1)}, ..., X_t^{(n)})dX_t^{(i)} \\ + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \theta}{\partial x_i x_j}(t, X_t^{(1)}, ..., X_t^{(n)})dX_t^{(i)}dX_t^{(j)}. \end{aligned} \tag{2.17}$$

In particular, the two-dimensional Itô's formula is given by

$$d\theta(t, X, Y) = \theta_t dt + \theta_x dX + \theta_y dY + \frac{1}{2}\theta_{xx} dX dX + \theta_{xy} dX dY + \frac{1}{2}\theta_{yy} dY dY. \tag{2.18}$$

Corollary 2.18 (Itô product rule)**.** *Let $X(t)$ and $Y(t)$ be two Itô processes. Then*

$$d\left(X(t)Y(t)\right) = Y(t)dX(t) + X(t)dY(t) + dX(t)dY(t). \tag{2.19}$$

2.3 Expectation Identity under Two Different Measures

Since it is very hard to compute the American spread option under the risk-neutral measure, it is necessary to change it to a different measure. Therefore we need to know how to convert the conditional expectation under a different measure.

Definition 2.19 (Radon-Nikodym derivative)**.** Give a measurable space (X, Σ), if a σ-finite measure ν on (X, Σ) is absolutely continuous with respect to a σ-finite measure μ on (X, Σ), then there is a measurable function $f : X \to [0, \infty)$, such that for any measurable subset

$A \subset X$:

$$\nu(A) = \int_A f(x) d\mu. \tag{2.20}$$

where $f := \frac{d\nu}{d\mu}$ is the Radon-Nikodym derivative.

In the context of the probability theory, we have

Definition 2.20. In $(\Omega, \mathcal{F})$, a probability space, given two equivalent probability measures $P, \widetilde{P}$, i.e., $P(A) = 0$ if and only if $\widetilde{P}(A) = 0$ for every $A \in \mathcal{F}$. Then denote $Z = \frac{d\widetilde{P}}{dP}$ and

$$\widetilde{P}(A) = \int_A Z(\omega) dP(\omega) \text{ for all } A \in \mathcal{F}. \tag{2.21}$$

Proposition 2.21. *$Z(\omega)$ has the following properties:*

(i) $P(Z > 0) = 1$,

(ii) $EZ = 1$,

(iii) for any random variable X,

$$\widetilde{E}X = E[ZX], \tag{2.22}$$

where $\widetilde{E}$ is the expectation under measure $\widetilde{P}$.

Definition 2.22. Given a filtration $\{\mathcal{F}(t), 0 \leq t \leq T\}$ and in the above setting, the Radon-Nikodym derivative process is defined as

$$Z(t) = E[Z|\mathcal{F}(t)], \ 0 \leq t \leq T. \tag{2.23}$$

Note that it is a martingale because of iterated conditioning: for $0 \leq s \leq t \leq T$

$$\begin{aligned} E[Z(t)|\mathcal{F}(s)] &= E[E[Z|\mathcal{F}(t)]|\mathcal{F}(s)] \\ &= E[Z|\mathcal{F}(s)] \\ &= Z(s). \end{aligned}$$

Lemma 2.23. *Let Y be an $\mathcal{F}(t)$-measurable random variable, $0 \le t \le T$. Then*

$$\widetilde{E}Y = E[YZ(t)]. \tag{2.24}$$

This is the expectation identity for two measures P and $\widetilde{P}$.

Lemma 2.24. *Let s and t satisfying $0 \le s \le t \le T$ and let Y be an $\mathcal{F}(t)$-measurable random variable. Then*

$$\widetilde{E}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)} E[YZ(t)|\mathcal{F}(s)]. \tag{2.25}$$

This is the conditional expectation identity for two measures P and $\widetilde{P}$. We will use this result later when we price the American spread option under the forward measure.

2.4 Girsanov Theorem and Martingale Representation

Girsanov Theorem is a powerful tool to find a new Brownian motion if we need to change the measure.

Theorem 2.25 (Girsanov Theorem for one-dimensional Brownian motion)**.** *([36, Theorem 5.2.3, p.212]). Let $(\Omega, \mathcal{F}, P)$ be a probability space, $W(t), 0 \le t \le T$, be a Brownian motion, $\{\mathcal{F}(t), 0 \le t \le T\}$, be a filtration for this Brownian motion, and $\Theta(t), 0 \le t \le T$, be an adapted process. Assume that*

$$E\int_0^t \Theta^2(u)du < \infty, \quad E\int_0^T \Theta^2(u)Z^2(u)du < \infty,$$

where $Z(t)$ is an exponential process which is defined by

$$Z(t) = \exp\left\{-\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du\right\}. \tag{2.26}$$

Set the random variable $Z := Z(T)$. Then $EZ = 1$ and the process $\widetilde{W}(t), 0 \le t \le T$, defined *by*

$$\widetilde{W}(t) := W(t) + \int_0^t \Theta(u)du \tag{2.27}$$

is a Brownian motion under the probability measure $\widetilde{P}$ *given by*

$$\widetilde{P}(A) := \int_A Z(\omega)dP(\omega), \quad \forall A \in \mathcal{F}. \tag{2.28}$$

Theorem 2.26 (Girsanov Theorem for multi-dimensional Brownian motion)**.** *([36, Theorem 5.4.1, p.224]). Let* $(\Omega, \mathcal{F}, P)$ *be a probability space,* $W(t) = (W_1(t), ..., W_n(t)), 0 \le t \le T$*, be an* n*-dimensional Brownian motion where* n *is a positive integer,* $\{\mathcal{F}(t), 0 \le t \le T\}$*, be a filtration for this Brownian motion, and* $\Theta(t) = (\Theta_1(t), ..., \Theta_n(t)), 0 \le t \le T$*, be an* n*-dimensional adapted process. Assume that*

$$E\int_0^t ||\Theta(u)||^2 du < \infty, \quad E\int_0^T ||\Theta(u)||^2 Z^2(u)du < \infty,$$

where $Z(t)$ *is an exponential process which is defined by*

$$Z(t) = \exp\left\{-\int_0^t \Theta(u) \cdot dW(u) - \frac{1}{2}\int_0^t ||\Theta(u)||^2 du\right\}. \tag{2.29}$$

Set the random variable $Z := Z(T)$*. Then* $EZ = 1$ *and the process* $\widetilde{W}(t), 0 \le t \le T$*, defined by*

$$\widetilde{W}(t) := W(t) + \int_0^t \Theta(u)du \tag{2.30}$$

is a n*-dimensional Brownian motion under the probability measure* $\widetilde{P}$ *given by*

$$\widetilde{P}(A) := \int_A Z(\omega)dP(\omega), \quad \forall A \in \mathcal{F}. \tag{2.31}$$

In (2.29), the Itô integral is

$$\int_0^t \Theta(u) \cdot dW(u) = \int_0^t \sum_{i=1}^n \Theta_i(u)dW_i(u) = \sum_{i=1}^n \int_0^t \Theta_i(u)dW_i(u).$$

And $||\Theta(u)||$ denotes the Euclidean norm

$$||\Theta(u)|| = \left(\sum_{i=1}^{n} \Theta_i^2(u)\right)^{\frac{1}{2}}.$$

Martingale Representation Theorem will help us to derive the partial differential equation for the American spread option later.

Theorem 2.27 (Martingale Representation Theorem for one-dimensional Brownian motion)**.** *([36, Theorem 5.3.1, p.221]) Let $(\Omega, \mathcal{F}, P)$ be a probability space, T be a fixed positive time, $W(t), 0 \leq t \leq T$ be a Brownian motion, and let $\{\mathcal{F}(t), 0 \leq t \leq T\}$ be the filtration generated by this Brownian motion. Let $M(t), 0 \leq t \leq T$ be a martingale with respect to this filtration. Then there is an adapted process $\Gamma(u), 0 \leq u \leq T$ such that*

$$M(t) = M(0) + \int_0^t \Gamma(u) dW(u), \quad 0 \leq t \leq T. \tag{2.32}$$

Theorem 2.28 (Martingale Representation Theorem for multi-dimensional Brownian motion)**.** *([36, Theorem 5.4.2, p.225]) Let $(\Omega, \mathcal{F}, P)$ be a probability space, T be a fixed positive time, $W(t) = (W_1(t), ..., W_n(t)), 0 \leq t \leq T$ be an n-dimensional Brownian motion with a positive integer n, and let $\{\mathcal{F}(t), 0 \leq t \leq T\}$ be the filtration generated by this Brownian motion. Let $M(t)$ be a martingale with respect to this filtration. Then there is an n-dimensional adapted process $\Gamma(u), 0 \leq t \leq T$ such that*

$$M(t) = M(0) + \int_0^t \Gamma(u) \cdot dW(u), \quad 0 \leq t \leq T. \tag{2.33}$$

Remark that the differential form of the martingale representation is

$$dM(t) = \Gamma(t) \cdot dW(t), \quad 0 \leq t \leq T. \tag{2.34}$$

On the other hand, if a stochastic process is given by (2.33). Then $M(t)$ is a martingale because Itô integral is a martingale. Therefore the stochastic process $M(t)$ is a martingale if and only if the dt term in the differential form of $M(t)$ is equal to 0.

2.5 Markov Property

We conclude this chapter by introducing the Markov property for the solutions to the stochastic differential equations.

Definition 2.29. Let $(\Omega, \mathcal{F}, P)$ be a probability space, let T be a fixed positive number, and let $\{\mathcal{F}(t), 0 \leq t \leq T\}$, be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $X(t), \quad 0 \leq t \leq T$. Assume that for all $0 \leq s \leq t \leq T$ and for every nonnegative , Borel-measurable function f, there is another Borel-measurable function g such that

$$E[f(X(t))|\mathcal{F}(s)] = g(X(s)). \tag{2.35}$$

Then we say that the X is a Markov process.

Theorem 2.30. *Let $X(u), u \geq 0$, be a solution to the stochastic differential equation*

$$dX(u) = \beta(u, X(u))du + \gamma(u, X(u))dW(u), \tag{2.36}$$

with initial condition given at time 0. Then for $0 \leq t \leq T$ and any Borel-measurable function h, there exists a Borel-measurable function $g(t,x)$ such that

$$E[h(X(T))|\mathcal{F}(t)] = g(t, X(t)). \tag{2.37}$$

Corollary 2.31. *Solutions to stochastic differential equations are Markov processes.*

Chapter 3. Risk-Neutral Pricing

In this chapter, we will present the risk-neutral measure and the risk-neutral pricing formula. We will also review the fundamental theorems of asset pricing which provide necessary and sufficient conditions for a financial market to satisfy two assumptions.

3.1 Stock Representation under the Risk-Neutral Measure

Consider the probability space $(\Omega, \mathcal{F}, P)$, and a Brownian motion $W(t), 0 \leq t \leq T$. Further, let $\{\mathcal{F}(t), 0 \leq t \leq T\}$ be a filtration for $W(t)$. Assume that a stock price is a generalized geometric Brownian motion determined by the following stochastic differential equation.

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t), 0 \leq t \leq T. \tag{3.1}$$

Here $\alpha(t)$ is called the mean rate of return of the stock and $\sigma(t)$ is called the volatility of the stock and they are adapted processes. Also assume $\sigma(t)$ is almost surely not zero for all $t \in [0, T]$. The equation above can be solved explicitly to get

$$S(t) = S(0) \exp\left\{\int_0^t \sigma(s)dW(s) + \int_0^t (\alpha(s) - \frac{1}{2}\sigma^2(s))ds\right\}. \tag{3.2}$$

In fact, define $f(x) = \ln(x)$, then by Itô's formula

$$\begin{aligned} df(S(t)) = d(\ln(S(t))) &= \frac{1}{S(t)}dS(t) + \frac{1}{2}\left(-\frac{1}{S^2(t)}\right)dS^2(t) \\ &= \alpha(t)dt + \sigma(t)dW(t) - \frac{1}{2}\sigma^2(t)dt \\ &= \sigma(t)dW(t) + (\alpha(t) - \frac{1}{2}\sigma^2(t))dt. \end{aligned}$$

So

$$\ln(S(t)) = \ln(S(0)) + \int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right)ds. \tag{3.3}$$

Thus,

$$S(t) = S(0)\exp\left\{\int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right) ds\right\}.$$

Definition 3.1. Suppose that $r(t)$ is an adapted interest rate process. The discount process is defined by

$$D(t) := e^{-\int_0^t r(s)ds}. \tag{3.4}$$

Its reciprocal is the money market account

$$M(t) := \frac{1}{D(t)} = e^{\int_0^t r(s)ds}. \tag{3.5}$$

The differential form of the discount process is

$$dD(t) = -r(t)D(t)dt. \tag{3.6}$$

In fact, let $I(t) = \int_0^t r(s)ds$, so $dI(t) = r(t)dt, \quad dIdI = 0$, let $f(x) = e^{-x}, f'(x) = -f(x), f''(x) = f(x)$.

By Itô's formula,

$$\begin{aligned} dD(t) &= df(I(t)) \\ &= f'(I(t))dI(t) + \frac{1}{2}f''(I(t))dI(t)dI(t) \\ &= -f(I(t))r(t)dt \\ &= -r(t)D(t)dt. \end{aligned}$$

Multiplying the stock price $S(t)$ with the discount process $D(t)$, we get the discounted stock price:

$$D(t)S(t) = S(0)\exp\left\{\int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - r(s) - \frac{1}{2}\sigma^2(s)\right) ds\right\}. \tag{3.7}$$

The discounted stock price satisfies the following stochastic differential equation.

$$\begin{aligned} d(D(t)S(t)) &= (\alpha(t) - r(t))D(t)S(t)dt + \sigma(t)D(t)S(t)dW(t) \\ &= \sigma(t)D(t)S(t)[\Theta(t)dt + dW(t)]. \end{aligned}$$

where we let

$$\Theta(t) := \frac{\alpha(t) - r(t)}{\sigma(t)}. \tag{3.8}$$

Applying Girsanov Theorem and letting

$$d\widetilde{W}(t) = \Theta(t)dt + dW(t), \tag{3.9}$$

we have

$$d(D(t)S(t)) = \sigma(t)D(t)S(t)d\widetilde{W}(t). \tag{3.10}$$

Let

$$\widetilde{P}(A) := \int_A Z(\omega)dP(\omega), \quad \forall A \in \mathcal{F}, \tag{3.11}$$

where

$$Z = \exp\left\{-\int_0^T \Theta(u)dW(u) - \frac{1}{2}\int_0^T \Theta^2(u)du\right\}.$$

as defined in the Girsanov Theorem (Theorem (2.25)).

Definition 3.2. A probability measure $\widetilde{P}$ is called a *risk-neutral measure* if

- Measure $\widetilde{P}$ is equivalent to the actual measure P (i.e., $\forall A \in \mathcal{F}$, $P(A) = 0$ if and only if $\widetilde{P}(A) = 0$), and
- under $\widetilde{P}$, the discounted stock price $D(t)S(t)$ is a martingale .

Therefore the probability measure $\widetilde{P}$ defined in (3.11) is a risk-neutral measure.

Under the risk-neutral measure $\widetilde{P}$,

$$dS(t) = r(t)S(t)dt + \sigma(t)S(t)d\widetilde{W}(t). \tag{3.12}$$

Note that the mean rate of return is equal to the interest rate $r(t)$, and the stochastic differential equation can be solved explicitly to get

$$S(t) = S(0)\exp\left\{\int_0^t \sigma(s)d\widetilde{W}(s) + \int_0^t (r(s) - \frac{1}{2}\sigma^2(s))ds\right\}. \tag{3.13}$$

An inverse problem is, given an asset $S(t)$, is there a stochastic differential equation in the form of (3.12) to which $S(t)$ is a solution? The answer is yes and provided by the following theorem.

Theorem 3.3 (Stochastic representation of assets)**.** *([36, Theorem 9.2.1, p.377]) Let* $(\Omega, \mathcal{F}, \widetilde{P})$ *be a probability space with a Brownian motion* $\widetilde{W}(t), t \geq 0$ *under the risk-neutral measure* $\widetilde{P}$*. Let* $S(t)$ *be a strictly positive price process for a non-dividend-paying asset, then there exists a volatility process* $\sigma(t)$ *such that*

$$\frac{dS(t)}{S(t)} = r(t)dt + \sigma(t)d\widetilde{W}(t). \tag{3.14}$$

This equation is equivalent to

$$\frac{d(D(t)S(t))}{D(t)S(t)} = \sigma(t)d\widetilde{W}(t). \tag{3.15}$$

or

$$\frac{D(t)S(t)}{S(0)} = \exp\left\{\int_0^t \sigma(u)d\widetilde{W}(t) - \frac{1}{2}\int_0^t ||\sigma(u)||^2 du\right\}. \tag{3.16}$$

3.2 Hedging

The fundamental idea behind no-arbitrage pricing is the hedging strategy (or replicating strategy), introduced by Black and Scholes in their celebrated article [3]. That is to reproduce the payoff of a derivative security by trading in the underlying asset (stock) and the money market account.

Definition 3.4. The portfolio is a pair of processes $\phi(t)$ and $\psi(t)$ which describe respectively the number of units of stock and of money market account which we hold at time t. The processes can take positive or negative values.

Definition 3.5. If $(\phi(t), \psi(t))$ is a portfolio with stock price $S(t)$ and money market account price $M(t)$, then $(\phi(t), \psi(t))$ is self-financing if and only if

$$dX(t) = \phi(t)dS(t) + \psi(t)dM(t), \tag{3.17}$$

where $X(t) = \phi(t)S(t) + \psi(t)M(t)$ is the value of the portfolio.

In other words, a portfolio is self-financing if and only if the change in its value only depends on the change of the asset prices.

Definition 3.6. Suppose that we have a risk-less money market account $M(t)$ and a risky security $S(t)$ with volatility $\sigma(t)$, and a payoff $V(T)$ up to time T. A hedging strategy (or replicating strategy) for $V(T)$ is a self-financing portfolio $(\phi(t), \psi(t))$ such that

$$X(T) = \phi(T)S(T) + \psi(T)M(T) = V(T).$$

The reason why we care about the hedging strategy is that not only can it replicate the payoff of the option at the expiration date but also the value of the hedging portfolio is equal to the value of the option at any time before the expiration date, i.e.,

$$X(T) = V(T) \Leftrightarrow X(t) = V(t), \quad \forall\ 0 \leq t \leq T. \tag{3.18}$$

To see this, we assume there is a moment $t < T$, such that $X(t) > V(t)$. Then an investor can buy one unit of the option at time t and sell one unit of the portfolio including $\Delta_1(t)$ units of $S(t)$ and $\Delta_2(t)$ units of $M(t)$. Because it is guaranteed that $X(T) = V(T)$ at the expiration date, the option bought and the portfolio sold by the investor will cancel with each other at time T. Thus the investor could make the profit $V(t) - X(t)$ without any risk. This contradicts with our basic assumption that the financial market is arbitrage-free. On the other hand, if $X(t) < V(t)$, the investor can buy one unit of the portfolio and sell one unit of the option to achieve the same effect.

3.3 Risk-Neutral Pricing Formula

Definition 3.7. A market model is complete if every derivative security can be hedged.

We will always assume that the market model we deal with is complete. In other word, for a given payoff $V(T)$ of the derivative security, we are able to choose an initial capital $X(0)$ and a portfolio strategy $(\phi(t), \psi(t))$, $0 \leq t \leq T$, such that

$$X(T) = V(T). \tag{3.19}$$

It can be proved that

$$d\left(D(t)X(t)\right) = \phi(t) d\left(D(t)S(t)\right). \tag{3.20}$$

In fact, let $f(x) = \frac{1}{x}$, then $f'(x) = -\frac{1}{x^2}$, $f''(x) = \frac{2}{x^3}$. Apply Itô's formula,

$$\begin{aligned} dM(t) &= d\left(f(D(t))\right) \\ &= f'(D(t))dD(t) + \frac{1}{2}f''(D(t))(dD(t))^2 \\ &= -\frac{1}{D^2(t)}(-r(t)D(t)dt) \\ &= \frac{r(t)}{D(t)}dt \\ &= r(t)M(t)dt. \end{aligned}$$

Here we used (3.6). Then since the portfolio $X(t)$ is self-financing, we have

$$\begin{aligned} dX(t) &= \phi(t)dS(t) + \psi(t)dM(t) \\ &= \phi(t)(\alpha(t)S(t)dt + \sigma(t)S(t)dW(t)) + r(t)\psi(t)M(t)dt \\ &= \phi(t)(\alpha(t)S(t)dt + \sigma(t)S(t)dW(t)) + r(t)(X(t) - \phi(t)S(t))dt \\ &= r(t)X(t)dt + \phi(t)(\alpha(t) - r(t))S(t)dt + \phi(t)\sigma(t)S(t)dW(t) \\ &= r(t)X(t)dt + \phi(t)\sigma(t)S(t)[\Theta(t)dt + dW(t)] \\ &= r(t)X(t)dt + \phi(t)\sigma(t)S(t)d\widetilde{W}(t). \end{aligned}$$

Then by Itô's product rule,

$$\begin{aligned} d(D(t)X(t)) &= dD(t)X(t) + D(t)dX(t) + dD(t)dX(t) \\ &= -r(t)D(t)X(t)dt + D(t)[r(t)X(t)dt + \phi(t)\sigma(t)S(t)d\widetilde{W}(t)] \\ &= \phi(t)\sigma(t)D(t)S(t)d\widetilde{W}(t) \\ &= \phi(t)d(D(t)S(t)). \end{aligned}$$

This completes the proof of (3.20).

Then according to the stock representation (3.10),

$$d\left(D(t)X(t)\right) = \phi(t)\sigma(t)D(t)S(t)d\widetilde{W}(t). \tag{3.21}$$

The corresponding integral form is

$$D(t)X(t) = D(0)X(0) + \int_0^t \phi(s)\sigma(s)D(s)S(s)d\widetilde{W}(s). \tag{3.22}$$

Since the Itô integral in the above equation is a martingale under $\widetilde{P}$, $D(t)X(t)$ is a martingale under $\widetilde{P}$ and

$$D(t)X(t) = \widetilde{E}[D(T)X(T)|\mathcal{F}(t)]. \tag{3.23}$$

The value $X(t)$ of the hedging portfolio is the value needed at time t in order to reproduce the payoff $V(T)$ of the derivative security. Therefore, we define it as the price $V(t)$ of the derivative security at time t. Thus the above equation becomes

$$D(t)V(t) = \widetilde{E}[D(T)V(T)|\mathcal{F}(t)]. \tag{3.24}$$

Since $D(t) = e^{-\int_0^t r(s)ds}$, we see

$$V(t) = \widetilde{E}\left[e^{-\int_t^T r(s)ds}V(T)|\mathcal{F}(t)\right]. \tag{3.25}$$

We refer to both (3.24) and (3.25) as the risk-neutral pricing formula.

3.4 Fundamental Theorem of Asset Pricing

We conclude this chapter by giving the Fundamental Theorems of asset pricing.

In the asset pricing, it is assumed that the market models do not admit arbitrage. An arbitrage is a way of trading with zero initial capital, which produces a positive probability to make money without any chance to lose money. The mathematical model for arbitrage is stated in the following definition.

Definition 3.8. An arbitrage is a self-financing portfolio value process $X(t)$ satisfying $X(0) = 0$ and also satisfying for some time $T > 0$,

$$P\{X(T) \geq 0\} = 1, P\{X(T) > 0\} > 0. \tag{3.26}$$

Theorem 3.9 (First fundamental Theorem of asset pricing)**.** *([19, Theorem 2.7, p.25]) The market model contains no arbitrage opportunities if and only if there exists a risk-neutral probability measure.*

We also assume in the asset pricing that a market model is complete.

Definition 3.10. A market model is complete if every derivative security can be hedged.

Theorem 3.11 (Second fundamental Theorem of asset pricing). *([36, Theorem 5.4.9, p.232]) Consider a market model that has a risk-neutral probability measure. The model is complete if and only if the risk-neutral probability measure is unique.*

Chapter 4. Stochastic Interest Rate

In this chapter, we will survey two stochastic interest rate models: Vasicek model and CIR model. They will be used to valuate the American options and American spread options later. In the case of constant interest rate, the risk-neutral pricing formula can be used to derive the partial differential equation for the option price. However, if the interest rate is stochastic, it is difficult to use the associated risk-neutral pricing formula to derive the partial differential equation for the option price. Instead we will price the option using the delta-hedging technique. This method requires the knowledge of the zero-coupon bond $B(t,T)$, which will be studied in section 4.2.

4.1 Interest Rate Models

In this section, we consider two interest rate models. The first model is called the Vasicek model which is determined by the following stochastic differential equation:

$$dr(t) = (\alpha - \beta r(t))dt + \sigma d\widetilde{W}(t), \tag{4.1}$$

where α, β, σ are positive constants, $\widetilde{W}(t)$ is a Brownian motion under the risk-neutral measure $\widetilde{P}$. Here α/β is called the long term mean of the interest rate $r(t)$, β is called the speed of reversion of the interest rate $r(t)$ and σ is called the volatility of the interest rate $r(t)$. This equation can be solved explicitly to get

$$r(t) = r(0)e^{-\beta t} + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma \int_0^t e^{-\beta(t-s)} d\widetilde{W}(s). \tag{4.2}$$

The expectation and variance are given by

$$E[r(t)] = r(0)e^{-\beta t} + \frac{\alpha}{\beta}(1 - e^{-\beta t}), \tag{4.3}$$

$$Var[r(t)] = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}). \tag{4.4}$$

The Vasicek model is a mean-reverting model in the sense that

$$E[r(t)] \to \frac{\alpha}{\beta} \text{ as } t \to \infty. \tag{4.5}$$

As time elapses, the interest rate $r(t)$ tends to drift towards its long-term mean $\frac{\alpha}{\beta}$. The variance converges to a finite value

$$Var[r(t)] \to \frac{\sigma^2}{2\beta} \text{ as } t \to \infty. \tag{4.6}$$

The Vasicek model has the disadvantage that $r(t)$ can be negative with a positive probability, which generally is not true in the financial market.

The second interest rate model is called the Cox-Ingersoll-Ross (CIR) model, which is determined by the following stochastic differential equation:

$$dr(t) = (\alpha - \beta r(t))dt + \sigma\sqrt{r(t)}d\widetilde{W}(t), \tag{4.7}$$

where α, β, σ are positive constants, $\widetilde{W}(t)$ is a Brownian motion under the risk-neutral measure $\widetilde{P}$. As in the Vasicek model, α/β is called the long term mean of the interest rate $r(t)$, β is called the speed of reversion of the interest rate $r(t)$ and σ is called the volatility of the interest rate $r(t)$.

This equation can be solved implicitly to get

$$r(t) = r(0)e^{-\beta t} + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma\int_0^t e^{-\beta(t-s)}\sqrt{r(s)}d\widetilde{W}(s). \tag{4.8}$$

The expectation and variance are given by

$$E[r(t)] = r(0)e^{-\beta t} + \frac{\alpha}{\beta}(1 - e^{-\beta t}), \tag{4.9}$$

$$Var[r(t)] = r(0)\frac{\sigma^2}{\beta}(e^{-\beta t} - e^{-2\beta t}) + \frac{\alpha\sigma^2}{2\beta^2}(1 - e^{-\beta t})^2. \tag{4.10}$$

CIR model is also a mean-reverting model since

$$E[r(t)] \to \frac{\alpha}{\beta} \text{ as } t \to \infty, \tag{4.11}$$

and the variance also converges to a finite value

$$Var[r(t)] \to \frac{\alpha\sigma^2}{2\beta^2} \text{ as } t \to \infty. \tag{4.12}$$

Intuitively, if $r(t)$ is very small, so is $\sigma\sqrt{r(t)}$. It is reasonable to neglect the random effect of the diffusion term. In another word, $r(t)$ is mainly affected by the drift term $\alpha - \beta r(t)$ which is close to the positive constant α since $r(t)$ is small at this time, as a consequence $r(t)$ will be pulled back from going negative and will remain non-negative.

4.2 Zero-Coupon Bond $B(t, T)$

In this section, we study the zero-coupon bond and derive the stochastic differential equation for $B(t, T)$ under the Vasicek and the CIR model.

Definition 4.1. A zero-coupon bond $B(t, T)$, $0 \leq t \leq T$ is a debt security that does not pay interest (a coupon) but is traded at a deep discount. A zero-coupon bond is bought at a price lower than its face value, with the face value repaid at the time of maturity. In the context of option pricing, we assume the face value is $B(T, T) = 1$. T is the maturity of the zero-coupon bond. According to the risk-neutral pricing formula (3.25), the zero-coupon bond price at time t maturing at T is given by [1]

$$B(t,T) = \widetilde{E}[e^{-\int_t^T r(s)ds} | \mathcal{F}(t)]. \tag{4.13}$$

In particular, under the Vasicek and the CIR models, the zero-coupon bond $B(t,T)$ can be expressed as an exponential of an affine function of $r(t)$.

Theorem 4.2. *If the interest rate $r(t)$ is determined by the Vasicek model, then the bond price is of the form*

$$B(t,T) = e^{-r(t)C(t,T) - A(t,T)}, \tag{4.14}$$

where $C(t,T), A(t,T)$ are deterministic functions of t and T given by

$$\begin{aligned}
C(t,T) &= \int_t^T e^{-\int_t^s \beta dv} ds = \frac{1}{\beta}(1 - e^{-\beta(T-t)}), \\
A(t,T) &= \int_t^T \left(\alpha C(s,T) - \frac{1}{2}\sigma^2 C^2(s,T) \right) ds \\
&= \left(\frac{\alpha}{\beta} - \frac{\sigma^2}{2\beta^2} \right)(T-t) + \frac{1}{\beta}\left(-\frac{\alpha}{\beta} + \frac{\sigma^2}{\beta^2} \right)\left(1 - e^{-\beta(T-t)}\right) \\
&\quad - \frac{\sigma^2}{4\beta^3}\left(1 - e^{-2\beta(T-t)}\right).
\end{aligned} \tag{4.15}$$

If the interest rate $r(t)$ is determined by the CIR model, then the bond price is of the form

$$B(t,T) = e^{-r(t)C(t,T) - A(t,T)}, \tag{4.16}$$

[1]Note that if the interest rate is not a constant, the zero-coupon bond should be a function of $r(t)$. We use $B(t,T)$ to denote the zero-coupon bond only for convention, it does not mean B only depends on t and T.

where $C(t,T), A(t,T)$ *are deterministic functions of* t *and* T *given by*

$$\begin{aligned} C(t,T) &= \frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t))+\frac{1}{2}\beta\sinh(\gamma(T-t))}, \\ A(t,T) &= -\frac{2\alpha}{\sigma^2}\ln\left[\frac{\gamma e^{\frac{1}{2}\beta(T-t)}}{\gamma\cosh(\gamma(T-t))+\frac{1}{2}\beta\sinh(\gamma(T-t))}\right] \\ \gamma &= \frac{1}{2}\sqrt{\beta^2+2\sigma^2}. \end{aligned} \tag{4.17}$$

Next we study the stochastic differential equation for the zero-coupon bond $B(t,T)$ under two interest rate models. Recall that for both the Vasicek and CIR models, the zero-coupon bond is of the form

$$B(t,T) = e^{-C(t,T)r(t)-A(t,T)}.$$

Firstly, suppose that $r(t)$ is determined by the Vasicek model. Let $f(x)=e^x$ then $f'(x)=f''(x)=e^x$. Apply Itô's formula

$$\begin{aligned} dB(t,T) &= f'(-C(t,T)r(t)-A(t,T))d(-C(t,T)r(t)-A(t,T)) \\ &\quad + \frac{1}{2}f''(-C(t,T)r(t)-A(t,T))(d(-C(t,T)r(t)-A(t,T)))^2 \\ &= B(t,T)d(-C(t,T)r(t)-A(t,T)) \\ &\quad + \frac{1}{2}B(t,T)(d(-C(t,T)r(t)-A(t,T)))^2. \end{aligned} \tag{4.18}$$

Here

$$\begin{aligned} d(-C(t,T)r(t)-A(t,T)) &= -r(t)dC(t,T)-C(t,T)dr(t)-dA(t,T) \\ &= [-r(t)C'(t,T)-C(t,T)(\alpha-\beta r(t))-A'(t,T)]dt \\ &\quad - \sigma C(t,T)d\widetilde{W}(t). \end{aligned} \tag{4.19}$$

And

$$(d(-C(t,T)r(t)-A(t,T)))^2 = \sigma^2C^2(t,T)dt. \tag{4.20}$$

Plugging (4.19) and (4.20) to (4.18), we get

$$\frac{dB(t,T)}{B(t,T)} = \left[-r(t)C'(t,T) - C(t,T)(\alpha - \beta r(t)) - A'(t,T)dt + \frac{1}{2}\sigma^2 C^2(t,T)\right] dt$$
$$- \sigma C(t,T)d\widetilde{W}(t). \tag{4.21}$$

A simple computation using (4.15) can show that the drift term (the square bracket part) of the above stochastic differential equation is equal to $r(t)$. Hence, for the Vasicek model, we have the stochastic differential equation of the zero-coupon bond under the risk-neutral measure

$$\frac{dB(t,T)}{B(t,T)} = r(t)dt - \sigma C(t,T)d\widetilde{W}(t). \tag{4.22}$$

Similarly, for the CIR model, the stochastic differential equation is given by

$$\frac{dB(t,T)}{B(t,T)} = r(t)dt - \sigma C(t,T)\sqrt{r(t)}d\widetilde{W}(t). \tag{4.23}$$

Plugging in the associated values of $C(t,T)$ for Vasicek model and CIR model respectively to the stochastic differential equations above, we obtain the volatility of the zero-coupon bond $B(t,T)$:

For the Vasicek model, the volatility of $B(t,T)$ is

$$\sigma_V = -\frac{\sigma}{\beta}(1 - e^{-\beta(T-t)}). \tag{4.24}$$

Then

$$dB(t,T) = r(t)B(t,T)dt + \sigma_V B(t,T)d\widetilde{W}(t). \tag{4.25}$$

For the CIR model, the volatility of $B(t,T)$ is

$$\sigma_C = -\sigma\sqrt{r(t)}\frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t)) + \frac{1}{2}\beta\sinh(\gamma(T-t))}, \tag{4.26}$$

where $\gamma = \frac{1}{2}\sqrt{\beta^2 + 2\sigma^2}$. Then

$$dB(t,T) = r(t)B(t,T)dt + \sigma_C B(t,T)d\widetilde{W}(t). \tag{4.27}$$

Chapter 5. American Option

In this chapter, we will incorporate stochasticity into the interest rate and price the American option under the Vasicek model and CIR model. Firstly we will derive the partial differential equation for the corresponding European option. Then we will formulate the free boundary problem associated with the American option. In order to go around the free boundary and implement the numerical computation, we will convert the free boundary problem to the linear complementarity conditions for the American option. We will conclude this chapter by numerical computation using finite difference method.

5.1 Introduction

In order to understand the American option, we need to introduce the European option. An European option is a contract which gives its owner the right but not the obligation to buy (call option) or sell (put option) the underlying asset $S(t), 0 \leq t \leq T$ at a specified price K (strike price), on a specified date T (expiration date). The payoff of the European option is the value of the option on the expiration date, which is $V(T) = (S(t) - K)^+$ for the call option and $V(T) = (K - S(t))^+$ for the put option. Thus by the risk-neutral pricing formula (3.25), the European option price at time t is given by

$$V(t) = \widetilde{E}\left[e^{-r(T-t)}V(T)|\mathcal{F}(t)\right], \tag{5.1}$$

where r is the constant interest rate and the expectation is taken under the risk-neutral measure $\widetilde{P}$.

The American option is more flexible than the European option. It can be exercised any time before or on the expiration date. The precise definition is given as follows.

Definition 5.1. Let $0 \leq t \leq T$, $x \geq 0$ be given. Assume $S(t) = x$. The price at time t of the American option expiring at time T is defined to be

$$V(t,x) = \max_{\tau \in \mathcal{T}_{[t,T]}} \widetilde{E}[e^{-r(\tau-t)}V(\tau)|S(t)=x]. \tag{5.2}$$

where $V(\tau) = (S(\tau) - K)^+$ for the call option and $V(\tau) = (K - S(\tau))^+$ for the put option. $\mathcal{T}_{[t,T]}$ denotes the set of stopping times for the filtration $\{\mathcal{F}_t(u), t \le u \le T\}$, taking values in $[t,T]$ or taking the value ∞, where $\mathcal{F}_t(u), t \le u \le T$ denotes the σ-algebra generated by the process $S(v), t \le v \le u$.

From now on, we will consider the American put option. Pricing the American call option is similar.

The key of pricing the American option is to determine if we need to hold the option or exercise the option immediately at time t. It is shown in [36, p.356–361] that the optimal exercise policy for the American option is of the form: "Exercise the put option as soon as $S(t)$ falls to a certain level $\mathcal{L}$." In other words, the owner of the American put option should wait until the underlying asset price falls to a certain level $\mathcal{L}$ before exercising. This is a defining property of the American option, i.e., the American option price determined by this optimal exercise policy is the same as the one in Definition 5.1.

By the discussion above, we can divide the set $\{(t,x) : 0 \le t \le T, x \ge 0\}$ into two regions, once the underlying asset price falls into the stopping region $\mathcal{S}$, the owner of the put option should exercise it immediately to obtain the immediate payoff value (intrinsic value). Thus the stopping region can be characterized as

$$\mathcal{S} = \{(t,x) : V(t,x) = (K-x)^+\}. \tag{5.3}$$

On the other hand, if $S(t)$ remains in the continuation region $\mathcal{C}$, the owner should wait and the current value of the American put option is the same as the current value of the corresponding European put option. Thus the continuation region can be characterized as

$$\mathcal{C} = \{(t,x) : V(t,x) = \text{the corresponding European put option price}\}. \tag{5.4}$$

The level $\mathcal{L}$ mentioned in the optimal exercise policy is the free boundary between $\mathcal{S}$ and $\mathcal{C}$ which is denoted by

$$\mathcal{L} : L(T-t,x) = 0. \tag{5.5}$$

Since in the stopping region, the American put option price has already been determined and equals to the intrinsic value, we just need to consider the option price in the continuation region. In the next section, we will price the corresponding European put option.

5.2 The Corresponding European Put Option

In this section, We derive the partial differential equation for the corresponding European put option with stochastic interest rate based on Fang's approach in [18]. Firstly we will set up the stochastic differential equations for the underlying asset and the interest rate. Then we will use the delta-hedging technique to derive the partial differential equation for the put option.

In the Black-Scholes framework, the underlying asset $S_1(t)$ is a geometric Brownian motion which satisfies the following stochastic differential equation.

$$dS(t) = rS(t)dt + \sigma_1 S(t) d\widetilde{W}_1(t),$$

where r is the interest rate, σ_1 is the volatility of $S_1(t)$, and they are both positive constants. $\widetilde{W}_1(t)$ is a Brownian motion under the risk-neutral measure.

Now we consider stochastic interest rate process $r(t)$. Then the underlying asset price satisfies the following stochastic differential equation.

$$dS(t) = r(t)S(t)dt + \sigma_1 S(t)d\widetilde{W}_1(t), \tag{5.6}$$

where the interest rate $r(t)$ satisfies one of the following models we mentioned in the previous chapter under the risk-neutral measure: [1]

- Vasicek model

$$dr(t) = (\alpha - \beta r(t))dt + \sigma_3 d\widetilde{W}_3(t) \tag{5.7}$$

- CIR model

$$dr(t) = (\alpha - \beta r(t))dt + \sigma_3 \sqrt{r(t)}d\widetilde{W}_3(t) \tag{5.8}$$

Here the two Brownian motions are assumed to have the correlation ρ_2:

$$d\widetilde{W}_1(t)d\widetilde{W}_3(t) = \rho_2 dt. \tag{5.9}$$

Note that the constants $\alpha, \beta, \sigma_3, \rho_2$ are all positive. Equation (5.9) is the differential form for the cross variation of $\widetilde{W}_1, \widetilde{W}_3$

$$[\widetilde{W}_1, \widetilde{W}_3](t) = \lim_{||\Pi|| \to 0} \sum_{i=0}^{n-1} [\widetilde{W}_1(t_{i+1}) - \widetilde{W}_1(t_i)][\widetilde{W}_3(t_{i+1}) - \widetilde{W}_3(t_i)] = \rho_2 t, \tag{5.10}$$

where $\Pi = \{t_0, ..., t_n : 0 = t_0 \leq t_1 \leq \cdots \leq t_n = t\}$ is a partition of the interval $[0, t]$.

The key of delta-hedging technique is to find a hedging portfolio of the option

$$X(t) = \Delta_1(t)S(t) + \Delta_2(t)B(t,T) + [X(t) - \Delta_1(t)S(t) - \Delta_2(t)B(t,T)],$$

such that $X(T) = V(T)$. Here we invest $\Delta_1(t)$ units of stock $S(t)$, $\Delta_2(t)$ units of zero-coupon bond $B(t,T)$ and invest the rest in the money market account. We also assume that

[1] The reason we use index $1, 3$ here instead of $1, 2$ is for the consistency with the next chapter.

the portfolio $X(t)$ is self-financing, i.e., $dX(t) = \Delta_1(t)dS(t) + \Delta_2(t)dB(t,T) + r(t)[X(t) - \Delta_1(t)S(t) - \Delta_2(t)B(t,T)]dt$. We have seen in (3.18) that the value of the hedging portfolio is equal to the value of the option at any time before the expiration date, i.e., $X(t) = V(t), \ \forall\ 0 \le t \le T$.

With the preparation above, we will derive the partial differential equation for the put option in three steps:

Step 1 We compute the evolution of the portfolio under the Vasicek model and the CIR model.

For the Vasicek model,

$$\begin{aligned} dX(t) &= \Delta_1(t)dS(t) + \Delta_2(t)dB(t,T) \\ &\quad + r(t)[X(t) - \Delta_1(t)S(t) - \Delta_2(t)B(t,T)]dt \\ &= \Delta_1(t)[r(t)S(t)dt + \sigma_1 S(t)d\widetilde{W}_1(t)] \\ &\quad + \Delta_2(t)[r(t)B(t,T)dt + \sigma_V B(t,T)d\widetilde{W}_3(t)] \\ &\quad + r(t)[X(t) - \Delta_1(t)S(t) - \Delta_2(t)B(t,T)]dt \\ &= r(t)X(t)dt + \Delta_1(t)\sigma_1 S(t)d\widetilde{W}_1(t) + \Delta_2(t)\sigma_V B(t,T)d\widetilde{W}_3(t). \end{aligned}$$

Here we used (4.25) in the second line above.

Replace $X(t)$ by $V(t)$, we have

$$dV(t) = r(t)V(t)dt + \Delta_1(t)\sigma_1 S(t)d\widetilde{W}_1(t) + \Delta_2(t)\sigma_V B(t,T)d\widetilde{W}_3(t). \tag{5.11}$$

For the CIR model,

$$\begin{aligned}
dX(t) &= \Delta_1(t)dS(t) + \Delta_2(t)dB(t,T) \\
&\quad + r(t)[X(t) - \Delta_1(t)S(t) - \Delta_2(t)B(t,T)]dt \\
&= \Delta_1(t)[r(t)S(t)dt + \sigma_1 S(t)d\widetilde{W}_1(t)] \\
&\quad + \Delta_2(t)[r(t)B(t,T)dt + \sigma_C B(t,T)d\widetilde{W}_3(t)] \\
&\quad + r(t)[X(t) - \Delta_1(t)S(t) - \Delta_2(t)B(t,T)]dt \\
&= r(t)X(t)dt + \Delta_1(t)\sigma_1 S(t)d\widetilde{W}_1(t) + \Delta_2(t)\sigma_C B(t,T)d\widetilde{W}_3(t).
\end{aligned}$$

Here we used (4.27) in the second line above.

Replace $X(t)$ by $V(t)$, we have

$$dV(t) = r(t)V(t)dt + \Delta_1(t)\sigma_1 S(t)d\widetilde{W}_1(t) + \Delta_2(t)\sigma_C B(t,T)d\widetilde{W}_3(t). \tag{5.12}$$

Step 2 By the Itô's formula, we can compute the evolution of the option price $V(t)$.

For the Vasicek model,

$$\begin{aligned}
dV(t,S(t),r(t)) &= V_t dt + V_S dS(t) + V_r dr(t) \\
&\quad + \frac{1}{2}V_{SS}dS(t)dS(t) + \frac{1}{2}V_{rr}dr(t)dr(t) + V_{Sr}dS(t)dr(t) \\
&= V_t dt + V_S[r(t)S(t)dt + \sigma_1 S(t)d\widetilde{W}_1(t)] \\
&\quad + V_r[(\alpha - \beta r(t))dt + \sigma_3 d\widetilde{W}_3(t)] \\
&\quad + \frac{1}{2}\sigma_1^2 S^2(t)V_{SS}dt + \frac{1}{2}\sigma_3^2 V_{rr}dt + \sigma_1\sigma_3\rho_2 S(t)V_{Sr}dt \\
&= [V_t + r(t)S(t)V_S + (\alpha - \beta r(t))V_r + \frac{1}{2}\sigma_1^2 S^2(t)V_{SS} \\
&\quad + \frac{1}{2}\sigma_3^2 V_{rr} + \sigma_1\sigma_3\rho_2 S(t)V_{Sr}]dt \\
&\quad + \sigma_1 S(t)V_S d\widetilde{W}_1(t) + \sigma_3 V_r d\widetilde{W}_3(t).
\end{aligned} \tag{5.13}$$

For the CIR model,

$$
\begin{aligned}
dV(t,S(t),r(t)) &= V_t dt + V_S dS(t) + V_r dr(t) \\
&\quad + \frac{1}{2}V_{SS}dS(t)dS(t) + \frac{1}{2}V_{rr}dr(t)dr(t) + V_{Sr}dS(t)dr(t) \\
&= V_t dt + V_S[r(t)S(t)dt + \sigma_1 S(t)d\widetilde{W_1}(t)] \\
&\quad + V_r[(\alpha - \beta r(t))dt + \sigma_3\sqrt{r}d\widetilde{W_3}(t)] \\
&\quad + \frac{1}{2}\sigma_1^2 S^2(t)V_{SS}dt + \frac{1}{2}\sigma_3^2 rV_{rr}dt + \sigma_1\sigma_3\rho_2\sqrt{r}S(t)V_{Sr}dt \\
&= [V_t + r(t)S(t)V_S + (\alpha - \beta r(t))V_r \\
&\quad + \frac{1}{2}\sigma_1^2 S^2(t)V_{SS} + \frac{1}{2}\sigma_3^2 rV_{rr} \\
&\quad + \sigma_1\sigma_3\rho_2\sqrt{r}S(t)V_{Sr}]dt + \sigma_1 S(t)V_S d\widetilde{W_1}(t) + \sigma_3\sqrt{r}V_r d\widetilde{W_3}(t). \quad (5.14)
\end{aligned}
$$

Step 3 Equating the evolutions of the portfolio (5.11, 5.12) with the evolutions of the option price (5.13, 5.14) respectively to get:

(a)For the Vasicek model,

$$\Delta_1(t) = V_S(t), \quad \Delta_2(t) = \frac{\sigma_3 V_r}{\sigma_V B(t,T)}, \quad (5.15)$$

and

$$
\begin{aligned}
r(t)V(t) &= V_t(t) + r(t)S(t)V_S(t) \\
&\quad + (\alpha - \beta r(t))V_r + \frac{1}{2}\sigma_1^2 S^2(t)V_{SS}(t) + \frac{1}{2}\sigma_3^2 V_{rr}(t) + \sigma_1\sigma_3\rho_2 S(t)V_{Sr}(t) \quad (5.16)
\end{aligned}
$$

Replacing $S(t), r(t)$ by x, r and simplifying the equation above, we get the desired partial differential equation

$$rV = V_t(t) + rxV_x + (\alpha - \beta r)V_r + \frac{1}{2}\sigma_1^2 x^2(t)V_{xx} + \frac{1}{2}\sigma_3^2 V_{rr} + \sigma_1\sigma_3\rho_2 xV_{xr}. \quad (5.17)$$

(b)For the CIR model,

$$\Delta_1(t) = V_S(t), \quad \Delta_2(t) = \frac{\sigma_3\sqrt{r}V_r}{\sigma_C B(t,T)}, \tag{5.18}$$

and

$$\begin{aligned} r(t)V(t) &= V_t(t) + r(t)S(t)V_S(t) \\ &+ (\alpha - \beta r(t))V_r + \frac{1}{2}\sigma_1^2 S^2(t)V_{SS}(t) + \frac{1}{2}\sigma_3^2 r(t)V_{rr}(t) + \sigma_1\sigma_3\rho_2\sqrt{r(t)}S(t)V_{Sr}(t) \end{aligned} \tag{5.19}$$

Replacing $S(t), r(t)$ by x, r and simplifying the equation above, we get the desired partial differential equation

$$rV = V_t(t) + rxV_x + (\alpha - \beta r)V_r + \frac{1}{2}\sigma_1^2 x^2(t)V_{xx} + \frac{1}{2}\sigma_3^2 rV_{rr} + \sigma_1\sigma_3\rho_2\sqrt{r}xV_{xr}. \tag{5.20}$$

5.3 Free Boundary Problem of American Put Option

In this section, we will formulate the free boundary problem of the American put option with stochastic interest rate. According to the optimal exercise policy, we know there exists an early exercise boundary $\mathcal{L}$ such that the American put option needs to be immediately exercised if the underlying asset price falls to $\mathcal{L}$. Although the boundary $\mathcal{L}$ is unknown, we can derive the conditions for the value of the option in the stopping region and the continuation region respectively. Furthermore, we can give the terminal and boundary conditions for the American put option.

Similar to the case of the constant interest rate, the set $\{(t,x,r) : 0 \le t \le T, x \ge 0, r \ge 0\}$ can be divided into two regions, the stopping region

$$\mathcal{S} = \{(t,x,r) : V(t,x,r) = (K - x)^+\} \tag{5.21}$$

and the continuation region

$$\mathcal{C} = \{(t, x, r) : V(t, x, r) = \text{the corresponding European put option price}\}. \tag{5.22}$$

The free boundary between $\mathcal{S}$ and $\mathcal{C}$ is the one mentioned in the optimal exercise policy

$$\mathcal{L} : L(T - t, x, r) = 0. \tag{5.23}$$

Correspondingly, it is optimal to exercise immediately to obtain the intrinsic value of the American put option when (t, x, r) is in the stopping region $\mathcal{S}$. While in the continuation region $\mathcal{C}$, it is optimal to hold the American put option. Thus in the continuation region, the American put option price is equal to the corresponding European put option price and satisfies the partial differential equations we derived in the previous section.

Next we consider the terminal and boundary conditions for the American put option. Firstly, the payoff on the expiration date is the terminal condition

$$V(T, x, r) = (K - x)^+. \tag{5.24}$$

Secondly, we assume that the option price V and the derivatives V_x, V_r are continuous. Since $V(t, x, r) = K - x$ is the option price in the stopping region $\mathcal{S}$, the left-hand derivative with respect to x on $\mathcal{L}$ is $V_x(t, x-, r) = -1$ and the left-hand derivative with respect to r on $\mathcal{L}$ is $V_r(t, x, r-) = 0$. Then the option price satisfies the smooth-pasting condition

$$V_x(t_0, x_0, r_0) = \lim_{x \to x_0^+} V_x(t_0, x, r_0) = \lim_{x \to x_0^-} V_x(t_0, x, r_0) = -1, \quad \forall (t_0, x_0, r_0) \in \mathcal{L} \tag{5.25}$$

$$V_r(t_0, x_0, r_0) = \lim_{r \to r_0^+} V_r(t_0, x_0, r) = \lim_{r \to r_0^-} V_r(t_0, x_0, r) = 0, \quad \forall (t_0, x_0, r_0) \in \mathcal{L}. \tag{5.26}$$

Furthermore, we have the asymptotic conditions [2]

$$\lim_{x\to\infty} V(t,x,r) = 0 \tag{5.27}$$

$$\lim_{r\to\infty} V(t,x,r) = 0. \tag{5.28}$$

Equation (5.27) is set due to the fact that if the underlying asset price is large then the put option is unlikely to be exercised and gains no value. For Equation (5.28), notice that the underlying asset satisfies the stochastic differential equation

$$\frac{dS(t)}{S(t)} = r(t)dt + \sigma_1 d\widetilde{W}_1(t).$$

A large interest rate will induce large increment of the asset price within a small period of time. Thus S may approach infinity, which is the case in Equation (5.27).

Finally, if the interest rate is approaching zero, there is no time value. The current option price should equal to its payoff on the expiration date, which is unknown beforehand. However, we can still find this boundary condition by letting $r = 0$ in Equation (5.17), Equation (5.20) and directly solving the partial differential equations. This will be done after we simplify the partial differential equations.

5.4 Simplification of the Free Boundary Problem

In this section, we will simplify the partial differential equations (5.17) and (5.20) by change of variables. Equipped with the simplified partial differential equations, we can reformulate the free boundary problem of the American put option.

Let $\tau = T - t$, $x = e^u$ and $h(\tau, u, r) = V(t, x, r)$. Then the partial derivatives of V can

[2] As $x \to 0$, option price is in the stopping region, so $V = (K - x)^+$.

be expressed as

$$V_t = h_\tau \frac{d\tau}{dt} = -h_\tau \tag{5.29}$$

$$V_x = h_u \frac{du}{dx} = \frac{1}{x} h_u \tag{5.30}$$

$$V_r = h_r \tag{5.31}$$

$$V_{rr} = h_{rr} \tag{5.32}$$

$$V_{xr} = \frac{1}{x} h_{ur} \tag{5.33}$$

$$V_{xx} = \frac{\partial}{\partial x}(\frac{1}{x} h_u) = -\frac{1}{x^2} h_u + \frac{1}{x} h_{uu} \frac{du}{dx} = -\frac{1}{x^2} h_u + \frac{1}{x^2} h_{uu} \tag{5.34}$$

Plugging (5.29)–(5.34) to the partial differential equation (5.17), we have the simplified partial differential equation for the Vasicek model

$$rh + h_\tau = (r - \frac{1}{2}\sigma_1^2) h_u + (\alpha - \beta r) h_r + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_3^2 h_{rr} + \sigma_1 \sigma_3 \rho_2 h_{ur} \tag{5.35}$$

And we let

$$\mathcal{Q} h = rh + h_\tau - [(r - \frac{1}{2}\sigma_1^2) h_u + (\alpha - \beta r) h_r + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_3^2 h_{rr} + \sigma_1 \sigma_3 \rho_2 h_{ur}] \tag{5.36}$$

Plugging (5.29)–(5.34) to the partial differential equation (5.20), we have the simplified partial differential equation for the CIR model

$$rh + h_\tau = (r - \frac{1}{2}\sigma_1^2) h_u + (\alpha - \beta r) h_r + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_3^2 r h_{rr} + \sigma_1 \sigma_3 \rho_2 \sqrt{r} h_{ur} \tag{5.37}$$

And we let

$$\mathcal{Q} h = rh + h_\tau - [(r - \frac{1}{2}\sigma_1^2) h_u + (\alpha - \beta r) h_r + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_3^2 r h_{rr} + \sigma_1 \sigma_3 \sqrt{r} \rho_2 h_{ur}] \tag{5.38}$$

Now we formulate the free boundary problem associated with the function $h(\tau, u, r)$. We firstly make the following change of variables for the American put option.

$$V(t,x,r) = h(\tau,u,r) \tag{5.39}$$

$$\tau = T - t \tag{5.40}$$

$$x = e^u. \tag{5.41}$$

And we define the payoff function associated with $h(\tau,u,r)$ as

$$g(\tau,u,r) := (K - e^u)^+, \tag{5.42}$$

Then the free boundary problem associated with the function $h(\tau,u,r)$ is stated as follows: In the stopping region $\mathcal{S}$ [3]

$$\mathcal{Q}h > 0 \tag{5.43}$$

$$h(\tau,u,r) = g(\tau,u,r). \tag{5.44}$$

In the continuation region $\mathcal{C}$

$$\mathcal{Q}h = 0 \tag{5.45}$$

$$h(\tau,u,r) > g(\tau,u,r). \tag{5.46}$$

The initial condition is

$$h(0,u,r) = g(0,u,r) = (K - e^u)^+. \tag{5.47}$$

The smooth-pasting conditions are

$$h_u(\tau_0,u_0,r_0) = \lim_{u\to u_0^+} h_u(\tau_0,u,r_0) = \lim_{u\to u_0^-} h_u(\tau_0,u,r_0) = -e^{u_0}, \quad \forall(\tau_0,u_0,r_0) \in \mathcal{L} \tag{5.48}$$

[3] Replacing h by g in 5.35 and 5.37, we get $\mathcal{Q}h > 0$.

$$h_r(\tau_0, u_0, r_0) = \lim_{r \to r_0^+} h_r(\tau_0, u_0, r) = \lim_{r \to r_0^-} h_r(\tau_0, u_0, r) = 0, \quad \forall (\tau_0, u_0, r_0) \in \mathcal{L}. \tag{5.49}$$

The asymptotic conditions [4] are

$$\lim_{u \to \infty} h(\tau, u, r) = 0 \tag{5.50}$$

$$\lim_{r \to \infty} h(\tau, u, r) = 0. \tag{5.51}$$

Finally, if the interest rate $r = 0$, then

$$h(\tau, u, 0) = \max \left\{ K\Phi\left(\frac{\ln K - u + \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right) - e^u\Phi\left(\frac{\ln K - u - \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right), (K - e^u)^+ \right\}, \tag{5.52}$$

where Φ is the cumulative distribution function of the standard normal distribution. The derivation is in Appendix A and Appendix B.

5.5 Linear Complementarity Conditions

The major difficulty of pricing American option is that the early exercise boundary $\mathcal{L}$ is unknown. Although we can compute numerically the partial differential equations in the free boundary problem in the previous section, it is difficult to track the early exercise boundary $\mathcal{L}$. On the other hand, the function h in the previous section satisfies the linear complementarity conditions that can help us go around the free boundary problem.

In practice, when we use numerical schemes to compute the option price, it is impossible to implement infinity on computers. Instead, we consider the spatial variable u in the interval $-u^- \leq u \leq u^+$, where u^-, u^+ are large positive numbers. Similarly, r is in the interval $0 \leq r \leq r^+$, where r^+ is a large positive number.

[4] As $u \to -\infty$, option price is in the stopping region, so $h = g$.

The linear complementarity conditions for the function h is:

$$\begin{aligned} &\mathcal{Q}h \cdot (h(\tau,u,r) - g(\tau,u,r)) = 0 \\ &\mathcal{Q}h \geq 0, \quad h(\tau,u,r) \geq g(\tau,u,r) \end{aligned} \tag{5.53}$$

with the initial and boundary conditions

$$\begin{aligned} &h(0,u,r) = g(0,u,r) = (K - e^u)^+ \\ &h(\tau,-u^-,r) = g(\tau,-u^-,r) = K \\ &h(\tau,u^+,r) = 0 \\ &h(\tau,u,0) = \\ &\quad \max\left\{K\Phi\left(\frac{\ln K - u + \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right) - e^u\Phi\left(\frac{\ln K - u - \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right), (K - e^u)^+\right\} \\ &h(\tau,u,r^+) = 0. \end{aligned} \tag{5.54}$$

5.6 Finite Difference Method And Implementation

The finite difference discretization of the the partial derivatives of h are as follows:

- Forward difference

$$h_\tau \approx \frac{h_{i,k}^{m+1} - h_{i,k}^m}{\Delta\tau} + O(\Delta\tau) \tag{5.55}$$

- Central difference

$$h_u \approx \frac{h_{i+1,k}^m - h_{i-1,k}^m}{2\Delta u} + O((\Delta u)^2) \tag{5.56}$$

$$h_r \approx \frac{h_{i,k+1}^m - h_{i,k-1}^m}{2\Delta r} + O((\Delta r)^2) \tag{5.57}$$

- Symmetric central difference

$$h_{uu} \approx \frac{h^m_{i+1,k} - 2h^m_{i,k} + h^m_{i-1,k}}{(\Delta u)^2} + O((\Delta u)^2) \tag{5.58}$$

$$h_{rr} \approx \frac{h^m_{i,k+1} - 2h^m_{i,k} + h^m_{i,k-1}}{(\Delta r)^2} + O((\Delta r)^2) \tag{5.59}$$

$$h_{ur} \approx \frac{h^m_{i+1,k+1} - h^m_{i-1,k+1} - h^m_{i+1,k-1} + h^m_{i-1,k-1}}{4\Delta u \Delta r} + O((\Delta u)^2 + (\Delta r)^2) \tag{5.60}$$

Let N, M, R be large positive numbers, then

$$\Delta\tau = \frac{T}{N}$$
$$\Delta u = \frac{u^+ + u^-}{M-1}$$
$$\Delta r = \frac{r^+}{R-1}$$

and

$$\tau = 0 + m\Delta\tau = m\Delta\tau, \quad 1 \le m \le N$$

$$u = -u^- + (i-1)\Delta u, \quad 1 \le i \le M$$

$$r = 0 + (k-1)\Delta r = (k-1)\Delta r, \quad 1 \le k \le R.$$

Plugging Equations (5.55)–(5.60) into the PDE (5.35), we obtain the discretization of the partial differential equation for the Vasicek model.

$$
\begin{aligned}
rh_{i,k}^{m} + \frac{h_{i,k}^{m+1} - h_{i,k}^{m}}{\Delta\tau} + O(\Delta\tau) = (r - \frac{1}{2}\sigma_1^2)\left[\frac{h_{i+1,k}^{m} - h_{i-1,k}^{m}}{2\Delta u} + O((\Delta u)^2)\right] \\
+ (\alpha - \beta r)\left[\frac{h_{i,k+1}^{m} - h_{i,k-1}^{m}}{2\Delta r} + O((\Delta r)^2)\right] \\
+ \frac{1}{2}\sigma_1^2\left[\frac{h_{i+1,k}^{m} - 2h_{i,k}^{m} + h_{i-1,k}^{m}}{(\Delta u)^2} + O((\Delta u)^2)\right] \\
+ \frac{1}{2}\sigma_3^2\left[\frac{h_{i,k+1}^{m} - 2h_{i,k}^{m} + h_{i,k-1}^{m}}{(\Delta r)^2} + O((\Delta r)^2)\right] \\
+ \sigma_1\sigma_3\rho_2\left[\frac{h_{i+1,k+1}^{m} - h_{i-1,k+1}^{m} - h_{i+1,k-1}^{m} + h_{i-1,k-1}^{m}}{4\Delta u\Delta r} + O((\Delta u)^2 + (\Delta r)^2)\right]. \quad (5.61)
\end{aligned}
$$

Let

$$
\begin{aligned}
A &= \frac{1}{2}\frac{\Delta\tau}{\Delta u}\left(r - \frac{1}{2}\sigma_1^2\right) \\
B &= \frac{1}{2}\frac{\Delta\tau}{\Delta r}(\alpha - \beta r) \\
C &= \frac{1}{2}\sigma_1^2\frac{\Delta\tau}{(\Delta u)^2} \\
D &= \frac{1}{2}\sigma_3^2\frac{\Delta\tau}{(\Delta r)^2} \\
E &= \frac{1}{4}\sigma_1\sigma_3\rho_2\frac{\Delta\tau}{\Delta u\Delta r}.
\end{aligned}
$$

Ignoring the higher order terms, we obtain the finite difference scheme of h for the Vasicek model

$$
\begin{aligned}
h_{i,k}^{m+1} = Eh_{i+1,k+1}^{m} + (A + C)h_{i+1,k}^{m} - Eh_{i+1,k-1}^{m} \\
+ (B + D)h_{i,k+1}^{m} + (1 - r\Delta\tau - 2C - 2D)h_{i,k}^{m} + (-B + D)h_{i,k-1}^{m} \\
- Eh_{i-1,k+1}^{m} + (-A + C)h_{i-1,k}^{m} + Eh_{i-1,k-1}^{m}. \quad (5.62)
\end{aligned}
$$

Plugging Equations (5.55)–(5.60) into the PDE (5.37), we obtain the discretization of the partial differential equation for the CIR model.

$$\begin{aligned} rh_{i,k}^m + \frac{h_{i,k}^{m+1} - h_{i,k}^m}{\Delta\tau} + O(\Delta\tau) &= (r - \frac{1}{2}\sigma_1^2)\left[\frac{h_{i+1,k}^m - h_{i-1,k}^m}{2\Delta u} + O((\Delta u)^2)\right] \\ &+ (\alpha - \beta r)\left[\frac{h_{i,k+1}^m - h_{i,k-1}^m}{2\Delta r} + O((\Delta r)^2)\right] \\ &+ \frac{1}{2}\sigma_1^2\left[\frac{h_{i+1,k}^m - 2h_{i,k}^m + h_{i-1,k}^m}{(\Delta u)^2} + O((\Delta u)^2)\right] \\ &+ \frac{1}{2}\sigma_3^2 r\left[\frac{h_{i,k+1}^m - 2h_{i,k}^m + h_{i,k-1}^m}{(\Delta r)^2} + O((\Delta r)^2)\right] \\ + \sigma_1\sigma_3\rho_2\sqrt{r}&\left[\frac{h_{i+1,k+1}^m - h_{i-1,k+1}^m - h_{i+1,k-1}^m + h_{i-1,k-1}^m}{4\Delta u\Delta r} + O((\Delta u)^2 + (\Delta r)^2)\right]. \quad (5.63) \end{aligned}$$

Let

$$\begin{aligned} A &= \frac{1}{2}\frac{\Delta\tau}{\Delta u}\left(r - \frac{1}{2}\sigma_1^2\right) \\ B &= \frac{1}{2}\frac{\Delta\tau}{\Delta r}(\alpha - \beta r) \\ C &= \frac{1}{2}\sigma_1^2\frac{\Delta\tau}{(\Delta u)^2} \\ D &= \frac{1}{2}\sigma_3^2 r\frac{\Delta\tau}{(\Delta r)^2} \\ E &= \frac{1}{4}\sigma_1\sigma_3\rho_2\sqrt{r}\frac{\Delta\tau}{\Delta u\Delta r}. \end{aligned}$$

Ignoring the higher order terms, we obtain the finite difference scheme of h for the CIR model.

$$\begin{aligned} h_{i,k}^{m+1} &= Eh_{i+1,k+1}^m + (A + C)h_{i+1,k}^m - Eh_{i+1,k-1}^m \\ &+ (B + D)h_{i,k+1}^m + (1 - r\Delta\tau - 2C - 2D)h_{i,k}^m + (-B + D)h_{i,k-1}^m \\ &- Eh_{i-1,k+1}^m + (-A + C)h_{i-1,k}^m + Eh_{i-1,k-1}^m. \quad (5.64) \end{aligned}$$

To make the notation simpler, we denote both the right hand side of equation (5.62) and equation (5.64) as $\mathcal{H}h^m$. Then the finite difference scheme of h is abbreviated as

$$h_{i,k}^{m+1} = \mathcal{H}h^m. \tag{5.65}$$

To let h satisfy the linear complementarity conditions, we need to modify the finite difference scheme above to be

$$h_{i,k}^{m+1} = \max\left\{\mathcal{H}h^m, (K - e^u)^+\right\}. \tag{5.66}$$

That is to say for each time step, h takes the maximum between the value of the corresponding European put option and the immediate payoff if the early exercise is optimal.

Eventually, by the change of variables

$$V(t) = h(\tau, u, r)$$
$$\tau = T - t$$
$$x = e^u.$$

we can get the option price at the current time $V(0)$.

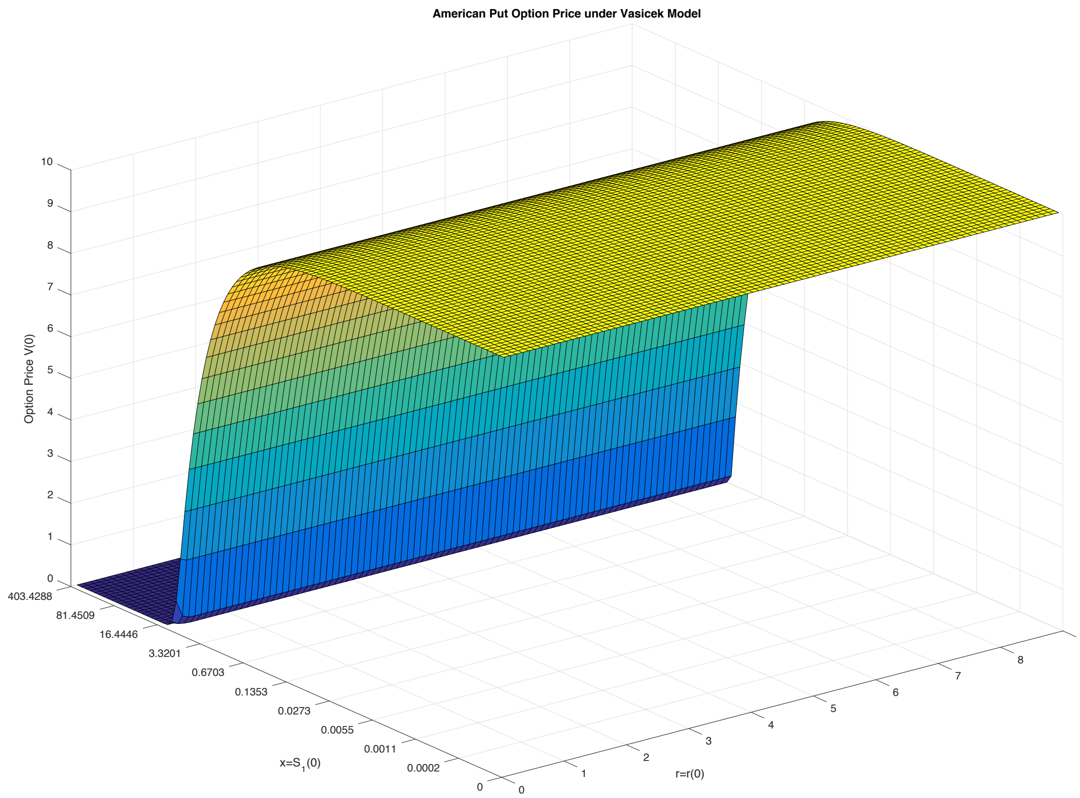

Figure 5.1: American Put Option under Vasicek model

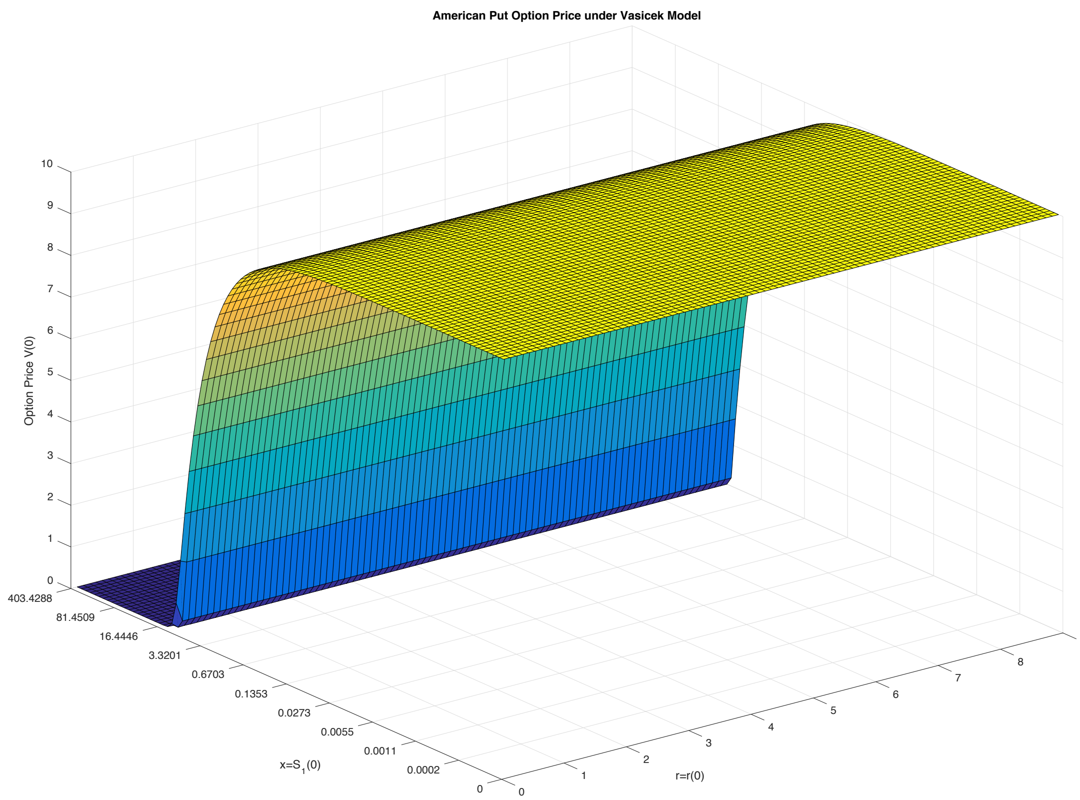

Figure 5.2: American Put Option under CIR model

Chapter 6. American Spread Option

In this chapter, we price the American spread option with stochastic interest rate under the Vasicek model and CIR model. We extend the idea and method used in the previous chapter for pricing the American option to price the American spread option. Similar to Chapter 5, we firstly derive the partial differential equation for the corresponding European spread option. Then we formulate the free boundary problem associated with the American spread option. In order to go around the free boundary and implement numerical computation, we formulate the linear complementarity conditions for the American spread option. We conclude this chapter by numerical computation using finite difference method.

6.1 Introduction

An European spread option has two underlying assets $S_1(t)$ and $S_2(t)$. The payoff on the expiration date depends on the difference of the two assets, which is $V(T) = ((S_1(T) - S_2(T)) - K)^+$ for the spread call option and $V(T) = (K - (S_1(T) - S_2(T)))^+$ for the spread put option. By the risk-neutral pricing formula (3.25), the European spread option price at time t is given by

$$V(t) = \widetilde{E}[e^{-r(T-t)}V(T)|\mathcal{F}(t)], \tag{6.1}$$

where r is the constant interest rate and the expectation is taken under the risk-neutral measure.

The American spread option can be exercised any time before or on the expiration date. The definition is given as follows.

Definition 6.1. Let $0 \leq t \leq T$, $x \geq 0$ and $y \geq 0$ be given. Assume $S_1(t) = x$ and $S_2(t) = y$. The price at time t of the American spread option expiring at time T is defined to be

$$V(t, x, y) = \max_{\tau \in \mathcal{T}_{[t,T]}} \widetilde{E}[e^{-r(\tau-t)}V(\tau))|S_1(t) = x, S_2(t) = y]. \tag{6.2}$$

where $V(\tau) = ((S_1(t) - S_2(t)) - K)^+$ for the spread call option and $V(\tau) = (K - (S_1(t) - S_2(t)))^+$ for the spread put option, $\mathcal{T}_{[t,T]}$ denotes the set of stopping times for the filtration $\{\mathcal{F}_t(u), t \leq u \leq T\}$, taking values in $[t, T]$ or taking value ∞. Here $\mathcal{F}_t(u), t \leq u \leq T$ denotes the σ-algebra generated by the process $S_1(v_1), t \leq v_1 \leq u$ and $S_2(v_2), t \leq v_2 \leq u$.

From now on, we will consider the American spread put option. Pricing the American spread call option is similar.

As the American put option, there is also an optimal exercise policy for the American spread put option. The option price determined by the optimal exercise policy is the same as the one in Definition 6.1. The optimal exercise policy is of the form: "Exercise the spread put option as soon as $(S_1(t) - S_2(t))$ falls to a certain level $\mathcal{L}$." In other words, the owner of the American spread put option should wait until the difference of the underlying assets falls to a certain level $\mathcal{L}$ before exercising.

By the discussion above, we divide the set $\{(t, x, y) : 0 \leq t \leq T, x \geq 0, y \geq 0\}$ into two regions, once the difference of the underlying asset prices falls into the stopping region $\mathcal{S}$, the owner of the spread put option should exercise it immediately to obtain the immediate payoff value. Thus the stopping region can be characterized as

$$\mathcal{S} = \{(t, x, y, r) : V(t, x, y, r) = (K - (x - y))^+\}, \tag{6.3}$$

On the other hand, if $S(t)$ remains in the continuation region $\mathcal{C}$, the owner should wait and then the value of the American spread put option is the same as that of the corresponding European spread put option. Thus the continuation region can be characterized as

$$\mathcal{C} = \{(t, x, y, r) : V(t, x, y, r) = \text{the corresponding European spread put option}\}, \tag{6.4}$$

The level $\mathcal{L}$ mentioned in the optimal exercise policy is the free boundary between $\mathcal{S}$ and $\mathcal{C}$ which is denoted by

$$\mathcal{L} : L(T - t, x, y) = 0. \tag{6.5}$$

We have already known that the American spread put option is equal to the intrinsic value in the stopping region. So in the next section, we will derive the American spread put option in the continuation region, which is equal to the corresponding European spread put option.

6.2 The Corresponding European Spread Put Option

In this section, we will derive the partial differential equation for the corresponding European spread put option with stochastic interest rate using the delta-hedging method. In the first part of this section, we will set up the stochastic differential equations for the two underlying asset and the interest rate. In the second part, we will find a hedging portfolio for the American spread put option and derive the partial differential equation for the option.

Assume that the underlying assets $S_1(t), S_2(t)$ satisfy the following stochastic differential equations

$$dS_1(t) = r(t)S_1(t)dt + \sigma_1 S_1(t)d\widetilde{W_1}(t) \tag{6.6}$$

$$dS_2(t) = r(t)S_2(t)dt + \sigma_2 S_2(t)d\widetilde{W_2}(t), \tag{6.7}$$

where the volatilities σ_1, σ_2 are positive constants and $\widetilde{W_1}(t), \widetilde{W_2}(t)$ are two Brownian motions under the risk-neutral measure. The stochastic interest rate $r(t)$ satisfies one of the following models under the risk-neutral measure:

- Vasicek model

$$dr(t) = (\alpha - \beta r(t))dt + \sigma_3 d\widetilde{W_3}(t) \tag{6.8}$$

- CIR model

$$dr(t) = (\alpha - \beta r(t))dt + \sigma_3 \sqrt{r(t)}d\widetilde{W_3}(t) \tag{6.9}$$

where the constants α, β, σ_3 are positive and $\widetilde{W_3}(t)$ is a Brownian motion under the risk-neutral measure.

The three Brownian motions are assumed to have the following cross variation relations

$$d\widetilde{W}_1(t)d\widetilde{W}_2(t) = \rho_1 dt \tag{6.10}$$

$$d\widetilde{W}_1(t)d\widetilde{W}_3(t) = \rho_2 dt \tag{6.11}$$

$$d\widetilde{W}_2(t)d\widetilde{W}_3(t) = \rho_3 dt, \tag{6.12}$$

where the correlations $0 \le \rho_1, \rho_2, \rho_3 \le 1$.

We want to find a hedging portfolio which is of the form:

$$X(t) = \Delta_1(t)S_1(t)+\Delta_2(t)S_2(t)+\Delta_3(t)B(t,T)+[X(t)-\Delta_1(t)S_1(t)-\Delta_2(t)S_2(t)-\Delta_3(t)B(t,T)],$$

such that $X(T) = V(T)$. Here we invest $\Delta_1(t)$ share of stock $S_1(t)$, $\Delta_2(t)$ share of stock $S_2(t)$ and $\Delta_3(t)$ share of zero-coupon bond $B(t,T)$ and invest the rest in the money market account. We also assume that the portfolio $X(t)$ is self-financing, i.e., $dX(t) = \Delta_1(t)dS_1(t)+\Delta_2(t)dS_2(t)+\Delta_3(t)dB(t,T)+r(t)[X(t)-\Delta_1(t)S_1(t)-\Delta_2(t)S_2(t)-\Delta_3(t)B(t,T)]dt$. Moreover, (3.18) guarantees that the value of the hedging portfolio is equal to the value of the option at any time before the expiration date, i.e., $X(t) = V(t), \forall 0 \le t \le T$.

With the preparation above, we can derive the partial differential equation for the put option in three steps:

Step 1 We compute the evolution of the portfolio under the Vasicek model and the CIR model.

For the Vasicek model,

$$
\begin{aligned}
dX(t) &= \Delta_1(t)dS_1(t) + \Delta_2(t)dS_2(t) + \Delta_3(t)dB(t,T) \\
&\quad + r(t)[X(t) - \Delta_1(t)S_1(t) - \Delta_2(t)S_2(t) - \Delta_3(t)B(t,T)]dt \\
&= \Delta_1(t)[r(t)S_1(t)dt + \sigma_1 S_1(t)d\widetilde{W}_1(t)] \\
&\quad + \Delta_2(t)[r(t)S_2(t)dt + \sigma_2 S_2(t)d\widetilde{W}_2(t)] \\
&\quad + \Delta_3(t)[r(t)B(t,T)dt + \sigma_V B(t,T)d\widetilde{W}_3(t)] \\
&\quad + r(t)[X(t) - \Delta_1(t)S_1(t) - \Delta_2(t)S_2(t)d\widetilde{W}_2(t) - \Delta_3(t)B(t,T)]dt \\
&= r(t)X(t)dt + \Delta_1(t)\sigma_1 S_1(t)d\widetilde{W}_1(t) \\
&\quad + \Delta_2(t)\sigma_2 S_2(t)d\widetilde{W}_2(t) + \Delta_3(t)\sigma_V B(t,T)d\widetilde{W}_3(t).
\end{aligned}
$$

Here we used (4.25) in the second equation above.

Replace $X(t)$ by $V(t)$, we have

$$
\begin{aligned}
dV(t) &= r(t)V(t)dt \\
&\quad + \Delta_1(t)\sigma_1 S_1(t)d\widetilde{W}_1(t) + \Delta_2(t)\sigma_2 S_2(t)d\widetilde{W}_2(t) + \Delta_3(t)\sigma_V B(t,T)d\widetilde{W}_3(t). \quad (6.13)
\end{aligned}
$$

For the CIR model,

$$
\begin{aligned}
dX(t) &= \Delta_1(t)dS_1(t) + \Delta_2(t)dS_2(t) + \Delta_3(t)dB(t,T) \\
&\quad + r(t)[X(t) - \Delta_1(t)S_1(t) - \Delta_2(t)S_2(t) - \Delta_3(t)B(t,T)]dt \\
&= \Delta_1(t)[r(t)S_1(t)dt + \sigma_1 S_1(t)d\widetilde{W}_1(t)] \\
&\quad + \Delta_2(t)[r(t)S_2(t)dt + \sigma_2 S_2(t)d\widetilde{W}_2(t)] \\
&\quad + \Delta_3(t)[r(t)B(t,T)dt + \sigma_V B(t,T)d\widetilde{W}_3(t)] \\
&\quad + r(t)[X(t) - \Delta_1(t)S_1(t) - \Delta_2(t)S_2(t)d\widetilde{W}_2(t) - \Delta_3(t)B(t,T)]dt \\
&= r(t)X(t)dt + \Delta_1(t)\sigma_1 S_1(t)d\widetilde{W}_1(t) \\
&\quad + \Delta_2(t)\sigma_2 S_2(t)d\widetilde{W}_2(t) + \Delta_3(t)\sigma_C B(t,T)d\widetilde{W}_3(t).
\end{aligned}
$$

Here we used (4.27) in the second equation above.

Replace $X(t)$ by $V(t)$, we have

$$
\begin{aligned}
dV(t) &= r(t)V(t)dt \\
&\quad + \Delta_1(t)\sigma_1 S_1(t)d\widetilde{W}_1(t) + \Delta_2(t)\sigma_2 S_2(t)d\widetilde{W}_2(t) + \Delta_3(t)\sigma_C B(t,T)d\widetilde{W}_3(t). \quad (6.14)
\end{aligned}
$$

Step 2 By the Itô's formula, we can compute the evolution of the option price $V(t)$:

For the Vasicek model,

$$
\begin{aligned}
dV(t, S(t), r(t)) &= V_t dt + V_{S_1} dS_1(t) + V_{S_2} dS_2(t) + V_r dr(t) \\
&\quad + \frac{1}{2} V_{S_1 S_1} dS_1(t) dS_1(t) + \frac{1}{2} V_{S_2 S_2} dS_2(t) dS_2(t) \\
&\quad + \frac{1}{2} V_{rr} dr(t) dr(t) + V_{S_1 S_2} dS_1(t) dS_2(t) \\
&\quad + V_{S_1 r} dS_1(t) dr(t) + V_{S_2 r} dS_2(t) dr(t) \\
&= V_t dt + V_{S_1} [r(t) S_1(t) dt + \sigma_1 S_1(t) d\widetilde{W}_1(t)] \\
&\quad + V_{S_2} [r(t) S_2(t) dt + \sigma_2 S_2(t) d\widetilde{W}_2(t)] \\
&\quad + V_r [(\alpha - \beta r(t)) dt + \sigma_3 d\widetilde{W}_3(t)] \\
&\quad + \frac{1}{2} \sigma_1^2 S_1^2(t) V_{S_1 S_1} dt + \frac{1}{2} \sigma_2^2 S_2^2(t) V_{S_2 S_2} dt + \frac{1}{2} \sigma_3^2 V_{rr} dt \\
&\quad + \sigma_1 \sigma_2 \rho_1 S_1(t) S_2(t) V_{S_1 S_2} dt + \sigma_1 \sigma_3 \rho_2 S_1(t) V_{S_1 r} dt \\
&\quad + \sigma_2 \sigma_3 \rho_3 S_2(t) V_{S_2 r} dt \\
&= [V_t + r(t) S_1(t) V_{S_1} + r(t) S_2(t) V_{S_2} + (\alpha - \beta r(t)) V_r \\
&\quad + \frac{1}{2} \sigma_1^2 S_1^2(t) V_{S_1 S_1} + \frac{1}{2} \sigma_2^2 S_2^2(t) V_{S_2 S_2} + + \frac{1}{2} \sigma_3^2 V_{rr} \\
&\quad + \sigma_1 \sigma_2 \rho_1 S_1 S_2 V_{S_1 S_2} + \sigma_1 \sigma_3 \rho_2 S_1(t) V_{S_1 r} \\
&\quad + \sigma_2 \sigma_3 \rho_3 S_2(t) V_{S_2 r}] dt + \sigma_1 S_1(t) V_{S_1} d\widetilde{W}_1(t) \\
&\quad + \sigma_2 S_2(t) V_{S_2} d\widetilde{W}_2(t) + \sigma_3 V_r d\widetilde{W}_3(t). \qquad (6.15)
\end{aligned}
$$

For the CIR model,

$$\begin{aligned}
dV(t,S(t),r(t)) &= V_t dt + V_{S_1} dS_1(t) + V_{S_2} dS_2(t) + V_r dr(t) \\
&\quad + \frac{1}{2} V_{S_1S_1} dS_1(t) dS_1(t) + \frac{1}{2} V_{S_2S_2} dS_2(t) dS_2(t) \\
&\quad + \frac{1}{2} V_{rr} dr(t) dr(t) + V_{S_1S_2} dS_1(t) dS_2(t) \\
&\quad + V_{S_1r} dS_1(t) dr(t) + V_{S_2r} dS_2(t) dr(t) \\
&= V_t dt + V_{S_1}[r(t)S_1(t)dt + \sigma_1 S_1(t) d\widetilde{W}_1(t)] \\
&\quad + V_{S_2}[r(t)S_2(t)dt + \sigma_2 S_2(t) d\widetilde{W}_2(t)] \\
&\quad + V_r[(\alpha - \beta r(t))dt + \sigma_3 \sqrt{r(t)} d\widetilde{W}_3(t)] \\
&\quad + \frac{1}{2}\sigma_1^2 S_1^2(t) V_{S_1S_1} dt + \frac{1}{2}\sigma_2^2 S_2^2(t) V_{S_2S_2} dt \\
&\quad + \frac{1}{2}\sigma_3^2 r(t) V_{rr} dt + \sigma_1\sigma_2\rho_1 S_1(t) S_2(t) V_{S_1S_2} dt \\
&\quad + \sigma_1\sigma_3\rho_2 S_1(t)\sqrt{r(t)} V_{S_1r} dt + \sigma_2\sigma_3\rho_3 S_2(t)\sqrt{r(t)} V_{S_2r} dt \\
&= [V_t + r(t)S_1(t)V_{S_1} + r(t)S_2(t)V_{S_2} + (\alpha - \beta r(t))V_r \\
&\quad + \frac{1}{2}\sigma_1^2 S_1^2(t) V_{S_1S_1} + \frac{1}{2}\sigma_2^2 S_2^2(t) V_{S_2S_2} + + \frac{1}{2}\sigma_3^2 r(t) V_{rr} \\
&\quad + \sigma_1\sigma_2\rho_1 S_1 S_2 V_{S_1S_2} + \sigma_1\sigma_3\rho_2 S_1(t)\sqrt{r(t)} V_{S_1r} \\
&\quad + \sigma_2\sigma_3\rho_3 S_2(t)\sqrt{r(t)} V_{S_2r}]dt + \sigma_1 S_1(t) V_{S_1} d\widetilde{W}_1(t) \\
&\quad + \sigma_2 S_2(t) V_{S_2} d\widetilde{W}_2(t) + \sigma_3 V_r d\widetilde{W}_3(t).
\end{aligned} \tag{6.16}$$

Step 3 Equating the evolutions of the portfolio (6.13, 6.14) with the evolutions of the option price (6.15 6.16) respectively to get:

(a)For the Vasicek model,

$$\Delta_1(t) = V_{S_1}(t), \quad \Delta_2(t) = V_{S_2}(t), \quad \Delta_3(t) = \frac{\sigma_3 V_r(t)}{\sigma_V B(t,T)}, \tag{6.17}$$

and

$$
\begin{aligned}
r(t)V(t) = V_t(t) &+ r(t)S_1(t)V_{S_1(t)} + r(t)S_2(t)V_{S_2(t)} + (\alpha - \beta r(t))V_r \\
&+ \frac{1}{2}\sigma_1^2 S_1^2(t)V_{S_1S_1}(t) + \frac{1}{2}\sigma_2^2 S_2^2(t)V_{S_2S_2}(t) + \frac{1}{2}\sigma_3^2 V_{rr}(t) \\
&+ \sigma_1\sigma_2\rho_1 S_1(t)S_2(t)V_{S_1S_2}(t) + \sigma_1\sigma_3\rho_2 S_1(t)V_{S_1r}(t) \\
&+ \sigma_2\sigma_3\rho_3 S_2(t)V_{S_2r}(t).
\end{aligned}
$$

Replacing $S(t), r(t)$ by x, r and simplifying the equation above, we get the desired partial differential equation

$$
\begin{aligned}
rV = V_t + rxV_x + ryV_y + (\alpha - \beta r)V_r + \frac{1}{2}\sigma_1^2 x^2 V_{xx} \\
+ \frac{1}{2}\sigma_2^2 y^2 V_{yy} + \frac{1}{2}\sigma_3^2 V_{rr} + \sigma_1\sigma_2\rho_1 xyV_{xy} + \sigma_1\sigma_3\rho_2 xV_{xr} + \sigma_2\sigma_3\rho_3 yV_{yr}.
\end{aligned}
\tag{6.18}
$$

(b)For the CIR model,

$$
\Delta_1(t) = V_{S_1}(t), \quad \Delta_2(t) = V_{S_2}(t), \quad \Delta_3(t) = \frac{\sigma_3\sqrt{r}V_r(t)}{\sigma_C B(t,T)}, \tag{6.19}
$$

and

$$
\begin{aligned}
r(t)V(t) = V_t(t) &+ r(t)S_1(t)V_{S_1(t)} + r(t)S_2(t)V_{S_2(t)} + (\alpha - \beta r(t))V_r \\
&+ \frac{1}{2}\sigma_1^2 S_1^2(t)V_{S_1S_1}(t) + \frac{1}{2}\sigma_2^2 S_2^2(t)V_{S_2S_2}(t) + \frac{1}{2}\sigma_3^2 r(t)V_{rr}(t) \\
&+ \sigma_1\sigma_2\rho_1 S_1(t)S_2(t)V_{S_1S_2}(t) + \sigma_1\sigma_3\rho_2 S_1(t)\sqrt{r(t)}V_{S_1r}(t) \\
&+ \sigma_2\sigma_3\rho_3 S_2(t)\sqrt{r(t)}V_{S_2r}(t).
\end{aligned}
$$

Replacing $S(t), r(t)$ by x, r and simplifying the equation above, we get the desired partial differential equation

$$rV = V_t + rxV_x + ryV_y + (\alpha - \beta r)V_r + \frac{1}{2}\sigma_1^2 x^2 V_{xx} + \frac{1}{2}\sigma_2^2 y^2 V_{yy} + \frac{1}{2}\sigma_3^2 r V_{rr} + \sigma_1\sigma_2\rho_1 xyV_{xy} + \sigma_1\sigma_3\rho_2 x\sqrt{r}V_{xr} + \sigma_2\sigma_3\rho_3 y\sqrt{r}V_{yr}. \quad (6.20)$$

6.3 Free Boundary Problem of American Spread Put Option

In this section, we formulate the free boundary problem of the American spread put option with stochastic interest rate. According to the optimal exercise policy, we know there exists an early exercise boundary $\mathcal{L}$ such that the American spread put option needs to be exercised immediately if the underlying asset price falls to $\mathcal{L}$. Although the boundary $\mathcal{L}$ is unknown, we can derive the conditions for the value of the option in the stopping region and the continuation region respectively. Furthermore, we can give the terminal and boundary conditions for the American spread put option.

By the optimal exercise policy of the American spread put option, the set $\{(t, x, y, r) : 0 \leq t \leq T, x \geq 0, y \geq 0, r \geq 0\}$ can be divided into two regions, the stopping region

$$\mathcal{S} = \{(t, x, y, r) : V(t, x, y, r) = (K - (x - y))^+\}, \quad (6.21)$$

and the continuation region

$$\mathcal{C} = \{(t, x, y, , r) : V(t, x, y, r) = \text{the corresponding European spread put option}\}. \quad (6.22)$$

The free boundary between $\mathcal{S}$ and $\mathcal{C}$ is the one mentioned in the optimal exercise policy

$$\mathcal{L} : L(T - t, x, y, r) = 0. \quad (6.23)$$

Correspondingly, it is optimal to exercise immediately to obtain the intrinsic value when (t, x, y, r) is in the stopping region $\mathcal{S}$, while in the continuation region $\mathcal{C}$, it is optimal to hold the American spread put option. Thus in the continuation region, the American spread put option price is equal to the corresponding European spread put option price and satisfies the partial differential equations we derived in the previous section.

Next we consider the terminal and boundary conditions for the American spread put option. Firstly, the payoff on the expiration date is the terminal condition

$$V(T, x, y, r) = (K - (x - y))^+. \tag{6.24}$$

Secondly, we assume that the option price V and the derivatives V_x, V_y, V_r are continuous. Since $V(t, x, y, r) = K - (x - y)$ is the option price in the stopping region $\mathcal{S}$, the left-hand derivative with respect to x on $\mathcal{L}$ is $V_x(t, x-, y, r) = -1$, the left-hand derivative with respect to y on $\mathcal{L}$ is $V_y(t, x, y-, r) = 1$ and the left-hand derivative with respect to r on $\mathcal{L}$ is $V_r(t, x, y, r-) = 0$. Then the option price satisfies the smooth-pasting conditions

$$V_x(t_0, x_0, y_0, r_0) = \lim_{x \to x_0^+} V_x(t_0, x, y_0, r_0) = \lim_{x \to x_0^-} V_x(t_0, x, y_0, r_0) = -1, \quad \forall (t_0, x_0, y_0, r_0) \in \mathcal{L} \tag{6.25}$$

$$V_y(t_0, x_0, y_0, r_0) = \lim_{y \to y_0^+} V_x(t_0, x_0, y, r_0) = \lim_{y \to y_0^-} V_y(t_0, x_0, y, r_0) = 1, \quad \forall (t_0, x_0, y_0, r_0) \in \mathcal{L} \tag{6.26}$$

$$V_r(t_0, x_0, y_0, r_0) = \lim_{r \to r_0^+} V_r(t_0, x_0, y_0, r) = \lim_{r \to r_0^-} V_r(t_0, x_0, y_0, r) = 0, \quad \forall (t_0, x_0, y_0, r_0) \in \mathcal{L}. \tag{6.27}$$

Furthermore, we have the asymptotic conditions

$$\lim_{x\to\infty} V(t,x,y,r) = 0 \tag{6.28}$$

$$\lim_{x\to 0} V(t,x,y,r) = (K+y)^+ \tag{6.29}$$

$$\lim_{y\to\infty} V(t,x,y,r) = y \tag{6.30}$$

$$\lim_{y\to 0} V(t,x,y,r) = AMput(t,x,r) \tag{6.31}$$

$$\lim_{r\to\infty} V(t,x,y,r) = 0. \tag{6.32}$$

Equation (6.28) is set due to the fact that if the first underlying asset price is large then the spread put option is unlikely to be exercised and gains no value. Equation (6.30) is set because if the second underlying asset is large, the option price will be dominated by it. Equation (6.29) is reasonable because as x approaches 0, the difference of the two underlying assets $x - y$ is negative and is in the stopping region. In Equation (6.31), $AMput(t,x,r)$ is the American put option price with only one underlying asset, which is the result of the previous chapter.

Finally, if the interest rate is approaching zero, we can find the boundary condition by letting $r = 0$ in Equation (6.18), Equation (6.20) and directly solving the partial differential equations. This will be done after we simplify the partial differential equations.

6.4 Simplification of the Free Boundary Problem

In this section, we will simplify the partial differential equations (6.18) and (6.20) by change of variables. We will then reformulate the free boundary problem of the American spread put option using the simplified partial differential equations.

Let $\tau = T - t$, $x = e^u$, $y = e^v$ and $h(\tau,u,v,r) = V(t,x,y,r)$. Then the partial derivatives of f can be expressed as

$$V_t = h_\tau \frac{d\tau}{dt} = -h_\tau \tag{6.33}$$

$$V_x = h_u \frac{du}{dx} = \frac{1}{x} h_u \tag{6.34}$$

$$V_y = h_v \frac{dv}{dy} = \frac{1}{y} h_v \tag{6.35}$$

$$V_r = h_r \tag{6.36}$$

$$V_{rr} = h_{rr} \tag{6.37}$$

$$V_{xr} = \frac{1}{x} h_{ur} \tag{6.38}$$

$$V_{yr} = \frac{1}{y} h_{vr} \tag{6.39}$$

$$V_{xy} = \frac{1}{xy} h_{uv} \tag{6.40}$$

$$V_{xx} = \frac{\partial}{\partial x}(\frac{1}{x} h_u) = -\frac{1}{x^2} h_u + \frac{1}{x} h_{uu} \frac{du}{dx} = -\frac{1}{x^2} h_u + \frac{1}{x^2} h_{uu} \tag{6.41}$$

$$V_{yy} = \frac{\partial}{\partial y}(\frac{1}{y} h_v) = -\frac{1}{y^2} h_v + \frac{1}{y} h_{vv} \frac{dv}{dy} = -\frac{1}{y^2} h_v + \frac{1}{y^2} h_{vv}. \tag{6.42}$$

Plugging (6.33)–(6.42) to Equation (6.18), we have the the simplified partial differential equation for the Vasicek model

$$\begin{aligned} rh + h_\tau = {} & (r - \frac{1}{2}\sigma_1^2) h_u + (r - \frac{1}{2}\sigma_2^2) h_v + (\alpha - \beta r) h_r \\ & + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_2^2 h_{vv} + \frac{1}{2}\sigma_3 h_{rr} + \sigma_1\sigma_2\rho_1 h_{uv} + \sigma_1\sigma_3\rho_2 h_{ur} + \sigma_2\sigma_3\rho_3 h_{vr}. \end{aligned} \tag{6.43}$$

We let

$$\begin{aligned} \mathcal{Q}h := rh + h_\tau - \Big[& (r - \frac{1}{2}\sigma_1^2) h_u + (r - \frac{1}{2}\sigma_2^2) h_v + (\alpha - \beta r) h_r \\ & + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_2^2 h_{vv} + \frac{1}{2}\sigma_3 h_{rr} + \sigma_1\sigma_2\rho_1 h_{uv} + \sigma_1\sigma_3\rho_2 h_{ur} + \sigma_2\sigma_3\rho_3 h_{vr} \Big]. \end{aligned}$$

Plugging (6.33)–(6.42) to Equation (6.20), we have the the simplified partial differential equation for the CIR model

$$rh + h_\tau = (r - \frac{1}{2}\sigma_1^2)h_u + (r - \frac{1}{2}\sigma_2^2)h_v + (\alpha - \beta r)h_r \\ + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_2^2 h_{vv} + \frac{1}{2}\sigma_3 r h_{rr} + \sigma_1\sigma_2\rho_1 h_{uv} + \sigma_1\sigma_3\rho_2\sqrt{r}h_{ur} + \sigma_2\sigma_3\sqrt{r}\rho_3 h_{vr}. \tag{6.44}$$

We let

$$\mathcal{Q}h := rh + h_\tau - \Big[(r - \frac{1}{2}\sigma_1^2)h_u + (r - \frac{1}{2}\sigma_2^2)h_v + (\alpha - \beta r)h_r \\ + \frac{1}{2}\sigma_1^2 h_{uu} + \frac{1}{2}\sigma_2^2 h_{vv} + \frac{1}{2}\sigma_3 r h_{rr} + \sigma_1\sigma_2\rho_1 h_{uv} + \sigma_1\sigma_3\rho_2\sqrt{r}h_{ur} + \sigma_2\sigma_3\rho_3\sqrt{r}h_{vr}\Big].$$

Now we formulate the free boundary problem associated with the function $h(\tau, u, v, r)$. We firstly make the following change of variables for the American spread put option.

$$V(t, x, y, r) = h(\tau, u, v, r) \tag{6.45}$$

$$\tau = T - t \tag{6.46}$$

$$x = e^u \tag{6.47}$$

$$y = e^v. \tag{6.48}$$

We define the payoff function associated with $h(\tau, u, v, r)$ as

$$g(\tau, u, v, r) := (K - (e^u - e^v))^+. \tag{6.49}$$

We state the free boundary problem associated with the function $h(\tau, u, v, r)$ as follows: In the stopping region $\mathcal{S}$

$$\mathcal{Q}h > 0 \tag{6.50}$$

$$h(\tau, u, v, r) = g(\tau, u, v, r) = \frac{(K - (e^u - e^v))^+}{B(\tau, r)}. \tag{6.51}$$

In the continuation region $\mathcal{C}$

$$\mathcal{Q}h = 0 \tag{6.52}$$

$$h(\tau,u,v,r) > g(\tau,u,v,r) = \frac{(K-(e^u-e^v))^+}{B(\tau,r)}. \tag{6.53}$$

The initial condition is

$$h(0,u,v,r) = g(0,u,v,r) = (K-(e^u-e^v))^+. \tag{6.54}$$

The smooth-pasting conditions are

$$h_u(\tau_0,u_0,v_0,r_0) = \lim_{u\to u_0^+} h_u(\tau_0,u,v_0,r_0) = \lim_{u\to u_0^-} h_u(\tau_0,u,v_0,r_0) = -e^{u_0}, \quad \forall(\tau_0,u_0,v_0,r_0)\in\mathcal{L} \tag{6.55}$$

$$h_v(\tau_0,u_0,v_0,r_0) = \lim_{v\to v_0^+} h_u(\tau_0,u_0,v,r_0) = \lim_{v\to v_0^-} h_v(\tau_0,u_0,v,r_0) = e^{v_0}, \quad \forall(\tau_0,u_0,v_0,r_0)\in\mathcal{L} \tag{6.56}$$

$$h_r(\tau_0,u_0,v_0,r_0) = \lim_{r\to r_0^+} h_r(\tau_0,u_0,v_0,r) = \lim_{r\to r_0^-} h_r(\tau_0,u_0,v_0,r) = 0, \quad \forall(tau_0,u_0,v_0,r_0)\in\mathcal{L}. \tag{6.57}$$

The asymptotic conditions are

$$\lim_{u\to\infty} h(\tau,u,v,r) = 0 \tag{6.58}$$

$$\lim_{u\to-\infty} h(\tau,u,v,r) = K + e^v \tag{6.59}$$

$$\lim_{v\to\infty} h(\tau,u,v,r) = e^v \tag{6.60}$$

$$\lim_{v\to-\infty} h(\tau,u,v,r) = AMput(\tau,u,r) \tag{6.61}$$

$$\lim_{r\to\infty} h(\tau,u,v,r) = 0. \tag{6.62}$$

Finally, if the interest rate $r = 0$, then

$$h(\tau, u, v, 0) = \max\{(K - (e^u - e^v))^+,$$
$$\frac{1}{4\pi\tau} e^{\frac{1}{4}(\sigma_1^2+\sigma_2^2)\tau + \frac{1}{2}(u+v)} \int_{\mathbb{R}^2} e^{-\frac{(\frac{\sqrt{2}u}{\sigma_1} - \xi)^2 + (\frac{\sqrt{2}v}{\sigma_2} - \eta)^2}{4\tau} - \frac{1}{2\sqrt{2}}(\sigma_1\xi + \sigma_2\eta)} \left[K - (e^{\frac{\sigma_1}{\sqrt{2}}\xi} - e^{\frac{\sigma_2}{\sqrt{2}}\eta})\right]^+ d(\xi, \eta)\}. \tag{6.63}$$

The derivation is in Appendix C.

6.5 Linear Complementarity Conditions

In this section, we convert the free boundary problem associated with h in the previous section to linear complementarity conditions for the function h. The linear complementarity conditions helps to go around the free boundary and implement numerical computation for the American spread put option.

In practice, we consider the spatial variable u, v in the interval $-u^- \leq u \leq u^+$ and $-v^- \leq v \leq v^+$, where u^-, u^+, v^-, v^+ are large positive numbers. Similarly, we consider r in the interval $0 \leq r \leq r^+$, where r^+ is a large positive number.

The linear complementarity problem for the function h is:

$$\begin{aligned} &\mathcal{Q}h \cdot (h(\tau, u, v, r) - g(\tau, u, v, r)) = 0 \\ &\mathcal{Q}h \geq 0, \quad h(\tau, u, v, r) \geq g(\tau, u, v, r), \end{aligned} \tag{6.64}$$

with the initial and boundary conditions [1]

$$\begin{aligned}
&h(0,u,v,r) = g(0,u,v,r) = (K-(e^u-e^v))^+\\
&h(\tau,-u^-,v,r) = g(\tau,-u^-,v,r) = \frac{K+e^v}{B(\tau,r)}\\
&h(\tau,u^+,v,r) = 0\\
&h(\tau,u,-v^-,r) = AMput(\tau,u,r)\\
&h(\tau,u,v^+,r) = \frac{e^{v^+}}{B(\tau,r)}\\
&h(\tau,u,v,0) = \max\{e^{A(\tau,T)}(K-(e^u-e^v))^+,\\
&\frac{1}{4\pi\tau}e^{\frac{1}{4}(\sigma_1^2+\sigma_2^2)\tau+\frac{1}{2}(u+v)}\int_{\mathbb{R}^2} e^{-\frac{(\frac{\sqrt{2}u}{\sigma_1}-\xi)^2+(\frac{\sqrt{2}v}{\sigma_2}-\eta)^2}{4\tau}-\frac{1}{2\sqrt{2}}(\sigma_1\xi+\sigma_2\eta)}\left[K-(e^{\frac{\sigma_1}{\sqrt{2}}\xi}-e^{\frac{\sigma_2}{\sqrt{2}}\eta})\right]^+ d(\xi,\eta)\}\\
&h(\tau,u,v,r^+) = 0.
\end{aligned} \tag{6.65}$$

6.6 Finite Difference Method and Implementation

The finite difference discretization of the the partial derivatives of h are as follows:

- Forward difference

$$h_\tau \approx \frac{h^{m+1}_{i,j,k}-h^m_{i,j,k}}{\Delta\tau} + O(\Delta\tau) \tag{6.66}$$

- Central difference

$$h_u \approx \frac{h^m_{i+1,j,k}-h^m_{i-1,j,k}}{2\Delta u} + O((\Delta u)^2) \tag{6.67}$$

$$h_v \approx \frac{h^m_{i,j+1,k}-h^m_{i,j-1,k}}{2\Delta v} + O((\Delta v)^2) \tag{6.68}$$

$$h_r \approx \frac{h^m_{i,j,k+1}-h^m_{i,j,k-1}}{2\Delta r} + O((\Delta r)^2) \tag{6.69}$$

[1]In practice, due to the complexity of the double integral, we use $h(\tau,u,v,0) = (K-(e^u-e^v))^+$ to estimate the the boundary condition when $r=0$.

- Symmetric central difference

$$h_{uu} \approx \frac{h^m_{i+1,j,k} - 2h^m_{i,j,k} + h^m_{i-1,j,k}}{(\Delta u)^2} + O((\Delta u)^2) \tag{6.70}$$

$$h_{vv} \approx \frac{h^m_{i,j+1,k} - 2h^m_{i,j,k} + h^m_{i,j-1,k}}{(\Delta v)^2} + O((\Delta v)^2) \tag{6.71}$$

$$h_{rr} \approx \frac{h^m_{i,j,k+1} - 2h^m_{i,j,k} + h^m_{i,j,k-1}}{(\Delta r)^2} + O((\Delta r)^2) \tag{6.72}$$

$$h_{uv} \approx \frac{h^m_{i+1,j+1,k} - h^m_{i-1,j+1,k} - h^m_{i+1,j-1,k} + h^m_{i-1,j-1,k}}{4\Delta u \Delta v} + O((\Delta u)^2 + (\Delta v)^2) \tag{6.73}$$

$$h_{ur} \approx \frac{h^m_{i+1,j,k+1} - h^m_{i-1,j,k+1} - h^m_{i+1,j,k-1} + h^m_{i-1,j,k-1}}{4\Delta u \Delta r} + O((\Delta u)^2 + (\Delta r)^2) \tag{6.74}$$

$$h_{vr} \approx \frac{h^m_{i,j+1,k+1} - h^m_{i,j-1,k+1} - h^m_{i,j+1,k-1} + h^m_{i,j-1,k-1}}{4\Delta v \Delta r} + O((\Delta v)^2 + (\Delta r)^2) \tag{6.75}$$

Let N, M, W, R be large positive numbers, then

$$\Delta \tau = \frac{T}{N}$$
$$\Delta u = \frac{u^+ + u^-}{M - 1}$$
$$\Delta v = \frac{v^+ + v^-}{W - 1}$$
$$\Delta r = \frac{r^+}{R - 1},$$

and

$$\tau = 0 + m\Delta\tau = m\Delta\tau, \quad 1 \leq m \leq N$$

$$u = -u^- + (i-1)\Delta u, \quad 1 \leq i \leq M$$

$$v = -v^- + (i-1)\Delta v, \quad 1 \leq j \leq W$$

$$r = 0 + (k-1)\Delta r = (k-1)\Delta r, \quad 1 \leq k \leq R.$$

Plugging Equations (6.66)–(6.75) into Equation (6.43), we obtain the discretization of the partial differential equation for the Vasicek model.

$$
\begin{aligned}
rh^m_{i,j,k} + \frac{h^{m+1}_{i,j,k} - h^m_{i,j,k}}{\Delta\tau} + O(\Delta\tau) = (r - \frac{1}{2}\sigma_1^2) \left[\frac{h^m_{i+1,j,k} - h^m_{i-1,j,k}}{2\Delta u} + O((\Delta u)^2)\right] \\
(r - \frac{1}{2}\sigma_2^2) \left[\frac{h^m_{i,j+1,k} - h^m_{i,j-1,k}}{2\Delta v} + O((\Delta v)^2)\right] \\
+ (\alpha - \beta r) \left[\frac{h^m_{i,j,k+1} - h^m_{i,j,k-1}}{2\Delta r} + O((\Delta r)^2)\right] \\
+ \frac{1}{2}\sigma_1^2 \left[\frac{h^m_{i+1,j,k} - 2h^m_{i,j,k} + h^m_{i-1,j,k}}{(\Delta u)^2} + O((\Delta u)^2)\right] \\
+ \frac{1}{2}\sigma_2^2 \left[\frac{h^m_{i,j+1,k} - 2h^m_{i,j,k} + h^m_{i,j-1,k}}{(\Delta v)^2} + O((\Delta v)^2)\right] \\
+ \frac{1}{2}\sigma_3^2 \left[\frac{h^m_{i,j,k+1} - 2h^m_{i,j,k} + h^m_{i,j,k-1}}{(\Delta r)^2} + O((\Delta r)^2)\right] \\
+ \sigma_1\sigma_2\rho_1 \left[\frac{h^m_{i+1,j+1,k} - h^m_{i-1,j+1,k} - h^m_{i+1,j-1,k} + h^m_{i-1,j-1,k}}{4\Delta u\Delta v} + O((\Delta u)^2 + (\Delta v)^2)\right] \\
+ \sigma_1\sigma_3\rho_2 \left[\frac{h^m_{i+1,j,k+1} - h^m_{i-1,j,k+1} - h^m_{i+1,j,k-1} + h^m_{i-1,j,k-1}}{4\Delta u\Delta r} + O((\Delta u)^2 + (\Delta r)^2)\right] \\
+ \sigma_2\sigma_3\rho_3 \left[\frac{h^m_{i,j+1,k+1} - h^m_{i,j-1,k+1} - h^m_{i,j+1,k-1} + h^m_{i,j-1,k-1}}{4\Delta v\Delta r} + O((\Delta v)^2 + (\Delta r)^2)\right]. \quad (6.76)
\end{aligned}
$$

Let

$$\begin{aligned}
A &= \frac{1}{2}\frac{\Delta\tau}{\Delta u}\left(r - \frac{1}{2}\sigma_1^2\right) \\
B &= \frac{1}{2}\frac{\Delta\tau}{\Delta v}\left(r - \frac{1}{2}\sigma_2^2\right) \\
C &= \frac{1}{2}\frac{\Delta\tau}{\Delta r}(\alpha - \beta r) \\
D &= \frac{1}{2}\sigma_1^2\frac{\Delta\tau}{(\Delta u)^2} \\
E &= \frac{1}{2}\sigma_2^2\frac{\Delta\tau}{(\Delta v)^2} \\
F &= \frac{1}{2}\sigma_3^2\frac{\Delta\tau}{(\Delta r)^2} \\
G &= \frac{1}{4}\sigma_1\sigma_2\rho_1\frac{\Delta\tau}{\Delta u\Delta v} \\
H &= \frac{1}{4}\sigma_1\sigma_3\rho_2\frac{\Delta\tau}{\Delta u\Delta r} \\
I &= \frac{1}{4}\sigma_2\sigma_3\rho_3\frac{\Delta\tau}{\Delta v\Delta r}
\end{aligned}$$

Ignoring the higher order terms, we obtain the finite difference scheme of h for the Vasicek model

$$\begin{aligned}
h_{i,j,k}^{m+1} = {} & Gh_{i+1,j+1,k}^m + Hh_{i+1,j,k+1}^m + (A+D)h_{i+1,j,k}^m - Hh_{i+1,j,k-1}^m \\
& - Gh_{i+1,j-1,k}^m + Ih_{i,j+1,k+1}^m + (B+E)h_{i,j+1,k}^m - Ih_{i,j+1,k-1}^m \\
& + (C+F)h_{i,j,k+1}^m + (1 - r\Delta\tau - 2D - 2E - 2F)h_{i,j,k}^m + (-C+F)h_{i,j,k-1}^m \\
& - Ih_{i,j-1,k+1}^m + (-B+E)h_{i,j-1,k}^m + Ih_{i,j-1,k-1}^m - Gh_{i-1,j+1,k}^m \\
& - Hh_{i-1,j,k+1}^m + (-A+D)h_{i-1,j,k}^m + Hh_{i-1,j,k-1}^m + Gh_{i-1,j-1,k}^m \quad (6.77)
\end{aligned}$$

Plugging Equations (6.66)–(6.75) into Equation (6.44), we obtain the discretization of the partial differential equation for the CIR model.

$$
\begin{aligned}
rh^m_{i,j,k} + \frac{h^{m+1}_{i,j,k} - h^m_{i,j,k}}{\Delta\tau} + O(\Delta\tau) = (r - \frac{1}{2}\sigma_1^2)\left[\frac{h^m_{i+1,j,k} - h^m_{i-1,j,k}}{2\Delta u} + O((\Delta u)^2)\right] \\
(r - \frac{1}{2}\sigma_2^2)\left[\frac{h^m_{i,j+1,k} - h^m_{i,j-1,k}}{2\Delta v} + O((\Delta v)^2)\right] \\
+ (\alpha - \beta r)\left[\frac{h^m_{i,j,k+1} - h^m_{i,j,k-1}}{2\Delta r} + O((\Delta r)^2)\right] \\
+ \frac{1}{2}\sigma_1^2\left[\frac{h^m_{i+1,j,k} - 2h^m_{i,j,k} + h^m_{i-1,j,k}}{(\Delta u)^2} + O((\Delta u)^2)\right] \\
+ \frac{1}{2}\sigma_2^2\left[\frac{h^m_{i,j+1,k} - 2h^m_{i,j,k} + h^m_{i,j-1,k}}{(\Delta v)^2} + O((\Delta v)^2)\right] \\
+ \frac{1}{2}\sigma_3^2 r\left[\frac{h^m_{i,j,k+1} - 2h^m_{i,j,k} + h^m_{i,j,k-1}}{(\Delta r)^2} + O((\Delta r)^2)\right] \\
+ \sigma_1\sigma_2\rho_1\left[\frac{h^m_{i+1,j+1,k} - h^m_{i-1,j+1,k} - h^m_{i+1,j-1,k} + h^m_{i-1,j-1,k}}{4\Delta u\Delta v} + O((\Delta u)^2 + (\Delta v)^2)\right] \\
+ \sigma_1\sigma_3\rho_2\sqrt{r}\left[\frac{h^m_{i+1,j,k+1} - h^m_{i-1,j,k+1} - h^m_{i+1,j,k-1} + h^m_{i-1,j,k-1}}{4\Delta u\Delta r} + O((\Delta u)^2 + (\Delta r)^2)\right] \\
+ \sigma_2\sigma_3\rho_3\sqrt{r}\left[\frac{h^m_{i,j+1,k+1} - h^m_{i,j-1,k+1} - h^m_{i,j+1,k-1} + h^m_{i,j-1,k-1}}{4\Delta v\Delta r} + O((\Delta v)^2 + (\Delta r)^2)\right]. \quad (6.78)
\end{aligned}
$$

Let

$$
\begin{aligned}
A &= \frac{1}{2}\frac{\Delta\tau}{\Delta u}\left(r - \frac{1}{2}\sigma_1^2\right)\\
B &= \frac{1}{2}\frac{\Delta\tau}{\Delta v}\left(r - \frac{1}{2}\sigma_2^2\right)\\
C &= \frac{1}{2}\frac{\Delta\tau}{\Delta r}(\alpha - \beta r)\\
D &= \frac{1}{2}\sigma_1^2\frac{\Delta\tau}{(\Delta u)^2}\\
E &= \frac{1}{2}\sigma_2^2\frac{\Delta\tau}{(\Delta v)^2}\\
F &= \frac{1}{2}\sigma_3^2 r\frac{\Delta\tau}{(\Delta r)^2}\\
G &= \frac{1}{4}\sigma_1\sigma_2\rho_1\frac{\Delta\tau}{\Delta u\Delta v}\\
H &= \frac{1}{4}\sigma_1\sigma_3\rho_2\sqrt{r}\frac{\Delta\tau}{\Delta u\Delta r}\\
I &= \frac{1}{4}\sigma_2\sigma_3\rho_3\sqrt{r}\frac{\Delta\tau}{\Delta v\Delta r}
\end{aligned}
$$

Ignoring the higher order terms, we obtain the finite difference scheme of h for the CIR model.

$$
\begin{aligned}
h_{i,j,k}^{m+1} = {} & Gh_{i+1,j+1,k}^m + Hh_{i+1,j,k+1}^m + (A+D)h_{i+1,j,k}^m - Hh_{i+1,j,k-1}^m\\
& - Gh_{i+1,j-1,k}^m + Ih_{i,j+1,k+1}^m + (B+E)h_{i,j+1,k}^m - Ih_{i,j+1,k-1}^m\\
& + (C+F)h_{i,j,k+1}^m + (1 - r\Delta\tau - 2D - 2E - 2F)h_{i,j,k}^m + (-C+F)h_{i,j,k-1}^m\\
& - Ih_{i,j-1,k+1}^m + (-B+E)h_{i,j-1,k}^m + Ih_{i,j-1,k-1}^m - Gh_{i-1,j+1,k}^m\\
& - Hh_{i-1,j,k+1}^m + (-A+D)h_{i-1,j,k}^m + Hh_{i-1,j,k-1}^m + Gh_{i-1,j-1,k}^m \qquad (6.79)
\end{aligned}
$$

To make the notation simpler, we denote the right hand side of Equation (6.77) and Equation (6.79) as $\mathcal{H}h^m$. Then the finite difference scheme of h is abbreviated by

$$h_{i,k}^{m+1} = \mathcal{H}h^m. \qquad (6.80)$$

In order to make h satisfy the linear complementarity conditions, we need to modify the finite difference scheme above to be

$$h_{i,k}^{m+1} = \max\left\{\mathcal{H}h^m, [K - (e^u - e^v)]^+\right\}. \tag{6.81}$$

That is to say for each time step, h takes the maximum between the value of the corresponding European spread put option and the immediate payoff if the early exercise is optimal.

Eventually, by change of variables

$$V(t, x, y, r) = h(\tau, u, v, r)$$
$$\tau = T - t$$
$$x = e^u$$
$$y = e^v.$$

we can get the option price at the current time $V(0)$.

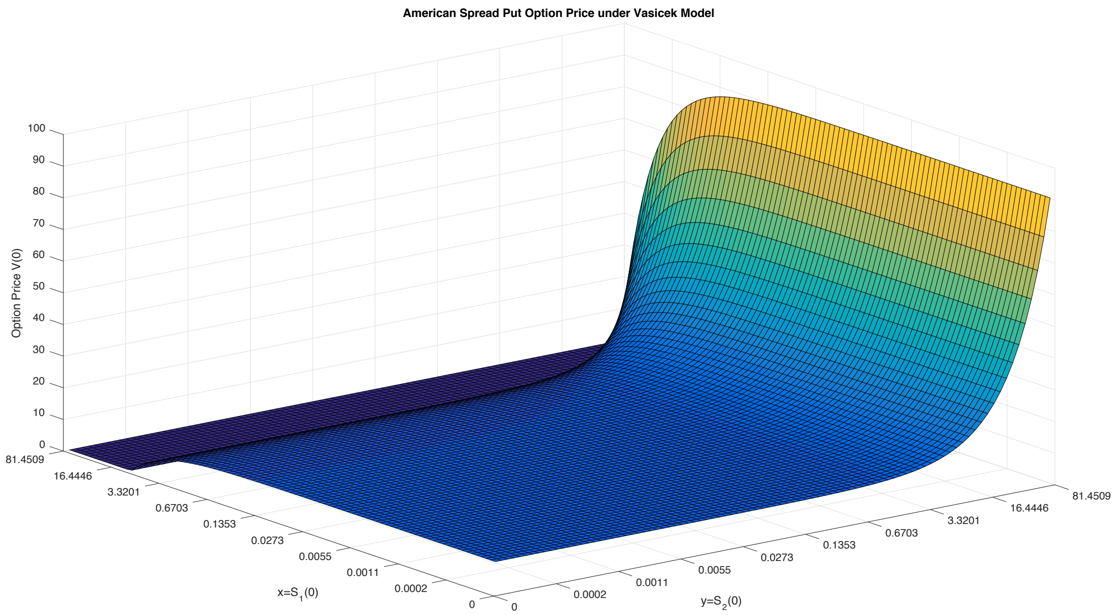

Figure 6.1: American Spread Put Option under Vasicek model (r(0)=0.05)

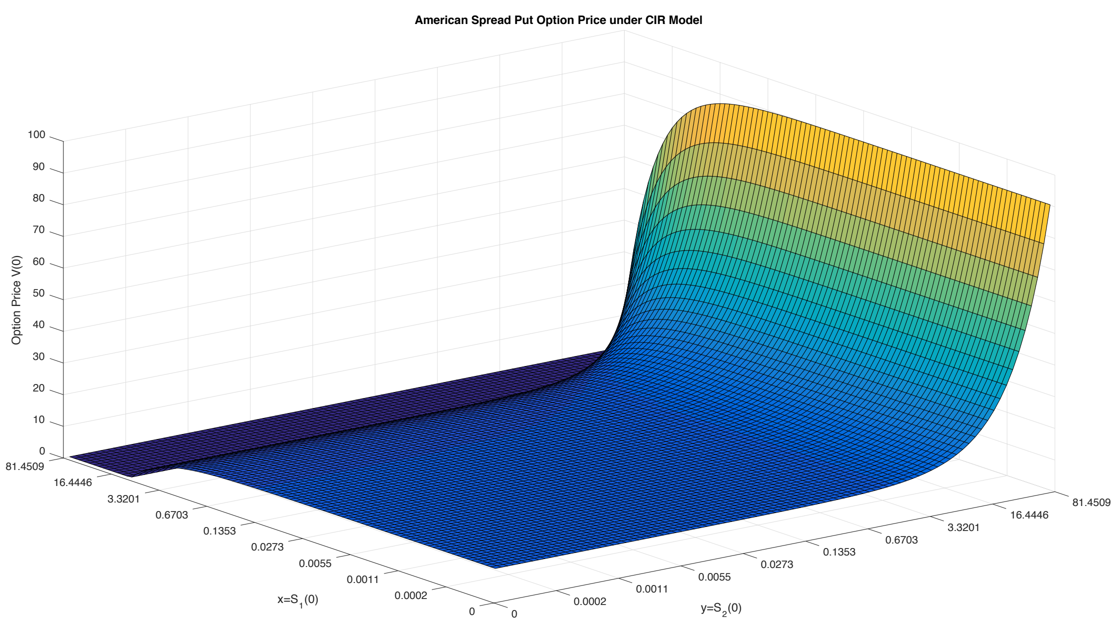

Figure 6.2: American Spread Put Option under CIR model (r(0)=0.05)

Chapter 7. Monte Carlo Methods

In this chapter, we do the numerical computation based on the Monte Carlo simulation for the American spread put option with stochastic interest rate. In section 7.1, we compute numerically the stochastic interest rate $r(t)$ and the underlying assets $S_1(t), S_2(t)$. In section 7.2, we apply the regression-based method in [17] to get an upper bound of the option price and apply the Longstaff and Schwartz method introduced in [31] to get a lower bound of the option price. In section 7.3, we use the dual method to get a tighter upper bound of the option price, so that we have a smaller estimation range for the option price.

7.1 Simulation of $r(t)$, $S_1(t)$ and $S_2(t)$

In this section, we compute numerically the stochastic interest rate $r(t)$ and the underlying assets $S_1(t)$, $S_2(t)$. All the simulations in this section will be done under the risk-neutral measure.

We divide the time interval $0 \leq t \leq T$ into $N-1$ subintervals with equal length $\Delta t = \frac{T}{N-1}$. Denote the endpoints as $t_1, t_2, ..., t_N$. Let $r_i := r(t_i)$, $i = 1, ..., N$ and apply the Euler scheme for the interest rate processes.

For the Vasicek model

$$dr(t) = (\alpha - \beta r(t))dt + \sigma_3\rho_2 d\widetilde{B_1}(t) + \sigma_3\rho_4 d\widetilde{B_2}(t) + \sigma_3\rho_5 d\widetilde{B_3}(t), \tag{7.1}$$

its Euler scheme is

$$r_{i+1} = r_i + (\alpha - \beta r_i)\Delta t + \sigma_3\rho_2\Delta\widetilde{B_1}(t_i) + \sigma_3\rho_4\Delta\widetilde{B_2}(t_i) + \sigma_3\rho_5\Delta\widetilde{B_3}(t_i). \tag{7.2}$$

For the CIR model

$$dr(t) = (\alpha - \beta r(t))dt + \sigma_3\rho_2\sqrt{r(t)}d\widetilde{B_1}(t) + \sigma_3\rho_4\sqrt{r(t)}d\widetilde{B_2}(t) + \sigma_3\rho_5\sqrt{r(t)}d\widetilde{B_3}(t), \tag{7.3}$$

its Euler scheme is

$$r_{i+1} = r_i + (\alpha - \beta r_i)\Delta t + \sigma_3\rho_2\sqrt{r_i}\Delta\widetilde{B_1}(t_i) + \sigma_3\rho_4\sqrt{r_i}\Delta\widetilde{B_2}(t_i) + \sigma_3\rho_5\sqrt{r_i}\Delta\widetilde{B_3}(t_i). \qquad (7.4)$$

Here $i = 1, ..., N-1$. $\Delta B(t_i) = B(t_{i+1}) - B(t_i) = \sqrt{\Delta t}w$ follows a normal distribution with mean 0 and variance Δt, w is a standard normal random variable.

A way to refine the Euler scheme is the Milstein scheme which requires to use the Itô's formula to expand the diffusion term to get a higher order accuracy. It does not improve the Vasicek model because the diffusion term is constant and has no way to be expanded. Therefore we will only consider the case of the CIR model.

The integral form of the CIR model is

$$\begin{aligned} r(t_{i+1}) = r(t_i) &+ \int_{t_i}^{t_{i+1}} (\alpha - \beta r(t))dt \\ &+ \sigma_3\rho_2 \int_{t_i}^{t_{i+1}} \sqrt{r(t)}d\widetilde{B_1}(t) + \sigma_3\rho_4 \int_{t_i}^{t_{i+1}} \sqrt{r(t)}d\widetilde{B_2}(t) + \sigma_3\rho_5 \int_{t_i}^{t_{i+1}} \sqrt{r(t)}d\widetilde{B_3}(t) \qquad (7.5) \end{aligned}$$

Instead of taking the left-end point approximation for the integrals in the Euler scheme, we are going to find a better approximation for $\sqrt{r(t)}$ on the interval $[t_i, t_{i+1}]$.

Let $f(x) = \sqrt{x}$, then $f'(x) = \frac{1}{2\sqrt{x}}$ and $f''(x) = -\frac{1}{4x^{3/2}}$. By the Itô's formula we obtain

$$\begin{aligned} d(\sqrt{r(t)}) &= df(r(t)) = f'(r(t))dr(t) + \frac{1}{2}f''(r(t))(dr(t))^2 \\ &= \frac{1}{2\sqrt{r(t)}}dr(t) - \frac{\sigma_3^2 r(t)}{8(r(t))^{3/2}}(\rho_2^2 + \rho_4^2 + \rho_5^2)dt \\ &= \left[\frac{\alpha - \beta r(t)}{2\sqrt{r(t)}} - \frac{\sigma_3^2}{8\sqrt{r(t)}}\right]dt + \frac{\sigma_3}{2\sqrt{r(t)}}(\rho_2 d\widetilde{B_1}(t) + \rho_4 d\widetilde{B_2}(t) + \rho_5 d\widetilde{B_3}(t)) \end{aligned}$$

Integrate both sides on $[t_i, t],\ t_i \le t \le t_{i+1}$ and apply the Euler scheme to obtain

$$\sqrt{r(t)} = \sqrt{r(t_i)} + \left[\frac{\alpha - \beta r(t_i)}{2\sqrt{r(t_i)}} - \frac{\sigma_3^2}{8\sqrt{r(t_i)}}\right](t - t_i)$$
$$+ \frac{\sigma_3}{2\sqrt{r(t_i)}}\left[\rho_2(\widetilde{B_1}(t) - \widetilde{B_1}(t_i)) + \rho_4(\widetilde{B_2}(t) - \widetilde{B_2}(t_i)) + \rho_5(\widetilde{B_3}(t) - \widetilde{B_3}(t_i))\right] \quad (7.6)$$

We get rid of the second term above because it is of order $O(t - t_i)$ while the stochastic term is just of order $O(\sqrt{t - t_i})$. Thus we have

$$\sqrt{r(t)} = \sqrt{r(t_i)}$$
$$+ \frac{\sigma_3}{2\sqrt{r(t_i)}}\left[\rho_2(\widetilde{B_1}(t) - \widetilde{B_1}(t_i)) + \rho_4(\widetilde{B_2}(t) - \widetilde{B_2}(t_i)) + \rho_5(\widetilde{B_3}(t) - \widetilde{B_3}(t_i))\right] \quad (7.7)$$

Then

$$\int_{t_i}^{t_{i+1}} \sqrt{r(t)}\, d\widetilde{B_1}(t) = \int_{t_i}^{t_{i+1}} \Bigg\{ \sqrt{r(t_i)}$$
$$+ \frac{\sigma_3}{2\sqrt{r(t_i)}}\left[\rho_2(\widetilde{B_1}(t) - \widetilde{B_1}(t_i)) + \rho_4(\widetilde{B_2}(t) - \widetilde{B_2}(t_i)) + \rho_5(\widetilde{B_3}(t) - \widetilde{B_3}(t_i))\right]\Bigg\} d\widetilde{B_1}(t) \quad (7.8)$$

The mixed integrals, i.e.,

$$\int_{t_i}^{t_{i+1}} (\widetilde{B_2}(t) - \widetilde{B_2}(t_i))\, d\widetilde{B_1}(t) \text{ and } \int_{t_i}^{t_{i+1}} (\widetilde{B_3}(t) - \widetilde{B_3}(t_i))\, d\widetilde{B_1}(t) \quad (7.9)$$

are called Levy area terms which are hard to simulate. Thus we get rid of the Levy area terms to obtain

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} \sqrt{r(t)}\, d\widetilde{B_1}(t) &= \sqrt{r(t_i)}\Delta\widetilde{B_1} + \frac{\sigma_3\rho_2}{2\sqrt{r(t_i)}} \int_{t_i}^{t_{i+1}} (\widetilde{B_1}(t) - \widetilde{B_1}(t_i))\, d\widetilde{B_1}(t) \\
&= \sqrt{r(t_i)}\Delta\widetilde{B_1} + \frac{\sigma_3\rho_2}{2\sqrt{r(t_i)}} \left(\frac{1}{2}\Delta\widetilde{B_1}^2 - \frac{1}{2}\Delta t\right) \\
&= -\frac{\sigma_3\rho_2}{4\sqrt{r(t_i)}}\Delta t + \sqrt{r(t_i)}\Delta\widetilde{B_1} + \frac{\sigma_3\rho_2}{4\sqrt{r(t_i)}}\Delta\widetilde{B_1}^2
\end{aligned} \quad (7.10)$$

Similarly we have

$$\int_{t_i}^{t_{i+1}} \sqrt{r(t)} d\widetilde{B_2}(t) = -\frac{\sigma_3\rho_4}{4\sqrt{r(t_i)}}\Delta t + \sqrt{r(t_i)}\Delta\widetilde{B_2} + \frac{\sigma_3\rho_4}{4\sqrt{r(t_i)}}\Delta\widetilde{B_2}^2 \tag{7.11}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{r(t)} d\widetilde{B_3}(t) = -\frac{\sigma_3\rho_5}{4\sqrt{r(t_i)}}\Delta t + \sqrt{r(t_i)}\Delta\widetilde{B_3} + \frac{\sigma_3\rho_5}{4\sqrt{r(t_i)}}\Delta\widetilde{B_3}^2 \tag{7.12}$$

Plugging the three integrals above to Equation (7.5) and apply the left-end point approximation to the drift term, we get the Milstein scheme for the CIR interest rate model:

$$\begin{aligned} r_{i+1} = r_i &+ \left(\alpha - \beta r_i - \frac{\sigma_3^2}{4\sqrt{r_i}}\right)\Delta t \\ &+ \sigma_3\sqrt{r_i}(\rho_2\Delta\widetilde{B_1} + \rho_4\Delta\widetilde{B_2} + \rho_5\Delta\widetilde{B_3}) + \frac{\sigma_3^2}{4\sqrt{r_i}}(\rho_2^2\Delta\widetilde{B_1}^2 + \rho_4^2\Delta\widetilde{B_2}^2 + \rho_5^2\Delta\widetilde{B_3}^2) \end{aligned} \tag{7.13}$$

For the underlying assets

$$dS_1(t) = r(t)S_1(t)dt + \sigma_1 S_1(t)d\widetilde{B_1}(t) \tag{7.14}$$

$$dS_2(t) = r(t)S_2(t)dt + \sigma_2\rho_1 S_2(t)d\widetilde{B_1}(t) + \sigma_2\sqrt{1-\rho_1^2}S_2(t)d\widetilde{B_2}(t). \tag{7.15}$$

We can solve the stochastic differential equations explicitly. In fact, let $f(x) = \ln x$ then

$f'(x) = \frac{1}{x}, f''(x) = -\frac{1}{x^2}$. Apply Itô's formula,

$$\begin{aligned}
d(f(S_1(t))) &= f'(S_1(t))dS_1 + \frac{1}{2}f''(S_1(t))dS_1dS_1 \\
&= \frac{1}{S_1(t)}[r(t)S_1(t)dt + \sigma_1 S_1(t)d\widetilde{B_1}(t)] + \frac{1}{2}\left(-\frac{1}{S_1^2(t)}\right)\sigma_1^2 S_1^2(t)dt \\
&= \left(r(t) - \frac{1}{2}\sigma_1^2\right)dt + \sigma_1 d\widetilde{B_1}(t)
\end{aligned}$$

$$\begin{aligned}
d(f(S_2(t))) &= f'(S_2(t))dS_2 + \frac{1}{2}f''(S_2(t))dS_2dS_2 \\
&= \frac{1}{S_2(t)}\left[r(t)S_2(t)dt + \sigma_2\rho_1 S_2(t)d\widetilde{B_1}(t) + \sigma_2\sqrt{1-\rho_1^2}S_2(t)d\widetilde{B_2}(t)\right] \\
&\quad + \frac{1}{2}\left(-\frac{1}{S_2^2(t)}\right)\left(\sigma_2^2\rho_1^2 S_2^2(t)dt + \sigma_2^2(1-\rho_1^2)S_2^2(t)dt\right) \\
&= \left(r(t) - \frac{1}{2}\sigma_2^2\right)dt + \sigma_2\rho_1 d\widetilde{B_1}(t) + \sigma_2\sqrt{1-\rho_1^2}d\widetilde{B_2}(t).
\end{aligned}$$

Apply the integral

$$\begin{aligned}
\ln S_1(t) &= \ln S_1(0) + \int_0^t \left(r(s) - \frac{1}{2}\sigma_1^2\right)ds + \int_0^t \sigma_1 d\widetilde{B_1}(s) \\
\ln S_2(t) &= \ln S_2(0) + \int_0^t \left(r(s) - \frac{1}{2}\sigma_2^2\right)ds + \int_0^t \sigma_2\rho_1 d\widetilde{B_1}(s) + \int_0^t \sigma_2\sqrt{1-\rho_1^2}d\widetilde{B_2}(s),
\end{aligned}$$

and then take the exponential to both sides to obtain

$$S_1(t) = S_1(0)\exp\left\{\int_0^t r(s)ds - \frac{1}{2}\sigma_1^2 t + \sigma_1\widetilde{B_1}(t)\right\} \tag{7.16}$$

$$S_2(t) = S_2(0)\exp\left\{\int_0^t r(s)ds - \frac{1}{2}\sigma_2^2 t + \sigma_2\rho_1\widetilde{B_1}(t) + \sigma_2\sqrt{1-\rho_1^2}\widetilde{B_2}(t)\right\}. \tag{7.17}$$

Under the same sample paths of Wiener processes as in the computation of $r(t)$ above, the underlying assets are

$$S_1(t_i) = S_1(0)\exp\left\{\int_0^{t_i} r(s)ds - \frac{1}{2}\sigma_1^2 t_i + \sigma_1 \sum_{j=0}^{i-1} \Delta\widetilde{B_1}(t_i)\right\} \tag{7.18}$$

$$S_2(t_i) = S_2(0)\exp\left\{\int_0^{t_i} r(s)ds - \frac{1}{2}\sigma_2^2 t_i + \sigma_2\rho_1 \sum_{j=0}^{i-1} \Delta\widetilde{B_1}(t_i) + \sigma_2\sqrt{1-\rho_1^2}\sum_{j=0}^{i-1} \Delta\widetilde{B_2}(t_i)\right\}, \tag{7.19}$$

or recursively,

$$S_1(t_{i+1}) = S_1(t_i)\exp\left\{\int_{t_i}^{t_{i+1}} r(s)ds - \frac{1}{2}\sigma_1^2 \Delta t_i + \sigma_1 \Delta\widetilde{B_1}(t_i)\right\} \tag{7.20}$$

$$S_2(t_{i+1}) = S_2(t_i)\exp\left\{\int_{t_i}^{t_{i+1}} r(s)ds - \frac{1}{2}\sigma_2^2 \Delta t_i + \sigma_2\rho_1 \Delta\widetilde{B_1}(t_i) + \sigma_2\sqrt{1-\rho_1^2}\Delta\widetilde{B_2}(t_i)\right\}. \tag{7.21}$$

Use the Trapezoidal rule to approximate the integral

$$\int_0^{t_i} r(s)ds \approx \frac{\Delta t}{2}\left(r_1 + 2r_2 + 2r_3 + \cdots + 2r_{i-2} + 2r_{i-1} + r_i\right). \tag{7.22}$$

Alternatively, we can apply the Euler scheme to the underlying assets:

$$S_1(t_{i+1}) = S_1(t_i) + r(t_i)S_1(t_i)\Delta t + \sigma_1 S_1(t_i)\Delta\widetilde{B_1}(t_i) \tag{7.23}$$

$$S_2(t_{i+1}) = S_2(t_i) + r(t_i)S_2(t_i)\Delta t + \sigma_2\rho_1 S_2(t_i)\Delta\widetilde{B_1}(t_i) + \sigma_2\sqrt{1-\rho_1^2}S_2(t_i)\Delta\widetilde{B_2}(t_i). \tag{7.24}$$

And the Milstein scheme:

For the integral form of the underlying assets

$$S_1(t_{i+1}) = S_1(t_i) + \int_{t_i}^{t_{i+1}} r(t)S_1(t)dt + \sigma_1\int_{t_i}^{t_{i+1}} S_1(t)d\widetilde{B_1}(t)$$

$$S_2(t_{i+1}) = S_2(t_i) + \int_{t_i}^{t_{i+1}} r(t)S_2(t)dt + \sigma_2\rho_1\int_{t_i}^{t_{i+1}} S_2(t)d\widetilde{B_1}(t) + \sigma_2\sqrt{1-\rho_1^2}\int_{t_i}^{t_{i+1}} S_2(t)d\widetilde{B_2}(t)$$

For integral $\int_{t_i}^{t_{i+1}} S_1(t)d\widetilde{B_1}(t)$, we apply the Euler scheme and get rid of the $(t-t_i)$ term and obtain

$$\begin{aligned}\int_{t_i}^{t_{i+1}} S_1(t)d\widetilde{B_1}(t) &= \int_{t_i}^{t_{i+1}} \left\{S_1(t_i) + \sigma_1 S_1(t_i)\left(\widetilde{B_1}(t) - \widetilde{B_1}(t_i)\right)\right\} d\widetilde{B_1}(t) \\ &= S_1(t_i)\Delta\widetilde{B_1} + \sigma_1 S_1(t_i)\left(\frac{1}{2}\Delta\widetilde{B_1}^2 - \frac{1}{2}\Delta t\right) \\ &= -\frac{1}{2}\sigma_1 S_1(t_i)\Delta t + S_1(t_i)\Delta\widetilde{B_1}^2 + \frac{1}{2}\sigma_1 S_1(t_i)\Delta\widetilde{B_1}^2.\end{aligned}$$

Similarly,

$$\int_{t_i}^{t_{i+1}} S_2(t)d\widetilde{B_1}(t) = -\frac{1}{2}\sigma_2\rho_1 S_2(t_i)\Delta t + S_2(t_i)\Delta\widetilde{B_1} + \frac{1}{2}\sigma_2\rho_1 S_2(t_i)\Delta\widetilde{B_1}^2 \tag{7.25}$$

and

$$\int_{t_i}^{t_{i+1}} S_2(t)d\widetilde{B_2}(t) = -\frac{1}{2}\sigma_2\sqrt{1-\rho_1^2}S_2(t_i)\Delta t + S_2(t_i)\Delta\widetilde{B_2} + \frac{1}{2}\sigma_2\sqrt{1-\rho_1^2}S_2(t_i)\Delta\widetilde{B_2}^2. \tag{7.26}$$

Plugging back to the original expressions and simplifying, we have

$$S_1(t_{i+1}) = S_1(t_i) + (r(t_i) - \frac{1}{2}\sigma_1^2)S_1(t_i)\Delta t + \sigma_1 S_1(t_i)\Delta\widetilde{B_1} + \frac{1}{2}\sigma_1^2 S_1(t_i)\Delta\widetilde{B_1}^2 \tag{7.27}$$

$$\begin{aligned}S_2(t_{i+1}) = S_2(t_i) &+ (r(t_i) - \frac{1}{2}\sigma_2^2)S_2(t_i)\Delta t + \sigma_2\rho_1 S_2(t_i)\Delta\widetilde{B_1} \\ &+ \sigma_2\sqrt{1-\rho_1^2}S_2(t_i)\Delta\widetilde{B_2} + \frac{1}{2}\sigma_2^2\rho_1^2 S_2(t_i)\Delta\widetilde{B_1}^2 + \frac{1}{2}\sigma_2^2(1-\rho_1^2)S_2(t_i)\Delta\widetilde{B_2}^2\end{aligned} \tag{7.28}$$

7.2 Regression-Based Method

7.2.1 Introduction. Due to the early exercise characteristic of the American spread options, the owner needs to determine whether to exercise or hold the options at the current time. This suggests that we can use a dynamic programming formulation as follows:

$$\widetilde{V}_m(x) = \widetilde{h}_m(x)$$
$$\widetilde{V}_{i-1}(x) = \max\{\widetilde{h}_{i-1}(x), E[D_{i-1,i}(X_i)\widetilde{V}_i(X_i)|X_{i-1} = x]\}, i = 1, ..., m.$$

where $\widetilde{V}_i$ is the option price at time t_i, $\widetilde{h}_i$ is the payment function which indicates the intrinsic value of the option at time t_i and $D_{i-1,i} = e^{-\int_{t_{i-1}}^{t_i} r(s)ds}$ is the discount factor from t_{i-1} to t_i. X_i is the underlying assets' price vector at time t_i, in the spread option case, $X_i = (S_1(t_i), S_2(t_i))$. That is to say, to obtain the American option price, one needs to compare the intrinsic value if the option is exercised immediately with the corresponding European option price if the option will be held. The whole formulation is based on a backward induction mechanism.

Although the formulation above is used in practical implementation, it is not as convenient as the following formulation for introducing the algorithm:

$$V_m(x) = h_m(x)$$
$$V_{i-1}(x) = \max\{h_{i-1}(x), E[V_i(X_i)|X_{i-1} = x]\}, i = 1, ..., m.$$

where we applied the replacement

$$h_i(x) = D_{0,i}(x)\widetilde{h}_i(x), \quad i = 1, ..., m$$

$$V_i(x) = D_{0,i}(x)\widetilde{V}_i(x), \quad i = 0, 1, ..., m.$$

For the simplicity of the expression, we denote the conditional expectation in the formulation above as

$$C_i(x) = E[V_{i+1}(X_{i+1})|X_i = x], \quad i = 0, ..., m-1.$$

This is the continuation value if we want to hold the option. Note that at the terminal time

m, the continuation value $C_m \equiv 0$ if the option is not exercised. Thus,

$$V_i(x) = \max\{h_i(x), C_i(x)\}, \quad i = 1, ..., m.$$

The key to approximate the option price V_i is to approximate the continuation value C_i. Similar to the idea of Taylor expansion and Fourier series expansion, we assume there are some basis functions $\phi_j : \mathbb{R}^2 \to \mathbb{R}$ and constants $\theta_{ij},\ j = 1, ..., L$, such that

$$C_i(x, \theta_i) = E[V_{i+1}(X_{i+1})|X_i = x] = \sum_{j=1}^{L} \theta_{ij}\phi_j(x). \tag{7.29}$$

By denoting

$$\theta_i^\top = (\theta_{i1}, ..., \theta_{iL}), \quad \phi(x) = (\phi_1(x), ..., \phi_L(x))^\top,$$

we can simplify the continuation value as

$$C_i(x, \theta_i) = \theta_i^\top \phi(x).$$

Under this hypothesis, the coefficient vector can be solved by

$$\theta_i = (E[\phi(X_i)\phi(X_i)^\top])^{-1} E[\phi(X_i)V_{i+1}(X_{i+1})].$$

In the practical implementation, however, the previous equation cannot be guaranteed. The accuracy of the approximation depends on the diversity of the functions ϕ. For the constants θ_i, we use the arithmetic average to estimate the expectations, therefore

$$\widehat{\theta}_i = (\frac{1}{M}\sum_{j=1}^{M} \phi(X_i^{(j)})\phi(X_i^{(j)})^\top)^{-1}(\frac{1}{M}\sum_{j=1}^{M} \phi(X_i^{(j)})V_{i+1}(X_i^{(j)})^\top), \tag{7.30}$$

where $\widehat{\theta}_i$ is the estimator for θ. We will use this convention later on. So

$$\widehat{C}_i(x) = \widehat{\theta}_i^\top \phi(x). \tag{7.31}$$

We can summarize it as the *algorithm for the regression-based method*:

- Simulate M independent paths

$$\{X_1^{(j)}, ..., X_N^{(j)}\}, \quad j = 1, ..., M,$$

  where

$$X_i^{(j)} = (S_1^{(j)}(i), S_2^{(j)}(i)).$$

- At the expiry ($i = N$), set

$$\widehat{V}_N^{(j)} = h_N(X_N^{(j)}), \quad j = 1, ..., M.$$

- For $i = N - 1, ..., 1$,
  - Calculate $\widehat{\theta}_i$ as in (7.30).
  - Calculate $\widehat{C}_i$ as in (7.31).
  - Set

$$\widehat{V}_i^{(j)} = \max\{h_i(X_i^{(j)}), \widehat{C}_i(X_i^{(j)})\}, \quad j = 1, ..., M. \tag{7.32}$$

- Set $\widehat{V}_0 = \dfrac{1}{M} \sum_{j=1}^M \widehat{V}_1^{(j)}$. This algorithm gives a high-biased estimator in the sense that

$$E[\widehat{V}_0] \geq V_0.$$

This approach is introduced in [37] and is showed in [38] that $\widehat{V}_0 \to V_0(X_0)$ as $M \to \infty$ if the assumption of the basis function representation holds for all $i = 1, ..., N-1$. In [31], Longstaff and Schwartz use their interleaving estimator

$$\widehat{V}_i^{(j)} = \begin{cases} h_i(X_i^{(j)}), & h_i(X_i^{(j)}) \geq \widehat{C}_i(X_i^{(j)}) \\ \widehat{V}_{i+1}^{(j)}, & h_i(X_i^{(j)}) < \widehat{C}_i(X_i^{(j)}) \end{cases} \tag{7.33}$$

in place of Equation (7.32). This algorithm produces a low-biased estimator in the sense that

$$E[\widehat{V}_0] \leq V_0,$$

and it has been proved to have almost sure convergence as $M \to \infty$.

7.2.2 Numerical Implementations. Now we can implement the two algorithms in the last section to estimate American spread option prices under two stochastic interest models, we can also compare these results with the constant interest rate case. Throughout this section, we set $S_1(0) = 50$, $S_2(0) = 40$, $r(0) = 0.05$, $\alpha = 1$, $\beta = 2$, $\sigma_1 = 0.2$, $\sigma_2 = 0.2$, $\sigma_3 = 0.3$, $\rho_1 = 0.5$, $\rho_2 = 0.4$, $\rho_3 = 0.3$, $T = 0.5$ and $K = 10$. We use M to denote the number of independent paths for the underlying assets (S_1, S_2), N the number of chances to exercise the option before the maturity date(exercise times for short). In other words, we discretize the time interval $[0, T]$ into N points and simulate on these points. We do this for two reasons:(1)In the real world, because of the existence of the transaction time and the rest period of Exchanges, the chance for exercising the option or any other financial derivatives is not continuous. (2)When $N \to \infty$, the simulation on the discretized points will approach the simulation on the time interval.

We will compare the price value varying $M = \{1000, 5000, 10000, 20000, 50000, 100000\}$ and $N = \{5, 10, 100\}$.

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	3.0927	0.1255	3.8143	0.2223	5.2369	2.4099
5000	3.1508	0.5289	3.7764	1.1492	5.3238	12.1812
10000	3.1977	0.9898	3.7557	2.1463	5.2875	23.4154
20000	3.1814	1.8829	3.7504	4.6652	5.3314	48.1620
50000	3.1456	4.6373	3.7328	10.8806	5.3628	118.7494
100000	3.1471	9.3768	3.7379	22.9579	5.3338	234.8219

Table 7.1: Monte Carlo with high-biased estimator under constant interest rate model ((S_1, S_2)–explicit)

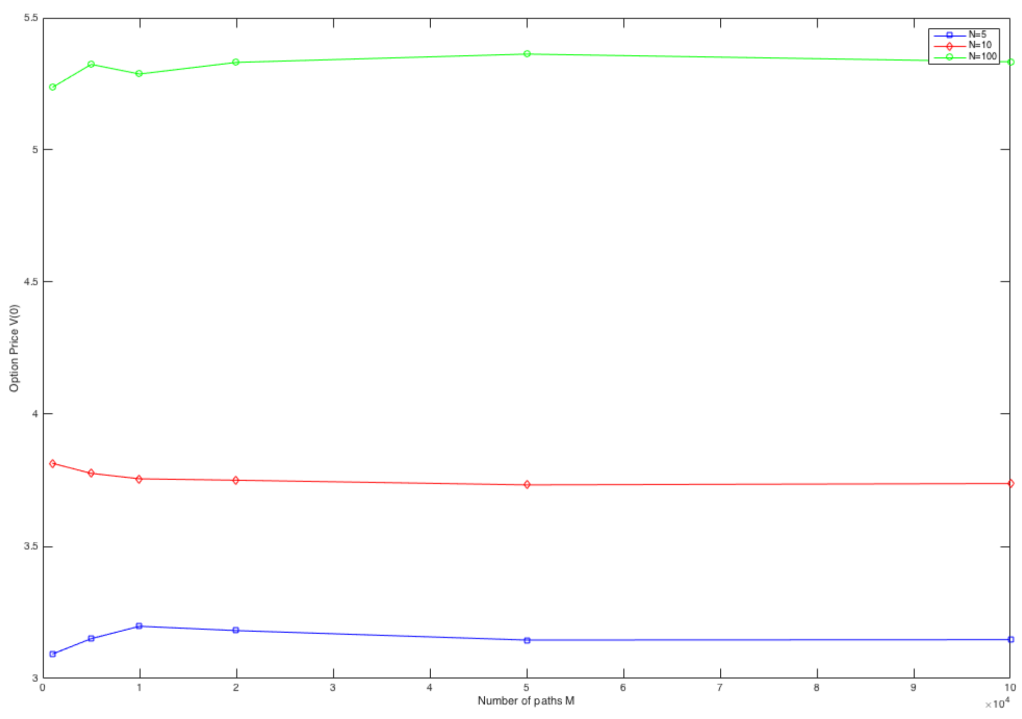

Figure 7.1: 2D-Monte Carlo with high-biased estimator under constant interest rate model ((S_1, S_2)–explicit)

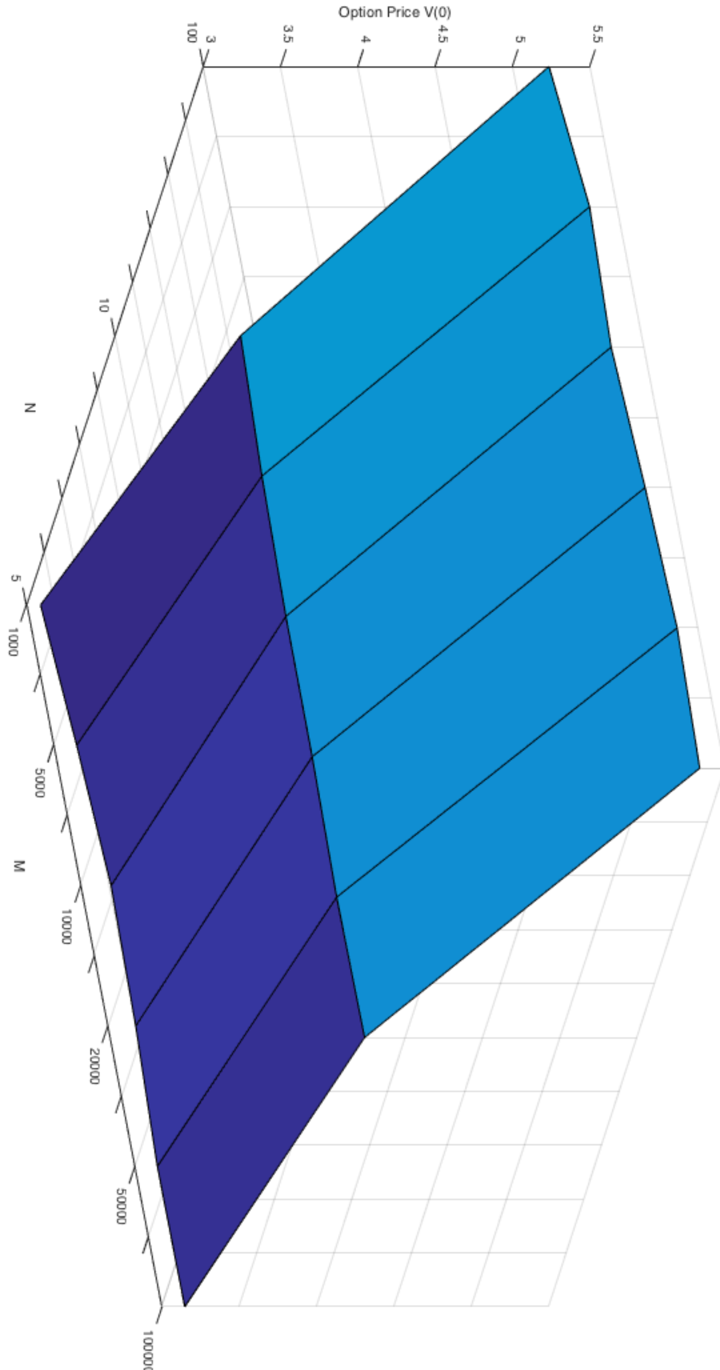

Figure 7.2: 3D-Monte Carlo with high-biased estimator under constant interest rate model $((S_1, S_2)$–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.3853	0.1472	2.5129	0.2036	2.3352	2.2232
5000	2.4594	0.5414	2.4338	0.9867	2.2111	10.7613
10000	2.4446	0.8963	2.3859	1.9456	2.2365	21.5603
20000	2.4605	1.7567	2.3774	3.9288	2.1778	43.0085
50000	2.4650	4.3240	2.3934	9.7510	2.2489	110.1174
100000	2.4483	8.6268	2.4138	19.3397	2.2382	219.3565

Table 7.2: Monte Carlo with Low-biased estimator under constant interest model ((S_1, S_2)–explicit)

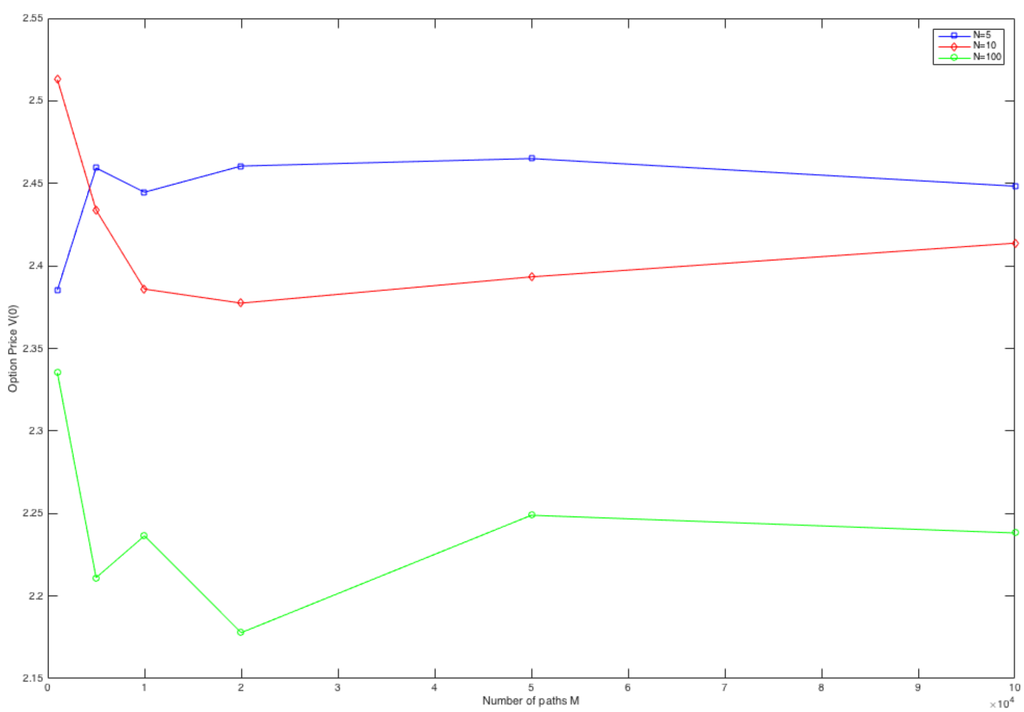

Figure 7.3: 2D-Monte Carlo with low-biased estimator under constant interest model ((S_1, S_2)–explicit)

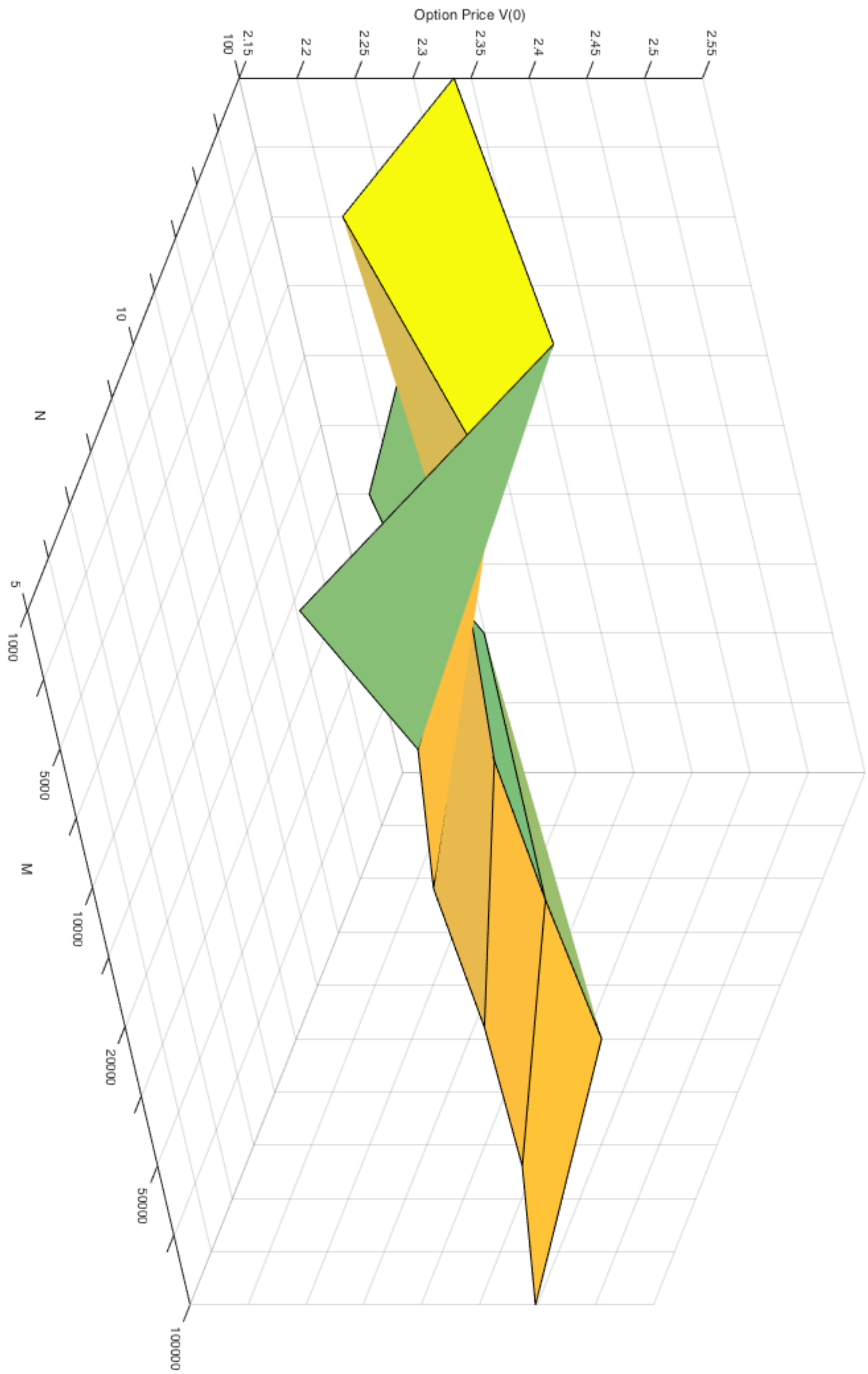

Figure 7.4: 3D-Monte Carlo with low-biased estimator under constant model $((S_1, S_2)$–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.7402	0.0946	3.5949	0.2012	5.5629	2.2166
5000	2.9749	0.4608	3.2586	1.0335	5.1939	10.8081
10000	2.9205	0.8923	3.4921	1.9456	5.6241	21.6271
20000	2.9183	1.7449	3.4418	3.8575	5.3490	43.1369
50000	2.8483	4.4333	3.5885	9.6949	5.3909	111.7667
100000	2.8641	9.1737	3.6523	19.4059	5.5515	223.2383

Table 7.3: Monte Carlo with high-biased estimator under Vasicek model (r–Euler scheme, (S_1, S_2)–explicit)

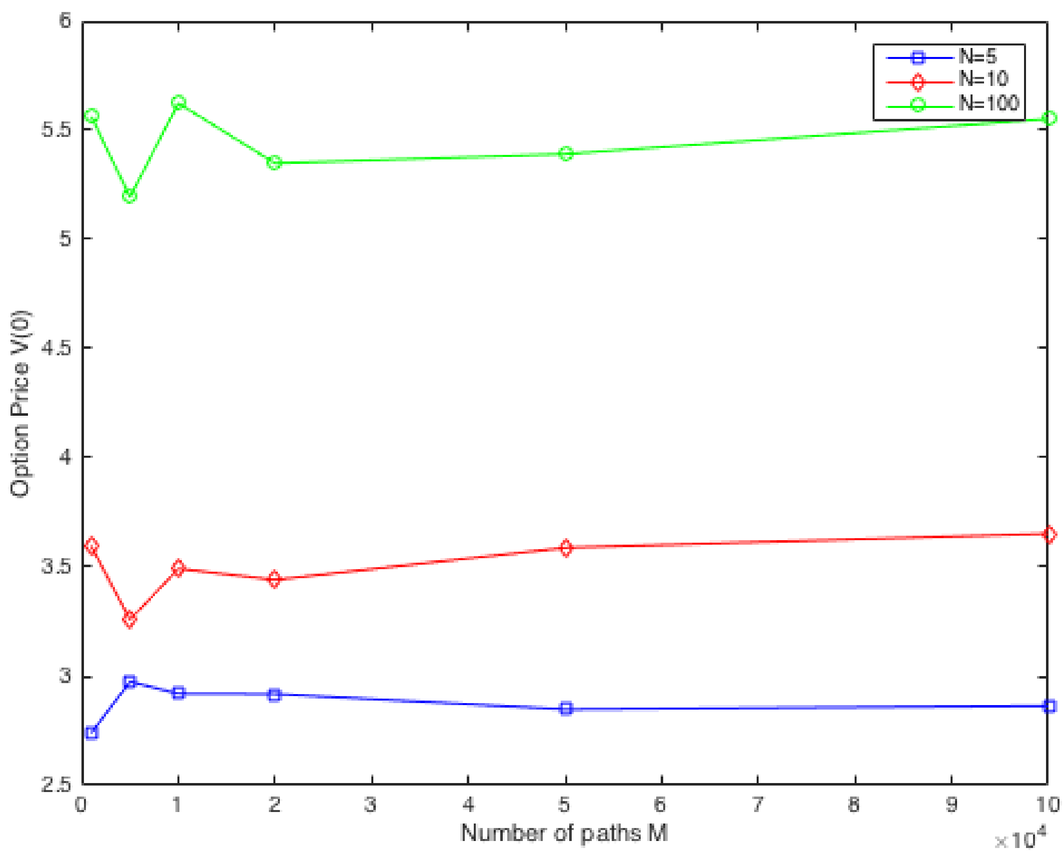

Figure 7.5: 2D-Monte Carlo with high-biased estimator under Vasicek model (r–Euler scheme, (S_1, S_2)–explicit)

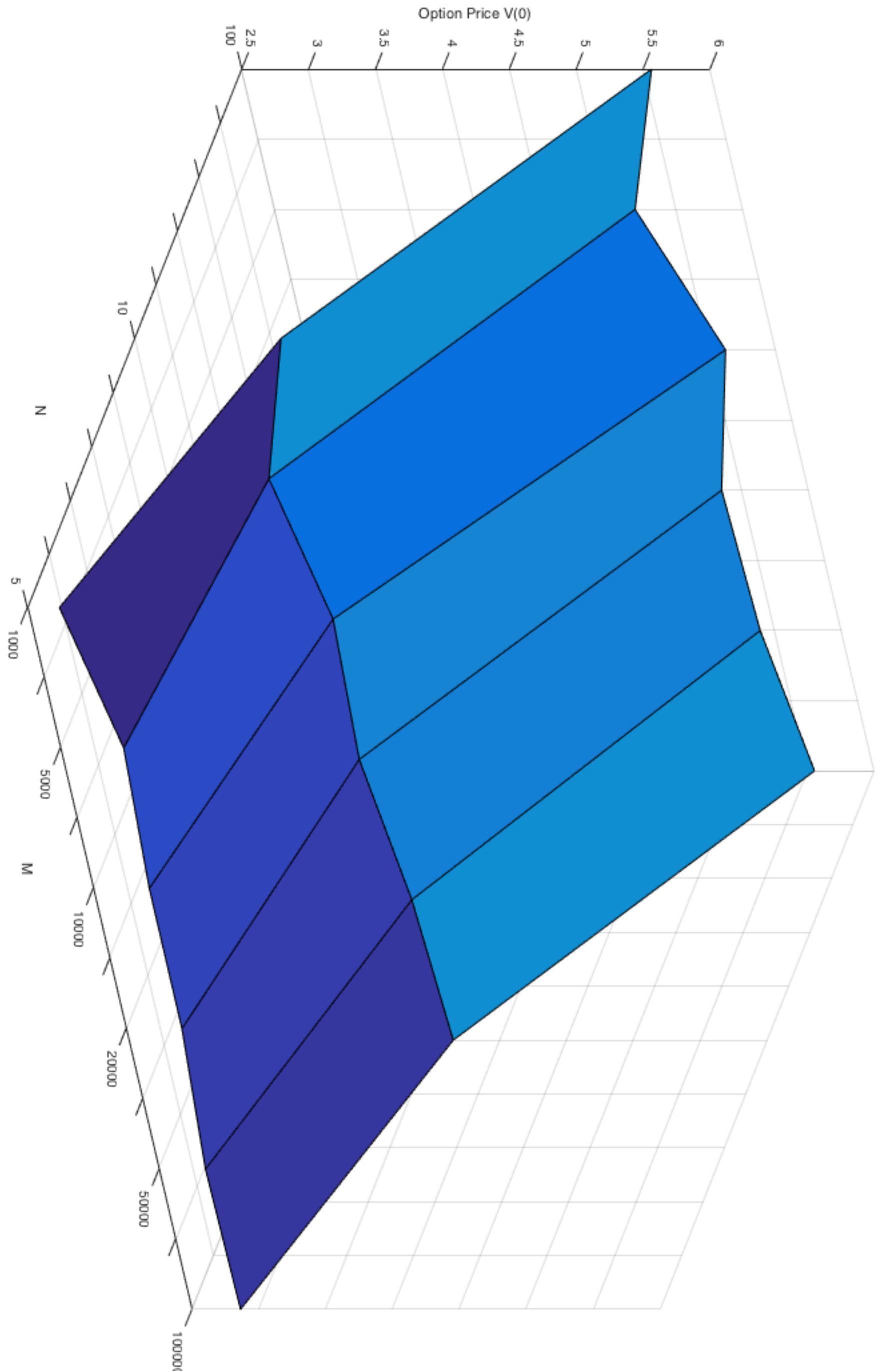

Figure 7.6: 3D-Monte Carlo with high-biased estimator under Vasicek model (r–Euler scheme, (S_1, S_2)–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.1169	0.1223	2.3142	0.2073	2.1809	2.1876
5000	2.2230	0.4782	2.3401	1.0415	2.0573	11.3607
10000	2.3006	0.9250	2.2682	2.1062	2.0499	22.2239
20000	2.2894	1.7803	2.2292	3.9393	2.0860	46.9122
50000	2.2626	4.3754	2.2102	9.7568	2.0862	115.2364
100000	2.2735	9.0513	2.2368	19.7889	2.0800	227.3569

Table 7.4: Monte Carlo with low-biased estimator under Vasicek model (r–Euler scheme, (S_1, S_2)–explicit)

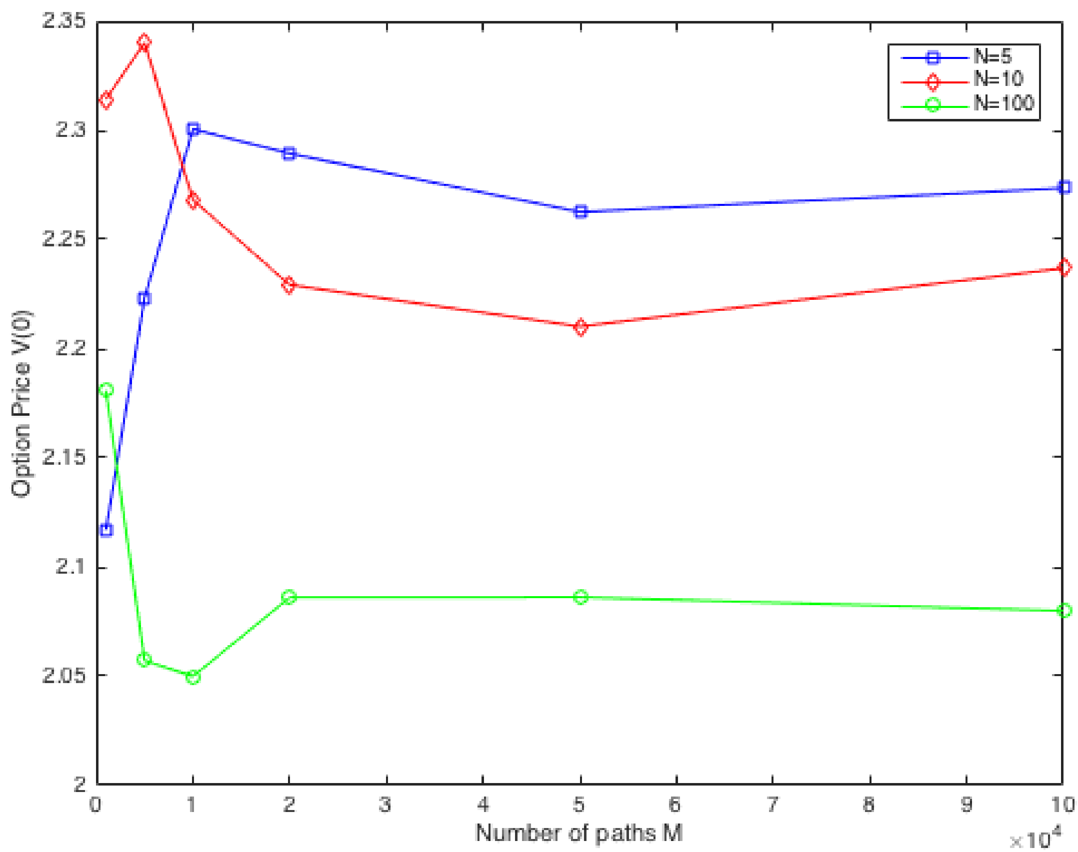

Figure 7.7: 2D-Monte Carlo with low-biased estimator under Vasicek model (r–Euler scheme, (S_1, S_2)–explicit)

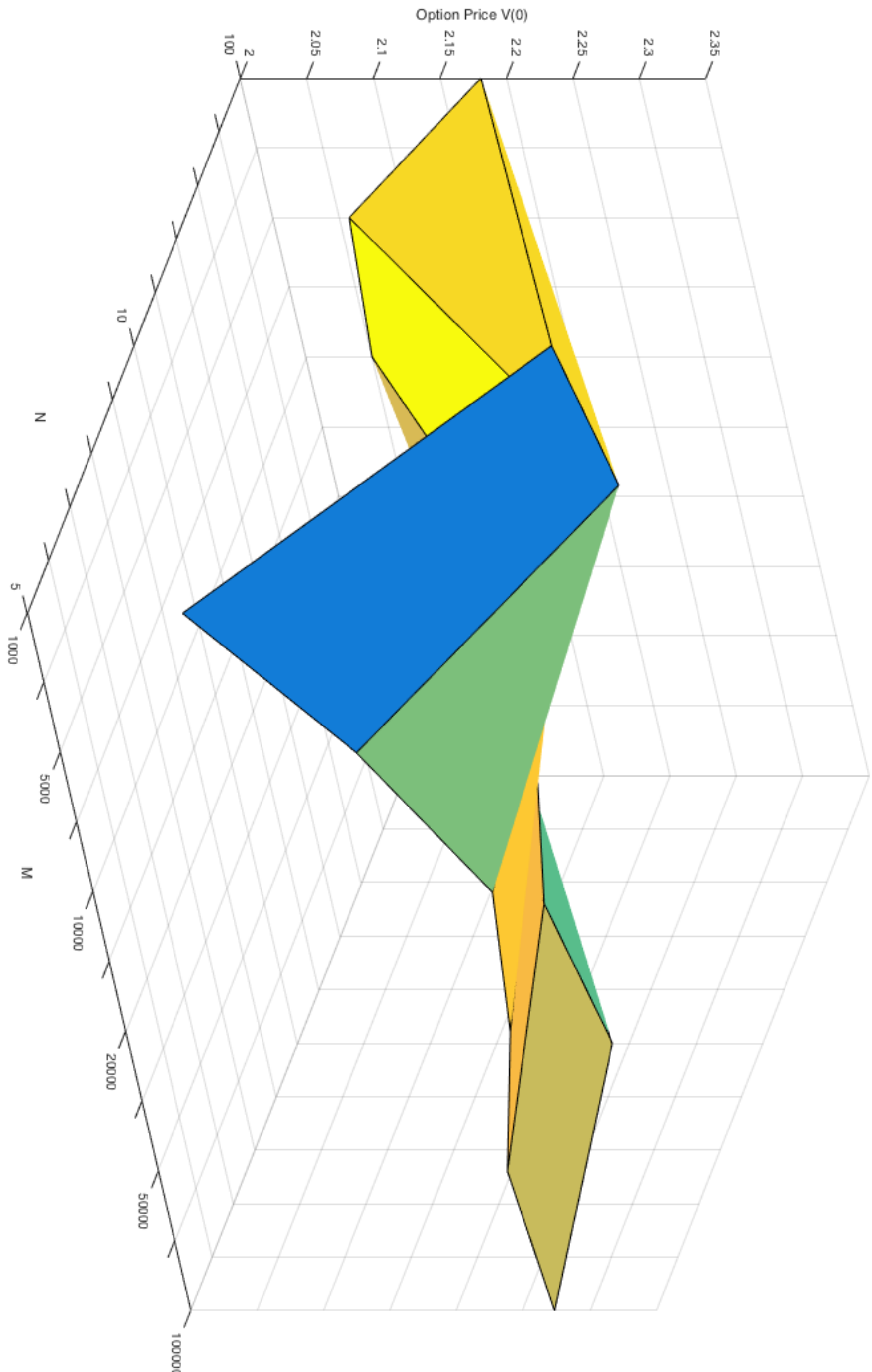

Figure 7.8: 3D-Monte Carlo with low-biased estimator under Vasicek model (r–Euler scheme, (S_1, S_2)–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.1169	0.1438	3.3812	0.2148	5.1030	2.1691
5000	2.9484	0.4689	3.4577	1.0042	5.1567	11.2310
10000	2.9431	1.0193	3.5206	2.1162	5.1639	21.9891
20000	2.9472	1.7882	3.5456	3.9196	5.1681	43.5762
50000	2.8988	4.5008	3.5056	9.8785	5.1168	106.7266112.8242
100000	2.9625	8.8873	3.5195	19.8227	5.1756	239.8470

Table 7.5: Monte Carlo with high-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–explicit)

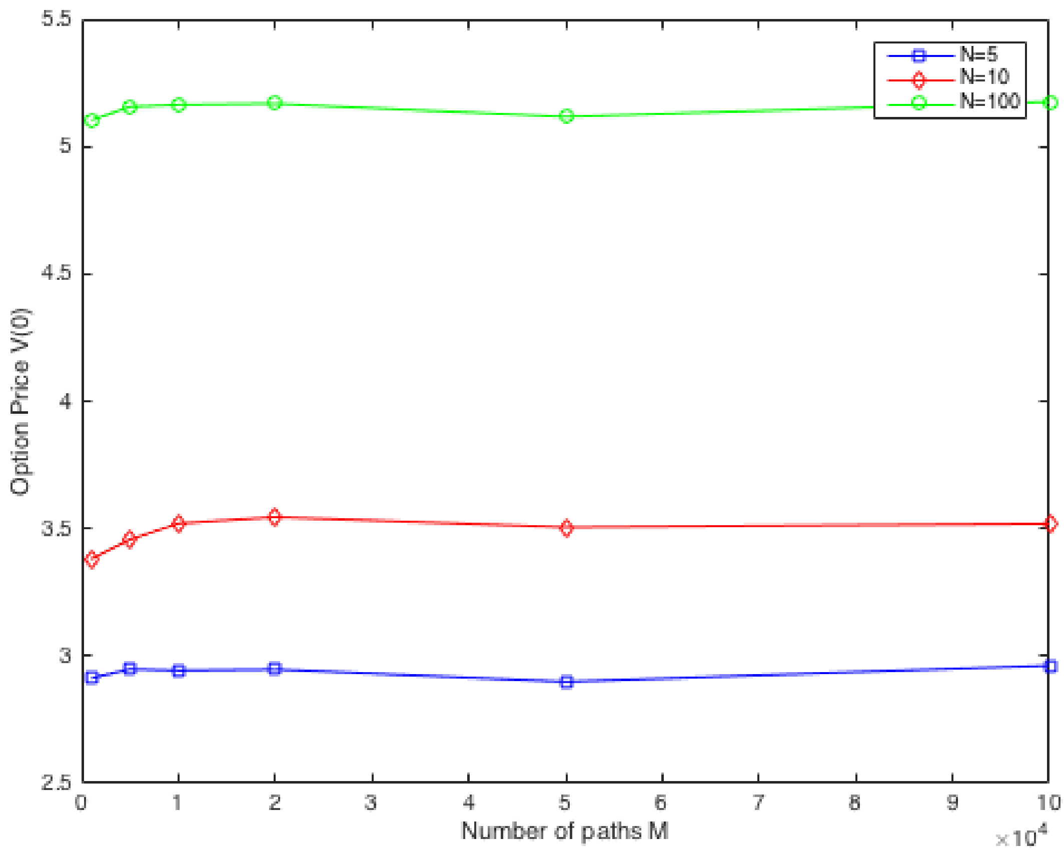

Figure 7.9: 2D-Monte Carlo with high-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–explicit)

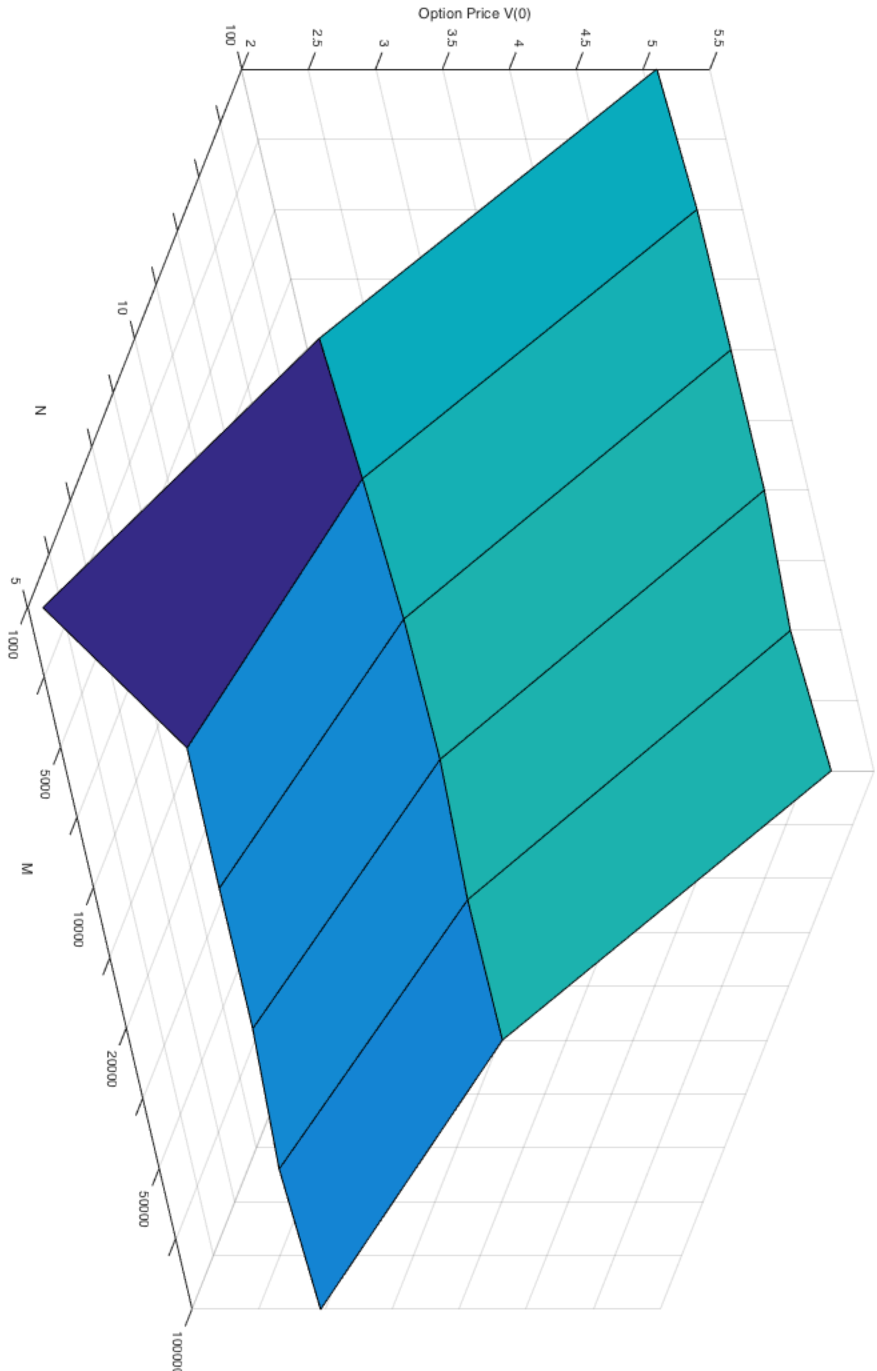

Figure 7.10: 3D-Monte Carlo with high-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.3970	0.1564	2.2730	0.2038	1.9817	2.1077
5000	2.2849	0.5157	2.1384	0.9662	2.1314	10.6454
10000	2.2933	0.9249	2.2523	1.9465	2.1091	21.4584
20000	2.2744	1.7118	2.2190	4.2281	2.0822	43.3503
50000	2.2893	4.1916	2.2306	10.4662	2.0784	105.1366
100000	2.2981	8.4950	2.2525	19.9697	2.0732	214.4767

Table 7.6: Monte Carlo with low-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–explicit)

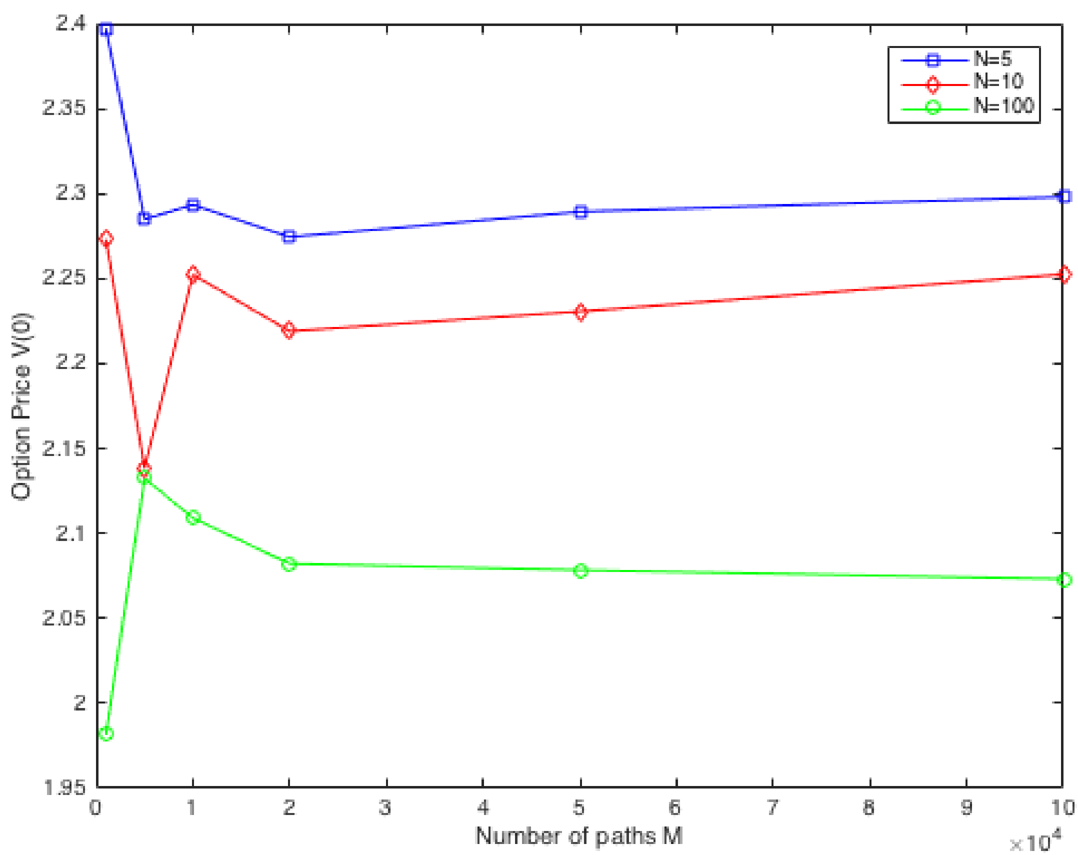

Figure 7.11: 2D-Monte Carlo with low-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–explicit)

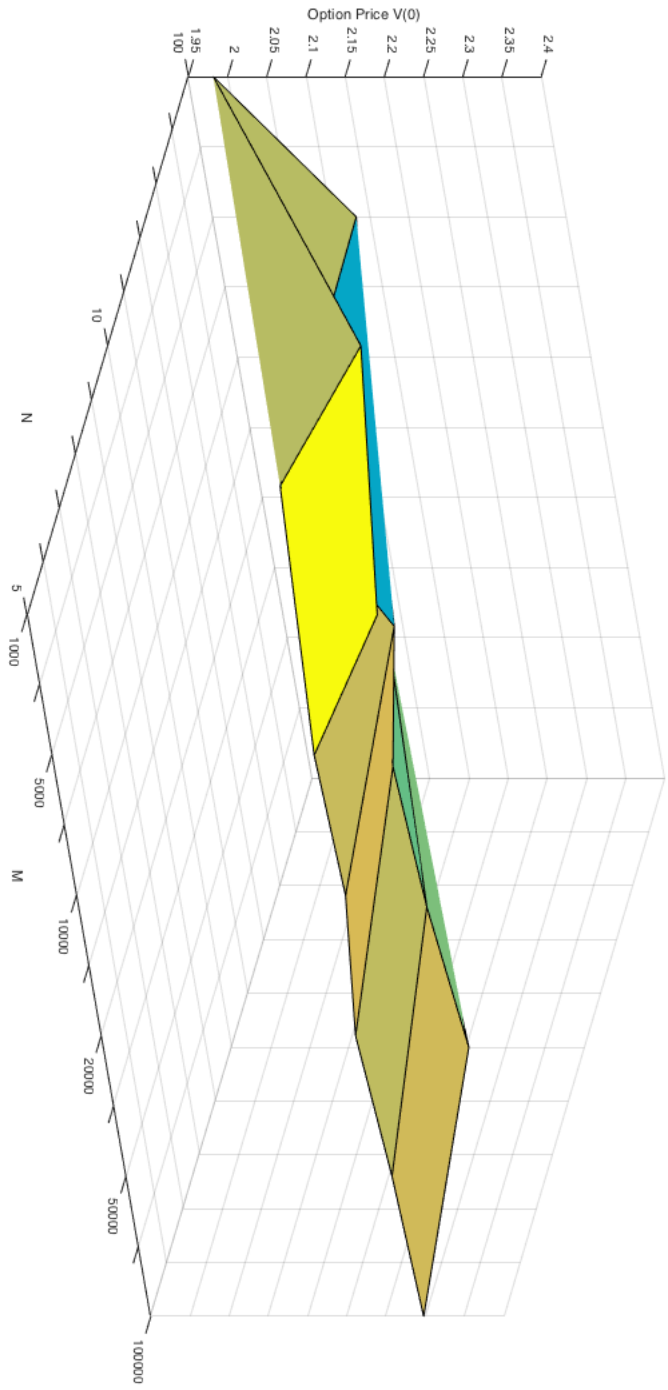

Figure 7.12: 3D-Monte Carlo with low-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.9322	0.1528	3.4227	0.2092	4.9908	2.1547
5000	2.8865	0.5148	3.4744	1.0097	5.1440	10.7550
10000	2.9054	0.9154	3.5727	1.9469	5.0964	21.6158
20000	2.8978	1.7245	3.4940	3.8606	5.2265	43.5733
50000	2.9780	4.2444	3.5054	9.7363	5.1239	106.3768
100000	2.9289	8.9010	3.5017	19.3355	5.0567	211.8942

Table 7.7: Monte Carlo with high-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–explicit)

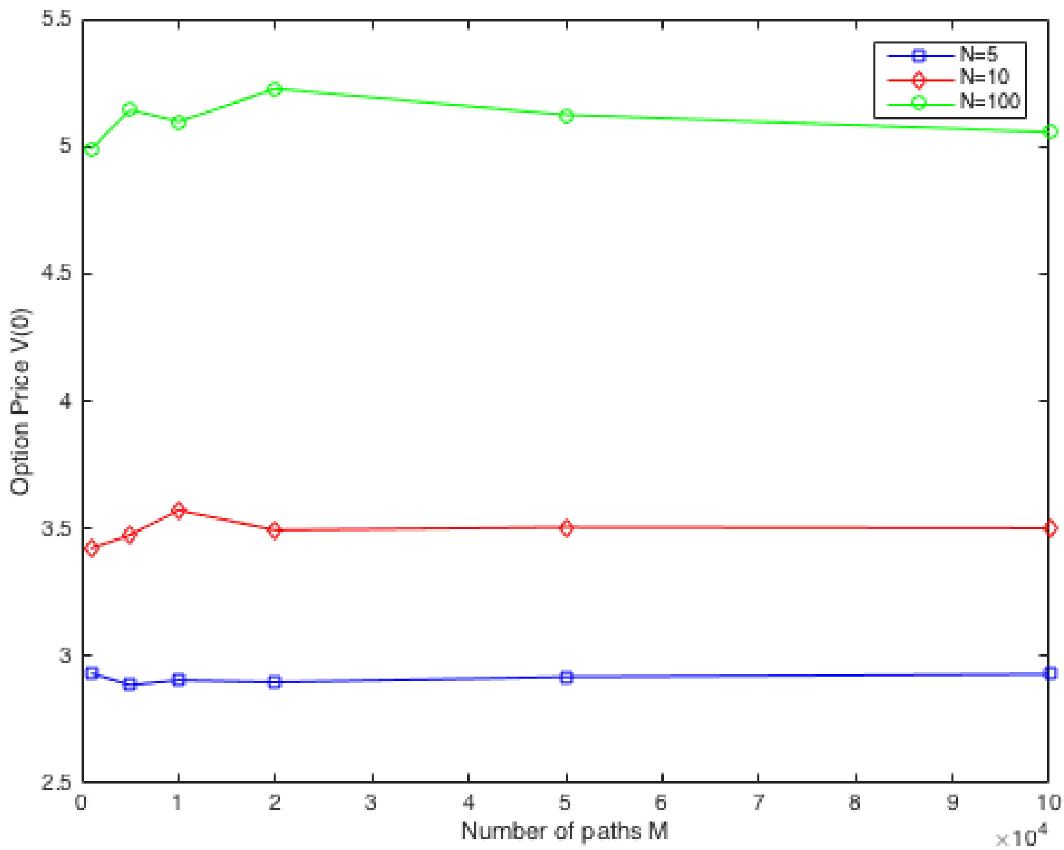

Figure 7.13: 2D-Monte Carlo with high-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–explicit)

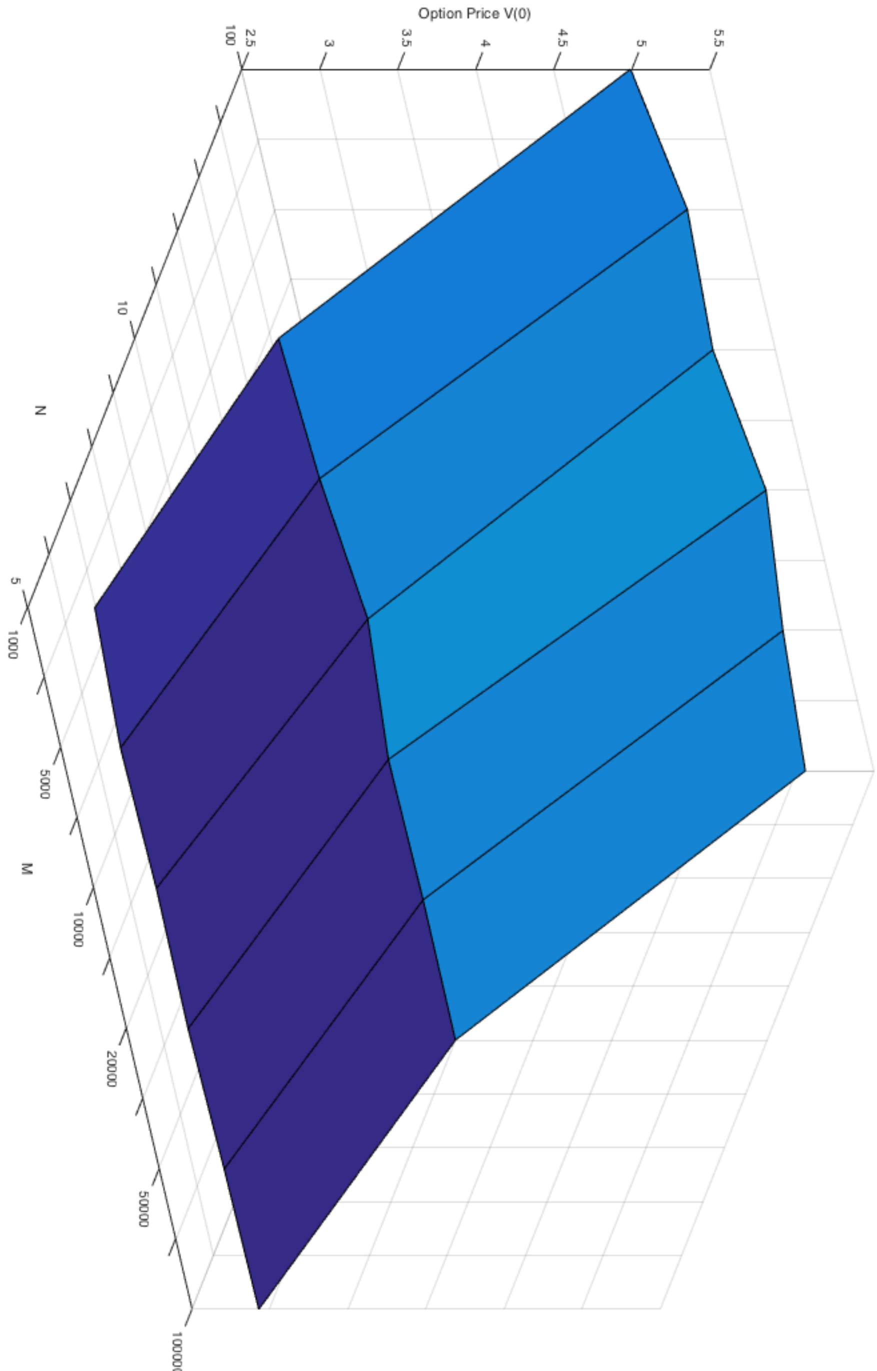

Figure 7.14: 3D-Monte Carlo with high-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.3149	0.1841	2.4115	0.2301	2.1297	2.3252
5000	2.2239	0.4954	2.2546	1.1066	2.0592	11.2984
10000	2.3214	0.9679	2.2402	2.3431	2.1294	23.4278
20000	2.2950	1.8823	2.2429	4.1519	2.0692	45.6496
50000	2.2880	4.7307	2.2429	10.4549	2.0837	114.6067
100000	2.2866	9.7160	2.2341	21.3634	2.0687	230.5335

Table 7.8: Monte Carlo with low-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–explicit)

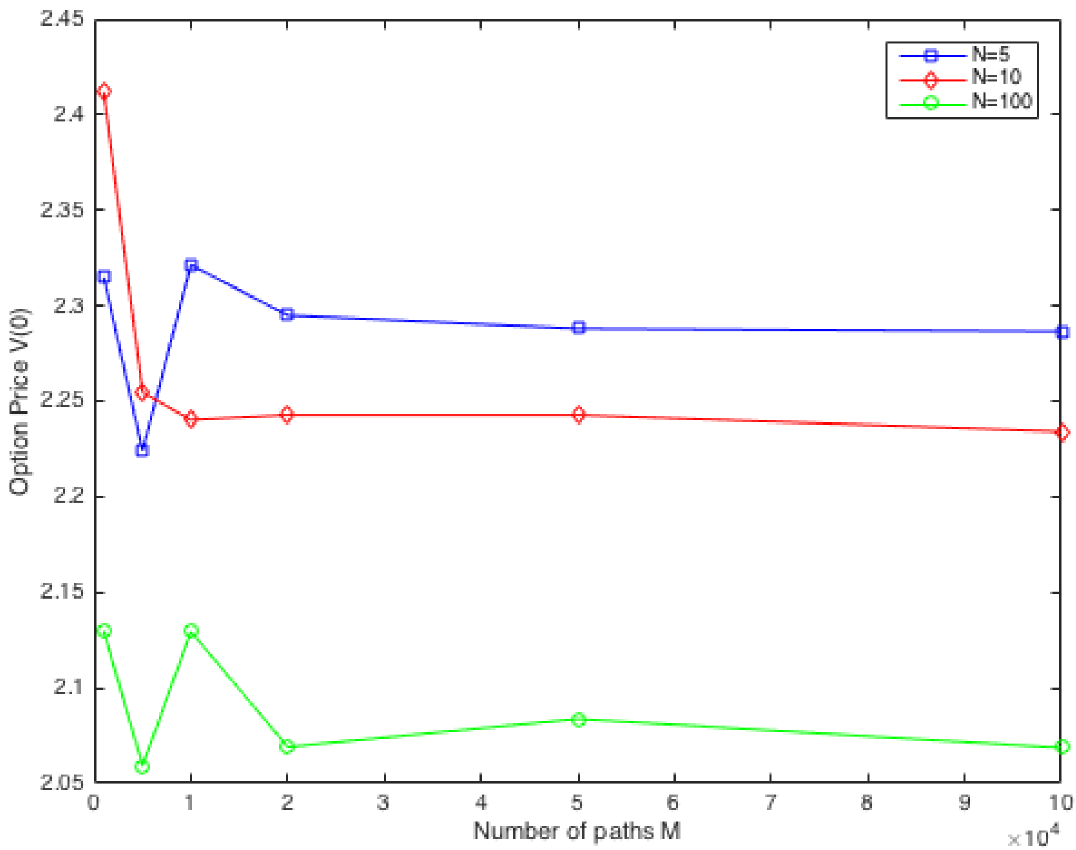

Figure 7.15: 2D-Monte Carlo with low-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–explicit)

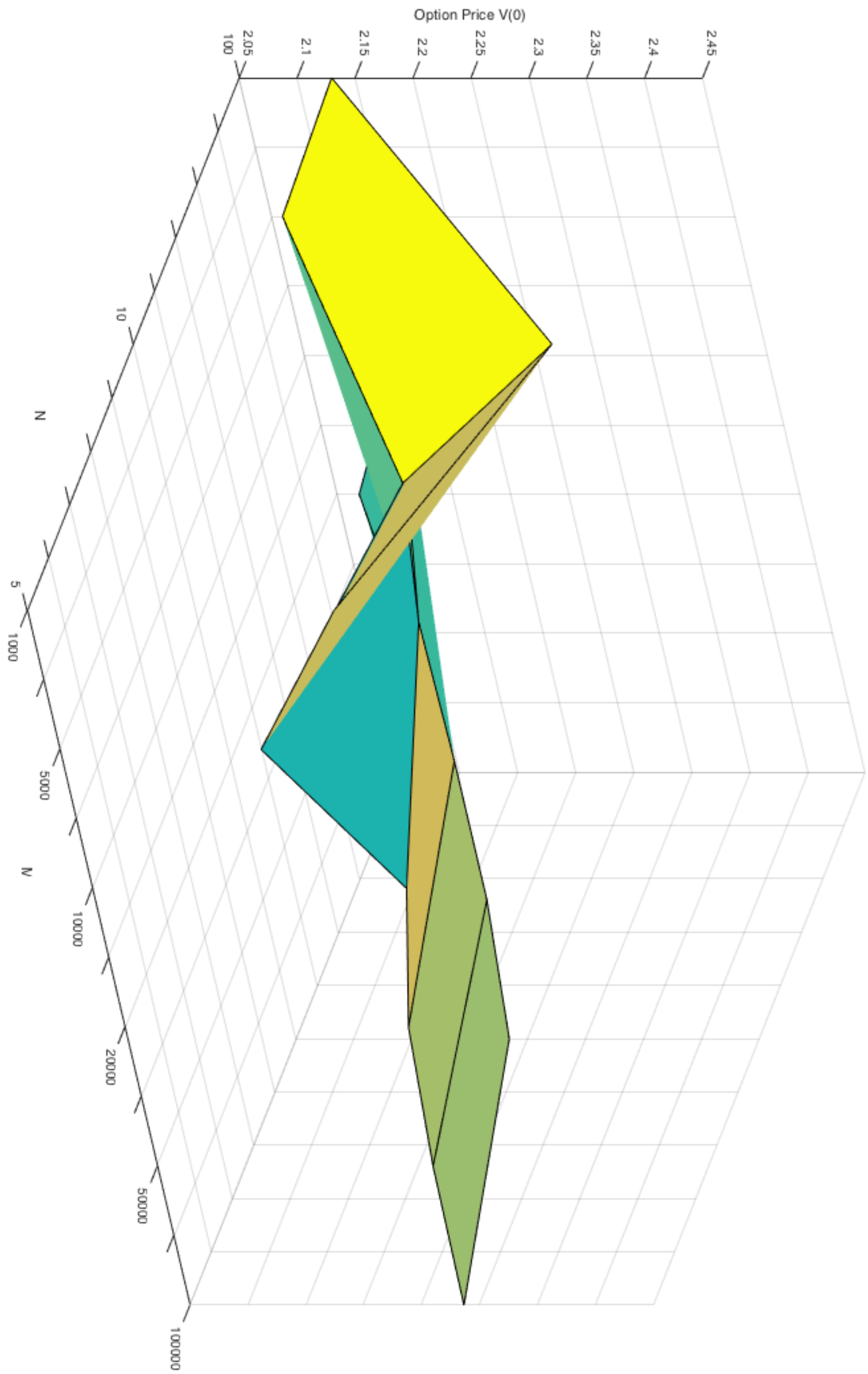

Figure 7.16: 3D-Monte Carlo with low-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–explicit)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.8826	0.1200	3.4103	0.2054	5.1653	2.1792
5000	2.9687	0.5360	3.4503	1.0177	5.1649	10.8793
10000	2.9554	0.9085	3.5253	2.0352	5.2292	21.9725
20000	2.8959	1.7278	3.4961	4.0847	5.1405	43.5237
50000	2.9262	4.5254	3.4500	9.7267	5.2651	109.6005
100000	2.8825	8.8693	3.4471	19.6880	5.1634	220.0782

Table 7.9: Monte Carlo with high-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–Euler scheme)

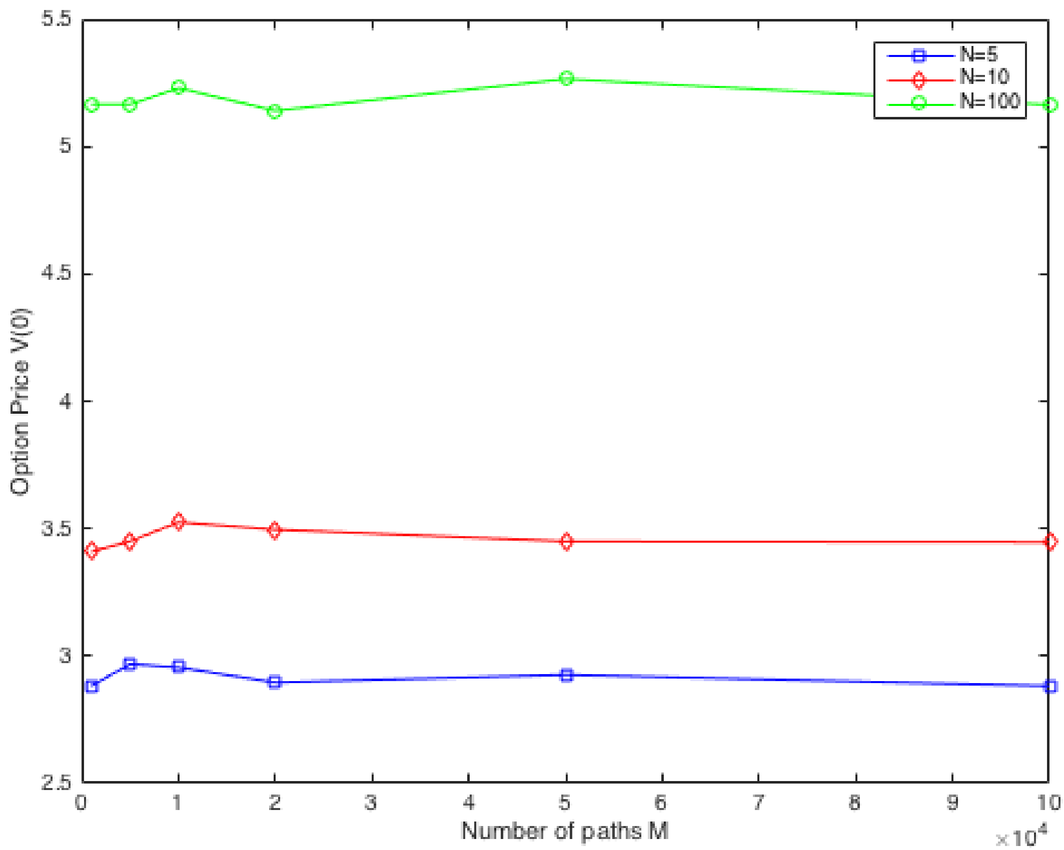

Figure 7.17: 2D-Monte Carlo with high-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–Euler scheme)

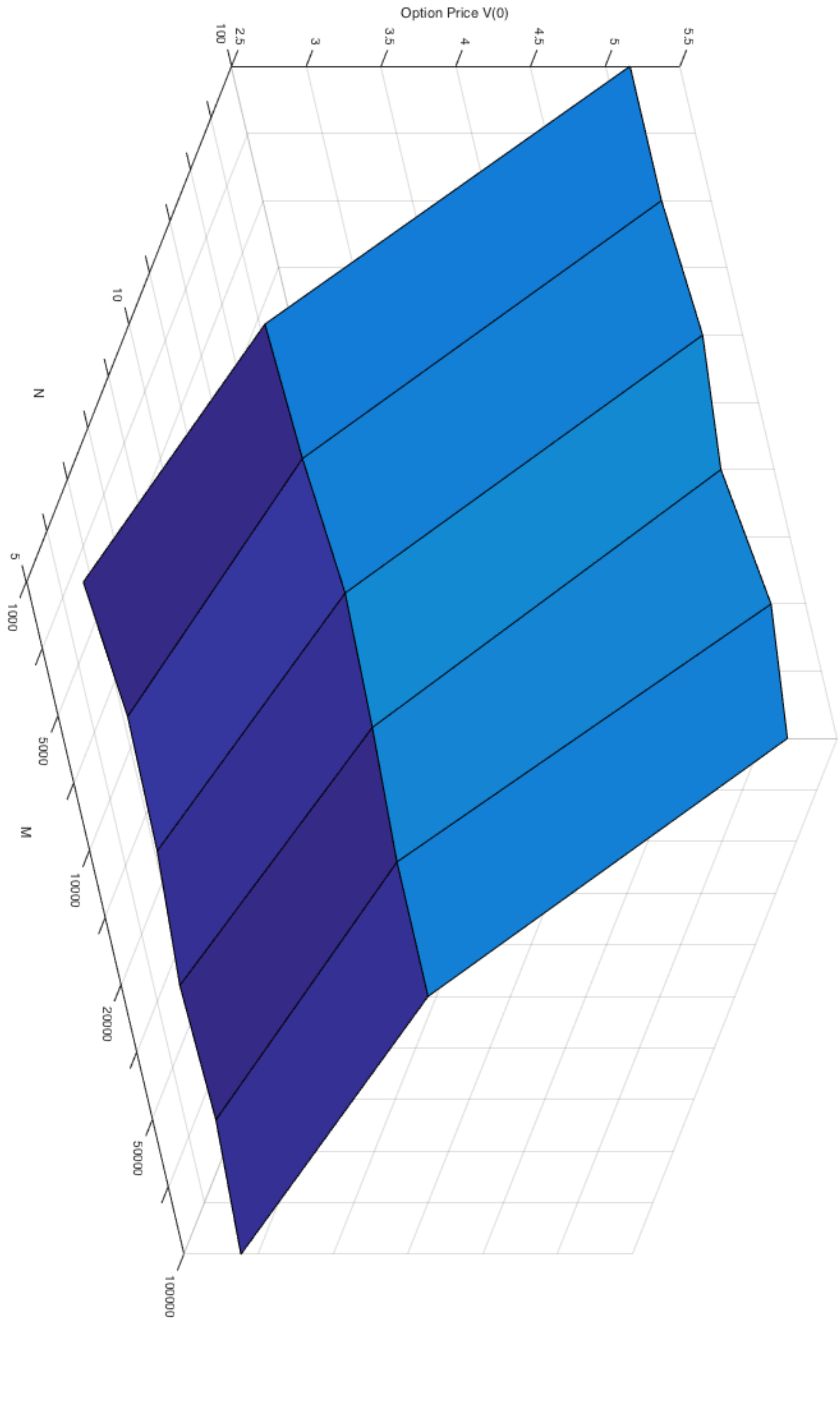

Figure 7.18: 3D-Monte Carlo with high-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–Euler scheme)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.2257	0.1317	2.4608	0.2056	2.1102	2.1735
5000	2.2581	0.4509	2.2306	0.9845	2.0429	10.9243
10000	2.2717	0.8994	2.2305	1.9841	2.1032	21.8706
20000	2.2612	1.8355	2.1962	3.9093	2.0875	44.6741
50000	2.2635	4.5275	2.1865	9.8917	2.0683	109.4263
100000	2.2480	8.9134	2.2321	19.7539	2.0916	224.3356

Table 7.10: Monte Carlo with low-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–Euler scheme)

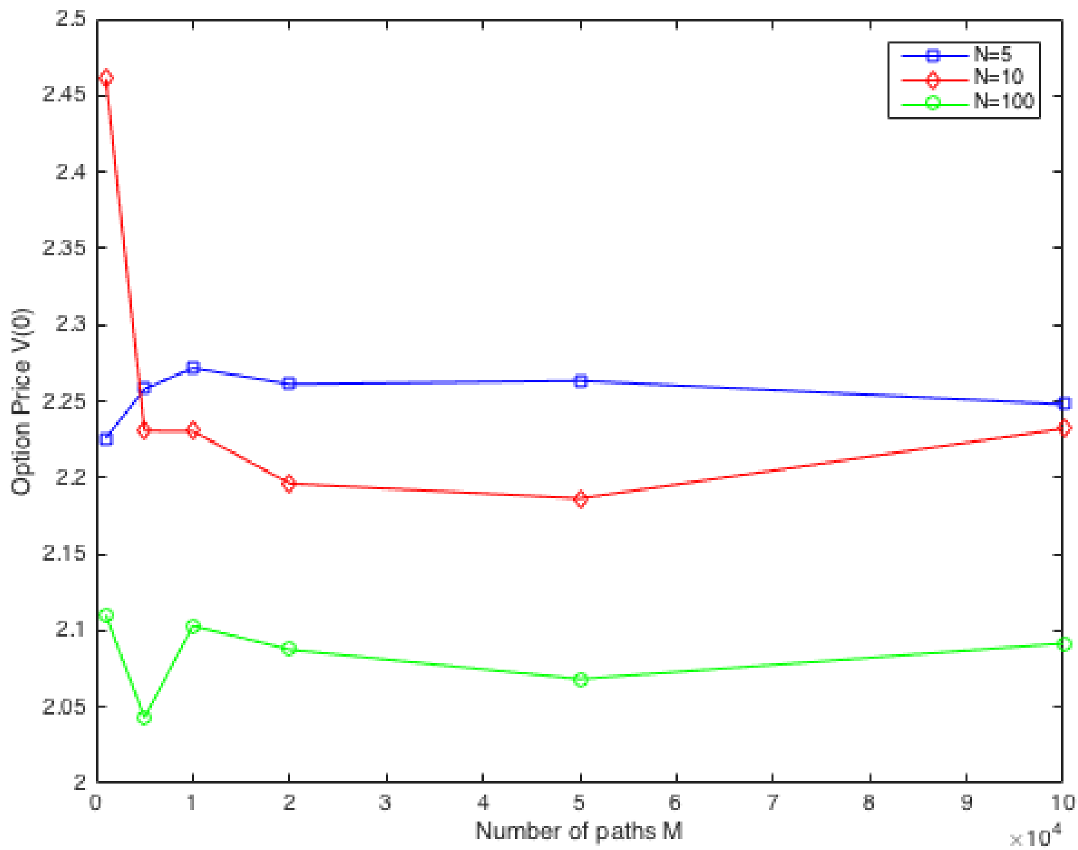

Figure 7.19: 2D-Monte Carlo with low-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–Euler scheme)

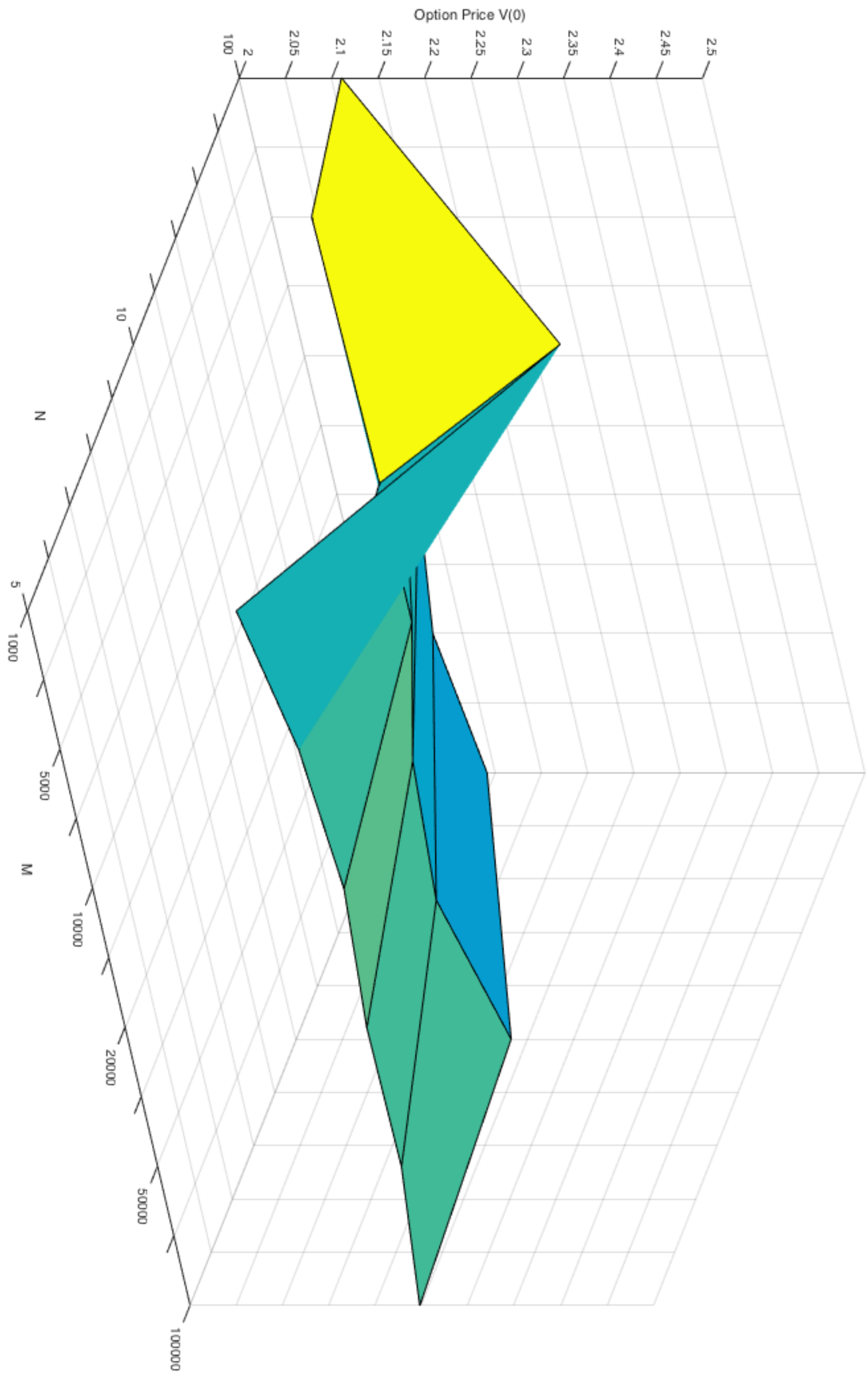

Figure 7.20: 3D-Monte Carlo with low-biased estimator under CIR model (r–Euler scheme, (S_1, S_2)–Euler scheme)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.7524	0.1889	3.6754	0.2199	5.0451	2.2230
5000	2.8346	0.5150	3.4649	1.0403	5.1030	11.1932
10000	2.8863	0.9322	3.4761	2.0637	5.0849	22.7213
20000	2.8701	1.7854	3.4919	4.0568	5.1585	45.1394
50000	2.8593	4.5222	3.5394	10.2142	5.0493	111.2069
100000	2.8446	9.1876	3.5143	20.9338	5.1514	222.0717

Table 7.11: Monte Carlo with high-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–Milstein scheme)

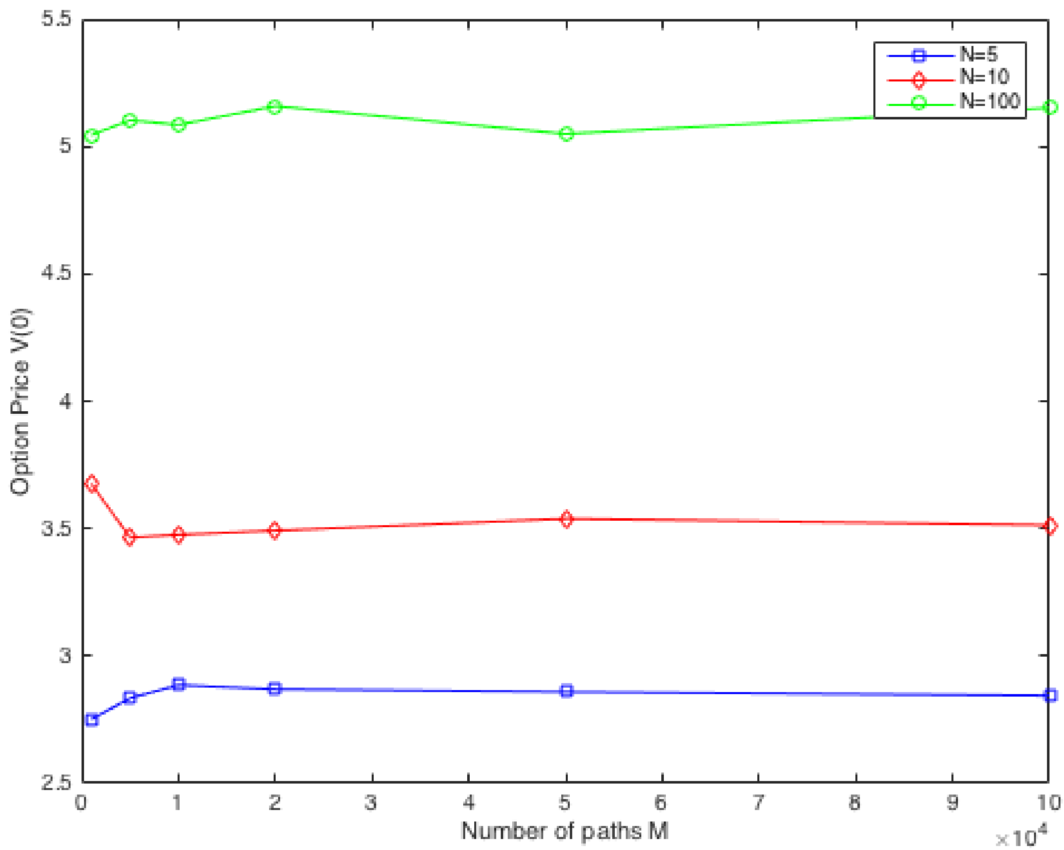

Figure 7.21: 2D-Monte Carlo with high-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–Milstein scheme)

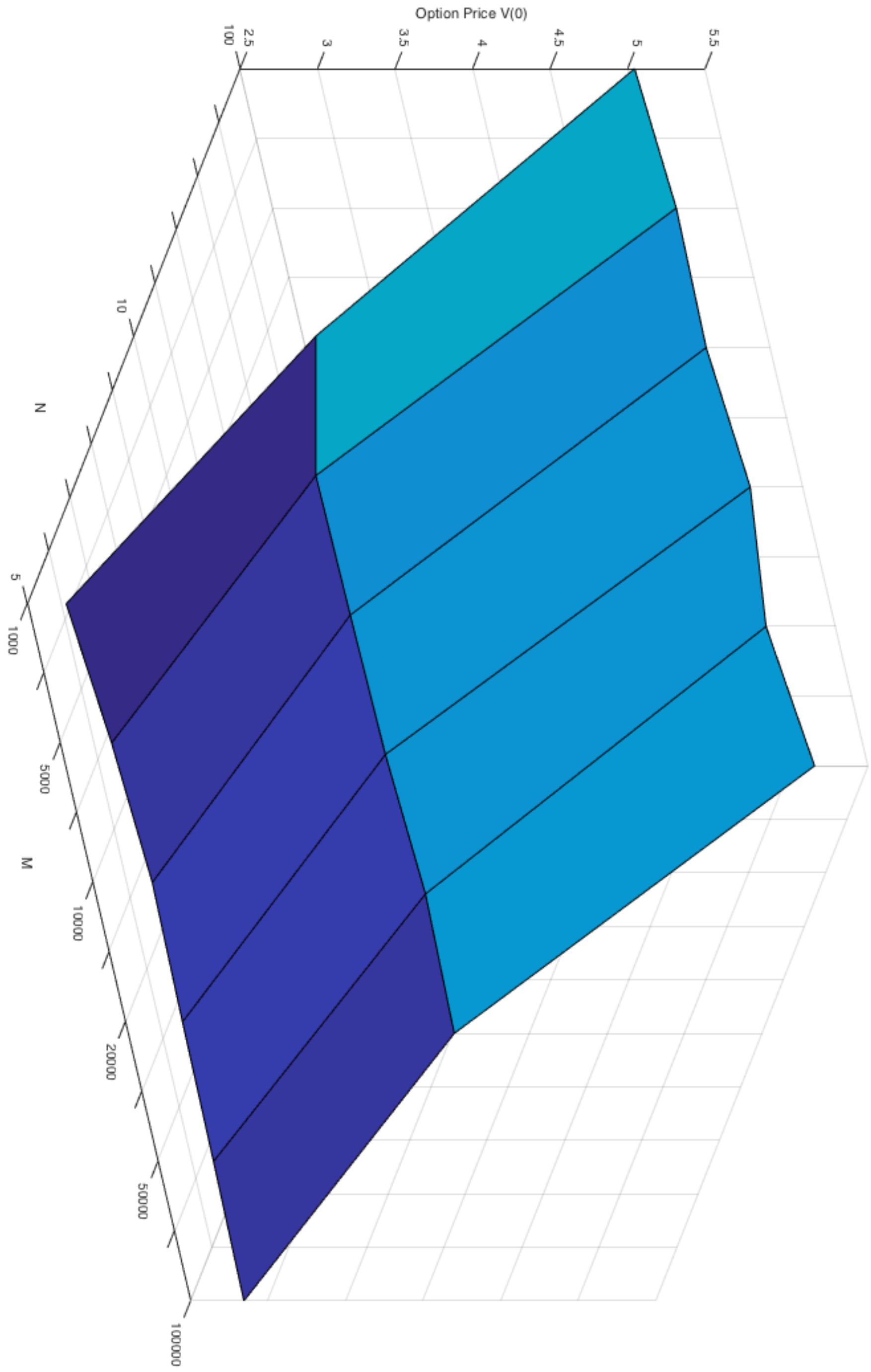

Figure 7.22: 3D-Monte Carlo with high-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–Milstein scheme)

	N=5		N=10		N=100	
M	Price	Time	Price	Time	Price	Time
1000	2.4004	0.1909	2.3235	0.1983	2.1748	2.2148
5000	2.2205	0.5030	2.1996	0.9608	2.0870	11.298410.5842
10000	2.2067	0.8709	2.2238	1.9696	2.0932	21.0577
20000	2.2480	1.7025	2.2171	3.7999	2.1081	42.2822
50000	2.2537	4.1982	2.2351	9.5597	2.0862	105.0757
100000	2.2590	8.4957	2.2157	19.0516	2.0680	210.0973

Table 7.12: Monte Carlo with low-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–Milstein scheme)

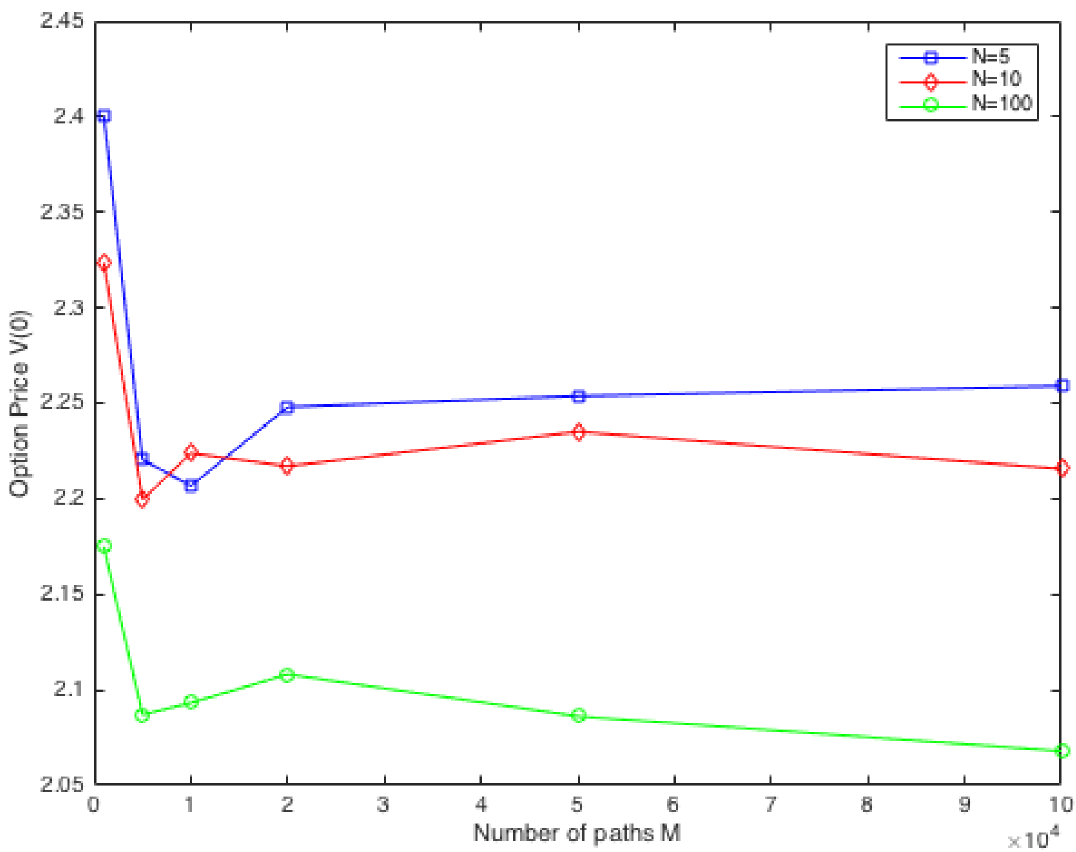

Figure 7.23: 2D-Monte Carlo with low-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–Milstein scheme)

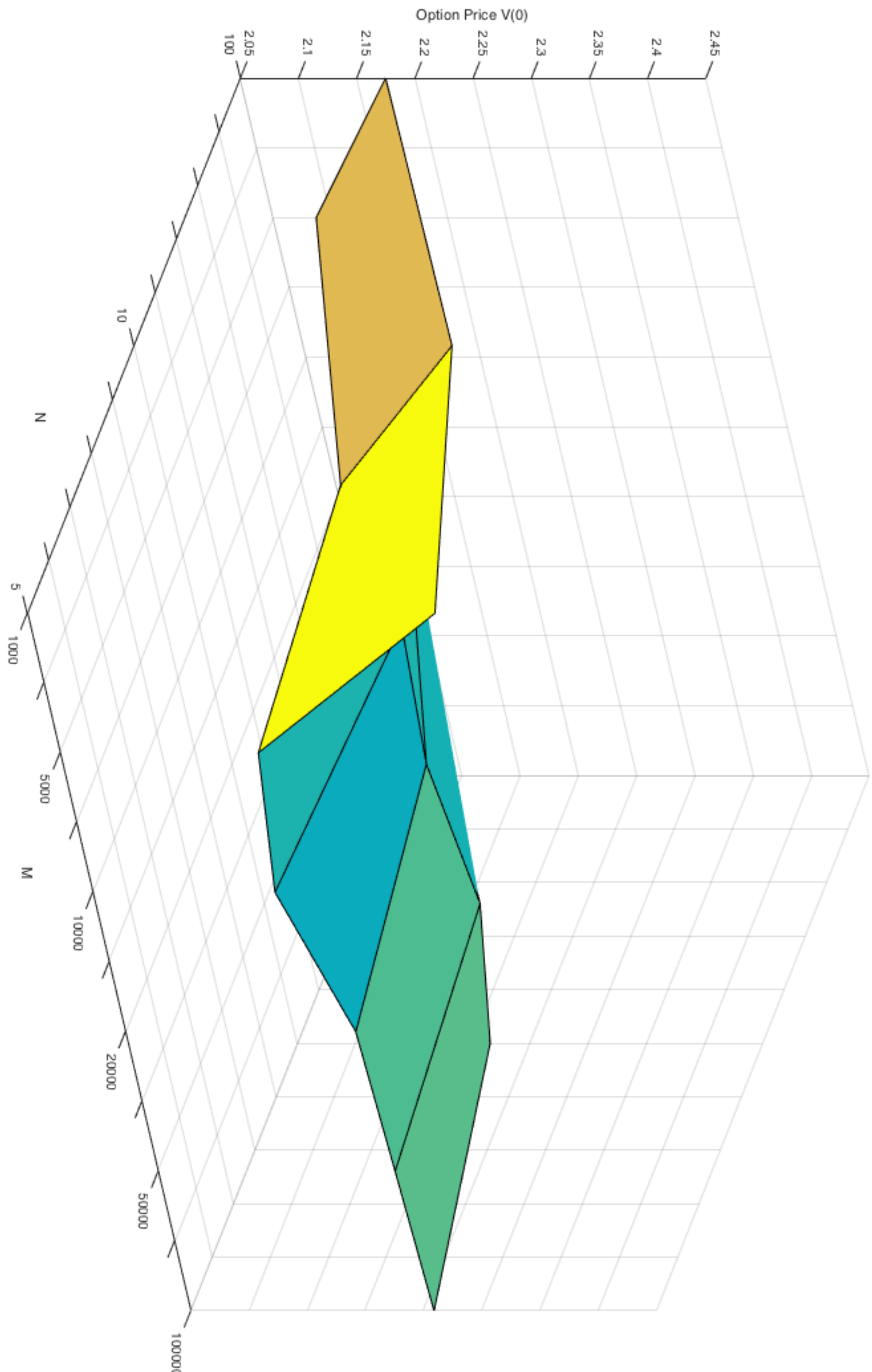

Figure 7.24: 3D-Monte Carlo with low-biased estimator under CIR model (r–Milstein scheme, (S_1, S_2)–Milstein scheme)

Based on these data, we can summarize the differences and characteristics among these models:

In the case of comparing the high-biased and low-biased estimators under these three interest rate models: (1) Under each interest rate model, the average of the price data for the high-biased estimator is bigger than that for the low-biased estimator. This result coincides with the description of the two estimators which was showed before, i.e., $E[\widehat{V}_0] \geq V_0$ for the high-biased estimator and $E[\widehat{V}_0] < V_0$ for the low-biased estimator. (2) For each high-biased estimator under these three interest rate models, the option price V is increasing as the number of the exercise chances N is increasing; on the other hand, for each low-biased estimator under these three interest rate models, the option price V is decreasing as N is increasing. (3) For each high-biased estimator under these three interest rate models, those three lines(blue, red, green) which indicate the evolution of the option prices have a trend of divergence as the number of the independent paths M increases; on the other hand, for each low-biased estimator under these three interest models, those three lines have a trend of convergence as M increases. This suggests that the low-biased estimators are more stable than the high-biased estimators.

In the case of comparing the behaviors of the estimators under the constant interest rate and the stochastic interest rate models: Both the high-biased and the low-biased estimator under the stochastic interest models give relatively close results, while the estimator in the constant interest rate model produces option prices far bigger than those in the stochastic interest rate models. The reason behind this phenomenon comes from the stochastic interest models

$$dr(t) = (\alpha - \beta r(t))dt + \cdots = \beta(\alpha/\beta - r(t))dt + \cdots .$$

In the parenthesis, α/β is the long-term mean of the interest rate. Recall that we set it as $\alpha/\beta = 1/2 = 0.5$ in the implementation above. However, the initial interest rate value was set as 0.05 for all the models. Thus the stochastic interest rates will evolve towards their long term means while the constant interest rate remains unchanged. The increasing of the interest rate leads to a rapid percentage growth of the underlying assets and then bigger difference between two assets $S_1 - S_2$. In other words, we may have smaller payoffs $(K - (S_1 - S_2))^+$ under the stochastic interest rate models than that under the constant interest model.

7.3 Dual Method

7.3.1 Introduction. In the last section, we have used the regression method to derive a high-biased estimator for the option price. The purpose of this section is to find a tighter upper bound for the option price. Together with the low estimator, we can determine a small interval that the option price lives in.

Unlike the regression method which is always seeking for the maximum elements among the intrinsic value and the continuous value, the duality method requires us to solve a max-min question:

$$V_0(X_0) = \sup_{\tau} E[h_\tau(X_\tau)] \leq \inf_{M} E[\max_{k=1,...,N}\{h_k(X_k) - M_k\}] \tag{7.34}$$

where the infimum is taken over all martingales with initial value 0. It is not hard to see how to get this inequality. Again we use the index k to express the exercise times $t_k, \quad k = 1, ..., N$. Let $\{M_k, \quad k = 0, 1, ..., N\}$ be a discrete martingale with $M_0 = 0$. Then

$$E[h_\tau(X_\tau)] = E[h_\tau(X_\tau) - M_\tau] \leq E[\max_{k=1,...,N}(h_k(X_k) - M_k)]$$

for stopping time τ taking values in the exercise times $\{t_1, ..., t_N\}$. Taking infimum on both sides to obtain

$$E[h_\tau(X_\tau)] \leq \inf_{M} E[\max_{k=1,...,N}(h_k(X_k) - M_k)]$$

Then take the supremum over all the stopping times, we get the thing we desired.

The key reason that Equation (7.34) can be used to practically approximate V_0 is that the minimum can be achieved by a martingale which makes it to an equality. The martingale is given by

$$M_i = \Delta_1 + \cdots + \Delta_i$$

$i = 1, ..., N$ with $M_0 = 0$. Here

$$\Delta_i = V_i(X_i) - E[V_i(X_i)|X_{i-1}] \tag{7.35}$$

$i = 1, ..., N$. It is a martingale since

$$\begin{aligned} E[M_i|X_{i-1}] &= E[\Delta_1 + \cdots + \Delta_i|X_{i-1}] \\ &= E[\Delta_1 + \cdots + \Delta_{i-1}|X_{i-1}] + E[\Delta_i|X_{i-1}] \\ &= E[\Delta_1 + \cdots + \Delta_{i-1}|X_{i-1}] \\ &= E[M_{i-1}|X_{i-1}] \\ &= M_{i-1} \end{aligned}$$

$i = 1, ..., N$. The third equation holds because

$$\begin{aligned} E[\Delta_i|X_{i-1}] &= E[V_i(X_i) - E[V_i(X_i)|X_{i-1}]|X_{i-1}] \\ &= E[V_i(X_i)|X_{i-1}] - E[E[V_i(X_i)|X_{i-1}]X_{i-1}] \\ &= E[V_i(X_i)|X_{i-1}] - E[V_i(X_i)|X_{i-1}] \\ &= 0. \end{aligned}$$

The last equation holds because M_{i-1} is $\sigma\{X_{i-1}\}$-measurable.

We use induction to show that

$$V_i(X_i) = \max\{h_i(X_i), h_{i+1}(X_{i+1}) - \Delta_{i+1}, ..., h_N(X_N) - \sum_{j=N}^{i+1} \Delta_j\} \tag{7.36}$$

$i = 1, ..., N$. We have seen that $V_m(X_m) = h_m(X_m)$. Now if it holds for i, then

$$\begin{aligned} V_{i-1}(X_{i-1}) &= \max\{h_{i-1}(X_{i-1}), E[V_i(X_i)|X_{i-1}]\} \\ &= \max\{h_{i-1}(X_{i-1}), V_i(X_i) - \Delta_i\} \end{aligned}$$

Since Equation (7.36) holds for i, extending $V_i(X_i)$ will complete the induction. In particular,

$$V_1(X_1) = \max_{k=1,...,N}(h_k(X_k) - (M_k - \Delta_1))$$

Thus by Equation (7.35),

$$V_0(X_0) = E[V_1(X_1)|X_0] = V_1(X_1) - \Delta_1 = \max_{k=1,...,N}(h_k(X_k) - M_k).$$

So the question of computing the price V_0 becomes computing the martingale $\{M_k\}$, or equivalently computing

$$\Delta_i = V_i(X_i) - C_{i-1}(X_{i-1}).$$

The No.1 principle of the duality formulation is that $\{\widehat{M}_k\}$, the approximation to the martingale M_k, is still a martingale. If it is true, then according to Equation (7.34),

$$\max_{k=1,...,N}(h_k(X_k) - \widehat{M}_k)$$

is a valid upper bound for the option price. The martingale principle is equivalent to

$$E[\widehat{\Delta}_i|X_{i-1}] = 0. \tag{7.37}$$

Now for

$$\widehat{\Delta}_i = \widehat{V}_i(X_i) - \widehat{C}_{i-1}(X_{i-1}), \tag{7.38}$$

we can approximate the first term using the technique in the last section, namely

$$\widehat{V}_i(X_i) = \max\{h_i(X_i), \widehat{C}_i(X_i)\}.$$

But for the second term, if we continue using the regression method to approximate the continuation value, the martingale principle would fail. However, since

$$\widehat{C}_{i-1}(X_{i-1}) = E[\widehat{V}_i(X_i)|X_{i-1}].$$

We can simply use average to approximate the conditional expectation. Fix X_{i-1}, we generate n independent random variable $\{X_i^{(j)}\}, \quad j = 1, ..., M$ of the next exercise time. Take the average

$$\frac{1}{M}\sum_{j=1}^{M}\widehat{V}_i(X_i^{(j)})$$

to approximate

$$\widehat{\Delta}_i = \widehat{V}_i(X_i) - \frac{1}{M}\sum_{j=1}^{M}\widehat{V}_i(X_i^{(j)})$$

To show it satisfies our big principle

$$\begin{aligned}
E[\widehat{\Delta}_i|X_{i-1}] &= E[\widehat{V}_i(X_i) - \frac{1}{M}\sum_{j=1}^{M}\widehat{V}_i(X_i^{(j)})|X_{i-1}] \\
&= E[\widehat{V}_i(X_i)|X_{i-1}] - E[\frac{1}{M}\sum_{j=1}^{M}\widehat{V}_i(X_i^{(j)})|X_{i-1}] \\
&= E[\widehat{V}_i(X_i)|X_{i-1}] - \frac{1}{M}\sum_{j=1}^{M}E[\widehat{V}_i(X_i^{(j)})|X_{i-1}] \\
&= E[\widehat{V}_i(X_i)|X_{i-1}] - \frac{1}{M} \times M E[\widehat{V}_i(X_i^{(j)})|X_{i-1}] \\
&= E[\widehat{V}_i(X_i)|X_{i-1}] - E[\widehat{V}_i(X_i^{(j)})|X_{i-1}] \\
&= 0.
\end{aligned}$$

Here we used the fact that $\{X_i^{(j)}\}$ and X_i are independent identical distributions conditioned on X_{i-1}.

7.3.2 Numerical implementations. We now implement the dual algorithm in the last section to approximate the American spread option prices under two stochastic interest models. Again we will compare these results with the constant interest rate case. We keep using the same value for the inputs except that we set $M = \{10, 50, 100, 200, 500, 1000\}$ and $N = \{5, 10, 20\}$. Recall the other parameters are set as $S_1(0) = 50$, $S_2(0) = 40$, $r(0) = 0.05$, $\alpha = 1$, $\beta = 2$, $\sigma_1 = 0.2$, $\sigma_2 = 0.2$, $\sigma_3 = 0.3$, $\rho_1 = 0.5$, $\rho_2 = 0.4$, $\rho_3 = 0.3$, $T = 0.5$ and $K = 10$.

From Table 7.13, 7.14 and 7.15, we see overall that the American option prices are decreasing as M and N are increasing. In other word, the upper bound is getting tighter and tighter if have a finer grid.

The dual method under CIR model did much better than that under the other two models. As M and N increases, the prices under the CIR model is more stable than the prices under the other models.

	N=5		N=10		N=20	
M	Price	Time	Price	Time	Price	Time
10	1.4725	0.0247	5.4344	0.0496	6.0062	0.1646
50	4.5122	0.0334	2.3355	0.1124	0.6696	0.3624
100	0.9313	0.0429	3.0445	0.1568	3.1882	0.5682
200	2.1713	0.0720	3.0414	0.2532	6.1207	0.9814
500	2.5148	0.1365	3.5693	0.5391	3.8729	2.4356
1000	7.3304	0.2354	3.1994	1.0811	2.7829	4.7259

Table 7.13: Duality under constant interest rate model $((S_1(t), S_2(t))$-explicit)

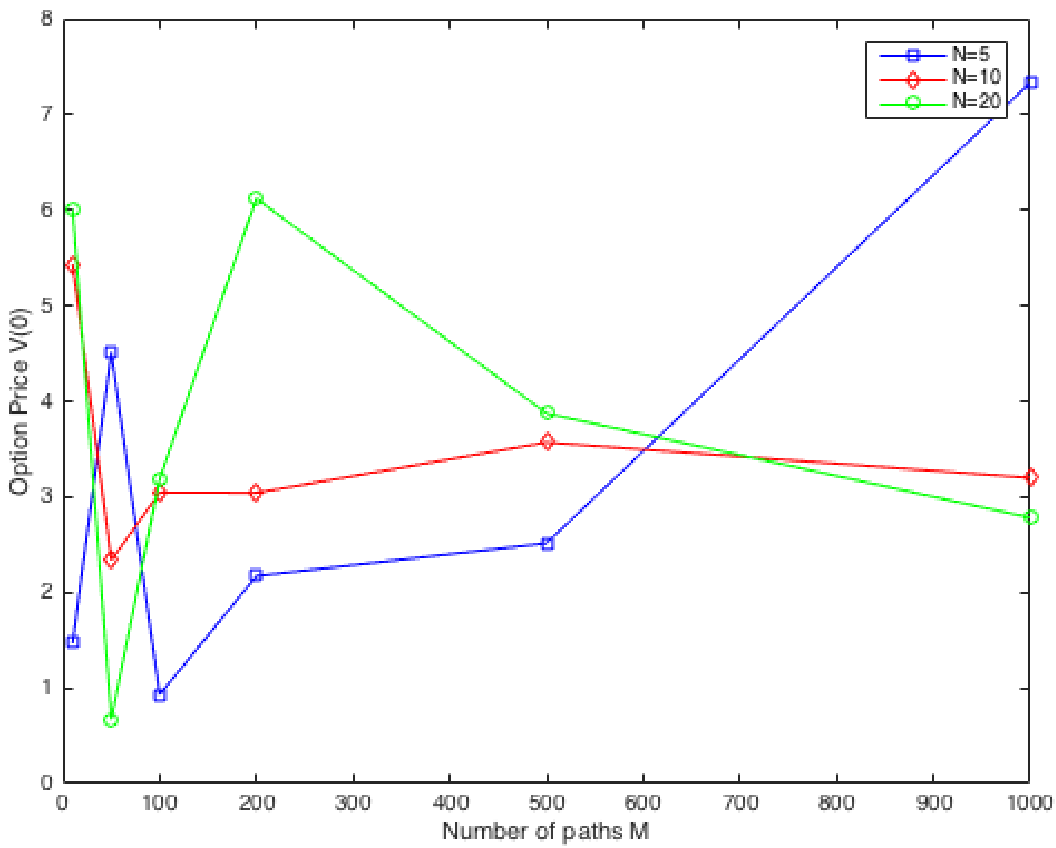

Figure 7.25: 2D-Duality under constant interest rate model $((S_1(t), S_2(t))$-explicit)

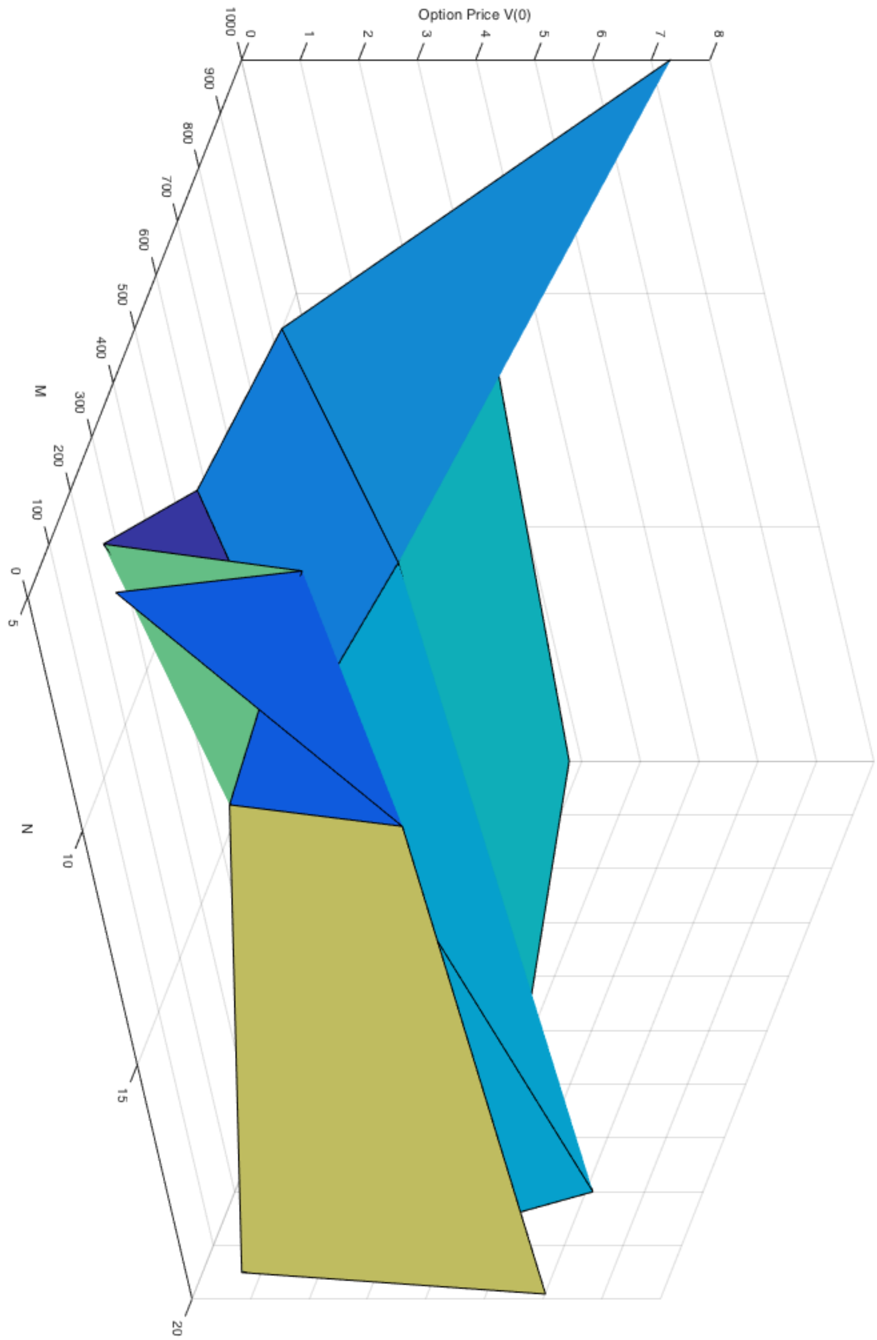

Figure 7.26: 3D-Duality under constant interest rate model $((S_1(t), S_2(t))$-explicit)

	N=5		N=10		N=20	
M	Price	Time	Price	Time	Price	Time
10	4.8967	0.0229	2.6760	0.0679	4.2786	0.2394
50	3.3871	0.0481	6.3426	0.1701	2.8119	0.3896
100	2.2351	0.0592	-0.9121	0.2087	4.5666	0.5823
200	3.6658	0.0817	3.3954	0.2650	4.3678	0.9732
500	4.4409	0.1438	5.9959	0.5464	1.8340	2.1786
1000	1.9274	0.2639	2.4602	1.0144	2.9386	4.1563

Table 7.14: Duality under Vasicek model ($r(t)$-Euler, $(S_1(t), S_2(t))$-explicit)

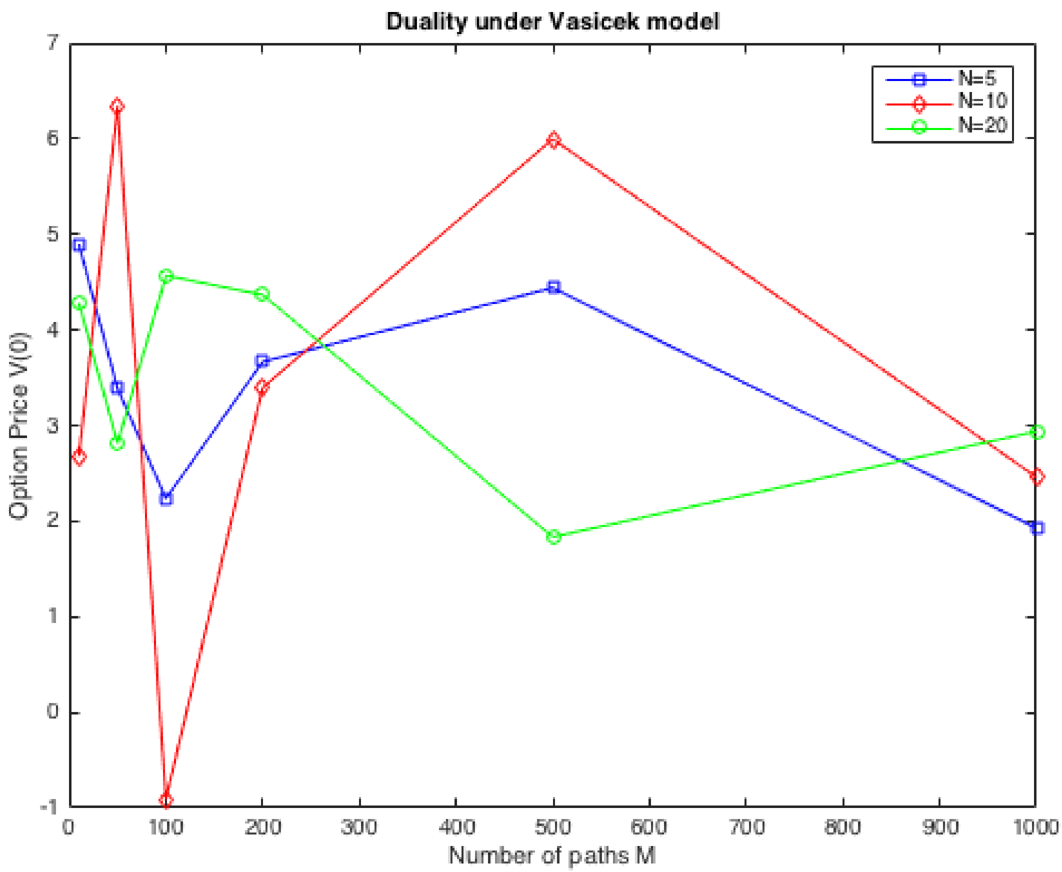

Figure 7.27: 2D-Duality under Vasicek model ($r(t)$-Euler, $(S_1(t), S_2(t))$-explicit)

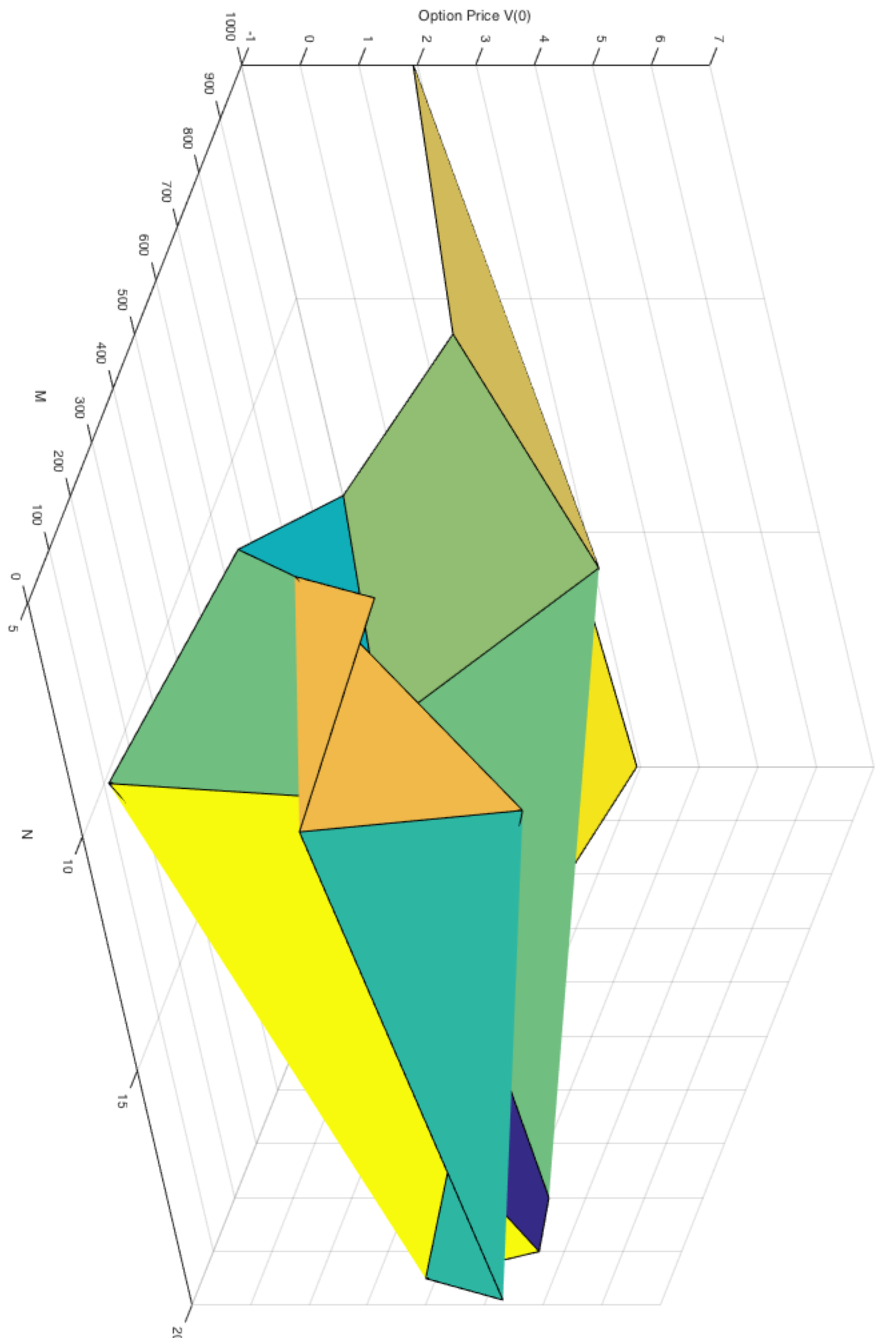

Figure 7.28: 3D-Duality under Vasicek model ($r(t)$-Euler, $(S_1(t), S_2(t))$-explicit)

	N=5		N=10		N=20	
M	Price	Time	Price	Time	Price	Time
10	7.8646	0.0393	-0.4352	0.0726	7.0401	0.2788
50	4.2871	0.0625	5.4728	0.1548	6.7649	0.3813
100	0.9778	0.0637	2.0144	0.1857	2.3866	0.5657
200	4.1056	0.0871	1.3412	0.2625	5.0917	1.0600
500	2.7548	0.1470	2.5185	0.5406	3.3313	2.1866
1000	3.9301	0.2439	2.5753	1.0055	2.9970	4.1989

Table 7.15: Duality under CIR model ($r(t)$-Euler, $(S_1(t), S_2(t))$-explicit)

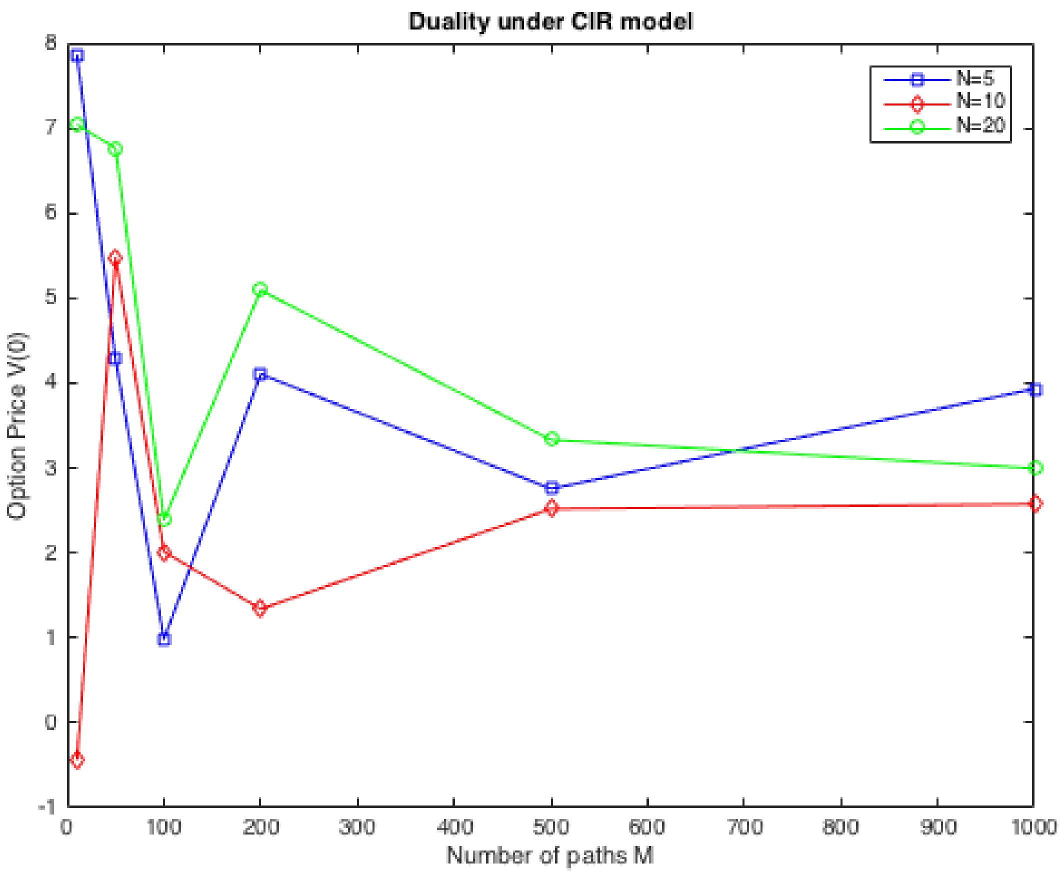

Figure 7.29: 2D-Duality under CIR model ($r(t)$-Euler, $(S_1(t), S_2(t))$-explicit)

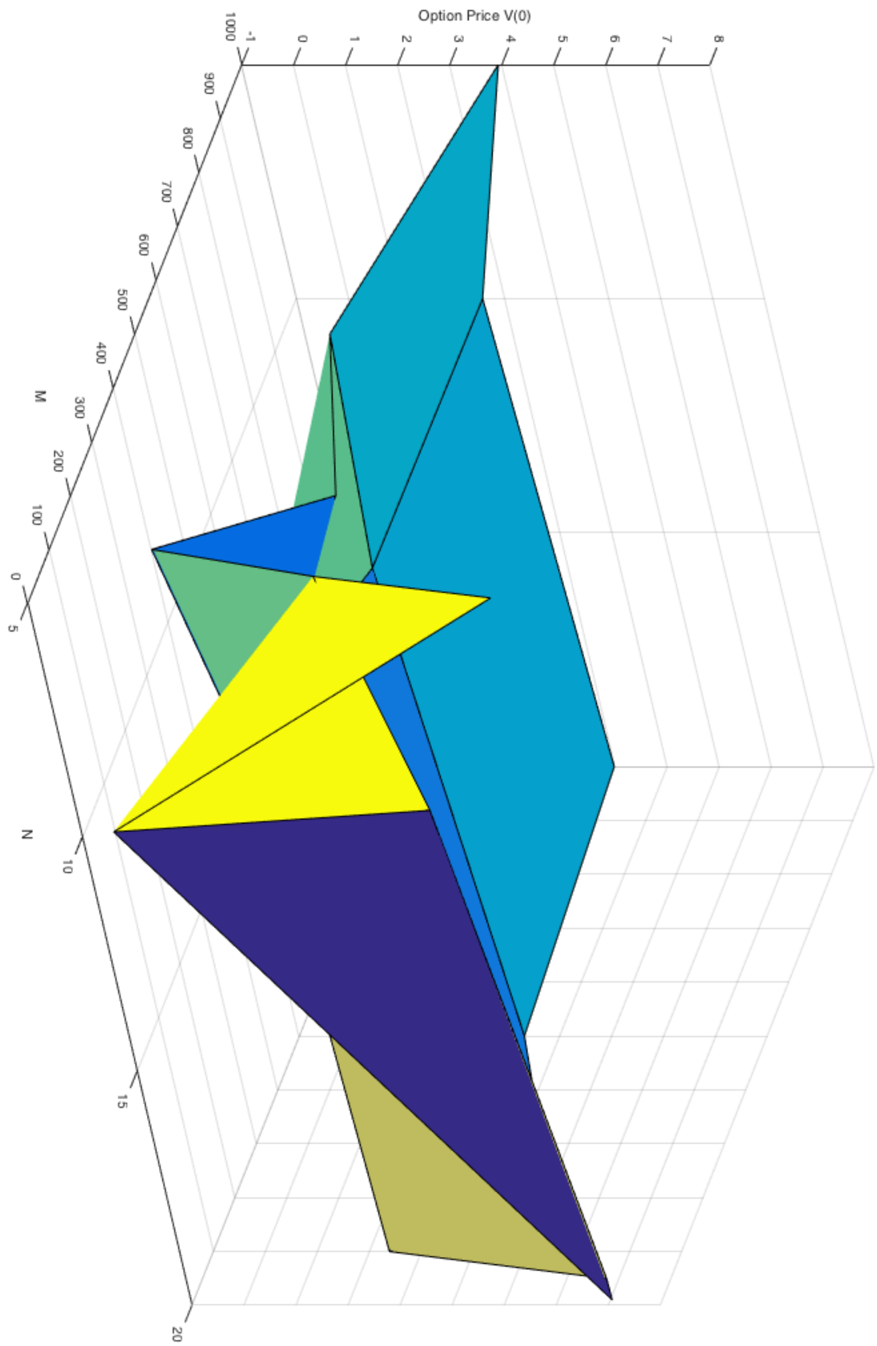

Figure 7.30: 3D-Duality under CIR model ($r(t)$-Euler, $(S_1(t), S_2(t))$-explicit)

	N=5		N=10		N=20	
M	Price	Time	Price	Time	Price	Time
10	2.0529	0.0547	2.2907	0.0529	8.9518	0.1799
50	7.5424	0.0388	2.1723	0.1582	3.9779	0.3996
100	0.6387	0.1045	3.4671	0.1820	2.7465	0.7043
200	4.7809	0.0979	4.3321	0.2839	2.4456	1.0762
500	4.0037	0.1576	2.6695	0.6007	5.9982	2.2830
1000	1.6641	0.3088	1.5114	1.0891	2.6819	4.6409

Table 7.16: Duality under CIR model ($r(t)$-Milstein, $(S_1(t), S_2(t))$-explicit)

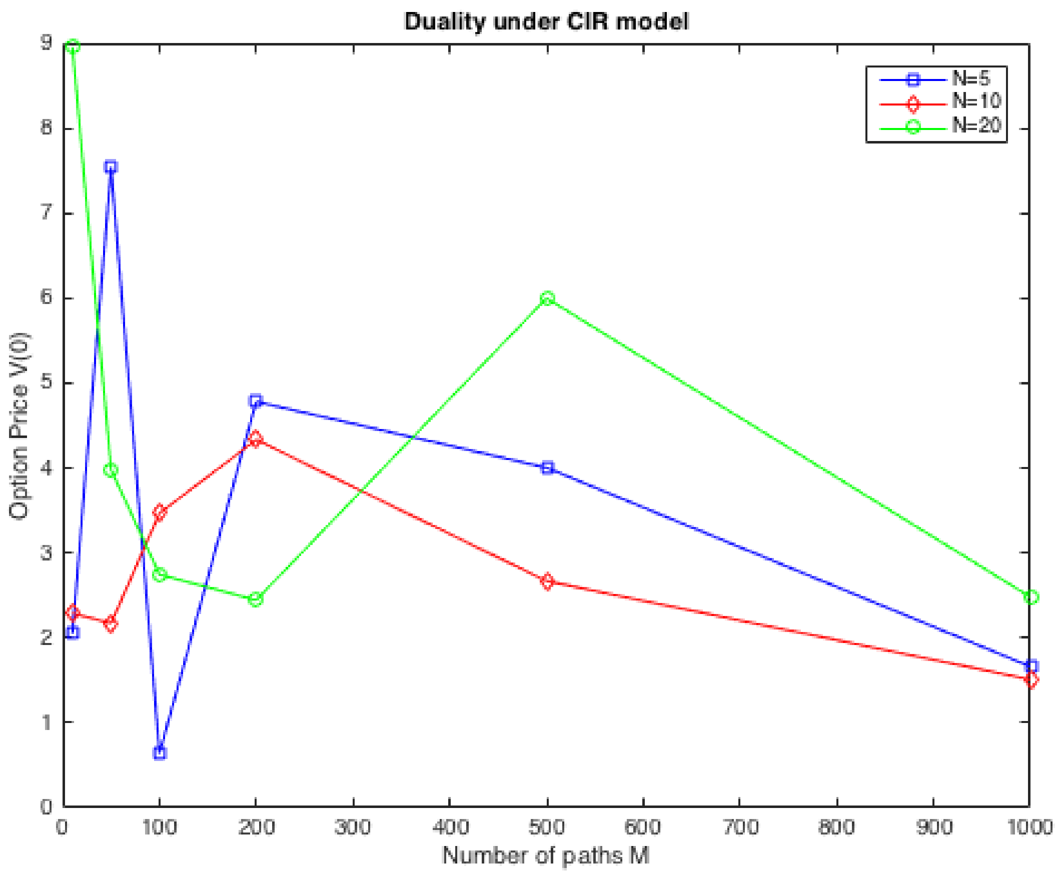

Figure 7.31: 2D-Duality under CIR model ($r(t)$-Milstein, $(S_1(t), S_2(t))$-explicit)

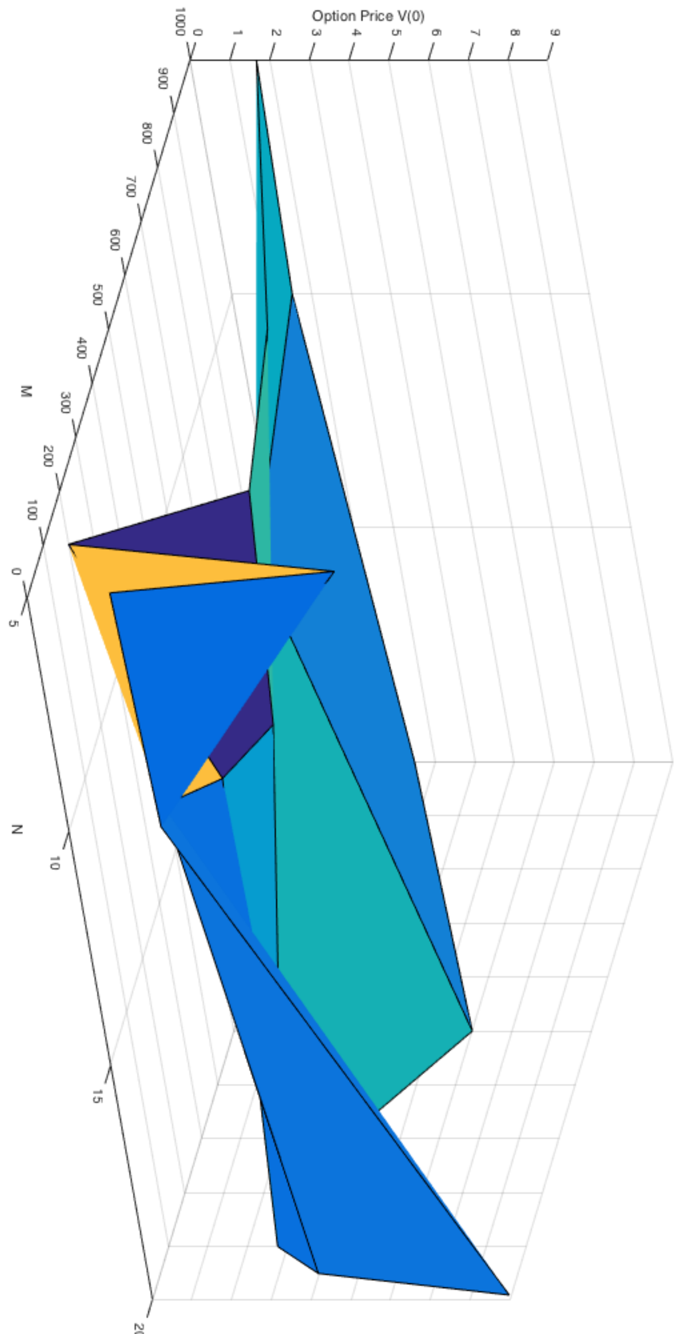

Figure 7.32: 3D-Duality under CIR model ($r(t)$-Milstein, $(S_1(t), S_2(t))$-explicit)

	N=5		N=10		N=20	
M	Price	Time	Price	Time	Price	Time
10	2.0529	0.0547	2.2907	0.0529	8.9518	0.1799
50	7.5424	0.0388	2.1723	0.1582	3.9779	0.3996
100	0.6387	0.1045	3.4671	0.1820	2.7465	0.7043
200	4.7809	0.0979	4.3321	0.2839	2.4456	1.0762
500	4.0037	0.1576	2.6695	0.6007	5.9982	2.2830
1000	1.6641	0.3088	1.5114	1.0891	2.6819	4.6409

Table 7.17: Duality under CIR model ($r(t)$-Euler, $(S_1(t), S_2(t))$-Euler)

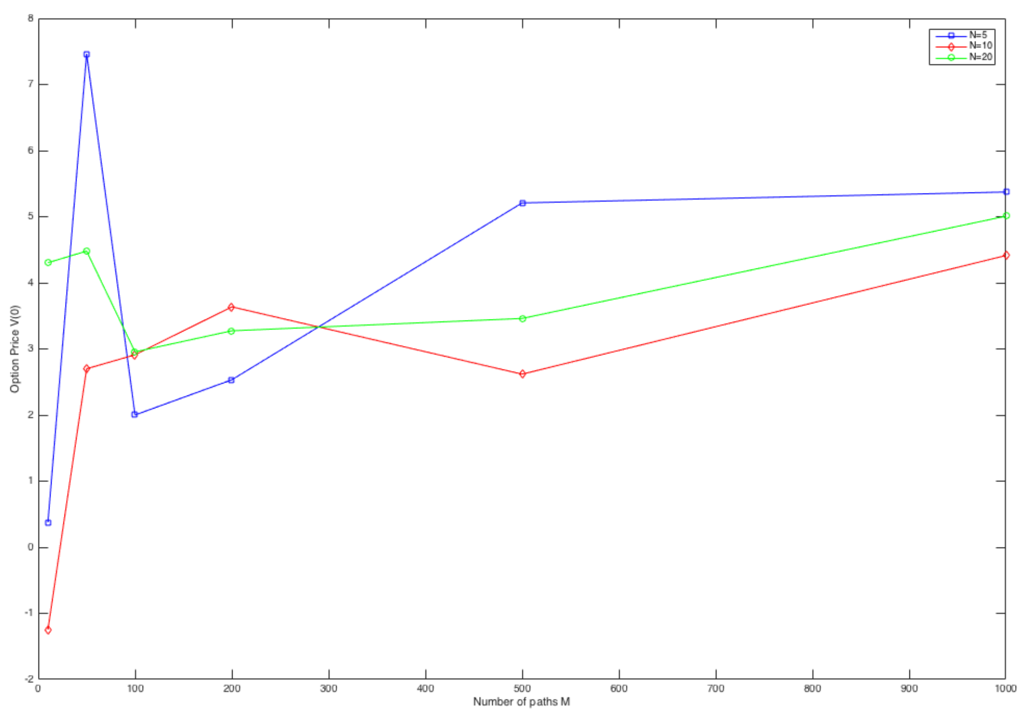

Figure 7.33: 2D-Duality under CIR model ($r(t)$-Euler, $(S_1(t), S_2(t))$-Euler)

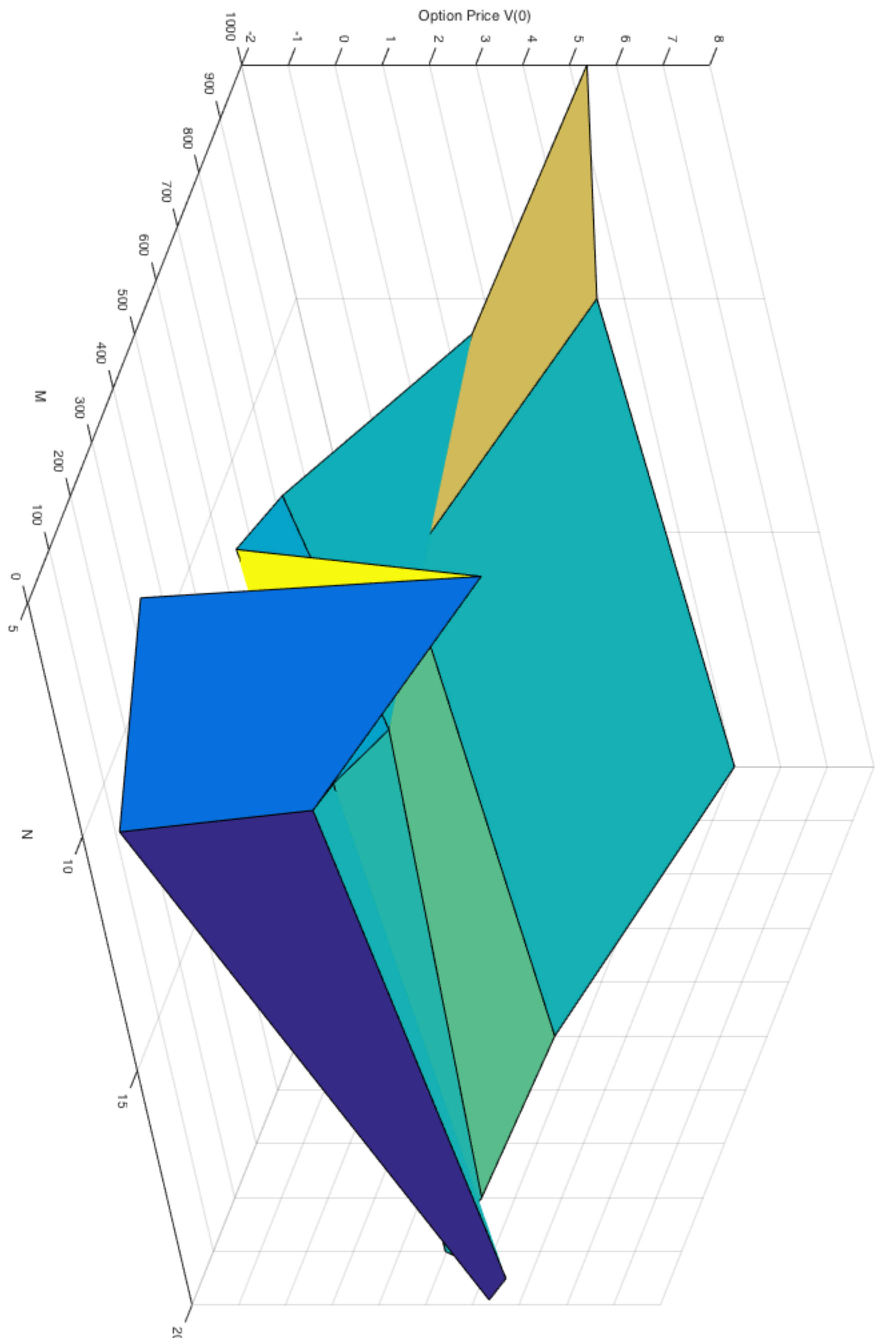

Figure 7.34: 3D-Duality under CIR model ($r(t)$-Euler, $(S_1(t), S_2(t))$-Euler)

	N=5		N=10		N=20	
M	Price	Time	Price	Time	Price	Time
10	2.0529	0.0547	2.2907	0.0529	8.9518	0.1799
50	7.5424	0.0388	2.1723	0.1582	3.9779	0.3996
100	0.6387	0.1045	3.4671	0.1820	2.7465	0.7043
200	4.7809	0.0979	4.3321	0.2839	2.4456	1.0762
500	4.0037	0.1576	2.6695	0.6007	5.9982	2.2830
1000	1.6641	0.3088	1.5114	1.0891	2.6819	4.6409

Table 7.18: Duality under CIR model ($r(t)$-Milstein, $(S_1(t), S_2(t))$-Milstein)

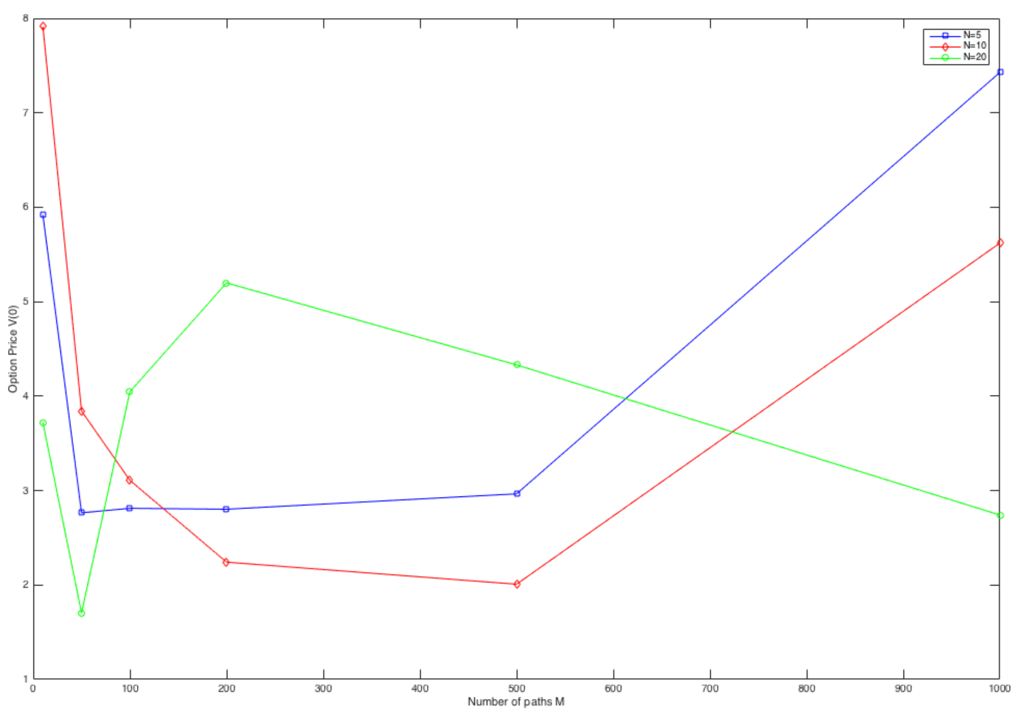

Figure 7.35: 2D-Duality under CIR model ($r(t)$-Milstein, $(S_1(t), S_2(t))$-Milstein)

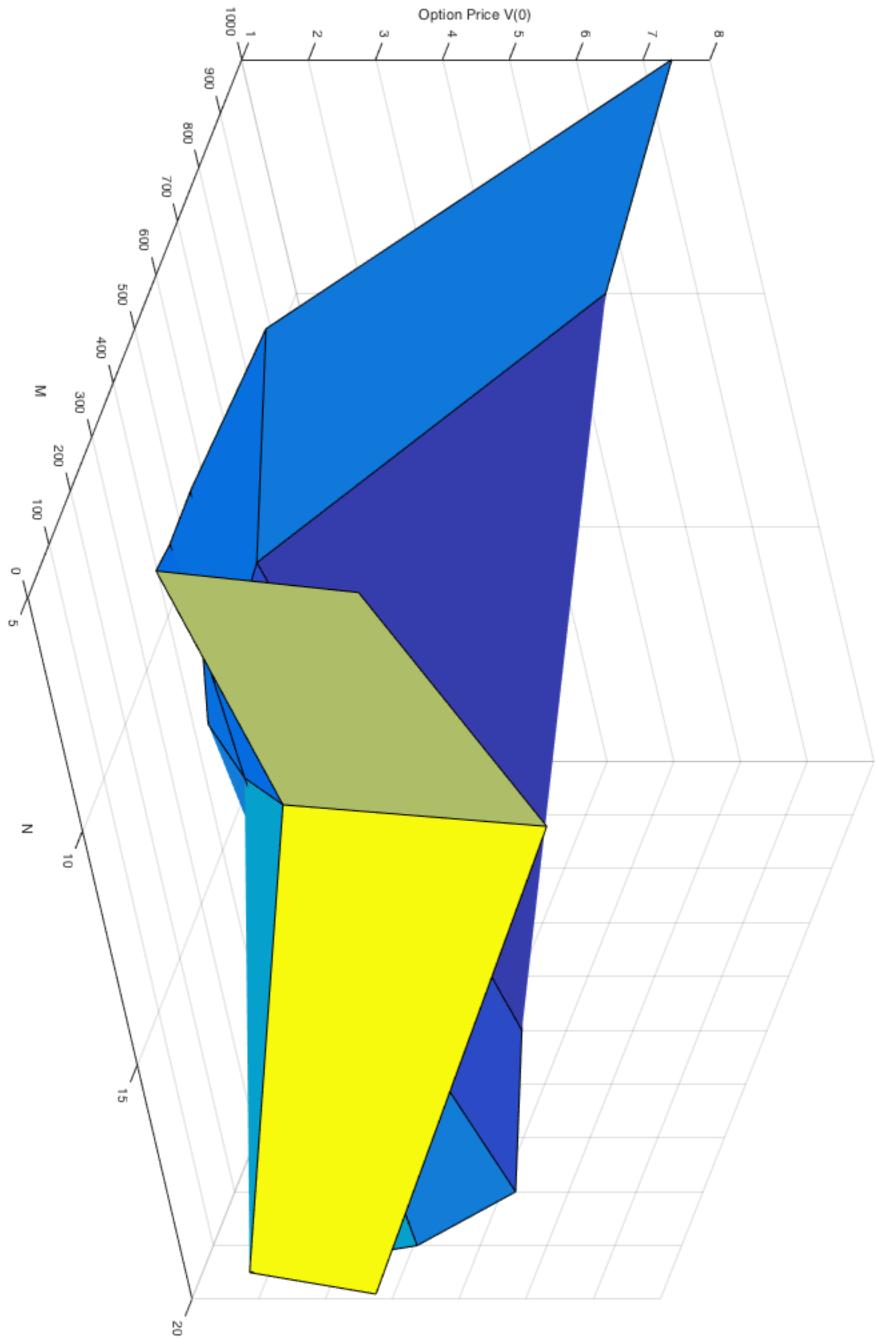

Figure 7.36: 3D-Duality under CIR model ($r(t)$-Milstein, $(S_1(t), S_2(t))$-Milstein)

7.4 The Comparison of Numerical Methods

Now we compare the four numerical methods we used to compute the American spread put option price: regression-based method with high-biased estimator, Longstaff and Schwartz method with low-biased estimator, the dual method and the finite difference method in Chapter 6. The comparisons will be done for the constant interest rate, Vasicek model and CIR model respectively. We set the exercise time number $N = 20$.

	Constant interest rate		
M	High	Low	Duality
10	6.5175	2.9314	6.0062
50	5.2964	2.6153	0.6696
100	4.7202	2.5023	3.1882
200	3.8585	2.5094	6.1207
500	4.1888	2.2346	3.8729
1000	4.1162	2.1706	2.7829

Table 7.19: Algorithms comparison with constant interest rate

	Vasicek model			
M	High	Low	Duality	PDE
10	3.5721	3.7459	4.2786	
50	3.0235	2.9648	2.8119	
100	3.9147	1.9686	4.5666	
200	4.4055	2.1023	4.3678	
500	4.2458	1.9719	1.8340	
1000	4.2563	2.3619	2.9386	2.6350

Table 7.20: Algorithms comparison under Vasicek model

	CIR model			
M	High	Low	Duality	PDE
10	3.7805	1.0033	7.0401	
50	3.6117	2.9259	6.7649	
100	4.8747	2.3451	2.3866	
200	4.1241	2.0881	5.0917	
500	3.8851	2.1799	3.3313	
1000	4.0216	2.2197	2.9970	2.6283

Table 7.21: Algorithms comparison under CIR model

For the Constant interest rate, Vasicek and the CIR models, the prices in the Duality columns are generally lower than the prices in the High estimator columns and are higher than the prices in the Low estimator columns. This is particularly obvious for the last two rows in Table 7.19, Table 7.20 and Table 7.21 when we have more independent paths M. For the stochastic interest rate cases, we can also compare the results in the last rows with the American spread put option prices which are obtained by the partial differential equation approach in the previous chapter. For the last row of Table 7.20, as we discussed, the real option price should be inside the interval $(2.3619, 2.9386)$. It happens that 2.6352, the option price obtained by the partial differential equation approach, is exactly in this interval. Similarly, for the last row of Table 7.21, 2.6283 is also located inside the interval $(2.2197, 2.9970)$. This, to some extent, demonstrates the reliability of both the partial differential equation approach and the direct Monte Carlo simulation method.

For the options of multidimensional underlying assets, the spread option is just a special kind. Instead of depending on the difference of the two underlying assets, we can construct many other payoffs:

- American max option

$$\widetilde{h}(S_1(t), S_2(t)) = (K - \max(S_1(t), S_2(t)))^+$$

- American min option

$$\widetilde{h}(S_1(t), S_2(t)) = (K - \min(S_1(t), S_2(t)))^+$$

- Asian-American spread option

$$\widetilde{h}(S_1(t), S_2(t)) = (K - (\frac{1}{t}\int_0^t S_1(u)du - S_2(t)))^+$$

or some option not existing in the world with payoff

$$\widetilde{h}(S_1(t), S_2(t)) = (K - S_1(t) \times S_2(t))^+.$$

Complicated form of payoff functions may results in difficult situation for option pricing.

Furthermore, for the underlying assets and interest rate models, we may incorporate stochasticity into the volatility and correlation parameters. It is well known that the implied volatility is not a constant, and the volatility smile phenomenon has be detected. Also, the dependence of the underlying assets should vary randomly as time elapses.

More techniques of numerical computation can be used: Fast Fourier Transform, stochastic mesh methods on Monte Carlo simulation, numerical integration methods and method of lines. Our goal is to find a more effective approximation algorithm which is less time consuming and more accurate.

Finally, we can further consider jump underlying asset processes. Since in the real world, the price of a tradable asset (like a stock) is given as a set of discrete numbers, it is important to consider discontinuous asset and their different behaviors.

Bibliography

[1] Andersen, L. and Broadie, M. (2004). Primal-dual simulation algorithm for pricing multidimensional American options. *Management Science*, 50(9):1222–1234.

[2] Baxter, M. and Rennie, A. (1996). *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge University Press, Cambridge, England.

[3] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

[4] Brealey, R. A., Myers, S. C., and Allen, F. (2011). *Principles of Corporate Finance. 10th edition*. McGraw-Hill, New York City, NY.

[5] Broadie, M. and Cao, M. (2008). Improved lower and upper bound algorithms for pricing American options by simulation. *Quantitative Finance*, 8(8):845–861.

[6] Broadie, M. and Glasserman, P. (1997). Pricing American-style securities by simulation. *Journal of Economic Dynamics and Control*, 21(8–9):1323–1352.

[7] Broadie, M. and Glasserman, P. (2004). A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance*, 7(4):41–67.

[8] Carmona, R. and Durrleman, V. (2003). Pricing and hedging spread options. *SIAM Review*, 45(4):627–685.

[9] Chen, S. (2010). *Asian Spread Option Pricing Models and Computation*. PhD thesis, Brigham Young University.

[10] Chiarella, C. and Ziveyi, J. (2011). Method of lines approach for pricing American spread options. *SSRN Electronic Journal. Available at: http://ssrn.com/abstract=2019353*.

[11] Chiarella, C. and Ziveyi, J. (2014). Pricing American options written on two underlying assets. *Quantitative Finance*, 14(3):409–426.

[12] Clement, E., Lamberton, D., and Protter, P. (2002). An analysis of a least squares regression algorithm for American option pricing. *Finance and Stochastics*, 6(4):449–471.

[13] Dempster, M. A. H. and Hong, S. S. G. (2002). *Mathematical Finance — Bachelier Congress 2000: Selected Papers from the First World Congress of the Bachelier Finance Society, Paris, June 29–July 1, 2000*, chapter Spread Option Valuation and the Fast Fourier Transform, pages 203–220. Springer-Verlag, Berlin, Heidelberg.

[14] Deng, S., Johnson, B., and Sogomonian, A. (2001). Exotic electricity options and the valuation of electricity generation and transmission assets. *Decision Support Systems*, 30(3):383–392.

[15] Durrett, R. (2010). *Probability: Theory and Examples. 4th edition*. Cambridge University Press, Cambridge, England.

[16] Evans, L. C. (2012). *An Introduction to Stochastic Differential Equations.* American Mathematical Soc., Providence, RI.

[17] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer-Verlag, New York.

[18] Haowen, F. (2012). European option pricing formula under stochastic interest rate. *Progress in Applied Mathematics*, 4(1):14–21.

[19] Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic intrgrals in the theory of continuous trading. *Stochastic Processes Appl.*, 11(3):215–260.

[20] Haugh, M. B. and Kogan, L. (2004). Pricing American options: a duality approach. *Operations Research*, 52(2):258–270.

[21] Hong, S. S. G. (2001). *Pricing and Hedging of Spread Options with Stochastic Component Correlation.* PhD thesis, University of Cambridge.

[22] Hu, Y. (2013). *American Spread Option Models and Valuation.* PhD thesis, Brigham Young University.

[23] Hull, J. (2011). *Options, Futures, and Other Derivatives. 8th edition.* Prentice-Hall, Englewood Cliffs, NJ.

[24] Hull, J. and White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3(4):573–92.

[25] Jackson, K. R., Jaimungal, S., and Surkov, V. (2008). Fourier space time-stepping for option pricing with lévy models. *Journal of Computational Finance*, 12(2):1–29.

[26] Johnson, R. L., Zulauf, C. R., Irwin, S. H., and Gerlow, M. E. (1991). The soybean complex spread: An examination of market efficiency from the viewpoint of a production process. *Journal of Futures Markets*, 11(1):25–37.

[27] Jones, F. J. (1981). Spreads: Tails, turtles and all that. *Journal of Futures Markets*, 11(4):565–596.

[28] Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus. 2nd edition.* Springer-Verlag, New York.

[29] Kloeden, P. E. and Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, New York.

[30] Kuo, H. (2006). *Introduction to Stochastic Integration.* Springer Science+Business Media, New York.

[31] Longstaff, F. A. and Schwartz, E. S. (2001). Valuing American options by simulation: a simple least-squares approach. *Review of Financial studies*, 14(1):113–147.

[32] Luo, Y. (2012). *Spread Option Pricing with Stochastic Interest Rate.* PhD thesis, Brigham Young University.

[33] Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183.

[34] Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications. 6th edition.* Springer-Verlag, Berlin Heidelberg.

[35] Rogers, L. C. G. (2002). Monte carlo valuation of American options. *Mathematical Finance*, 12(3):271–286.

[36] Shreve, S. E. (2004). *Stochastic Calculus for Finance* II *Continuous-Time Models.* Springer-Verlag, New York.

[37] Tsitsiklis, J. N. and Roy, B. V. (1999). Optimal stopping of markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Transactions on Automatic Control*, 44(10):1840–1851.

[38] Tsitsiklis, J. N. and Roy, B. V. (2001). Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks*, 12(4):694–703.

[39] Wilmott, P., Howison, S., and Dewynne, J. (1995). *The Mathematics of Financial Derivatives: A Student Introduction.* Cambridge University Press, Cambridge, England.

Appendix A. Derivation of the Value of American Option when the Interest Rate is Zero

When $r = 0$, function h only depends on τ, u, that is,

$$h(\tau, u, 0) = h(\tau, u).$$

The computation of $h(\tau, u)$ would be divided according to the continuous region and the stopping region.

In the continuous region the partial differential equations (5.35) and (5.37) both reduce to

$$h_\tau = \frac{1}{2}\sigma_1^2(h_{uu} - h_u) \tag{A.1}$$

The parabolic partial differential equation (A.1) can be further reduced to be heat equation. In fact, we can always express h as

$$h(\tau, u) = f(\tau)g(u)z(\tau, u) \tag{A.2}$$

Its partial derivatives are

$$h_\tau = f_\tau g z + f g z_\tau$$

$$h_u = f g_u z + f g z_u$$

$$h_{uu} = f g_{uu} z + 2 f g_u z_u + f g z_{uu}$$

Inserting these equations to Equation (A.1), we obtain

$$f_\tau g z + f g z_\tau = \frac{1}{2}\sigma_1^2(f g_{uu} z + 2 f g_u z_u + f g z_{uu} - f g_u z - f g z_u) \tag{A.3}$$

Now we decide to let f and g be of the exponential form

$$f(\tau) = c_1 \exp[\hat{f}(\tau)], \quad g(u) = c_2 \exp[\hat{g}(u)] \tag{A.4}$$

where c_1, c_2 are constants. Thus the derivatives become

$$\begin{aligned} f_\tau &= f\hat{f}_\tau \\ g_u &= g\hat{g}_u \\ g_{uu} &= g_u\hat{g}_u + g\hat{g}_{uu} = g\hat{g}_u^2 + g\hat{g}_{uu}. \end{aligned}$$

Inserting these derivatives into Equation (A.3) and simplify it to attain

$$z_\tau = \frac{1}{2}\sigma_1^2 z_{uu} + \frac{1}{2}\sigma_1^2(2\hat{g}_u - 1)z_u + \left[\frac{1}{2}\sigma_1^2(\hat{g}_{uu} + \hat{g}_u^2 - \hat{g}_u) - \hat{f}_\tau\right] z. \tag{A.5}$$

In order for the Equation (A.5) to be a heat equation, we need

$$\begin{aligned} &\frac{1}{2}\sigma_1^2(2\hat{g}_u - 1) = 0 \\ &\frac{1}{2}\sigma_1^2(\hat{g}_{uu} + \hat{g}_u^2 - \hat{g}_u) - \hat{f}_\tau = 0 \end{aligned}$$

which implies

$$\begin{aligned} \hat{f}(\tau) &= -\frac{1}{8}\sigma_1^2\tau + d_1 \\ \hat{g}(u) &= \frac{1}{2}u + d_2 \end{aligned}$$

where d_1, d_2 are constants.

Plugging these expressions back to Equation (A.2) to get

$$\begin{aligned} h(\tau,u) &= f(\tau)g(u)z(\tau,u) \\ &= c_1 e^{\hat{f}(\tau)} c_2 e^{\hat{g}(u)} z(\tau,u) \\ &= c_1 c_2 e^{-\frac{1}{8}\sigma_1^2\tau + d_1} e^{\frac{1}{2}u + d_2} z(\tau,u) \\ &= c_1 c_2 e^{d_1} e^{d_2} e^{-\frac{1}{8}\sigma_1^2\tau + \frac{1}{2}u} z(\tau,u) \\ &= c e^{-\frac{1}{8}\sigma_1^2\tau + \frac{1}{2}u} z(\tau,u). \end{aligned}$$

In the last equation, we let $c = c_1 c_2 e^{d_1} e^{d_2}$. Without loss of generality, we can set $c = 1$. Then the parabolic partial differential equation is reduced to the form of heat equation:

$$z_\tau = \frac{1}{2}\sigma_1^2 z_{uu} \tag{A.6}$$

via the separation of variables

$$h(\tau,u) = e^{-\frac{1}{8}\sigma_1^2\tau + \frac{1}{2}u} z(\tau,u). \tag{A.7}$$

From the initial condition (5.47) for h, we can induce the initial condition for z

$$z(0,u) = e^{-\frac{1}{2}u}(K - e^u)^+. \tag{A.8}$$

Next we are going to solve the heat equation with the initial value. This can be done by the Fourier transform and its inverse.

Definition A.1. The Fourier transform of function $f(x)$ is defined as

$$\mathcal{F}(f(x))(k) = \widetilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ikx} dx \tag{A.9}$$

Property A.2. *Denote $f^{(n)}(x)$ the n-th derivative of $f(x)$, then*

$$\mathcal{F}(f^{(n)}(x))(k) = (ik)^n \mathcal{F}(f(x))(k). \tag{A.10}$$

By Property (A.2), the Fourier transform of the heat equation (A.6) with respect to the spatial variable u is

$$\widetilde{z}_\tau = -\frac{1}{2}\sigma_1^2 k^2 \widetilde{z}. \tag{A.11}$$

Its solution is

$$\widetilde{z}(\tau, k) = \widetilde{z}(0, k)e^{-\frac{1}{2}\cdot\sigma_1^2 k^2 \tau} \tag{A.12}$$

In order to solve for $z(\tau, u)$, the inverse Fourier transform needs to be applied to equation (A.12).

Definition A.3. The inverse Fourier transform of the function $\widetilde{f}(k)$ is defined as

$$\mathcal{F}^{-1}(\widetilde{f}(k))(x) = f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widetilde{f}(k) e^{ikx} dk. \tag{A.13}$$

Theorem A.4 (Convolution Theorem)**.** *If*

$$(f * g)(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y)g(y)dy \tag{A.14}$$

then

$$\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g) \tag{A.15}$$

or

$$\widetilde{f * g} = \widetilde{f} \cdot \widetilde{g} \tag{A.16}$$

If we introduce z_1, z_2 such that

$$\mathcal{F}(z_1) = \widetilde{z_1} = e^{-\frac{1}{2}\sigma_1^2 k^2 \tau}$$
$$\mathcal{F}(z_2) = \widetilde{z_2} = \widetilde{z}(0, k)$$

then Equation (A.12) can be expressed as

$$\widetilde{z}(\tau, k) = \widetilde{z_1}(\tau, k)\widetilde{z_2}(\tau, k). \tag{A.17}$$

Therefore the Convolution Theorem implies that

$$z(\tau, u) = (z_1 * z_2)(\tau, u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} z_1(\tau, u - \xi) z_2(\tau, \xi) d\xi. \tag{A.18}$$

By a simple computation, we obtain

$$z_1 = \mathcal{F}^{-1}(\widetilde{z_1}) = \frac{1}{\sigma_1 \sqrt{\tau}} e^{-\frac{u^2}{2\sigma_1^2 \tau}}$$
$$z_2 = \mathcal{F}^{-1}(\widetilde{z_2}) = z(0, u).$$

Inserting them to Equation (A.18), we therefore have

$$z(\tau, u) = \frac{1}{\sigma_1 \sqrt{2\pi\tau}} \int_{\mathbb{R}} e^{-\frac{(u-\xi)^2}{2\sigma_1^2 \tau}} z(0, \xi) d\xi. \tag{A.19}$$

Plug the initial condition (A.8) into the equation above to get

$$z(\tau,u) = \frac{1}{\sigma_1\sqrt{2\pi\tau}}\int_{\mathbb{R}} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}}\left(Ke^{-\frac{1}{2}\xi} - e^{\frac{1}{2}\xi}\right)^+ d\xi \tag{A.20}$$

$$= \frac{1}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}}\left(Ke^{-\frac{1}{2}\xi} - e^{\frac{1}{2}\xi}\right) d\xi \tag{A.21}$$

$$= \frac{K}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}-\frac{1}{2}\xi} d\xi - \frac{1}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}+\frac{1}{2}\xi} d\xi \tag{A.22}$$

$$= KI_{-\frac{1}{2}} - I_{\frac{1}{2}}. \tag{A.23}$$

In (A.21), we used the fact that $Ke^{-\frac{1}{2}\xi} - e^{\frac{1}{2}\xi} > 0$ if and only if $u < \ln K$. In the last equation we denoted the two integrals as

$$I_{-\frac{1}{2}} = \frac{1}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}-\frac{1}{2}\xi} d\xi$$

$$I_{\frac{1}{2}} = \frac{1}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}+\frac{1}{2}\xi} d\xi.$$

As in the Appendix B, these two integrals can be computed and have a simpler form

$$I_a = e^{au+\frac{a^2\sigma_1^2\tau}{2}}\Phi\left(\frac{\ln K - u - a\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right) \tag{A.24}$$

where

$$\Phi(\zeta) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\zeta} e^{-\frac{\eta^2}{2}} d\eta \tag{A.25}$$

is the cumulative distribution function of the standard normal distribution.

Inserting (A.24) into (A.23), we have

$$z(\tau,u) = Ke^{-\frac{1}{2}u+\frac{\sigma_1^2\tau}{8}}\Phi\left(\frac{\ln K - u + \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right) - e^{\frac{1}{2}u+\frac{\sigma_1^2\tau}{8}}\Phi\left(\frac{\ln K - u - \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right). \tag{A.26}$$

Then by the separation of variables (A.7), we get

$$h(\tau,u) = K\Phi\left(\frac{\ln K - u + \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right) - e^{u}\Phi\left(\frac{\ln K - u - \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right). \tag{A.27}$$

While in the stopping region

$$h(\tau, u) = e^{A(\tau,T)}(K - e^u)^+. \tag{A.28}$$

Thus combine the results in these two regions to get

$$h(\tau, u, 0) = \max\left\{K\Phi\left(\frac{\ln K - u + \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right) - e^u\Phi\left(\frac{\ln K - u - \frac{1}{2}\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right), e^{A(\tau,T)}(K - e^u)^+\right\}. \tag{A.29}$$

Appendix B. Computation of a Class of Integrals

The task of this section is computing and simplifying

$$I_a = \frac{1}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-\frac{(u-\xi)^2}{2\sigma_1^2\tau}+a\xi}d\xi. \tag{B.1}$$

We complete the square for the exponent:

$$-\frac{(u-\xi)^2}{2\sigma_1^2\tau} + a\xi = -c_1(c_2 - \xi)^2 + c_3. \tag{B.2}$$

Expanding both sides gives

$$-\frac{\xi^2}{2\sigma_1^2\tau} + \left(\frac{u}{\sigma_1^2\tau} + a\right)\xi - \frac{u^2}{2\sigma_1^2\tau} = -c_1\xi^2 + 2c_1c_2\xi - c_1c_2^2 + c_3. \tag{B.3}$$

Equating the like terms to get

$$c_1 = \frac{1}{2\sigma_1^2\tau}$$
$$c_2 = u + a\sigma_1^2\tau$$
$$c_3 = au + \frac{a^2\sigma_1^2\tau}{2},$$

so

$$I_a = \frac{e^{c_3}}{\sigma_1\sqrt{2\pi\tau}}\int_{-\infty}^{\ln K} e^{-c_1(c_2-\xi)^2}d\xi. \tag{B.4}$$

Let $\eta = (c_2 - \xi)\sqrt{2c_1}$ then $d\eta = -\sqrt{2c_1}d\xi$. When $\xi \to -\infty$, $\eta \to \infty$ and when $\xi = \ln K$, $\eta = (c_2 - \ln K)\sqrt{2c_1}$. Thus

$$I_a = -\frac{e^{c_3}}{2\sigma_1\sqrt{c_1\pi\tau}}\int_{\infty}^{(c_2-\ln K)\sqrt{2c_1}} e^{-\frac{\eta^2}{2}}d\eta. \tag{B.5}$$

Finally let $\zeta = -\eta$ then $d\zeta = -d\eta$. Therefore,

$$I_a = \frac{e^{c_3}}{2\sigma_1\sqrt{c_1\pi\tau}}\int_{-\infty}^{(\ln K-c_2)\sqrt{2c_1}} e^{-\frac{\zeta^2}{2}}d\zeta \tag{B.6}$$
$$= \frac{e^{c_3}}{\sigma_1\sqrt{2c_1\tau}}\Phi((\ln K - c_2)\sqrt{2c_1}). \tag{B.7}$$

Notice that

$$e^{c_3} = e^{au+\frac{a^2\sigma_1^2\tau}{2}}$$
$$\sigma_1\sqrt{2c_1\tau} = 1$$
$$(\ln K - c_2)\sqrt{2c_1} = \frac{\ln K - u - a\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}.$$

Therefore

$$I_a = e^{au+\frac{a^2\sigma_1^2\tau}{2}}\Phi\left(\frac{\ln K - u - a\sigma_1^2\tau}{\sigma_1\sqrt{\tau}}\right). \tag{B.8}$$

Appendix C. Derivation of the Value of American Spread Option when the Interest Rate is Zero

When $r = 0$, function h only depends on τ, u, v, that is,

$$h(\tau, u, v, 0) = h(\tau, u, v).$$

The computation of $h(\tau, u, v)$ would be divided according to the continuous region and the stopping region.

In the continuous region the partial differential equations (6.43) and (6.44) both reduce to

$$h_\tau = \frac{1}{2}\sigma_1^2(h_{uu} - h_u) + \frac{1}{2}\sigma_2^2(h_{vv} - h_v) + \sigma_1\sigma_2\rho_1 h_{uv}. \tag{C.1}$$

The parabolic partial differential equation (C.1) can be further reduced to heat equation. In fact, we can always express h as

$$h(\tau, u, v) = f(\tau)g(u)p(v)z(\tau, u, v). \tag{C.2}$$

Its partial derivatives are

$$\begin{aligned}
h_\tau &= f_\tau gpz + fgpz_\tau \\
h_u &= fg_u pz + fgpz_u \\
h_v &= fgp_v z + fgpz_v \\
h_{uu} &= fg_{uu}pz + 2fg_u pz_u + fgpz_{uu} \\
h_{vv} &= fgp_{vv}z + 2fgp_v z_u + fgpz_{vv} \\
h_{uv} &= fg_u p_v z + fg_u pz_v + fgp_v z_u + fgpz_{uv}.
\end{aligned}$$

Inserting these equations to Equation (C.1), we obtain

$$\begin{aligned}f_\tau gpz + fgpz_\tau = \frac{1}{2}\sigma_1^2(fg_{uu}pz + 2fg_u pz_u + fgpz_{uu} - fg_u pz - fgpz_u)\\ + \frac{1}{2}\sigma_2^2(fgp_{vv}z + 2fgp_v z_u + fgpz_{vv} - fgp_v z - fgpz_v)\\ + \sigma_1\sigma_2\rho_1(fg_u p_v z + fg_u pz_v + fgp_v z_u + fgpz_{uv}). \quad \text{(C.3)}\end{aligned}$$

Now we decide to let f, g and p be of the exponential form

$$f(\tau) = c_1 \exp[\hat{f}(\tau)], \quad g(u) = c_2 \exp[\hat{g}(u)], \quad p(v) = c_3 \exp[\hat{p}(v)], \quad \text{(C.4)}$$

where c_1, c_2, c_3 are constants. Thus the derivatives become

$$\begin{aligned}f_\tau &= f\hat{f}_\tau\\ g_u &= g\hat{g}_u\\ g_{uu} &= g_u\hat{g}_u + g\hat{g}_{uu} = g\hat{g}_u^2 + g\hat{g}_{uu}\\ p_v &= p\hat{p}_v\\ p_{vv} &= p_v\hat{p}_v + p\hat{p}_{vv} = p\hat{p}_v^2 + p\hat{p}_{vv}.\end{aligned}$$

Inserting these derivatives into Equation (C.3) and simplify it we attain

$$\begin{aligned}z_\tau = \frac{1}{2}\sigma_1^2 z_{uu} + \frac{1}{2}\sigma_2^2 z_{vv} + \frac{1}{2}\sigma_1^2(2\hat{g}_u - 1)z_u + \frac{1}{2}\sigma_2^2(2\hat{p}_v - 1)z_v\\ + \left[\frac{1}{2}\sigma_1^2(\hat{g}_{uu} + \hat{g}_u^2 - \hat{g}_u) + \frac{1}{2}\sigma_2^2(\hat{p}_{vv} + \hat{p}_v^2 - \hat{p}_v) - \hat{f}_\tau\right] z. \quad \text{(C.5)}\end{aligned}$$

In order for the Equation (C.5) to be a heat equation, we need

$$\frac{1}{2}\sigma_1^2(2\hat{g}_u - 1) = 0$$
$$\frac{1}{2}\sigma_2^2(2\hat{p}_v - 1) = 0$$
$$\frac{1}{2}\sigma_1^2(\hat{g}_{uu} + \hat{g}_u^2 - \hat{g}_u) + \frac{1}{2}\sigma_2^2(\hat{p}_{vv} + \hat{p}_v^2 - \hat{p}_v) - \hat{f}_\tau = 0$$

which implies

$$\hat{g}(u) = \frac{1}{2}u + d_1$$
$$\hat{p}(v) = \frac{1}{2}v + d_2$$
$$\hat{f}(\tau) = -\frac{1}{8}(\sigma_1^2 + \sigma_2^2)\tau + d_3$$

where d_1, d_2, d_3 are constants.

Plugging these expressions back to Equation (C.2) to get

$$\begin{aligned} h(\tau, u, v) &= f(\tau)g(u)p(v)z(\tau, u, v) \\ &= c_3 e^{\hat{f}(\tau)} c_1 e^{\hat{g}(u)} c_2 e^{\hat{p}(v)} z(\tau, u, v) \\ &= c_1 c_2 c_3 e^{-\frac{1}{8}(\sigma_1^2+\sigma_2^2)\tau + d_3} e^{\frac{1}{2}u + d_1} e^{\frac{1}{2}v + d_2} z(\tau, u, v) \\ &= c_1 c_2 c_3 e^{d_1} e^{d_2} e^{d_3} e^{-\frac{1}{8}(\sigma_1^2+\sigma_2^2)\tau + \frac{1}{2}(u+v)} z(\tau, u, v) \\ &= c e^{-\frac{1}{8}(\sigma_1^2+\sigma_2^2)\tau + \frac{1}{2}(u+v)} z(\tau, u, v). \end{aligned}$$

In the last equation, we let $c = c_1 c_2 e^{d_1} e^{d_2}$. Without loss of generality, we can set $c = 1$. Then the parabolic partial differential equation is reduced to the form of heat equation:

$$z_\tau = \frac{1}{2}\sigma_1^2 z_{uu} + \frac{1}{2}\sigma_2^2 z_{vv} \tag{C.6}$$

via the separation of variables

$$h(\tau,u,v) = e^{-\frac{1}{8}(\sigma_1^2+\sigma_2^2)\tau+\frac{1}{2}(u+v)}z(\tau,u,v). \tag{C.7}$$

To make it prettier, let

$$\begin{aligned} u = \frac{\sigma_1}{\sqrt{2}}m, \quad v = \frac{\sigma_2}{\sqrt{2}}n \\ l(\tau,m,n) = z(\tau,u,v) \end{aligned}. \tag{C.8}$$

Then

$$l_\tau = \Delta l = l_{mm} + l_{nn}. \tag{C.9}$$

From the initial condition (6.54) for h, we can induce the initial condition for l

$$l(0,m,n) = e^{-\frac{1}{2\sqrt{2}}(\sigma_1 m+\sigma_2 n)}\left[K - (e^{\frac{\sigma_1}{\sqrt{2}}m} - e^{\frac{\sigma_2}{\sqrt{2}}n})\right]^+. \tag{C.10}$$

Next we are going to solve the heat equation with the initial value. This can be done by the two-dimensional Fourier transform and its inverse. (see [16])

$$l(\tau,m,n) = \frac{1}{4\pi\tau}\int_{\mathbb{R}^2} e^{-\frac{|(m,n)-(\xi-\eta)|^2}{4\tau}} l(0,\xi,\eta)d(\xi,\eta). \tag{C.11}$$

Then

$$z(\tau,u,v) = \frac{1}{4\pi\tau}\int_{\mathbb{R}^2} e^{-\frac{(\frac{\sqrt{2}}{\sigma_1}u-\xi)^2+(\frac{\sqrt{2}}{\sigma_2}v-\eta)^2}{4\tau}-\frac{1}{2\sqrt{2}}(\sigma_1\xi+\sigma_2\eta)}\left[K - (e^{\frac{\sigma_1}{\sqrt{2}}\xi} - e^{\frac{\sigma_2}{\sqrt{2}}\eta})\right]^+ d(\xi,\eta). \tag{C.12}$$

Finally,

$$h(\tau,u,v) = e^{-\frac{1}{8}(\sigma_1^2+\sigma_2^2)\tau+\frac{1}{2}(u+v)}z(\tau,u,v)$$
$$= \frac{1}{4\pi\tau}e^{-\frac{1}{8}(\sigma_1^2+\sigma_2^2)\tau+\frac{1}{2}(u+v)}\int_{\mathbb{R}^2} e^{-\frac{(\frac{\sqrt{2}}{\sigma_1}u-\xi)^2+(\frac{\sqrt{2}}{\sigma_2}v-\eta)^2}{4\tau}-\frac{1}{2\sqrt{2}}(\sigma_1\xi+\sigma_2\eta)}\left[K-(e^{\frac{\sigma_1}{\sqrt{2}}\xi}-e^{\frac{\sigma_2}{\sqrt{2}}\eta})\right]^+ d(\xi,\eta). \tag{C.13}$$

While in the stopping region

$$h(\tau,u,v) = (K-(e^u-e^v))^+. \tag{C.14}$$

Thus combine the results in these two regions we get

$$h(\tau,u,v,0) = \max\{(K-(e^u-e^v))^+,$$
$$\frac{1}{4\pi\tau}e^{\frac{1}{4}(\sigma_1^2+\sigma_2^2)\tau+\frac{1}{2}(u+v)}\int_{\mathbb{R}^2} e^{-\frac{(\frac{\sqrt{2}u}{\sigma_1}-\xi)^2+(\frac{\sqrt{2}v}{\sigma_2}-\eta)^2}{4\tau}-\frac{1}{2\sqrt{2}}(\sigma_1\xi+\sigma_2\eta)}\left[K-(e^{\frac{\sigma_1}{\sqrt{2}}\xi}-e^{\frac{\sigma_2}{\sqrt{2}}\eta})\right]^+ d(\xi,\eta)\}. \tag{C.15}$$