To my family
ACKNOWLEDGMENTS

First I would like to thank my advisor Dr. Yan for his kindly help and constant encouragement throughout the research process. I’d also like to thank Dr. Rao, Dr. Hager, Dr. McCullough and Dr. Qiu for their interest in my research. I appreciate the opportunity UF math department has offered me during those years. It’s been a pleasure and an honor to study here, and I’ve learnt a lot since I came here. Finally I’d like to thank my family for their constant support, both mentally and physically. I could not imagine to finish this without all the helps from those people that care about me.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION TO CIR MODEL AND SIMULATION</td>
<td>10</td>
</tr>
<tr>
<td>1.1</td>
<td>The Term Structure Of Interest Rates</td>
<td>10</td>
</tr>
<tr>
<td>1.2</td>
<td>Models Of The Short-Term Interest Rate</td>
<td>14</td>
</tr>
<tr>
<td>1.3</td>
<td>Introduction to CIR Model</td>
<td>17</td>
</tr>
<tr>
<td>1.4</td>
<td>Simulation Issue</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>EULER DISCRETIZATION AND RELATED ISSUES</td>
<td>23</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction To Euler-Maruyama Scheme</td>
<td>23</td>
</tr>
<tr>
<td>2.2</td>
<td>Convergence Order</td>
<td>24</td>
</tr>
<tr>
<td>2.3</td>
<td>Euler Scheme For CIR Process</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>CIR PROCESS EXACT SIMULATION</td>
<td>29</td>
</tr>
<tr>
<td>3.1</td>
<td>Traditional Simulation Method</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>Some Drawbacks In Simulating CIR Process</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Main Result</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>INTRODUCTION TO HESTON MODEL AND SIMULATION</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>Option Pricing Theory</td>
<td>37</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction To Stochastic Volatility Model</td>
<td>43</td>
</tr>
<tr>
<td>4.3</td>
<td>Introduction To Heston’s Model And Simulation Issue</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>SIMULATION RESULTS AND APPLICATIONS IN OPTION PRICING</td>
<td>53</td>
</tr>
<tr>
<td>5.1</td>
<td>Simulation Results For CIR Interest Rate Model</td>
<td>53</td>
</tr>
<tr>
<td>5.2</td>
<td>Simulation Results For Heston Model</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>CONCLUSION</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>BIOGRAPHICAL SKETCH</td>
<td>78</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-1</td>
<td>Parameter table for CIR interest model simulation.</td>
<td>54</td>
</tr>
<tr>
<td>5-2</td>
<td>Estimated expectation of interest rate at time $T = 1$ in test Case I. Numbers in parentheses are sample standard deviations.</td>
<td>54</td>
</tr>
<tr>
<td>5-3</td>
<td>Estimated expectation of interest rate at time $T = 1$ in test Case II. Numbers in parentheses are sample standard deviations.</td>
<td>55</td>
</tr>
<tr>
<td>5-4</td>
<td>Estimated expectation of interest rate at time $T = 1$ in test Case III. Numbers in parentheses are sample standard deviations.</td>
<td>55</td>
</tr>
<tr>
<td>5-5</td>
<td>Estimated bond price in test Case I. Numbers in parentheses are sample standard deviations.</td>
<td>58</td>
</tr>
<tr>
<td>5-6</td>
<td>Estimated bond price in test Case II. Numbers in parentheses are sample standard deviations.</td>
<td>58</td>
</tr>
<tr>
<td>5-7</td>
<td>Estimated bond price in test Case III. Numbers in parentheses are sample standard deviations.</td>
<td>58</td>
</tr>
<tr>
<td>5-8</td>
<td>Parameter table for Heston model simulation. In all cases $r = 0$, $V(0) = \theta$ and $X(0) = 100$.</td>
<td>62</td>
</tr>
<tr>
<td>5-9</td>
<td>Estimated European call option price in test Case I. Numbers in parentheses are sample standard deviations.</td>
<td>63</td>
</tr>
<tr>
<td>5-10</td>
<td>Estimated European call option price in test Case II. Numbers in parentheses are sample standard deviations.</td>
<td>66</td>
</tr>
<tr>
<td>5-11</td>
<td>Estimated European call option price in test Case III. Numbers in parentheses are sample standard deviations.</td>
<td>69</td>
</tr>
<tr>
<td>5-12</td>
<td>Estimated discrete Asian call option price in test Case III. Numbers in parentheses are sample standard deviations.</td>
<td>72</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Normal yield curve</td>
<td>13</td>
</tr>
<tr>
<td>1-2</td>
<td>Inverted yield curve</td>
<td>13</td>
</tr>
<tr>
<td>1-3</td>
<td>Three sample paths of different OU-processes with $\theta = 1, \mu = 1.2, \sigma = 0.3$</td>
<td>16</td>
</tr>
<tr>
<td>1-4</td>
<td>An example of CIR process with $\alpha = 0.01, b = 1$ and $\sigma = 0.1$</td>
<td>18</td>
</tr>
<tr>
<td>4-1</td>
<td>Implied Volatility Surface</td>
<td>44</td>
</tr>
<tr>
<td>5-1</td>
<td>Comparison between different schemes in estimating $E[R(1)]$ in test Case I</td>
<td>56</td>
</tr>
<tr>
<td>5-2</td>
<td>Comparison between different schemes in estimating $E[R(1)]$ in test Case II</td>
<td>56</td>
</tr>
<tr>
<td>5-3</td>
<td>Comparison between different schemes in estimating $E[R(1)]$ in test Case III</td>
<td>57</td>
</tr>
<tr>
<td>5-4</td>
<td>Comparison between different schemes in estimating $E(e^{-\int_0^t R(s)ds})$ in test Case I</td>
<td>59</td>
</tr>
<tr>
<td>5-5</td>
<td>Comparison between different schemes in estimating $E(e^{-\int_0^t R(s)ds})$ in test Case II</td>
<td>60</td>
</tr>
<tr>
<td>5-6</td>
<td>Comparison between different schemes in estimating $E(e^{-\int_0^t R(s)ds})$ in test Case III</td>
<td>60</td>
</tr>
<tr>
<td>5-7</td>
<td>Comparison between different schemes in estimating European option price in test Case I (K=70)</td>
<td>64</td>
</tr>
<tr>
<td>5-8</td>
<td>Comparison between different schemes in estimating European option price in test Case I (K=100)</td>
<td>64</td>
</tr>
<tr>
<td>5-9</td>
<td>Comparison between different schemes in estimating European option price in test Case I (K=140)</td>
<td>65</td>
</tr>
<tr>
<td>5-10</td>
<td>Comparison between different schemes in estimating European option price in test Case II (K=70)</td>
<td>67</td>
</tr>
<tr>
<td>5-11</td>
<td>Comparison between different schemes in estimating European option price in test Case II (K=100)</td>
<td>67</td>
</tr>
<tr>
<td>5-12</td>
<td>Comparison between different schemes in estimating European option price in test Case II (K=140)</td>
<td>68</td>
</tr>
<tr>
<td>5-13</td>
<td>Comparison between different schemes in estimating European option price in test Case III (K=70)</td>
<td>70</td>
</tr>
<tr>
<td>5-14</td>
<td>Comparison between different schemes in estimating European option price in test Case III (K=100)</td>
<td>70</td>
</tr>
</tbody>
</table>
5-15 Comparison between different schemes in estimating European option price in test Case III (K=140)  ................................................................. 71

5-16 Comparison between different schemes in estimating discrete Asian option price in test Case III (K=70)  ................................................................. 72

5-17 Comparison between different schemes in estimating discrete Asian option price in test Case III (K=100)  ................................................................. 73

5-18 Comparison between different schemes in estimating discrete Asian option price in test Case III (K=140)  ................................................................. 73
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

A FAST AND EXACT SIMULATION FOR CIR PROCESS

By

Anqi Shao

May 2012

Chair: Liqing Yan
Major: Mathematics

We present a fast and exact simulation method for the CIR process. Traditional simulation method relies on an algorithm to generate a non-central chi square random variable, which is quite slow when the degrees of freedom is less than $\frac{3}{1}$, and the non-centrality parameter is large. Thus it’s only applicable when people are just interested in simulating the process along a few time points, for example, European options prices. But for some exotic options which depends on the process on a lot of time points, for example Asian option prices, this method is very slow and inefficient. In this paper we analyze the algorithm to see its limitation, and propose a new algorithm which is much faster. This method enables fast and exact simulation of the CIR process on a large number of time grids. Numerical results on option pricing based on Heston model is rather encouraging, compared with other simulation schemes.
CHAPTER 1
INTRODUCTION TO CIR MODEL AND SIMULATION

This chapter briefly reviews basic theory about interest rates, the term structure of interest rates, and introduced the famous short term interest rate model: CIR model in finance, as well as how to simulate a CIR process in general. Two commonly used methods in the literature are mentioned.

1.1 The Term Structure Of Interest Rates

The term structure of interest rates, also known as the yield curve, measures the relationship among the yields on default-free securities that differ only in their term to maturity. The study of this functional relationship has long been of interest to economists. The term structure implies the market's anticipations of future events by offering a complete schedule of interest rates across time. Hence, an explanation of the term structure gives us a way to understand and extract this information. We can then use this information to make predictions about how changes in the underlying variables will affect the yield curve.

Interest rates and their dynamics provide probably the most computationally difficult part of the modern financial theory. The modern fixed income market includes not only bonds but all kinds of derivative securities sensitive to interest rates. Moreover interest rates are important in pricing all other market securities since they are used in time discounting. Interest rates are also important on the corporate level since most investment decisions are based on some expectations regarding alternative opportunities and the cost of capital-both depend on the interest rates.

The interest rate market is where the price of money is set-how much does it cost to have money tomorrow, money in a year, money in ten years? The price of money over a term depends not only on the length of the term, but also on the moment-to-moment random fluctuations of the interest rate market.
The most basic interest rate contract is an agreement to pay some money now in exchange for a promise of receiving a (usually) larger sum later. In general, the worth of such a contract will depend on factors other than just the time value of money, such as the credit worthiness of the promisor, etc. We are solely concerned with the time value of money for default-free borrowing.

The basic contract only requires two numbers to describe it—its length, or maturity, which records when we are to receive the later payment, and the ratio of the size of that payment to our initial payment. Suppose the maturity date is T, and we pay $P(0, T)$ initially to receive one dollar at time $T$. The promise of a dollar at time $T$ could be regarded as an asset, which will have some worth at any time $t$ before $T$. This asset is called a discount bond, and the price $P(0, T)$ is its price at time zero. But it can have a different price at any other time $t$ up to maturity $T$, call it, say, $P(t, T)$. This price $P(t, T)$, the value at time $t$ of receiving a dollar at time $T$, is a process in time.

Real markets do not have a single interest rate. Instead, they have bonds of different maturities, some paying coupons and others not paying coupons. From these bonds, yields to different maturities can be implied. In practice, from market data one can ultimately determine prices of zero-coupon bonds for a number of different maturity dates. Each of these bonds has a yield specific to its maturity, where yield is defined to be the constant continuously compounding interest rate over the lifetime of the bond that is consistent with its price. In our example, given a discount bond price $P(t, T)$ at time $t$, the yield $R(t, T)$ is given by:

$$R(t, T) = -\frac{\log P(t, T)}{T - t}.$$  

or equivalently,

$$P(t, T) = e^{-R(t, T)(T - t)}.$$  

In our simple example, we assume a zero-coupon bond with face value equal to 1. The formula above implies that capital equal to the price of the bond, invested at a
continuously compounded interest rate equal to the yield, would, over the lifetime of the bond, result in a final payment of the face value. In real life, instead of having a single interest rate, real markets have a *yield curve*, which one can regard either as a function of finitely many yields plotted versus their corresponding maturities or more often as a function of time obtained by interpolation from the finitely many maturity-yield pairs provided by the market.

The difference in yields at different maturities reflects market beliefs about future interest rates. If there is a possibility that rates might be higher in the future, long-term loans will have to charge a higher rate than short-term ones. Typically, the normal yield curve would increase with maturity. And just as its name indicates, this is the yield curve shape that forms during normal market conditions, wherein investors generally believe that there will be no significant changes in the economy, such as in inflation rates, and that the economy will continue to grow at a normal rate. During such conditions, investors expect higher yields for fixed income instruments with long-term maturities that occur farther into the future. In other words, the market expects long-term fixed income securities to offer higher yields than short-term fixed income securities. This is a normal expectation of the market because short-term instruments generally hold less risk than long-term instruments; the farther into the future the bond's maturity, the more time and, therefore, uncertainty the bondholder faces before being paid back the principal. To invest in one instrument for a longer period of time, an investor needs to be compensated for undertaking the additional risk.

But if current rates are high and expected to fall, the yield curve can become inverted and long bond yields will be less than short bonds. These yield curves are rare, and they form during extraordinary market conditions wherein the expectations of investors are completely the inverse of those demonstrated by the normal yield curve. In such abnormal market environments, bonds with maturity dates further into the future are expected to offer lower yields than bonds with shorter maturities. The inverted yield
curve indicates that the market currently expects interest rates to decline as time moves farther into the future, which in turn means the market expects yields of long-term bonds to decline.

You may be wondering why investors would choose to purchase long-term fixed-income investments when there is an inverted yield curve, which indicates that investors expect to receive less compensation for taking on more risk. Some investors, however, interpret an inverted curve as an indication that the economy will soon experience a slowdown, which causes future interest rates to give even lower yields. Before a slowdown, it is better to lock money into long-term investments at present prevailing yields, because future yields will be even lower.

Figure 1-1. Normal yield curve

Figure 1-2. Inverted yield curve

But the yield curve just give us an idea of the rate of borrowing for each term length. It would be convenient if we could get the current, or instantaneous cost of borrowing in a single number. What we can do is look at the current rate for instantaneous borrowing. That is, borrowing which is paid back (nearly) instantly. Suppose at time $t$ we borrow over the period from $t$ to $t + \Delta t$, where $\Delta t$ is a small time increment, the rate we get is the yield $R(t, t + \Delta t)$:

$$R(t, t + \Delta t) = -\frac{\log P(t, t + \Delta t)}{\Delta t}$$
As $\Delta t \to 0$, we call the limit the *instantaneous rate*, or *short rate*, $r_t$, which is given by both the expressions

$$r_t = R(t, t),$$

and

$$r_t = -\frac{\partial}{\partial T} \log P(t, t).$$

The short rate is not only an important process in the interest rate market, but many models are based exclusively on its behavior, with all the other bonds extrapolated from it.

### 1.2 Models Of The Short-Term Interest Rate

The theory of interest-rate modeling was originally based on the assumption of specific one-dimensional dynamics for the instantaneous rate process $r_t$. Modeling directly such dynamics is very convenient since all fundamental quantities (rates and bonds) are readily defined as the expectation of a functional of the process $r_t$.

The short term riskless interest rate is one of the most fundamental and important prices determined in financial markets. More models have been put forward to explain its behavior than for any other issue in finance. Many of the more popular models currently used by academic researchers and practitioners have been developed in a continuous-time setting, which provides a rich framework for specifying the dynamic behavior of the short-term riskless rate. A partial listing of these interest rate models includes those by Merton (1973) [38], Brennan and Schwartz (1982) [10], Vasicek (1977) [42], Dothan (1978) [21], Cox, Ingersoll, and Ross (1985) [17], Hull and White (1990) [29], Black, Fischer and Piotr Karasinski (1991) [8].

One of the earliest papers to tackle arbitrage-free pricing of bonds and interest-rate derivatives was Vasicek (1977) [42]. This paper is best known for the Vasicek model for the risk-free rate of interest, $R(t)$ described below. However, Vasicek also developed a more general approach to pricing which ties in with what we now refer to as the
risk-neutral-pricing approach. The model specifies that the instantaneous interest rate follows the stochastic differential equation:

$$dR(t) = \alpha(b - R(t))dt + \sigma dW(t), \quad R(0) = r_0.$$  

This dynamics has some interesting properties that make the model attractive. Mathematically, the equation is linear and can be solved explicitly, the distribution of the short rate is Gaussian, and both the expressions and the distributions of several useful quantities related to the interest-rate world are easily obtainable. We can solve the SDE to obtain:

$$R(t) = R(0)e^{-\alpha t} + b(1 - e^{-\alpha t}) + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW(s).$$

Vasicek’s model was the first one to capture mean reversion, an essential characteristic of the interest rate that sets it apart from other financial prices. As opposed to stock prices for instance, interest rates cannot rise indefinitely. This is because at very high levels they would hamper economic activity, prompting a decrease in interest rates. Similarly, interest rates can not decrease below 0. As a result, interest rates move in a limited range, showing a tendency to revert to a long run value. In particular, Vasicek’s model exhibits mean reversion, which means that if the interest rate is above the long run mean \(r > b\), then the coefficient \(\alpha\) makes the drift become negative so that the rate will be pulled down in the direction of \(r\). Likewise, if the rate is less than the long run mean \(r < b\), then the coefficient \(\alpha\) makes the drift become positive so that the rate will be pulled up in the direction of \(r\). The coefficient \(\alpha\) is, thus, the speed of adjustment of the interest rate towards its long run normal level. This feature is particularly attractive because without it, interest rates could drift permanently upward the way stock prices do and this is simply not observed in practice. This particular type of stochastic process is referred to as an Ornstein-Uhlenbeck process.
The main disadvantage is that, under Vasicek’s model, it is theoretically possible for the interest rate to become negative, an undesirable feature. Since from the distribution of $R(t)$, it’s easy to see that $R(t)$ is normally distributed with mean and variance given by:

$$E(R(t)) = R(0) e^{-\alpha t} + b(1 - e^{-\alpha t})$$

$$Var(R(t)) = \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha t})$$

This implies that, for each time $t$, the rate $R(t)$ can be negative with positive probability. The general equilibrium approach developed by Cox, Ingersoll and Ross (1985) [17] led to the introduction of a "square-root" term in the diffusion coefficient of the instantaneous short rate dynamics. The resulting model has been a benchmark for many years because of its analytical tractability and the fact that, contrary to the Vasicek (1977) [42] model, the instantaneous short rate is always positive. Hull and White (1990) [29] proposed an even more general model by considering a time-varying parameter in the Vasicek model. Later, Black and Karasinski (1991) [7] assumed that the logarithm $\ln(R(t))$ of the instantaneous short rate evolves according to a generalized Vasicek
model with time-dependent coefficients. In the next section, we’ll focus on the celebrated CIR short rate model.

1.3 Introduction to CIR Model

The Cox-Ingersoll-Ross (CIR) [17] model is a diffusion process suitable for modeling the term structure of interest rates. It was introduced in 1985 by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross as an extension of the Vasicek model. The simplest version of this model describes the dynamics of the interest rate $X(t)$ as a solution of the following stochastic differential equation (SDE):

$$dX(t) = \alpha(b - X(t))dt + \sigma\sqrt{X(t)}dW(t), \quad X(0) = x_0 \geq 0 \quad (1-1)$$

for $\alpha > 0$, $b > 0$, $\sigma > 0$ and a standard Brownian motion $W$. This process has some appealing properties from an applied point of view, for example, the interest rate stays non-negative, and is elastically pulled towards the long-term constant value $b$ at a speed controlled by $\alpha$ (mean-reverting). Those properties are attractive in modeling real-life interest rates. In particular, the condition

$$2\alpha b \geq \sigma^2$$

would ensure that the origin is inaccessible to the process, so that we can grant that $X(t)$ remains positive. Intuitively, when the rate is at a low level (close to zero), the standard deviation $\sigma\sqrt{X(t)}$ also becomes close to zero, which dampens the effect of the random shock on the rate. Consequently, when the rate gets close to zero, its evolution becomes dominated by the drift factor, which pushes the rate upwards (towards equilibrium).

The interest rate behavior implied by this structure thus has the following empirically relevant properties: (i) Negative interest rates are precluded. (ii) If the interest rate reaches zero, it can subsequently become positive. (iii) The absolute variance of the
interest rate increases when the interest rate itself increases. (iv) There is a steady state distribution for the interest rate.

Figure 1-4. An example of CIR process with $\alpha = 0.01$, $b = 1$ and $\sigma = 0.1$

The SDE (1-1) is not explicitly solvable, hence the tractability of the CIR model is not as good as the Vasicek model in this regard. Nevertheless, the transition density for the process is known. Based on results of Feller [25], Cox et al. [17] noted that the distribution of $X(t)$ given $X(u)$ for some $u < t$ is, up to a scale factor, a non-central chi-square distribution. In fact,

$$X(t)|X(u) \sim c \chi^2_d(\lambda)$$

where

$$c = \frac{\sigma^2 (1 - e^{-\alpha(t-u)})}{4\alpha}, \quad d = \frac{4b\alpha}{\sigma^2}, \quad \lambda = \frac{4\alpha e^{-\alpha(t-u)}}{\sigma^2(1 - e^{-\alpha(t-u)})} X(u).$$

Straightforward calculations give the expected value and variance of $X(t)$ as:

$$E(X(t)|X(u)) = X(u)e^{-\alpha(t-u)} + b(1 - e^{-\alpha(t-u)}),$$

$$V_{ar}(X(t)|X(u)) = X(u)\left(\frac{\sigma^2}{\alpha}(e^{-\alpha(t-u)} - e^{-2\alpha(t-u)}) + b\left(\frac{\sigma^2}{2\alpha}\right)(1 - e^{-\alpha(t-u)})^2.\right.$$
Although CIR model is mainly used in finance in modeling interest rates, it should be noted that this process has other financial applications. For example, the stochastic volatility of the stock price (Heston) \([27]\) and the credit spread (Brigo and Alfonsi) \([11]\).

1.4 Simulation Issue

Financial models usually specify the dynamics of the state variables, e.g., stock price, volatility and interest rate, as stochastic differential equations (SDE). If these SDEs yield closed form solutions, then Monte Carlo simulation can be used to generate an unbiased estimator of the price of a derivative security. Basically we generate many sample paths of the state variable and compute the payoff of the derivative for each path. Discounting and averaging over all paths gives an estimator of the derivative price. The error in the Monte Carlo estimator can be calculated using the central limit theorem and converges to zero as the number of sample paths used increases.

However, in most cases, the SDEs that define the dynamics of the state variables do not yield closed form solutions. In this case, it’s still possible to use Monte Carlo simulation to compute derivative prices. But first we have to discretize the time interval and simulating the state process dynamics on this discrete time grid. However, the approximation of continuous time processes by discrete time processes introduces bias into the simulation estimator.

One drawback of the CIR process is that the SDE (1-1) is not explicitly solvable. In practical usage of such models (e.g. to price options) we are often faced with the problem of simulating a CIR process. In general there are two ways to do it, namely, exact simulation methods and approximation schemes. There are pros and cons associated with each method. The drawback of exact simulation methods is the computation time that they require. Exact simulation in general requires more time than a simulation with approximation schemes (Up to a factor 10). Hence it should be used to compute expectations that depend on the values of the process at just a few fixed times. On the contrary, for expectations that depends on all the path (such
as integrals) discretization schemes should be preferred. On the other hand, the
drawback of approximation schemes in general is the bias they introduced into the
estimator. Since the magnitude of the bias is unknown, it’s difficult to obtain valid
confidence intervals. And another serious problem with discretization schemes of CIR
process, in general any square-root diffusion process, is the square-root itself, which has
unbounded derivatives near zero. Therefore, discretization schemes that (explicitly or
implicitly) involve the derivatives of the coefficients—even when they assure the positivity
of approximation—usually lose their accuracy near zero, especially, for large $\sigma$. The larger
$\sigma$ is, the more concentrated near zero the value distributions of CIR process are. There
are ways to get around this problem, we’ll talk about them in detail in Chapter 2.

The first method is the exact simulation method, which is based on the transition
probability density function of the CIR process. In Cox’s paper [17], it was noted that the
distribution of $X(t)$ given $X(u)$ for any $0 < u < t$ is a noncentral chi-squared distribution.
The transition law of $X(t)$ can be expressed as:

$$X(t)|X(u) \sim c \chi^2_d(\lambda)$$

(1–2)

where

$$c = \frac{\sigma^2 (1 - e^{-\alpha(t-u)})}{4\alpha}, \quad d = \frac{4b\alpha}{\sigma^2}, \quad \lambda = \frac{4\alpha e^{-\alpha(t-u)}}{\sigma^2 (1 - e^{-\alpha(t-u)})} X(u).$$

(1–3)

This says that, given $X(u)$, $X(t)$ is distributed as $\sigma^2 (1 - e^{-\alpha(t-u)})/(4\alpha)$ times a noncentral
chi-square random variable with $d$ degrees of freedom and non-centrality parameter $\lambda$.

Thus, we can sample from the distribution of $X(t)$ exactly, provided that we can sample from the noncentral chi-square distribution. However, the drawback of the
traditional method in drawing non-central chi-square distribution is that it’s slow,
especially in the case when $d < 1$ and $\lambda$ is relatively big. The exact reason will be analyzed in chapter 3, but when the realization of the process along a lot of discrete
time is desired, the old exact simulation method is not efficient. It’s only competitive
when one has to simulate the process just at one time (or few times), for example to compute European options prices with a Monte-Carlo algorithm. On the contrary, they are drastically too slow if one has to simulate along a time-grid, as it is the case to calculate path-dependent options prices.

The second one is the popular Euler discretization. Euler scheme can always be used to approximate the paths of the interest rate process on a discrete time grid. And it’s in general faster than the exact simulation, which will be mentioned later. The drawback is that it’s an approximation of continuous time process by discrete time process, hence introduces bias into the simulation estimator. And in practice, many time steps may be necessary to reduce the bias to an acceptable level. Another problem is that when discretizing a CIR process, or in general a square-root diffusion process, a simple Euler scheme may not be well-defined because it can lead to negative values for which the square root is not defined. Consider the following most straightforward Euler scheme on the time interval $[0, T]$ for a CIR process $X(t)$:

$$
\hat{X}(t_{i+1}) = \hat{X}(t_i) + \alpha (b - \hat{X}(t_i))[t_{i+1} - t_i] + \sigma \sqrt{\hat{X}(t_i)}(W(t_{i+1}) - W(t_i))
$$

with $\hat{X}(t_0) = x_0 > 0, \alpha > 0, b > 0, \sigma > 0$ can lead to negative values since the Gaussian increment is not bounded from below. Thus, this simple scheme is not well defined. To correct this problem, Deelstra and Delbaen [19] have proposed the "full truncation scheme":

$$
\hat{X}(t_{i+1}) = \hat{X}(t_i) + \alpha (b - \hat{X}(t_i)^+)[t_{i+1} - t_i] + \sigma \sqrt{\hat{X}(t_i)^+}(W(t_{i+1}) - W(t_i))
$$

while Diop proposed in [6] proposed the "reflection scheme":

$$
\hat{X}(t_{i+1}) = |\hat{X}(t_i) + \alpha (b - \hat{X}(t_i))[t_{i+1} - t_i] + \sigma \sqrt{\hat{X}(t_i)}(W(t_{i+1}) - W(t_i))|$$
Also, Alfonsi [1] [2] [3] presented several implicit schemes and higher order schemes. In the next chapter, we'll briefly review some basic properties of Euler scheme, and the problems with simulating an CIR process with Euler schemes.
CHAPTER 2
EULER EISCRETIZATION AND RELATED ISSUES

This section presents the simplest approximation, Euler scheme, to a continuous-time stochastic process. This method is always easy to implement and almost applicable everywhere, but it's not always accurate enough to meet practical needs. Most of the following introduction comes from Glasserman's book [26].

2.1 Introduction To Euler-Maruyama Scheme

The Euler-Maruyama Scheme is a method for the approximate numerical solution of a stochastic differential equation (SDE). It is a simple generalization of the Euler method for ordinary differential equations to stochastic differential equations. We consider a process \( X \) satisfying a stochastic differential equation of the form

\[
dX(t) = a(X(t))dt + b(X(t))dW(t)
\]  

(2–1)

with initial condition \( X(0) = x_0 \), some fixed real number. We won’t get into details about the uniqueness and existence conditions of the SDE, but essentially the coefficient functions \( a \) and \( b \) are assumed to satisfy some technical conditions. Given the functions \( a \) and \( b \), the stochastic process \( X(t) \) is a solution of the SDE (2-1) if \( X(t) \) solves the integral equation

\[
X(t) - X(0) = \int_0^t a(X(s))ds + \int_0^t b(X(s))dW(s)
\]

A famous example is the so called geometric Brownian motion, which is used to model the evolution of asset prices. It’s a process satisfying the following SDE:

\[
dS(t) = \mu S(t)dt + \sigma S(t)dW(t)
\]

We can actually solve this equation and get \( S(t) = S(0) \exp\{\sigma W(t) + (\mu - \frac{1}{2}\sigma^2)t\} \).
In practice, however, many SDE’s are not explicitly solvable like the above one. Hence we can’t always get an analytical solution \( X(t) \) to a given SDE. But we could always get approximate numerical solution of the given SDE. The basic idea is essentially from Euler Scheme for ordinary differential equation (ODE). In particular, let’s use \( \hat{X}(t) \) to denote a time-discretized approximation to \( X(t) \). Suppose we discretize the interval \([0, T]\): let \( \Delta t = T/N \) and \( t_n = n\Delta t, n = 0, 1, 2, ..., N \). The exact solution on the time grid would be

\[
X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_{n+1}} a(X(s))ds + \int_{t_n}^{t_{n+1}} b(X(s))dW(s)
\]

The Euler-Maruyama approximation on the time grid \( 0 = t_0 < t_1 < ... < t_N = T \) is defined by

\[
\hat{X}(t_{i+1}) = \hat{X}(t_i) + a(\hat{X}(t_i))[t_{i+1} - t_i] + b(\hat{X}(t_i))\sqrt{t_{i+1} - t_i}Z_{i+1},
\]

with \( Z_1, Z_2, ... \) independent, m-dimensional standard normal random vectors. It is a simple generalization of the Euler method for ordinary differential equations to stochastic differential equations. It is named after Leonhard Euler and Gisiro Maruyama. Implementation of this method is straightforward, when the functions \( a \) and \( b \) are easy to evaluate. It’s easy to see that as the time grid gets finer, the Euler-Maruyama approximation \( \hat{X}(t) \) would converge to the real solution \( X(t) \). In the next section we’ll talk about the convergence of Euler-Maruyama Scheme.

### 2.2 Convergence Order

Two broad categories of error of approximation are commonly used in measuring the quality of discretization methods: criteria based on the path-wise proximity of a discretized process to a continuous process, and criteria based on the proximity of the corresponding distributions. These are generally termed strong and weak criterion, respectively.
**Definition** (Strong error criterion) Given a sequence of discrete time approximation \( \{ \hat{X}(0), \hat{X}(h), \hat{X}(2h), \ldots \} \) to a continuous time process \( X \), where \( h = T/n \) for a fixed time \( T \). We say \( \hat{X} \) converges strongly to \( X \) if it converges in \( L^1 \), i.e., if

\[
\lim_{n \to \infty} E(|\hat{X}(nh) - X(T)|) = 0
\]

We say that a discretization \( \hat{X} \) has strong order of convergence \( \gamma > 0 \) if

\[
E\|\hat{X}(nh) - X(T)\| \leq ch^\gamma
\]

for some vector norm \( \| . \| \), some constant \( c \) and for \( h \) sufficiently small.

The strong error criterion measures the deviation between the individual values of \( X \) and the approximation \( \hat{X} \). In contrast, a weak error criterion looks like the following:

**Definition** (Weak error criterion) Given a sequence of discrete time approximation \( \{ \hat{X}(0), \hat{X}(h), \hat{X}(2h), \ldots \} \) to a continuous time process \( X \), where \( h = T/n \) for a fixed time \( T \). We say \( \hat{X} \) converges weakly to \( X \) if

\[
\lim_{n \to \infty} E(f(\hat{X}(nh))) = E(f(X(T)))
\]

We say that a discretization \( \hat{X} \) has weak order of convergence \( \gamma > 0 \) if

\[
|E(f(\hat{X}(nh))) - E(f(X(T))))| \leq ch^\gamma
\]

for some constant \( c \) and for \( h \) sufficiently small, for all \( f \) whose derivatives of order \( 0, 1, \ldots, 2\gamma + 2 \) are polynomially bounded. (A function \( g \) is polynomially bounded if \( |g(x)| \leq k(1 + |x|^q) \) for some constants \( k \) and \( q \) and all \( x \in \mathbb{R} \))

For applications in derivatives pricing, we really only care about the weak error criteria. We would like to ensure that option prices (which are expectations) computed from \( \hat{X}(t) \) are close to prices computed from \( X(t) \). We are not really concerned about the specific paths of the two processes.
Under modest conditions, even the simple Euler scheme converges as the time step $h$ decreases to zero. We compare different discretization schemes based on the rate at which they converge. A large value of $\gamma$ implies faster convergence to zero of the discretization error. It's often the case that for the same scheme, the strong order of convergence is smaller than the weak order of convergence. For example, the Euler scheme typically have a strong order of $1/2$, but it often achieves a weak order of $1$. Stronger conditions are required for the Euler scheme to have weak order $1$. In [34] the authors require the functions $a$ and $b$ be four times continuously differentiable with polynomially bounded derivatives. But good accuracy on smooth functions may not be directly relevant to our intended applications: the CIR process has a square root function, which is not Lipschitzian. This posed a problem for the ordinary Euler scheme. We'll discuss about this problem in the next section.

2.3 Euler Scheme For CIR Process

Using an Euler discretization to simulate CIR process gives rise to the problem that while the process itself is guaranteed to be nonnegative, the discretization is not. The main difficulty when discretizing the CIR process is located at $0$, where the square-root is not Lipschitzian. General schemes, such as the Euler scheme or the Milstein scheme are in general not well defined because they can lead to negative values for which the square root is not defined.

In Istvan and Miklos (2011) [30] the authors established the result of a convergence speed estimate for Euler schemes corresponding to SDEs with $1/2$-Holder continuous diffusion coefficients. This result could be applied to CIR process, since the diffusion coefficient of CIR process fails to be Lipschitz continuous near the origin. Hence classical results on the rate of strong convergence for the corresponding Euler scheme do not apply in this case. In general, they showed that the convergence rate for Euler schemes corresponding to SDEs with $1/2$-Holder continuous diffusion coefficients
(CIR as a special case) without any restrictions on the parameters should be $1/\ln n$. In particular, fix $T > 0$ and consider the SDE

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \quad X(0) = \xi$$

on the interval $[0, T]$, where $W(t)$ is a standard Brownian motion, $\xi$ is independent of $W(t)$, and the coefficients satisfy the following conditions:

**Assumption:** $\sigma, f, g : [0, T] \times R \to R$ are measurable; $g(t, \cdot)$ is monotone decreasing; $b = f + g$ and there exist $K > 0, \alpha \in [0, 1/2]$ and $\gamma \in (0, 1]$ such that for all $t \in [0, T]$ and $x, y \in \mathbb{R}$,

$$|\sigma(t, x) - \sigma(t, y)| \leq K|x - y|^{1+\alpha}, \quad |f(t, x) - f(t, y)| \leq K|x - y|,$$

$$|g(t, x) - g(t, y)| \leq K|x - y|^\gamma,$$

and

$$|b(t, 0)| + |\sigma(t, 0)| \leq K.$$

Under the above assumption, the following theorem gives the convergence order of Euler Scheme:

**Theorem:** Let the assumption above hold and let $E|\xi|^{1+2\alpha} < \infty$. Then there is a constant $C$ depending only on $K, T, \gamma$ and $E|\xi|^{1+2\alpha}$ such that

$$E|X(\tau) - X_n(\tau)| \leq \frac{C}{\ln n}, \quad \text{if } \alpha = 0,$$

$$E|X(\tau) - X_n(\tau)| \leq C \left( \frac{1}{n^\alpha} + \frac{1}{n^{\gamma/2}} \right), \quad \text{if } \alpha \in (0, \frac{1}{2}],$$

for all $n \geq 2$ and for every stopping time $\tau \leq T$.

Hence for CIR process, which is the case when $\alpha = 0$, without any restriction on the parameters, a slow rate of $O(1/\ln n)$ is established.
Also in Alfonsi’s paper [1] the author proposed several implicit discretization schemes for the CIR process. But for large values of \( \sigma \), for example, when \( \sigma^2 \gg 4\alpha b \), none of the scheme seems to be efficient due to large discretization bias. In the next chapter, we’ll look at a new method to simulate an CIR process exactly. The method is exact, hence there is no discretization errors.
In this section, we introduce the traditional exact simulation method of a CIR process, and show why the so called "Poisson method" is not efficient in some cases. And we find a new method to simulate a non-central chi-squared random variable, which is much more efficient than the Poisson method.

3.1 Traditional Simulation Method

The CIR process is a Markov process with continuous paths defined by the following stochastic differential equation (SDE):

\[ dX(t) = \alpha(b - X(t))dt + \sigma \sqrt{X(t)}dW(t), \quad X(0) = x_0 \geq 0 \]

for \( \alpha > 0, b > 0, \sigma > 0 \) and a standard Brownian motion \( W \). As already mentioned in chapter 1, the CIR process has some appealing properties and was originally used to model interest rate in finance. It’s well known that there is a way to get the exact distribution of the CIR process \( X(t) \), given \( X(u) \) for \( u < t \). In fact, the transition law of \( X(t) \) is a non-central chi-square distribution:

\[ X(t)|X(u) \sim c \chi_d^2(\lambda) \]

with degrees of freedom \( d = \frac{4b\alpha}{\sigma^2} \), \( c = \frac{\sigma^2(1 - e^{-\alpha(t-u)})}{4\alpha} \) and non-centrality parameter

\[ \lambda = \frac{4\alpha e^{-\alpha(t-u)}}{\sigma^2(1 - e^{-\alpha(t-u)})}X(u). \]

Hence, if we can sample from the non-central chi-squared distribution, we can sample from the distribution of \( X(t) \) exactly.

Let’s first introduce some basic facts about central chi-square and non-central chi-square distributions. If \( d \) is a positive integer and \( Z_1, \ldots, Z_d \) are independent \( N(0, 1) \) random variables, then the distribution of \( Z_1^2 + Z_2^2 + \ldots + Z_d^2 \) is called the (central) chi-square distribution with \( d \) degrees of freedom. The central chi-square random variable is just a special case of a non-central chi-square random variable, with centrality.
The symbol $\chi^2_d(0)$ denotes a random variable with this distribution; the prime in $\chi'_d(\lambda)$ emphasizes that this symbol refers to the noncentral case. The cumulative density function (CDF) of a central chi-square distribution is given by

$$P(\chi^2_d(0) \leq x) = \frac{1}{\Gamma(d/2)} \int_0^x e^{-z/2} z^{(d/2)-1} dz,$$

where $\Gamma(.)$ denotes the gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt$ and $\Gamma(n) = (n-1)!$ if $n$ is a positive integer. This expression defines a valid probability distribution for all $d > 0$ and thus extends the definition of $\chi^2_d(0)$ to non-integer $d$.

For integer $d$ and constants $a_1, \ldots, a_d$, the distribution of

$$\sum_{i=1}^d (Z_i + a_i)^2$$

is noncentral chi-square with $d$ degrees of freedom and non-centrality parameter $\lambda = \sum_{i=1}^d a_i^2$. Let $\chi'_d(\lambda)$ denote a non-central chi-squared random variable with $d$ degrees of freedom and non-centrality parameter $\lambda$. The probability density function (pdf) of this random variable is

$$g(\lambda, d, x) = e^{-\frac{x}{\lambda}} \sum_{j=0}^\infty \frac{(\lambda/2)^j}{j! \Gamma(d/2 + j)} x^{\frac{d}{2} + j - 1}.$$  \hspace{1cm} (3–2)

and its characteristic function (c.f.) is

$$E \exp \{it\chi'_d(\lambda)\} = \left(1 - 2it\right)^{-\frac{d}{2}} \exp\left\{\frac{\lambda}{1 - i2t} - \frac{\lambda}{2}\right\}.$$  \hspace{1cm} (3–3)

This easily implies its additivity, namely

$$\chi'_d(\lambda_1) + \chi'_d(\lambda_2) \overset{d}{=} \chi'_d(\lambda_1 + \lambda_2).$$

for two independent noncentral chi-square random variables with degrees of freedoms $d_1$, $d_2$ and noncentrality parameters $\lambda_1$, $\lambda_2$, respectively. Therefore, in particular, we have

$$\chi'_d(\lambda) \overset{d}{=} \chi'_1(\lambda) + \chi'_d(0), \text{ for } d > 1$$  \hspace{1cm} (3–3)
Therefore, to generate $\chi'_{d}(\lambda), d > 1$, we can generate $\chi^2_{d-1}(0)$ and an independent standard normal random variable $Z$ and set

$$\chi'_{d}(\lambda) = (Z + \sqrt{\lambda})^2 + \chi^2_{d-1}(0).$$

(3–4)

Thus, sampling from a noncentral chi-squared distribution is reduced to sampling from an ordinary chi-squared and an independent normal when $d > 1$. And this method is in general quite efficient.

But for any $0 < d < 1$, we can not shift $\lambda$ to a normal random variable $Z$. In this case, a noncentral chi-squared random variable can only be generated as an ordinary chi-squared random variable with a random degrees of freedom parameter. In order to see this, suppose $N$ is a Poisson random variable with mean $\lambda/2$, then

$$P(N = j) = e^{-\lambda/2}\frac{(\lambda/2)^j}{j!}, \quad j = 0, 1, 2, \ldots$$

Consider now a random variable $\chi^2_{d+2N}(0)$ with $N$ having this Poisson distribution. Conditional on $N = j$, the random variable has an ordinary (central) chi-square distribution with $d + 2j$ degrees of freedom:

$$P(\chi^2_{d+2N}(0) \leq x|N = j) = \frac{1}{2^{d/2+j}\Gamma(d/2+j)} \int_0^x e^{-z/2}z^{(d/2)+j-1}dz.$$  

The unconditional distribution is thus given by

$$\sum_{j=0}^{\infty} P(N = j)P(\chi^2_{d+2N}(0) \leq x|N = j) = \sum_{j=0}^{\infty} e^{-\lambda/2}\frac{(\lambda/2)^j}{j!}P(\chi^2_{d+2j}(0) \leq y),$$

which is precisely the non-central chi-square distribution with degrees of freedom $d$ and non-centrality parameter $\lambda$.

Therefore, we have the following relationship:

$$\chi'_{d}(\lambda) \overset{d}{=} \chi^2_{d+2N}(0), \quad \text{for } d < 1.$$  

(3–5)
We may therefore sample $\chi^2_d(\lambda)$ by first generating a Poisson random variable $N$ with mean $\lambda/2$ and then, conditional on $N$, sampling a chi-square random variable with $d + 2N$ degrees of freedom. This method can be used to sample $\chi^2_d(\lambda)$ when $0 < d < 1$, and in general is reasonably fast. Theoretically, being able to generate a noncentral chi-square random variable, we could generate a discrete time sample path of a CIR process based on the conditional distribution $X(t)|X(u)$. In other words, given an initial $X(0)$ and any time grid $0 = t_0 < t_1 < \ldots < t_N = T$, we are able to generate a non-biased sample of the CIR process on the time grid.

But we’ll see in the next section that it can only be used if we just need to generate one non-central chi-square random variable, with $\lambda$ not very big. In real life, numerous derivatives have a path-dependent feature. In other words, their prices usually depend on the underlying asset not only at one single time point, but many of them. In this case, we would have to simulate the CIR process along a time grid with many time points. This is not a serious problem when CIR process is used to model interest rate, since $d$ would typically be larger than 2. But when applied to model variance process in a stochastic volatility model, like Heston’s model, it poses serious problem since $\sigma$ is often much bigger than in the interest rate model. In this case, $d = \frac{4b\sigma}{\sigma^2}$ would typically be less than 1, and in many cases close to 0. This is exactly the situation where the "Poisson method" has to be used to generate a CIR process. Although the above method yields sample with no bias, practitioners often have to settle for Euler discretization methods to price path-dependent derivatives under Heston’s model, since the above algorithm is too slow for real time simulation. We’ll analyze in detail the reason why the traditional method to simulate a CIR process is not efficient in this case to price path-dependent securities in the next section. More efficient algorithm to generate CIR processes is needed in this situation.
3.2 Some Drawbacks In Simulating CIR Process

Let’s use the algorithm above to simulate a CIR process. In particular, we only consider the case when $d < 1$, since that’s when we have to use the “Poisson method” for simulation. Suppose we want to simulate a CIR process $X(t)$ along $n$ equally spaced time points on the time interval $[0, T]$. In this case, given an initial value $X(0)$, we could use the transition distribution of $X(t)$ to simulate a random vector distributed according to the law of $(X(T/n), X(2T/n), ..., X(T))$ inductively. Let $t_0, t_1, t_2, ..., t_n$ to denote the time points $0, T/n, 2T/n, ..., T$. Starting from $X(0)$, which is given, we could generate the whole sample path of the CIR process along the time grid $t_i$ iteratively.

In general, given $X(t_{i-1})$, the conditional distribution of $X(t_i)$ is $c X_d(t_{i-1})$, where $c = \frac{\sigma^2(1 - e^{-a\Delta t})}{4\alpha}$, $\lambda(i) = \frac{4\alpha e^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})}X(t_{i-1})$. Here $\Delta t = t_i - t_{i-1} = \frac{T}{n}$. It’s easy to see that the value of $\lambda(i)$ depends on $X(t_{i-1})$ and $\Delta t$. This fact turns out to be the major problem in simulating a CIR process with $d < 1$ using the Poisson method along many time points.

Suppose we want to use Monte Carlo method to approximate a path-dependent option (for example, Asian option or Lookback option), assuming the underlying asset price process follows the Heston stochastic volatility model. In particular, the payoff function of a standard European-style discrete Asian call option is

$$\left(\frac{1}{n+1} \sum_{i=0}^{n} S_t - K\right)^+$$

whereas the payoff function of a fixed strike discrete Lookback call option is

$$\left(\max_{0 \leq i \leq n} S_t - K\right)^+$$

Here $S_t$ denotes the underlying asset price at time $t$, $K$ is the strike price. A typical Monte Carlo method involves simulating the asset price on the time grid $0 = t_0 < t_1 < ... < t_n = T$ for $M$ times, where $M$ is typically much bigger compared with the number of time points $n$. For each sample of the asset price process, compute the corresponding
payoff function value. Then the Monte Carlo estimation of the option price is given by the mean value of the discounted payoff values along those different paths generated.

Thus, in order to compute the price of a path-dependent option using Monte Carlo method, we are interested in simulating an asset price process along a lot of time points. In particular, Heston’s model use a CIR process to model the evolution of stochastic volatility. Hence in order to simulate the asset price process, we first need to simulate a CIR process along those time points. As mentioned above, the fact that the value of \( \lambda(i) \) depends on \( X(t_{i-1}) \) and \( \Delta t \) poses a problem in simulating the CIR process. The reason is that when the needed number of time points \( n \) is big, \( \Delta t = T/n \) is very small. In this case \( \lambda(i) = \frac{4\alpha e^{-\alpha \Delta t}}{\sigma^2(1 - e^{-\alpha \Delta t})}X(t_{i-1}) \) would be huge since it’s a decreasing function of \( \Delta t \), and actually goes to \( \infty \) as \( \Delta t \to 0 \). This poses a serious problem for the simulation.

Because when \( d < 1 \), which is common for Heston model, in order to generate a non-central chi-square random variable \( \chi_d^2(\lambda(i)) \), we need to first generate a Poisson random variable with mean \( \lambda(i)/2 \). Then conditional on the value of the Poisson random variable \( N \), generate a central chi-square random variable \( \chi^2_{d+2N}(0) \). Normally, in order to calculate the price of a path-dependent option, the number of time points we need is big. In this case the algorithm will spend a huge amount of time generating Poisson random variables at each time point, since \( \lambda(i)/2 \) is huge. Hence, although theoretically this algorithm could still be used to simulate a CIR process at multiple time points, it’s not efficient if the number of time points involved is big. In this case people prefer to use Euler Scheme to generate the sample path instead, although it often generate results which are biased and not accurate. Because the Poisson method is simply too slow.

In the next section, we’ll introduce a new method to generate a non-central chi-square random variable, which avoids this problem.
3.3 Main Result

The following theorem provides a new method to generate a non-central chi square random variable based on a central chi square random variable.

**Theorem**  A noncentral Chi-square r.v $\chi_d^2(\lambda)$ can be expressed by

$$\chi_d^2(\lambda) \overset{d}{=} \chi_d^2(0) + \mathcal{Y}(\lambda, Z, \tilde{Z}, U)$$  \hfill (3–6)

where $\chi_d^2(0)$ has a gamma distribution $G(d/2, 2)$, $U$ has a uniform distribution on $[0, 1]$, $Z$ and $\tilde{Z}$ are standard normal random variables. All four random variables are independent, and

$$\mathcal{Y}(\lambda, Z, \tilde{Z}, U) = \begin{cases} 
0, & \text{if } \lambda + 2\ln(U) \leq 0. \\
(Z + \sqrt{\lambda + 2\ln(U)})^2 + \tilde{Z}^2, & \text{if } \lambda + 2\ln(U) > 0.
\end{cases}$$  \hfill (3–7)

**Proof:**  The probability density function $g(\lambda, d, x)$ of $\chi_d^2(\lambda)$ is given by

$$g(\lambda, d, x) = e^{-\frac{x}{2}} \sum_{j=0}^{\infty} \left(\frac{\lambda}{2}\right)^j \frac{e^{-\frac{x}{2}}}{\Gamma\left(\frac{d}{2} + j\right)} \frac{x^{\frac{d}{2}+j-1}}{2^{\frac{d}{2}+j}}.$$  

from this density function, we have

$$e^{\frac{\lambda}{2}} g(\lambda, d, x) = g(0, d, x) + \sum_{j=1}^{\infty} \left(\frac{\lambda}{2}\right)^j \frac{e^{-\frac{x}{2}}}{\Gamma\left(\frac{d}{2} + j\right)} \frac{x^{\frac{d}{2}+j-1}}{2^{\frac{d}{2}+j}}.$$  

where $g(0, d, x) = \frac{e^{-\frac{x}{2}}}{\Gamma\left(\frac{d}{2}\right)} \frac{x^{\frac{d}{2}-1}}{2^{\frac{d}{2}}}$ is the density function of $\chi_d^2(0)$. Now, differentiating both sides w.r.t $\lambda$ yields

$$\frac{\partial}{\partial \lambda} \left( e^{\frac{\lambda}{2}} g(\lambda, d, x) \right) = \frac{1}{2} \sum_{j=1}^{\infty} \left(\frac{\lambda}{2}\right)^{j-1} \frac{1}{(j-1)!} \frac{e^{-\frac{x}{2}}}{\Gamma\left(\frac{d}{2} + j\right)} \frac{x^{\frac{d}{2}+j-1}}{2^{\frac{d}{2}+j}}.$$  

$$= \frac{1}{2} \sum_{j=0}^{\infty} \left(\frac{\lambda}{2}\right)^j \frac{1}{j!} g(0, d + 2j, x)$$  

$$= \frac{1}{2} e^{\frac{x}{2}} g(\lambda, d + 2, x).$$
Therefore, integrating both side, we get the following relationship:

\[ e^{\frac{1}{2} \lambda} g(\lambda, d, x) = g(0, d, x) + \frac{1}{2} \int_{0}^{\lambda} e^{\frac{1}{2} t} g(t, d + 2, x) dt. \]

Multiplying both sides by \( e^{-\frac{1}{2} \lambda} \), then letting \( s = e^{\frac{1}{2}(t-\lambda)} \), we have:

\[
\begin{align*}
    g(\lambda, d, x) &= e^{-\frac{1}{2} \lambda} g(0, d, x) + \frac{1}{2} e^{-\frac{1}{2} \lambda} \int_{0}^{\lambda} e^{\frac{1}{2} t} g(t, d + 2, x) dt \\
    &= e^{-\frac{1}{2} \lambda} g(0, d, x) + \int_{e^{-\frac{1}{2} \lambda}}^{1} g(\lambda + 2 \ln s, d + 2, x) ds
\end{align*}
\]

from which we claim that

\[
\chi_d^2(\lambda) \overset{d}{=} \begin{cases} 
    \chi_d^2(0), & \text{if } \lambda + 2 \ln(U) \leq 0. \\
    \chi_{d+2}^2(\lambda + 2 \ln U), & \text{if } \lambda + 2 \ln(U) > 0.
\end{cases}
\]

By additivity, \( \chi_{d+2}^2(\lambda + 2 \ln U) \overset{d}{=} \chi_d^2(0) + (Z + \sqrt{\lambda + 2 \ln U})^2 + \bar{Z}^2 \), the theorem follows.

Using this theorem, we could generate a non-central chi-square random variable using a central chi-square random variable, plus possibly two independent Gaussian random variables and a uniform random variable. The most important thing is by generating the random variable this way, we avoid the problem of simulating a Poisson random variable in a CIR process simulation. And it’s much more efficient when we’re interested in simulating the CIR process along a large number of time points, since the algorithm doesn’t depend on the number of time points \( n \) the way Poisson random variable does in the old algorithm. The old algorithm will be very slow with large value of \( \lambda \), whereas the speed of this algorithm is independent to the value of \( \lambda \). Hence it’s both bias free and quick. This algorithm is most suitable in simulating stochastic volatility in Heston model.
This chapter briefly discusses basic option pricing theory. In particular, Heston’s stochastic volatility (SV) model is introduced and how the exact simulation method can be used to price options.

4.1 Option Pricing Theory

Options have existed for a long time. It wasn’t until publication of the Black-Scholes (1973) \[9\] option pricing formula that a theoretically consistent framework for pricing options became available. That framework was a direct result of work by Robert Merton as well as Black and Scholes. Since then the derivatives market has grown into a multi-trillion dollar industry. Options have become important to industry, particularly as they can be used to hedge out risk. Option pricing in itself has become an important research area. Numerous option pricing models were proposed, each one trying to generalize the Black-Scholes framework by relaxing one or more of the unrealistic restrictions of the original Black-Scholes model.

An Option is a major financial derivative, it gives the holder of that option the right, not the obligation to trade a fixed amount of underling asset at an agreed-upon price on the maturity date (European option) or any time on or before the maturity date (American option). A \textit{call option} gives the holder the right, but not obligation to buy, a \textit{put option} gives the holder the right, but not obligation to sell. There are two sides of every option contract. On one side is the investor who has taken the long position(i.e., has bought the option). On the other side is the investor who has taken a short position (i.e., has sold or written the option). The most straight forward financial derivatives are European call option and European put option. The price in the contract is known as the \textit{strike price}; the date in the contract is known as the \textit{maturity}. The payoff from a long
position in a European call option is:

$$\max(S(T) - K, 0)$$

where $T$ is the maturity, $S(T)$ is the stock price at maturity, and $K$ is the strike price.

Hence, the buyer of a European call option hopes the price of the underlying stock would rise above $K$ at the maturity, in which case he would use the right to buy the stock at $K$ and make a profit. But if the stock price would drop below $K$ at maturity, he doesn’t have the obligation to exercise the option. Similarly, the payoff from a long position in a European put option is:

$$\max(K - S(T), 0)$$

Although it’s fairly easy to understand the payoff structure of different types of options, to price them is quite a different story. The question of how options should be priced had been the subject of long intellectual debates, starting from the early sixties.

Option pricing theory traces its root to Bachelier (1900) who used Brownian motion to model options on French government bonds. This work anticipated by five years Einstein’s independent use of Brownian motion in physics. Somehow, the research didn’t pick up until 1960’s. In Samuelson (1965) [39], he considered long-term equity options, and used geometric Brownian motion to model the random behavior of the underlying stock. Based upon this, he modeled the random value of the option at exercise. Unfortunately, Samuelson’s formula was largely arbitrary. It offered no means for a buyer and seller with different risk aversions to agree on a price for an option.

In the seminal paper by Black and Scholes (1973) [9], they derived a partial differential equation for valuing claims contingent on a traded underlying asset. The equation is general. By applying different boundary conditions, it can be solved to price any such contingent claim. Black and Scholes applied the boundary conditions for a European call option on a non-dividend-paying stock and obtained their famous option pricing formula. The key idea behind the derivation was to perfectly hedge the option by
buying and selling the underlying asset in just the right way and consequently "eliminate risk". This hedge is called delta hedging and is the basis of more complicated hedging strategies.

In the original Black-Scholes model, they assumed that the derivative's underlying price follows a standard model for geometric Brownian motion:

\[ dS(t) = \mu S(t)dt + \sigma S(t)dW_t \]

where \( \mu \) is a constant drift(expected return) of the security price \( S(t) \), \( \sigma \) is the constant volatility, and \( W_t \) is a standard Brownian motion. This stochastic differential equation(SDE) has the following explicit solution:

\[ S(t) = S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \]

Under those assumptions, they derived a partial differential equation, now called the Black-Scholes equation, which governs the price of the option over time. The key idea behind the derivation was to perfectly hedge the option by buying and selling the underlying asset in just the right way and consequently "eliminate risk". Suppose we construct the following portfolio:

\[ \Pi = V(S, T) - \Delta S \]

Here \( V(S, T) \) denotes a value of one long option position and \( \Delta \) is the amount short in the underlying. The value of the hedged portfolio will change by

\[ d\Pi = dV(S, T) - \Delta dS \]

Using Ito's formula, we have

\[ dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt \]
Plug in, the change in the portfolio will be:

$$d \Pi = (\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + (\frac{\partial V}{\partial S} - \Delta) dS$$

Hence, The risk in the portfolio $\Pi$ is removed if

$$\frac{\partial V}{\partial S} - \Delta = 0$$

Therefore, the quantity $\Delta$ is chosen as

$$\Delta = \frac{\partial V}{\partial S}$$

If the portfolio is delta hedged continuously (dynamic hedging strategy), then a risk-less portfolio is constructed with a dynamics given by

$$d \Pi = (\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt$$

Since there is no stochastic term, $\Pi$ is a risk-free investment and hence must offer the same return as any other risk-free investment. Therefore, from the no-arbitrage condition

$$d \Pi = r \Pi dt$$

It is important to note that the portfolio $\Pi$ represents a self-financing, replicating and hedging strategy. It replicates a risk-free investment and it is hedged since it has no stochastic component. Making substitutions to equation it is obtained that

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)dt = r(V - S \frac{\partial V}{\partial S})dt$$

Simplifying the above equation, we showed that the payoff of any contingency claim $V(s, t)$ must satisfy the following partial differential equation(Black-Scholes equation):

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} = rV$$
Coupled with the appropriate terminal and boundary conditions, they got closed-form solutions for European call and put options. The value of a call option for a non-dividend paying underlying stock in terms of the Black-Scholes parameters is:

\[
C(S, t) = N(d_1)S - N(d_2)Ke^{-r(T-t)}
\]
\[
d_1 = \frac{\log \frac{S}{K} + (r + \frac{\sigma^2}{2})(T - t)}{\sigma \sqrt{T - t}}
\]
\[
d_2 = \frac{\log \frac{S}{K} + (r - \frac{\sigma^2}{2})(T - t)}{\sigma \sqrt{T - t}} = d_1 - \sigma \sqrt{T - t}
\]

The price of a corresponding put option is:

\[
P(S, t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S
\]

where \(N(.)\) is the cumulative density function of the standard normal distribution, \(T - t\) is the time to maturity, \(S\) is the spot price of the underlying asset, \(K\) is the strike price, \(r\) is the risk free rate, and \(\sigma\) is the volatility of returns of the underlying asset.

Alternatively, there is another way to find the price of derivatives, namely risk neutral pricing. This is one of the most important principles in derivative valuation. Risk neutral pricing is a powerful method for computing prices of derivative securities. The basic idea of risk neutral pricing is that given a realistic market model for the stock price process, we can construct a "risk free" portfolio that consists only the underlying stock and risk free bond. If there exists an equivalent measure \(Q\) under which the discounted stock price process is a martingale (the risk-neutral measure), then based on Martingale Representation Theorem, we could perfectly replicate any contingency claim based on the stock. The real world probability measure \(P\), surprisingly, has nothing to do with option pricing. Finding this particular equivalent measure \(Q\) relies on Girsanov theorem. After we get this perfect replicating scheme, it’s easy to argue, based on no arbitrage pricing theory, that the only arbitrage-free price of the claim at time zero is the expected value of the discounted payoff of the claim under the risk-neutral
measure $Q$. In particular, for an European call option, the payoff at maturity $T$ is given by $(S(T) - K)^+$, where $K$ is the strike price. Hence, risk-neutral pricing tells us that the price of the European call at time 0 is:

$$E^Q[e^{-rT}(S(T) - K)^+]$$

Since Black and Scholes published their seminal article on option pricing in 1973, there has been vast explosions of theoretical and empirical investigation on option pricing. People start to realize that the Black-Scholes model disagrees with reality in a number of ways, some significant. For example, empirical studies have shown that an asset's log-return distribution is non-Gaussian. It is characterized by heavy tails and high peaks (leptokurtic). There is also empirical evidence and economic arguments that suggest that equity returns and implied volatility are negatively correlated (also termed the leverage effect). This departure from normality plagues the Black-Scholes-Merton model with many problems. Another problem is that the Black-Scholes model assumes constant volatility. If the Black-Scholes model held, then the implied volatility for a particular stock would be the same for all strikes and maturities. In practice, the volatility surface (the three-dimensional graph of implied volatility against strike and maturity) is not flat. The typical shape of the implied volatility curve for a given maturity depends on the underlying instrument. This is the famous "volatility smile effects", which is a long-observed pattern in which at-the-money options tend to have lower implied volatilities than in- or out-of-the-money options. Finally, Black-Scholes model assumes continuous paths of the stock price, whereas in real life, jumps in equity prices frequently happen.

In the last two decades, option pricing has witnessed an explosion of new models that each relax some of the restrictive Black-Scholes (BS) (1973) assumptions. Examples include:
(i) the stochastic-interest-rate option models of Merton (1973) [38] and Amin and Jarrow (1992) [33];
(ii) the jump-diffusion/pure jump models of Bates (1991) [18], Madan and Chang (1996) [20], and Merton (1976) [37];
(iii) the constant-elasticity-of-variance (CEV) model of Cox and Ross (1976) [31];
(iv) the stochastic volatility models of Heston (1993) [27], Hull and White (1987a) [28], Melino and Turnbull (1990) [36], Scott (1987) [41], Stein and Stein (1991) [23], and Wiggins (1987) [43];

In the next section, we’ll focus on one of the approaches above, namely, stochastic volatility models.

4.2 Introduction To Stochastic Volatility Model

The volatility of a financial asset is defined as the standard deviation per unit of time of the continuously compounded asset returns. It is an important input to many financial decisions such as asset allocation, option pricing, and risk management.

Stochastic volatility models are one approach to resolve a shortcoming of the Black-Scholes model. In particular, Black-Scholes model assumes that the underlying volatility is constant over the life of the derivative, and unaffected by the changes in the price level of the underlying security. The observed market returns display non constant volatility, clustering and are not normal distributed. The implied volatility is the volatility of the underlying which, when substituted into Black-Scholes formula, gives a theoretical price equal to the market price. The implied volatilities of options in the market show dependence on strike and time to expiration and are not constant. If the Black-Scholes
model is correct, then all options on the same underlying asset should give the same implied volatility. However, Black-Scholes implied volatilities usually vary across strike prices and across maturities. Taking limitations of Black-Scholes into account, one might see the apparent benefits of using non-constant volatility models.

One of the appealing features of the Black-Scholes model is that it yields closed-form analytic formulae for many different kinds of options. It assumes a constant volatility, which is unobservable in the market. In practice, people often use implied volatility to measure option's relative value. Implied volatility is the volatility that, when used in a particular pricing model, yields a theoretical value for the option equal to the current market price of that option. Thus, if the assumption of constant volatility were right, the implied volatility using Black-Scholes model should be the same across all possible strikes and maturities. But extensive empirical evidences during the past decades have shown that it's not the case. In particular, from the following graph, it’s easy to see that the implied volatility surface is far from flat.

![Implied Volatility Surface](image)

Figure 4-1. Implied Volatility Surface

To model the stochastic evolution of volatility over time, an assumption about the stochastic process that governs its dynamics needs to be made. The specification of the process should be based on its ability to explain most of the empirical regularities
of volatility. Several empirical studies have investigated the time-series properties of volatility. The main findings includes:

(i) Fat tails. Since the early sixties, many empirical studies have found that the probability that extreme events will occur is greater than the corresponding probability calculated under the normal distribution. In other words, the empirical distribution of returns exhibits excess kurtosis; it accumulates more probability mass in the tails (i.e., "fat tails") than the normal distribution does. This is called a leptokurtic distribution.

(ii) Mean Reversion in Volatility. Empirical studies have found that volatility oscillates around a constant value. This phenomenon is termed "mean reversion," indicating that volatility tends to revert to a long-run mean.

(iii) Leverage Effect. In 1976, Fischer Black noted that there is a negative relationship between volatility and price changes. This phenomenon is termed "leverage effect." Christie (1982) [16] attributed the leverage effect to the fact that a drop in stock prices tends to increase the leverage of the firm, which in turns increases its risk as this is measured by volatility.

(iv) Clustering Effect. Any casual observation of financial time series reveals clusters of high and low volatility episodes. Mandelbrot (1963) [35] and Fama (1965) [24] reported evidence that periods of high (low) volatility are followed by periods of high (low) volatility. Mandelbrot has called this phenomenon "the clustering effect" of volatility. In either case, the sign of changes from one period to the next is unpredictable. Volatility clustering suggests the presence of autocorrelation in volatility changes.

After the stock market crash in 1987, it was evident that volatility of the stock returns could not be treated as a constant parameter. Hull and White's model (1987) [28] was one of the earliest and simplest stochastic volatility models that was introduced in the same year the stock market crashed. In particular, they considered the following dynamics:

$$dS(t) = \mu S(t) dt + \sqrt{V(t)} S(t) dW(t)$$
\[ dV(t) = aV(t)dt + bV(t)dW_2(t) \]

where \( dW_1^2 dW_2^2 = \rho dt \), and \( a, b \) are positive constants. In this case, the volatility \( \sigma_t = \sqrt{V(t)} \) is a geometric Brownian motion.

Scott (1989) [40] considered the case in which the logarithm of the volatility is a mean reverting process and this was further developed by Stein and Stein (1991) [23]. He assumed the stock price follows:

\[ dS(t) = \mu S(t)dt + \exp(V(t))S(t)dW_1(t) \]

\[ dV(t) = a(b - V(t))dt + cdW_2(t) \]

where \( dW_1(t)dW_2(t) = \rho dt \), and \( a, b \) and \( c \) are positive constants. In this case, the log-volatility \( V(t) = \log \sigma(t) \) is an Ornstein-Uhlenbeck (OU) process.

In 1993 Heston introduced a model where the volatility is related to a mean reverting square root process, commonly known as CIR process. The mean reverting square root process was first introduced by Cox, Ingersoll and Ross (1985) [17] to imitate the behavior of risk-free interest rate. Heston offers a stochastic volatility model that is not based on the BS formula. It provides a closed form solution for the price of a European call option when the spot asset is correlated with volatility (Heston 1993) [27]. These features made the Heston model popular for pricing plain vanilla options. In the next section we'll introduce Heston's model.

### 4.3 Introduction To Heston’s Model And Simulation Issue

Heston model (1993) is the most successful stochastic volatility model that attempts to capture the smile effect observed in implied volatilities of liquidly traded options, and to fulfill the gap of the unrealistic constant volatility assumed in the Black-Scholes model [9]. In Heston model, variances, not volatilities, are specified
to follow a mean-reverting square root process. This is a process that is widely applied in finance, for example, the CIR short rate model (1985) [17], and affine-structure model.

Since its inception in 1993, the Heston stochastic volatility model has received a growing attention among practitioners and academics. The model proposed by Heston [27] extends the Black and Scholes [9] model and includes it as a special case. Heston’s setting takes into account non-lognormal distribution of the assets returns, leverage effect, important mean-reverting property of volatility and it remains analytically tractable. It relaxes the constant volatility assumption in the classical Black-Scholes model by incorporating an instantaneous short term variance process. As such, a decent (though not all) number of smile and skew patterns can be built into volatility surfaces by a relatively restricted number of parameters.

The basic Heston’s model does not model stochastic volatilities directly, but stochastic variances. The process specifying the variance $V(t)$ is identical to the one that Cox-Ingersoll-Ross (1985) [17] apply for short interest rate, and is called mean-reverting square root process. The mean-reversion is a desired property for stochastic volatility or variance and is well documented by many empirical studies. Furthermore, it assumes that $S(t)$, the stock price process and the variance process evolve according to the following two SDEs under the risk-neutral measure:

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)[\rho dW_1(t) + \sqrt{1-\rho^2} dW_2(t)],$$

$$dV(t) = \kappa(\theta - V(t))dt + \sigma \sqrt{V(t)}dW_1(t).$$

The first equation gives the dynamics of the stock price: $S_t$ denotes the stock price at time $t$, $r$ is the risk neutral drift, $\sqrt{V(t)}$ is the volatility. The second equation gives the evolution of the variance which follows the square-root diffusion process: $\theta$ is the long-term level that variance gradually converges to, $\kappa$ represents the speed of mean reversion, and $\sigma$ is a parameter which determines the volatility of the variance process. $W_1(t)$ and $W_2(t)$ are two independent Brownian motion processes, and $\rho$ represents the
instantaneous correlation between the return process and the volatility process. If the parameters obey the condition \(2\kappa\theta \geq \sigma^2\) (known as the Feller condition) then the process \(V(t)\) is strictly positive. Typically, the correlation \(\rho\) is negative, pointing to the fact that a down-move in the stock price is correlated with an up-move in the volatility (leverage effect).

It is worthwhile mentioning that the variance process \(V(t)\) is noncentrally chi-square distributed. In fact,

\[
V(t)|V(s) \sim c\chi^2_d(\lambda)
\]

where

\[
c = \frac{\sigma^2(1 - e^{-\kappa(t-s)})}{4\kappa}, \quad d = \frac{4\kappa\theta}{\sigma^2}, \quad \lambda = \frac{4\kappa e^{-\kappa(t-s)}}{\sigma^2(1 - e^{-\kappa(t-s)})} V(s).
\]

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

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

\[
Var(V(t)|V(s)) = V(s)\left(\frac{\sigma^2}{\kappa}(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}) + \theta\frac{\sigma^2}{2\kappa}(1 - e^{-\kappa(t-s)})^2\right).
\]

If \(\kappa, \theta\) and \(\sigma\) satisfy the following condition:

\[
2\kappa\theta > \sigma^2, \quad V_0 > 0,
\]

it can be shown that variances \(V(t)\) are always positive and the variance process is then well-defined under the above condition.

Heston is the first one who proposes a closed-form price for a standard European call when using a square-root volatility process by inverting the characteristic function seen as a Fourier transform. But there is no close-form price for path-dependent exotic options, where the option payoff depends on the whole path of the underlying asset price. Many practical applications of models with Heston-dynamics involve the pricing
and hedging of path-dependent options, which, in turn, nearly always requires the introduction of Monte Carlo methods.

Given some arbitrary set of discrete times \( \{ t_i \}_{i=1}^N \), we want to generate random paths of the pair \((S(t), V(t))\) along those discrete times. This would be required, for instance, in the pricing of path-dependent securities with payout functions that depend on observations of \( S(t) \) at a given finite set of dates. In order to do that, it suffices to be able to generate \((S(t + \Delta), V(t + \Delta))\) given \((S(t), V(t))\). A few previously proposed techniques for updating \( S(t) \) and \( V(t) \) from time \( t \) to time \( t + \Delta \) include a "full truncation" scheme based on a naive Euler scheme, Kahl-Jackel discretization scheme involving an implicit Milstein scheme for the \( V \)-process and their "IJK" discretization for the stock process, Broadie-Kaya scheme based on an exact distribution of \( S(t) \), and Anderson’s QE scheme.

The naive Euler scheme with a "full truncation" fix would take the following form:

\[
\ln \hat{S}(t + \Delta) = \ln \hat{S}(t) + \left( r - \frac{1}{2} \hat{V}(t)^{+} \right) \Delta + \sqrt{\hat{V}(t)^{+}(\rho Z_1 + \sqrt{1 - \rho^2} Z_2)} \sqrt{\Delta},
\]

\[
\hat{V}(t + \Delta) = \hat{V}(t) + \kappa(\theta - \hat{V}(t)^{+}) \Delta + \sigma \sqrt{\hat{V}(t)^{+} Z_1 \sqrt{\Delta}},
\]

where \( Z_1 \) and \( Z_2 \) are independent standard Gaussian variables, and the notation \( x^{+} = \max(x, 0) \). Its main characteristic is that the process for \( V \) is allowed to go below zero, at which point the process for \( V \) becomes deterministic with an upward drift of \( \kappa \theta \).

The bias of the Monte Carlo estimator for the stock price using this naive Euler scheme becomes really big when \( \sigma \) is relatively big and \( 2 \kappa \theta << \sigma^2 \).

Kahl-Jackel (2006) [32] suggests discretizing the \( V \)-process using an implicit Milstein scheme, and coupled with their "IJK" discretization for the stock process. They proposed the following scheme:

\[
\ln \hat{S}(t + \Delta) = \ln \hat{S} + \left( r - \frac{1}{4} (\hat{V}(t + \Delta) + \hat{V}(t)) \right) \Delta + \rho \sqrt{\hat{V}(t) Z_1 \sqrt{\Delta}}
\]

\[
+ \frac{1}{2} \sqrt{\hat{V}(t + \Delta) + \hat{V}(t)} \sqrt{1 - \rho^2 \sqrt{Z_2 \sqrt{\Delta}}} + \frac{1}{4} \sigma \rho \Delta (Z_1^2 - 1),
\]
\[
\hat{V}(t + \Delta) = \frac{\hat{V}(t) + \kappa \theta \Delta + \sigma \sqrt{\hat{V}(t)} \sqrt{\Delta} + \frac{1}{4} \sigma^2 \Delta (Z_1^2 - 1)}{1 + \kappa \Delta}
\]

But one thing about this scheme is that it only results in positive paths for the \( V \) process if \( 4 \kappa \theta > \sigma^2 \). In practice, this restriction is rarely satisfied. Thus this scheme for \( V \) will produce negative values with substantial probability. To fix this, we use a similar approach like the "full truncation" Euler scheme. Basically whenever \( \hat{V}(t) \) drops below zero, we use the naive Euler scheme (deterministic with an upward drift of \( \kappa \theta \)) to get \( \hat{V}(t + \Delta) \). Otherwise, use the above scheme for \( \hat{V}(t) \) and make sure to use \( \hat{V}(t + \Delta)^+ \) and \( \hat{V}(t)^+ \) in sampling \( \ln \hat{S}(t + \Delta) \).

Broadie-Kaya (2006) [12] used the Poisson method to sample \( V(t + \Delta) \) directly. Basically, the scheme involves sampling from a Poisson distribution, and then sample from a central chi-square distribution with its degree of freedom parameter determined by the outcome of the Poisson draw. As we already mentioned in simulating CIR process, this method of sampling a non-central chi square random variable with degrees of freedom less than 1 is not efficient. After this, to obtain a bias-free scheme for sampling the stock price process, first integrate the SDE for \( V(t) \), and get

\[
V(t + \Delta) = V(t) + \int_t^{t+\Delta} \kappa (\theta - V(u)) du + \sigma \int_t^{t+\Delta} \sqrt{V(u)} dW_1(u)
\]

or we could write it as

\[
\int_t^{t+\Delta} \sqrt{V(u)} dW_1(u) = \sigma^{-1} (V(t + \Delta) - V(t) - \kappa \theta \Delta + \kappa \int_t^{t+\Delta} V(u) du).
\]

Now, by Ito’s formula, the log-stock price satisfies the following SDE:

\[
d \ln S(t) = (r - \frac{1}{2} V(t)) dt + \rho \sqrt{V(u)} dW_1(u) + \sqrt{1 - \rho^2} \sqrt{V(u)} dW_2(u)
\]

In integral form, we have

\[
\ln S(t + \Delta) = \ln S(t) + r \Delta + \frac{\rho}{\sigma} (V(t + \Delta) - V(t) - \kappa \theta \Delta)
\]
\[ (+\frac{\kappa \rho}{\sigma} - \frac{1}{2}) \int_t^{t+\Delta} V(u) du + \sqrt{1-\rho^2} \int_t^{t+\Delta} \sqrt{V(u)} dW_2(u) \]

after plugging in the expression of \( \int_t^{t+\Delta} \sqrt{V(u)} dW_2(u) \) from earlier. Since \( V(u) \) and \( W_2(u) \) are independent, the distribution of \( \int_t^{t+\Delta} \sqrt{V(u)} dW_2(u) \) given the path generated by \( V(t) \) is normal with mean 0 and variance \( \int_t^{t+\Delta} V(u) du \). Now it’s clear that conditional on \( V(t+\Delta) \) and \( \int_t^{t+\Delta} V(u) du \), the distribution of \( \ln X(t+\Delta) \) is Gaussian with easily computable first and second moments. They examined the characteristic function of \( \int_t^{t+\Delta} V(u) du \) and numerically Fourier-invert to get the distribution function, which is very complicated numerically. As mentioned in Anderson’s paper, rather than using Fourier methods, simply write

\[ \int_t^{t+\Delta} V(u) du \approx \Delta[\gamma_1 V(t) + \gamma_2 V(t+\Delta)] \]

for certain constants \( \gamma_1 \) and \( \gamma_2 \). There are multiple ways for setting \( \gamma_1 \) and \( \gamma_2 \), the simplest being the Euler-like setting: \( \gamma_1 = 1, \gamma_2 = 0 \). Or a central discretization:

\( \gamma_1 = \gamma_2 = \frac{1}{2} \).

We use our new method to simulate the path of \( V(t) \), and use the above approach instead of Broadie and Kaya’s approach to approximate \( \int_t^{t+\Delta} V(u) du \). After plugging in the approximation, we have:

\[
\ln \tilde{S}(t+\Delta) = \ln \tilde{S}(t) + r\Delta + \frac{\rho}{\sigma} (\tilde{V}(t+\Delta) - \tilde{V}(t)) - \kappa \theta \Delta + \Delta \left( \frac{\kappa \rho}{\sigma} - \frac{1}{2} \right) (\gamma_1 \tilde{V}(t) + \gamma_2 \tilde{V}(t+\Delta)) \\
+ \sqrt{\Delta} \sqrt{1-\rho^2} \sqrt{\gamma_1 \tilde{V}(t) + \gamma_2 \tilde{V}(t+\Delta)} Z \\
= \ln \hat{X}(t) + K_0 + K_1 \tilde{V}(t) + K_2 \tilde{V}(t+\Delta) + \sqrt{K_3 \tilde{V}(t) + K_4 \tilde{V}(t+\Delta)} Z
\]

where \( Z \) is a standard Gaussian random variable, independent of \( \tilde{V} \), and \( K_0, \ldots, K_4 \) are given by

\[
K_0 = (r - \frac{\rho \kappa \theta}{\sigma}) \Delta, \quad K_1 = \gamma_1 \Delta \left( \frac{\kappa \rho}{\sigma} - \frac{1}{2} \right) - \frac{\rho}{\sigma}.
\]
\[ K_2 = \gamma_2 \Delta \left( \frac{k^2}{\sigma} - \frac{1}{2} \right) + \frac{\rho}{\sigma}, \quad K_3 = \gamma_1 \Delta (1 - \rho^2), \quad K_4 = \gamma_2 \Delta (1 - \rho^2) \]

Our method to simulate the stock price follows the line of Anderson’s method. But instead of the QE scheme in his paper to approximate the distribution of volatility process, we use our new result to generate V-process precisely. And based on that, generate the path for stock price using the Euler scheme in Anderson’s paper.
CHAPTER 5
SIMULATION RESULTS AND APPLICATIONS IN OPTION PRICING

In this section, we show our simulation results on our new exact method. Specifically, we simulate a CIR process with different set of parameters using both our new method and several Euler schemes. We also simulate the Heston model for the stock price with different sets of parameters using both our new method and traditional Euler scheme. We’ll also use it to price European options and compare them with exact option prices.

5.1 Simulation Results For CIR Interest Rate Model

The CIR model describes the dynamics of the interest rate \( R(t) \) as a solution of the following stochastic differential equation (SDE):

\[
dR(t) = \alpha(b - R(t))dt + \sigma \sqrt{R(t)}dW(t), \quad R(0) = r_0 \geq 0
\]

(5–1)

for \( \alpha > 0, b > 0, \sigma > 0 \) and a standard Brownian motion \( W \).

For small \( R(t) \), the non-centrality parameter approaches zero, and the distribution of \( R(t + \Delta t) \) becomes proportional to that of an ordinary (central) chi-square distribution with \( d = 4\alpha b/\sigma^2 \) degrees of freedom. We recall that the density of a central chi-square distribution with \( d \) degrees of freedom is

\[
f(x; d) = \frac{1}{2^{d/2}\Gamma(d/2)}e^{-x/2}x^{d/2-1}.
\]

Thus, for parameters which satisfy the condition \( d = 4\alpha b/\sigma^2 << 2 \), the presence of the term \( x^{d/2-1} \) implies that, for small \( R(t) \), the density of \( R(t + \Delta t) \) will be very large around 0. This is the main reason that traditional Euler method fails to be accurate any more.

To test our exact method to simulate a CIR process, we use the following 3 different cases in the parameter table:
Table 5-1. Parameter table for CIR interest model simulation.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Case I</th>
<th>Case II</th>
<th>Case III</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>2</td>
<td>1.2</td>
<td>1</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>$b$</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>$R(0)$</td>
<td>0.3</td>
<td>0.1</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Here the degrees of freedom $d = 0.04, 0.11, 0.16$, respectively, all less than 1.

The time interval we consider is $[0, 1]$. We use Monte Carlo method to approximate $E(R(1))$, and compare the results from the new exact method, the Poisson method and two Euler Schemes (truncation and reflection). The theoretical mean $E(R(1)) = 0.310, 0.118, 0.067$, respectively. We run the Euler Schemes with time points $n = 1, 5, 7, 10, 14, 20, 30, 50$, and $m = 10000$ sample paths. The Poisson method and our new method, on the other hand, since they are exact, were run with $m = 10000$ sample paths and just $n = 1$ time point.

Table 5-2. Estimated expectation of interest rate at time $T = 1$ in test Case I. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>Euler1</th>
<th>Euler2</th>
<th>Poisson</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.624(0.008)</td>
<td>0.909(0.007)</td>
<td>0.320(0.011)</td>
<td>0.306(0.011)</td>
</tr>
<tr>
<td>1/5</td>
<td>0.559(0.011)</td>
<td>0.974(0.122)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/7</td>
<td>0.530(0.011)</td>
<td>0.905(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/10</td>
<td>0.525(0.012)</td>
<td>0.844(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/14</td>
<td>0.482(0.011)</td>
<td>0.786(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/20</td>
<td>0.480(0.012)</td>
<td>0.749(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/30</td>
<td>0.442(0.012)</td>
<td>0.707(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/50</td>
<td>0.421(0.012)</td>
<td>0.630(0.012)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5-3. Estimated expectation of interest rate at time $T = 1$ in test Case II. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>Euler1</th>
<th>Euler2</th>
<th>Poisson</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.217(0.003)</td>
<td>0.316(0.004)</td>
<td>0.119(0.004)</td>
<td>0.121(0.004)</td>
</tr>
<tr>
<td>1/5</td>
<td>0.202(0.004)</td>
<td>0.335(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/7</td>
<td>0.200(0.004)</td>
<td>0.321(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/10</td>
<td>0.194(0.004)</td>
<td>0.304(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/14</td>
<td>0.184(0.004)</td>
<td>0.291(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/20</td>
<td>0.177(0.004)</td>
<td>0.268(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/30</td>
<td>0.172(0.004)</td>
<td>0.249(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/50</td>
<td>0.167(0.004)</td>
<td>0.231(0.004)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-4. Estimated expectation of interest rate at time $T = 1$ in test Case III. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>Euler1</th>
<th>Euler2</th>
<th>Poisson</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.129(0.002)</td>
<td>0.186(0.001)</td>
<td>0.068(0.002)</td>
<td>0.066(0.002)</td>
</tr>
<tr>
<td>1/5</td>
<td>0.124(0.002)</td>
<td>0.211(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/7</td>
<td>0.118(0.002)</td>
<td>0.194(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/10</td>
<td>0.113(0.002)</td>
<td>0.188(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/14</td>
<td>0.110(0.002)</td>
<td>0.174(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/20</td>
<td>0.106(0.002)</td>
<td>0.169(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/30</td>
<td>0.100(0.002)</td>
<td>0.156(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/50</td>
<td>0.095(0.002)</td>
<td>0.143(0.003)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Obviously in those particular cases, there is a large bias with both Euler methods, since the interest rate process stays around 0 most of the time. The result from the new method is unbiased and much more accurate. The Poisson method is also very accurate in those cases, but when $n$ gets bigger, it would be much slower.
Figure 5-1. Comparison between different schemes in estimating $E[R(1)]$ in test Case I

Figure 5-2. Comparison between different schemes in estimating $E[R(1)]$ in test Case II
Figure 5-3. Comparison between different schemes in estimating \( E[R(1)] \) in test Case III

For the same set of parameters, we also tried to estimate the bond price \( E(e^{-\int_0^1 R(t) \, dt}) \). Here, since an integral is involved, we’ll approximate the integral by a summation, which requires us to simulate the process on a lot of time points. Since there is no analytical result for the expectation, we estimate it using the exact new method with \( 10^6 \) sample paths and \( 10^3 \) time points to estimate the integral. And the estimated bond prices are \( 0.821, 0.913, 0.948 \), respectively in those three cases. With \( m = 10000 \) sample paths and \( n = 10, 14, 20, 30, 50, 100 \) time points, we compared the performance of the Euler schemes, the Poisson method and the exact method:
Table 5-5. Estimated bond price in test Case I. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler1</th>
<th>Euler2</th>
<th>Poisson</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/10</td>
<td>0.749(0.003)</td>
<td>0.608(0.002)</td>
<td>0.817(0.002)</td>
<td>0.833(0.003)</td>
</tr>
<tr>
<td>1/14</td>
<td>0.749(0.003)</td>
<td>0.632(0.002)</td>
<td>0.821(0.002)</td>
<td>0.829(0.003)</td>
</tr>
<tr>
<td>1/20</td>
<td>0.767(0.003)</td>
<td>0.657(0.002)</td>
<td>0.819(0.002)</td>
<td>0.823(0.003)</td>
</tr>
<tr>
<td>1/30</td>
<td>0.771(0.003)</td>
<td>0.682(0.002)</td>
<td>0.820(0.002)</td>
<td>0.826(0.003)</td>
</tr>
<tr>
<td>1/50</td>
<td>0.776(0.003)</td>
<td>0.701(0.002)</td>
<td>0.825(0.002)</td>
<td>0.823(0.003)</td>
</tr>
<tr>
<td>1/100</td>
<td>0.780(0.003)</td>
<td>0.722(0.002)</td>
<td>0.819(0.003)</td>
<td>0.821(0.003)</td>
</tr>
</tbody>
</table>

Table 5-6. Estimated bond price in test Case II. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler1</th>
<th>Euler2</th>
<th>Poisson</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/10</td>
<td>0.876(0.002)</td>
<td>0.820(0.001)</td>
<td>0.913(0.001)</td>
<td>0.915(0.001)</td>
</tr>
<tr>
<td>1/14</td>
<td>0.879(0.002)</td>
<td>0.827(0.001)</td>
<td>0.912(0.001)</td>
<td>0.913(0.001)</td>
</tr>
<tr>
<td>1/20</td>
<td>0.884(0.002)</td>
<td>0.840(0.001)</td>
<td>0.911(0.001)</td>
<td>0.914(0.001)</td>
</tr>
<tr>
<td>1/30</td>
<td>0.886(0.002)</td>
<td>0.853(0.001)</td>
<td>0.913(0.001)</td>
<td>0.913(0.001)</td>
</tr>
<tr>
<td>1/50</td>
<td>0.891(0.002)</td>
<td>0.862(0.001)</td>
<td>0.912(0.001)</td>
<td>0.913(0.001)</td>
</tr>
<tr>
<td>1/100</td>
<td>0.893(0.001)</td>
<td>0.871(0.001)</td>
<td>0.912(0.001)</td>
<td>0.913(0.001)</td>
</tr>
</tbody>
</table>

Table 5-7. Estimated bond price in test Case III. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler1</th>
<th>Euler2</th>
<th>Poisson</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/10</td>
<td>0.921(0.001)</td>
<td>0.883(0.001)</td>
<td>0.949(0.001)</td>
<td>0.949(0.001)</td>
</tr>
<tr>
<td>1/14</td>
<td>0.925(0.001)</td>
<td>0.888(0.001)</td>
<td>0.947(0.001)</td>
<td>0.948(0.001)</td>
</tr>
<tr>
<td>1/20</td>
<td>0.928(0.001)</td>
<td>0.896(0.001)</td>
<td>0.948(0.001)</td>
<td>0.947(0.001)</td>
</tr>
<tr>
<td>1/30</td>
<td>0.930(0.001)</td>
<td>0.905(0.001)</td>
<td>0.948(0.001)</td>
<td>0.948(0.001)</td>
</tr>
<tr>
<td>1/50</td>
<td>0.933(0.001)</td>
<td>0.909(0.001)</td>
<td>0.949(0.001)</td>
<td>0.948(0.001)</td>
</tr>
<tr>
<td>1/100</td>
<td>0.936(0.001)</td>
<td>0.919(0.001)</td>
<td>0.948(0.001)</td>
<td>0.948(0.001)</td>
</tr>
</tbody>
</table>
From the above tables, it’s easy to see that both Euler schemes still have bias in pricing the bond price, whereas the Poisson method and our new exact method is much more accurate. But in terms of speed, Poisson method is much slower than our new exact method, even only 100 time points are used in the approximation of the integral. Our new exact method remains constant speed, no matter how many time points we’re interested in.

Figure 5-4. Comparison between different schemes in estimating $E\left(e^{-\int_0^1 R(t)dt}\right)$ in test Case I
Figure 5-5. Comparison between different schemes in estimating $E(e^{-\int_0^t R(t)dt})$ in test Case II

Figure 5-6. Comparison between different schemes in estimating $E(e^{-\int_0^t R(t)dt})$ in test Case III
5.2 Simulation Results For Heston Model

The Heston model assumes that $S_t$, the stock price and the variance processes evolve according to the following two SDEs under the risk-neutral measure:

$$dS_t = rS_t dt + \sqrt{V_t} S_t \left[ \rho dW_t^{(1)} + \sqrt{1 - \rho^2} dW_t^{(2)} \right],$$

$$dV_t = \kappa (\theta - V_t) dt + \sigma \sqrt{V_t} dW_t^{(1)}.$$

The first equation gives the dynamics of the stock price: $S_t$ denotes the stock price at time $t$, $r$ is the risk neutral drift, $\sqrt{V_t}$ is the volatility. The second equation gives the evolution of the variance which follows the square-root diffusion process: $\theta$ is the long-term level that variance gradually converges to. $\kappa$ represents the speed of mean reversion, and $\sigma$ is a parameter which determines the volatility of the variance process. $W_t^{(1)}$ and $W_t^{(2)}$ are two independent Brownian motion processes, and $\rho$ represents the instantaneous correlation between the return process and the volatility process.

To test our new exact method, we turn to the pricing of European options in the Heston model. This constitutes a standard test case, as prices can be computed with great precision either from the analytical result in chapter 4, or from the FFT method in Carr and Madan (1998) [14]. We consider a call option $C$ maturing at time $T$ with strike $K$; let the exact option price at time 0 be $C(0)$. We use our exact method to simulate $V(T)$, and use the approach in Anderson [4] to approximate $X(T)$ with $\hat{X}(T)$. Then we can get the approximation $\hat{C}(0)$ to the option price by computing the expectation

$$\hat{C}(0) = E((\hat{X}(T) - K)^+).$$

In order to estimate $\hat{C}(0)$, we use Monte Carlo methods. Specifically, we draw $N$ independent samples of $\hat{X}^{(1)}(T), \hat{X}^{(2)}(T), \ldots, \hat{X}^{(N)}(T)$ using an equidistant time-grid with
fixed step; then

\[ \hat{C}(0) \approx \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\hat{X}^{(i)}(T) - K^+). \]

And we use the FFT method in Carr and Madan (1998) [14] to compute exact price \( C(0) \). This value is used as a benchmark, and based on that, we compare our new method with traditional Euler scheme.

We consider three different sets of parameters, which are listed in the following table:

Table 5-8. Parameter table for Heston model simulation. In all cases \( r = 0, V(0) = \theta \) and \( X(0) = 100 \).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Case I</th>
<th>Case II</th>
<th>Case III</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma )</td>
<td>1</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>0.5</td>
<td>0.3</td>
<td>1</td>
</tr>
<tr>
<td>( \rho )</td>
<td>-0.9</td>
<td>-0.5</td>
<td>-0.3</td>
</tr>
<tr>
<td>( T )</td>
<td>10</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>( V(0), \theta )</td>
<td>0.04</td>
<td>0.04</td>
<td>0.09</td>
</tr>
</tbody>
</table>

For each set of parameters above, we compare the option price using three different strikes (\( K=70, K=100, K=140 \)). The analytic value is set to be a benchmark for comparison. And it is computed using the FFT method in Carr and Madan (1998) [14].

In our numerical results, we use the following discretization schemes: the "full truncation" Euler scheme; the Poisson method; and our new scheme. To keep the sample standard deviation low, all tests were run using a high number of paths, \( N = 10^5 \). And the values of time step \( \Delta \) range from \( 1/32 \) year to 1 year.
Table 5-9. Estimated European call option price in test Case I. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler</th>
<th>Poisson</th>
<th>New</th>
<th>Analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K=70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>39.722(0.118)</td>
<td>36.136(0.071)</td>
<td>36.126(0.071)</td>
<td>35.856</td>
</tr>
<tr>
<td>1/2</td>
<td>38.262(0.095)</td>
<td>35.937(0.071)</td>
<td>36.037(0.071)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>37.039(0.082)</td>
<td>35.905(0.071)</td>
<td>35.890(0.071)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>36.230(0.077)</td>
<td>35.997(0.071)</td>
<td>35.860(0.071)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>36.159(0.074)</td>
<td>35.852(0.071)</td>
<td>35.809(0.071)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>35.980(0.073)</td>
<td>35.899(0.071)</td>
<td>35.727(0.071)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>19.313(0.092)</td>
<td>13.308(0.042)</td>
<td>13.340(0.042)</td>
<td>13.007</td>
</tr>
<tr>
<td>1/2</td>
<td>16.738(0.066)</td>
<td>13.158(0.042)</td>
<td>13.160(0.042)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>15.145(0.054)</td>
<td>13.120(0.042)</td>
<td>13.186(0.042)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>14.208(0.048)</td>
<td>13.106(0.042)</td>
<td>13.065(0.042)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>13.618(0.045)</td>
<td>13.075(0.042)</td>
<td>13.138(0.042)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>13.302(0.043)</td>
<td>13.036(0.042)</td>
<td>12.990(0.042)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4.514(0.060)</td>
<td>0.273(0.008)</td>
<td>0.270(0.008)</td>
<td>0.330</td>
</tr>
<tr>
<td>1/2</td>
<td>2.245(0.033)</td>
<td>0.279(0.008)</td>
<td>0.281(0.008)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>1.043(0.018)</td>
<td>0.286(0.008)</td>
<td>0.301(0.008)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>0.565(0.011)</td>
<td>0.289(0.008)</td>
<td>0.294(0.008)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>0.401(0.009)</td>
<td>0.291(0.008)</td>
<td>0.292(0.008)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>0.343(0.009)</td>
<td>0.293(0.008)</td>
<td>0.309(0.008)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5-7. Comparison between different schemes in estimating European option price in test Case I (K=70)

Figure 5-8. Comparison between different schemes in estimating European option price in test Case I (K=100)
Estimated option price
Comparison between different schemes
Truncation Euler Scheme

Figure 5-9. Comparison between different schemes in estimating European option price in test Case I (K=140)
Table 5-10. Estimated European call option price in test Case II. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler</th>
<th>Poisson</th>
<th>New</th>
<th>Analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K=70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>41.369(0.219)</td>
<td>37.012(0.141)</td>
<td>37.058(0.140)</td>
<td>36.602</td>
</tr>
<tr>
<td>1/2</td>
<td>39.954(0.194)</td>
<td>37.047(0.140)</td>
<td>36.941(0.149)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>38.328(0.167)</td>
<td>36.981(0.139)</td>
<td>37.090(0.142)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>37.649(0.155)</td>
<td>36.908(0.142)</td>
<td>36.994(0.145)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>37.570(0.177)</td>
<td>37.127(0.139)</td>
<td>36.906(0.139)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>37.469(0.147)</td>
<td>36.836(0.138)</td>
<td>37.133(0.138)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23.931(0.263)</td>
<td>16.768(0.132)</td>
<td>16.543(0.130)</td>
<td>16.203</td>
</tr>
<tr>
<td>1/2</td>
<td>20.991(0.295)</td>
<td>16.637(0.148)</td>
<td>16.498(0.122)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>18.578(0.145)</td>
<td>16.578(0.127)</td>
<td>16.619(0.134)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>17.641(0.149)</td>
<td>16.620(0.134)</td>
<td>16.770(0.169)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>17.264(0.135)</td>
<td>16.443(0.165)</td>
<td>16.680(0.168)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>16.785(0.123)</td>
<td>16.386(0.122)</td>
<td>16.414(0.124)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11.727(0.265)</td>
<td>5.146(0.112)</td>
<td>5.118(0.100)</td>
<td>4.875</td>
</tr>
<tr>
<td>1/2</td>
<td>8.207(0.149)</td>
<td>5.203(0.143)</td>
<td>5.174(0.142)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>6.627(0.122)</td>
<td>5.184(0.133)</td>
<td>5.103(0.098)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>5.832(0.108)</td>
<td>5.174(0.145)</td>
<td>5.246(0.132)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>5.599(0.130)</td>
<td>5.101(0.102)</td>
<td>5.175(0.113)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>5.288(0.120)</td>
<td>5.011(0.110)</td>
<td>5.142(0.113)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5-10. Comparison between different schemes in estimating European option price in test Case II (K=70)

Figure 5-11. Comparison between different schemes in estimating European option price in test Case II (K=100)
Figure 5-12. Comparison between different schemes in estimating European option price in test Case II (K=140)
Table 5-11. Estimated European call option price in test Case III. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler</th>
<th>Poisson</th>
<th>New</th>
<th>Analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K=70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>41.456(0.243)</td>
<td>38.579(0.174)</td>
<td>38.970(0.196)</td>
<td>38.721</td>
</tr>
<tr>
<td>1/2</td>
<td>39.828(0.202)</td>
<td>38.986(0.192)</td>
<td>38.978(0.183)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>39.602(0.197)</td>
<td>38.678(0.182)</td>
<td>38.814(0.192)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>39.257(0.231)</td>
<td>38.643(0.215)</td>
<td>38.864(0.182)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>38.938(0.187)</td>
<td>38.884(0.179)</td>
<td>38.554(0.183)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>38.408(0.191)</td>
<td>38.732(0.187)</td>
<td>38.639(0.187)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>26.677(0.382)</td>
<td>22.188(0.150)</td>
<td>22.109(0.339)</td>
<td>21.748</td>
</tr>
<tr>
<td>1/2</td>
<td>23.913(0.184)</td>
<td>22.104(0.171)</td>
<td>21.559(0.167)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>23.019(0.177)</td>
<td>21.438(0.177)</td>
<td>21.559(0.163)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>22.271(0.168)</td>
<td>21.549(0.175)</td>
<td>21.582(0.160)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>22.125(0.164)</td>
<td>21.704(0.124)</td>
<td>21.723(0.156)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>21.877(0.161)</td>
<td>21.733(0.159)</td>
<td>21.692(0.169)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14.445(0.199)</td>
<td>9.665(0.123)</td>
<td>9.887(0.140)</td>
<td>9.981</td>
</tr>
<tr>
<td>1/2</td>
<td>12.350(0.175)</td>
<td>9.736(0.127)</td>
<td>9.656(0.130)</td>
<td></td>
</tr>
<tr>
<td>1/4</td>
<td>11.028(0.157)</td>
<td>10.012(0.130)</td>
<td>9.718(0.129)</td>
<td></td>
</tr>
<tr>
<td>1/8</td>
<td>10.705(0.150)</td>
<td>10.076(0.176)</td>
<td>10.020(0.185)</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>10.384(0.148)</td>
<td>9.865(0.156)</td>
<td>10.083(0.139)</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>10.081(0.148)</td>
<td>9.902(0.155)</td>
<td>9.817(0.127)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5-13. Comparison between different schemes in estimating European option price in test Case III (K=70)

Figure 5-14. Comparison between different schemes in estimating European option price in test Case III (K=100)
Figure 5-15. Comparison between different schemes in estimating European option price in test Case III (K=140)

From the above tables and graphs, we could see that the new method is more efficient and accurate than the "full truncation" Euler Scheme, which has a large bias in all three cases. And as expected, Poisson method is pretty accurate but takes a much longer time to yield the result than our new method. And as the time-step gets finer, it will become even much slower.

Finally, we look at price of discrete Asian options under Heston’s model. Using the parameters in case three, we use Monte Carlo method to estimate the price of a discrete Asian call option on time interval [0, 5]. Unfortunately there is no analytical formula for discrete Asian call option under Heston’s model. We therefore approximate the price using our Monte Carlo simulation with \( m = 10^6 \) and \( n = 1000 \), and use it as a benchmark. Once again the strike prices are \( K = 70, 100, 140 \). We would use \( m = 10000 \) and \( n = 12, 30, 52, 100, 252 \) time points in the simulation. In this case, Poisson method would be drastically too slow to use in simulation. Our new algorithm turns out to be both accurate and efficient.
Table 5-12. Estimated discrete Asian call option price in test Case III. Numbers in parentheses are sample standard deviations.

<table>
<thead>
<tr>
<th>Δ</th>
<th>Euler</th>
<th>Poisson</th>
<th>New</th>
<th>Analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K=70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/12</td>
<td>33.094(0.357)</td>
<td>32.672(0.339)</td>
<td>32.920(0.328)</td>
<td>32.907</td>
</tr>
<tr>
<td>1/30</td>
<td>32.952(0.349)</td>
<td>33.092(0.332)</td>
<td>33.307(0.352)</td>
<td></td>
</tr>
<tr>
<td>1/52</td>
<td>32.712(0.333)</td>
<td>33.099(0.340)</td>
<td>33.281(0.352)</td>
<td></td>
</tr>
<tr>
<td>1/100</td>
<td>32.543(0.326)</td>
<td>32.987(0.337)</td>
<td>32.885(0.340)</td>
<td></td>
</tr>
<tr>
<td>1/252</td>
<td>32.771(0.333)</td>
<td>32.876(0.334)</td>
<td>32.674(0.359)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/12</td>
<td>13.533(0.294)</td>
<td>13.004(0.278)</td>
<td>12.770(0.275)</td>
<td>12.707</td>
</tr>
<tr>
<td>1/30</td>
<td>12.887(0.259)</td>
<td>12.633(0.263)</td>
<td>12.819(0.264)</td>
<td></td>
</tr>
<tr>
<td>1/52</td>
<td>12.687(0.258)</td>
<td>13.024(0.306)</td>
<td>12.471(0.293)</td>
<td></td>
</tr>
<tr>
<td>1/100</td>
<td>12.580(0.250)</td>
<td>12.905(0.274)</td>
<td>12.738(0.270)</td>
<td></td>
</tr>
<tr>
<td>1/252</td>
<td>12.952(0.267)</td>
<td>12.802(0.267)</td>
<td>12.904(0.261)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K=140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/12</td>
<td>3.645(0.192)</td>
<td>3.741(0.245)</td>
<td>3.360(0.218)</td>
<td>3.122</td>
</tr>
<tr>
<td>1/30</td>
<td>3.543(0.200)</td>
<td>3.253(0.183)</td>
<td>3.031(0.186)</td>
<td></td>
</tr>
<tr>
<td>1/52</td>
<td>3.081(0.180)</td>
<td>2.759(0.139)</td>
<td>3.179(0.170)</td>
<td></td>
</tr>
<tr>
<td>1/100</td>
<td>2.895(0.159)</td>
<td>3.095(0.183)</td>
<td>3.320(0.244)</td>
<td></td>
</tr>
<tr>
<td>1/252</td>
<td>3.259(0.179)</td>
<td>3.231(0.202)</td>
<td>3.098(0.184)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5-16. Comparison between different schemes in estimating discrete Asian option price in test Case III (K=70)
Figure 5-17. Comparison between different schemes in estimating discrete Asian option price in test Case III (K=100)

Figure 5-18. Comparison between different schemes in estimating discrete Asian option price in test Case III (K=140)
In this paper, a fast and efficient way of simulating CIR process is introduced. We proved a theorem that could be used to generate a non-central Chi-square random variable. And the way we generate it avoids the problem with the traditional method that involves Poisson random variable generation. Our new method is exact compared with Euler schemes, and fast compared with Poisson method. Especially in the case when $d << 1$ and we need to simulate a CIR process along many time points, traditional Poisson method is very slow since the non-centrality parameter is big. On the other hand, our new method for generating a CIR process is independent of the number of time points needed in the simulation, thus much quicker than the Poisson method. We proposed several numerical tests to test our algorithm and the results have shown that our method is very efficient and suitable to be used in this particular case.
REFERENCES


BIOGRAPHICAL SKETCH

Anqi Shao was born in Urumqi, China in 1982. He moved to Yantai with his family in 1994 and attended No. 1 high school there. He went to the department of mathematics in 1999, and got his bachelor's degree from Shandong University in 2003. He then went to graduate school at Lehigh University in 2004, and transferred to University of Florida in 2005. He received Master of Science in mathematics in 2008 and Master of Statistics in 2009. He completed his doctorate of philosophy in the area of mathematical finance in the summer of 2012.