University of Tartu

Faculty of Science and Technology

Institute of Mathematics and Statistics

Joseph Haske

Option Pricing Using Stochastic Volatility Models

Actuarial and Financial Engineering

Master's Thesis (30 ECTS)

Supervisor: Toomas Raus

Tartu 2022

Option Pricing Using Stochastic Volatility Models

Master's Thesis

Joseph Haske

Abstract. The purpose of this thesis is to explore stochastic volatility models to price American and European options. The two methods used are both based on a quadrinomial tree, but the first uses an Ornstein-Uhlenbeck process and the Monte Carlo method with a quadrinomial recombining tree and the second uses the Heston model and a tree-based approach that combines a grid and bilinear interpolation to estimate the option price. The thesis is split into four chapters. In the first chapter, it gives an overview of options, option pricing models, and numerical methods. The second chapter discusses the quadrinomial recombining tree, and the third presents the tree-based approach that uses a grid and bilinear interpolation. Finally the fourth, presents the results of both methods and then compares their performance and flexibility.

CERCS research specialization: P160 Statistics, operations research, programming, financial and actuarial mathematics.

Keywords: Options, option pricing, stochastic volatility models, recombining tree methods.

Optsioonide hindamine stohhastilise volatiilsuse mudelite korral

Magistritöö

Joseph Haske

Lühikokkuvõte. Selle lõputöö eesmärk on uurida Euroopa ja Ameerika optsioonide hindamist stohhastilise volatiilsuse korral. Töös vaadeldakse kahte kirjanduses välja pakutud meetodit optsioonide hindamiseks, kui volatiilsust kirjeldatakse Ornstein-Uhlenbecki ja või selle erijuhu Hestoni mudeliga. Mõlemad meetodid kasutavad rekombineerivat nelja haruga alusvara hinnapuud, kuid esimene meetod kasutab täiendavalt Monte Carlo meetodit ning teine meetod kasutab rekombineeruvuse tagamiseks bilineaarset interpolatsiooni. Lõputöö on jagatud neljaks peatükiks. Esimeses peatükis antakse ülevaade optsioonidest, standardsest Black-Scholesi mudelist, numbrilistest meetoditest optsiooni hindamiseks ning ülevaade stohhastilise volatiilsuse mudelitest. Töö teises ja kolmandas peatükis vaadeldakse ülal nimetatud kahte meetodit optsiooni hinna leidmiseks stohhastilise volatiilsuse korral. Töö neljandas peatükis tuuakse numbriliste eksperimentide tulemused nende meetodite kohta ning võrreldakse meetodite ajamahukust ning täpsust.

CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Märksõnad: Optsioonid, optsioonide hindamine, stohhastilised volatiilsuse mudelid, rekombineerivad hinnapuud.

Contents

Introducton

Option pricing is a cornerstone of financial engineering. From the time that the publications by Black and Scholes (1973) and Merton (1973) were published, the Black-Scholes option pricing model was widely used until the 1987 stock market crash when it was noticed that volatility isn't constant across the range of options. Since then, many avenues have been explored to improve model accuracy. Key among these are non-constant volatility models. Some models use a deterministic function on stock price to determine the volatility, however those explored in this thesis use a stochastic process to model volatility.

A common class of processes used to model stochastic volatility are mean-reverting processes. The volatility process used in this thesis is the Ornstein-Uhlenbeck process. One specific case of this process is the Heston model, which is also used in this thesis. The main idea of mean-reverting models is that the stochastic volatility process has a mean level that it may deviate from; but, after the deviation, it will always trend back toward its mean.

Since many of the non-constant and stochastic volatility models don't have closed form solutions or only have such solutions under certain payoff conditions, numerical methods of computation must be used. The three major classes in financial engineering are the Monte Carlo method, finite difference methods, and binomial recombining trees. This thesis uses elements of all three to estimate option prices using two different calculation methods.

The first method uses a quadrinomial recombining tree, or a tree with four successors to each node, and the Monte Carlo method. It was originally presented by Florescu and Viens (2008). In many methods involving stochastic volatility, the distribution of both the stock price and volatility are assumed to be normally distributed. However, in this method, the underlying volatility distribution is estimated using a genetic algorithm and historical data. It then goes on to use this distribution to construct a quadrinomial tree, which is used to calculate the price of the option.

The second method, uses only estimated parameters to construct a grid of values for price and volatility. Bilinear interpolation is to estimate option values anywhere within the grid. Similarly to tree methods, the volatility and price may only go up or down when calculating the successors of each grid point to find its value. This method was originally presented by Vellekoop and Nieuwenhuis (2009) and is based on the Heston model (Heston 1993).

The thesis starts in chapter one with an overview of options, option pricing models and numerical methods. It covers the main differences between constant, non-constant, and stochastic volatility models and the motivations and for each. The models presented are the Black-Scholes model, the constant elasticity of variance model, mean-reverting stochastic volatility models. It also discusses the Heston model which is a special case of the mean-reverting model. Additionally, gives a brief overview of the numerical simulation methods mentioned above, the Monte Carlo method, finite difference methods, and the binomial model.

The next two chapters discuss the two methods outlined above. They present a detailed discussion of the parts of each model. The discussion includes the equations that make up the foundation of the model, and build a description of how each method works from that foundation. Each of these methods have been implemented using the C programming language which was chosen due to its speed, low overhead, and simplicity.

The results of the numerical experiments are laid out in chapter four. This chapter discusses the main results for each method using the parameters presented in the respective papers. Then, it finishes by presenting a direct comparison of the two models and their relative flexibility, accuracy, and performance.

1 Background on Options and Their Pricing

The purpose of this chapter is to provide a background of options and how they are priced. As such, the papers by Black and Scholes (1973), Merton (1973), and Heston (1993) are all heavily referenced since they provide the framework that the methods explored in this thesis are based on. Since they provide a collection of results from various sources, the textbooks by Baxter and Rennie (1996), and Bingham and Kiesel (2004) are also used.

1.1 Option Definitions

An **option** is a financial instrument giving the owner the right, but not the obligation, to buy or sell an asset on or before a specified date at a specified price. The specified date is called the **expiration** or **maturity date**, and the specified price is the **exercise** or **striking price**. An option with the right to buy an asset is called a **call option**, and one with the right to sell an asset is called a **put option**. An option which can only be exercised on the expiration date is called a **European option**, whereas an option which can be exercised on or before the expiration date is called an **American option**.

1.2 Option Price

In order to give the payoff of an option, first, remember that the owner of the option has the right, but not the obligation to buy or sell an asset. Therefore, consider a call option with exercise price E on a stock with price $S_T = S(T)$ at time T. Note that throughout this paper S_T and $S(T)$ will be used interchangeably to clarify or simplify notation as needed. If $S_T < E$ the owner would not exercise the option and the option expires worthless, but if $S_T > E$ then the owner exercises the option and immediately sells the stock the revenue is $S_T - E$. And therefore, the payoff of a call option C with exercise price E at its expiration date T is given by

$$C(S_T, E) = \max\{S_T - E, 0\}.$$

Through the same logic the payoff of a put option P with exercise price E at its expiration date T can be obtained, and is

$$P(S_T, E) = \max\{E - S_T, 0\}.$$

To price options for which the payoff depends only on one underlying asset, it is enough to consider a model with two securities. The first, the underlying asset with price $S(t)$, and the second, a risk-free investment, such as a bond, with price $B(t)$. In this model, it is assumed that short-selling and borrowing of money is not limited. Normally assets, in particular stocks, are bought at one price $S(0)$ and sold at a later time T for another, hopefully higher, price $S(T)$. Short selling a stock is the process of selling stock that isn't owned, but rather borrowed, by the investor. In this regard, they own negative shares of the stock for which they immediately receive $S(0)$, and which, they must later buy back at a later time T for price $S(T)$.

In this model, the other important assumption is that arbitrage strategies do not exist. Arbitrage is, in essence, a way to make money without any risk. To present this idea in concrete terms the example from Bingham and Kiesel (2004) can be used. Take a simple market with one bond with price $B(t)$, one stock with price $S(t)$, and one European call option with exercise price $E = 1$ and price $C(t)$ on the stock. In this simple example, there are two tradeable time moments, time $t = 0$ and the expiration time of the option $t = T$. The current prices of the instruments are $B(0) = 1,\ S(0) = 1,\ C(0) = 0.2$. At time T, the stock may move up or down by a known amount from $S(0)$, meaning there are two possible values of $S(T)$. However, the bond appreciates at the risk-free rate $r = 0.25$ regardless of what the market does. Therefore, at time T the value of the bond is $B(T) = 1.25$). The value of the stock and its option if the stock goes up in price is $S_u(T) = 1.75$, and $C_u(T) = 0.75$. However, if the stock price goes down down the prices are $S_d(T) = 0.75$ and $C_d(T) = 0$. Consider the portfolio $(1.8, -3, 4)$. This ordered triple denotes the purchase of 1.8 bonds, the short sale of 3 stocks, and the purchase of 4 options. A profit of $-1.8+3-4(0.2) = 0.4$ is obtained upon creation of this portfolio. At time T, if the price of the stock goes up the value of the portfolio is $1.8(1.25)-3(1.75)+4(0.75) = 0$, and likewise if the price goes down the value is $1.8(1.25)-3(0.75)+0 = 0$. This represents an arbitrage opportunity, because the profit of 0.4 was obtained at $t = 0$ with no future risk to the investor.

To price the option, consider an arbitrage-free market model where it is not possible to conduct an arbitrage strategy. It is well known that the model is arbitrage free, if there exists a risk-neutral martingale measure $\mathbb{Q}$ so that

$$\mathbb{E}_{\mathbb{Q}}(S_{t'}|S_t) = S_t e^{r(t'-t)} \tag{1.1}$$

for all $t' > t$ where r is the continuously compounded annual risk-free interest rate. Under this martingale measure, a risk neutral investor will be indifferent between buying the stock or investing the same amount at the risk-free rate. Then, the price V of a European call option at time t is given by $V_C = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}(C(S_T, E) \mid S_t)$ and for a European put option $V_P = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}(P(S_T, E) \mid S_t)$. It is worth noting that the value of an American option is always at least as high as than that of a European option, meaning $V_{C_A}(t) \geq V_{C_E}(t)$ and $V_{P_A}(t) \geq V_{P_E}(t)$. This difference in price is due to the ability to exercise the American option before its expiration.

Another useful tool in option pricing is the put-call parity condition. This can be obtained by making two hypothetical portfolios. The first is made by buying a call option and selling a put option at prices $V_C(S_t, t)$ and $V_P(S_t, t)$. The payoffs at time T with exercise price E are $C(S_T, E) = \max\{S_T - E, 0\}$ and $P(S_T, E) = \max\{E - S_t, 0\}$, so the value of the portfolio at time T is $C_T - P_T = S_T - E$. The second portfolio is made by buying one share now at price S_t which will be worth S_T at time T and selling E zero-coupon bonds for price B_t now which will each be worth $B_T = 1$ at time T. The value of the second portfolio is also $S_T - E$ at time T. Therefore, the discounted present values of the two portfolios at time t must be the same under the theory of rational pricing, and

$$V_C(S_t, t) - V_P(S_t, t) = S_t - EB_t. \tag{1.2}$$

1.3 Stock Price Behavior

In order to model stock price, the standard equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{1.3}$$

is used; and in it, μ is the drift, σ is the volatility of the rate of return, and dW_t is a change in Brownian motion. The stochastic process W is a standard Brownian motion (Wiener process) on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ if:

- $W_0 = 0$ with probability 1
- The increments of W are normally distributed as $W_{t+u} - W_t \sim N(0, u)$
- The non-intersecting increments are independently distributed. That is $W_{t+u} - W_t$ and $W_{s+v} - W_s$ independent random variables if $[t, t+u] \cap [s, s+v] = \emptyset$
- The Brownian motion W_t is continuous at time t with probability 1

Using Ito's lemma it can be shown that for the log-price then

$$d(ln(S_t)) = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t$$

holds. And denoting $X_t = ln(S_t)$ yields

$$dX_t = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t.$$

1.4 Black-Scholes Market Model

The Black-Scholes model was presented by Black and Scholes in 1973 gives an arbitrage free price for European options; and it has the following assumptions:

- Price of the underlying stock follows the equation defined in (1.3),
- The short-term interest rate $r(t)$ is known.
- The rate of return of the stock price has a known variance $\sigma(t)$.
- The stock does not pay dividends.
- The option is a European option.
- There are no transaction costs associated with buying or selling the stock or option.
- It is possible to borrow any fraction of the price of the security at the short-term interest rate.
- There aren't any penalties to short selling.

Given the above assumptions, Black and Scholes (1973) proved, using Ito's lemma, that if the price of the underlying asset is $S(t)$ then the option price $V(S(t),t)$ at time t satisfies the second order partial differential equation

$$\frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 V(S(t),t)}{\partial S(t)^2} + rS(t)\frac{\partial V(S(t),t)}{\partial S(t)} - rV(S(t),t) + \frac{\partial V(S(t),t)}{\partial t} = 0. \tag{1.4}$$

This partial differential equation has infinitely many solutions. In order to get a unique solution, use the fact that the option price at time T is equal to the payoff. For example, in the case of a European call option $V_C(S_T,T) = \max\{S_T - E, 0\}$ equation (1.4) has the solution

$$V_C(S_t,t) = S_t N(d_1) - Ee^{-r(T-t)}N(d_2), \tag{1.5}$$

where

$$d_1 = \frac{ln(S_t/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

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

In this set of equations, N is the standard normal distribution cumulative distribution function. Similarly to the call option, for a put option with payoff $V_P(S_T,T) = \max(E - S_T, 0)$,

$$V_P(S_t,t) = Ee^{-r(T-t)}N(-d_2) - S_t N(-d_1)$$

may be obtained, where the equations for d_1 and d_2 are the same as in (1.5).

1.5 Numerical Estimation Methods

The Black-Scholes model provides a nice closed-form solution for European options with known functions $r(t)$ and $\sigma(t)$. However, for American and exotic options, and for models with stochastic volatility, numerical methods must be used to obtain the price of the option. In this section three of the most well known methods for numerical estimation are presented as well as the advantages and drawbacks of each.

1.5.1 Monte Carlo Method

For option pricing, the Monte Carlo method simulates a large number of paths N through the stochastic differential equation for price and uses an average of the results to obtain the price of the option. To do this, start by computing the price of the underlying asset at time T using

$$(S_T)_n = S_0 e^{(r+\frac{\sigma^2}{2})T+\sigma W_T}$$

for runs $n = 1, 2, ..., N$. The value of the option at time T is equal to the payoff, so for a call option $(V_C(S_T, T))_n = \max\{(S_T)_n - E, 0\}$ for $n = 1, 2, ..., N$. This gives the estimate for the risk-neutral expectation of

$$\mathbb{E}_{\mathbb{Q}}(V(S_T, T)) = \frac{1}{N} \sum_{n=1}^{N} (V_C((S_T)_n, T))_n.$$

Discounting this price to time $t = 0$, gives the price of the option,

$$V(S_0, 0) = e^{-rT} \mathbb{E}_{\mathbb{Q}}(V(S_T, T)).$$

The size of N is the main factor in determining the accuracy of this method. The Monte Carlo method's largest drawback is that it can be very computationally expensive for processes which converge slowly. However, the method is very flexible and works well for problems in high dimensions. (Glasserman, 2003).

1.5.2 Finite Difference Methods

The finite difference methods are based around the fact that

$$\frac{\partial v(x,t)}{\partial x} = \lim_{\Delta x \to 0} \frac{v(x + \Delta x, t) - v(x,t)}{\Delta x}$$

and

$$\frac{\partial^2 v(x,t)}{\partial x^2} = \lim_{\Delta x \to 0} \frac{v(x + \Delta x, t) - 2v(x,t) + v(x - \Delta x, t)}{\Delta x^2}.$$

Using these formulas, alongside a set of initial boundary conditions, allows the estimation $v(x,t)$ within a range of values in the form of a grid. To get values between grid-points, interpolation is used. The choice of accurate boundary conditions and values, and small enough values of Δx and Δt are required to obtain accurate estimates. Unlike with the Monte Carlo method, determining the error is not straightforward and can only be estimated. However, this type of method works for a range of situations and with appropriate boundary conditions can be very accurate. (Thomas, 1995).

1.5.3 Binomial Model

The binomial model is a method of approximating the stochastic process of the stock price by allowing it to move in two directions, up or down. In order to start take the time period which starts at time $t = 0$ and ends at time $t = T$. Split that period up into M pieces such that $\Delta t = \frac{T}{M}$. And now let $t_m = m\Delta t$ for $m = 0, ..., M$. At each t_m, the asset price $S_m = S(t_m)$ may move up by a factor of u or down by a factor of d. Take

$$S_{m+1} = \begin{cases} uS_m & \text{with probability } q \\ dS_m & \text{with probability } 1 - q \end{cases} \tag{1.6}$$

to be the formula for the price at S_{m+1} given S_m. The model is arbitrage free if the risk-neutral probability measure $\mathbb{Q}$ exists such that the return of the risky asset is the same as the return of a risk-less asset at risk-free interest rate r. So,

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{S_{m+1} - S_m}{S_m} \;\middle|\; S_m\right] = e^{r\Delta t}. \tag{1.7}$$

Since $\mathbb{E}_{\mathbb{Q}}\left[\frac{S_{m+1}-S_m}{S_m} \;\middle|\; S_m\right] = \frac{\mathbb{E}_{\mathbb{Q}}[S_{m+1} \mid S_m]-S_m}{S_m}$ and

$$\mathbb{E}_{\mathbb{Q}}\left[S_{m+1} \mid S_m\right] = (qu + (1 - q)d)S_m, \tag{1.8}$$

it follows that

$$qu + (1 - q)d = e^{r\Delta t}. \tag{1.9}$$

From this q may be written as

$$q = \frac{e^{r\Delta t} - d}{u - d}.$$

In order to guarantee that the variance of the rate of the return of the binomial model corresponds to the actual, historical, rate of return of the risky asset,

$$Var_{\mathbb{Q}}\left[\frac{S_{m+1}-S_m}{S_m} \;\middle|\; S_m\right] = \sigma^2\Delta t + \mathcal{O}(\Delta t) \tag{1.10}$$

must hold. Here, $\mathcal{O}(\Delta t)$ means that as $\Delta t \to 0$ then

$$\frac{\mathcal{O}(\Delta t)}{\Delta t} \to 0.$$

Using the property $Var(x) = \mathbb{E}x^2 - (\mathbb{E}x)^2$ and equations (1.8) and (1.9) gives

$$Var_{\mathbb{Q}}\left[\frac{S_{m+1}-S_m}{S_m} \;\middle|\; S_m\right] = \frac{1}{S_m^2} Var_{\mathbb{Q}}\left[S_{m+1} \mid S_m\right] \tag{1.11}$$

and

$$Var_{\mathbb{Q}}\left[S_{m+1} \mid S_m\right] = qu^2S_m^2 + (1-q)d^2S_m^2 - (quS_m + (1-q)dS_m)^2.$$

From these two equations and equation (1.10),

$$qu^2 + (1-q)d^2 = e^{2r\Delta t} + \sigma^2\Delta t + \mathcal{O}(\Delta t) \tag{1.12}$$

is obtained for determining u and d. Since there is only one equation with two unknowns, u and d are not uniquely determined. However, one possibility is to take $d = 1/u$. Using this additional condition it is possible to show that equation (1.12) is satisfied by

$$u = e^{\sigma\sqrt{\Delta t}}.$$

This model makes it possible to find the value of the option with expiration T at time $t = 0$. At time T, the price has $M+1$ different values $S_{M,j} = S_0u^jd^{M-j}$, $j = 0, 1, ..., M$. Since the price at time T is equal to the payoff,

$$V(S_{M,j}, t_M) = P(S_{M,j}, t_M)$$

may be written. Now, the option price may be computed recursively for times $t_{M-1}, t_{M-2}, ...t_0$, and

$$V(S_{m,j}, t_m) = e^{-r\Delta t}\left(qV(S_{m+1,j+1}, t_{m+1}) + (1-q)V(S_{m+1,j}, t_{m+1})\right) \tag{1.13}$$

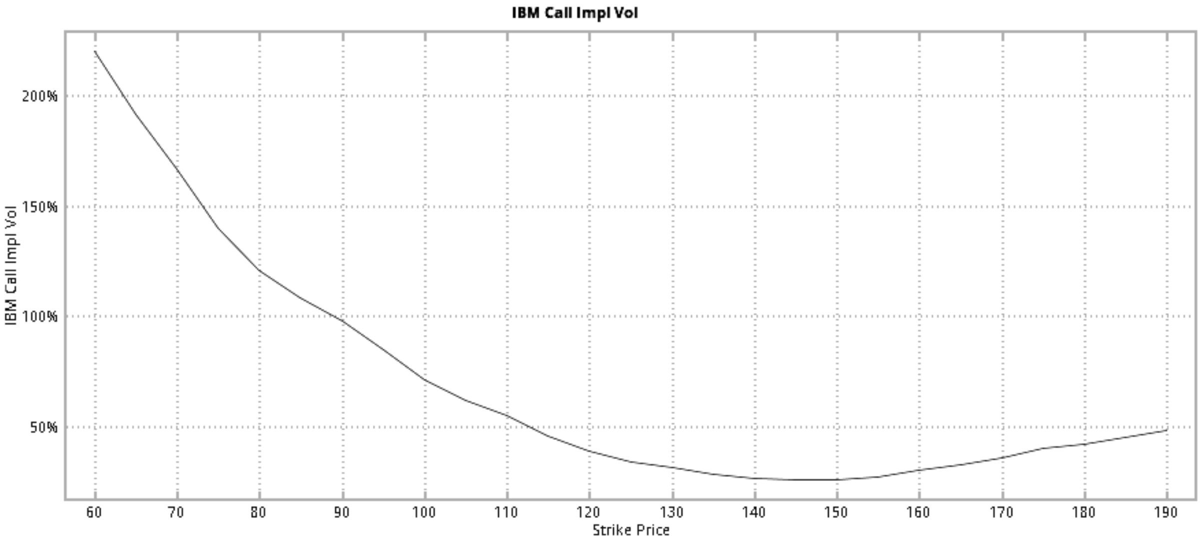

Figure 1.1: The Implied Volatility for IBM Call Options Expiring 17Jun2022 on 16May2022 Courtesy of Thinkorswim® Desktop

holds in the case of European options. The price of the option at time $t = 0$ is $V(0) = V(S_0, 0)$.

In the case of American options, for the same recursive time steps find $W(S_{m,j}, t_m) = e^{-r\Delta t}\left(qV(S_{m+1,j+1}, t_{m+1}) + (1-q)V(S_{m+1,j}, t_{m+1})\right)$ for each node which corresponds to holding this option until t_{m+1} as in equation (1.13). However, for American options, the ability to exercise early must also be checked and therefore,

$$V(S_{m,j}, t_m) = \max\{W(S_{m,j}, t_m), P(S_{m,j}, t_m)\}.$$

And again, the price of the option at time $t = 0$ is $V(0) = V(S_0, 0)$.

1.6 Non-Constant Volatility and the Volatility Smile

The volatility smile is characterized by a skew in the implied volatility of options across a range of strikes and is pictured in figure 1.1. The future volatility is not a known value and therefore must be estimated. The Black-Scholes formula may be inverted using numerical methods to calculate implied volatility which is often denoted as σ_{imp}, and so, the **implied volatility** is the estimate of future volatility implied by current option prices. The book, *The Volatility Smile*, by Derman and Miller (2016) provides an excellent reference to this concept, and it is the main reference for this section. Up until this skew was noticed, the

Black-Scholes model provided a great way to value options as it provides an arbitrage free method to find an option's value with a closed form solution. Prior to the 1987 stock market crash, the implied volatility of the market agreed. However, after the crash, it was noticed that a U-shape was forming in the implied volatility graph, the volatility smile. This was because the market in 1987 saw the first intra-day drop of $>20\%$ in market prices since 1929. Meaning, it was the first major drop in the stock market since Black, Scholes, and Merton published their papers in 1973, and so, this was the first test of their models under dire market conditions. This crash showed that the constant volatility assumption of the Black-Scholes model was insufficient as market volatility is asymmetric, that is large downward swings are much more likely than upward ones. From this, firms realized that low-strike puts should cost more than high-strike calls due to their higher probability. This new insight was reflected in the formation of the volatility smile. There are several ways to attempt to model this behavior, one is using heuristics and the experience of the trader. This is the least quantitative of the methods. The next is to attempt to model the changes to the implied volatility surface, rather than the underlying asset. However, the method focused on in this thesis is to find a more complicated model which contains more than a geometric Brownian motion to generate the price path of the underlying. One class of models of this type are known as local volatility models. One that is well known and studied is the constant elasticity of variance (CEV) model. These models consider volatility to be a deterministic function on stock price, and retain much of the simplicity of the Black-Scholes model. The CEV model will be discussed in the next section. The next class of models are stochastic volatility models, the basis of this thesis. These model volatility as a separate stochastic process from price, although the movement of the two may be correlated. While there are other classes of models, such as jump-diffusion models, they are left out for brevity. The main idea behind improved model accuracy is not to determine whether or not the market price for high liquidity American or European style options is correct and attempting to arbitrage against them, but rather for determining the price of exotic options with low liquidity. Unlike high liquidity standard options these lower liquidity options may not have a set market price, so it's crucial for firms to be able to accurately price them. Another need for models which better capture market dynamics is to improve the accuracy of hedging. Hedging is the process of taking offsetting positions to reduce the market risk of the main position, in this case selling or purchasing the underlying stock to reduce the market risk of the option position.

1.7 The Constant Elasticity of Variance Model

For non-constant volatility, a slight modification to equation (1.3) must be made to obtain

$$dS_t = \mu S_t dt + \sigma(S_t, t) S_t dW_t. \tag{1.14}$$

An example of such an equation is the CEV model was proposed by Cox in 1975 and further explored by Beckers (1980). The model has $\sigma(S_t, t) = \sigma S^{\beta/2-1}$ giving

$$dS_t = \mu S_t dt + \sigma S_t^{\beta/2} dW_t. \tag{1.15}$$

The name comes from the fact that the elasticity of the volatility of the price $\epsilon = \frac{d\sigma^2/\sigma^2}{dS/S}$ is constant and proportional to the elasticity parameter of the volatility β.

1.8 Stochastic Volatility Models

While non-constant volatility models do a good job of capturing the volatility smile, they miss out on the fact that stock price tends to be bursty. That is, it has periods of a relatively constant volatility followed by periods of higher or lower volatility. While stochastic volatility models have more parameters to estimate than a non-constant volatility model, the added parameters allow them to capture added complexity of this behavior.

1.9 Mean-Reverting Stochastic Volatility Models

For mean-reverting stochastic volatility models, the paper by Fouque, Papanicolaou, and Sircar (2000) is used. The main features of this class of stochastic model are that volatility is positive and mean-reverting, and volatility shocks are negatively correlated with asset price shocks. For this class of models a mean-reverting Ornstein-Uhlenbeck (OU) process is used to model stock prices, and is given by

$$dS_t = \mu S_t dt + \sigma(Y_t) S_t dW_t,$$
$$dY_t = \alpha(\nu - Y_t)dt + \psi(Y_t)d\hat{Z}_t,$$
$$\hat{Z}_t := \rho W_t + \sqrt{1-\rho^2} Z_t.$$

The process Y_t has several parameters, ν is the mean-level of Y_t, α is the rate of mean-reversion of this process, the function $\psi(Y_t)$ is the volatility of Y_t, and ρ is the correlation between the price and volatility shocks. The functions $\sigma(Y_t)$ and $\psi(Y_t)$ may be chosen in a way that suits the specific problem. Additionally, it is common to let $\psi(Y_t) = \beta\sqrt{Y_t}$.

1.10 Heston Model

One specific type of mean-reverting stochastic volatility model explored in this thesis is the Heston model. This volatility process is a special case of the volatility of the OU process, and is given by

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{Y_t} S_t dW_t, \\ dY_t &= \alpha\left(\nu - Y_t\right) dt + \beta\sqrt{Y_t} dZ_t, \end{aligned} \tag{1.16}$$

where the equivalents of the OU process are $\sigma(Y_t) = \sqrt{Y_t}$, and $\psi(Y_t) = \beta\sqrt{Y_t}$. In this model the Brownian motions W and Z may be correlated by the correlation coefficient ρ.

2 Quadrinomial Recombining Tree

This chapter is based on the work of Florescu and Viens (2008). They considered a mean-reverting stochastic volatility model based on the OU process

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma(Y_t) S_t dW_t, \\ dY_t &= \alpha(\nu - Y_t)dt + \psi(Y_t)dZ_t \end{aligned} \tag{2.1}$$

to model price and volatility, and using the parameters they estimated. In this particular case, the Brownian motions W and Z are uncorrelated. For the log-price $X_t = ln(S_t)$,

$$dX_t = \left(\mu - \frac{\sigma^2(Y_t)}{2}\right) dt + \sigma(Y_t)dW_t \tag{2.2}$$

holds. The following algorithm is used to model stochastic volatility and estimate the volatility distribution of the actual volatility process given a set of historical observations. This portion of the method happens in two steps. First, a mutation step, where paths of the model are simulated the objective, or market, probability measure $\mathbf{P}$ and discretized versions of equations (2.1) and (2.2). And second, from those simulated paths, it assigns probabilities of occurring to each of the paths using their relative closeness to the realized historical price.

Using the output of the density function, it then samples values of Y to use in a quadrinomial recombining tree to model the price process. At each time t, four successors to the current price of the stock are calculated, from each of these four successors, and so on. Each of these nodes is assigned a probability and so at the end, expected present value can be calculated. The tree is recombining by construction in order to limit the total number of nodes and calculations. The Monte Carlo method calculates the price of the option using N quadrinomial trees.

2.1 Estimating the Filtered Stochastic Volatility Distribution

For this section, it is assumed that the coefficients μ, ν, and α and the functions $\sigma(y)$ and $\psi(y)$ are already known or have been estimated. Additionally, the functions $\sigma(y)$ and

$\psi(y)$ must twice differentiable with bounded derivatives of all orders up to two. Here, the formulas are taken from those in the paper by Florescu and Viens (2008), and so, $\sigma(y) = e^{-|y|}$ and $\psi(y) = \beta$.

To estimate the density function of the stochastic volatility process, assume the historical log-prices $x_{t_1}, ..., x_{t_K}$ are known. From these, the probabilities are estimated as

$$p_k\left(dy\right) = \mathbf{P}\left[Y_{t_k} \in dy | X_{t_1} = x_{t_1}, ..., X_{t_k} = x_{t_k}\right]$$

for $k = 1, 2, ..., K$ where $\mathbf{P}$ is the objective probability measure. This is a filtered stochastic volatility process. The distribution estimation is done by a genetic-type algorithm with a two-step iteration: a mutation step and a selection step. For the historical data, let $h = t_{k+1} - t_k$.

Step 1: Let $X_{t_0} = x_0$ and $Y_{t_0} = \nu$, where x_0 is the initial historical stock price at t_0 and ν is the long run average variance of the volatility process Y_t. Let n be the number of paths being simulated and M be the number of steps simulated in each path between any two times t and $t + h$.

Mutation Step: Each mutation step generates an (X'_{t_1}, Y'_{t_1}) pair performed by iterating through discretizations of the OU process in (2.1) and (2.2),

$$\begin{aligned} dX_t &= \left(r - \frac{\sigma^2(Y_t)}{2}\right) dt + \sigma(Y_t) dW_t, \\ dY_t &= \alpha(\nu - Y_t) dt + \psi(Y_t) dZ_t. \end{aligned} \tag{2.3}$$

The discretized functions are obtained using the Euler Method, which is a common and straightforward way to do this. Taking the general equation

$$dX_t = a(X_t)dt + b(X_t)dW_t,$$

the Euler Method's time-discretized approximation with M steps between times t and $t + h$ is

$$X_{i+1} = X_i + a(X_i)\frac{h}{M} + b(X_i)\sqrt{\frac{h}{M}}U_{i+1}$$

where a and b are functions of X_m and $i = 0, ..., M-1$ and $U_{i+1} \sim N(0,1)$. The error for this method of the drift term is of the order $\mathcal{O}(\frac{h}{M})$ and the error of the diffusion term is of the order $\mathcal{O}(\sqrt{\frac{h}{M}})$ (Glasserman, 2003). Starting with $X'_{t_0,0} = x_0$ and $Y'_{t_0,0} = \nu$, applying the Euler method to equation (2.3) yields

$$\begin{aligned} X'_{t_0,i+1} &= X'_{t_0,i} + \frac{h}{M}\left(\mu - \frac{\sigma^2(Y'_{t_0,i})}{2}\right) + \sqrt{\frac{h}{M}}\sigma(Y'_{t_0,i})U_i, \\ Y'_{t_0,i+1} &= Y'_{t_0,i} + \frac{h}{M}\alpha(\nu - Y'_{t_0,i}) + \sqrt{\frac{h}{M}}\psi(Y'_{t_0,i})U'_i. \end{aligned} \tag{2.4}$$

Here, U and U' are iid standard normally distributed random numbers. Let $X'_{t_1} := X'_{t_0,M}$ and $Y'_{t_1} := Y'_{t_0,M}$. The mutation step is performed n times to obtain, $\{(X'^j_{t_1}, Y'^j_{t_1})\}$, $j = 1, ..., n$.

Selection Step: Start by letting ϕ_n be a function in $L^1(R)$. Here, the same formula as given by Florescu and Viens (2008) is used,

$$\phi_n(x) = \begin{cases} \sqrt[3]{n}(1 - |x\sqrt[3]{n}|) & if \ -\frac{1}{\sqrt[3]{n}} < x < \frac{1}{\sqrt[3]{n}} \\ 0 & otherwise. \end{cases}$$

Each selection step creates a discrete probability measure for the set of $Y'^j_{t_1}$, $j \in \{1, ..., n\}$ by using the following procedure. Let $C = \sum_{j=1}^{n} \phi_n(X'^j_{t_1} - x_1)$. If $C > 0$ then

$$\Phi_1^n(Y) = \begin{cases} \frac{1}{C}\sum_{j=1}^{n} \phi_n(X'^j_{t_1} - x_1) & if \ Y = Y'^j_{t_1} \\ 0 & otherwise. \end{cases} \tag{2.5}$$

The value of C is chosen such that Φ_1^n is a discrete probability measure for Y'_{t_1} giving an estimation of the distribution of Y_{t_1} from which values $\{Y^j_{t_1}\}_{j=\overline{1,n}}$ can be sampled. If $C = 0$, then none of the values of $X'^j_{t_1}$ are sufficiently close to x_1 to occur with a positive probability. Therefore, the value of n will need to be reduced to increase the distance between $X'^j_{t_1}$ and x_1 that gives a positive probability.

Steps 2 to K: For $k = 2, 3, ..., K$ and starting with the distribution Φ^n_{k-1} generated at the end of the last selection step, generate n values of $Y^j_{t_{k-1}}$, $j = 1, ..., n$. Using the ordered pair $\left(X_{t_{k-1}}, Y^j_{t_{k-1}}\right)$ apply formula (2.4) to obtain the Euler method estimation of n pairs,

$\{(X'^j_{t_k}, Y'^j_{t_k})\}_{j=1,...,n}$. Using these pairs, apply the selection step to obtain the estimated probability distribution for Y_{t_k},

$$\Phi^n_k(Y) = \begin{cases} \frac{1}{C}\sum_{j=1}^n \phi_n(X'^j_{t_k} - x_k) & if\ Y = Y'^j_{t_k} \\ 0 & otherwise, \end{cases} \tag{2.6}$$

and again, $C = \sum_{j=1}^n \phi_n\left(X'^j_{t_i} - x_i\right)$. As in $k = 1$, Φ^n_k is a probability measure for Y'_{t_k} giving an estimation of the distribution of Y_{t_k} from which values $\{Y^j_{t_k}\}_{j=\overline{1,n}}$ can be sampled.

Output: The output of the algorithm is the last estimated discrete probability distribution Φ^n_K. This probability distribution is the estimate of $p_k(dy)$, which is the distribution of the process Y_t given the historical log-prices $x_{t_1}, ..., x_{t_K}$. In order to create the quadrinomial tree, only Φ^n_K is used. So, given equation (2.6),

$$\hat{p}_j = \frac{1}{C}\phi_n(X'^j_{t_K} - x_{t_K}) \tag{2.7}$$

and

$$\hat{Y}_j = Y'^j_{t_K} \tag{2.8}$$

are defined.

2.2 Option Valuation With a Quadrinomial Tree

To create the quadrinomial tree, first split the time interval $[0, T]$ into N subintervals, such that $\Delta t = \frac{T}{N}$. Then, using the final filtered stochastic volatility distribution Φ^n_K with the properties defined in (2.7) and (2.8) from the output of Section 2.1, for each time period $i\Delta t$, sample values of $Y_i,\ i = 1, ..., N$. Thus, in each time step $i\Delta t$ the value Y_i is used to construct the successors of each node. Begin with price $x = ln(S_0)$ as the base node of the tree. For this section, it is assumed that appropriate value of the short-term interest rate r and the desired exercise price of the option E have been chosen. Starting with x, find its four successors x_1, x_2, x_3, and x_4 as depicted in figure 2.1.

In order to construct the successors, consider a grid of points of the form $l\sigma(Y_i)\sqrt{\Delta t}$ were l is an integer value. This means that the value of the current node x will fall somewhere

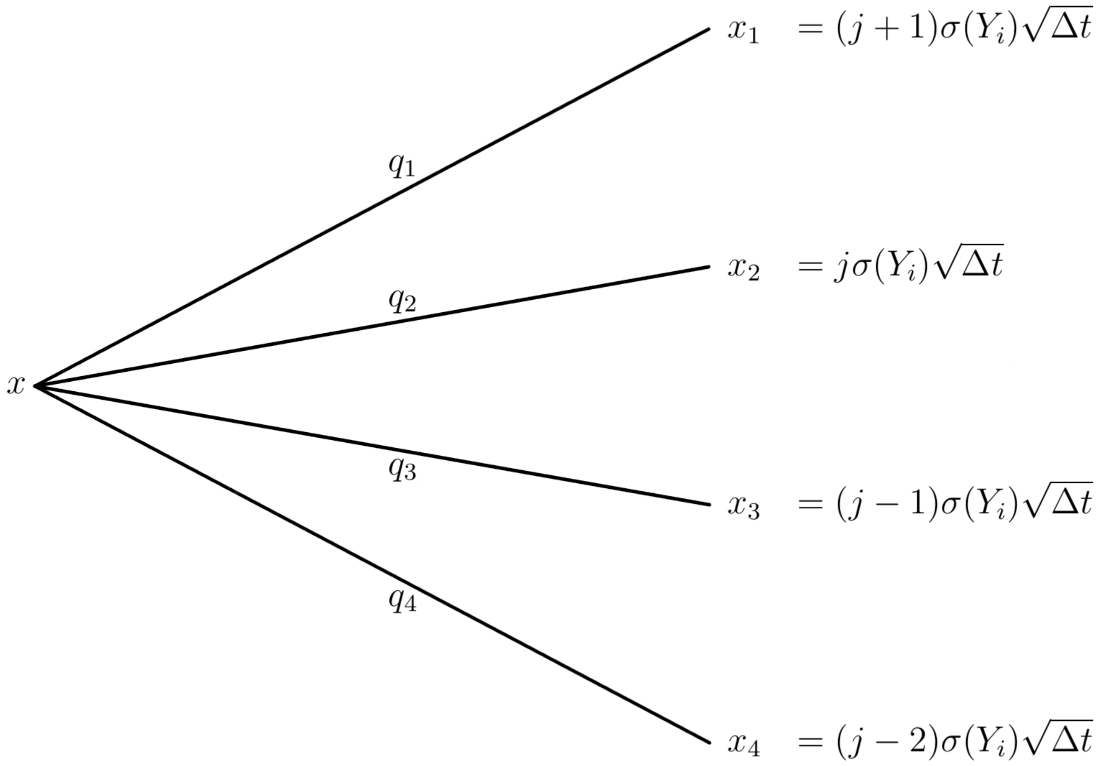

Figure 2.1: The successors of each point an their risk-neutral probabilities

in this grid. Since the values of l limits the number of values the successors may take, this will also make the tree recombining. In order to guarantee that x falls between the successors x_2 and x_3, let j be the smallest value which makes $x_2 > x$, that is

$$j = inf\{l \in \mathbb{N} \mid l\sigma(Y_i)\sqrt{\Delta t} \geq x\}.$$

The value of j may be calculated using the inequality $x \leq j\sigma(Y_i)\sqrt{\Delta t}$. Since $\sigma(Y_i)$ and $\sqrt{\Delta t}$ are known, the smallest $j \in \mathbb{N}$ that satisfies this inequality is found using the ceiling function,

$$j = \left\lceil \frac{x}{\sigma(Y_i)\sqrt{\Delta t}} \right\rceil.$$

Now, let δ be the distance from the node x and its closest successor on the grid and b be its standardized value. Therefore, if $j\sigma(Y_i)\sqrt{\Delta t}$ is closer to x then $\delta = x - j\sigma(Y_i)\sqrt{\Delta t}$. But if $(j-1)\sigma(Y_i)\sqrt{\Delta t}$ is closer, then $\delta = x - (j-1)\sigma(Y_i)\sqrt{\Delta t}$. Now, let $b = \delta / \left(\sigma\left(Y_i\right)\sqrt{\Delta t}\right)$. From here, the algorithm may take one of two paths, the first being if x is closer to $j\sigma(Y_i)\sqrt{\Delta t}$ and the second if x is closer to $(j-1)\sigma(Y_i)\sqrt{\Delta t}$.

If x_2 is closer: That is $x_2 - x \leq x - x_3$, then we define $\delta = x - j\sigma(Y_i)\sqrt{\Delta t}$. In order to solve for the range of values of δ, start with the maximum. By definition $x \leq x_2$ and the definition δ, it is clear that $\delta \leq 0$. For the minimum value start with the inequality $x_2 - x \leq x - x_3$ and expand the terms x_2 and x_3 to get

$$x \geq \frac{j\sigma(Y_i)\Delta t + (j-1)\sigma(Y_i)\Delta t}{2}.$$

With a small amount of algebra

$$x \geq j\sigma(Y_i)\Delta t - \frac{\sigma(Y_i)\Delta t}{2}$$

may be reached. Now using the definition of δ,

$$\delta \geq -\frac{\sigma(Y_i)\Delta t}{2}$$

is obtained. So, $\delta \in \left[-\frac{\sigma(Y_i)\Delta t}{2}, 0\right]$. Substituting the endpoint values of δ into the equation for b directly yields $b \in \left[-\frac{1}{2}, 0\right]$. The formulation of the tree given in figure 2.1 provides a good starting point, but it doesn't account for the drift of the process. So, in order to obtain the convergence of the mean of the increment to the drift of the process X_t, add the drift quantity to each of the successors giving the set of equations

$$\begin{cases} x_1 &= (j+1)\sigma(Y_i)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_2 &= j\sigma(Y_i)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_3 &= (j-1)\sigma(Y_i)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_4 &= (j-2)\sigma(Y_i)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t. \end{cases} \tag{2.9}$$

Now that values have been obtained for each of the successors x_1, x_2, x_3, and x_4, their risk-neutral probabilities q_1, q_2, q_3, and q_4 must be calculated. These probabilities have the properties

$$0 \leq q_i \leq 1, i = 1, 2, 3, 4 \tag{2.10}$$

and

$$\sum_{i=1}^{4} q_i = 1. \tag{2.11}$$

Notice that $\delta = x - j\sigma(Y_i)\sqrt{\Delta t}$ and therefore $j\sigma(Y_i)\sqrt{\Delta t} = x - \delta$ so the set of equations in (2.9) can be rewritten as

$$\begin{cases} x_1 - x & = \sigma(Y_i)\sqrt{\Delta t} - \delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_2 - x & = -\delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_3 - x & = -\sigma(Y_i)\sqrt{\Delta t} - \delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_4 - x & = -2\sigma(Y_i)\sqrt{\Delta t} - \delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t. \end{cases}$$

According to (2.3), define $\Delta x = x_i - x,\ i = 1, 2, 3, 4$ with expected value

$$\mathbb{E}_{\mathbb{Q}}\left[\Delta x | Y_i\right] = \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t. \tag{2.12}$$

From this, and given

$$\Delta x' = \Delta x - \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t,$$

it follows that $\mathbb{E}_{\mathbb{Q}}\left[\Delta x' \mid Y_i\right] = 0$. Therefore terms of the form

$$\left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t,$$

can be ignored. Now, expanding the expected value of $\Delta x'$ to include the probabilities q_1, q_2, q_3, and q_4 yields

$$\left(\sigma(Y_i)\sqrt{\Delta t} - \delta\right) q_1 + (-\delta)q_2 + \left(-\sigma(Y_i)\sqrt{\Delta t} - \delta\right) q_3 + \left(-2\sigma(Y_i)\sqrt{\Delta t} - \delta\right) q_4 = 0. \tag{2.13}$$

This simplifies to

$$\sigma(Y_i)\sqrt{\Delta t}(q_1 - q_3 - 2q_4) - \delta(q_1 + q_2 + q_3 + q_4) = 0.$$

Which, by utilizing equation (2.11), the above may be rewritten as

$$\sigma(Y_i)\sqrt{\Delta t}\,(q_1 - q_3 - 2q_4) - \delta = 0, \tag{2.14}$$

or, with the q terms isolated on the left hand side

$$q_1 - q_3 - 2q_4 = \frac{\delta}{\sigma(Y_i)\sqrt{\Delta t}}. \tag{2.15}$$

Next, switching focus to the variance of Δx and $\Delta x'$,

$$Var_{\mathbb{Q}}\left[\Delta x \mid Y_i\right] = Var_{\mathbb{Q}}\left[\Delta x' \mid Y_i\right] = \sigma^2(Y_i)\Delta t.$$

Utilizing the property $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$,

$$\begin{aligned}\left(\sigma(Y_i)\sqrt{\Delta t} - \delta\right)^2 q_1 + (-\delta)^2 q_2 + \left(-\sigma(Y_i)\sqrt{\Delta t} - \delta\right)^2 q_3 + \left(-2\sigma(Y_i)\sqrt{\Delta t} - \delta\right)^2 q_4 \\ -\mathbb{E}\left[\Delta x' | Y_i\right]^2 = \sigma^2(Y_i)\Delta t\end{aligned} \tag{2.16}$$

is obtained. Substituting for $\mathbb{E}\left[\Delta x' \mid Y_i\right]$ using (2.14) yields

$$\begin{aligned}\sigma^2(Y_i)\Delta t = (\sigma^2(Y_i)\Delta t)(q_1 + q_3 + 4q_4) + (2\delta\sigma(Y_i)\sqrt{\Delta t})(-q_1 + q_3 + 2q_4) + \\ \delta^2(q_1 + q_3 + q_4) + \delta^2 q_2 - \left[\sigma(Y_i)\sqrt{\Delta t}(q_1 - q_3 - 2q_4) - \delta\right]^2.\end{aligned}$$

By expanding the last term of the above equation and utilizing (2.11), it becomes:

$$\begin{aligned}\sigma^2(Y_i)\Delta t = (\sigma^2(Y_i)\Delta t)(q_1 + q_3 + 4q_4) + (2\delta\sigma(Y_i)\sqrt{\Delta t})(-q_1 + q_3 + 2q_4) + \\ \delta^2 - \left[\sigma^2(Y_i)\Delta t(q_1 - q_3 - 2q_4)^2 + (2\delta\sigma(Y_i)\sqrt{\Delta t})(-q_1 + q_3 + 2q_4) + \delta^2\right].\end{aligned}$$

Canceling like terms yields

$$\sigma^2(Y_i)\Delta t(q_1 + q_3 + 4q_4) - \sigma^2(Y_i)\Delta t(q_1 - q_3 - 2q_4)^2 = \sigma^2(Y_i)\Delta t.$$

Continue by dividing by $\sigma^2(Y_i)\Delta t$ to obtain

$$(q_1 + q_3 + 4q_4) - (q_1 - q_3 - 2q_4)^2 = 1.$$

Substituting (2.15) in for the second term on the left hand side, this simplifies to

$$q_1 + q_2 + 4q_3 = 1 + \frac{\delta^2}{\sigma^2(Y_i)\Delta t}. \tag{2.17}$$

Taking into consideration equations (2.10), (2.11), (2.15), and (2.17) the following system of equations for determining the risk-neutral probabilities appears as

$$\begin{cases} q_1 + q_2 + q_3 + q_4 = 1 \\ q_1 - q_3 - 2q_4 = \frac{\delta}{\sigma(Y_i)\sqrt{\Delta t}} = b \\ q_1 + q_3 + 4q_4 = 1 + \frac{\delta^2}{\sigma^2(Y_i)\Delta t} = 1 + b^2. \end{cases}$$

This system of equations has three equations and four unknowns. Therefore, it has infinitely many solutions and the risk-neutral probabilities are not unique. Let $q := q_4$ be a fixed value and solve for q_1, q_2, and q_3 with respect to q and b. This gives:

$$\begin{cases} q_1 + q_2 + q_3 = 1 - q \\ q_1 - q_3 = b + 2q \\ q_1 + q_3 = 1 + b^2 - 4q. \end{cases}$$

Adding together the second and third equations gives

$$q_1 = \frac{1}{2}(1 + b + b^2) - q,$$

and by subtracting them

$$q_3 = \frac{1}{2}(1 - b + b^2) - 3q$$

is obtained. Rearranging the first equation in the system gives

$$q_2 = 1 - q - q_1 - q_3,$$

into which the values of q_1 and q_3 may be substituted which yields

$$q_2 = 1 - q - \left(\frac{1}{2}(1 + b + b^2) - q\right) - \left(\frac{1}{2}(1 - b + b^2) - 3q\right).$$

By simplifying this expression,

$$q_2 = 3q - b^2$$

is obtained. Taking together the results of q_1, q_2, and q_3 gives:

$$\begin{cases} q_1 = \frac{1}{2}(1+b+b^2) - q \\ q_2 = 3q - b^2 \\ q_3 = \frac{1}{2}(1-b+b^2) - 3q. \end{cases} \tag{2.18}$$

All that remains is to determine the valid values of q which satisfy (2.10), (2.11), and (2.18). It is known that q, q_1, q_2, and q_3 must be between 0 and 1, so for the first equation of (2.18), begin with the inequality

$$0 \leq \frac{1}{2}(1+b+b^2) - q \leq 1.$$

Next, isolate q to obtain

$$-\frac{1}{2}(1-b-b^2) \leq q \leq \frac{1}{2}(1+b+b^2).$$

To find the maximum and minimum over the range $b \in \left[-\frac{1}{2}, 0\right]$, start with $\frac{dq}{db} = \frac{1}{2} + b$ and $\frac{d^2q}{db^2} = 1$. Therefore, both functions have an absolute minimum at $b = -\frac{1}{2}$. For $b > -\frac{1}{2}$, $\frac{dq}{db} > 0$, therefore the function has a strictly positive slope and is strictly increasing over the range of $b \in \left(-\frac{1}{2}, 0\right]$. Therefore, the maximum over this range will be at $b = 0$ and the maximum and minimum points in the range may be taken from the endpoints. So, given $b \in \left[-\frac{1}{2}, 0\right]$, with $b = 0$ gives $-\frac{1}{2} \leq q \leq \frac{1}{2}$ and for $b = -\frac{1}{2}$, $-\frac{5}{8} \leq q \leq \frac{3}{8}$. Taking the most limiting conditions from these inequalities and (2.10) gives $0 \leq q \leq \frac{3}{8}$. Now, taking the second equation of (2.18) the inequality,

$$0 \leq 3q - b^2 \leq 1$$

may be built. And, a quick simplification yields

$$\frac{b^2}{3} \leq q \leq \frac{1+b^2}{3}.$$

Using similar reasoning as before, $\frac{dq}{db} = \frac{2}{3}b$ for both equations. From this it is easy to see that the absolute minimum of this function occurs at $b = 0$. Using this minimum and $\frac{dq}{db}$, it follows that the function is strictly decreasing over the range of $b \in \left[-\frac{1}{2}, 0\right)$. Therefore over this range, the maximum occurs at $b = -\frac{1}{2}$. Solving for $b = 0$ gives $0 \leq q \leq \frac{1}{3}$, and $b = -\frac{1}{2}$ yields $\frac{1}{12} \leq q \leq \frac{5}{12}$. Taking these inequalities with the previously obtained $0 \leq q \leq \frac{3}{8}$ gives $\frac{1}{12} \leq q \leq \frac{3}{8}$. Finally, take the third equation of (2.18) and it is clear that

$$0 \leq \frac{1}{2}(1 - b + b^2) - 3q \leq 1.$$

With some algebra, this yields

$$\frac{1}{6}(-1 - b + b^2) \leq q \leq \frac{1}{6}(1 - b + b^2).$$

From $\frac{dq}{db} = \frac{1}{3}b - \frac{1}{6}$, the absolute minimum of the function is at $b = \frac{1}{2}$. Since $\frac{dq}{db} < 0$ for $b \in \left[-\frac{1}{2}, \frac{1}{2}\right)$, it is clear that the function is strictly decreasing over the range $b \in \left[-\frac{1}{2}, 0\right]$, and so, the maximum and minimum values occur at the endpoints, $b = -\frac{1}{2}$ and $b = 0$ respectively. Solving for $b = 0$ gives $-\frac{1}{6} \leq q \leq \frac{1}{6}$, and $b = -\frac{1}{2}$ gives $-\frac{1}{24} \leq q \leq \frac{7}{24}$. From these and $\frac{1}{12} \leq q \leq \frac{3}{8}$, it's easy to see that $q \in \left[\frac{1}{12}, \frac{1}{6}\right]$.

If x_3 is closer: That is $x_2 - x > x - x_3$, define $\delta = x - (j-1)\sigma(Y_i)\sqrt{\Delta t}$. In order to solve for the bounds of δ, start with similar logic as in the previous section, it is easy to see that $\delta \geq 0$. For the maximum value start with the inequality $x_2 - x \geq x - x_3$, isolate x and expand the terms x_2 and x_3 to get

$$x \leq \frac{[(j-1)+1]\sigma(Y_i)\Delta t + (j-1)\sigma(Y_i)\Delta t}{2}.$$

Now, move $x - (j-1)\sigma(Y_i)\Delta t$ to the left hand side of the equation and

$$x - (j-1)\sigma(Y_i)\Delta t \leq \frac{\sigma(Y_i)\Delta t}{2}$$

is obtained. Using the definition of δ, this may be rewritten as

$$\delta \leq \frac{\sigma(Y_i)\Delta t}{2}.$$

So, $\delta \in \left[0, \frac{\sigma(Y_i)\Delta t}{2}\right]$. Substituting δ into the equation for b directly yields $b \in \left[0, \frac{1}{2}\right]$. As before, in order to simplify the convergence of the mean of the increment to the drift of the process X_t, add the drift quantity to each of the successors giving equation (2.9). In order to calculate the probabilities of the successors notice that $\delta = x - (j-1)\sigma(Y_i)\sqrt{\Delta t}$ and therefore $(j-1)\sigma(Y_i)\sqrt{\Delta t} = x - \delta$ so the set of equations in (2.9) can be rewritten as

$$\begin{cases} x_1 - x &= 2\sigma(Y_i)\sqrt{\Delta t} - \delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_2 - x &= \sigma(Y_i)\sqrt{\Delta t} - \delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_3 - x &= -\delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t \\ x_4 - x &= -\sigma(Y_i)\sqrt{\Delta t} - \delta + \left(r - \frac{\sigma^2(Y_i)}{2}\right)\Delta t. \end{cases} \tag{2.19}$$

Using equation (2.12), similarly to (2.13),

$$\left(2\sigma(Y_i)\sqrt{\Delta t} - \delta\right) q_1 + (\sigma(Y_i)\sqrt{\Delta t} - \delta)q_2 + (-\delta)\, q_3 + \left(-\sigma(Y_i)\sqrt{\Delta t} - \delta\right) q_4 = 0$$

may be written. This simplifies to

$$\sigma(Y_i)\sqrt{\Delta t}(2q_1 + q_2 - q_4) - \delta(q_1 + q_2 + q_3 + q_4) = 0,$$

which, by the property $q_1 + q_2 + q_3 + q_4 = 1$ and with some rearrangement of terms, can be written as

$$2q_1 + q_2 - q_4 = \frac{\delta}{\sigma(Y_i)\sqrt{\Delta t}}. \tag{2.20}$$

Now for variance similarly to (2.16), but using (2.19),

$$\left(2\sigma(Y_i)\sqrt{\Delta t} - \delta\right)^2 q_1 + (\sigma(Y_i)\sqrt{\Delta t} - \delta)^2 q_2 + (-\delta)^2 q_3 + \left(-\sigma(Y_i)\sqrt{\Delta t} - \delta\right)^2 q_4$$
$$-\mathbb{E}\left[\Delta x' | Y_i\right]^2 = \sigma^2(Y_i)\Delta t$$

is written. By substituting for and expanding the $\mathbb{E}_{\mathbb{Q}}\left[\Delta x'|Y_i\right]^2$ term, this becomes

$$\sigma^2(Y_i)\Delta t = (\sigma^2(Y_i)\Delta t)(4q_1 + q_2 + q_4) - (2\delta\sigma(Y_i)\sqrt{\Delta t})(2q_1 + q_2 - q_4) +$$
$$\delta^2 - \left[\sigma^2(Y_i)\Delta t(2q_1 + q_2 - q_4)^2 - (2\delta\sigma(Y_i)\sqrt{\Delta t})(2q_1 + q_2 - q_4) + \delta^2\right].$$

Performing the same cancellations as in the previous section yields

$$(4q_1 + q_2 + q_4) - (2q_1 + q_2 - q_4)^2 = 1.$$

After some algebra and substituting in δ,

$$4q_1 + q_2 + q_4 = 1 + \frac{\delta^2}{\sigma^2(Y_i)\Delta t} \tag{2.21}$$

is reached. Taking the definition of b and equations (2.20) and (2.21) the following system of equations appears as

$$\begin{cases} q_1 + q_2 + q_3 + q_4 = 1 \\ 2q_1 + q_2 - q_4 = \frac{\delta}{\sigma(Y_i)\sqrt{\Delta t}} = b \\ 4q_1 + q_2 + q_4 = 1 + \frac{\delta^2}{\sigma^2(Y_i)\Delta t} = 1 + b^2. \end{cases}$$

There are four unknowns and three equations and therefore an infinite number of solutions exist. Now denote $q := q_1$ and solve for the other probabilities with respect to q and b,

$$\begin{cases} q_2 = \frac{1}{2}(1 + b + b^2) - 3q \\ q_3 = 3q - b^2 \\ q_4 = \frac{1}{2}(1 - b + b^2) - q. \end{cases} \tag{2.22}$$

To determine the possible values of q, the same method as before is used. And as before, an inspection of the first and second order derivatives will show that the functions are strictly increasing or strictly decreasing for $b \in \left[0, \frac{1}{2}\right]$. So, the endpoints of b may be used to determine the maximum and minimum values of q. For the first equation,

$$0 \le \frac{1}{2}(1 + b + b^2) - 3q \le 1$$

is obtained. A small amount of algebra yields

$$\frac{1}{6}(-1 + b + b^2) \le q \le \frac{1}{6}(1 + b + b^2).$$

Given $b \in \left[0, \frac{1}{2}\right]$, with $b = 0$ gives $-\frac{1}{2} \le q \le \frac{1}{2}$ and for $b = \frac{1}{2}$, $-\frac{5}{8} \le p \le \frac{3}{8}$. Now taking the second equation of (2.22),

$$0 \le 3q - b^2 \le 1$$

is built. From this isolate q to obtain

$$\frac{b^2}{3} \leq q \leq \frac{1+b^2}{3}.$$

Solving for $b = 0$ gives $0 \leq q \leq \frac{1}{3}$, and $b = \frac{1}{2}$ yields $\frac{1}{12} \leq q \leq \frac{5}{12}$. Finally, take the third equation of (2.22) and write

$$0 \leq \frac{1}{2}(1 - b + b^2) - q \leq 1.$$

Using a similar process as earlier, this may be rewritten as

$$\frac{1}{2}(-1 - b + b^2) \leq q \leq \frac{1}{2}(1 - b + b^2).$$

Solving for $b = 0$ gives $-\frac{1}{6} \leq q \leq \frac{1}{6}$, and $b = \frac{1}{2}$ gives $-\frac{1}{24} \leq q \leq \frac{7}{24}$. Once again, from these six inequalities, it's easy to see that $q \in \left[\frac{1}{12}, \frac{1}{6}\right]$.

In order to calculate the price of an option, q is chosen such that $q \in \left[\frac{1}{12}, \frac{1}{6}\right]$. For each discrete time moment $i = 0, 1, ...N$ such that $t_i = i\Delta t$ there are m_i values of x and each may be denoted $x_{i,k}$ for $k = 1, 2, ...m_i$. Starting with time t_0, $m_0 = 1$ and $x_{0,1} = x_0$ with four successors belonging to the grid defined by $j\sigma(Y_i)\sqrt{\Delta t}$ at time t_1, $x_{1,1}$, $x_{1,2}$, $x_{1,3}$, and $x_{1,4}$. Each of which has four successors belonging to the grid defined by $j\sigma(Y_i)\sqrt{\Delta t}$ at time t_2, and so on. Since the four successors of each point $x_{i,k}$ are defined on a grid, they may be shared between parent nodes and this property makes the tree recombining and limits its growth rate. The probabilities of the successor nodes given the parent node are given by (2.18) and (2.22). At time t_N, value of the option is equal to the payoff of each call option at $x_{N,k}$ is calculated,

$$P(x_{N,k}, N) = \max\{e^{x_{N,k}} - E, 0\}.$$

From here, backwards induction is used for each time step $i = N-1, ..., 0$ and the price of the option at each node x is $V(x, i)$ is given according to the values of its successors $V(x_j, i+1)$ for $j = 1, 2, 3, 4$. For European options, $V^E(x, i) = e^{-r\Delta t} \sum_{j=1}^{4} V(x_j, i+1)q_j$. However, for American options, the option to exercise early must also be checked, so, $V^A(x, i) = \max\{e^{-r\Delta t} \sum_{j=1}^{4} V(x_j, i+1)q_j, \max\{e^x - E, 0\}\}$. In the case of both European and American options the price of the option is given at time $i = 0$ by $V^E(x, 0)$ and $V^A(x, 0)$ respectively.

The tree valuation is done M times, after which the results are averaged to obtain the Monte-Carlo estimation of the value of the option. In order to show that the tree constructed by this section converges to the process for dX_t in (2.3) with the values of Y being sampled from the estimated probability distribution generated from section 2.1, the following theorem was proved by Florescu and Viens (2008).

THEOREM 2.2.1. *Consider the quadrinomial tree, with nodes indexed by the log values of a stock, defined by the successors x_1, x_2, x_3, and x_4 of a value x as in (2.9), with probabilities q_1, q_2, q_3, and q_4 given by the relations (2.18) (resp. (2.22)), with $q = q_4$ (resp. $q = q_1$) when x_4 is furthest from the parent value x (resp. x_1 is furthest). For any fixed $q \in \left[\frac{1}{12}, \frac{1}{6}\right]$, these probabilities define a martingale measure on the paths of the tree. Furthermore, the Markov chain defined on the vertices of the tree under any such measure on the tree, defined by the relations (2.18) and (2.22), converges in distribution to the continuous process (2.3) as the time interval $h = \Delta t \to 0$ and the number of filtering particles $n \to \infty$.*

3 Tree-Based Method Using the Heston Model

This chapter is based on the work of Vellekoop and Nieuwenhuis (2009). Starting with the Heston model for a logarithmic stock price process $X_t = ln(S_t)$ and a squared volatility process Y under the risk-neutral probability measure $\mathbb{Q}$, this method creates a tree-like grid of possible values and then uses bilinear interpolation and backwards induction to estimate the price of the option. The specific version of the Heston model used in this chapter is

$$\begin{aligned} dX_t &= (r - \frac{1}{2}Y_t)dt + \sqrt{Y_t}dW_t, \\ dY_t &= \alpha(\nu - Y_t)dt + \beta\sqrt{Y_t}dZ_t \end{aligned} \tag{3.1}$$

and allows for correlated Brownian motions W_t and Z_t. The correlation coefficient in this model is ρ. Due to the fact that ν is the long term mean of the volatility process Y and $\sigma = \sqrt{Y_t}$, unlike in the previous method, ν must be positive in order to have a real solution. For each time moment, the method estimates the maximum and minimum possible values or X_t and Y_t, and then creates a grid of values. After the grid is created, an expected value for each node is calculated in a method using themes from both finite difference and tree methods. Using this method both the price of American and European options can be found.

3.1 Grid Setup

In this section the main parameters of the grid used to calculate the option price are calculated. Starting with the Heston model in (3.1), split the time period $t = 0, ..., T$ into $m \in \{\mathbb{N} > 0\}$ pieces so that $\Delta t = \frac{T}{m}$. Now, using the Euler method, define the discrete time stochastic process

$$\begin{aligned} X_{k+1} &= X_k + (r - \frac{1}{2}Y_k)\Delta t + z^1_{k+1}\sqrt{Y_k^+\Delta t}, \\ Y_{k+1} &= Y_k + \alpha(\nu - Y_k^+)\Delta t + z^2_{k+1}\beta\sqrt{Y_k^+\Delta t}, \end{aligned} \tag{3.2}$$

where k represents the index of each X_k and Y_k for $k = 0, 1, 2, ..., m-1$ and z_k^1, $z_k^2 \in \{-1, 1\}$ are iid random variables distributed according to

$$\mathbb{Q}(z_k^1 = i, z_k^2 = j) = \frac{1}{4}(1 + ij\rho), \;\; i, j \in \{-1, 1\}. \tag{3.3}$$

Here, Y_k^+ is the non-negative portion of Y_k, or $Y_k^+ = max\{Y_k, 0\}$. In an unconstrained model, the number of calculations is 4^m. So, in order to limit the number of calculations, a grid is defined. Start by obtaining the maximum and minimum values of x_k and y_k at each time k by taking

$$\begin{aligned} x_k^{max} &= \max\{x \mid \mathbb{Q}(X_k = x) > 0\}, \\ x_k^{min} &= \min\{x \mid \mathbb{Q}(X_k = x) > 0\}, \\ y_k^{max} &= \max\{y \mid \mathbb{Q}(Y_k = y) > 0\}, \end{aligned}$$

and

$$y_k^{min} = \min\{y \mid \mathbb{Q}(Y_k = y) > 0\}.$$

Now, choose a number of points in the price and volatility directions at each time step to be m_x and $m_y \in \{\mathbb{N} > 0\}$, then define

$$\Delta x_k = \frac{(x_k^{max} - x_k^{min})}{m_x}$$

and

$$\Delta y_k = \frac{(y_k^{max} - y_k^{min})}{m_y}.$$

Using the two previous sets of equations define a set of ordered pairs for x and y,

$$\hat{S} = \{(x_k^{min} + i\Delta x_k, y_k^{min} + j\Delta y_k) \mid i = 0, ..., m_x, j = 0, ..., m_y\}. \tag{3.4}$$

In order to obtain option prices between grid points bilinear interpolation is needed. The linear interpolation formula for a function $f(x)$ is defined as

$$f(x) = c_0 f(x_0) + c_1 f(x_1),$$

where

$$c_0 := 1 - \frac{x - x_0}{x_1 - x_0}$$

and

$$c_1 := \frac{x - x_0}{x_1 - x_0}.$$

From their definitions it is easy to see that restricting $x_0 \leq x \leq x_1$ that $0 \leq c_i \leq 1,\ i \in \{0, 1\}$. It is also clear that $c_0 + c_1 = 1$. Now consider bilinear interpolation for the function $f(x, y)$. First, interpolation is done in the x-direction for y_0,

$$f(x, y_0) = \left(1 - \frac{x - x_0}{x_1 - x_0}\right) f(x_0, y_0) + \frac{x - x_0}{x_1 - x_0} f(x_1, y_0),$$

and then for y_1,

$$f(x, y_1) = \left(1 - \frac{x - x_0}{x_1 - x_0}\right) f(x_0, y_1) + \frac{x - x_0}{x_1 - x_0} f(x_1, y_1).$$

Next, using the same definitions for c_0 and c_1, interpolate these formulas in the y-direction,

$$f(x, y) = \left(1 - \frac{y - y_0}{y_1 - y_0}\right) \left[c_0 f(x_0, y_0) + c_1 f(x_1, y_0)\right] + \frac{y - y_0}{y_1 - y_0} \left[c_0 f(x_0, y_1) + c_1 f(x_1, y_1)\right].$$

Now, let $c_{i,0} = (1 - \frac{y - y_0}{y_1 - y_0}) c_i$ and $c_{i,1} = \frac{y - y_0}{y_1 - y_0} c_i$ for $i \in \{0, 1\}$, and

$$f(x, y) = c_{0,0} f(x_0, y_0) + c_{1,0} f(x_1, y_0) + c_{0,1} f(x_0, y_1) + c_{1,1} f(x_1, y_1) \tag{3.5}$$

is obtained. Given that $0 \leq c_i \leq 1,\ i \in \{0, 1\}$ is the result of restricting $x_0 \leq x \leq x_1$, placing the additional restriction of $y_0 \leq y \leq y_1$ guarantees that $0 \leq c_{i,j} \leq 1,\ i, j \in \{0, 1\}$. Calculating the sum of $c_{i,j}$ for $i, j \in \{0, 1\}$ is done by first writing

$$\left(1 - \frac{y - y_0}{y_1 - y_0}\right)\left(1 - \frac{x - x_0}{x_1 - x_0}\right) + \left(1 - \frac{y - y_0}{y_1 - y_0}\right)\left(\frac{x - x_0}{x_1 - x_0}\right)$$
$$+ \left(\frac{y - y_0}{y_1 - y_0}\right)\left(1 - \frac{x - x_0}{x_1 - x_0}\right) + \left(\frac{y - y_0}{y_1 - y_0}\right)\left(\frac{x - x_0}{x_1 - x_0}\right).$$

Next, factor out the terms containing y to obtain

$$\left(1 - \frac{y - y_0}{y_1 - y_0}\right)\left[\left(1 - \frac{x - x_0}{x_1 - x_0}\right) + \left(\frac{x - x_0}{x_1 - x_0}\right)\right] + \left(\frac{y - y_0}{y_1 - y_0}\right)\left[\left(1 - \frac{x - x_0}{x_1 - x_0}\right) + \left(\frac{x - x_0}{x_1 - x_0}\right)\right].$$

From before, it is known that $c_0 + c_1 = 1$, so

$$\left(1 - \frac{y - y_0}{y_1 - y_0}\right) + \left(\frac{y - y_0}{y_1 - y_0}\right)$$

may be written, and is clearly equal to 1. Therefore $c_{0,0} + c_{1,0} + c_{0,1} + c_{1,1} = 1$, allowing these to be considered as probabilities.

3.2 Grid Calculation

Now that the mathematical background and grid setup has been given, the grid used for calculating the discounted expected value of the option needs to be set up. First, the number of time steps $m \in \{\mathbb{N} > 0\}$ and the mesh sizes for price and volatility are chosen such that $m_x, m_y \in \{\mathbb{N} > 0\}$ are chosen. Given starting values for volatility squared $Y_0^2 = \sigma^2$ and log-price of the stock $X_0 = ln(S_0)$ iterate over equation (3.2) so that for each discrete time k the maximum and minimum values of X_{k+1} and Y_{k+1} are denoted x_{k+1}^{max}, x_{k+1}^{min}, y_{k+1}^{max}, and y_{k+1}^{min}. In order to define the grid of prices and volatilities $\hat{S}_k$, first the values of Δx_k and Δy_k are calculated as

$$\Delta x_k = \frac{x_k^{max} - x_k^{min}}{m_x}$$

and

$$\Delta y_k = \frac{y_k^{max} - y_k^{min}}{m_y}.$$

Using these values, the grid

$$\hat{S}_k = \{(x_k^{min} + i\Delta x_k, y_k^{min} + j\Delta y_k) \mid i = 0, ..., m_x, j = 0, ..., m_y\}$$

may be built.

3.3 Calculate Option Value

Calculate the option values during the final time step for either a call or put option with either

$$V_C(S_T, T) = \max\{S_T - E, 0\}$$

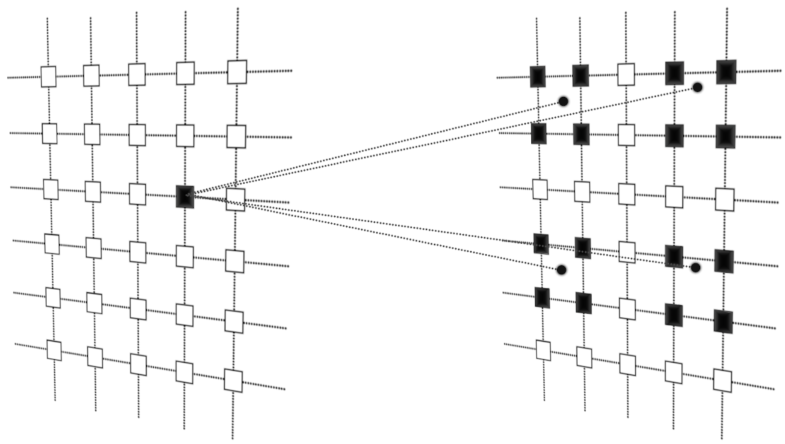

Figure 3.1: The successors of a point within the grid

or

$$V_P(S_T, T) = \max\{E - S_T, 0\}.$$

Now work backwards through the grid starting at $k = m - 1$. From each node in $\hat{S}_k$ at time step k calculate the four successors for each grid point, $\hat{S}_k(x, y)$. This is done using equation (3.2), and the four successors are denoted $(x^{z1}_{k+1},\ y^{z2}_{k+1})$ for $z1, z2 \in \{-1, 1\}$. Note that here the superscripts $z1$ and $z2$ representing the values of z^1_{k+1} and z^2_{k+1} used in equation (3.2).

For each of the four successors, $(x^{z1}_{k+1},\ y^{z2}_{k+1})$, find the four grid points in $\hat{S}_{k+1}$ which surround it. For example, for $(x^{-1}_{k+1},\ y^{-1}_{k+1})$, the point x_0 is the largest point x^S_{k+1}, where the superscript S represents that the point is aligned with the grid, which is smaller than x^1_{k+1}, which may be written as

$$x_0 = \max\{x^{-1}_{k+1} \geq x^S_{k+1}\}.$$

Similarly,

$$y_0 = \max\{y^{-1}_{k+1} \geq y^S_{k+1}\}.$$

The points above $(x_{k+1}^{-1},\ y_{k+1}^{-1})$ are x_1 and y_1 and are written similarly. That is,

$$x_1 = \min\{x_{k+1}^{-1} \leq x_{k+1}^{S}\}$$

and

$$y_1 = \min\{y_{k+1}^{-1} \leq y_{k+1}^{S}\}.$$

And so, each of the four successors $(x_{t+1}^{z1},\ y_{t+1}^{z2})$ has four points around it, $\hat{S}_{t+1}(x_0,\ y_0)$, $\hat{S}_{t+1}(x_1,\ y_0)$, $\hat{S}_{t+1}(x_0,\ y_1)$, and $\hat{S}_{t+1}(x_1,\ y_1)$.

So, instead of the initial discrete process (X_k, Y_k) given by (3.2) with 4 successors for each point, consider the new process $(\tilde{X}_k, \tilde{Y}_k)$ with 16 successors for each point. The probabilities of the successors are given by

$$c_{k,l}\mathbb{Q}(z^1 = i, z^2 = j)$$

for $k, l \in \{0, 1\}$ and $i, j \in \{-1, 1\}$. In order to show that the new process can be used to price the option, the following theorem was proved in the paper by Vellekoop and Nieuwenhuis (2009).

THEOREM 3.3.1. *Assume that:*

$$\lim_{m\to\infty} \Delta t^m = 0, \quad \lim_{m\to\infty} \max_{k=1,\ldots,m-1} \frac{\Delta y_k^m}{\Delta t^m} = 0, \quad \lim_{m\to\infty} \max_{k=1,\ldots,m-1} \frac{\Delta x_k^m}{\Delta t^m} = 0,$$

where $\Delta t^m = \frac{T}{m}$*. Then the process* $(\tilde{X}_k,\ \tilde{Y}_k)$ *converges in distribution to the continuous process* $(X,\ Y)$ *defined in (3.1).*

For each point $(x_{t+1}^{z1},\ y_{t+1}^{z2})$, use bilinear interpolation as defined in equation (3.5) on the option values of the surrounding grid points $V(\hat{S}_{t+1}(x_0,\ y_0), t+1)$, $V(\hat{S}_{t+1}(x_1,\ y_0), t+1)$, $V(\hat{S}_{t+1}(x_0,\ y_1), t+1)$, and $V(\hat{S}_{t+1}(x_1,\ y_1), t+1)$ to get an estimation of the value of the option $V((x_{t+1}^{z1},\ y_{t+1}^{z2}), t+1)$. This gives an estimation of the value of the option at the points $(x_{t+1}^{z1},\ y_{t+1}^{z2})$ of

$$\begin{aligned} V((x_{t+1}^{z1},\ y_{t+1}^{z2}), t+1) =& c_{0,0}V(\hat{S}_{t+1}(x_0,\ y_0), t+1) + c_{1,0}V(\hat{S}_{t+1}(x_1,\ y_0), t+1) + \\ & c_{0,1}V(\hat{S}_{t+1}(x_0,\ y_1), t+1) + c_{1,1}V(\hat{S}_{t+1}(x_1,\ y_1), t+1). \end{aligned}$$

In order to calculate the value of the option of the grid point V$(\hat{S}(x_t, y_t), t)$, each bilinear interpolated successor option value $V((x_{t+1}^{z1},\ y_{t+1}^{z2}), t+1)$ happens with the probability defined in equation (3.3), and the value of the option at point $\hat{S}(x_t, y_t)$ is

$$V(\hat{S}(x_t, y_t), t) = e^{-r\Delta t} \sum_{i \in \{-1,1\}} \sum_{j \in \{-1,1\}} V((x_{t+1}^{i},\ y_{t+1}^{j}), t+1)\mathbb{Q}(z^1 = i, z^2 = j). \tag{3.6}$$

Iterate from time $t = T - \Delta t$ to $t = 0$, and at time $t = 0$, the price of a European call option is given by $V(\hat{S}(x_0, y_0), 0)$. An American option may be exercised early, and so first find

$$W(\hat{S}(x_t, y_t), t) = \sum_{i \in \{-1,1\}} \sum_{j \in \{-1,1\}} V((x_{t+1}^{i},\ y_{t+1}^{j}), t+1)\mathbb{Q}(z^1 = i, z^2 = j). \tag{3.7}$$

And to find the price of an American option,

$$V_C^A(\hat{S}(x_t, y_t), t) = \max\{e^{-r\Delta t} W(\hat{S}(x_t, y_t), t),\ \max(e^{x_t} - E, 0)\}$$

is used. Iterate from time $t = T - \Delta t$ to $t = 0$, and at time $t = 0$, the price of the American call option is $V_C^A(\hat{S}(x_0, y_0), 0)$.

4 Numerical Experiments

In this chapter the results of the numerical experiments performed are presented. All times are generated on a system with a Core i7-4500U processor and 8 GB of RAM. The algorithms were implemented first in Python, then due to their slow performance, converted to C. The performance of the Python code, for larger numbers of calculations, was on the order of minutes instead of seconds. Since performance accurately measured on the scale of seconds, all times presented will be in seconds. Since C only has the ability to generate uniformly distributed random numbers, generation of normally distributed numbers W_1 and W_2 was done using the Box-Muller Transform (Muller 1959). Therefore,

$$W_1 = \sqrt{-2ln(U_1)}cos(2\pi U_2),$$
$$W_2 = \sqrt{-2ln(U_1)}sin(2\pi U_2),$$

where U_1 and $U_2 \sim U(0,1)$ are independent uniformly distributed random variables between 0 and 1. The compilation was done using GCC for a Linux system. The generation of Black-Scholes prices was done in Python using the NumPy library. In this chapter, when referring to accuracy, this means accurate to the Black-Scholes model or exact values given in the respective original paper. A useful quote to keep in mind is, "Just as we have only language to describe language's flaws, so we have become accustomed to using the BSM [Black-Scholes-Merton] language to describe the violations of BSM." (Derman and Miller, 2016) This means that even though the constant volatility assumption of the Black-Scholes model has its shortcomings, it still provides a good baseline for comparison and for this reason it is used as a marker of relative accuracy.

4.1 Quadrinomial Recombining Tree

In order to check the performance of each part of the algorithm, the validation must be performed in two steps. First, the performance of the second step independent of the first step, and then the algorithm as a whole. In this method, the quadrinomial tree portion produces a similar value as the Black-Scholes price for the same volatility. It will be shown that the parameters given by Florescu and Viens (2008) do not produce the correct result. The average volatility in this section are calculated for the estimated filtered stochastic volatility distribution as

$$\bar{\sigma} = \sum_{y \in Y} e^{-|y|} p(y) \tag{4.1}$$

where the set Y is the set of volatility values estimated during the first part of the algorithm and $p(y)$ is the probability of that value of $y \in Y$ occurring in the estimated distribution. In each subsection, an alternate set of parameters are given which give a more accurate result. As in the paper, 100 simulations are done using the Monte Carlo method for each result. In some tables, the bid and ask prices of the options from Florescu and Veins (2008) are given. The **bid** is the highest price an investor is willing to pay to buy the option. The **ask** is the lowest price an investor is willing to sell the option at.

4.1.1 IBM Stock

For the IBM data start with The starting stock value is $S_0 = 83.70$, the price of the stock on July 19, 2005, and $T = 42/252$. Table 4.1 gives an overview of the results for the options expiring in September 2005. In the column QT Const Vol, volatility wasn't estimated from the probability distribution, but held constant at $\sigma = 0.234$, the value given in Florescu and Viens (2008) which they based on historical data. The risk-neutral probability q of the furthest node from x in figure 2.1 is chosen as 0.135. The choice of 0.135 is based on the paper by Florescu and Veins(2008). However, it will be shown later that the method is agnostic to the value of q within the range defined in section 2.2, $q \in \left[\frac{1}{12}, \frac{1}{6}\right]$.

Strike	Bid	Ask	BS Price	QT Const Vol	FV Price	Run 1	Run 2
60	23.8	24.0	24.0424	24.0431	24.1654	24.0424	24.0454
70	13.9	14.1	14.1733	14.1532	14.1654	14.1750	14.2857
75	9.0	9.2	9.5431	9.5220	9.1700	9.5424	9.8385
80	4.6	4.8	5.6188	5.6068	5.5738	5.6112	6.1152
85	1.6	1.65	2.8130	2.8232	1.0571	2.8019	3.3854
90	0.35	0.4	1.1809	1.2085	0.1123	1.1727	1.6632
95	0.1	0.15	0.4154	0.4426	0.0040	0.4138	0.7278

Table 4.1: IBM: Initial Quadrinomial Tree Results

In the columns Run 1 and Run 2, the estimation of the filtered stochastic probability

Strike	Bid	Ask	BS Price	Mod ν 1	Mod ν 2	Mod ν, β 1	Mod ν, β 2
60	23.8	24.0	24.0424	24.0420	24.0420	24.0421	24.0422
70	13.9	14.1	14.1733	14.1136	14.1251	14.1454	14.1639
75	9.0	9.2	9.5431	9.2858	9.3502	9.4465	9.5130
80	4.6	4.8	5.6188	5.0452	5.2106	5.4309	5.5624
85	1.6	1.65	2.8130	2.1043	2.3134	2.5889	2.7465
90	0.35	0.4	1.1809	0.6492	0.7981	1.0025	1.1264
95	0.1	0.15	0.4154	0.1498	0.2148	0.3142	0.3835

Table 4.2: IBM: Quadrinomial Tree Results with Modified Parameters

	Original Parameters			Mod ν			Mod ν, β		
q	E=60	80	95	E=60	80	95	E=60	80	95
1/12	24.0430	5.7858	0.5156	24.0420	5.1847	0.2001	24.0421	5.4687	0.3332
17/192	24.0430	5.7857	0.5156	24.0420	5.1846	0.2002	24.0421	5.4686	0.3333
3/32	24.0430	5.7856	0.5157	24.0420	5.1845	0.2003	24.0421	5.4686	0.3333
19/192	24.0430	5.7855	0.5158	24.0420	5.1845	0.2004	24.0421	5.4686	0.3333
5/48	24.0430	5.7853	0.5159	24.0420	5.1844	0.2004	24.0421	5.4684	0.3333
7/64	24.0430	5.7852	0.5160	24.0420	5.1844	0.2005	24.0421	5.4685	0.3333
11/96	24.0431	5.7851	0.5161	24.0420	5.1843	0.2006	24.0421	5.4685	0.3333
23/192	24.0431	5.7850	0.5162	24.0420	5.1843	0.2007	24.0421	5.4684	0.3333
1/8	24.0431	5.7849	0.5163	24.0420	5.1843	0.2007	24.0421	5.4684	0.3333
25/192	24.0431	5.7848	0.5164	24.0420	5.1842	0.2008	24.0421	5.4683	0.3333
13/96	24.0431	5.7847	0.5165	24.0420	5.1842	0.2009	24.0421	5.4683	0.3333
9/64	24.0431	5.78496	0.5166	24.0420	5.1841	0.2010	24.0421	5.4683	0.3333
7/48	24.0431	5.7845	0.5167	24.0420	5.1841	0.2010	24.0421	5.4682	0.3333
29/192	24.0431	5.7843	0.5168	24.0420	5.1840	0.2011	24.0421	5.4682	0.3333
5/32	24.0431	5.7842	0.5169	24.0420	5.1840	0.2012	24.0421	5.4682	0.3333
31/192	24.0431	5.7841	0.5170	24.0420	5.1839	0.2013	24.0421	5.4681	0.3333
1/6	24.0431	5.7840	0.5171	24.0420	5.1839	0.2014	24.0421	5.4681	0.3333

Table 4.3: IBM: Quadrinomial Tree Results with Different Values of q

distribution was done with historical daily data from April 18, 2004 to July 18, 2005 and the parameters presented in Florescu and Veins (2008). That is the parameters for the model are $\alpha = 11.85566$, $\nu = 0.9345938$, $\beta = 4.13415$, $\mu = 0.04588$, and $r = 0.0343$ with $M = 300$ discrete time steps and $n = 1000$ paths for stock price and volatility. The Run 1 column gives the estimation of the price for the whole algorithm for one run. The average volatility given by equation (4.1) for this run was 0.1996. The estimation of the filtered stochastic volatility distribution took 2.6969 seconds and starts at the beginning of the program, including the time it takes to calculate $X_t = ln(S_t)$, and ends once the distribution is estimated. The Monte-Carlo method took 0.6303 seconds and starts with the allocation of the memory for the calculations, including the generation of Y values for each run, and ends once the sum of the runs is completed for all strikes, the division by the number of runs is done inline with the print statement. The set of Y values for each run is used across the range of strikes. While the results of this run look promising, they don't show the whole picture. The results vary quite widely. For another run, the values were as given in the column Run 2 where the average volatility was 0.2307. They also don't match the results from the paper by Florescu and Viens (2008) which are shown in the FV Price column.

As the article had inconsistent parameter notations when giving the values of the paraemeters, and the average volatility in the case of run 1 differs quite a bit from 0.234, the constant rate used in the Black Scholes calculation, consider the following two cases. In the first case, $\nu = ln(0.234)$, and in the second $\nu = ln(0.234)$ and $\beta = 1$. This is done to analyze how the results change as a result of changing these parameters. In the first case, the volatility lowers to between 0.1462 and 0.1604 and the results are closer to the market prices. Changing β as well brings the results back up toward the Black-Scholes price. The average volatilities for these runs were 0.2144 and 0.2249. In both cases, the difference in average volatility between the runs was smaller, 0.0142 and 0.0105, than the original parameters, which had a difference of 0.0311 between the two runs.

In table 4.3, evenly spaced values of $q \in \left[\frac{1}{12}, \frac{1}{6}\right]$ in increments of $\frac{1}{192}$ were checked for $E = 60, 80, 95$. It is clear that for the case of the available data, that the choice of q given the same set of volatilities makes a difference of less than one cent across the range of strike prices.

4.1.2 S&P 500 Index

For the S&P 500, a similar result to the IBM results is obtained. That is, the parameters given in the paper by Florescu and Viens (2008) do not give accurate results, however, by changing ν to $ln(0.13)$, the constant volatility of the stock, accurate results can be reached. And again, the performance of the algorithm is relatively independent from the choice of q. Here, the tables give a brief overview of these facts, but they also investigate the how the model changes with different amounts of historical data.

In table 4.4 there is an overview of several results. The first is the BS Price, which is the Black-Scholes price with $\sigma = 0.13$. The Est Price column uses $\nu = -4.38$, the value

Strike	Bid	Ask	BS Price	FV Price	Est Price	Mod ν	3 Mo Est	1 Mo Est
700	435.9	437.9	440.7351	440.8361	440.7351	440.7351	440.7351	440.7351
1005	131.9	133.9	136.1130	135.8376	136.0859	136.1148	136.1149	136.1127
1100	42.4	44.4	46.7735	43.8620	41.1951	46.8897	46.8958	46.7981
1135	17.1	18.6	23.2710	19.1077	6.4631	23.4432	23.4509	23.3099
1140	14.3	15.8	20.6695	16.4364	2.6238	20.8421	20.8514	20.7119
1175	2.7	3.0	7.7807	4.6066	0.0000	7.9166	7.9261	7.8108
1225	0.15	0.2	1.1934	0.3398	0.0000	1.2375	1.2411	1.1989

Table 4.4: S&P 500: Comparison of Parameters and Amount of Historical Data

	Original Parameters			Modified ν		
q	E=700	1140	1225	E=700	1140	1225
1/12	440.7351	2.6458	0.0000	440.7351	20.6894	1.1922
17/192	440.7351	2.6458	0.0000	440.7351	20.6893	1.1923
3/32	440.7351	2.6457	0.0000	440.7351	20.6892	1.1925
19/192	440.7351	2.6457	0.0000	440.7351	20.6891	1.1927
5/48	440.7351	2.6457	0.0000	440.7351	20.6890	1.1928
7/64	440.7351	2.6457	0.0000	440.7351	20.6889	1.1930
11/96	440.7351	2.6457	0.0000	440.7351	20.6888	1.1931
23/192	440.7351	2.6456	0.0000	440.7351	20.6886	1.1933
1/8	440.7351	2.6456	0.0000	440.7351	20.6885	1.1935
25/192	440.7351	2.6456	0.0000	440.7351	20.6884	1.1936
13/96	440.7351	2.6456	0.0000	440.7351	20.6883	1.1938
9/64	440.7351	2.6456	0.0000	440.7351	20.6882	1.1940
7/48	440.7351	2.6455	0.0000	440.7351	20.6881	1.1942
29/192	440.7351	2.6455	0.0000	440.7351	20.6880	1.1943
5/32	440.7351	2.6455	0.0000	440.7351	20.6878	1.1945
31/192	440.7351	2.6455	0.0000	440.7351	20.6877	1.1947
1/6	440.7351	2.6455	0.0000	440.7351	20.6876	1.1948

Table 4.5: S&P: Quadrinomial Tree Results with Different Values of q

provided in Florescu and Viens (2008). With this the model underestimates the average volatility of the S&P 500 when compared to the constant volatility of 0.13. The typical average volatility for the original parameters is between 0.013 and 0.014. For this estimate and the rest of the subsection the rest of the parameters were: $S_0 = 1139.93$, $T = 29/252$, $\alpha = 50$, $\beta = 1$, $\mu = 0.04$, and $r = 0.01$. And again, for this subsection $M = 300$ and $n = 1000$. To generate the values using daily data from January 1st, 1999 to April 21st, 2004 took 55.1945 seconds for the estimation of the filtered stochastic volatility distribution.

In the Mod ν column, $\nu = ln(0.13) = -2.040221$, or the logarithm of the constant volatility. This corrects the severe underestimation of the volatility and brings the average to approximately 0.13. As expected, the time it took to get the values was similar to the previous result, taking 55.3296 seconds for the volatility distribution estimation.

The 3 Mo Est column continues using $\nu = -2.040221$, but also uses only historical data from January 21st, 2004 to April 21st, 2004 to estimate the filtered stochastic probability distribution, and the 1 Mo Est column reduces that further to March 21st, 2004 to April 21st, 2004. The times to generate the distributions are 2.5985 and 0.8655 seconds respectively. When annualized volatilities were calculated on the data starting January 1st, 1999 using

$$\sigma = \sqrt{\frac{252}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2},$$

$\sigma = 0.2082$ was obtained. For the data starting January 21st, 2004, $\sigma = 0.1242$. And finally that beginning February 21st, 2004, $\sigma = 0.1272$. From this, it seems as though it is logical to say that, in this case, the genetic algorithm doesn't need more than a month of daily data to be accurate. As seen by the differences in the choice of ν it is clear that it is what drives the average volatility of the distribution and therefore the estimated prices of the options.

The last thing to evaluate is the stability of q as it relates to the S&P 500, and this is done in table 4.5. The first thing to notice is that the value of q affects the value of the option at the deep in-the-money strike of 700 by less than 1/100 of a cent. Once again, the choice of q makes a difference of less than one cent in the valuation of the option across the range of strikes.

From these results, it has been shown that the quadrinomial tree method works well. But, since the parameters from the original paper were unclear due to some confusion with the notation, it is difficult to give an exact estimate of how accurate the parameters and results are.

4.1.3 Implementation Comparison

The tree model is unique in the fact that European options present an interesting implementation option, which is, instead of calculating all of the probabilities and then working backward through time, using the law of total probability to carry the probabilities forward through the tree and using the conditional probabilities of the final nodes with their values to calculate the price. So for, an option with maturity date T split into M periods, discrete time m is $t = m\Delta t$, $m = 0, 1, ..., M$. If there are N nodes in the tree at discrete time step m and K nodes at $m-1$,

$$\mathbb{Q}(x_m^n) = \sum_{s \in \{x_{m-1}^k\}} \mathbb{Q}(x_m^n \mid s)\mathbb{Q}(s)$$

for each node x_m^n in $m = 1, 2, ..., M$, $k = 1, ..., K$, and $n = 1..., N$.

In calculating the successor values, it is possible to calculate only those that are needed, however, another consideration with this approach is calculating all values of $j \in [j_{min}, j_{max}]$ between its maximum and minimum values at the current time step, recall that $j = inf\{l \in \mathbb{N} | l\sigma(Y_i)\sqrt{\Delta t} \geq x\}$. So,

$$j_{max} = \left\lceil \frac{x_{max}}{\sigma(Y_i)\sqrt{\Delta t}} \right\rceil + 1$$

and

$$j_{min} = \left\lceil \frac{x_{min}}{\sigma(Y_i)\sqrt{\Delta t}} \right\rceil - 2$$

where $x_{min} = \min\{x_m^n\}$ and $x_{max} = \max\{x_m^n\}$.

Version	Unoptimized	Optimized	10000 MC Sims	N=1000	N=2000
Minimum	1.4981	0.5774	56.2485	71.0865	322.7444
Brute Force	1.3886	0.5504	54.8037	71.7764	284.8044
Amer Comp	1.6995	0.5881	61.5190	70.0989	268.8641

Table 4.6: Algorithm Implementation Comparison

This means that there are three main ideas to test, a European option only method which calculates only the needed values, a European option only method which calculates for all possible values of j, and using the traditional tree approach which is compatible with American options as well. Here they'll be called the minimal method, the brute-force method, and the American option compatible method. Table 4.6 shows a comparison of the times, in seconds, between the algorithms. The Unoptimized column shows the times with no compiler optimizations, the optimized column uses the standard compiler optimizations, and the 10000 MC Sims column uses 10000 Monte Carlo simulations instead of the previous 100, as in the first two. The times gathered in this subsection are based on the original parameters for the IBM stock, and while using the modified parameters make execution time faster, the same relationships between the times hold.

The final two columns of table 4.6 show how the methods compare when increasing the size of the tree, when the number of time steps in the tree increased to 1000 and then to 2000. The American option compatible method is the slowest when the depth of the tree is small because of the overhead of setting up the tree and extra logic. Even when the number of Monte Carlo simulations is increased, the time difference is still there due to the tree needing to be set up each time. However, when the depth of the tree is increased, the difference in the number of calculations for the expected value goes up. The effect of those extra calculations are seen in the minimum method since it contains the added logic of the tree to calculate only the required values and the extra calculations of carrying the probabilities forward. The brute force method, in contrast, saves time by avoiding many of the logical statements needed by the other two methods. In terms of memory overhead, the minimal and brute-force methods require 16 bytes represent each node. Since past nodes aren't required, they may be ignored and so the maximum memory required depends only on the time moment with the highest number of nodes. The American option compatible method requires the most memory requiring 80 bytes for each node, and requiring that they are persistent throughout the generation of the tree. The American option compatible method may be able to be improved by adding features and ideas from the brute-force method, but this is beyond the scope of this thesis.

4.2 Tree-Based Method Using the Heston Model

This method works for both American and European options, and the accuracy of the results from Vellekoop and Nieuwenhuis (2009) were able to be duplicated. The values obtained in this thesis and those in the original paper, at most, differ by three cents for European options and four cents for American options. But, the magnitude of error (the absolute difference between the calculated an exact price) is approximately the same as in the original paper, and the values seem to converge to the same numbers. The results are shown in tables 4.7 and 4.8. The average times are calculated using the run times of the five starting prices of the underlying, $S_0 = \{8, 9, 10, 11, 12\}$ with each set of parameters m_x, m_y, and m, where m is the number of discrete time points between $t = 0$ and $t = T$. Additionally, m_x is the number of price points in the grid at time t and m_y is the number of volatility points. In this case the parameters $E = 10$, $\alpha = 5$, $\nu = 0.16$, $\beta = 0.9$, $\rho = 0.1$, and $r = 0.1$ are taken from Vellekoop and Nieuwenhuis (2009). The exact values are also taken from the paper and are known solutions to the Heston model with these parameters. The average times are taken by running the simulations for one set of values m_x, m_y, and m, and dividing by five, the number of starting values, to obtain an average time required to calculate one price with that set of parameters. In the tables, as expected, as m_x, m_y,

m_x	m_y	m	$S_0 = 8$	9	10	11	12	Avg Time
125	6	25	1.8435	1.0382	0.4838	0.2011	0.0823	0.0013
250	12	35	1.8420	1.0417	0.4893	0.2024	0.0813	0.0074
500	24	50	1.8410	1.0440	0.4927	0.2041	0.0810	0.0405
1000	48	71	1.8403	1.0454	0.4954	0.2055	0.0806	0.2254
Exact			1.8389	1.0483	0.5015	0.2082	0.0804	

Table 4.7: European Put Option

m_x	m_y	m	$S_0 = 8$	9	10	11	12	Avg Time
125	6	25	1.9925	1.0934	0.5010	0.2060	0.0839	0.0017
250	12	35	1.9946	1.0977	0.5076	0.2076	0.0829	0.0091
500	24	50	1.9963	1.1007	0.5106	0.2094	0.0826	0.0518
1000	48	71	1.9974	1.1028	0.5136	0.2109	0.0822	0.2914

Table 4.8: American Put Option

and m are increased, the estimation by the model seems to converge to the exact price of the options.

4.3 Model Comparison

Now that the models' individual results have been presented, a comparison is in order. First, the volatility of the quadrinomial tree method is

$$\sigma_1(Y_t) = e^{-|Y_t|}.$$

for And, for the tree-based method,

$$\sigma_2(y_t) = \sqrt{y_t}.$$

It is clear that σ_1 and σ_2 are not equivalent and therefore the stochastic processes for Y_t and y_t are not interchangeable. Therefore, a transformation is needed. Starting with

$\sigma_1(Y_t) = \sigma_2(y_t)$, from which

$$|Y_t| = ln\left(\frac{1}{\sqrt{y_t}}\right) \tag{4.2}$$

and

$$y_t = e^{-2|Y_t|} \tag{4.3}$$

may be obtained. Since the tree-based method is based on a squared volatility and the quadrinomial tree isn't, equation (4.3) is quite intuitive. It also shows how easy it is to substitute any OU process into this model. On the other hand, equation (4.2) provides some interesting challenges. In order for the value to be defined under the square root and to avoid dividing by zero, $0 < y_t$ must be true and for the logarithm to be positive as required by the absolute value, $y_t \leq 1$ must be met. However, these conditions are not guaranteed by the process y_t. Even if a formula could be found such that these restrictions could be relaxed and historical data for the underlying was available, two parameters present problems. First, a non-zero correlation coefficient ρ is not allowed in the quadrinomial tree model. And second, μ isn't used in the tree-based model so it would need to be first estimated. This means that the tree-based model parameters cannot be used in the quadrinomial tree method. Since it is only possible to use the quadrinomial tree parameters in the tree-based method, but not the other way around, the results of this investigation will be explored.

			$m_x = 125$	250	500	1000
			$m_y = 6$	12	24	48
Strike	BS Price	QT Price	$m = 25$	35	50	71
60	24.024	24.0422	24.0420	24.0396	24.0402	24.0411
70	14.1733	14.1632	14.2311	14.2019	14.1964	14.1950
75	9.5431	9.5103	9.6617	9.6102	9.5964	9.6913
80	5.5188	5.5572	5.7906	5.7176	5.6961	5.6863
85	2.8130	2.7401	3.0058	2.9230	2.8969	2.8874
90	1.1809	1.1219	1.3571	1.2813	1.2588	1.2500
95	0.4154	0.3805	0.5493	0.4909	0.4757	0.4691
Est Vol Time		2.5794				
MC Time		0.3951				
Total Time		3.1218	0.0125	0.0679	0.3858	2.2003

Table 4.9: IBM: Comparison Results

Using the discussion above the models can be directly compared using the IBM parameters $\alpha = 11.85566$, $\nu = ln(0.234)$, $\beta = 1$, $T = 42/252$, $r = 0.0343$, and $\rho = 0$ the results are in table 4.9. The BS Price column contains the Black-Scholes price for comparison. The

			$m_x = 125$	250	500	1000
			$m_y = 6$	12	24	48
Strike	BS Price	QT Price	$m = 25$	35	50	71
700	440.7351	440.7351	440.7174	440.7200	440.7235	440.7264
1005	136.1130	136.1183	136.0783	136.0859	136.0944	136.1009
1100	46.7735	47.0392	46.7645	46.7606	46.7908	46.8163
1135	23.2710	23.6549	23.3007	23.2832	23.3042	23.3392
1140	20.6695	21.0575	20.6918	20.6676	20.7169	20.7320
1175	7.7807	8.0954	7.8034	7.7788	7.8079	7.8365
1225	1.1934	1.3020	1.1770	1.1819	1.1990	1.2148

Table 4.10: S&P Quadrinomal Tree Model

next column, QT Price, shows the prices generated using the quadrinomial tree method, and the last four columns give the tree-based prices with the given values of m_x, m_y, and m.

For the S&P the parameters $\alpha = 50$, $\nu = ln(0.13)$, $\beta = 1$, $T = 29/252$, $r = 0.01$, and $\rho = 0$ are used. The results are in table 4.10 and the layout is the same as the table for IBM.

From these two sets of results it is possible to conclude that given the proper transformation of the volatility process, it is possible to use the parameters of the quadrinomial tree in the tree-based method. Therefore, the tree-based method is more flexible, in that it can more easily take in parameters of different volatility models, it doesn't require the estimation of μ, and it also allows for the correlation coefficient ρ of the Brownian motions of price and volatility to be non-zero.

From a time perspective, the quadrinomial tree performs slower than the tree-based method. However, if it is assumed that the situation allows the filtered stochastic volatility distribution may be estimated in advance, then the methods are approximately even, if the tree-based method uses $m_x = 500$, $m_y = 24$, and $m = 50$. As shown in table 4.7, the difference between the accuracy of those values and the next are less than 0.004 with the error from the exact answer being on the same order of magnitude, it is safe to say that either method may be used with approximately the same execution time. However, if the distribution may not be calculated in advance, it is clear that the tree-based method is faster.

Conclusion

Option pricing will continue to play an important role in financial engineering. Stochastic volatility models have been and remain a popular method for pricing all types of options. The goal of this thesis was to implement and conduct numerical experiments with two stochastic volatility models and compare their performance. In this regard, it was found that the tree-based model performed faster than the quadrinomial tree. It was found that the quadrinomial tree seemed to be more accurate than the tree-based method for the IBM stock, but for the S&P 500 index the tree-based method seemed to be more accurate. Even though the accuracy of the parameters used in the quadrinomial tree was under question, it seems that if the two methods are given the same set of parameters, they give comparable results.

The main issue with the results section of this thesis is that the results of the quadrinomial tree method do not match the results of the paper by Florescu and Viens (2008). However, there was some doubt about the accuracy of their parameters, as the notation in the original paper was inconsistent the associated section. Additionally, the sources used for historical data may not have been the same, and so, the historical prices may have been different.

For the quadrinomial tree method, it was shown that the way the tree is implemented has a significant effect on the speed of the algorithm, and that for certain grids, it may be faster to calculate all of the possible values within the grid instead of only those that are necessary. The thesis was also able to show that, for some models, it is possible to use the estimated parameters of another model with the correct transformation on the volatility process Y. Using this the accuracy and speed of the models can be directly compared.

Similarly to the three implementations of the quadrinomial tree that were discussed, future research could investigate how to optimize these methods for different types of exotic options which leverage the restrictions on their payoffs.

References

Baxter, Martin, Andrew Rennie. 1996. *Financial Calculus: An introduction to derivative pricing.* Cambridge: The Press Syndicate of the University of Cambridge.

Beckers, Stan. 1980. "The Constant Elasticity of Variance Model and Its Implications for Option Pricing." *The Journal of Finance* 35, no. 3: 661-73. http://www.jstor.org/stable/2327490.

Bingham, Nicholas, Kiesel Rüdiger. 2004. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives.* 2nd ed. London: Springer-Verlag. https://doi.org/10.1007/978-1-4471-3856-3.

Black, Fischer, and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81, no. 3: 637–54. http://www.jstor.org/stable/1831029.

Derman, Emanuel, and Michael B. Miller. 2016. *The Volatility Smile.* New Jersey: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119289258.

Florescu, Ionuţ, and Frederi G. Viens. 2008. "Stochastic volatility: option pricing using a multinomial recombining tree." *Applied Mathematical Finance* 15, no. 2: 151-81.

Fouque, Jean-Pierre, George Papanicolaou, and K. Ronnie Sircar. 2000. "Mean-reverting stochastic volatility." *International Journal of Theoretical & Applied Finance* 4: 23-39.

Glasserman, Paul. 2003. *Monte Carlo Methods in Financial Engineering.* New York: Springer. https://doi.org/10.1007/978-0-387-21617-1_2.

Heston, Steven L. 1993. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *The Review of Financial Studies* 6, no. 2: 327–43. http://www.jstor.org/stable/2962057.

Lim, Kian Guan. 2011. *Probability and Finance Theory.* 1st ed. Singapore: World Scientific Publishing Co. Pte. Ltd. https://doi.org/10.1142/7781.

Linetsky, Vadim, and Rafael Mendoza. 2010. "The constant elasticity of variance model." *Encyclopedia of Quantitative Finance*: 328-34.

Merton, Robert. 1973. "Theory of Rational Option Pricing." *The Bell Journal of Economics and Management Science* 4, no. 1 (Spring): 141-83.
http://www.jstor.org/stable/3003143.

Muller, Mervin E. 1959. "A comparison of methods for generating normal deviates on digital computers." *Journal of the ACM* 6, no. 3: 378-83.

Thomas, J. W. 1995. *Numerical Partial Differential Equations: Finite Difference Methods.* New York: Springer-Verlag.
https://doi.org/10.1007/978-1-4612-0569-2.

Vellekoop, Michael, Hans Nieuwenhuis. 2009. "A tree-based method to price American options in the Heston model" *The Journal of Computation Finance* 13, no. 1: 1-21.

Wilmott, Paul, Jeff Dewynne, and Sam Howison. 1994. *Option Pricing: Mathematical models and computation.* Oxford Financial Press.

Non-exclusive licence to reproduce thesis and make thesis public

I, Joseph Haske,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, Option Pricing Using Stochastic Volatility Models, supervised by Toomas Raus.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Joseph Haske
16/05/2022