TÉCNICO LISBOA

Numerical Methods for Stochastic Differential Equations

with Applications to Finance

Matilde Lopes Rosa

Thesis to obtain the Master of Science Degree in

Mathematics and Applications

Supervisor: Prof. Carlos José Santos Alves

Examination Committee

Chairperson: Prof. António Manuel Pacheco Pires
Supervisor: Prof. Carlos José Santos Alves
Members of the Committee: Prof. Svilen Stanimirov Valtchev

May 2016

Resumo

A precificação de derivados financeiros é usada por investidores com o objetivo de aumentar os retornos esperados e minimizar o risco associado a um investimento. As opções, em particular, oferecem benefícios como risco limitado e alavancagem, sendo alvo de muita pesquisa no sentido de desenvolver modelos que prevejam o preço correto das opções.

O modelo de precificação de opções mais celebrado é o modelo de Black-Scholes, mas sendo reduzido à equação do calor, usa algumas hipóteses restritivas, levando a precificações erradas. Assim, desenvolveram-se extensões do modelo de Black-Scholes, com variação da volatilidade e da taxa de juro. Nesta tese, revendo a abordagem clássica, com esquemas de diferenças finitas para a equação do calor, aplicamos depois métodos numéricos para equações diferenciais estocásticas - esquemas de Euler-Maruyama e Milstein, implementados em simulações de Monte Carlo, que serão o principal foco desta tese.

Começamos por considerar modelos de preços dos ativos em que a volatilidade e a taxa de juro são coeficientes que dependem do tempo e de seguida consideramos também coeficientes que dependem do preço do ativo. Verificando-se condições suficientes nos coeficientes, as aproximações numéricas vão convergir no sentido fraco e forte, para o preço das opções europeias. Finalmente, implementamos o modelo de Heston com volatilidade estocástica, para o qual também obteremos boas taxas de convergência.

Palavras-chave: EDE's, Euler-Maruyama, Milstein, modelo de Black-Scholes, Monte Carlo, diferenças finitas

Abstract

The pricing of financial derivatives is used to help investors to increase the expected returns and minimise the risk associated with an investment. Options, in particular, offer benefits such as limited risk and leverage and for this reason much research has been dedicated to the development of models that accurately price options. The most celebrated model for option pricing is the Black-Scholes model, however when reduced to the heat equation, the simplified assumptions lead to the mispricing of options. Thus, extensions of this model have been developed, that consider a variation of the volatility and of the interest rate. In this work, we review the classical approach using finite difference schemes for the heat equation, and then we apply numerical schemes for stochastic differential equations - Euler-Maruyama and Milstein schemes, using Monte Carlo simulations, which will be the main focus of this thesis.

We start by considering asset models where the volatility and the interest rate are time-dependent coefficients and then extend these models to coefficients that also depend on the price of the underlying asset. We will see that as long as sufficient conditions on the coefficients are satisfied, the numerical approximations will converge in weak and strong sense, to the price of European call options. We conclude with an implementation of the Heston model for stochastic volatility, showing that will also present good convergence rates.

Keywords: SDE's, Euler-Maruyama, Milstein, Black-Scholes model, Monte Carlo, finite differences

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

ODE Ordinary Differential Equation

PDE Partial Differential Equation

SDE Stochastic Differential Equation

Chapter 1

Introduction

First published by Fischer Black and Myron Scholes in 1973 [4], and developed further by Robert Merton in the same year [25], the Black-Scholes model remains one of the most widely used models in option pricing. The Black-Scholes model gives the theoretical value for European call and put options.

Despite its wide use, it relies on several assumptions on the market whose correctness is questionable. The most delicate of these assumptions is that the stock returns follow a normal distribution with constant volatility. Empirical studies (see for instance[24], [28]) show that the distribution of the stock returns have high peeks and longer tails than those of the normal distribution and that volatility is usually negatively correlated to the price of the underlying stock, tending to increase when asset prices decrease. In order to overcome these restrictions, several extensions to the Black-Scholes model have been proposed since.

A popular approach to account for the fact that volatility changes over time is to consider that its behaviour follows a stochastic process itself. The most celebrated model for stochastic volatility is Heston's (1993) [15], however this was not the first stochastic volatility model to be applied in the pricing of options. Before Heston's model there were other important works such as the Hull and White model in 1987 [18], Scott [31] or Wiggins [33].

In 1976, Merton was the first to apply a jump-diffusion in option pricing [26] to approximate the movement of stock prices subject to occasional discontinuous breaks, which addresses the issue of "long tails". In 2002 Kou [23] proposed a jump-diffusion model with a double exponential distribution which makes it possible to obtain analytical solutions for many path-dependent options.

Models that relax the assumption of constant interest rate have also been subject of analysis. In 1977, Vasicek's model [32] was the first to consider interest rates as mean reverting processes and later Hull and White [19] extended this model by introducing time-dependent parameters.

For most of these models there is not an analytical solution available and when there is, the formulas often exhibit integrals that become intractable or are just complicated and difficult to evaluate without resorting to numerical methods to find the price of the derivatives.

Finite difference methods are a popular approach in option pricing due to their easy implementation and flexibility. They work by replacing the continuous derivatives in the partial differential equation by finite differences. They were first used to compute numerical solutions for options by Schwartz [30] in 1987, and further developed by Courtadon [8] in 2002, who proposed a more accurate approximation. Finite difference methods are specially suitable for low dimension problems, where the payoff depends on a single underlying asset, such as the one dimensional Black-Scholes model and for the pricing of simple path-dependent options or early exercise options, like American options. However, the pricing of more complex path dependent options, multi-dimensional models or models with more complicated features,

such as time-dependent parameters, is better handled by Monte Carlo simulations.

The Monte Carlo method was first used in the pricing of European options by Boyle [6]. The essence of this method is to simulate the stochastic differential equation that describes the behaviour of the underlying asset to generate paths for the asset price, compute the payoff of each random path, average the payoffs and then discount them back to the present. The concept behind this rationale is that, according to the law of large numbers, the average over a large enough number of samples approximately equals their mean. One of its main disadvantages is its computational inefficiency when compared to other numerical methods, as it requires many sample paths to produce accurate results.

In this thesis we study numerical methods to approximate the value of options in extensions of the Black-Scholes model with time-dependent parameters. We will start by studying the classical Black-Scholes model and approximate de Black-Scholes PDE with finite difference schemes. Next we follow Higham [16] and consider numerical methods for SDE's, in particular the Euler-Maruyama and Milstein methods, which we will implement using Monte Carlo simulations. For these simulations we will start with the classical Black-Scholes asset model and later relax the assumption of constant volatility and interest rate and assume that the underlying asset follows a geometric Brownian motion with time-dependent coefficients. Finally we will assume stochastic volatility and simulate the Heston model.

More specifically, this work is organised as follows: Chapter 2 starts with the concept of option as a special case of a financial derivative. It then provides some background on probability theory and stochastic calculus and introduces the asset model which the Black-Scholes model uses.

In chapter 3 we will look at the classical Black-Scholes model, including the derivation of the Black-Scholes equation, its transformation into the heat equation and the derivation of an exact solution. We also provide some insight on the limitations of this model. The chapter concludes with some extensions to the classical Black-Scholes model that attempt to overcome some of its restrictions.

In chapter 4 we will look at how to solve the Black-Scholes equation using a finite difference approach by transforming the Black Scholes PDE into the heat equation. We will discuss three different finite differences schemes and provide the respective algorithms. We close the chapter with a brief discussion on the restrictions of these methods.

Chapter 5 discusses numerical methods to approximate stochastic differential equations. We will look at two different discretisation schemes and how to implement them in Monte Carlo simulations. Some theory about convergence, consistency and stability of the methods will be provided as well.

Chapter 6 shows the results of the implementation of the algorithms described in the two previous chapters.

Chapter 7 closes this thesis with a summary of the results obtained in the numerical experiments and gives a brief overview of topics of interest that are out of the scope of this work.

Chapter 2

Stochastic differential equations and asset pricing

2.1 Financial derivatives

As the name suggests, financial derivatives are contracts whose value is derived from the value of one or several underlying entities. These entities can be assets (cash, bonds, equity, indexes, commodities, interest rates and exchange rates) or events such as a corporation default, a natural disaster or even the weather. These contracts agree on the transaction of a specified amount of assets or cash flows at a previously agreed price or pricing procedure, at a specific date or dates or during specific periods of time or events. Financial derivatives are used for several purposes including risk management, hedging, speculation, arbitrage and access to assets or markets that are hard to trade. In this work we will focus on options (specifically European plain vanilla options), although other common derivatives such as futures and swaps are also common.

2.1.1 Options

An option is a contract that gives its holder the right, but not the obligation, to buy or sell the underlying asset S at an agreed price, the strike price K, at a specified time or during a specified period of time. The time until the expiration date is called the time-to-maturity. The main thing that distinguishes options from other financial derivatives is that the holder has no obligation of exercising the option.

Options exist mainly in two forms: call options and put options. A call option gives its holder the right to buy the underlying asset at a certain price within a specific period of time. A put option gives the holder the right to sell the underlying asset at a certain price within a specific period of time. We may have two kinds of positions in an option: a long position or a short position. Long positions are those that are owned and short positions are those that are owed. Therefore a long call gives the holder the right to buy whereas a short call gives the seller (or writer) the obligation to sell if the option is exercised. On the other hand, a long put gives the holder right to sell whereas a short put gives the writer the obligation to buy the option in case it is exercised.

There are several types of options but the two most commonly traded are American and European options. American options can be exercised at any time during the specified period of time that goes from the date of purchase to the expiration of the contract. In the mean time the holder can also choose between keeping the option and selling it. European options are more restrictive regarding the exercise date as they can only be exercised at the maturity. But in the mean time the holder can also choose to

keep the option or to sell it.

The type of options we have described in this section are the simplest, most standard version of options and are known as *vanilla options* or *plain vanilla options*, in analogy with the ice cream flavour vanilla, which is one of the most common flavours. In short, this term can be used to describe any standard option.

European Options

In this work we will restrict our attention to European options. We have seen that these options can only be exercised at the maturity, which means that the value of the option at expiration only depends on the price of the underlying asset at that date. This is why European options are easier to value than American options. The value of the option at maturity is called payoff.

For a European call the payoff function is given by

$$h(S_T) = \max\{S_T - K, 0\} \tag{2.1}$$

where h denotes the payoff of the option, S_T the stock price at the maturity and K the strike price. This means that if the price of the underlying asset is higher than the strike price of the option the holder of the option can make a profit of $S - K > 0$, therefore he will exercise the option. In this situation we say that the option is in-the-money. On the other hand, if the price of the underlying asset is below the strike price it means that the option would be worthless, as the holder would have to pay more than the market price to exercise the option. Therefore the holder would let the option expire without exercising it and its payoff would be zero. In this case the option is said to be out-of-the-money. Finally, if the price of the underlying asset equals the strike price of the option it would not make a difference whether or not the holder exercises the option and the option is said to be at-the-money.

For a European put the payoff function is given by

$$h(S_T) = \max\{K - S_T, 0\} \tag{2.2}$$

In this case, if the price of the underlying asset is higher than the strike price of the option, it would be useless for the holder to exercise the option, as he would be selling the underlying asset at lower price than the market price. If, however, the price of the underlying asset is below the strike price, the holder could make a profit of $K - S > 0$, as he would be selling the underlying asset at a price above the market price. Again, if the price of the underlying asset and the strike price are equal, it makes no difference what the holder decides.

2.1.2 Option value

There are several concepts which are crucial to value an option. The first is related with the time value of money, the idea that a given amount of money available in the present time is more valuable than the same amount available in future dates due to its potential earning capacity. If money can earn interest, any amount of money is worth more the sooner it is received, as it can be invested into more money in future instants.

The second is related with no arbitrage, the concept that assets must be appropriately priced and thus if a portfolio of assets can be replicated with another portfolio of assets with the same cash flows under all scenarios, the two portfolios must have the same price. In practice, if two portfolios of assets with the same value have different prices, the portfolio with the lowest price can be bought and the portfolio with the highest price can be sold, until both portfolios trade at the same price in the market. As a

consequence and in the presence of no arbitrage, investors cannot earn risk-free returns without taking risk.

An illustrative example is an important relationship in finance: the put-call parity that defines the relationship between the prices of a European call and a European put with the same underlying asset, strike price and maturity. Indeed, when a European call and a European put share the same underlying asset, strike price and maturity, the payoffs at maturity of a portfolio A composed by a European call and a bond with the value of the strike price and portfolio B, composed by a European put and the underlying stock are the same, that is, $C_T + K = P_T + S_T \Leftrightarrow Max(S_T - K, 0) + K = Max(K - S_T, 0) + S_T$.

As European options cannot be exercised before maturity, the time independent value of both portfolios must be equal to the discounted value of both portfolios or:

$$C_t + Ke^{-r(T-t)} = P_t + S_t \tag{2.3}$$

when r, the bond interest rate is assumed constant. Equation (2.3) is called Put-Call-Parity.

We have already discussed some important concepts to value an option, and in the previous section we had already seen the value of an option at the maturity. But the real goal when valuing an option is to know its price at the present time, and therefore choosing what rate to use to discount the value to the present, is an important task. The challenge that this poses has to do with the fact that different investors have different sensibilities to the risk. This would lead to different discount rates, which would account for each investor's preferences but unfortunately these are very hard to quantify. However, it is possible to derive arguments to overcome this. In this work we will look at two different approaches for pricing an option, which are equivalent. The first is based on the replication of the underlying asset, and we will look into it when we derive the classical Black-Scholes equation. The other uses risk neutral probabilities, which we will discuss when we introduce the Monte Carlo method.

2.2 Stochastic essentials

2.2.1 Basic concepts

In this section we will define some important basic concepts of probability and stochastic calculus. The following definitions are based on the works [27] and [3].

Definition 2.2.1. (ς-field). A ς-field $\mathcal{F}$ (on Ω) is a collection of subsets of Ω satisfying the following conditions:

- It is not empty: $\emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$
- If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
- If $A_1, A_2, ... \in \mathcal{F}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ and $\cap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 2.2.2. (Probability measure). A function that assigns a number $P(A)$ to each set A in the ς-field $\mathcal{F}$ is called a probability measure on $\mathcal{F}$, provided that the following conditions are satisfied:

- $P(A) \geq 0$, for every $A \in \mathcal{F}$
- $P(\Omega) = 1$
- If $A_1, A_2, ...$ are disjoint sets in $\mathcal{F}$ then $P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$

Definition 2.2.3. (Probability space). A probability space is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ where: Ω is a set, $\mathcal{F}$ is a ς-field of subsets of Ω, and P is a probability measure on $\mathcal{F}$.

Definition 2.2.4. (Random variable). A random variable on the probability space $(\Omega, \mathcal{F}, P)$ is a real-valued fuction X such that for all real a and b, the sets $\{\omega : a \leq X(\omega) \leq b\}$ are in $\mathcal{F}$.

Definition 2.2.5. (Probability distribution function). The collection of probabilities

$$F_X(x) = P(X \leq x) = P(\{\omega : X(\omega) \leq x\}), \quad x \in \mathbb{R} = (-\infty, \infty)$$

is the probability distribution function F_X of X.

Most continuous distributions of interest have a density f_X:

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy, \quad x \in \mathbb{R} \tag{2.4}$$

where $f_X(x) \geq 0$ for every $x \in \mathbb{R}$ and $\int_{-\infty}^{\infty} f_X(y)dy = 1$

Definition 2.2.6. (Expectation, variance and moment). The expectation of a random variable X with density f_X is given by

$$\mu = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{2.5}$$

The variance of X is defined as

$$\sigma^2 = var(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx \tag{2.6}$$

The k-th central moment of X $(k > 0)$ is defined as

$$E\left[(X-\mu)^k\right] = \int_{-\infty}^{\infty} (x-\mu)^k f_X(x)dx \tag{2.7}$$

Definition 2.2.7. (Skewness and kurtosis). The skewness of X is the third standardised central moment of X

$$skew(X) = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \int_{-\infty}^{\infty} \frac{(x-\mu)^3}{\sigma^3} f_X(x)dx \tag{2.8}$$

The kurtosis of X is the fourth standardised central moment of X

$$kurt(X) = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \int_{-\infty}^{\infty} \frac{(x-\mu)^4}{\sigma^4} f_X(x)dx \tag{2.9}$$

Definition 2.2.8. (Stochastic process). A stochastic process X is a collection of random variables $(X_t, t \in T) = (X_t(\omega), t \in T, \omega \in \Omega)$, defined on Ω and T is an interval if X is accontinuous-time and a finite coutable set if X is a discrete-time process.
A stochastic process is a function of two variables. For a fixed instant of time t, it is a random variable: $X_t = X_t(\omega)$, $\omega \in \Omega$. For a fixed outcome $\omega \in \Omega$, it is a function of time: $X_t = X_t(\omega)$, $t \in T$. This function is called a sample path of the process X.

Definition 2.2.9. (Conditional expectation). A random variable Z is called the conditional expectation of X given the ς-field $\mathcal{F}$ (we write $Z = E(X|\mathcal{F})$) if

- Z does not contain more information than that contained in $\mathcal{F}$: $\varsigma(Z) \subset \mathcal{F}$
- Z satisfies the relation $E(XI_A) = E(ZI_A)$ for all $A \in \mathcal{F}$

Definition 2.2.10. (Brownian motion). A stochastic process $W = (W_t, t \in [0, \infty])$ is called a Brownian motion or a Wiener process if the following conditions are satisfied:

1. It starts at zero: $W_0 = 0$ with probability 1
2. it has independent and stationary increments (see [[27], pp. 30] for the definition)
3. $W_t - W_s \sim \mathcal{N}(0, t-s)$, for $0 \leq s \leq t$
4. it has continuous sample paths: "no jumps".

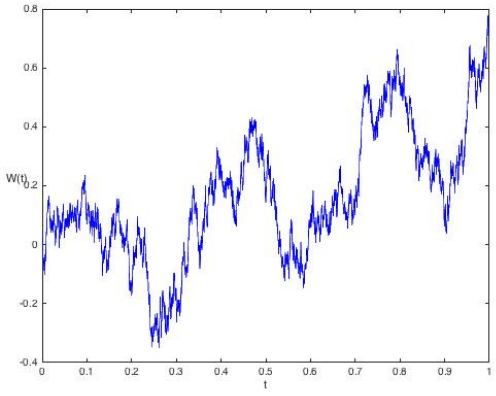


Figure 2.1: Discretised Brownian path

Definition 2.2.11. (Geometric Brownian motion). A stochastic process X_t is said to follow a geometric Brownian motion if it satisfies the following stochastic differential equation[1]

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{2.10}$$

where μ and σ are constants and W_t is a Brownian motion. The coefficients μX_t and σX_t are called the drift and diffusion coefficients, respectively.

Theorem 2.2.1. (Strong law of large numbers). Let $R_1, R_2, ...$ be independent random variables on a given probability space. Assume uniformly bounded fourth central moments, that is, for some positive real number M, $E[|R_i - E(R_i)|^4] \leq M$, for all i. Let $S_n = R_1 + ... + R_n$. Then[2]

$$\frac{S_n - E(S_n)}{n} \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty$$

In particular, if $E(R_i) = m$ for all i, then $E(S_n) = nm$; hence

$$\frac{S_n}{n} \xrightarrow{a.s.} m$$

Theorem 2.2.2. (Central limit theorem). For each n, let $R_1, R_2, ..., R_n$ be independent random variables on a given probability space. Assume that the R_i all have the same density function f with finite mean m and and finite variance $\sigma^2 > 0$, and finite third moment as well. Let $S_n = R_1 + ... + R_n$ and

$$T_n = \frac{S_n - E(S_n)}{\sigma(S_n)}$$

(where $\sigma(S_n)$ is the standard deviation of S_n) so that T_n has mean 0 and variance 1. Then $T_1, T_2, ...$ converge in distribution to a random variable that is normal with mean 0 and variance 1.

[1] The definition of stochastic differential equation is introduced in the next section
[2] a.s. stands for almost surely

Definition 2.2.12. (Filtration). Let $(\mathcal{F}_t, t \geq 0)$ be a collection of ς-fields on Ω. $(\mathcal{F}_t, t \geq 0)$ is called a filtration if $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s \leq t$.

Definition 2.2.13. (Adapted process). The stochastic process $\{Y_t, t \geq 0\}$ is said to be adapted to the filtration $(\mathcal{F}_t, t \geq 0)$ if $\varsigma(X_t) \subset \mathcal{F}_t$ for all $t \geq 0$.

Definition 2.2.14. (Martingale). A stochastic process $(X_t, t \geq 0)$ is called a continuous-time martingale with respect to the filtration $(\mathcal{F}_t, t \geq 0)$, if

1. $E|X_t| < \infty$, for all $t \geq 0$
2. X is adapted to $(\mathcal{F}_t)$
3. $E(X_t|\mathcal{F}_s) = X_s$ for all $0 \leq s < t$

The following definition is based on [21].

Definition 2.2.15. (Risk neutral measure). Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with filtration $\{\mathcal{F}_t, t = 0, 1, ..., T\}$, a probability measure $\mathbb{Q}$ is said to be risk neutral if

1. $\mathbb{Q}$ is equivalent to $\mathbb{P}$, that is, $\mathbb{P}(A) > 0$ if and only if $\mathbb{Q}(A) > 0$, $\forall A \in \mathcal{F}_t$
2. The price of every security in the market (using the risk free rate as numeraire) is a martingale with respect to probability measure $\mathbb{Q}$

2.2.2 Stochastic integration

In order to understand Stochastic differential equations we must first introduce the notion of Ito integral. This is important because in this work we will define the SDE's in terms of this integral, altough they can also be defined using Stratonovich integral (see remark 2.2.3).

Consider the following integral with respect to the Brownian motion W_s

$$I(X) = \int_0^T X(s)dW_s \tag{2.11}$$

where $X(s)$ is a function or a stochastic process defined on $[0, T]$.

We have seen that the sample paths of a Brownian motion are continuous but almost surely nowhere differentiable, which means that the integral defined in (2.11) cannot be a Riemann or Lebesgue integral because there is no process $W'_t = \frac{dWt}{dt}$. Moreover, as the sample paths are not of bounded variation[3] it cannot be interpreted as a Riemann–Stieltjes integral either. Althout we cannot define the integral (2.11) on a path-by-path basis in the Riemann-Stieltjes sense we can definite it as the limit of a sequence of Riemann-Stieltjes sums. This leads us to concept of Itô integral.

Definition 2.2.16. (Itô's integral). Let X be a stochastic process locally square-integrable, i.e. $\int_0^T E[X_t^2]dt < \infty$ and adapted to the filtration generated by the Brownian motion W_t in $[0, T]$. If $0 = t_0 < ... < t_n = T$ is a partition of the interval $[0, T]$, the Itô integral can be defined as

$$\int_0^T X_t dW_t = \lim_{n\to\infty} \sum_{i=1}^{n} X_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) \tag{2.12}$$

[3] Note that a deterministic function f is of bounded variation if it can be expressed as the difference between two non-decreasing functions, i.e. $\sup_\tau = \sum_{i=1}^n |f(t_i) - f(t_{i-1})| < \infty$

The idea behind this definition is that by requiring X_t to be an adapted process, at time t X only depends on information up to that instant. Because of this the Itô integral must always be evaluated at the left end point of the interval, while in Riemann integrals one can choose from any point of the interval.

Remark 2.2.3. It is also possible to use another definition of stochastic integral, introduced by Stratonovich (see [27]). The definition can be found in appendix A. However, in this work we will only consider the Itô integral definition as it is more intuitive from a financial point of view. As we will see in the next chapter, the Black-Scholes model assumes that there are no arbitrage opportunities and this property is satisfied as the Itô integral is a martingale.

Now that we have introduced the Itô integral we can introduce the stochastic chain rule, which is very useful to derive solutions of SDE's. The following lemma can be found in [5].

Lemma 2.2.4. (Itô's lemma. Univariate case). If $X(t)$ is a stochastic process that satisfies $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$ and $f \in C^2$, then $Y(t) = f(X(t))$ is also a stochastic process that satisfies

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu(t, X_t)\frac{\partial f}{\partial X} + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2 f}{\partial X^2}\right)dt + \sigma(t, X_t)\frac{\partial f}{\partial X}dW \tag{2.13}$$

where all the partial derivatives of f are evaluated at (t, X_t).

2.2.3 Stochastic differential equations

Stochastic differential equations are differential equations that have at least one term that is a stochastic process. If we consider an ordinary differential equation

$$\frac{dX(t)}{dt} = a(t, X(t)), \quad X(0) = X_0 \tag{2.14}$$

the easiest way to introduce randomness is to consider the initial codition as a random process. Then the solution of equation (2.14) would become a stochastic process $(X_t, t \in [0, T])$. These types of equations are called random differential equations and can be solved as deterministic differential equations. However, we are interested in another way of introducing randomness in equation (2.14). This is achieved by adding a random noise term. Equation (2.14) then becomes

$$dX(t, \omega) = a(t, X(t, \omega))dt + b(t, X(t, \omega))dW(t), \quad X(0, \omega) = X_0 \tag{2.15}$$

where ω denotes that $X = X(t, \omega)$ is a random variable, a and b are deterministic functions and $W = (W_t, t \geq 0)$ is a Brownian motion. Therefore the randomness of the solution X is due both to the randomness in the initial condition and the Brownian motion. From now on we will drop the ω from the notation and will write equation (2.15) as

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \quad X(0) = X_0 \tag{2.16}$$

Equation (2.16) is the differential form of the equation

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s, \quad 0 \leq t \leq T \tag{2.17}$$

where the first integral on the right-hand side is a Riemann integral, and the second one is an Itô stochastic integral.

There are two main types of solutions to a SDE, weak and strong solutions. In this work we will focus on the second type and provide a theorem for the existence and uniqueness of a strong solution to a SDE, which can be found in Kloeden and Platen [22].

Theorem 2.2.5. (Existence and uniqueness). Let a and b be two functions defined on $\mathbb{R}^n \times [t_0, T]$ that satisfy the following conditions

1. (Measurability): $a = a(t, x)$ and $b = b(t, x)$ are jointly L^2-measurable in $(t, x) \in [t_0, T] \times \mathbb{R}$
2. (Lipschitz condition): There exists a constant $K > 0$ such that $|a(t, x) - a(t, y)| \leq K|x - y|$ and $|b(t, x) - b(t, y)| \leq K|x - y|$, for all $t \in [t_0, T]$ and $x, y \in \mathbb{R}$
3. (Linear growth bound): There exists a constant $K > 0$ such that $|a(t, x)^2| \leq K^2(1 + |x|^2)$ and $|b(t, x)^2| \leq K^2(1 + |x|^2)$, for all $t \in [t_0, T]$ and $x, y \in \mathbb{R}$
4. (Initial condition): X_{t_0} is $\mathcal{F}$-measurable with $E\left(|X_{t_0}|^2\right) < \infty$

Then the stochastic differential equation (2.16) has a pathwise unique strong solution X_t on $[t_0, T]$ with

$$\sup_{t_0 \leq t \leq T} E\left(|X_{t_t}|^2\right) < \infty$$

.

2.3 Asset price model

In this section we will look at the stochastic differential equation that describes the price movement of an asset that is based on a Brownian motion. The asset price is the price of the underlying stock. We will consider the model suggested by Samuelson and later used by Black and Scholes. In this model the stock price S_t evolves according to the stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.18}$$

which is called a geometric Brownian motion and where μ is the constant expected return on the stock, and $\sigma > 0$ is the constant volatility of the stock.

Usually instead of the stock price itself we consider the stock return, which can be defined as the change in the price divided by the stock original value, i.e. $\frac{dS}{S}$. Equation (2.18) can be written in terms of the stock return as:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{2.19}$$

Figure 2.2 shows a sample path of a geometric Brownian motion over $[0, 1]$ with $N = 5000$ and time step $\delta t = \frac{1}{N}$ described by equation (2.22).

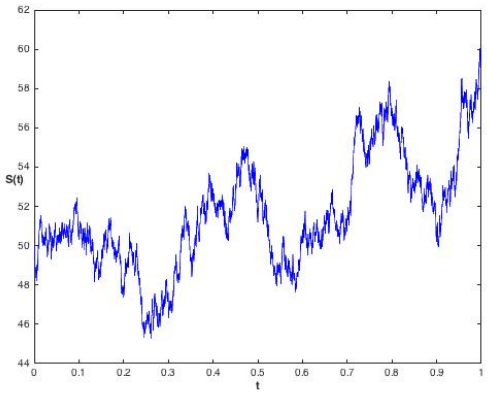


Figure 2.2: A sample path of a geometric Brownian Motion defined by the stochastic differential equation (2.22) with $S_0 = 50$, $\mu = 0.02$ and $\sigma = 0.25$

The idea behind this model is that there are two parts responsible for the evolution of the stock return. One is a deterministic component that contributes μdt in time dt. The other accounts for the random changes in the asset prices which is the component σdW_t.

It is possible to derive an explicit solution to equation (2.19). We compute $d \log S_t$ by applying lemma 2.2.4. with $f(t, S) = \log S$, $\mu(t, S) = \mu S$ and $\sigma(t, S) = \sigma S$:

$$d \log S_t = \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dW_t \tag{2.20}$$

Integrating with respect to t yields

$$\log S_t - \log S_0 = \left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W_t \tag{2.21}$$

The stock price is then given explicitly by

$$S_t = S_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t} \tag{2.22}$$

which means that the return has a lognormal distribution or the logarithm of return has a normal distribution

$$\log \left(\frac{S_t}{S_0} \right) \sim \mathcal{N} \left[\left(\mu - \frac{1}{2}\sigma^2 \right) t, \sigma^2 t \right] \tag{2.23}$$

Chapter 3

Black-Scholes model

The Black-Scholes model is one of the most popular models for pricing European call and put options. In this chapter we will look at this model in its original formulation, key assumptions, limitations and extensions.

3.1 Classic Black-Scholes equation

The Black-Scholes equation was first introduced by Fischer Black and Myron Scholes in the early 1970's and describes the price evolution of financial instruments such as stocks, which can be used to determine the price of European call and put options.

3.1.1 Assumptions

In order to derive the Black-Scholes equation we first have to look at some assumptions:

1. The price of the underlying asset follows a lognormal distribution, with a constant expected return and volatility, which arises from equation (2.18)
2. Short-selling is possible
3. There are no transaction costs regarding the purchase or selling of the underlying asset or the option
4. The underlying asset pays no dividend.
5. There are no riskless arbitrage opportunities
6. Trading of the underlying stock is continuous
7. The risk-free interest rate is constant during the entire period

3.1.2 Derivation of the Black-Scholes equation

We start by considering the classical Black-Scholes formulation as described for instance in Hull [17]. The following notation concerns the Black-Scholes model with constant parameters:

Table 3.1: Notation for the classical Black Scholes model

S_t	stock price at time t
μ	expected return of the stock
σ	stock's volatility
W	Brownian motion
r	risk free rate
K	strike price

For this derivation we follow the steps in [13] using the approach based on the concept of delta hedging, that is, on the idea of eliminating (or reducing) a portfolio's exposure to the price of an underling asset. Before deriving the equation we will first see the argument behind the Black-Scholes equation. In order to do this there are some important financial concepts we must first introduce.

In [13] a *portfolio* is defined as a pair (a_t, b_t) of adapted processes specifying the number of units held at time t of the underlying asset and the riskless bond, respectively. Furthermore, if we let B_t be the value of the bond at time t and S_t be the value of the stock at time t, the *value of the portfolio* at time t is

$$\Pi_t = a_t S_t + b_t B_t \tag{3.1}$$

Now, if we construct the portfolio such that its value at the maturity almost surely equals the payoff $h(S_T)$ of the derivative

$$a_T S_T + b_T B_T = h(S_T) \tag{3.2}$$

the portfolio is said to be a *replicating portfolio*, because it replicates the value of the derivative at time T.

Moreover, we can say that a portfolio is a *self-financing* portfolio if its changes in value are due only to changes in stock and bond prices. This means that there are no cash flows coming in or out, for example in order to buy more units of stock we would have to sell bonds to pay for it. In differential form this is expressed as

$$d\Pi_t = a_t dS_t + b_t dB_t \tag{3.3}$$

Having been through the essential financial concepts we can now derive the Black-Scholes equation. The goal is to find a self-financing portfolio (a_t, b_t) that replicates the payoff of a derivative in order to hedge the risk. Here our derivative is a European option with payoff function $h(S_t)$. At this point it is not necessary to specify whether it is a call or a put option. We will denote the price of the option by $V(S_t, t)$. Despite not knowing if it is even possible to find such a function $V(S_t, t)$ we will assume that the second order derivative of V with respect to S and the first order derivative of V with respect to t must be continuous in the domain $D_V = \{(S, t) : S \geq 0, \ \ 0 \leq t \leq T\}$, so that we can apply the Itô lemma. Furthermore, in order to eliminate the possibility of arbitrage opportunities we must require that

$$a_T S_T + b_T B_T = V(S_t, t) \tag{3.4}$$

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

Assume now that the portfolio consists of N_t short positions of the option, with price $V(S_t, t)$, which we will simply denote by V_t, and a_t units of the underlying stock. The change in the value of the portfolio is then $a_t dS_t - N_t dV_t$. In order for the portfolio to be riskless, its change must be equal to the change in

a riskless asset, that is

$$a_t dS_t - N_t dV_t = r(a_t S_t - N_t V_t)dt \tag{3.5}$$

Recall that the Black-Scholes model assumes that the price of the underlying stock follows a geometric Brownian motion like (2.18). Since $V(S_t, t)$ is a function of S_t we can apply lemma 2.2.4, Itô's lemma, with $Y(t) = V(S_t, t)$, to find the value of dV in terms of r and σ. This yields

$$dV = \left(\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S\frac{\partial V}{\partial S}dW \tag{3.6}$$

where r and σ are constant parameters and W is a Wiener process.

Substituting equations (3.6) and (2.18) into equation (3.5) we get

$$\begin{aligned} &a_t(\mu S_t dt + \sigma S_t dW_t) - N_t\left[\left(\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S_t\frac{\partial V}{\partial S}dW_t\right] = \\ &= r(a_t S_t - N_t V_t)dt \end{aligned} \tag{3.7}$$

Since we want the portfolio to be riskless, the terms containing dW_t must vanish from the equation.

Proposition 1. Setting the coefficients of dW_t to zero yields

$$a_t = N_t\frac{\partial V}{\partial S} \tag{3.8}$$

and

$$a_t S_t(\mu - r) + rN_t V_t = N_t\left(\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right) \tag{3.9}$$

Proof. Regarding (3.8), we set

$$a_t\sigma S_t dW_t - N_t\sigma S_t\frac{\partial V}{\partial S}dW_t = 0$$

Rearranging

$$a_t\sigma S_t dW_t = N_t\sigma S_t\frac{\partial V}{\partial S}dW_t$$

Dividing by $\sigma S_t dW_t$ leads to (3.8).

Regarding (3.9), when the dW_t terms are eliminated we get

$$a_t\mu S_t dt - N_t\left(\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt = ra_t S_t dt - rN_t V_t dt$$

Dividing by dt and rearranging we get equation (3.9). □

Proposition 2. Substituting equation (3.8) into equation (3.9) we obtain the Black-Scholes PDE

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - rV = 0 \tag{3.10}$$

Proof.

$$N_t\frac{\partial V}{\partial S}S_t(\mu - r) + rN_t V_t = N_t\left(\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right) \Leftrightarrow \tag{3.11}$$

$$N_t\mu S_t\frac{\partial V}{\partial S} - N_t rS_t\frac{\partial V}{\partial S} + rN_t V_t - N_t\mu S_t\frac{\partial V}{\partial S} = N_t\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right) \tag{3.12}$$

Dividing by N_t and rearranging gives

$$\mu S_t \frac{\partial V}{\partial S} - \mu S_t \frac{\partial V}{\partial S} = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + rS_t \frac{\partial V}{\partial S} - rV_t$$

Noticing that the terms that have μ vanish and rearranging again leads to equation (3.10). □

Note that this choice of portfolio, made only of options and units of stocks, allows us to eliminate the risk to determine the price V of the option and the ratio $\frac{a_t}{N_t}$. In fact, equation (3.8) is called *Delta*, and gives the number of units of the underlying asset that must be used to replicate the option. Hence the strategy name, delta hedging.

It is also worth mentioning that the fact that the terms containing μ are eliminated during the derivation and therefore the Black-Scholes PDE does not contain any μ, means that the price of the options is independent of the investor's attitude towards risk. Since the price of the derivative does not depend on the risk preferences we can therefore use any set of risk preferences to value the option. In particular, we can assume that all investors are risk neutral, that is, that they always demand only the risk free rate of interest as the average expected return on investment and do not require a risk premium. This allows to take $\mu = r$. The argument for this is that as long as we can find a replicating portfolio for the option we can eliminate all the risk.

3.1.3 Boundary and final conditions for European options

In the previous section we showed how to derive the Black-Scholes PDE. Without any further conditions this equation has many different solutions. In option pricing it is essential that the price of the option is unique, otherwise arbitrage opportunities would be possible. In order to find a unique solution for the equation we must now define some boundary conditions. The conditions presented in this section are for vanilla European options. Through this work, whenever it is not necessary to specify whether the option is a call or a put we will use V to designate the value of the option. When a distinction is necessary we will use C for the value of a call option and P for the value of a put option.

We have already seen that the final condition for a call or a put option is just its payoff at the maturity. Hence the terminal value of a European call is given by

$$C(S,T) = \max\{S - K, 0\} \tag{3.13}$$

Moreover, the value of a right to buy an asset worth 0 is also 0, so we have the following boundary condition

$$C(0,t) = 0 \tag{3.14}$$

However, as the value of the asset becomes larger, the option is more likely to be exercised, and the value of the strike price becomes less important until it can actually be neglected. This gives us

$$C(S,t) \to S \ \ as \ \ S \to \infty \tag{3.15}$$

For a put option, recall that its value at the maturity is given by

$$P(S,T) = \max\{K - S, 0\} \tag{3.16}$$

Furthermore, when the value of the underlying asset is 0, the final payoff of a put will be equal to

the strike price K. So, to know the value of the put at time t we just need to discount K to the present, which yields

$$P(0,t) = Ke^{-r(T-t)} \tag{3.17}$$

Finally, as the value of the stock increases the option is less likely to be exercised, which means that it will expire worthless

$$P(S,t) \to 0 \ \ as \ \ S \to \infty \tag{3.18}$$

3.1.4 Black-Scholes transformation to the heat equation

The Black-Scholes equation can be transformed into the heat equation by simply making a change of variables. This can be very convenient as the heat equation and its solution have been subject of deep study. Therefore, one may want to transform the heat equation solution back to get the solution to the Black-Scholes equation. In what follows we will consider a European call only, because the results for a put option can be easily derived using put-call-parity.

Consider again the Black-Scholes PDE, given in (3.10). Since now we are considering the specific case of a European call, we will denote the price of the call by C, instead of V. This yields

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC = 0 \tag{3.19}$$

Note that the domain of C is $D_C = \{(S,t) : S > 0, \quad 0 \le t \le T\}$.

Further, consider the following change of variables

$$S = Ke^x, \quad t = T - \frac{2\tau}{\sigma^2}, \quad C(S,t) = Kv(x,\tau)$$

This change of variables ensures that the domain of the new variable $v(x,\tau)$ is $D_v = \{(x,\tau) : -\infty < x < \infty, \ \ 0 \le \tau \le \frac{\sigma 2}{2}T\}$, and using the chain rule we can obtain the partial derivatives of C

$$\frac{\partial C}{\partial t} = K\frac{\partial v}{\partial \tau}\frac{\partial \tau}{\partial t} = -\frac{1}{2}\sigma^2 K\frac{\partial v}{\partial \tau} \tag{3.20}$$

$$\frac{\partial C}{\partial S} = K\frac{\partial v}{\partial x}\frac{\partial x}{\partial S} = \frac{K}{S}\frac{\partial v}{\partial x} \tag{3.21}$$

$$\frac{\partial^2 C}{\partial S^2} = \frac{\partial}{\partial S}\left(\frac{K}{S}\frac{\partial v}{\partial x}\right) = -\frac{K}{S^2}\frac{\partial v}{\partial x} + \frac{K}{S}\frac{\partial^2 v}{\partial x^2}\frac{1}{S} = \frac{K}{S^2}\left(\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x}\right) \tag{3.22}$$

Substituting the partial derivatives (3.20) through (3.22) into equation (3.19) gives

$$-\frac{1}{2}\sigma^2\frac{\partial v}{\partial \tau} + rS\frac{K}{S}\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2 S^2\frac{K}{S^2}\left(\frac{\partial^2 v}{\partial x} - \frac{\partial v}{\partial x}\right) - rKv = 0 \tag{3.23}$$

Dividing equation (3.23) by K and setting $c = \frac{2r}{\sigma^2}$ we get the following equation

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + (c-1)\frac{\partial v}{\partial x} - cv \tag{3.24}$$

We now need an initial condition for the problem. Under this change of variables the initial condition corresponds to the terminal condition before the transformation, because at $t = T$ we have $\tau = 0$. The

initial condition (3.13) under the transformation of variables becomes

$$v(x,0) = \max\{e^x - 1, 0\} \tag{3.25}$$

In order for equation (3.24) to look more like the heat equation, the terms $\frac{\partial v}{\partial x}$ and v must vanish from the equation. This can be achieved by a change of variables. We set

$$v(x,\tau) = e^{\alpha x + \beta \tau} u(x,\tau) \tag{3.26}$$

for some constants α and β that we will have to determine. Substituting equation (3.26) into equation (3.24) and dividing by $e^{\alpha x + \beta \tau} u(x,\tau)$ yields

$$\beta u + \frac{\partial u}{\partial \tau} = \alpha^2 u + 2\alpha \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + (c-1)\left(\alpha u + \frac{\partial u}{\partial x}\right) \tag{3.27}$$

To eliminate u from equation (3.27) we just have to set $\beta = \alpha^2 + (c-1)\alpha - c$ and to eliminate $\frac{\partial u}{\partial x}$ we set $0 = 2\alpha + (c-1)$. This yields

$$\alpha = -\frac{1}{2}(c-1) \tag{3.28}$$

$$\beta = -\frac{1}{4}(c+1)^2 \tag{3.29}$$

Substituting equation (3.28) and (3.29) into equation (3.26) leads to

$$v(x,\tau) = e^{-\frac{1}{2}(c-1)x - \frac{1}{4}(c+1)^2 \tau} u(x,\tau) \tag{3.30}$$

where $u(x,\tau)$ satisfies the heat equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < \infty \quad \text{and} \quad \tau \in [0,T] \tag{3.31}$$

with the initial condition

$$u(x,0) = u_0(x) = \max\{e^{\frac{1}{2}(c+1)x} - e^{\frac{1}{2}(c-1)x}, 0\} \tag{3.32}$$

Note that the domain of u is D_v, which is a more restrictive domain than the usual domain of the heat equation: $x \in \mathbb{R}$ and $\tau \geq 0$. However, as the time of the option ends at the maturity T it is only natural that τ is bounded as well. On the other hand, despite S being positive, the domain of x is the whole real axis.

Earlier in this chapter, we have defined the boundary and final conditions for European call and put options. We will now define them again under this change of variables.

Proposition 3. The initial and boundary conditions for a European call option are given by

$$\begin{aligned} &u(x,0) = \max\{e^{\frac{1}{2}(c+1)x} - e^{\frac{1}{2}(c-1)x}, 0\} \\ &\lim_{x \to -\infty} u(x,\tau) = 0 \\ &\lim_{x \to +\infty} u(x,\tau) - e^{\frac{1}{2}(c+1)x + \frac{1}{4}(c+1)^2 \tau} = 0 \end{aligned} \tag{3.33}$$

Proof. We start by noting that $C(S,t) = Kv(x,\tau)$ so, plugging in equation (3.30), it yields

$$C(S,t) = Ke^{-\frac{1}{2}(c-1)x-\frac{1}{4}(c+1)^2\tau}u(x,\tau) \tag{3.34}$$

For the initial condition, considering (3.13), we get

$$\begin{aligned} C(S,T) = \max\{Ke^x - K, 0\} = Ke^{-\frac{1}{2}(c-1)x}u(x,0) \Leftrightarrow \\ u(x,0) = \max\{e^{\frac{1}{2}(c+1)x} - e^{\frac{1}{2}(c-1)x}, 0\} \end{aligned}$$

For the first boundary condition note that $C = 0 \Leftrightarrow u = 0$, so it is imediate that

$$\lim_{x\to-\infty} u(x,\tau) = 0$$

For the other boundary condition we have that $C(S,t) - S \to 0$ as $S \to +\infty$ is equivalent to

$$Ke^{-\frac{1}{2}(c-1)x-\frac{1}{4}(c+1)^2\tau}u(x,\tau) - Ke^x \to 0 \quad \text{as} \quad e^x \to +\infty$$

. Rearranging we get

$$\lim_{x\to+\infty} u(x,\tau) - e^{\frac{1}{2}(c+1)x+\frac{1}{4}(c+1)^2\tau} = 0$$

□

Proposition 4. The initial and boundary conditions for a European put option are given by

$$\begin{aligned} u(x,0) &= \max\{e^{\frac{1}{2}(c-1)x} - e^{\frac{1}{2}(c+1)x}, 0\} \\ \lim_{x\to-\infty} u(x,\tau) - e^{\frac{1}{2}(c-1)x+\frac{1}{4}(c-1)^2\tau} &= 0 \\ \lim_{x\to+\infty} u(x,\tau) &= 0 \end{aligned} \tag{3.35}$$

Proof. Similarly to the previous proof, we start by noticing that

$$P(S,t) = Ke^{-\frac{1}{2}(c-1)x-\frac{1}{4}(c+1)^2\tau}u(x,\tau) \tag{3.36}$$

Considering (3.16), the initial condition becomes

$$\begin{aligned} P(S,T) = \max\{K - Ke^x\} = Ke^{-\frac{1}{2}(c-1)x}u(x,0) \Leftrightarrow \\ \Leftrightarrow u(x,0) = \max\{e^{\frac{1}{2}(c-1)x} - e^{\frac{1}{2}(c+1)x}\} \end{aligned}$$

For the first boundary condition, using (3.17) and (3.36) leads to

$$\lim_{x\to-\infty} Ke^{-\frac{1}{2}(c-1)x-\frac{1}{4}(c+1)^2\tau}u(x,\tau) = Ke^{-r(T-t)}$$

Dividing by K and rearranging equation (3.37) we get

$$\lim_{x\to-\infty} u(x,\tau) = e^{-r(T-t)+\frac{1}{2}(c-1)x+\frac{1}{4}(c+1)^2\tau} \tag{3.37}$$

Noting that $(T-t) = \frac{2\tau}{\sigma^2}$ and recalling that $c = \frac{2r}{\sigma^2}$ leads to

$$\lim_{x\to-\infty} u(x,\tau) = e^{-\tau c+\frac{1}{2}(c-1)x+\frac{1}{4}(c+1)^2\tau}$$

Finally, rearranging the square leads to the result we wanted to prove.

For the other boundary condition note that $P = 0 \Leftrightarrow u = 0$, so it is immediate that

$$\lim_{x \to +\infty} u(x, \tau) = 0$$

□

3.1.5 Exact solution to the Black-Scholes equation

For European call and put options it is possible to compute the exact solution to the Black-Scholes formula. There is more then one way to arrive at this solution. One approach is to solve equation (3.10) subjected to the boundary conditions derived in the previous section. Another way is to use risk-neutral valuation. In this work we will present the proof for a European call using the first approach and we will follow [34].

Proposition 5. For a European call the exact solution is given by

$$C(S, t) = \mathcal{N}(d_1)S - \mathcal{N}(d_2)Ke^{-r(T-t)} \tag{3.38}$$

where $\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-z^2}{2}} dz$ is the normal cumulative distribution function and

$$d_1 = \frac{\log(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{3.39}$$

$$d_2 = \frac{\log(S/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{3.40}$$

Proof. Recall that the solution to the heat equation $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ with $-\infty < x < +\infty$, $\tau > 0$ and initial condition $u(x, 0) = u_0(x)$ is given by

$$u(x, \tau) = \int_{-\infty}^{\infty} u_0(s) f(x - s, \tau) ds \tag{3.41}$$

where $f(x - s, \tau) = e^{-(x-s)^2/4\tau}$ is the heat kernel. We can write equation (3.41) as

$$u(x, \tau) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} u_0(s) e^{-(x-s)^2/4\tau} ds \tag{3.42}$$

where $u_0(x)$ is given by (3.32).

To evaluate the integral we use the transformation of variables $y = \frac{(s-x)}{\sqrt{2\tau}}$ which means that $dy = \left(\frac{1}{\sqrt{2\tau}}\right) ds$. Making this substitution, equation (3.42) becomes

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_0 \left(y\sqrt{2\tau} + x\right) e^{-\frac{y^2}{2}} dy \tag{3.43}$$

Notice that the function u_0 is strictly positive when the argument is also strictly positive and zero otherwise. Since we are interested in evaluating the integral when $u_0 > 0$ we must require that $y > -\frac{x}{\sqrt{2\tau}}$. Equation (3.43) becomes

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{\frac{1}{2}(c+1)(x+y\sqrt{2\tau})} e^{\frac{-y^2}{2}} dy - \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{\frac{1}{2}(c-1)(x+y\sqrt{2\tau})} e^{\frac{-y^2}{2}} dy \tag{3.44}$$

We denote the first and second integrals I_1 and I_2, respectively, and we will evaluate them separately. To evaluate I_1 we start by completing the square in the exponent. The exponent is

$$\begin{aligned} I_1 &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{\frac{1}{2}(c+1)(x+y\sqrt{2\tau})} e^{\frac{-y^2}{2}} dy \\ &= \frac{e^{\frac{1}{2}(c+1)x}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{\frac{1}{4}(c+1)^2\tau} e^{-\frac{1}{2}\left(y-\frac{1}{2}(c+1)\sqrt{2\tau}\right)^2} dy \\ &= \frac{e^{\frac{1}{2}(c+1)x+\frac{1}{4}(c+1)^2\tau}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-\frac{1}{2}\left(y-\frac{1}{2}(c+1)\sqrt{2\tau}\right)^2} dy \end{aligned} \tag{3.45}$$

Now we make another change of variables with $z = y - \frac{1}{2}(c+1)\sqrt{2\tau}$, which means that $dz = dy$. Substituting in 3.45 yields

$$I_1 = \frac{e^{\frac{1}{2}(c+1)x+\frac{1}{4}(c+1)^2\tau}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}-\frac{1}{2}(c+1)\sqrt{2\tau}}^{\infty} e^{-\frac{1}{2}z^2} dz \tag{3.46}$$

Finally, using the symmetry of the integral, equation (3.46) can be expressed as the cumulative distribution function for the normal distribution. If we denote

$$d_1 = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(c+1)\sqrt{2\tau} \tag{3.47}$$

we can write (3.46) as

$$I_1 = e^{\frac{1}{2}(c+1)x+\frac{1}{4}(c+1)^2\tau}\mathcal{N}(d_1) \tag{3.48}$$

where

$$\mathcal{N}(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{-\frac{1}{2}z^2} dz \tag{3.49}$$

is the cumulative distribution function for the normal distribution.

The evaluation of I_2 is identical except that $(c+1)$ is replaced by $(c-1)$ throughout.

Going back to equation (3.44) we get

$$u(x,\tau) = e^{\frac{1}{2}(c+1)x+\frac{1}{4}(c+1)^2\tau}\mathcal{N}(d_1) - e^{\frac{1}{2}(c-1)x+\frac{1}{4}(c-1)^2\tau}\mathcal{N}(d_2) \tag{3.50}$$

where d_1 is given as in (3.47) and $d_2 = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(c-1)\sqrt{2\tau}$. Now we must transform back the variables using $x = \log\left(\frac{S}{K}\right)$, $\tau = \frac{1}{2}\sigma^2(T-t)$ and $C(S,t) = Kv(x,\tau)$. Plugging this change of variables in equation (3.50) yields

$$C(S,t) = \mathcal{N}(d_1)S - \mathcal{N}(d_2)Ke^{-r(T-t)}$$

where d_1 and d_2 are give in (3.39) and (3.40), respectively.

□

Proposition 6. For a European put the exact solution is given by

$$P(S,t) = \mathcal{N}(-d_2)Ke^{-r(T-t)} - \mathcal{N}(-d_1)S \tag{3.51}$$

where d_1 and d_2 are given by (3.39) and (3.40), respectively.

Proof. The proof is straightforward using the put-call-parity (2.3).

$$
\begin{aligned}
P(S,t) &= C(S,t) - S + Ke^{-r(T-t)} = \\
&= \mathcal{N}(d_1)S - S + Ke^{-r(T-t)} - \mathcal{N}(d_2)Ke^{-r(T-t)} = \\
&= -(1-\mathcal{N}(d_1))S + (\mathcal{N}(d_2))Ke^{-r(T-t)} = \\
&= -\mathcal{N}(-d_1)S + \mathcal{N}(-d_2)Ke^{-r(T-t)} = \\
&= \mathcal{N}(-d_2)Ke^{-r(T-t)} - \mathcal{N}(-d_1)S
\end{aligned}
$$

□

3.1.6 Black-Scholes model limitations

We have seen that the Black-Scholes model is based on several assumptions, which do not hold in the real world. The model assumes that the stock returns follow a normal distribution, which is the same as saying that the stock prices themselves follow a lognormal distribution. However, in reality this is never the case. Real world distributions exhibit skewness and kurtosis, i.e. have heavier tails than those of the normal distribution. This means that extreme market movements are more likely to occur than as suggested by the normal distribution, which leads to the Black-Scholes model underpricing or overpricing an option.

Another implication of the Brownian motion is that volatility should be known and constant in time. This, however, does not happen in practice, and in fact volatility is the only parameter in the Black-Scholes model that cannot be directly observed from the market. The most common way to estimate volatility for a given underlying asset is to plug in the market price of the options and other observed parameters in Black-Scholes model to get the implied volatility. The implied volatility is the volatility of the underlying asset that when substituted in the Black-Scholes option pricing equation would equate the market option price. For a call option, it can be written as

$$
C(S_t, K, \tau, r, \sigma_{imp}(K,\tau)) = C_t^*(K,\tau) \tag{3.52}
$$

where $C(S_t, K, \tau, r, \sigma_{imp}(K,\tau))$ is the Black-Scholes call price and $C_t^*(K,\tau)$ is the market option price at time t.

If volatility were constant, as assumed in the Black-Scholes model, then for a fixed maturity, implied volatility should be constant regardless of the strike price. What happens in reality is that most markets have persistent patterns of volatilities varying by strike price. In some markets, like currency markets, when plotting the implied volatilities against the strike prices we get an U-shaped curve which resembles a smile, hence this particular pattern is called volatility smile. In equity markets (options on stock), the pattern is usually downward sloping and referred to as volatility skew. In this case options out-of-the-money have lower implied volatilities than options at-the-money. Furthermore, if we consider the dynamics between the implied volatility and the stock price, empirical studies show that they are negatively correlated.

Another unrealistic assumption is that of the interest rate being constant. The interest rate plays an important role in the pricing of the option as it is used to discount the payoff of the option to the present so it is crucial to take into account the behaviour of the interest rate when building an option pricing model.

3.2 Extensions to the Black-Scholes model

In order to overcome the Black-Scholes model limitations, many models have been proposed over time. As it is not possible to capture every aspect of the markets in a single model, we will focus on some important constraints. We will start by relaxing the Black-Scholes assumption of constant volatility and interest rate by allowing them to be time-varying deterministic functions. Later we will also consider that volatility follows its own stochastic process, using the model developed by Heston.

3.2.1 Time-dependent parameters

We will start by assuming that both the volatility and the expected return on the stock are continuous deterministic functions of time. The stock price dynamics are similar to the ones of the Black-Scholes model but now the stock price follows the stochastic differential equation

$$dS_t = \mu(t)S_t dt + \sigma(t)S_t dW_t \tag{3.53}$$

where σ and μ are continuous deterministic functions.

To obtain the explicit solution to equation (3.53) we proceed as in the previous sections. Here we compute $d\log S_t$ by applying lemma 2.2.4. with $f(t,S) = \log S$, $\mu(t,S) = \mu(t)S$ and $\sigma(t,S) = \sigma(t)S$:

The stock price is then given explicitly by

$$S_T = S_t e^{\int_t^T (\mu(x) - \frac{1}{2}\sigma^2(x))dx + \int_t^T \sigma(x)dW_x} \tag{3.54}$$

which means that even when both the interest rate and volatility are functions of time, the log of the stock returns has a normal distribution

$$\log\left(\frac{S_T}{S_t}\right) \sim \mathcal{N}\left[\left(\bar{\mu} - \frac{1}{2}\bar{\sigma}^2\right)(T-t), \bar{\sigma}^2(T-t)\right] \tag{3.55}$$

where $\bar{\mu} = \frac{1}{T-t}\int_t^T \mu(x)dx$ and $\bar{\sigma}^2 = \frac{1}{T-t}\int_t^T \sigma^2(x)dx$.

Like in the time-dependent volatility model, the derivation of the Black-Scholes equation is similar to the one with constant parameters. The following result can be found in Wilmott [34].

Proposition 7.

$$\frac{\partial V}{\partial t} + r(t)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2(t)S^2\frac{\partial^2 V}{\partial S^2} - r(t)V = 0 \tag{3.56}$$

Proof. Again we assume that the portfolio consists of N_t short positions of the option, with price $V(t,S_t)$, which we will simply denote by V_t, and a_t units of the underlying stock. The change in the value of the portfolio is then $a_t dS_t - N_t dV_t$. In order to the portfolio to be riskless, its change must be equal to the change in a riskless asset, that is

$$a_t dS_t - N_t dV_t = r(t)(a_t S_t - N_t V_t)dt \tag{3.57}$$

Note that equation (3.57) is the same equation as (3.5) exept that now r is a function of time. Indeed, the crucial step in this derivation is to notice that the expected return on the stock, μ is a function of the risk-free rate, plus a component that depends on the investor's sensivity to risk, both of which are not constant in the real world. So, if we assume that the expected return on stock is time-dependent we also have to assume that the risk-free rate is time-dependent, hence the result above.

Since $V(t,S_t)$ is a function of S_t and S_t follows the geometric Brownian motion described by equation

(3.54), we are still in the conditions of lemma 2.2.4, which allows for the coefficients to be functions of time. Therefore we can apply lemma 2.2.4, Itô's lemma with $Y(t) = V(t, S_t)$, to find the value of dV in terms of $\mu(t)$ and $\sigma(t)$. This yields

$$dV = \left(\frac{\partial V}{\partial t} + \mu(t) S \frac{\partial V}{\partial S} + \frac{\sigma^2(t)}{2} S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma(t) S \frac{\partial V}{\partial S} dW \tag{3.58}$$

where μ and σ are time-dependent functions and W is a Wiener process. Substituting equations (3.58) and (3.53) into equation (3.57) we get

$$\begin{aligned} &a_t(\mu(t)S_t dt + \sigma(t)S_t dW_t) - N_t \left[\left(\frac{\partial V}{\partial t} + \mu(t) S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2(t) S_t^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma(t) S_t \frac{\partial V}{\partial S} dW_t \right] = \\ &= r(t)(a_t S_t - N_t V_t) dt \end{aligned} \tag{3.59}$$

Since we want the portfolio to be riskless, dW_t must be set to zero. Setting $dW_t = 0$ yields

$$a_t = N_t \frac{\partial V}{\partial S} \tag{3.60}$$

and

$$a_t S_t \left[\mu(t) - r(t) \right] + r(t) N_t V_t = N_t \left(\frac{\partial V}{\partial t} + \mu(t) S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2(t) S_t^2 \frac{\partial^2 V}{\partial S^2} \right) \tag{3.61}$$

just like in the case of constant parameters. Finally, substituting equation (3.60) into equation (3.61) we obtain the Black-Scholes PDE for time-dependent coefficients

$$\frac{\partial V}{\partial t} + r(t) S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2(t) S^2 \frac{\partial^2 V}{\partial S^2} - r(t) V = 0 \tag{3.62}$$

□

Remark 3.2.1. Note that in this model, when pricing a put option, we must account for the fact that r appears explicitly in equation (3.17). So, when r is a function of time, the correct condition is

$$P(0, t) = K e^{-\int_t^T r(\tau) d\tau} \tag{3.63}$$

Having seen that by considering a geometric Brownian motion with continuous time-dependent coefficients we can derive similar results to those with constant parameters we are now interested in analysing the case where the coefficients have jump discontinuities. This is of interest because in the real world there are often jumps in the volatility or interest rate of the stock, which are not accounted for in the classical Black-Scholes model.

Consider again equation (3.53) but now assume that $r(t)$ and $\sigma(t)$ may have some discontinuity points, which correspond to the jumps, and are continuous elsewhere. We can see from theorem 2.2.5 that without Lipschitz continuous coefficients we cannot guarantee the existence and uniqueness of a solution to the SDE.

3.2.2 Stochastic Volatility - Heston model

We now consider the case where the variance of the stochastic process is a random variable itself. Although there are several models for stochastic volatility, in this work we will focus exclusively on the Heston model.

The Heston model extends the Black-Scholes model as it assumes that volatility is a mean-reverting

stochastic process. This is an important assumption because in financial markets volatility will tend to move to the average volatility over time, otherwise we could have assets with volatility going near zero or tending to infinity. This model also accounts for the fact that the stock returns follow a non-Gaussian distribution and for the leverage effect, i.e. negative correlation between asset prices and volatility. For these reasons, the Heston model has become a benchmark among the stochastic volatility models for pricing options. Its popularity also comes from the fact that it is possible to compute a closed form solution of a European call which is easy to evaluate except for an integral, which has to be numerically evaluated.

The price dynamics of the Heston model are similar to those of the Black-Scholes model but they also include a stochastic behaviour for the volatility process. The price and variance dynamics are given by

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dW_t^S \tag{3.64}$$

$$d\nu_t = k(\theta - \nu_t)dt + \xi\sqrt{\nu_t} dW_t^\nu \tag{3.65}$$

$$dW_t^S dW_t^\nu = \rho dt \tag{3.66}$$

where the parameters are shown in table 3.2.

Table 3.2: Notation for the Heston model

$S(t)$	stock price at time t
$\nu(t)$	instantaneous variance
μ	rate of return of the asset
θ	long-run mean
k	rate at which V_t reverts to θ
ξ	volatility of volatility

and W_t^S and W_t^ν are Wiener processes with correlation ρ.

As we did for the Black-Scholes model, we will now derive the partial differential equation of the Heston model. This derivation is based on the derivations done in [13] and [29]. The derivation of the Black Scholes PDE is similar to the one for the constant parameters. However, in order to have a riskless portfolio we now need another derivative asset written on the same underlying asset to account for the new source of randomness introduced by the stochastic volatility. The two derivatives differ by the maturity date or the strike price.

Consider a portfolio Π consisting of one short position of option V, a_t shares of the stock S and M_t long positions of the other option U. The value of the portfolio is

$$\Pi_t = a_t S_t + M_t U - V \tag{3.67}$$

Assuming that the portfolio is self-financing, the change in the portfolio value is

$$d\Pi_t = a_t dS + M_t dU - dV \tag{3.68}$$

Again we want a riskless portfolio, so the change in its value must equal the change in a riskless asset,

that is

$$a_t dS + M_t dU - dV = r(a_t S + M_t U - V)dt \tag{3.69}$$

Note that $V(S_t, t, \nu_t)$ and $U(t, S_t, \nu_t)$ are functions of S_t and ν_t, both of which follow a geometric Brownian motion. Therefore we can apply lemma A.0.1, Itô's lemma for the bivariate case with $f(X(t), Y(t)) = V(S(t), \nu(t))$, $X(t) = S(t)$ and $Y(t) = \nu(t)$. Also, recall that the drift coefficient of $d\nu$ is given by $k(\theta - \nu_t)$. For simplicity, and also because it will eventually vanish from the equations, we will denote the drift of $d\nu$, μ_ν in what follows.

This leads to the following expression for dV

$$\begin{aligned} dV = &\left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \mu_\nu \frac{\partial V}{\partial \nu} + \frac{1}{2}\nu^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 V}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 V}{\partial S \partial \nu}\right)dt + \\ &+ \nu S\frac{\partial V}{\partial S}dW_S + \xi\nu\frac{\partial V}{\partial \nu}dW_\nu \end{aligned} \tag{3.70}$$

Applying Itô's lemma to dU leads to the same expression as (3.70), but in U. Substituting (3.70), the respective expression for dU and (3.64) into equation (3.69) and rearranging the left-hand side, yields

$$\begin{aligned} d\Pi_t = a_t dS + M_t dU - dV = r(a_t S + M_t U - V)dt \Leftrightarrow \\ \Leftrightarrow \Bigg[a_t\mu S + M_t\left(\frac{\partial U}{\partial t} + \mu S\frac{\partial U}{\partial S} + \mu_\nu\frac{\partial U}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 U}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 U}{\partial S\partial\nu}\right) - \\ - \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \mu_\nu\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 V}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 V}{\partial S\partial\nu}\right)\Bigg]dt + \\ + \left[a_t\nu S + M_t\nu S\frac{\partial U}{\partial S} - \nu S\frac{\partial V}{\partial S}\right]dW^S + \left[M_t\xi\nu\frac{\partial U}{\partial \nu} - \xi\nu\frac{\partial V}{\partial \nu}\right]dW^\nu = \\ = r(a_t S + M_t U - V)dt \end{aligned} \tag{3.71}$$

In order to eliminate the risk caused by movements in the stock and volatility, the terms containing dW^S and dW^ν must vanish from equation (3.71). So we set the coefficients of dW^S and dW^ν to zero. This leads to

$$\frac{\partial V}{\partial S} = a_t + M_t\frac{\partial U}{\partial S} \tag{3.72}$$

and

$$\frac{\partial V}{\partial \nu} = M_t\frac{\partial U}{\partial \nu} \tag{3.73}$$

Proposition 8. Substituing a_t and M_t, given in (3.72) and (3.73), respectively, into equation (3.71) and rearranging it yields

$$\begin{aligned} \frac{\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\nu S^2\frac{\partial^2 V}{\partial S^2} + \rho\xi\nu S\frac{\partial^2 V}{\partial S\partial\nu} + \frac{1}{2}\xi^2\nu\frac{\partial^2 V}{\partial\nu^2} - rV}{\frac{\partial V}{\partial \nu}} = \\ = \frac{\frac{\partial U}{\partial t} + rS\frac{\partial U}{\partial S} + \frac{1}{2}\nu S^2\frac{\partial^2 U}{\partial S^2} + \rho\xi\nu S\frac{\partial^2 U}{\partial S\partial\nu} + \frac{1}{2}\xi^2\nu\frac{\partial^2 U}{\partial\nu^2} - rU}{\frac{\partial U}{\partial \nu}} \end{aligned} \tag{3.74}$$

Proof. We start by substituting a_t into (3.69) because its expression depends on M_t

$$\begin{aligned}
d\Pi_t = &\left(\frac{\partial V}{\partial S} - M_t\frac{\partial U}{\partial S}\right)\mu S dt + \\
&+ M_t\left(\frac{\partial U}{\partial t} + \mu S\frac{\partial U}{\partial S} + \mu_\nu\frac{\partial U}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 U}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 U}{\partial S\partial\nu}\right)dt - \\
&- \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \mu_\nu\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 V}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 V}{\partial S\partial\nu}\right)dt = \\
&= r\left[\left(\frac{\partial V}{\partial S} - M_t\frac{\partial U}{\partial S}\right)S + M_t U - V\right]dt
\end{aligned}$$

Now that we have substituted a_t we can replace M_t. We also divide by dt.

$$\begin{aligned}
&\left(\frac{\partial V}{\partial S} - \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial U}{\partial \nu}}\frac{\partial U}{\partial S}\right)\mu S + \\
&+ \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial U}{\partial \nu}}\left(\frac{\partial U}{\partial t} + \mu S\frac{\partial U}{\partial S} + \mu_\nu\frac{\partial U}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 U}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 U}{\partial S\partial\nu}\right) - \\
&- \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \mu_\nu\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 V}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 V}{\partial S\partial\nu}\right) = \\
&= r\left[\left(\frac{\partial V}{\partial S} - \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial U}{\partial \nu}}\frac{\partial U}{\partial S}\right)S + \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial U}{\partial \nu}}U - V\right]
\end{aligned}$$

Now we separate the terms that depend on $\frac{\partial V}{\partial \nu}/\frac{\partial U}{\partial \nu}$, from those that do not:

$$\begin{aligned}
&-\mu S\frac{\partial V}{\partial S} + \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \mu_\nu\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 V}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 V}{\partial S\partial\nu}\right) + r\frac{\partial V}{\partial S} - V = \\
&= \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial U}{\partial \nu}}\left[-\mu S\frac{\partial U}{\partial S} + \left(\frac{\partial U}{\partial t} + \mu S\frac{\partial U}{\partial S} + \mu_\nu\frac{\partial U}{\partial \nu} + \frac{1}{2}\nu^2 S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2\nu^2\frac{\partial^2 U}{\partial \nu^2} + \rho\nu\xi S\frac{\partial^2 U}{\partial S\partial\nu}\right)\right] + \\
&+ \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial U}{\partial \nu}}\left[r\frac{\partial U}{\partial S} - U\right]
\end{aligned} \tag{3.75}$$

Finally, arranging the equation so that the left-hand side only depends on V and the right-hand side only depends on U and noticing that the terms that depend on μ vanish, leads to (3.74). □

In the above equation (3.75), the left-hand side is a function of V only and the right-hand side becomes a function of U only. This implies that both sides of the equation can be written as a function f, of the independent variables S, ν and t. Following [29] that follows Heston [15], we specify this function as

$$f(S,\nu,t) = -k(\theta - \nu) + \lambda(S,\nu,t) \tag{3.76}$$

where $\lambda(S,\nu,t)$ is the price of volatility risk. Substituting $f(S,\nu,t)$ into the right-hand side of equation (3.74), we get

$$\frac{\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\nu S^2\frac{\partial^2 V}{\partial S^2} + \rho\xi\nu S\frac{\partial^2 V}{\partial S\partial\nu} + \frac{1}{2}\xi^2\nu\frac{\partial^2 V}{\partial \nu^2} - rV}{\frac{\partial V}{\partial \nu}} = -k(\theta - \nu) + \lambda(S,\nu,t) \tag{3.77}$$

re-arranging it produces the Heston PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\nu S^2\frac{\partial^2 V}{\partial S^2} + \rho\xi\nu S\frac{\partial^2 V}{\partial \nu\partial S} + \frac{1}{2}\xi^2\nu\frac{\partial^2 V}{\partial \nu^2} - rV + rS\frac{\partial V}{\partial S} + [k(\theta - \nu) - \lambda(S,\nu,t)]\frac{\partial V}{\partial \nu} = 0 \tag{3.78}$$

Chapter 4

The finite difference method for the Black-Scholes equation

In this chapter we will use finite difference schemes to approximate the Black-Scholes partial differential equation when reduced to the heat equation, as we have seen in section 3.1.4. The idea in finite difference methods is to find a solution for the differential equation by approximating every partial derivative numerically. We will follow [34] and [1]

4.1 General concepts

4.1.1 Definitions

Consider the following initial-boundary value problem with homogeneous Dirichlet boundary condition for the heat equation

$$\frac{\partial u}{\partial \tau}(x,\tau) = \frac{\partial^2 u}{\partial x^2}(x,\tau)$$

where $u(x,\tau)$ is defined in $-\infty < x < \infty$ and $0 \leq \tau \leq \frac{1}{2}\sigma^2 T$

$$\begin{cases} \frac{\partial u}{\partial \tau}(x,\tau) = \frac{\partial^2 u}{\partial x^2}(x,\tau) & (x,\tau) \in (x_{N^-}, x_{N^+}) \times (\tau_0, \tau_M) \\ u(x,0) = u_0(x) & x \in (x_{N^-}, x_{N^+}) \\ u(x_{N^-},\tau) = 0 & \tau \in (\tau_0, \tau_M) \\ u(x_{N^+},\tau) = 0 & \tau \in (\tau_0, \tau_M) \end{cases} \tag{4.1}$$

Where the first equation is the heat equation, and the other equations correspond to the boundary conditions.

For the finite difference approximation we need to define a rectangular region in the domain of u and partition it to form a mesh of equally spaced points. The discretisation steps Δx and $\Delta \tau$ are defined as

$$\Delta x = \frac{x_n - x_{N^-}}{n}, \qquad n = 0, 1, ..., N-1, N$$

$$\Delta \tau = \frac{\tau_m - \tau_0}{m}, \qquad m = 0, 1, ..., M-1, M$$

where $\tau_0 = 0$, $x_0 = x_{N^-}$ and $x_N = x_{N^+}$.

4.2 Numerical methods

4.2.1 θ schemes

A θ scheme is a convex combination of an explicit and an implicit scheme, which takes the form θscheme $= (1-\theta)$explicit $+ \theta$implicit, where θ is a parameter in [0,1]. Therefore, for the value $\theta = 0$ we recover the explicit scheme, for $\theta = 1$ the fully implicit scheme and for $\theta = \frac{1}{2}$ we recover the Crank-Nicolson scheme. Moreover, when $\theta \neq 0$ we have an implicit scheme.

For the explicit scheme we consider forward differences in time and second order central differences in space

$$\partial_\tau u(x_n, \tau_m) = \frac{u_{n,m+1} - u_{n,m}}{\Delta\tau} + O(\Delta\tau) \tag{4.2}$$

$$\partial_x u(x_n, \tau_m) = \frac{u_{n+1,m} - 2u_{n,m} + u_{n-1,m}}{(\Delta x)^2} + O((\Delta x)^2) \tag{4.3}$$

Substituting equations (4.2) and (4.3) into the heat equation and ignoring the terms $O(\Delta\tau)$ and $O((\Delta x)^2)$ we obtain the explicit scheme

$$u_{n,m} = \chi u_{n+1,m} + (1-2\chi)u_{n,m} + \chi u_{n-1,m}, \quad \chi = \frac{\Delta\tau}{(\Delta x)^2} \tag{4.4}$$

For the fully implicit scheme we also use forward differences in time and second order central differences in space

$$\partial_\tau u(x_n, \tau_{m+1}) = \frac{u_{n,m+1} - u_{n,m}}{\Delta\tau} + O(\Delta\tau) \tag{4.5}$$

$$\partial_x u(x_n, \tau_{m+1}) = \frac{u_{n+1,m+1} - 2u_{n,m+1} + u_{n-1,m+1}}{(\Delta x)^2} + O((\Delta x)^2) \tag{4.6}$$

Substituting equations (4.5) and (4.6) into the heat equation and ignoring the terms $O(\Delta\tau)$ and $O((\Delta x)^2)$ we obtain the fully implicit scheme

$$u_{n,m+1} = -\chi u_{n+1,m+1} + (1+2\chi)u_{n,m+1} - \chi u_{n-1,m+1}, \quad \chi = \frac{\Delta\tau}{(\Delta x)^2} \tag{4.7}$$

As the Crank-Nicolson scheme is a weighted average of an explicit scheme and an implicit scheme, the time and space differences are given by

$$\partial_\tau u(x_n, \tau_{m+\frac{1}{2}}) = \frac{u_{n,m+1} - u_{n,m}}{\Delta\tau} + O((\Delta\tau)^2) \tag{4.8}$$

$$\partial_x u(x_n, \tau_{m+\frac{1}{2}}) = \frac{u_{n+1,m} - 2u_{n,m} + u_{n-1,m} + u_{n+1,m+1} - 2u_{n,m+1} + u_{n-1,m+1}}{2(\Delta x)^2} + O((\Delta x)^2) \tag{4.9}$$

Substituting equations (4.8) and (4.9) into the heat equation and ignoring the terms $O((\Delta\tau)^2)$ and $O((\Delta x)^2)$ we obtain the Cranck-Nicolson scheme

$$u_{n,m+1} - \frac{1}{2}\chi(u_{n+1,m+1} - 2u_{n,m+1}) = u_{n,m} + \frac{1}{2}\chi(u_{n+1,m} - 2u_{n,m} + u_{n-1,m}), \quad \chi = \frac{\Delta\tau}{(\Delta x)^2} \tag{4.10}$$

We can write the above finite difference schemes in matrix form

$$M_{I,1-\theta}u_{m+1} = M_{E,\theta}u_m + b_{m+\theta}$$

where

$$M_{I,1-\theta} = \begin{bmatrix} 1+2\chi\theta & -\theta\chi & 0 & \cdots & 0 \\ -\theta\chi & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\theta\chi \\ 0 & \cdots & 0 & -\theta\chi & 1+2\chi\theta \end{bmatrix}$$

$$M_{E,\theta} = \begin{bmatrix} 1-2\chi(1-\theta) & (1-\theta)\chi & 0 & \cdots & 0 \\ (1-\theta)\chi & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & (1-\theta)\chi \\ 0 & \cdots & 0 & (1-\theta)\chi & 1-2\chi(1-\theta) \end{bmatrix}$$

$$u_m = \begin{bmatrix} u_{N^-+1,m} \\ \vdots \\ u_{0,m} \\ \vdots \\ u_{N^+-1,m} \end{bmatrix}$$

$$b_{m+\theta} = \begin{bmatrix} \chi(1-\theta)u_{N^-,m} + \chi\theta u_{N^-,m+1} \\ 0 \\ \vdots \\ 0 \\ \chi(1-\theta)u_{N^+,m} + \chi\theta u_{N^+,m+1} \end{bmatrix}$$

$b_{m+\theta}$ is the vector with the boundary conditions.
In this work we consider three θ schemes which correspond to the values $0, \frac{1}{2}, 1$ of θ.

Convergence, consistency and stability of the θ methods

It can be shown [1] that a scheme is unconditionally stable for $\theta \geq \frac{1}{2}$ and for $\theta < \frac{1}{2}$ it is stable if $(1-2\theta)\chi \leq \frac{1}{2}$ holds, where $\chi = \frac{\Delta\tau}{(\Delta x)^2}$. Furthermore, a scheme is consistent of order 2 when $\theta = \frac{1}{2}$ and consistent of order 1 in the other cases ($\theta = 0$ and $\theta = 1$), which implies convergence of the same order as long as the scheme is stable as well.

4.3 Numerical experiments

Implementation

In this section we discuss the algorithms used to carry out the experiments in this chapter and chapter 6. Although the algorithms were implemented using *Matlab*, these are generic algorithms that can be implemented in other programming languages such as *Fortran*, and therefore are not optimised for a

particular language.

We divide the space interval $[0, x]$ in N subintervals of equal length Δx and divide the time interval $[0, \tau]$ in M subintervals of equal length $\Delta\tau$. This results in a grid of $(N+1) \times (M+1)$ points. Recall that $u(x, \tau)$ is defined on $-\infty < x < \infty$ and $0 \leq \tau \leq \frac{1}{2}\sigma^2 T$, which means that we have to choose some endpoints, x_{N^-} and x_{N^+}, to limit the x domain. There is no criterion for choosing these points, however a price range too small will result in a poor approximation at the boundaries and a range too large will lead to unecessary computations.

In the end we wish to have the value of the option V at $t = 0$ and $S = S_0$, so it is convenient to keep track of these points throughout the computations. For that purpose we define $x_0 = \log\left(\frac{S_0}{K}\right)$ and N^- and N^+ as the number of points to the left and to the right of x_0, respectively.

For the implementation of the algorithms we considered a matrix U of dimensions $(2N+1) \times (M+1)$, where we stored the values of $u(x, \tau)$ for each of the mesh points. As *Matlab* does not allow for an index $i = 0$ we start our indexes at 1.

The explicit scheme is quite straightforward, to compute the unknown values of $u_{m+1}(x, \tau)$ we only need to use known values of u_m , which does not require solving a system of equations. Algorithm 1 shows the implementation of this scheme. However, in the implicit schemes we need to solve a linear system of equations at each time step. The matrix $M_{I,1-\theta}$ is positive definite, thus it is invertible, nevertheless inverting a matrix is computationally costly so a more efficient method to solve the system is to factorise the matrix using for instance LU decomposition. This decomposition is unique if we require that the diagonal of L or U consists of ones. *Matlab* has a built in routine that implements this factorisation with the diagonal of L consisting of ones which we used in our implementation. In order to avoid mistaking matrix U from the heat equation for the matrix U from the LU decomposition, we named the latter as Up. Algorithms 2 and 3 show the implementation of the fully implicit scheme and the Crank-Nicolson scheme, respectively.

It is also worth mentioning that in our experiments we compute the relative error by using equations (3.38) and (3.51) for the true value of call and put options, respectively.

Algorithm 1 Explicit scheme for option pricing

1. Define r, σ, T, S_0, K, $\Delta\tau$, Δx, N^-, N^+
2. Set $\tau_{max} = \frac{1}{2}\sigma^2 T$, $M = \frac{\tau_{max}}{\Delta\tau}$, $N = N^- + N^+$, $x_0 = \log\left(\frac{S_0}{K}\right)$, $x_{N^-} = x_0 - N^-\Delta x$, $x_{N^+} = x_0 + N^+\Delta x$, $\chi = \frac{\Delta\tau}{\Delta x^2}$, $c = \frac{2r}{\sigma^2}$, $\alpha = -\frac{1}{2}(c-1)$, $\beta = -\frac{1}{4}(c+1)^2$
3. Define boundary conditions in x and in τ

```
for j = 1 : M do
    for i = 2 : N do
        U(i, j+1) = χU(i+1, j) + (1 − 2χ)U(i, j) + χU(i−1, j)
    end for
end for
```

4. Transform the equation back to V, at time 0: $V(S_0, t_0) = KU(N^- + 1, M + 1)e^{\alpha x_0 + \beta\tau_{max}}$

Algorithm 2 Implicit scheme for option pricing

1. Define r, σ, T, S_0, K, $\Delta\tau$, Δx, N^-, N^+
2. Set $\tau_{max} = \frac{1}{2}\sigma^2 T$, $M = \frac{\tau_{max}}{\Delta\tau}$, $N = N^- + N^+$, $x_0 = \log\left(\frac{S_0}{K}\right)$, $x_{N^-} = x_0 - N^-\Delta x$, $x_{N^+} = x_0 + N^+\Delta x$, $\chi = \frac{\Delta\tau}{\Delta x^2}$, $c = \frac{2r}{\sigma^2}$, $\alpha = -\frac{1}{2}(c-1)$, $\beta = -\frac{1}{4}(c+1)^2$
3. Define boundary conditions in x and in τ
4. Construct matrix $M_{I,0}$, as defined in section 4.2
5. LU decomposition of matrix $M_{I,0}$

1: **for** $j = 2 : M + 1$ **do**
2: Solve $Ly = M_{E,0}u_m + b_{m+1}$:
3: Substitute first interior node:
4: $U(2, j) = U(2, j-1) + \chi U(1, j)$
5: **for** $i = 3 : N - 1$ **do**
6: Substitute the next interior nodes until $N - 2$:
7: $U(i, j) = U(i, j-1) - L(i-1, 1 : i-2)U(2 : i-1, j)$
8: **end for**
9: Substitute last interior node:
10: $U(N, j) = U(N, j-1) + \chi U(N+1, j)L(N-1, 1 : N-2)U(2 : N-1, j)$
11: Solve $\mathrm{Up}u_{m+1} = y$:
12: Substitute last interior node:
13: $U(N, j) = \frac{U(N,j)}{Up(N-1,N-1)}$
14: **for** $i = N : 2$ **do**
15: Substitute the remaining nodes:
16: $U(i, j) = \frac{U(i,j) - Up(i-1,i:N-1)U(i+1:N,j)}{Up(i-1,i-1)}$
17: **end for**
18: **end for**

6. Transform the equation back to V, at time 0: $V(S_0, t_0) = KU(N^- + 1, M + 1)e^{\alpha x_0 + \beta\tau_{max}}$

Algorithm 3 Crank-Nicolson scheme for option pricing

1. Define r, σ, T, S_0, K, $\Delta\tau$, Δx, N^-, N^+
2. Set $\tau_{max} = \frac{1}{2}\sigma^2 T$, $M = \frac{\tau_{max}}{\Delta\tau}$, $N = N^- + N^+$, $x_0 = \log\left(\frac{S_0}{K}\right)$, $x_{N^-} = x_0 - N^-\Delta x$, $x_{N^+} = x_0 + N^+\Delta x$, $\chi = \frac{\Delta\tau}{\Delta x^2}$, $c = \frac{2r}{\sigma^2}$, $\alpha = -\frac{1}{2}(c-1)$, $\beta = -\frac{1}{4}(c+1)^2$
3. Define boundary conditions in x and in τ
4. Construct matrix $M_{I,\frac{1}{2}}$, as defined in section 4.2
5. LU decomposition of matrix $M_{I,\frac{1}{2}}$

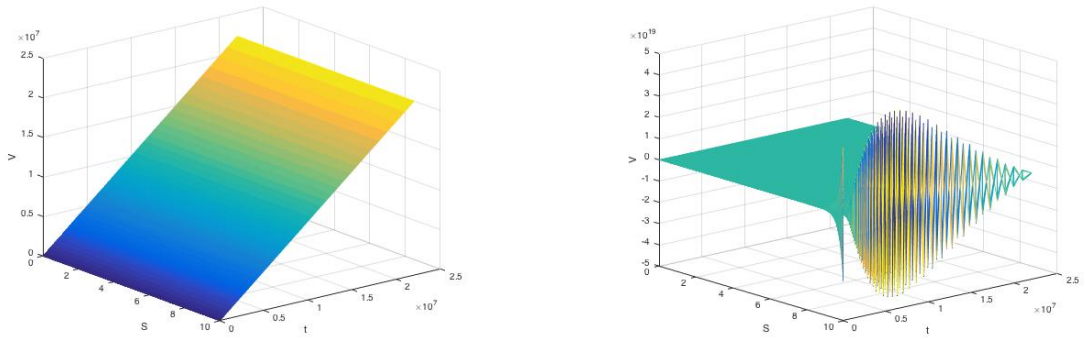
1: **for** $j = 2 : M+1$ **do**
2: Solve $Ly = M_{E,\frac{1}{2}}u_m + b_{m+\frac{1}{2}}$:
3: Substitute first interior node:
4: $U(2,j) = U(2,j-1) + \frac{1}{2}\chi\left[U(3,j-1) - 2U(2,j-1) + U(1,j-1)\right] + \frac{1}{2}\chi U(1,j)$
5: **for** $i = 3 : N-1$ **do**
6: Substitute the next interior nodes until $N-2$:
7: $U(i,j) = U(i,j-1) + \frac{1}{2}\chi\left[U(i+1,j-1) - 2U(i,j-1) + U(i-1,j-1)\right] - L(i-1,1:i-2)U(2:i-1,j)$
8: **end for**
9: Substitute last interior node:
10: $U(N+1,j) = U(N+1,j-1) + \frac{1}{2}\chi\left[U(N+2,j-1) - 2U(N+1,j-1) + U(N,j-1)\right] + \frac{1}{2}\chi U(N+2,j) - L(N,1:N-1)U(2:N,j)$
11: Solve $\mathrm{Up}u_{m+1} = y$:
12: Substitute last interior node:
13: $U(N+1,j) = \frac{U(N+1,j)}{Up(N,N)}$
14: **for** $i = N : 2$ **do**
15: Substitute the remaining nodes:
16: $U(i,j) = \frac{U(i,j) - Up(i-1,i:N)U(i+1:N+1,j)}{Up(i-1,i-1)}$
17: **end for**
18: **end for**

6. Transform the equation back to V, at time 0: $V(S_0,t_0) = KU(N^- + 1, M+1)e^{\alpha x_0 + \beta\tau_{max}}$

Stability condition for the explicit method

In order for the explicit method to be stable it must verify the stability condition: $\chi \leq \frac{1}{2}$.

Figure 3.1 compares the approximation of the value of a call option using the explicit scheme when the stability condition is verified and when it is not.



(a) Stable explicit method

(b) Unstable explicit method

Figure 4.1: Approximation of V(S,t) with a stable and an unstable explicit scheme

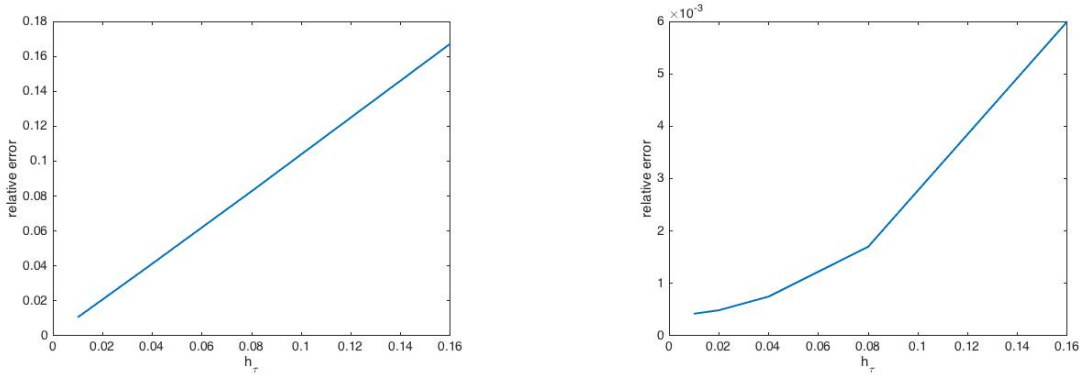
In this experiment we fixed $\Delta x = 0.05$ and used $\Delta\tau = 0.001$ to have a stable scheme. To break the stability condition we set $\Delta\tau = 0.0013$.

Convergence as we increase the mesh spacing

As we have seen in the previous section, the fully implicit and explicit schemes have an order of convergence of $O(\Delta\tau) + O(\Delta x^2)$, while the Crank-Nicolson scheme has an order of convergence of $O(\Delta\tau^2) + O(\Delta x^2)$. This means that for small values of $\Delta\tau^2$ the error in Crank-Nicolson scheme will be similar to the error in the explicit scheme because the dominant term will be $O(\Delta x^2)$ as well. For this reason, there is no great advantage in using implicit schemes when imposing the stability condition $\frac{\Delta\tau}{\Delta x^2} \leq \frac{1}{2}$. However, in order to compare the schemes, we imposed the condition in all off the schemes in the numerical experiences carried out in chapter 5.

Furthermore, if we fix Δx and increase $\Delta\tau$, we see that the Crank-Nicolson scheme will have quadratic convergence and the implicit scheme will have linear convergence, as the terms $O(\Delta\tau^2)$ and $O(\Delta\tau)$ will dominate in each method, respectively. Usually when studying the error in the approximations we should consider all the points in the domain, however, in this work we will only look at the error in the point $V(S_0, t_0)$. We do this not only because the goal of option pricing is to know the present value of the option and hence the point $V(S_0, t_0)$ is the point we are interested in, but also because the function $V(S, t)$ takes both null values and very large values, which could cause oscillations in the error.

Figure 4.2 shows the evolution of the relative error of V at the point (S_0, t_0) for the fully implicit scheme and the Crank-Nicolson scheme. To obtain those results we set $\Delta x = 0.05$ and increased $\Delta\tau$ by a factor of 2, starting with $\Delta\tau = 0.01$.



(a) Linear convergence for the implicit method

(b) Quadratic convergence for the Crank-Nicolson method

Figure 4.2: Evolution of the relative error for the fully implicit scheme and the Crank-Nicolson scheme

4.3.1 Limitations of finite difference schemes for PDE's in option pricing

Finite difference methods for PDE's are widely used in option pricing. They are easy to implement and are a competitive choice for low dimension problems, such as the one-dimension Black-Scholes model presented in this chapter. The pricing of early exercise options, like American options, is also well handled by these methods.

Despite their many advantages, they also have several drawbacks. The most obvious one is that they can only be used when it is possible to derive a PDE for the model. This is for instance the case of models that are based on the Brownian motion. However, if we want to depart from the assumption that the underlying asset follows a geometric Brownian motion we will have models for which there is no PDE. We have seen in chapter 3 some extensions to the classical Black-Scholes model, for all of which it is possible to derive a PDE. However, the introduction of time-dependent parameters or even dependence on the underlying asset makes the models more complex. When the coefficients of the model are constant or depend only on time, it is possible to reduce the PDE to the heat equation and even find an exact solution but when the coefficients also depend on the underlying asset this no longer possible. In this case the Black-Sholes equation must be solved directly. In 2011, Cen and Le [7] proposed a robust and accurate finite difference method for this case but results on this topic is still scarce. For some stochastic volatility models, like the Heston model, it is possible to derive a PDE, as we have seen, and there is extensive work on the approximation of this PDE by finite difference schemes. For jump diffusion models it is also possible to derive a partial integro-differential equation, which can be approximated by finite difference schemes, see for instance Duffy [9].

Nonetheless, for complex models the Monte Carlo method seems a more suitable choice as it can easily accommodate changes on the asset model. In this sense, the next chapter will present numerical methods for SDE's and show how to implement them using a Monte Carlo approach.

Chapter 5

Numerical methods for stochastic differential equations

This chapter follows partially the article *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations* by J. Higham [16].

5.1 Euler-Maruyama method

The Euler-Maruyama method is a generalisation of the Euler method for ordinary differential equations to stochastic differential equations and may be applied to an equation of the form

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t), \quad X(0) = X_0, \quad 0 \leq t \leq T. \tag{5.1}$$

where a and b are scalar functions and the initial condition $X(0)$ is a random variable.

Equation (5.1) is a way of representing the integral equation

$$X(t) = X(0) + \int_0^T a(s, X(s))ds + \int_0^T b(s, X(s))dW_s \tag{5.2}$$

where the solution $X(t)$ is a random variable for each t and the last integral is an Itô integral.

To find an approximate solution on the interval $[0, T]$ we discretise it into L equal subintervals of width Δt and approximate X values

$$X_0 < X_1 < ... < X_L$$

at the respective t points

$$0 = \tau_0 < \tau_1 < \tau_2 < ... < \tau_L = T$$

The explicit method takes the form

$$X_j = X_{j-1} + a(\tau_{j-1}, X_{j-1})\Delta t + b(\tau_{j-1}, X_{j-1})\Delta W_j \qquad j = 1, 2, ..., L \tag{5.3}$$

where X_j denotes our approximation, $\Delta t = \frac{T}{L}$ and $\Delta W_j = W(\tau_j) - W(\tau_{j-1})$.
We can rewrite equation (5.3) as

$$X_j = X_{j-1} + \int_{\tau_{j-1}}^{\tau_j} a(s, X(s))ds + \int_{\tau_{j-1}}^{\tau_j} b(s, X(s))dW(s) \tag{5.4}$$

by applying the Euler-Maruyama method to equation (5.2). In this form we can easily see that the integrals are discretised at the left end points. This is an essential detail because, as we have seen in chapter 2, when we introduced the stochastic integration, Itô's integral is only defined at the left end points. In addition, notice that if $b \equiv 0$ then we have the Euler method for ODE's.

Like in the case of ODE's, it is also possible to define implicit schemes. These are required for stiff SDE's, that is, for SDE's that are unstable for some numerical methods. The fully implicit scheme is defined as

$$X_j = X_{j-1} + a(\tau_j, X_j)\Delta t + b(\tau_j, X_j)\Delta W_j \qquad j = 1, 2, ..., L \tag{5.5}$$

However, this scheme has unbounded random variables because of the implicitness in the diffusion term. This gives rise to some issues when applying fully implicit schemes to obtain strong approximations of solutions of SDE's because finite absolute moments generally will not exist for fully implicit strong approximations and therefore it would not make sense to make a strong convergence analysis. For further details refer to [22].

For this reason, in this thesis we will only consider semi-implicit schemes, that is, schemes in which only the non-random coefficients are implicit.

Consider now the semi-implicit Euler scheme

$$X_j = X_{j-1} + a(\tau_j, X_j)\Delta t + b(\tau_{j-1}, X_{j-1})\Delta W_j \qquad j = 1, 2, ..., L \tag{5.6}$$

Like in the deterministic case, we can also define a family of semi-implicit Euler schemes for SDE's

$$X_j = X_{j-1} + \left[\theta a(\tau_j, X_j) + (1-\theta)a(\tau_{j-1}, X_{j-1})\right]\Delta t + b(\tau_{j-1}, X_{j-1})\Delta W_j \qquad j = 1, 2, ..., L \tag{5.7}$$

where $\theta \in [0, 1]$ represents the degree of implicitness. When $\theta = 0$ we have the explicit scheme (5.3), when $\theta = 1$ we recover the fully semi-implicit scheme (5.6) and when $\frac{1}{2}$ we have a generalisation of the deterministic trapezoidal method.

We now address one crucial question in numerical methods for SDE's, that is how to discretise the Brownian motion.

Discretised Brownian motion

Since we are in discrete time we also need to consider discretised Brownian motion in $[0, T]$. We divide the interval in N subintervals of equal spacing δt, where $\delta t = \frac{T}{N}$. Also, represent by $\mathcal{N}(0, 1)$ the standard normal distribution. Each ΔW_j is a random variable of the form

$$\Delta W_j = z_i\sqrt{\delta t} \tag{5.8}$$

where z_i is sampled from $\mathcal{N}(0, 1)$.

Note that we have different timesteps for the Brownian motion and the numerical method. In fact we will set $\Delta t = R\delta t$, where R is an integer ≥ 1, to study the strong order of convergence in our numerical experiments. Otherwise, if we increased both the Brownian motion and the numerical method time step, the coarse approximation of the Brownian motion would reflect on the error as well and it would not be possible to clearly see how the change in time step affects the strong convergence of the method. On the other hand, as we only need the mean of the solution to study weak convergence it is not important which time step we use to sample the Brownian motion. For this reason in our numerical experiments

we will use the same time step for the numerical approximation and for the Brownian motion.

5.2 Convergence and consistency

Now that we have introduced the Euler-Maruyama scheme for SDE's we want to determine if the method converges to the true solution, that is, if as we decrease the mesh spacing the numerical solution approaches the exact solution. The definition of convergence of numerical methods for SDE's is very similar to that of ODE's. However, because of the random component we have more than one way of analysing convergence. In the first one we are interested in computing the difference between the approximate and exact solutions at specific mesh points, therefore this type of convergence is path dependent. In the second we are only interest in convergence in distribution, that is, in approximating the expectations of the Ito process. In this section we will follow Higham [16] and Kloeden and Platen [22].

5.2.1 Strong convergence and consistency

The strong order of convergence gives the rate at which the mean of the errors decreases as the time step tends to zero. In our numerical experiences we are interested in the error only at maturity, T.

Definition 5.2.1. (Strong convergence). A general time discrete approximation converges strongly to the solution at time T if

$$\lim_{\Delta t \to 0} E[|X(T) - \tilde{X}(T)|] = 0 \tag{5.9}$$

where E denotes the expected value and $\tilde{X}(T)$ is the approximation of $X(t)$ at time $t = T$ computed with constant step Δt.

Further, we denote the error at final time T in the strong sense as

$$e_{\Delta t}^{\text{strong}} := E[|X(T) - \tilde{X}(T)|] \tag{5.10}$$

In order to be able to compare the accuracy of different numerical schemes we must introduce the concept of rate of convergence, which is similar to the concept for ODE's.

Definition 5.2.2. (Strong order of convergence). A general time discrete approximation is said to strongly converge with order γ at time T if there exists a constant C such that

$$e_{\Delta t}^{\text{strong}} \leq C \Delta t^{\gamma} \tag{5.11}$$

Note that when the diffusion coefficient is zero and the initial value is deterministic, the above definition is reduced to convergence criterion for ODE's. In fact the convergence criterion for SDE's is just a generalisation of the convergence criterion for ODE's.

It can be show that under conditions of theorem 2.2.5 on a and b , the family of Euler schemes has a strong order of convergence of $\frac{1}{2}$.

Now that we have studied the accuracy of the numerical schemes, we wish to study the consistency of the schemes, just like we would for ODE's, that is, we want to know if the truncation errors vanish as the time step goes to zero.

Definition 5.2.3. (Strong consistency). A discrete time approximation $\tilde{X}$ corresponding to a time discretisation $(\tau)_{\Delta t} = \{\tau_j : j = 0, 1, ...\}$ with constant step size Δt is strongly consistent if there exists a

non-negative function $c = c(\Delta t)$ with

$$\lim_{\Delta t \to 0} c(\Delta t) = 0 \tag{5.12}$$

such that

$$E\left[\left|E\left(\frac{\tilde{X}_{j+1} - \tilde{X}_j}{\Delta t}\middle|\mathcal{F}_{\tau_j}\right) - a(\tau_j, \tilde{X}_j)\right|^2\right] \leq c(\Delta t) \tag{5.13}$$

and

$$E\left(\frac{1}{\Delta t}\left|\tilde{X}_{j+1} - \tilde{X}_j - E\left(\tilde{X}_{j+1} - \tilde{X}_j|\mathcal{F}_{\tau_j}\right) - b(\tau_j, \tilde{X}_j\Delta W_j)\right|^2\right) \leq c(\Delta t) \tag{5.14}$$

for all fixed values $\tilde{X}_j = x$ and $j = 0, 1, ...$ and where $\{\mathcal{F}_\tau, \tau \geq 0\}$ is a preassigned increasing family of ς-fields.

Condition (5.13) requires that the difference between the mean of the increment of the approximation and that of the Ito process converge to zero. In fact, when the random coefficient is zero, this is equivalent to the definition of consistency in the deterministic case. Condition (5.14) says that the variance of the difference between the random parts of the approximation and the Itô process should also converge to zero. This means that strong consistency gives an indication of the pathwise closeness.

5.2.2 Weak convergence and consistency

The strong order of convergence is quite demanding to implement, as it requires that the whole path is known. However, we do not always need that much information, if we are interested, for instance, in just knowing the probability distribution of the solution $X(t)$. In this case it would suffice to know the rate the at which error of the means decreases, as the time step tends to zero.

Definition 5.2.4. (Weak convergence). A method has weak convergence at time T if

$$\lim_{\Delta t \to 0} |E(f(X(T))) - E(f(\tilde{X}(T)))| = 0 \tag{5.15}$$

for all functions f in the polynomial class. Moreover, f needs to be smooth and display polynomial growth.

Like we did for the strong convergence we define the error at the final time T as

$$e_{\Delta t}^{\text{Weak}} := |E[f(X(T))] - E[f(\tilde{X}(T))]| \tag{5.16}$$

Definition 5.2.5. (Weak order of convergence). A method is said to weakly converge with order γ at time T if there exists a constant C such that

$$e_{\Delta t}^{\text{Weak}} \leq C\Delta t^\gamma \tag{5.17}$$

It can be shown that the family of Euler methods has weak order of convergence 1, if the conditions on a and b, of theorem 2.2.5, are met.

As we did in the strong case, we now give a definition of weak consistency.

Definition 5.2.6. (Weak consistency). A discrete time approximation $\tilde{X}$ corresponding to a time discretisation ($\tau_{\Delta t} = \{\tau_j : j = 0, 1, ...\}$) with constant step size Δt is weakly consistent if there exists a non-negative function $c = c(\Delta t)$ with

$$\lim_{\Delta t \to 0} c(\Delta t) = 0 \tag{5.18}$$

such that

$$E\left[\left|E\left(\frac{\tilde{X}_{j+1}-\tilde{X}_j}{\Delta t}\Big|\mathcal{F}_{\tau_j}\right)-a(\tau_j,\tilde{X}_j)\right|^2\right]\leq c(\Delta t) \tag{5.19}$$

and

$$E\left(\left|E\left(\frac{1}{\Delta t}(\tilde{X}_{j+1}-\tilde{X}_j)(\tilde{X}_{j+1}-\tilde{X}_j)^\top\Big|\mathcal{F}_{\tau_j}\right)-b(\tau_j,\tilde{X}_j)b(\tau_j,\tilde{X}_j)^\top\right|^2\right)\leq c(\Delta t) \tag{5.20}$$

for all fixed values $\tilde{X}_j = x$ and $j = 0, 1, ...$ and where $\{\mathcal{F}_\tau, \tau \geq 0\}$ is a preassigned increasing family of ς-fields and where $\top$ denotes the transpose of the vector.

It is worth noting that condition (5.19) is the same as condition (5.13) in the definition of strong consistency. However, condition (5.20) is different from condition (5.14) as it only requires the variance of the increment of the approximation to be close to the variance of the Itô process, while the definition of strong consistency required that the variance of the difference between the increments of the approximation and the Itô process vanish.

5.3 Milstein method

The Milstein's method uses Itô's lemma to add the second order term to the Euler-Maruyama scheme and increase the approximation's accuracy to 1. It has both strong and weak order of convergence equal to 1, under the usual assumptions on a and b, which means that it will converge to the true solution faster than the Euler-Maruyama method, as the time step goes to zero.

The Milstein's method is also applied to an equation of the form (5.1). The method takes the form

$$X_j = X_{j-1} + a(\tau_{j-1}, X_{j-1})\Delta t + b(\tau_{j-1}, X_{j-1})\Delta W_j + \frac{1}{2}b(\tau_{j-1}, X_{j-1})b'(\tau_{j-1}, X_{j-1})\left[\Delta W_j^2 - \Delta t\right] \tag{5.21}$$

where $b' = \dfrac{\partial b}{\partial X}$

In practice the Milstein scheme is just a variation of the Euler-Maruyama method in which we add the term $\frac{1}{2}bb'[\Delta^2 W_j - \Delta t]$. And like we had for the Euler-Maruyama we can also have an implicit Milstein scheme. However, for the same reasons, we will only consider a semi-implicit scheme.

$$\begin{aligned} X_j = X_{j-1} &+ a(\tau_j, X_j)\Delta t + b(\tau_{j-1}, X_{j-1})\Delta W_j + \\ &+ \frac{1}{2}b(\tau_{j-1}, X_{j-1})b'(\tau_{j-1}, X_{j-1})\left[\Delta W_j^2 - \Delta t\right] \end{aligned} \tag{5.22}$$

where $b' = \dfrac{\partial b}{\partial X}$.

As with the Euler-Maruyama schemes we can also define a family of semi-implicit Milstein schemes

$$\begin{aligned} X_j = X_{j-1} &+ \left[(1-\theta)a(\tau_j, X_j) + \theta a(\tau_{j-1}, X_{j-1})\right]\Delta t + b(\tau_{j-1}, X_{j-1})\Delta W_j + \\ &+ \frac{1}{2}b(\tau_{j-1}, X_{j-1})b'(\tau_{j-1}, X_{j-1})\left[\Delta W_j^2 - \Delta t\right] \end{aligned} \tag{5.23}$$

where again $\theta \in [0, 1]$ is the degree of implicitness. When $\theta = 0$ we recover the explicit scheme (5.21), when $\theta = 1$ we obtain the fully semi-implicit scheme (5.22) and when $\theta = \frac{1}{2}$ we obtain the generalisation of the deterministic trapezoidal method.

5.4 Numerical stability

So far we have seen the accuracy of the Euler-Maruyama and Milstein schemes. Now we are interested in seeing the solution behaviour in the long term as well as its sensitivity to small changes in the initial values. Recall that usually, when we talk about stability, we talk about the stability of an equilibrium position, i.e. $X(t) \equiv 0$. Following Higham [16] we will define two types of stability and we will restrict our attention to the linear SDE

$$dX(t) = aX(t)dt + bX(t)dW(t) \tag{5.24}$$

where a and b are allowed to be complex.

For $X_0 \neq 0$ with probability 1, we define *mean square stability* as

$$\lim_{t\to\infty} E\left[X(t)^2\right] = 0 \tag{5.25}$$

Note that the solution of equation (5.24) is given by

$$X(t) = X(0)e^{\left(a-\frac{1}{2}b^2\right)t+bW(t)} \tag{5.26}$$

So, it is easy to verify that for the linear test equation (5.24), (5.25) is equivalent to saying that a and b of equation (5.24) satisfy $Re(a) + \frac{1}{2}|b|^2 < 0$.

For $X_0 \neq 0$ with probability 1 we define *asymptotic stability* as

$$\lim_{t\to\infty} |X(t)^2| = 0, \quad \text{with probability 1} \tag{5.27}$$

Similarly, considering the linear test equation (5.24), (5.27) is equivalent to saying that the solution of equation (5.24) satisfies $Re\left(a - \frac{1}{2}b^2\right) < 0$.

Further, by looking at both definitions, (5.25) and (5.27), it is straightforward that mean-square stability implies asymptotic stability, but not the other way around.

Now that we have seen what conditions are necessary for equation (5.24) to be mean-square stable and asymptotic stable, we are interested in finding conditions that ensure that the numerical methods are also stable. Since explicit methods are usually more unstable than implicit methods, we will focus on the stability analysis of the explicit methods only. Recall from ODE's that in order to a numerical scheme to be stable, both the differential equation and the numerical scheme have to be stable. The same applies for SDE's. So, now we assume that a and b are such that the solution of (5.24) is mean-square and asymptotic stable. Applying the explicit Euler-Maruyama scheme (5.3) to equation (5.24) and using simple properties of the expected value we get the following condition on a and b, for which the Euler-Method is mean-square stable

$$\lim_{j\to\infty} E\left[X_j^2\right] = 0 \Leftrightarrow |1 + a\Delta t|^2 + |b|^2\Delta t < 1 \tag{5.28}$$

Finding the conditions for which the Euler-Maruyama method is asymptotic stable requires the application of the strong law of large numbers and the law of iterated logarithm, which leads to the following condition

$$\lim_{j\to\infty} |X_j^2| = 0, \quad \text{with probability 1} \Leftrightarrow E\left[\log|1 + a\Delta t + b\sqrt{\Delta t}\mathcal{N}(0,1)|\right] < 0 \tag{5.29}$$

For a more detailed explanation on how to derive the above condition, see Higham [**?**]. Now we derive

conditions for the stability of the Milstein method. Similarly to what we did for the Euler-Maruyama scheme, we apply the explicit Milstein method (5.21) to equation (5.24) and using simple properties of the expected value we get

$$\lim_{j\to\infty} E\left[X_j^2\right] = 0 \Leftrightarrow |1 + a\Delta t|^2 + |b|^2\Delta t < 1 \tag{5.30}$$

which is the same condition as in the Euler-Maruyama scheme. In order to the Milstein scheme to be asymptotic stable, the following condition has to hold

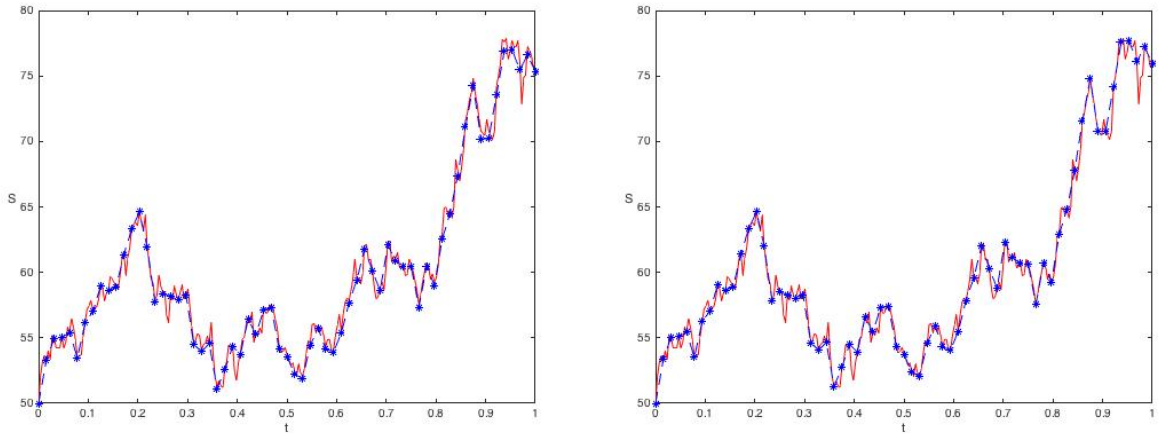
$$\lim_{j\to\infty} |X_j^2| = 0, \quad \text{with probability 1} \Leftrightarrow E\left[\log|1 + a\Delta t + b\sqrt{\Delta t}\mathcal{N}(0,1) + \frac{1}{2}b^2\sqrt{\Delta t}\mathcal{N}(0,1)|\right] < 0 \tag{5.31}$$

5.5 Numerical experiments

In this section we apply the Euler-Maruyama (5.3) and Milstein (5.21) methods to the Black-Scholes asset model equation (2.18).

The *matlab* codes used for the numerical experiments carried out in this section are based on the ones used in Higham [16] but modified for the Black-Scholes equation. The Brownian motion was discretised as discussed in 5.1, i.e. $\Delta W = \sqrt{\delta t} z_i$. The variables z_i were computed using the pseudorandom generator *randn* from *Matlab*, which produces an independent pseudorandom number from the standard normal distribution. To generate the random paths we created an array with dimensions $1 \times N$ using $randn(1, N)$ and scaled by $\sqrt{\delta t}$. In order to be able to repeat the experiments, *Matlab* allows to set the initial state of the random number generator. This means that by using the same initial state, the same pseudorandom sequence is generated.

Figure 5.1 shows the true and the approximate solution of equation (2.18) using the Euler-Maruyama and Milstein methods. The Brownian motion is sampled on the interval $[0, 1]$ and has a time step of $\delta t = 2^{-8}$, while the time step for the approximations is $\Delta t = 2^{-7}$. The values of the parameters are $\mu = 0.06$, $\sigma = 0.25$ and $S_0 = 50$.



(a) True solution and Euler-Maruyama approximation

(b) True solution and Milstein approximation

Figure 5.1: True solution (in red) and approximations (in blue) using Euler-Maruyama and Milstein methods.

5.5.1 Order of convergence

Now our goal is to experimentally confirm the order of convergence of these numerical methods.

To test strong convergence, we fix the spacing intervals for the Brownian motion and use eight different time steps $\Delta t = 2^{p-1}\delta t$ for the numerical methods, with p ranging from 1 to 8. The Brownian motion is sampled on the interval $[0, 1]$ and has a time step of $\delta t = 2^{-11}$. The values of the parameters we used are $\mu = 0.06$, $\sigma = 0.25$, $S_0 = 50$ and $K = 100$. Because it is computationally heavy to compute the strong convergence, as it has to store the whole path and not just the end point, we only sample over 5000 paths, which we denote by M.

The errors were computed as in (5.10). For the true solution we used equation (2.22), and computed it for each random path.

Table 5.1 clearly shows a strong order of convergence of $\frac{1}{2}$ for the Euler-Maruyama and an order of 1 for the Milstein method. In fact, Euler-Maruyama requires cutting the time step by a factor of 4 to reduce the error to half whereas Milstein only requires cutting it by a factor of 2.

Table 5.1: Strong convergence of Euler-Maruyama and Milstein schemes for $\mu = 0.06$, $\sigma = 0.25$, $S_0 = 50$, $T = 1$ and $M = 5000$

Δt	Error values Euler-Maruyama	Error values Milstein
2^{-4}	0.466616	0.041683
2^{-5}	0.336610	0.021142
2^{-6}	0.236714	0.010645
2^{-7}	0.167614	0.005340
2^{-8}	0.117789	0.002685
2^{-9}	0.083629	0.001338
2^{-10}	0.058646	0.000673
2^{-11}	0.041627	0.000334

To test weak convergence we only need the end points of the sample paths, so we compute them all simultaneously. We compute 500 000 paths.

Since weak convergence only requires the average of the solution, one way to compute the average of the true solution would be to simply use the fact that $E[S(T)] = S_0 e^{rT}$, where E denotes the expected value, which is the approach used by Higham in [16]. However, in our experiments we chose to compute the true solution, by using equation (2.22), for each random path (simultaneously), and then take the average. The reason for this was the fact that using $E[S(T)] = S_0 e^{rT}$ required a larger number of sample paths to obtain convergence than computing the average of the true solution, which increased the computational time.

In table 5.2 we can see that both Euler-Maruyama and Milstein have an experimental weak order of convergence of 1. Indeed, to reduce the error by half we need to cut the time step by a factor of 2. The errors were computed as in (5.16) with $f(X) = X$.

Table 5.2: Weak convergence of Euler-Maruyama and Milstein schemes for $\mu = 0.06$, $\sigma = 0.25$, $S_0 = 50$, $T = 1$ and $M = 500\ 000$

Δt	Error values Euler-Maruyama	Error values Milstein
2^{-4}	0.007110	0.006015
2^{-5}	0.003127	0.003020
2^{-6}	0.001713	0.001492
2^{-7}	0.001467	0.000746
2^{-8}	0.000812	0.000382
2^{-9}	0.000396	0.000189
2^{-10}	0.000107	0.000093
2^{-11}	0.000033	0.000047

Figure 5.2 displays the plotted data of tables 5.1 and 5.2 on a loglog scale. If we consider (5.11) and (5.17) and the inequalities hold with approximate equality, we can take the logs and obtain, respectively

$$\log(e_{\Delta t}^{\text{Strong}}) \approx \log(C) + \frac{1}{2}\log(\Delta t) \tag{5.32}$$

$$\log(e_{\Delta t}^{\text{Weak}}) \approx \log(C) + \log(\Delta t) \tag{5.33}$$

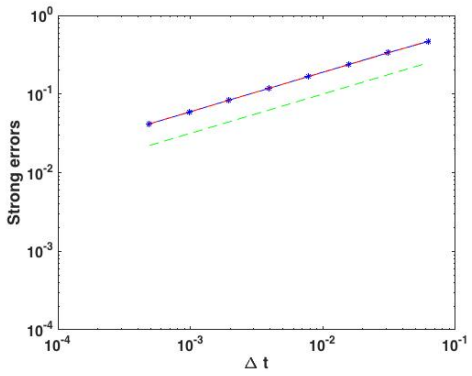
Expressions (5.32) and (5.33) are represented by the blue asterik. Further, assuming that for some constants C and q the relations $e_{\Delta t}^{\text{Strong}} = C\Delta t^q$ and $e_{\Delta t}^{\text{Weak}} = C\Delta t^q$ hold, such that

$$\log(e_{\Delta t}^{\text{Strong}}) = \log(C) + q\log(\Delta t) \tag{5.34}$$

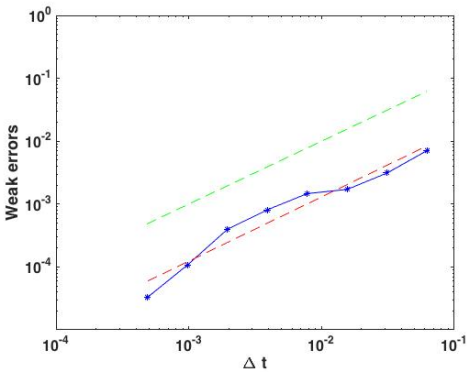
and

$$\log(e_{\Delta t}^{\text{Weak}}) = \log(C) + q\log(\Delta t) \tag{5.35}$$

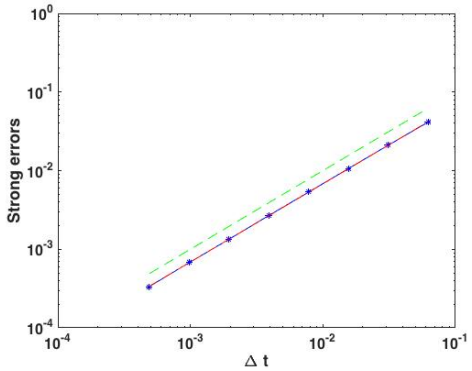
we can build a linear regression using a least squares fit for $\log(C)$ and q. These regressions are represented by the red dashed lines. The green dashed line is a reference slope with the theoretical order of convergence, which is $\frac{1}{2}$ for the strong convergence of the Euler-Maruyama method and 1 in the other cases.



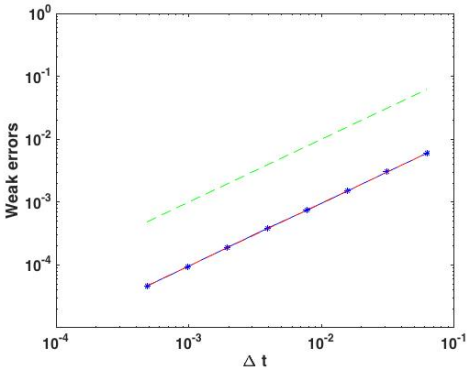
(a) Strong convergence - Euler-Maruyama ($q = 0.5003$)



(b) Weak convergence - Euler-Maruyama ($q = 1.0217$)



(c) Strong convergence - Milstein ($q = 0.9949$)



(d) Weak convergence - Milstein ($q = 1.0007$)

Figure 5.2: Strong and weak convergence for the Euler-Maruyama and Milstein methods. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

From the estimated regressions we can further confirm that both the experimental strong ($q = 0.5003$) and weak ($q = 1.0217$) orders of convergence of the Euler-Maruyama method agree with the theoretically predicted values. Moreover, the Milstein scheme also has both experimental strong ($q = 0.9949$ and weak ($q = 1.0007$) orders of convergence in good accordance with the theoretical values.

5.5.2 Numerical stability

We now want to experimentally test the numerical stability of the explicit Euler-Maruyama and Milstein schemes. For our test equation we will use

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{5.36}$$

where μ and σ are constant real parameters. As we have seen in section 5.4, the equilibrium solution of the SDE is asymptotically stable if

$$\mu - \frac{1}{2}\sigma^2 < 0 \tag{5.37}$$

which allows μ to be positive. However, the equilibrium solution is only stable in the mean-square sense if

$$\mu + \frac{1}{2}\sigma^2 < 0 \tag{5.38}$$

For this reason we will verify the stability of the schemes using $\mu > 0$ for the asymptotic stability and $\mu < 0$ to test mean-square stability. Further, in order to the explicit Euler-Maruyama scheme to be stable, in addition to the above mentioned conditions, the parameters have to verify $E\left[\log|1 + a\Delta t + b\sqrt{\Delta t}\mathcal{N}(0,1)|\right] < 0$ for asymptotic stability and $|1 + a\Delta t|^2 + |b|^2\Delta t < 1$ for mean-square stability.

Following Higham [16], to test mean-square stability we solve equation (5.36) with a constant initial value $S_0 = 1$ over $T = [0, 20]$. For parameters we chose $\mu = -2$ and $\sigma = 0.45$, which clearly satisfy condition (5.38). We apply Euler-Maruyama (5.3) to equation (5.36) and sample through 500 000 random paths using three different step sizes $\Delta t = \{1, \frac{1}{2}, \frac{1}{4}\}$. To test asymptotic stability we chose $\mu = 0.01$ and $\sigma = 0.45$, which satisfy condition (5.37) but not (5.38). Since asymptotic stability concerns a probability 1 event we consider only one random path on $T = [0, 1000]$ and used the same three time steps for the mean-square stability.

Figure 5.3 shows the test of mean-square and asymptotic stability for the explicit Euler-Maruyama method.

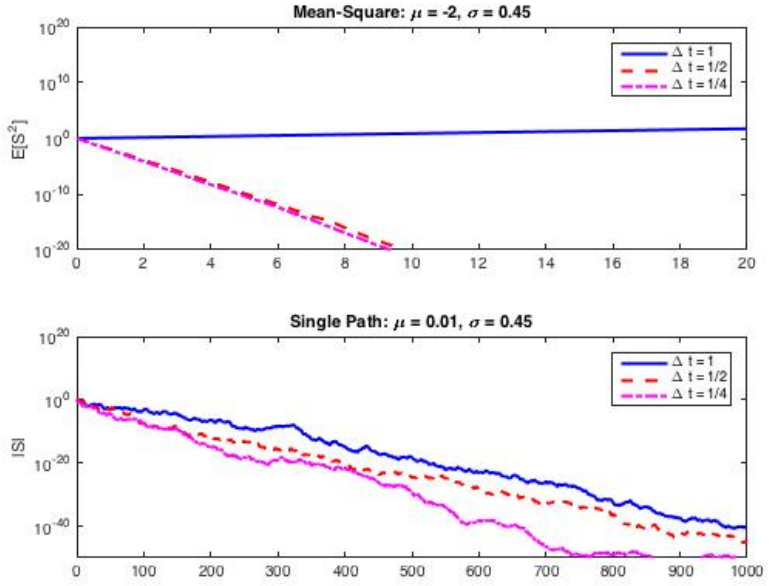


Figure 5.3: Test for mean-square and asymptotic stability of explicit Euler-Maruyama method

From figure 5.3 it is clear that condition (5.28) is only satisfied for the smaller time steps, so the Euler-Maruyama method is only stable in the mean square sense for $\Delta t \in \{\frac{1}{2}, \frac{1}{4}\}$. On the other hand, all the time steps seem to satisfy condition (5.29), so the solution seems to approach zero for all the time steps.

To test the stability of the Milstein scheme we chose the same test parameters and initial condition. As the condition for asymptotic stability of the scheme is the same as that of Euler-Maruyama scheme, it is only natural that for the same choice of parameters, the method is stable as well. In what concerns asymptotic stability, the Milstein scheme is also stable for all the time steps.

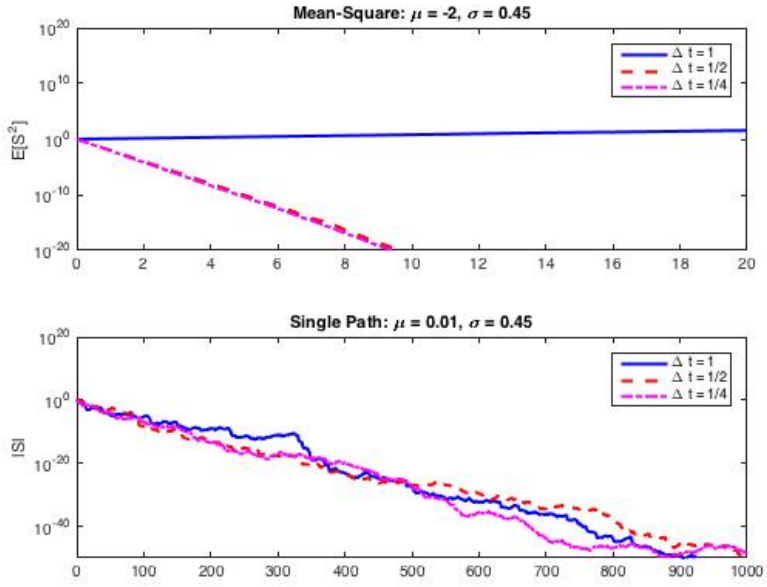


Figure 5.4: Test for mean-square and asymptotic stability of explicit Milstein method

5.6 Monte Carlo method for European option pricing

The idea of the Monte Carlo method when used in option pricing is to estimate the value of an option by simulating a large number of sample values of S_T, calculate the payoffs and find the estimated option price as the average of the discounted simulated payoffs. In fact, according to the law of large numbers this average will converge to the exact option value. To approximate the price of the underlying asset at maturity, S_T, we will implement Euler-Maruyama and Milstein schemes.

In option pricing, the Monte Carlo method uses the risk neutral valuation, that is, uses $\mu = r$. We have already discussed an argument using replicating strategies, however this involves many calculations so now we will see a simpler approach in which we use a change in probability measure to make the stock earn the risk-free rate in mean. It is worth noticing that in this approach we do not assume that the investors are risk neutral, here the idea is that risk premium adjustments are reflected in the differences between the risk-neutral probabilities and the "real world" probabilities. Actually, Girsanov's theorem A.0.2 allows us to account for the risk premiums of the investors in the Brownian motion process itself and see this as a new Brownian motion under the change of probability measure. We denote by $\mathbb{P}$ the probability measure of the "real" world and by $\mathbb{Q}$ the risk neutral probability measure. Therefore, under the risk neutral measure, the value of an option at time t is given by

$$V(S_t, t) = e^{-r(T-t)} E^{\mathbb{Q}}[h(S_T, T)] \tag{5.39}$$

where $h(S_T, T)$ is the expected payoff of the option. Under the new probability measure the discounted price of the option is a martingale, which is why it is also called *equivalent martingale measure.*

It is worth noting that the way (5.39) is defined, is only valid when interest rates are independent of the price of the underlying asset, because when they are dependent they must be estimated as well, and

therefore equation (5.39) should be

$$V(S_t, t) = E^{\mathbb{Q}}[h(S_0, 0)] \tag{5.40}$$

where $E^{\mathbb{Q}}[h(S_0, 0)]$ is the expected discounted payoff under the risk neutral measure. In fact, equations (5.39) and (5.40) are equivalent when the interest rate is constant or depends only on time.

5.6.1 Constant parameters

We start by introducing Monte Carlo simulations for the classic Black-Scholes asset model, in which the drift and diffusion coefficients are constant, although we will later simulate models with time and asset price dependent coefficients. Further, we will restrict our attention to vanilla European call options only, as the results for put options can be easily derived using the put-call-parity (2.3).

Consider the asset price model defined by the following stochastic differential equation

$$dS_t = rS_t dt + \sigma S_t dW_t \tag{5.41}$$

where r is the constant expected return on the asset under the risk neutral valuation, σ is the constant volatility and W is a Brownian motion.

We have just seen that the value of a call option at time t, with constant interest rate, is just its payoff at the maturity discounted to time t under the risk neutral measure. So, it suffices to know the value of the option at its expiration and discount that value with the factor $e^{-r(T-t)}$ to obtain the price of the option. In this work we are interested in knowing the price of the option at the present, that is, for $t = 0$, so expression (5.39) simplifies to

$$V(S_0, 0) = e^{-rT} E^{\mathbb{Q}}[\max\{S(T) - K, 0\}] \tag{5.42}$$

where $E^{\mathbb{Q}}[\max\{S(T) - K, 0\}]$ is the expected payoff of the call under the risk neutral measure.

To estimate the expected payoff of the option at the maturity we take the arithmetic mean

$$\bar{h}(S_T, T) = \frac{1}{M} \sum_{i=1}^{M} h(S_T^{i,L}, T) \tag{5.43}$$

where h is the payoff of the option and $S_T^{i,L}$ is the approximate value of S_T over the i-th sample path using L time steps.

In this thesis the Monte Carlo algorithm for simulating the price of a European call option uses Euler-Maruyama or Milstein methods to approximate the price of the underlying stock. Before presenting the algorithm we will first consider the discretisation of equation (5.41). Applying the Euler-Maruyama scheme (5.3) to equation (5.41), we get the following discretisation

$$S_j = S_{j-1} + rS_{j-1}\Delta t + \sigma S_{j-1}\sqrt{\Delta t} z_i, \quad j = 1, 2, ..., L \quad \text{and} \quad i = 1, ..., M \tag{5.44}$$

where z_i is a random variable sampled from $\mathcal{N}(0, 1)$, L is the number of time steps and M is the number of sample paths.

Applying the Milstein scheme (5.21) to equation (5.41) the discretisation is as follows

$$S_j = S_{j-1} + rS_{j-1}\Delta t + \sigma S_{j-1}\sqrt{\Delta t} z_i + \frac{1}{2}\sigma^2 S_{j-1}\left[(\sqrt{\Delta t} z_i)^2 - \Delta t\right], \quad j = 1, 2, ..., L \text{ and } i = 1, ..., M \tag{5.45}$$

Algorithm 4 is a Monte Carlo algorithm to simulate the price of a vanilla European call, assuming that the behaviour of the asset price is described by equation (5.41). For the approximation of the asset price at each time step we can use equation (5.44) or (5.48).

Algorithm 4 Monte Carlo for European option pricing with constant volatility and interest rate

1. Define r, σ, T, L, S_0, Δt, K

   1: **for** $j = 1 : L$ **do**
   2: Compute M random paths of a Brownian motion
   3: Implement Euler-Maruyama or Milstein methods to approximate the values $S_i(T)$,
   4: $i = 1, 2..., M$, for each random path
   5: **end for**

2. Compute the payoff of the option at the maturity $h_i(S_T^i, T) = \max\{S_T^i - K, 0\}$ for each random path $i = 1, 2, ..., M$

3. Approximate the expected value of $h(S_T, T)$ by calculating the average as in equation (5.43)

4. Calculate the present value of the estimator of step 3 to obtain the price V of the option: $\bar{V}(S_0, T) = e^{-rT}\bar{h}(S_T, T)$

Error and rate of convergence

The estimator given in equation (5.43) has two types of error associated with it: a statistical error and a discretisation error. The first one is a consequence of the central limit theorem and decreases at the rate $O(1/\sqrt{M})$ (see [14]). In fact, the standard error of the simulation estimator is $\sigma/\sqrt{M}$, where σ is the standard deviation of the stock price. The second error comes from the fact that we are discretising a continuous time process. Since we are pricing European options and these are path independent, because the price for the buyer depends exclusively on the terminal price, we are only interested in weak error estimates, which are given by $|E[V(S_T)] - E[V(S_T^{i,n})]|$. Hence, the Monte Carlo method using Euler-Maruyama or Milstein methods with weak order of convergence, has an order of convergence of $O(1/\sqrt{M}) + O(\Delta t)$, where Δt is the time step used in the discretisation scheme.

We run algorithm 4 for different values of M in order to experimentally verify that the standard error of the simulation estimator $\sigma/\sqrt{M}$ decreases at the rate of $O(1/\sqrt{M})$. Table 5.3 displays the values of the call options obtained by the Euler-Maruyama and Milstein methods and the standard error, for the different values of M. The values of the parameters used are $\sigma = 0.3$, $r = 0.07$, $S_0 = 80$, $K = 100$ and $T = 1$. For the time step we used 2^{-10}.

Table 5.3: Monte Carlo method convergence rate

M	Euler-Maruyama	Weak error	Milstein	Weak error	$\frac{\sigma}{\sqrt{M}}$
100	7.8484	0.0003343	7.8487	0.0000304	0.03000
400	6.1288	0.0017259	6.1305	0.0000230	0.01500
1600	5.7885	0.0006906	5.7892	0.0000229	0.00750
6400	5.3594	0.0003690	5.3597	0.0000220	0.00375
25600	5.0497	0.0000647	5.0497	0.0000212	0.00188
102400	4.9466	0.0000329	4.9466	0.0000210	0.00094
409600	5.0088	0.0000572	5.0089	0.0000213	0.00047
1638400	5.0172	0.0000341	5.0172	0.0000214	0.00023

where the standard error is given by $\sigma/\sqrt{M}$.

It is clear that when we multiply M by 4, $\sigma/\sqrt{M}$ decreases by half, which is theoretically correct.

5.6.2 Time-dependent parameters

The Monte Carlo method described in the previous section for constant parameters can be easily extended to time-dependent parameters. In fact, the algorithm is very similar and for this reason it is not shown here. The main difference is that as the interest rate and the volatility are functions of time, we need to compute them at each time step. It is also worth noting that in order to discount the option price to the present we must use as discount factor $e^{\int_0^T r(x)dx}$ and therefore we need to solve the integral in the exponent. The present value of the option is therefore given by

$$V(S_0, 0) = e^{-\int_0^T r(x)dx} E^{\mathbb{Q}}[h(S_T, T)] \tag{5.46}$$

where $E^{\mathbb{Q}}[h(S_T, T)]$ is approximated as (5.43). Note that the integral in exponent does not depend on S so we do not need to compute its expected value.

The time-dependent asset model is described by the following equation

$$dS_t = r_t S_t dt + \sigma_t S_t dW_t \tag{5.47}$$

where r_t and σ_t are deterministic, continuous functions of time. Applying the Euler-Maruyama scheme (5.3) to equation (5.47) we obtain the following discretisation

$$S_j = S_{j-1} + r_{j-1} S_{j-1} \Delta t + \sigma_{j-1} S_{j-1} \sqrt{\Delta t} z_i, \quad j = 1, 2, ..., L \quad \text{and} \quad i = 1, ..., M \tag{5.48}$$

where z_i is a random variable sampled from $\mathcal{N}(0, 1)$, L is the number of time steps, M is the number of sample paths, $r_j = r(t_j)$ and $\sigma_j = \sigma(t_j)$.

Applying the Milstein scheme (5.21) to equation (5.47) the discretisation is as follows

$$S_j = S_{j-1} + r_{j-1} S_{j-1} \Delta t + \sigma_{j-1} S_{j-1} \sqrt{\Delta t} z_i + \frac{1}{2} \sigma_{j-1}^2 S_{j-1} \left[(\sqrt{\Delta t} z_i)^2 - \Delta t \right], \quad j = 1, 2, ..., L \text{ and } i = 1, ..., M \tag{5.49}$$

Recall that we concluded section 3.2.1 by discussing why would it be interesting to consider time-dependent parameters with jump discontinuities and that the general theory for SDE's would not apply

in that case. In fact, we have seen in section 2.2.3 that the drift and diffusion coefficients have to be globally Lipschitz, which clearly is not the case when we allow these functions to be discontinuous. As this condition is not satisfied, we cannot guarantee that the Euler and Milstein schemes will converge. Nonetheless we will see this experimentally in section 6.2.3. Although the algorithm is not included, it is important to explain how the price of the option is discounted to the present when we consider coefficients with a jump discontinuity. Since a jump discontinuity in the drift coefficient means that we will have two different values for the interest rate, we will have to discount the price in two steps, that is,

$$V(S_0, 0) = \prod_{n=1}^{2} e^{-r_n(t_n - t_{n-1})} E^{\mathbb{Q}}[h(S_T, T)] \tag{5.50}$$

5.6.3 Time and asset price dependent parameters

We now consider a more complex model, where both the expected rate on return and the volatility depend on time and on the price of the underlying stock. The asset model that describes the dynamics of the underlying asset is the following

$$dS_t = r(t, S_t)S_t dt + \sigma(t, S_t)S_t dW_t \tag{5.51}$$

where $r(t, S_t)$ and $\sigma(t, S_t)$ are continuous, deterministic functions of time and of the underlying asset price.

We will approach this equation in two different ways. We will start by discretising it using explicit Euler-Maruyama and Milstein schemes and after we will use semi-implicit schemes with $\theta = \frac{1}{2}$.

Applying the Euler-Maruyama scheme (5.3) to equation (5.51) we obtain the following discretisation

$$S_j = S_{j-1} + r(\tau_{j-1}, S_{j-1})S_{j-1}\Delta t + \sigma(\tau_{j-1}, S_{j-1})S_{j-1}\sqrt{\Delta t}z_i, \quad j = 1, 2, ..., L \quad \text{and} \quad i = 1, ..., M \tag{5.52}$$

where z_i is a random variable sampled from $\mathcal{N}(0, 1)$, L is the number of time steps and M is the number of sample paths.

Applying the Milstein scheme (5.21) to equation (5.51) the discretisation is as follows

$$\begin{aligned} S_j = {} & S_{j-1} + r(\tau_{j-1}, S_{j-1})S_{j-1}\Delta t + \sigma(\tau_{j-1}, S_{j-1})S_{j-1}\sqrt{\Delta t}z_i + \\ & + \frac{1}{2}\sigma(\tau_{j-1}, S_{j-1})\frac{\partial \sigma}{\partial S}S_{j-1}\left[(\sqrt{\Delta t}z_i)^2 - \Delta t\right], \quad j = 1, 2, ..., L \text{ and } i = 1, ..., M \end{aligned} \tag{5.53}$$

Now that the interest rate and the volatility depend on the price of the underlying asset, we will have different values for the interest rate and for the volatility at each time step and for each different sample path. As we have previously discussed, this means that it is no longer equivalent to compute the present value of the option as the discounted expected payoffs or as the expectation of the discounted payoffs, because the discount rate has to be estimated as well. Therefore in this model, we compute the present value of the option as the expected value of the discounted payoffs, that is

$$V(S_0, 0) = E^{\mathbb{Q}}\left[h(S_0, 0)\right] \tag{5.54}$$

We will now see how we can approximate the expected value in (5.54).

We start by computing the payoff of the option for each sample path

$$h(S_T^i, T) = \max\{S_T^i - K, 0\}, \quad i = 1, 2, ..., M \tag{5.55}$$

where M is the number of sample paths.

Since the discount rates will be different for each sample path we need to discount the payoff for each sample path one at time, using the corresponding discount rate for each time interval. The payoffs are discounted to the present in the following way

$$h(S_0^i, 0) = \prod_{n=0}^{L-1} e^{-r_n(t_{n+1}-t_n)} h(S_T^{i,L}, T), \quad i = 1, 2, ..., M \tag{5.56}$$

where L is the number of time steps. Only when the payoff of the option is discounted to the present, for each sample path, can we take the mean, yielding

$$V(S_0, 0) = \frac{1}{M} \sum_{i=1}^{M} h(S_0^{i,L}, 0) \tag{5.57}$$

where $S_T^{i,L}$ is the approximate value of S_T over the *i-th* sample path using L time steps. Note that this is a different estimator from the estimator (5.43) introduced in section 5.6.1.

We now consider the semi-implicit schemes. Recall that when there is implicitness in the coefficients it requires to solve an additional equation at each time step, which is computationally heavier. One approach is to use a predictor-corrector scheme. We start by defining the predictor, which is the solution of an explicit scheme at level n, which in our case is the explicit Euler-Maruyama scheme (5.3). Applying this scheme to equation (5.51) leads to the following discretisation

$$\bar{S}_j = S_{j-1} + r(\tau_j, S_j) S_j \Delta t + \sigma(\tau_{j-1}, S_{j-1}) S_{j-1} \sqrt{\Delta t} z_i \tag{5.58}$$

For the corrector scheme we use the generalisation of the deterministic trapezoidal method, which is the scheme (5.7) with $\theta = \frac{1}{2}$. Applying it to equation (5.51) yields

$$S_j = S_{j-1} + \frac{1}{2} \left[r(\tau_j, \bar{S}_j) \bar{S}_j + r(\tau_{j-1}, S_{j-1}) S_{j-1} \right] \Delta t + \sigma(\tau_{j-1}, S_{j-1}) S_{j-1} \sqrt{\Delta t} z_i \tag{5.59}$$

In our experiments we will also apply the predictor-corrector method using the explicit Milstein method and the corresponding generalisation of the deterministic trapezoidal method.

Applying the explicit Milstein scheme (5.21) to equation (5.51) we get the predictor scheme

$$\begin{aligned} \bar{S}_j = S_{j-1} &+ r(\tau_{j-1}, S_{j-1}) S_{j-1} \Delta t + \sigma(\tau_{j-1}, S_{j-1}) S_{j-1} \sqrt{\Delta t} z_i + \\ &+ \frac{1}{2} \sigma(\tau_{j-1}, S_{j-1}) \frac{\partial \sigma}{\partial S} S_{j-1} \left[(\sqrt{\Delta t} z_i)^2 - \Delta t \right] \end{aligned} \tag{5.60}$$

and applying the family of semi-implicit Milstein schemes (5.23) with $\theta = \frac{1}{2}$ we get the corrector scheme

$$\begin{aligned} S_j = S_{j-1} &+ \frac{1}{2} \left[r(\tau_j, \bar{S}_j) \bar{S}_j + r(\tau_{j-1}, S_j - 1) S_{j-1} \right] \Delta t + \sigma(\tau_{j-1}, S_{j-1}) S_{j-1} \sqrt{\Delta t} z_i + \\ &+ \frac{1}{2} \sigma(\tau_{j-1}, S_{j-1}) \frac{\partial \sigma}{\partial S} S_{j-1} \left[(\sqrt{\Delta t} z_i)^2 - \Delta t \right] \end{aligned} \tag{5.61}$$

Algorithm 5 is Monte Carlo algorithm to simulate the price of a vanilla European call, using Euler or Milstein predictor-corrector scheme to approximate the value of the underlying asset at each time step. Here function f is used to designate the predictor scheme, which can be (5.58) or (5.60), and function g designates the respective corrector schemes (5.59) and (5.61).

Algorithm 5 Monte Carlo for option pricing when $r(t, S_t)$ and $\sigma(t, S_t)$ are deterministic functions of time and of the stock price

1. Define T, L, S_0, Δt, K, M

for $i = 1 : M$ **do**
  Define $dW = \sqrt{\Delta t} z_i$
  Initialize $t = 0$, $W = 0$, $S^i_{temp} = S_0$, $\sigma_0 = \sigma(S_0, t_0)$ and $r_0 = r(S_0), t_0$
  for $j = 1 : L$ **do**
    Define $t = t + \Delta t$, $W = W + dW$, $\bar{S}^i_j = f(S^i_{temp})$, $S^i_j = g(\bar{S}^i_j, S^i_{temp})$, abserror$=|\bar{S}^i_j - S^i_j| > \epsilon$
      while abserror$> \epsilon$ **do**
       $\sigma_j = \sigma(\bar{S}^i_j, t)$, $r_j = r(\bar{S}^i_j, t)$
        $\bar{S}^i_j = S^i_j$, $S^i_j = g(\bar{S}^i_j, S^i_{temp})$, abserror$=|\bar{S}^i_j - S^i_j|$
      end while
    $S^i_{temp} = S^i_j$
  end for
  Compute the payoff of the option, $h_i(\tilde{S}^i_T, T) = \max\{\tilde{S}^i_T - K, 0\}$ for each sample path
  Discount $h_i(\tilde{S}^i_T, T)$ at each time step, using the corresponding value of r for that time step, until the present time.
end for

2. Take the mean of the discounted payoffs to get the price V of the option : $\bar{V}(S_0, 0) = \frac{1}{M} \sum_{i=1}^{M} h(S_0^{i,n}, 0)$

5.6.4 Heston model

In the Heston model we need to discretise not only the stochastic process $\{S_t\}_{t\geq 0}$ but also the stochastic process $\{\nu_t\}_{t\geq 0}$.

Applying the Euler-Maruyama scheme (5.3) to equations equations (3.64) and (3.65) we get the following discretisations

$$S_t = S_{t-1} + rS_{t-1}\Delta t + \sqrt{\nu_{t-1}} S_{t-1} \sqrt{\Delta t} Z^s_t \tag{5.62}$$

$$\nu_t = \nu_{t-1} + k(\theta - \nu_{t-1})\Delta t + \xi \sqrt{\nu_{t-1}} \sqrt{\Delta t} Z^\nu_t \tag{5.63}$$

where $\{Z^s_t\}_{t\geq 0}$ and $\{Z^\nu_t\}_{t\geq 0}$ are standard normal random variables with correlation ρ. These variables can be expressed as a function of independent standard random variables

$$Z^s_t = Z^1_t \tag{5.64}$$

$$Z^\nu_t = \rho Z^1_t + \sqrt{1 - \rho^2} Z^2_t \tag{5.65}$$

where $\{Z^1_t\}_{t\geq 0}$ and $\{Z^2_t\}_{t\geq 0}$ are two independent standard normal random variables. Writing the variables in this form will be useful in the implementations of the schemes.

Applying the Milstein scheme (5.21) to equations (3.64) and (3.65) we get the following discretisations

$$S_t = S_{t-1} + rS_{t-1}\Delta t + \sqrt{\nu_{t-1}}S_{t-1}\sqrt{\Delta t}Z_t^s + \frac{1}{2}\nu_{t-1}S_{t-1}\Delta t\left[(Z_t^s)^2 - 1\right] \tag{5.66}$$

$$\nu_t = \nu_{t-1} + k(\theta - \nu_{t-1})\Delta t + \xi\sqrt{\nu_{t-1}}\sqrt{\Delta t}Z_t^\nu + \frac{1}{4}\xi^2\Delta t\left[(Z_t^\nu)^2 - 1\right] \tag{5.67}$$

The discretisation of the Heston model gives rise to several issues. First, when simulating the bivariate process (S_t, ν_t), as ν_t follows a Cox–Ingersoll–Ross (CIR) process (see [17]), the use of discretisation schemes such as Euler and Milstein, might generate negative values of ν_t. The fullfilement of the condition

$$\frac{2k\theta}{\xi^2} > 1 \tag{5.68}$$

known as the *Feller condition*, ensures that ν is strictly positive. This condition is essential to obtain good discretisations, however, in practice it is rarely satisfied because ξ usually takes large values [**?**]. One approach to overcome this problem is to fix the negative values of ν as they arise. There are at least two ways to do this: the *absorption assumption*, where if $\nu < 0$ then $\nu = 0$ or the *reflection assumption*, where if $\nu < 0$ then $\nu = |\nu|$. The drawback to the first approach is that it creates zero variances, which is not realistic. The drawback to the second approach is that it transforms low volatilities into high volatilities. Nonetheless, in our experiments we will adopt the second approach.

Another issue that arises when we discretise the Heston Model is that it also violates the Lipschitz condition due to the square root in the diffusion coefficient, which means that the convergence of the numerical schemes is not guaranteed.

Algorithm 6 provides the basic steps to approximate the price of a European call option using the Heston model discretised with the Euler-Maruyama or the Milstein schemes.

Algorithm 6 Monte Carlo for option pricing - Heston model

1. Define T, L, S_0, Δt, K, ν_0, θ, k, ξ, r

   1: **for** $j = 1 : L$ **do**
   2: Compute M random paths of a Brownian motion
   3: Generate two independent random variables Z_1 and Z_2 and define $Z_t^s = Z_t^1$ and $Z_t^\nu = \rho Z_t^1 + \sqrt{1-\rho^2}Z_t^2$
   4: Implement Euler-Maruyama or Milstein methods to approximate the values $\nu_i(T)$ for each random path
   5: Replace ν_t by $|\nu_t|$ to avoid negative volatilities
   6: Implement Euler-Maruyama or Milstein methods to approximate the values $S_i(T)$, $i = 1, 2..., M$, for each random path
   7: **end for**

2. Compute the payoff of the option $h_i(S_T^i, T) = \max\{S_T^i - K, 0\}$ for each random path $i = 1, 2, ..., M$

3. Approximate the expected value $h(S_T, T)$ by calculating the average as in equation (5.43)

4. Calculate the present value of the estimator of step 3 to get the price V of the option: $\bar{V}(S_0, T) = e^{-rT}\bar{h}(S_T, T)$

Chapter 6

Numerical experiments

6.1 Finite difference methods for PDE's

In this section we approximate the Black-Scholes formula in the heat equation form, as described in section 3.1.4. The idea is to approximate the function $u(x, \tau)$ and then convert it back to $V(S, t)$ to recover the price of the option and compute the relative error at $t = t = 0$. Again, we will only compute the relative error at the point $V(S_0, t_0)$.

We start by replicating the numerical experiments carried out in M.M. Fernandes thesis [12, p. 36] and compare the results, which are displayed on table 6.1. Next we repeat the same experiments for different values of the parameters. Those results will be later compared to the results obtained using numerical methods for stochastic differential equations to approximate the price of the underlying asset.

Table 6.1 shows the approximate price of call and put European options as well as the relative errors at the point $V(S_0, t_0)$. The values of the parameters are: $S_0 = 1000$, $T = 10$, $r = 0.1$, $\sigma = 0.4$ and $K = 500, 1000$ and 1500, which are the same values as the ones used in [12].

The results presented in this section were obtained by implementing the algorithms described in section 4.3. Note that the results were not computed by using the put-call-parity (2.3).

In order to be able to compare the three schemes we consider the same mesh spacing for all the methods, taking into account that the stability condition of the explicit method requires that $\chi \leq \frac{1}{2}$. In [12] the steps used were $\Delta x = 0.08$ and $\Delta\tau = 0.0032$ but in this work we considered $\Delta x = 0.05$ and $\Delta\tau = 0.001$, which might explain the slightly better results.

Like in [12], the best results are the ones obtained with the Crank-Nicolson and explicit schemes, which is theoretically correct. As discussed in section 4.2, imposing the condition $\chi \leq \frac{1}{2}$ on the implicit methods is not advantageous because the error in the Crank-Nicolson method will be close to that of the explicit method.

Table 6.1: Relative errors in the approximation of the function V using finite differences, for $S_0 = 1000$, $T = 10$, $r = 0.1$, $\sigma = 0.4$ and $K = \{500, 1000, 1500\}$

Scheme	Strike price	Approx. call price	True call price	Relative error	Approx. put price	True put price	Relative error
Explicit		835.48443		0.000445	19.79759		0.000090
Implicit	500	836.75167	835.8561	0.001071	19.78241	19.7958	0.000677
C-Nicolson		836.11744		0.000313	19.79000		0.000294
Explicit		715.19723		0.000545	83.45011		0.000195
Implicit	1000	716.45353	715.5869	0.001211	83.42401	83.4664	0.000507
C-Nicolson		715.82478		0.000332	83.43707		0.000351
Explicit		624.18058		0.000616	176.3732		0.000065
Implicit	1500	625.45168	624.5655	0.001418	176.36190	176.3847	0.000129
C-Nicolson		624.81552		0.000400	176.36753		0.000097

Table 6.2 displays the approximate prices for European call and put options as well as the relative errors, at the point $V(S_0, t_0)$. The values of the parameters are the same as the ones used in the next section to allow comparisons between finite difference schemes and Monte Carlo simulations. In our choice of parameters we always use $T = 1$ to allow faster computations, especially when doing Monte Carlo simulations. For the time and space steps we chose again $\Delta\tau = 0.001$ and $\Delta x = 0.05$, respectively. Also note that in this second experiment we vary the initial stock price instead of the strike price.

Table 6.2: Relative errors in the approximation of the function V using finite differences, for $T = 1$, $r = 0.07$, $\sigma = 0.3$, $K = 100$, and $S_0 = \{80, 100, 120\}$

Scheme	Initial stock price	Approx. call price	True call price	Relative error	Approx. put price	True put price	Relative error
Explicit		5.02058		0.001585	18.26275		0.000588
Implicit	80	5.02862	5.01263	0.003189	18.26121	18.25201	0.000504
C-Nicolson		5.02441		0.002350	18.26180		0.000536
Explicit		15.19732		0.000866	8.44019		0.001147
Implicit	100	15.15414	15.21050	0.003706	8.38503	8.44988	0.007675
C-Nicolson		15.17587		0.002277	8.41275		0.004394
Explicit		30.29774		0.000491	3.54131		0.005380
Implicit	120	30.27447	30.28288	0.000278	3.50365	3.522260	0.005282
C-Nicolson		30.28600		0.000103	3.52239		0.000036

It is worth noticing that when comparing the relative errors from table 6.1 with those from table 6.2, we see that, despite the space and time steps being the same, the results from table 6.1 are slightly better. This has to do with the fact that in table 6.2 we used an expiration date of one year, while on table 6.1 we used 10 years. Since the domain of τ is limited, this means that for the approximations on table 6.2 the maximum number of time steps were only 45, while in table 6.1 we were able to compute 800 time steps.

6.2 Monte Carlo simulations

6.2.1 Constant parameters

In this section we implement Euler-Maruyama and Milstein methods as Monte Carlo methods to approximate the price V of a European option, using the algorithm 4 described in section 5.6.1.

Since the payoff of a European option does not depend on the path by which the strike prices are reached, this type of options are path independent. This means that we are only interested in the value of the stock at the maturity. This allows us to compute all the sample paths simultaneously and store only the last value of S, for each path. For this reason it was possible to run 500 000 sample paths without it being too computationally costly.

The Brownian motion was discretised as in 5.1, i.e. $\Delta W = \sqrt{\Delta t} z_i$ and the variables z_i were computed as discussed in 5.5.1, except that in the experiments carried out in this chapter we used the arbitrarily chosen initial state 10 but we could have used any other state.

Table 6.3 displays the prices of call and put European options approximated using the Euler-Maruyama (5.3) and the Milstein (5.21) schemes. The weak relative errors are computed as

$$\frac{|E[h(S_0, 0)] - E[h(\tilde{S}_0, 0)]|}{|E[h(S_0, 0)]|} \tag{6.1}$$

where $h(\tilde{S}_0, 0)$ is the discounted payoff of the approximated prices of the underlying asset and $h(S_0, 0)$ is the payoff of the exact prices, computed using the exact solution to (5.41) which is given by (2.22). It is worth noting that we chose to compute the weak errors only, because in order to compute the strong errors, due to computational costs, we would have to use less sample paths in the approximation of the call. This would then lead to higher errors and we could wrongly conclude that the finite difference schemes for the Black-Scholes PDE would produce more accurate results than the Monte Carlo simulations, whereas with a higher number of sample paths the accuracy of the Monte Carlo simulations would increase.

The values of the parameters are the same as the ones used in section 6.1. Here we are also interested in analysing how the initial stock price influences the approximation, so we chose three different values for S_0 and fixed the strike price. For the discretisation schemes we used a time step $\Delta t = 2^{-7}$.

Table 6.3: Relative errors in the approximation of the function V, for $T = 1$, $r = 0.07$, $\sigma = 0.3$, $K = 100$, $S_0 = \{80, 100, 120\}$ and $M = 500\ 000$

Scheme	Initial stock price	Approx call price	True call price	Relative error	Approx put price	True put price	Relative error
Euler-M	80	5.02412	5.03275	0.001716	18.2234	18.2289	0.000303
Milstein		5.02731		0.001082	18.2250		0.000215
Euler-M	100	15.23101	15.23283	0.000119	8.4203	8.4182	0.000245
Milstein		15.22552		0.000480	8.4128		0.000642
Euler-M	120	30.32419	30.32099	0.000105	3.5034	3.4956	0.002247
Milstein		30.31428		0.000221	3.4911		0.001265

We observe that the Euler-Maruyama and the Milstein schemes have similar performances. It is also worth noting that in the approximation of a call option, a smaller initial stock price leads to greater relative errors, and for a put option the inverse situation happens. We compute the relative error because it should not be sensitive to the magnitude of the exact solution. Indeed, notice that for a call, a lower

initial stock price means a lower true value of the option and for a put a lower initial stock price means a higher true value of the option. Recall that the payoff of a call option is given by (2.1), which means that for values of S_T lower than K the option is worthless. Therefore, for a call, the lower the initial stock price, the more difficult is to profit from the option and hence the option price is lower.

6.2.2 Comparison of results

In this section we compare the results obtained with the finite difference schemes for the Black-Scholes PDE with the ones obtained with Monte Carlo simulations. These results are displayed in tables 6.2 and 6.3 respectively.

The finite difference schemes and the Monte Carlo methods have similar performances, although the statistical results from the Monte Carlo simulations seem slightly better. In fact, we used 500 000 sample paths in the Monte Carlo simulations, which suggests that the error associated with the standard error should be in the order of 10^{-3} and the one associated with the discretisation should be in the order of 10^{-3} as we used a time step of 0.0078. On the other hand, in the finite difference schemes, as we used a time step of 0.001 and a space step of 0.05, the errors should be in the order of 10^{-3}, even for the Crank-Nicolson scheme, as the time term becomes too small. However, the Monte Carlo method converges rather slowly and needs a high number of paths to produce good results. Nonetheless, it allows to relax some assumptions, such as the constant parameters, as we will see in the next section.

6.2.3 Time-dependent parameters

In this section we drop the Black-Scholes assumption that the parameters are constant. We start with the simplest model, where we assume that the volatility is a positive deterministic function of time and then extend this assumption to the interest rate.

Time-dependent volatility

This asset model is a particular case of model (3.53) that we have seen in section 3.2.1. In this model, however, the stock price follows a stochastic differential equation of the form

$$dS_t = rS_t dt + \sigma(t) S_t dW_t \tag{6.2}$$

whose explicit solution is given by $S_T = S_t e^{r(T-t) - \frac{1}{2}\int_t^T \sigma^2(x)dx + \int_t^T \sigma(x) dW_x}$. As the explict solution depends itself on stochastic integrals, in our numerical experiments we will study the convergence of the numerical schemes by using an approximation of the true solution with a small time step.

We start by considering a sinusoidal function because market volatility tends to exhibit oscillatory behaviour. We chose as function $\sigma(t) = 0.1 + 0.3t + 0.03\sin(30t)$ to have $0.1 \leq \sigma \leq 0.3704$, which are reasonable values for the volatility.

Tables 6.4 and 6.5 show the strong and weak errors, respectively, in the approximation of the value $V(S_0, t_0)$ of a call option at time $t_0 = 0$, with initial stock price $S_0 = 80$, strike price $K = 100$ and maturity $T = 1$. The approximations were computed for seven different time steps using Monte Carlo simulations with 5000 sample paths in table 6.4 and 500 000 sample paths in table 6.5. The implementation is carried out as discussed in section 5.6.2. To approximate the price of the stock at each time step we used the explicit Euler-Maruyama (5.3) and Milstein (5.21) methods applied to equation (6.2). The first and third columns show the price of the call computed using each of these methods, respectively.

As the exact solution is not known, to compute the strong and weak errors we used an approximation of the true solution using a small time step of 2^{-11} as the reference solution. The strong error was then computed as the expectation of the absolute value of the difference of the discounted payoffs for an approximation with a time step 2^{-11} and approximations with time steps Δt, where $\Delta t = \{2^{-3}, 2^{-4}, ..., 2^{-9}\}$ that is

$$e^{strong} = E[|h(\tilde{S}_0, 0) - h(\tilde{S}_0, 0)^{\Delta t}]| \tag{6.3}$$

where $h(\tilde{S}_0, 0)$ is the discounted payoff computed using the reference time step of 2^{-11}.

The reference values for the true solution are 3.9247 and 3.9252 for the strong convergence and 4.0065 and 4.0066 for the weak convergence, using the Euler-Maruyama and Milstein schemes, respectively.

As strong convergence is computationally costlier we only sampled through 5000 sample paths. The weak error was computed over 500 000 sample paths, by taking the absolute value of the difference between the expected values of the discounted payoffs, for an approximation with a time step 2^{-11} and approximations with time steps Δt, where $\Delta t = \{2^{-3}, 2^{-4}, ..., 2^{-9}\}$, i.e.

$$e^{weak} = |E[h(\tilde{S}_0, 0)] - E[h(\tilde{S}_0, 0)]^{\Delta t}| \tag{6.4}$$

Table 6.4: Strong errors in the approximation of the function V, for $T = 1$, $r = 0.07$, $\sigma = 0.1 + 0.3t + 0.03\sin(30t)$, $S_0 = 80$ $K = 100$, $M = 5000$

Time step	Euler Maruyama	Strong error	Milstein	Strong error
2^{-3}	4.1579	0.888660	4.2479	0.823665
2^{-4}	4.0273	0.556461	4.0738	0.472084
2^{-5}	3.9720	0.327195	3.9904	0.251241
2^{-6}	3.9467	0.204558	3.9636	0.127581
2^{-7}	3.9322	0.127688	3.9429	0.062061
2^{-8}	3.9270	0.076604	3.9327	0.029086
2^{-9}	3.9257	0.047783	3.9283	0.013087

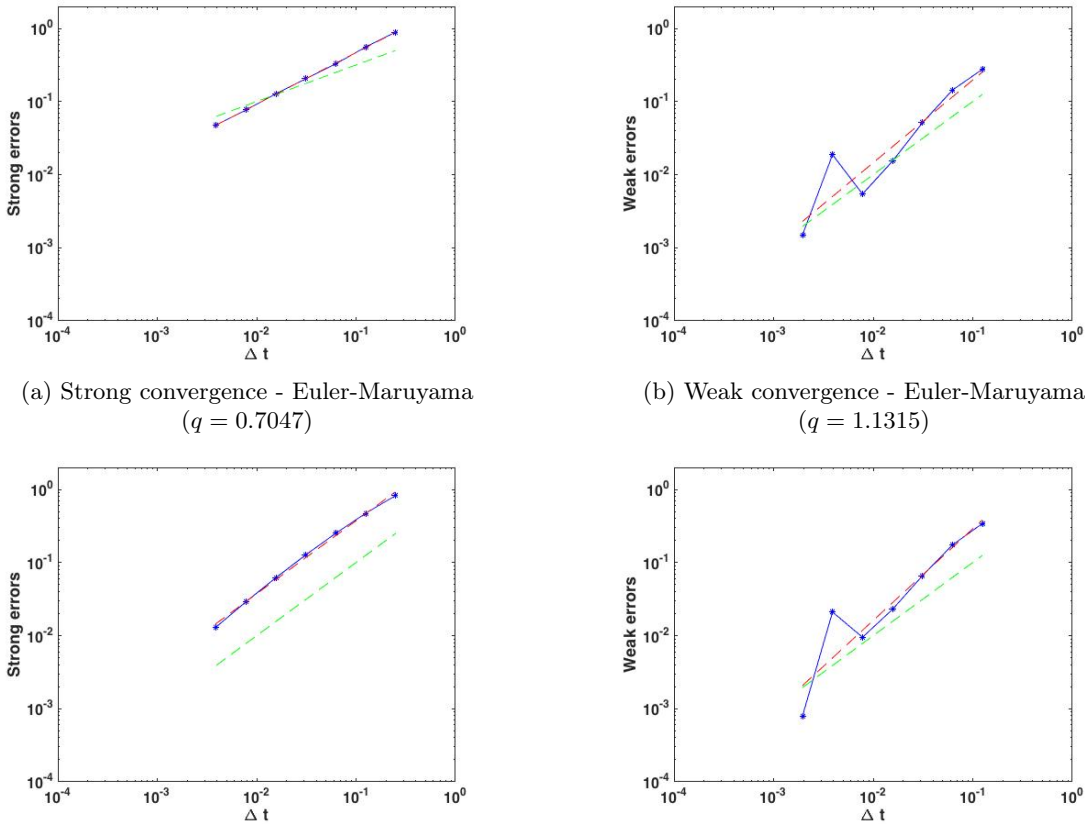
Table 6.5: Weak errors in the approximation of the function V, for $T = 1$, $r = 0.07$, $\sigma = 0.1 + 0.3t + 0.03\sin(30t)$, $S_0 = 80$ $K = 100$, $M = 500\ 000$

Time step	Euler Maruyama	Weak error	Milstein	Weak error
2^{-3}	4.2827	0.276142	4.3480	0.341385
2^{-4}	4.1499	0.143376	4.1804	0.173759
2^{-5}	4.0577	0.051156	4.0726	0.065942
2^{-6}	4.0218	0.015275	4.0295	0.022842
2^{-7}	4.0120	0.005427	4.0161	0.009499
2^{-8}	4.0252	0.018716	4.0276	0.021013
2^{-9}	4.0050	0.001500	4.0058	0.000796

By studying the strong error in table 6.4, we see that the Euler-Maruyama exhibits an experimental strong order of convergence above $\frac{1}{2}$, because when we cut the time step by a factor of 4 the error reduces to more than half. On the other hand, in order to reduce the error by half in the Milstein method

we only need to cut the time step by a factor of 2, which means that the experimental strong order of convergence of the Milstein method is 1. From table 6.5, an analysis of the weak error suggests that both the Euler-Maruyama and the Milstein methods also exhibit a weak order of convergence of 1, so both orders of convergence of the Euler-Maruyama and the Milstein methods agree with the theory.

Figure 6.1 shows the plotted values of the strong and weak errors for the Euler-Maruyama and Milstein methods on tables 6.4 and 6.5, and an estimated regression using a least squares fit.



(a) Strong convergence - Euler-Maruyama ($q = 0.7047$)

(b) Weak convergence - Euler-Maruyama ($q = 1.1315$)

(c) Strong convergence - Milstein ($q = 0.9995$)

(d) Weak convergence - Milstein ($q = 1.2544$)

Figure 6.1: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on tables 6.4 and 6.5. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

As expected, the estimated regression for the strong convergence of the Euler-Maruyama scheme has a slope $q = 0.7047$, which is above the theoretically predicted value and for the Milstein scheme is $q = 0.9995$. Furthermore, the estimated regression for weak convergence of the Milstein scheme is also close to the reference value, with a slope $q = 1.2544$ slightly above the theoretical value. Moreover, the estimated regression for the weak convergence of the Euler-Maruyama scheme has a slope $q = 1.1315$, which suggests that the experimental weak order of convergence of the method is also slightly above the theoretical value of 1. Overall, we see that the Milstein method converges faster than the Euler-Maruyama in both weak and strong sense.

Time-dependent risk-free rate and volatility

We now introduce in the model a new time-dependent parameter, the risk-free rate. In section 3.2.1. we have seen that in this model the price of the underlying stock follows a stochastic differential equation

of the form $dS_t = r(t)S_t dt + \sigma(t)S_t dW_t$, and that its explicit solution also depends on stochastic integrals. So, we will once again study the convergence of the numerical methods by using an approximation of the true solution with a small time step.

We start by considering $r(t)$ and $\sigma(t)$ continuous functions. We used the same function for the volatility as in the previous model and for the interest rate we also chose a sinusoidal function because the interest rate can also exhibit oscillatory behaviour. We chose the function $r(t) = 0.01 + 0.03t + 0.03\sin(60t)$, to have $0.01 \leq r \leq 0.0309$, which are reasonable values for the risk free rate.

Tables 6.6 and 6.7 show the strong and weak errors, repectively, in the approximation of the price $V(S_0, t_0)$ of a call option at time $t_0 = 0$, with initial stock value $S_0 = 80$, strike price $K = 100$ and maturity $T = 1$. The approximations were computed for seven different time steps using a Monte Carlo simulation with 5000 sample paths in table 6.6 and 500 000 sample paths in table 6.7 just like in the previous model. The price of the underlying asset is also approximated using explicit Euler-Maruyama and Milstein schemes, that is, using discretisations (5.48) and (5.49), which we discussed in section 5.6.2. The strong and weak errors are computed as in the previous model.

The reference values for the true solution are 3.0264 and 3.0265 for the strong convergence and 3.0873 and 3.0876 for the weak convergence, using the Euler-Maruyama and Milstein schemes, respectively.

Table 6.6: Strong errors in the approximation of the function V, for $T = 1$, $r = 0.01 + 0.03t + 0.03\sin(60t)$, $\sigma = 0.1 + 0.3t + 0.03\sin(30t)$, $S_0 = 80$ $K = 100$, $M = 5000$

Time step	Euler Maruyama	Strong error	Milstein	Strong error
2^{-3}	3.3863	0.799844	3.5017	0.766539
2^{-4}	3.1081	0.462963	3.1652	0.396126
2^{-5}	3.0736	0.269586	3.0954	0.213058
2^{-6}	3.0497	0.171053	3.0676	0.108745
2^{-7}	3.0354	0.105628	3.0463	0.052480
2^{-8}	3.0298	0.062341	3.0351	0.024700
2^{-9}	3.0278	0.039143	3.0300	0.011073

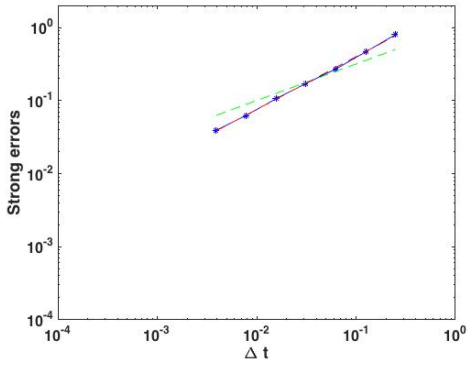
Table 6.7: Weak errors in the approximation of the function V, for $T = 1$, $r = 0.01 + 0.03t + 0.03\sin(60t)$, $\sigma = 0.1 + 0.3t + 0.03\sin(30t)$, $S_0 = 80$ $K = 100$, $M = 500\ 000$

Time step	Euler Maruyama	Weak error	Milstein	Weak error
2^{-3}	2.6417	0.445632	2.7320	0.355596
2^{-4}	2.8055	0.281786	2.8480	0.239561
2^{-5}	2.9468	0.140531	2.9674	0.120144
2^{-6}	3.0179	0.069379	3.0283	0.059306
2^{-7}	3.0604	0.026856	3.0661	0.021514
2^{-8}	3.0968	0.009478	3.0998	0.012216
2^{-9}	3.0922	0.004896	3.0933	0.005684

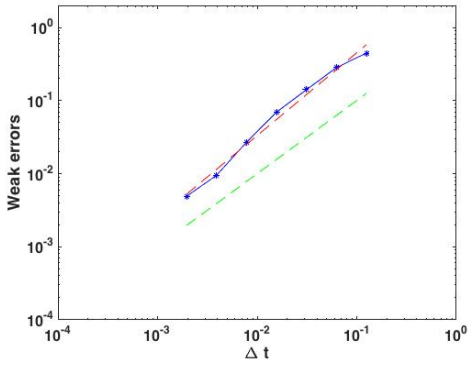
Table 6.6 shows an experimental strong order of convergence above the theoretical value $\frac{1}{2}$ for the Euler-Maruyama scheme. On the other hand, the experimental weak order of convergence for the Milstein method is in good accordance with the theoretical value. From table 6.7 we see that both schemes also

converge in the weak sense with good accordance with the theoretically predicted values.

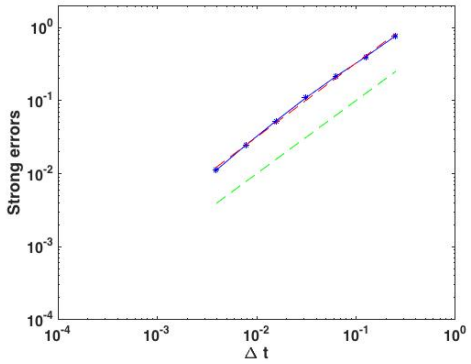
The strong and weak errors in tables 6.6 and 6.7, respectively, are plotted in figure 6.2 on a loglog scale along with a linear regression estimated using a least squares fit, which is represented by the red dashed line.



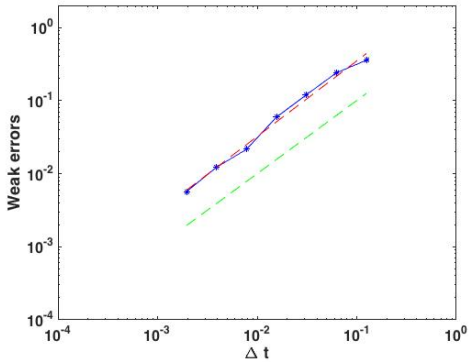
(a) Strong convergence - Euler-Maruyama ($q = 0.7213$)



(b) Weak convergence - Euler-Maruyama ($q = 1.1321$)



(c) Strong convergence - Milstein ($q = 1.0131$)



(d) Weak convergence - Milstein ($q = 1.0346$)

Figure 6.2: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on tables 6.6 and 6.7. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

Figure 6.2 further confirms that the experimental strong order of convergence of the Euler-Maruyama scheme is above the value predicted by theory and that the strong order of convergence for the Mislstein scheme is as predicted by theory. Indeed, the slope of the estimated linear regression for the strong convergence is $q = 0.7213$ for the Euler-Maruyama scheme and $q = 01.0131$ for the Milstein scheme. Further, the slope of the estimated linear regression, which gives the experimental weak order of convergence, is $q = 1.1321$ for the Euler-Maruyama scheme and $q = 1.0346$ for the Milstein scheme. This means that, for both schemes, the experimental orders of convergence are in good accordance with the theory.

We now assume that both r and σ have a jump discontinuity at $t = 0.3$.

$$r(t) = \begin{cases} 0.04 & t < 0.3 \\ 0.03 & t \geq 0.3 \end{cases} \tag{6.5}$$

$$\sigma(t) = \begin{cases} 0.2 & t < 0.3 \\ 0.3 & t \geq 0.3 \end{cases} \tag{6.6}$$

Tables 6.8 and 6.9 display the strong and weak errors, respectively, in the approximation of the price $V(S_0, t_0)$ of a call option at time $t_0 = 0$, with initial stock value $S_0 = 80$, strike price $K = 100$ and maturity $T = 1$. The approximations were computed for seven different time steps using a Monte Carlo simulation with 5000 sample paths in table 6.6 and 500 000 sample paths in table 6.7 just like in the previous models. The price of the underlying asset is also approximated using explicit Euler-Maruyama and Milstein schemes, that is, using discretisations (5.48) and (5.49), which we discussed in section 5.6.2. The strong and weak errors are computed as in (6.3) and (6.4), respectively.

The reference values for the true solution are 2.7032 and 2.7036 for the strong convergence and 3.4279 and 3.4281 for the weak convergence, using the Euler-Maruyama and Milstein schemes, respectively.

Table 6.8: Strong errors in the approximation of the function V, for $T = 1$, $S_0 = 80$ $K = 100$, $M = 5000$

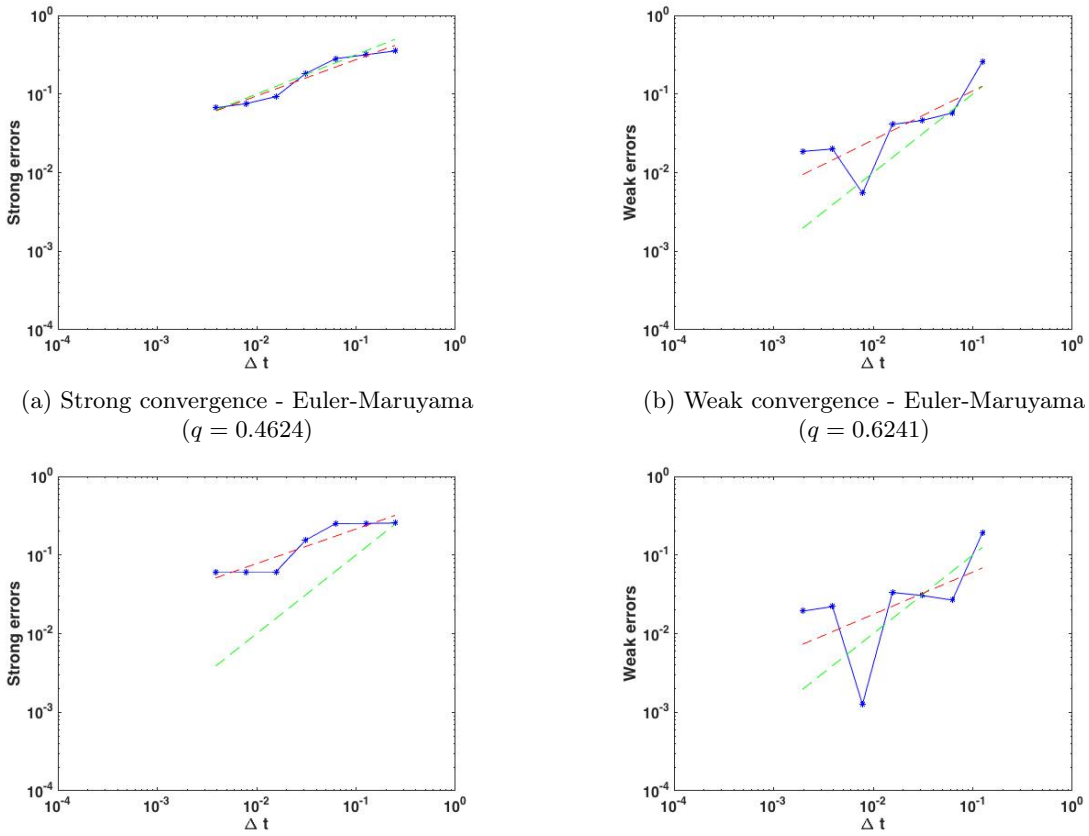
Time step	Euler Maruyama	Strong error	Milstein	Strong error
2^{-3}	2.5401	0.354397	2.6222	0.255246
2^{-4}	2.5927	0.315582	2.6361	0.252661
2^{-5}	2.6278	0.281531	2.6431	0.251627
2^{-6}	2.6744	0.183233	2.6849	0.155290
2^{-7}	2.6942	0.092611	2.7010	0.061147
2^{-8}	2.6992	0.075368	2.7019	0.061106
2^{-9}	2.7004	0.066978	2.7024	0.061096

Table 6.9: Weak errors in the approximation of the function V, for $T = 1$, $S_0 = 80$ $K = 100$, $M = 500\ 000$

Time step	Euler Maruyama	Weak error	Milstein	Weak error
2^{-3}	3.1713	0.256658	3.2370	0.191150
2^{-4}	3.3705	0.057423	3.4014	0.026699
2^{-5}	3.3818	0.046191	3.3976	0.030497
2^{-6}	3.3869	0.041042	3.3948	0.033365
2^{-7}	3.4225	0.005486	3.4269	0.001259
2^{-8}	3.4479	0.019969	3.4502	0.022103
2^{-9}	3.4466	0.018628	3.4475	0.019328

As discussed in section 3.2.1, with the interest rate and the volatility as discontinuous functions, there is no guarantee that the discretisation methods will converge. In fact, by looking at the results of tables 6.8 and 6.9, even when we consider smaller time steps, the methods do not seem to converge in the weak sense nor in the strong sense, except for the weak convergence in the Euler-Maruyama method. The strong and weak errors are plotted on Figure 6.3, as well as the usual estimated linear regression, which further supports this conclusion. Besides the results shown in this work, we also tested the convergence for different initial states of the random number generator to be sure that the methods did not fail to converge only for a specific random sequence.

Figure 6.2 shows the plotted strong and weak errors in tables 6.6 and 6.7, respectively, on a loglog scale along with a linear regression estimated using a least squares fit, which is represented by the red dashed line.



(a) Strong convergence - Euler-Maruyama ($q = 0.4624$)

(b) Weak convergence - Euler-Maruyama ($q = 0.6241$)

(c) Strong convergence - Milstein ($q = 0.5379$)

(d) Weak convergence - Milstein ($q = 0.4402$)

Figure 6.3: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on tables 6.6 and 6.7. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

As expected, the experimental weak order of convergence is $q = 0.6241$ for the Euler-Maruyama method and $q = 0.4402$ for the Milstein method, which means that in both schemes the strong orders of convergence are far below the theoretically predicted values. Moreover, the experimental strong order of convergence is $q = 0.4624$ for the Euler-Maruyama scheme and $q = 0.5379$ for the Milstein scheme. Although the experimental weak order of convergence of the Euler-Maruyama scheme does not seem too far from the value predicted by theory the result is not significant. The fact that neither of the methods converges in weak or strong sense is not unexpected, as the convergence of the schemes is not guaranteed for non-Lipschitz coefficients.

6.2.4 Time and asset price dependent parameters

We now look at a model where the expected rate on return and the volatility depend both on time and on the price of the underlying asset.

Consider the following function for the interest rate

$$r(t, S_t) = \frac{0.05}{1+t} + \frac{0.05}{1+S_t} \tag{6.7}$$

$$\sigma(t, S_t) = \frac{0.2}{1+t} + \frac{0.2}{1+S_t} \tag{6.8}$$

The choice for the volatility and interest rate functions comes from the fact that both the volatility

and the interest rate are usually negatively correlated with the price of the undelying asset.
We will start by implementing the model using explicit schemes.

Tables 6.10 and 6.11 show the strong and weak errors, respectively, in the approximation of the price $V(S_0, t_0)$ of a call option at time $t_0 = 0$, with initial stock value $S_0 = 80$, strike price $K = 100$ and maturity $T = 1$. The approximations were computed for seven different time steps using a Monte Carlo simulation with 5000 sample paths in table 6.10 and 50 000 sample paths in table 6.11. Recall that for models in which the coefficients also depend on the price of the underlying asset, we need to compute one sample path at time. For this reason it becomes too computationally costly to sample through a large number of samples paths, hence here we only sample through 50 000 paths when we study the weak convergence. To approximate the price of the stock at each time step we used the discretisations (5.52) and (5.53) and the strong and weak errors are computed as in (6.3) and (6.4), respectively.

The reference values for the true solution are 0.5812 and 0.5814 for the strong convergence and 0.5533 and 0.5534 for the weak convergence, using the Euler-Maruyama and Milstein schemes, respectively.

Table 6.10: Strong errors in the approximation of the function V, for $T = 1$, $r(t, S_t) = \frac{0.05}{(1+t)} + \frac{0.05}{(1+S_t)}$ $\sigma(t, S_t) = \frac{0.2}{(1+t)} + \frac{0.2}{(1+S_t)}$, $S_0 = 80$ $K = 100$, $M = 5000$

Time step	Euler Maruyama	Strong error	Milstein	Strong error
2^{-3}	0.4461	0.135291	0.4718	0.109656
2^{-4}	0.5120	0.069598	0.5258	0.055687
2^{-5}	0.5473	0.035530	0.5536	0.027803
2^{-6}	0.5642	0.019697	0.5673	0.014140
2^{-7}	0.5729	0.011897	0.5746	0.006803
2^{-8}	0.5773	0.007135	0.5783	0.003175
2^{-9}	0.5796	0.004276	0.5801	0.001373

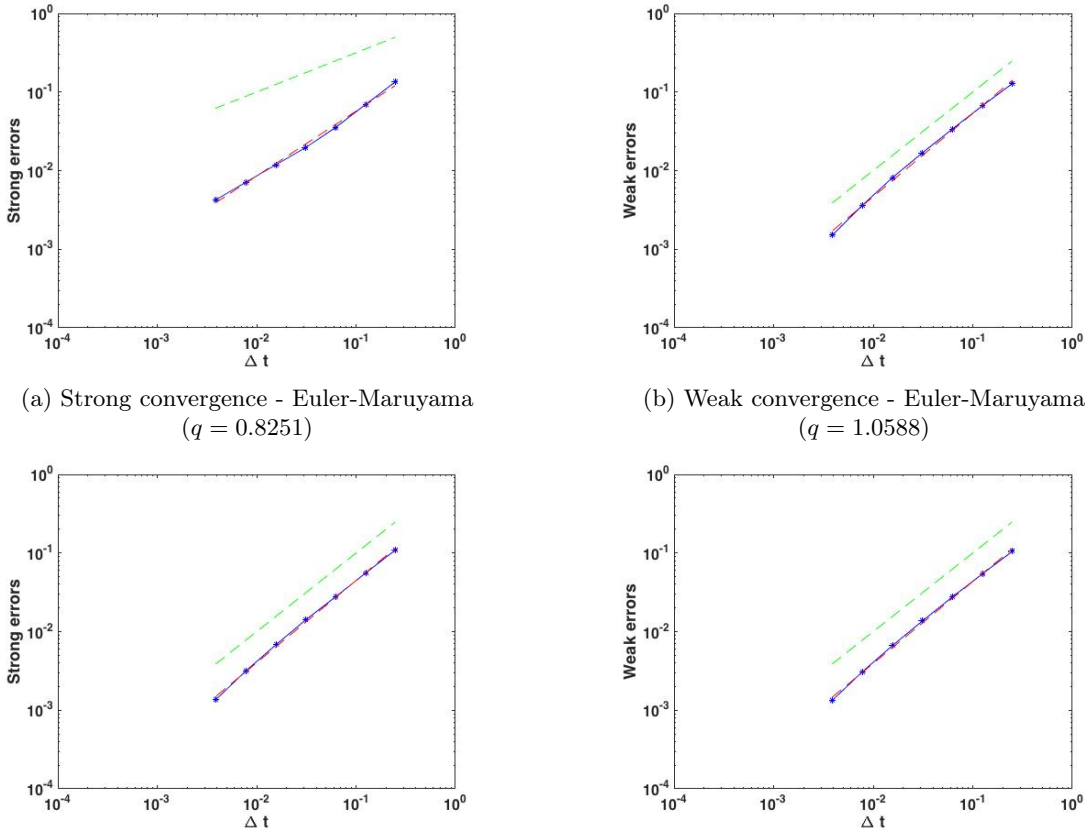
Table 6.11: Weak errors in the approximation of the function V, for $T = 1$, $r(t, S_t) = \frac{0.05}{(1+t)} + \frac{0.05}{(1+S_t)}$ $\sigma(t, S_t) = \frac{0.2}{(1+t)} + \frac{0.2}{(1+S_t)}$, $S_0 = 80$ $K = 100$, $M = 50\ 000$

Time step	Euler Maruyama	Weak error	Milstein	Weak error
2^{-3}	0.4250	0.128244	0.4482	0.105190
2^{-4}	0.4865	0.066814	0.4991	0.054281
2^{-5}	0.5194	0.033924	0.5257	0.027656
2^{-6}	0.5364	0.016864	0.5397	0.013664
2^{-7}	0.5451	0.008144	0.5467	0.006650
2^{-8}	0.5497	0.003639	0.5503	0.003095
2^{-9}	0.5518	0.001522	0.5520	0.001341

From the analysis of the weak errors in table 6.11 we see that if we cut the time step by a factor of 2 the weak error reduces to half, so both numerical approximations have an experimental weak order of convergence of 1. The same happens for the strong error in the Milstein scheme. Furthermore, in table 6.10, if we cut the time step by a factor of 4, the strong error decreases to more than half, therefore the experimental strong order of convergence of the Euler-Maruyama method is above the theoretical value

of $\frac{1}{2}$. This means that both experimental strong and weak orders of convergence agree with or are above the theoretically predicted values.

Figure 6.4 displays the strong and weak errors on tables 6.10 and 6.11, respectively, and a linear regression on a loglog scale, estimated by a least squares fit. As the estimated regressions seem to have the same slope as the reference lines, it further supports the above observation.



(a) Strong convergence - Euler-Maruyama ($q = 0.8251$)

(b) Weak convergence - Euler-Maruyama ($q = 1.0588$)

(c) Strong convergence - Milstein ($q = 1.0448$)

(d) Weak convergence - Milstein ($q = 1.0430$)

Figure 6.4: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on tables 6.10 and 6.11. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

In fact, the estimated regression for the strong convergence in the Euler-Maruyama method has a slope of $q = 0.8251$, which is above the theoretically predicted value and for the Milstein scheme it is $q = 1.0448$, which is in good accordance with the theoretical value. Moreover, the slope of the estimated regression for the weak convergence in the Euler-Maruyama scheme is $q = 1.0588$ and 1.0448 for the Milstein scheme, which means that the experimental weak order of convergence of both schemes agrees with the theoretical order of convergence.

Table 6.12 displays the strong and weak errors in the approximation of the value $V(S_0, t_0)$ of a call option at time $t_0 = 0$, with initial stock price $S_0 = 80$, strike price $K = 100$ and maturity $T = 1$. The approximations were computed for seven different time steps using a Monte Carlo simulation with 500 sample paths. The reason why we only sampled through 500 sample paths is that implicit models have to solve an additional equation at each time step, which is computationally costlier. To approximate the price of the stock at each time step we used the predictor-corrector method described in section 5.6.3, with the explicit Euler-Maruyama method (5.58) and the explicit Milstein method (5.60) as the predictors and

the generalisation of the deterministic trapezoidal method (5.59) and (5.61) as the corrector methods. The strong and weak errors were computed as in (6.3) and (6.4), respectively.

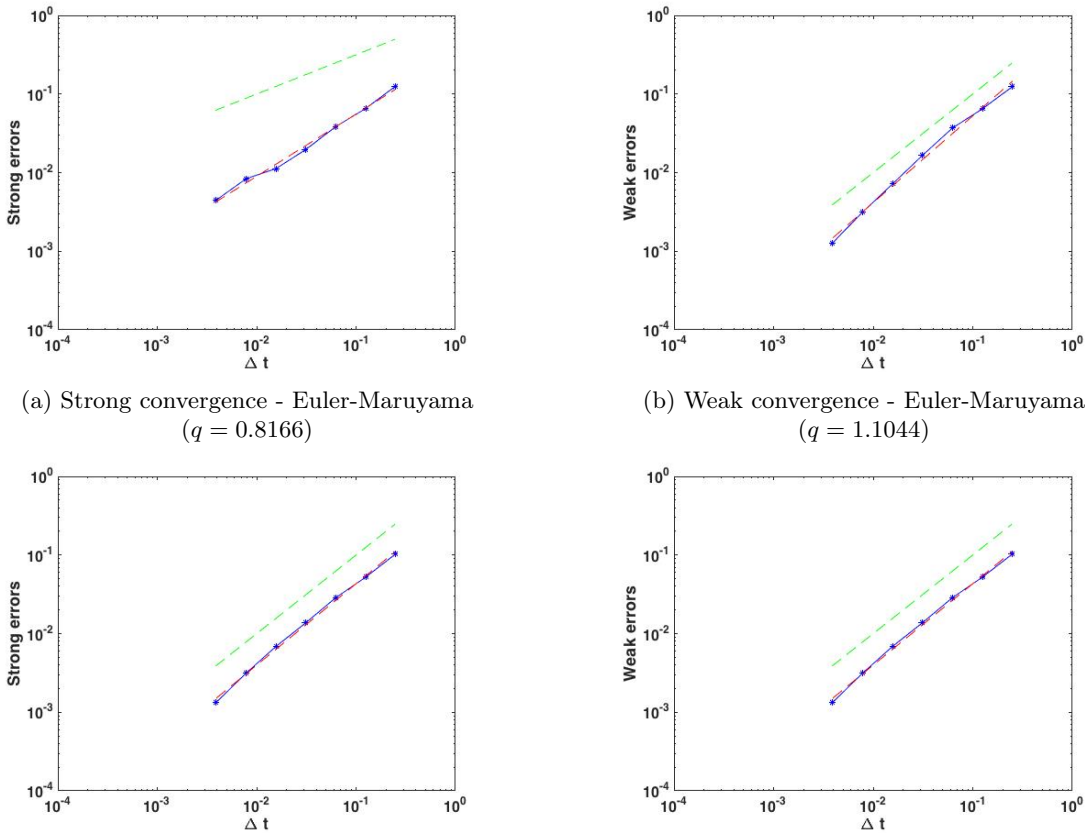
The reference values for the true solution are 0.6126 using the Euler-Maruyama scheme and 0.6125 using the Milstein scheme. Note that since we used the same number of sample paths for the weak and strong convergence, the reference values for the exact solution are the same for both types of convergence.

Table 6.12: Strong and weak errors in the approximation of the function V, for $T = 1$, $r(t, S_t) = \frac{0.05}{(1+t)} + \frac{0.05}{(1+S_t)}$ $\sigma(t, S_t) = \frac{0.2}{(1+t)} + \frac{0.2}{(1+S_t)}$, $S_0 = 80$ $K = 100$, $M = 500$

Time step	Euler Maruyama	Strong error	Weak error	Milstein	Strong error	Weak error
2^{-3}	0.4885	0.126005	0.124054	0.5096	0.103027	0.102956
2^{-4}	0.5477	0.065455	0.064821	0.5602	0.052290	0.052290
2^{-5}	0.5757	0.038474	0.036897	0.5839	0.028630	0.028630
2^{-6}	0.5960	0.019681	0.016564	0.5988	0.013755	0.013755
2^{-7}	0.6054	0.011222	0.007166	0.6056	0.006864	0.006864
2^{-8}	0.6094	0.008346	0.003130	0.6094	0.003130	0.003130
2^{-9}	0.6113	0.004456	0.001274	0.6112	0.001330	0.001330

The weak errors in table 6.12 approximately reduce to half when we cut the time step by a factor of two, which suggests that the weak order of convergence of the numerical methods should be close to one. The strong errors in the Euler-Maruyama predictor-corrector scheme seem to decrease to more than half when we cut the time step by a factor of 4, which means that the experimental strong order of convergence should be above the theoretical value of $\frac{1}{2}$. For the Milstein predictor-corrector method the strong order of convergence seems to correspond to the theoretical value of 1.

Figure 6.5 plots the strong and weak errors on table 6.12 on a loglog scale and a linear regression estimated using a least squares fit.



(a) Strong convergence - Euler-Maruyama ($q = 0.8166$)

(b) Weak convergence - Euler-Maruyama ($q = 1.1044$)

(c) Strong convergence - Misltein ($q = 1.0361$)

(d) Weak convergence - Milstein ($q = 1.0360$)

Figure 6.5: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on table 6.12. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

As suggested by table 6.12, the Euler-Maruyama predictor-corrector method has an experimental strong order of convergence of $q = 0.8166$ and an experimental weak order of convergence of $q = 1.1044$ while the Milstein predictor-corrector method has an experimental strong order of convergence $q = 1.0361$ and an experimental weak order of convergence of $q = 1.0360$. So, we can conclude that the orders of convergence of both numerical schemes are in good accordance with the theory.

6.2.5 Heston Model

In this section we implement the Heston model using algorithm 6 with the discretisations (5.62) and (5.63) for the stock price and the volatility, respectively.

One of the difficulties of the Heston model is how to choose the parameters, as they influence the shape of the volatility smile and can induce skewness in the distribution of the stock returns. Furthermore, the prices approximated by the model are quite parameter sensitive, so small changes in the parameters values lead to considerably different results. For this reason, as we will consider two scenarios, one when the Feller condition is fulfilled and another when it is not, we will only change θ and k to keep the two scenarios as comparable as possible. We start by considering a scenario in which the parameters satisfy the Feller condition.

Table 6.14 shows the value $V(S_0, t_0)$ of a call option at time $t_0 = 0$, with initial stock price $S_0 = 100$, strike price $K = 100$ and maturity $T = 1$, computed for seven different time steps using a Monte Carlo

simulation with 500 000 sample paths, like we discussed in section 5.6.4. To approximate the price of the stock at each time step we used the discretisations (5.62) and (5.63), which are the result of the application of explicit Euler-Maruyama (5.3) and Milstein (5.21) methods to equations (3.64) and (3.65). As in the previous models, the first and third columns show the price of the call computed using each of these methods. The strong and weak errors were computed as before, that is, as in (6.3) and (6.4), respectively.

The reference values for the true solution are 1.0512 and 0.1019 for the strong convergence and 1.0526 and 1.0705 for the weak convergence, using the Euler-Maruyama and Milstein schemes, respectively.

Recall that as strong convergence is more computationally costly we only sampled through 5000 sample paths, while the weak error was computed over 500 000 sample paths. With our choice of parameters, the Feller condition is 1.2245, so it is fulfilled.

Table 6.13: Strong errors in the approximation of the function V, for $T = 1$, $r = 0.0015$, $\nu_0 = 0.2$, $\theta = 0.2$, $\xi = 1.4$, $k = 6$, $\rho = -0.7$, $S_0 = 100$, $K = 100$, $M = 5000$

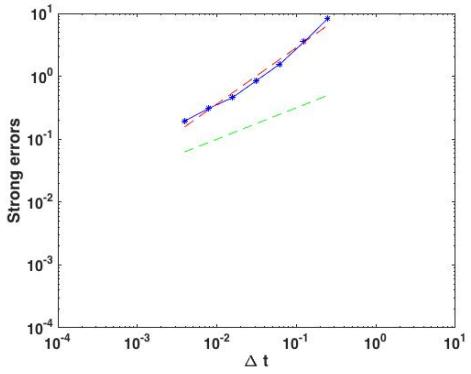
Time step	Euler Maruyama	Strong error	Milstein	Strong error
2^{-3}	9.3276	8.334688	6.6323	6.53033
2^{-4}	4.4870	3.535108	2.9038	2.815030
2^{-5}	2.4350	1.577938	1.2346	1.142377
2^{-6}	1.5922	0.848310	0.5842	0.482255
2^{-7}	1.2390	0.456420	0.3036	0.202031
2^{-8}	1.1111	0.305791	0.1791	0.089278
2^{-9}	1.0657	0.193762	0.1435	0.046574

Table 6.14: Weak errors in the approximation of the function V, for $T = 1$, $r = 0.0015$, $\nu_0 = 0.2$, $\theta = 0.2$, $\xi = 1.4$, $k = 6$, $\rho = -0.7$, $S_0 = 100$, $K = 100$, $M = 500\ 000$

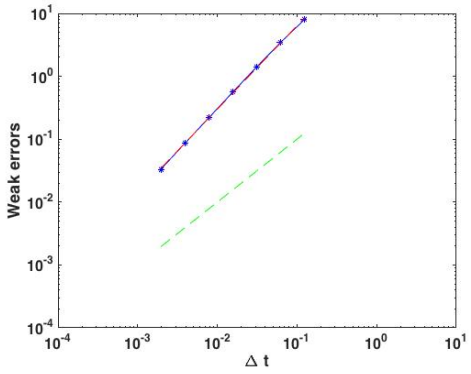
Time step	Euler Maruyama	Weak error	Milstein	Strong error
2^{-3}	9.1101	8.057556	5.7480	4.677503
2^{-4}	4.5196	3.467014	2.7992	1.728769
2^{-5}	2.4691	1.416493	1.8662	0.795762
2^{-6}	1.6123	0.559762	1.4836	0.413119
2^{-7}	1.2716	0.219021	1.2866	0.216124
2^{-8}	1.1382	0.085663	1.1826	0.112093
2^{-9}	1.0857	0.033121	1.1253	0.054868

Although there is no guarantee that the discretisation schemes will converge, as discussed in section 5.6.4, both schemes converge in strong and weak sense. In fact, the Euler-Maruyama strong and weak orders of convergence are above the value predicted by theory. On the other hand, the Milstein strong and weak order of convergence seems to be in good accordance with the theoretical value. In the previous models we have already seen that in what concerns weak convergence, the Euler-Maruyama and Milstein schemes have similar performances, although Milstein usually produces slightly better results. For the Heston model, not only is this true, but it is also worth noticing that Milstein scheme seems to generate significantly better results than the Euler-Maruyama for bigger time steps.

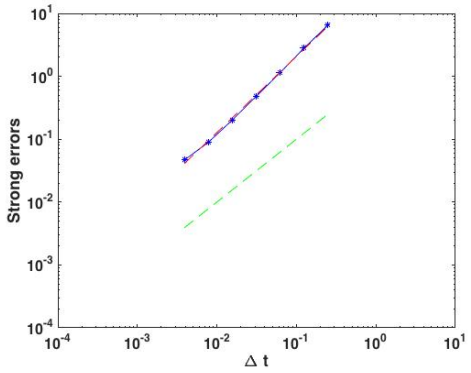
Figure 6.6 shows the plotted strong and weak errors of tables 6.13 and 6.14, respectively on a loglog scale, as well as the estimated linear regression using a least squares fit.



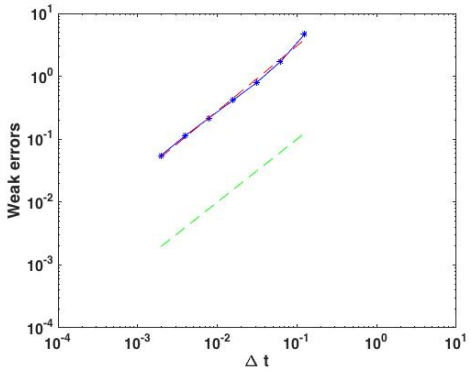
(a) Strong convergence - Euler-Maruyama ($q = 0.8976$)



(b) Weak convergence - Euler-Maruyama ($q = 1.3267$)



(c) Strong convergence - Milstein ($q = 1.2090$)



(d) Weak convergence - Milstein ($q = 1.0363$)

Figure 6.6: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on tables 6.13 and 6.14. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

The estimated linear regression for the strong convergence rate of Euler-Maruyama has a slope of $q = 0.8976$ and the estimated linear regression for the weak convergence rate is $q = 1.3267$, which confirms that this method converges with a higher order of convergence than as predicted by theory. On the other hand, the Milstein scheme has an experimental strong order of convergence $q = 1.2090$ and an experimental weak order of convergence $q = 1.0363$, which are in good accordance with theory. In fact, a closer analysis of table 6.14 shows that the Euler-Maruyama scheme performs poorly for bigger time steps than the Milstein scheme, but has greater accuracy when the discretisation is more refined.

As discussed in 5.6.4, in the real world the Feller condition is rarely met, so now we test the convergence of the numerical schemes for the Heston model when the Feller condition is not fulfilled. In order to have the Feller condition not met we only change the values of k, and θ. With the new choice of parameters the Feller condition does not hold with the value 0.0612. The results on tables 6.15 and 6.16 were computed in the same way as the ones on the previous table, with the same parameters expect for the three parameters mentioned above.

The reference values for the true solution are 0.0623 and 0.0543 for the strong convergence and 0.0775 and 0.0682 for the weak convergence, using the Euler-Maruyama and Milstein schemes, respectively.

Table 6.15: Strong errors in the approximation of the function V, for $T = 1$, $r = 0.0015$, $\nu_0 = 0.2$, $\theta = 0.02$, $\xi = 1.4$, $k = 3$, $\rho = -0.7$, $S_0 = 100$, $K = 100$, $M = 5000$

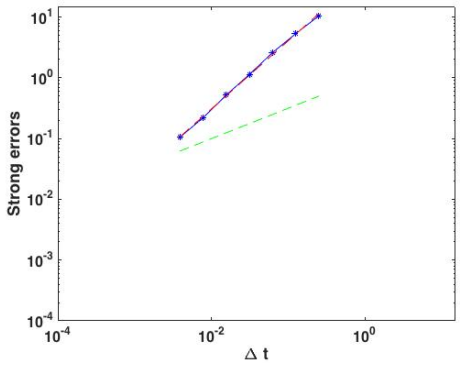
Time step	Euler Maruyama	Strong error	Milstein	Strong error
2^{-3}	10.5482	10.496545	11.6831	11.640063
2^{-4}	5.3910	5.342973	6.8751	6.834755
2^{-5}	2.6581	2.613339	3.8451	3.804123
2^{-6}	1.1919	1.144334	2.0245	1.984052
2^{-7}	0.5735	0.523174	1.0146	0.970061
2^{-8}	0.2781	0.221038	0.5214	0.480544
2^{-9}	0.1601	0.106036	0.2511	0.208773

Table 6.16: Weak errors in the approximation of the function V, for $T = 1$, $r = 0.0015$, $\nu_0 = 0.2$, $\theta = 0.02$, $\xi = 1.4$, $k = 3$, $\rho = -0.7$, $S_0 = 100$, $K = 100$, $M = 500\ 000$

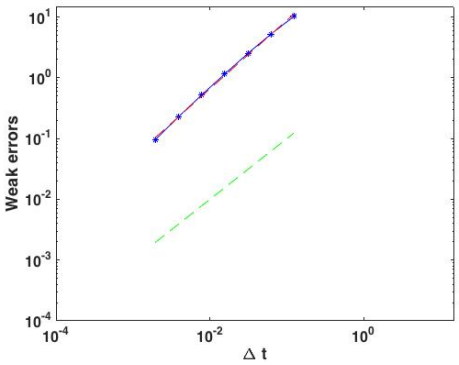
Time step	Euler Maruyama	Weak error	Milstein	Weak error
2^{-3}	10.4822	10.404665	7.3695	7.301245
2^{-4}	5.3590	5.281528	3.1849	3.116695
2^{-5}	2.6186	2.541109	1.4282	1.359932
2^{-6}	1.2402	1.162676	0.6740	0.605755
2^{-7}	0.5997	0.522204	0.3432	0.274940
2^{-8}	0.3059	0.228430	0.1949	0.126657
2^{-9}	0.1732	0.095724	0.1251	0.056838

From tables 6.15 and 6.16 it is clear that both methods converge in weak and strong sense with an experimental order of 1.

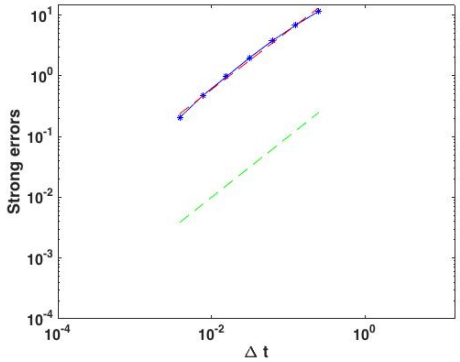
Figure 6.7 shows the plotted errors of tables 6.15 and 6.16 and an estimated linear regression on a loglog scale using a least squares fit.



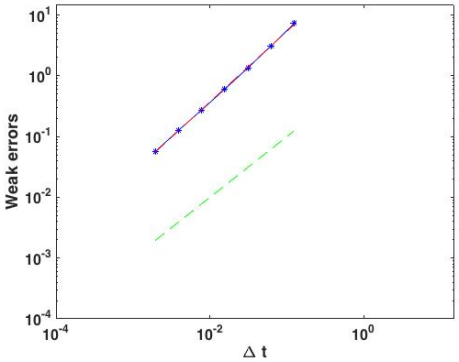
(a) Strong convergence - Euler-Maruyama ($q = 1.1214$)



(b) Weak convergence - Euler-Maruyama ($q = 1.1299$)



(c) Strong convergence - Milstein ($q = 0.9655$)



(d) Weak convergence - Milstein ($q = 1.1630$)

Figure 6.7: Strong and weak convergence for the errors of Euler-Maruyama and Milstein methods on tables 6.15 and 6.16. The green dashed lines represent the reference slope and the red dashed lines represent the least squares regression.

The estimated linear regression for the strong convergence rate of Euler-Maruyama has a slope of $q = 1.1214$ and the estimated linear regression for the weak convergence rate is $q = 1.1299$. Comparing with results when the Feller condition was met, now the strong order of convergence is even higher and the weak order of convergence is slightly lower. The Milstein method has an experimental strong order of convergence $q = 0.9655$ and weak order $q = 1.1630$, which are both higher than the results when the Feller condition was met. So, overall the methods converge with the order theoretically predicted, except for the Euler-Maruyama that has a much higher weak order of convergence. This is not surprising though, because when the Feller condition is not met, the variance can take negative values and we fix that by reflecting that value, which induces positive bias in the pricing of European options. Since the Euler-Maruyama discretisation usually produces more negative values of variance than the Milstein scheme [29], this has more influence on the performance of the Euler discretisation. Also in this case it is clear that the Milstein method is considerably more accurate than the Euler-Maruyama for larger time steps.

Chapter 7

Conclusions

The objective of this thesis was to approximate the price of European vanilla options by finite difference schemes for PDE's and by numerical methods for SDE's in Monte Carlo simulations. In the first approach, we transformed the Black-Scholes partial differential equation into the heat equation and used the Crank-Nicolson scheme and the forward and backward Euler schemes. In the second approach we studied several extensions to the Black-Scholes asset model and used the Euler-Maruyama and Milstein schemes to approximate the price of the underlying asset.

The approximations via finite difference schemes for PDE's were used to approximate the classical Black-Scholes equation, where the volatility and interest rate were constant parameters. The obtained results have good accuracy, with relative errors in the order of 10^{-3}. Moreover, the computation time was less than one second, even for the implicit schemes. However, the use of constant volatility and interest rate is not realistic and the relaxation of these assumptions is better handled by numerical methods for SDE's in Monte Carlo simulations than by finite difference schemes for PDE's.

We started our Monte Carlo simulations with the classical Black-Scholes asset model, with constant coefficients. When compared to the finite difference schemes the accuracy of the methods were very similar, although slightly better in the Monte Carlo simulations. This is not surprising, however, because we used half a million sample paths in our simulations. On the other hand the Monte Carlo simulations are very inefficient when compared to finite difference schemes, as it took almost one minute to run the algorithm and the results are not significantly better.

For the models without constant coefficients we did not have an explicit solution so we compared the orders of convergence using an approximation of the exact solution with a small time step. The model with volatility and interest rate as sinusoidal functions of time had a good accuracy when using the Euler-Maruyama and the Milstein methods, as the weak and strong orders of convergence were as predicted by the theory or above the theoretical values. However, when we tested the model with coefficients with jump discontinuities, both methods failed to converge as the coefficients were not Lipschitz continuous.

To test time and asset price dependency we used first explicit schemes and then predictor-corrector schemes. Both approaches had strong and weak orders of convergence in good accordance with the theoretically predicted values. However, it is worth emphasising that for the predictor-corrector schemes we only sampled through 500 sample paths, which is low number of sample paths for a Monte Carlo simulation. Nonetheless, the results were as good as those of the explicit schemes, which were sampled through 50 000 paths for the weak convergence and 5000 for the strong convergence.

Finally we tested the stochastic volatility model, the Heston model, using two different scenarios: one where the Feller condition was met and another when it was not met. In both scenarios we had good convergence results, but we actually obtained better results when the Feller condition was not met, which might be caused by the positive bias induced when reflecting negative variance values. Nonetheless, in

both scenarios, the convergence orders of the Euler-Maruyama scheme were considerably above the theoretically predicted values. However, the Heston model is rather sensitive to changes in the parameters and requires calibration to obtain more realistic results.

Overall, it is worth noticing that in all of the asset models the Euler-Maruyama and the Milstein methods produced similar results in what concerns weak convergence. For this reason it is natural to choose the Euler-Maruyama scheme for smaller time steps, as it simpler to implement and less burdensome and the Milstein scheme could be used for larger time steps. When it comes to strong convergence, the Milstein scheme is usually the first choice.

An important issue with the models that we presented in this work is that the computational cost increases both with the number of time steps and the number of sample paths used. And as the convergence rate of the Monte Carlo method is $O(1/\sqrt{M})$, it requires a large number of trials to produce accurate results, which makes this method very inefficient. This might not represent a problem in the pricing of path-independent options, such as the vanilla European options studied in this work, and when only weak convergence of the numerical approximations is required, but we have seen that in order to study strong convergence, sampling through more than 5000 random paths becomes too computationally demanding. Duffie and Glynn [11] propose an optimal allocation between the number of time steps and the number of Monte Carlo trials, where the number of time steps should be chosen proportional to the square-root of the number of sample paths for first order methods. With this allocation the convergence rate of the error, using the Euler-Maruyama discretisation is $O(M^{-\frac{1}{3}})$. However, the proportional constant is difficult to estimate. Another way to overcome the low convergence rate of Monte Carlo methods that has been gaining importance in the financial industry is to use variance control techniques, such as the use Quasi-Monte Carlo methods, which have a convergence rate of $O(1/M)$, see [10], but it also has several drawbacks.

Although the results obtained in the simulation of the extensions to the Black-Scholes model were good and the functions used for the interest rate and volatility are realistic, in order to have more accurate results the models should be calibrated to the market values. Moreover, it would also be interesting to develop new SDE's theory to account for non-Lipschitz drift and diffusion coefficients, as this would allow more flexibility in the choice of the functions. Research has been carried out in this sense, such as [2], which considers that the drift and diffusion coefficients of the SDE can have discontinuities but are still regular enough to guarantee the existence of a unique weak solution and uses an Exit method to deal with the possibility of a path hitting a point of discontinuity. Hutzenthaler, Jentzen and Kloeden [20] propose a modified explicit Euler method (tamed Euler) to deal with SDE's with non-globally Lipschitz continuous coefficients but assume that the drift coefficient is globally one-sided Lipschitz continuous and has an at most polynomially growing derivative and that the diffusion coefficient is Lipschitz continuous. Both these works rely on several assumptions on the coefficients, as do most of the works carried out so far, which leaves this subject open to improvement.

Appendix A

Lemma A.0.1. (Itô's lemma. Bivariate case). Let $X(t)$ and $Y(t)$ be two stochastic processes that satisfy $dX_t = \mu_x(t, X_t)dt + \sigma_x(t, X_t)dW_t^x$ and $dY_t = \mu_y(t, X_t)dt + \sigma_y(t, X_t)dW_t^y$, where W^x and W^y are two Wiener processes with correlation ρ . Further, let $f \in C^2$, then $A(t) = f(X(t), Y(t))$ is also a stochastic process and is given by

$$\begin{aligned} df(t, X_t, Y_t) = &\left[\frac{\partial f}{\partial t} + \mu_x(X_t, Y_t)\frac{\partial f}{\partial X} + \mu_y(X_t, Y_t)\frac{\partial f}{\partial Y} + \frac{1}{2}\sigma_x^2(X_t, Y_t)\frac{\partial^2 f}{\partial X^2} + \frac{1}{2}\sigma_y^2(X_t, Y_t)\frac{\partial^2 f}{\partial Y^2}\right] dt \\ &+ \left[\rho\sigma_x(X_t, Y_t)\sigma_y(X_t, Y_t)\frac{\partial^2 f}{\partial X \partial Y}\right] dt + \sigma_x(X_t, Y_t)\frac{\partial f}{\partial X}dW_t^x + \sigma_y(X_t, Y_t)\frac{\partial f}{\partial Y}dW_t^y \end{aligned} \tag{A.1}$$

where all the partial derivatives of f are evaluated at (t, X_t, Y_t).

Definition A.0.1. (Stratonovich integral). Let X be a stochastic process locally square-integrable, i.e. $\int_0^T E[X_t^2]dt < \infty$ and adapted to the filtration generated by the Brownian motion W_t in $[0, T]$. If $0 = t_0 < ... < t_n = T$ is a partition of the interval $[0, T]$, the Stratonovich integral can be defined as the unique mean square limit of the Riemann-Stieltjes sums, i.e.:

$$\int_0^T X_t \circ dW_t = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{X_{t_i} + X_{t_{i-1}}}{2}(W_{t_i} - W_{t_{i-1}})$$

Theorem A.0.2. (Girsanov's theorem). The following statements hold:

- The stochastic process
$$M_t = e^{-qW_t - \frac{1}{2}q^2 t}, \quad t \in [0, T]$$
(where q is a constant) is a martingale with respect to the natural Brownian filtration $\mathcal{F}_t = \sigma(W_s, s \leq t), \quad t \in [0, T]$, under the probability measure $\mathbb{P}$.

- The relation
$$\mathbb{Q}(A) = \int_A M_t(\omega)d\mathbb{P}(\omega), \quad A \in \mathcal{F}$$
defines a probability measure $\mathbb{Q}$ on $\mathcal{F}$, which is equivalent to $\mathbb{P}$.

- Under the probability measure $\mathbb{Q}$, the process $\tilde{W}$ defined by $\tilde{W}_t = W_t + qt$, $t \in [0, T]$, is a standard Brownian motion.

- The process $\tilde{W}$ is adapted to the filtration $\mathcal{F}_t$

The probability measure $\mathbb{Q}$ is called an equivalent martingale measure.

Bibliography

[1] Alves, C.J.S. (2008). *Numerical Analysis of Partial Differential Equations: An Introduction*. (In Portuguese)

[2] Arnold, S.K. (2007). *Approximation Schemes for SDEs with Discontinuous Coefficients*. PhD thesis. Swiss Federal Institute of Technology Zurich.

[3] Ash, R. (1970). *Basic Probability Theory*. Dover Publications.

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

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

[6] Boyle, M. (1977). Options: A Monte Carlo Approach. *Journal of Financial Economics*. Vol. 4, No 3, pp. 323-338.

[7] Cen, Z. and Le, A. (2011). A Robust and Accurate Finite Difference Method for a Generalized Black–Scholes Equation. *Journal of Computational and Applied Mathematics*. Vol. 235, No 13, pp. 3728–3733.

[8] Courtadon, G. (1982). A More Accurate Finite Difference Approximation for the Valuation of Options. *Journal of Financial and Quantitative Analysis*. Vol. 17, No 5, pp. 697-703.

[9] Duffy, D.J. (2006). *Finite Difference Methods in Financial Engineering*. John Wiley & Sons Ltd.

[10] Duffy, D.J. and Kienitz, J. (2009). *Monte Carlo Frameworks: Building Customisable High-performance C++ Applications*. Wiley.

[11] Duffie, D. and Glynn, P. (1995). Efficient Monte Carlo Simulations of Security Prices. *The Annals of Applied Probability*. Vol. 5, No 4, pp. 897-905.

[12] Fernandes, M.M. (2009). *Finite Differences Schemes for Pricing of European and American Options*. MA thesis. Instituto Superior Técnico. (in Portuguese)

[13] Fouque, J., Papanicolaou, G. and Sircar, K. (2000). *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press.

[14] Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer-Verlag New York.

[15] Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*. Vol. 6, No 2, pp. 327-343.

[16] Higham, D.J. (2006). An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*. Vol. 43, No 3, pp. 525-546.

[17] Hull, J. (2005). *Options,Futures And Other Derivatives* . Prentice Hall.

[18] Hull, J. and White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*. Vol. 42, No 2, pp. 281–300.

[19] Hull, J. and White, A. (1990). Pricing Interest-Rate-Derivative Securities. *Review of Financial Studies*. Vol. 3, No 4, pp. 573-592.

[20] Hutzenthaler, M., Jentzen, A. and Kloeden, P. (2012). Strong Convergence of an Explicit Numerical Method for SDE's with Nonglobally Lipschitz Continuous Coefficients *The Annals of Applied Probability*. Vol. 22, No 4, pp. 1611-1641.

[21] Kijima, M. (2013). *Stochastic Processes with Applications to Finance*. Chapman and Hall.

[22] Kloeden, P.E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag Berlin Heidelberg.

[23] Kou, S.G. (2002). A Jump-Diffusion Model for Option Pricing. *Management Science*. Vol. 48, No 8, pp. 1086–1101.

[24] Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*. Vol. 36, No 4, pp. 394-419.

[25] Merton, R.C. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*. Vol. 4, No 1, pp. 141-183.

[26] Merton, R.C. (1976). Option Pricing When Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*. Vol. 3, No 1-2, pp. 125-144.

[27] Mikosch, T. (1998). *Elementary Stochastic Calculus, with Finance in View* . World Scientific Publishing .

[28] Officer, R.R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association*. Vol. 67, No 340, pp. 807-812.

[29] Rouah, F.D. (2013). *The Heston Model and its Extensions in Matlab and C#*. Wiley.

[30] Schwartz, E. (1977). The Valuation of Warrants: Implementing a New Approach. *Journal of Financial Economics*. Vol. 4, No 1, pp. 79-93.

[31] Scott, L. (1987). Option Pricing When the Variance Changes Randomly: Theory, Estimation and an Application. *Journal of Financial and Quantitative Analysis*. Vol. 22, No 4, pp. 419–438.

[32] Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*. Vol. 5, No 2, pp. 177–188.

[33] Wiggins, J.B. (1987). Option Values Under Stochastic Volatility: Theory and Empirical Estimates.. *Journal of Financial Economics*. Vol. 19, No 2, pp. 51–372.

[34] Wilmott,P., Howison, S. and Dewynne,J. (1996). *The Mathematics of Financial Derivatives: A Student Introduction*. Cambridge University Press.