Comparison of Hedging Strategies in the Presence of Proportional Transaction Costs

Igor Jurievich Rodionov

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science.

26th June 2012

School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg.
Declaration

I declare that this dissertation is my own, unaided work, except where otherwise acknowledged. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

26th June 2012
Abstract

Complete market models are an idealisation. In reality, market frictions make any model incomplete, since it is not possible to eliminate all risk at an acceptable cost. This dissertation examines various methods of hedging European options in the presence of proportional transaction costs. Since trading on a continuous basis will incur infinite cost, all strategies result in a finite number of discrete trades. There are two distinct approaches, namely local-in-time and global-in-time methods. Local-in-time methods, such as the Leland algorithm, discretise time into fixed intervals and insist that trades take place at these times. This method uses the Black Scholes hedging approach with a modified volatility parameter. An alternative local-in-time method examined was the quadratic hedging which was developed by Schweizer. This approach minimises the quadratic cost function of the hedging strategy. In contrast, global-in-time methods employ a continuous time approach which optimally chooses between hedging (and paying a transaction cost) or not trading (and incurring hedge slippage). Davis, Panas and Zariphopoulou used a utility indifference pricing framework to derive a hedging technique. In particular, this was achieved by applying stochastic optimal control theory to derive a non-linear PDE. Monoyios proposed an alternative derivation using marginal utility. In both cases, the non-linear PDE is discretised to form a dynamic programming equation which is solved using a Markov chain approximation method. Since this is based on an augmented binomial tree, the approach is computationally expensive. Using asymptotic analysis on the above approach, Whalley and Wilmott derived an approximate analytical expression for the hedge rule. In order to evaluate the performance of these hedging methods, we generated a large number of geometric Brownian paths and hedged at-the-money European call options from inception to maturity. From this, histograms of the total hedging error (profit and loss) were computed, which allowed the comparison of VaR and expected shortfall values across all methods. Due to the computational complexity of these algorithms, it was necessary to compute the results using a parallel computing cluster. Our results show that the global-in-time methods perform favourably as expected.
For my parents,
Jury and Olga.
Acknowledgements

I would like to acknowledge my supervisor, Dr Thomas McWalter for your invaluable help, unwavering patience and continuous support that you have given me over the two years. This project would not have been possible without the countless hours we have spent together discussing various issues and topics. I am greatly indebted to Dr Coenraad Labuschagne for the financial support of the project. A special thank you is sent to Shun Pillay from the School of Computer Science for many hours spent helping me work on the cluster. I would also like to thank my fellow students and friends for the assistance and inspiration that you have shown me over this period. My family has given me infinite encouragement and I am truly grateful to you for being there, thank you. Last but not least, I would like to acknowledge everyone who has lent me their hand during this enlightening journey.
Contents

1. Introduction ......................................................... 1
   1.1 Strategies ...................................................... 3
       1.1.1 Wilmott Delta .............................................. 3
       1.1.2 Leland ...................................................... 4
       1.1.3 Quadratic Hedging ........................................ 4
       1.1.4 Utility Indifference Pricing .............................. 4
   1.2 Comparison ...................................................... 5
   1.3 Document Structure ........................................... 7

2. Discrete Hedging with Black Scholes ............................ 9
   2.1 Introduction .................................................... 9
   2.2 Black Scholes Model .......................................... 9
   2.3 Hedging Error .................................................. 12
   2.4 Gamma Hedging ................................................ 15
   2.5 Numerical Results ............................................ 17

3. Optimal Discrete Hedging ......................................... 19
   3.1 Introduction .................................................... 19
   3.2 Outline ......................................................... 19
   3.3 Derivation ...................................................... 20
   3.4 Pricing Equation .............................................. 23
   3.5 Numerical Results ............................................ 24

4. Discrete Hedging with Transaction Costs ........................ 29
   4.1 Introduction .................................................... 29
   4.2 The Leland Method ............................................ 29
   4.3 Criticism of Leland Strategy ................................. 32
   4.4 Numerical Results ............................................ 33

5. Quadratic Hedging ................................................ 36
   5.1 Introduction .................................................... 36
   5.2 Problem Formulation in Discrete Time ....................... 36
   5.3 Transaction Costs ............................................. 40
   5.4 Total Hedging Error .......................................... 43
   5.5 Numerical Implementation .................................... 44
   5.6 Modified Algorithm .......................................... 46
   5.7 Numerical Results ............................................ 48
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Distribution of hedging error for one long option</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>Distribution of hedging error for one long and one short option</td>
<td>17</td>
</tr>
<tr>
<td>3.1</td>
<td>Difference between Wilmott and Black Scholes delta</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Difference between Wilmott and Black Scholes price</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Difference between Black Scholes and Wilmott variances against drift</td>
<td>27</td>
</tr>
<tr>
<td>3.4</td>
<td>Difference between Black Scholes and Wilmott variances against hedging times</td>
<td>28</td>
</tr>
<tr>
<td>4.1</td>
<td>Total hedging error distributions for Leland strategy</td>
<td>35</td>
</tr>
<tr>
<td>5.1</td>
<td>Discretisation scheme for quadratic hedging approach</td>
<td>45</td>
</tr>
<tr>
<td>5.2</td>
<td>Effects of tree size on quadratic price</td>
<td>49</td>
</tr>
<tr>
<td>5.3</td>
<td>Effects of tree size on quadratic hedge ratio</td>
<td>50</td>
</tr>
<tr>
<td>5.4</td>
<td>Effects of hedging times on quadratic prices</td>
<td>50</td>
</tr>
<tr>
<td>6.1</td>
<td>Construction of the interpolated Markov process</td>
<td>59</td>
</tr>
<tr>
<td>7.1</td>
<td>Hedging bounds against maturity</td>
<td>78</td>
</tr>
<tr>
<td>7.2</td>
<td>Utility indifference prices against stock price</td>
<td>80</td>
</tr>
<tr>
<td>7.3</td>
<td>Utility indifference hedging bounds against stock price with $\gamma = 0.1$</td>
<td>81</td>
</tr>
<tr>
<td>7.4</td>
<td>Utility indifference hedging bounds against stock price with $\gamma = 1$</td>
<td>82</td>
</tr>
<tr>
<td>7.5</td>
<td>Utility indifference hedging bounds against stock price with $\gamma = 10$</td>
<td>83</td>
</tr>
<tr>
<td>7.6</td>
<td>Expected shortfall is plotted for the calibration of the gamma parameter</td>
<td>84</td>
</tr>
<tr>
<td>7.7</td>
<td>Value at risk is plotted for the calibration of the gamma parameter</td>
<td>85</td>
</tr>
<tr>
<td>7.8</td>
<td>Utility indifference bounds against risk aversion parameter</td>
<td>85</td>
</tr>
<tr>
<td>7.9</td>
<td>Utility indifference bounds against risk aversion parameter</td>
<td>86</td>
</tr>
<tr>
<td>7.10</td>
<td>Utility indifference bounds against volatility</td>
<td>86</td>
</tr>
<tr>
<td>7.11</td>
<td>Utility indifference bounds against number of time steps</td>
<td>87</td>
</tr>
<tr>
<td>8.1</td>
<td>Example of profit and loss distributions</td>
<td>99</td>
</tr>
<tr>
<td>8.2</td>
<td>Value at risk of PnL against hedging times</td>
<td>100</td>
</tr>
<tr>
<td>8.3</td>
<td>Expected shortfall of PnL against hedging times</td>
<td>100</td>
</tr>
<tr>
<td>8.4</td>
<td>Variance of PnL against hedging times</td>
<td>101</td>
</tr>
<tr>
<td>8.5</td>
<td>Mean of PnL against hedging times</td>
<td>101</td>
</tr>
<tr>
<td>8.6</td>
<td>Value at risk of PnL against proportional transaction costs</td>
<td>103</td>
</tr>
<tr>
<td>8.7</td>
<td>Expected shortfall of PnL against proportional transaction costs</td>
<td>103</td>
</tr>
<tr>
<td>8.8</td>
<td>Variance of PnL against proportional transaction costs</td>
<td>104</td>
</tr>
<tr>
<td>8.9</td>
<td>Mean of PnL against proportional transaction costs</td>
<td>104</td>
</tr>
</tbody>
</table>
8.10 Value at risk of PnL against volatility .......................... 105
8.11 Expected shortfall of PnL against volatility ........................ 106
8.12 Variance of PnL against volatility ................................. 106
8.13 Mean of PnL against volatility ................................. 107
List of Tables

4.1 Total hedging error statistics for Leland strategy .......................... 34
8.1 Scenario 1 parameters .................................................. 94
8.2 Scenario 2 parameters .................................................. 94
8.3 Scenario 3 parameters .................................................. 95
8.4 Time taken using 32 Matlab workers ................................. 95
8.5 Time taken using one Matlab worker ................................. 96
8.6 Scenario 1 prices and hedge ratios ................................... 97
8.7 Scenario 2 prices and hedge ratios ................................... 97
8.8 Scenario 3 prices and hedge ratios ................................... 97
Chapter 1

Introduction

One of the main problems studied in mathematical finance is the pricing of contingent claims. Initially this was done by either computing the discounted expected payoff of the claim or through economic reasoning which allowed market forces such as supply and demand to determine the price. In 1973, a new approach was developed in the seminal papers by Black and Scholes [6] and Merton [32]. For the first time, it was shown that an option could be replicated exactly by the use of a dynamic portfolio strategy under certain assumptions. The concept of hedging contingent claims is the subject of this dissertation. This introductory chapter aims to give a basic overview of the development of hedging contingent claims and introduce preliminary notation which is used throughout the dissertation.

A contingent claim \(X\) allows the holder to receive an amount at maturity \(T\) of the form

\[X = \Psi(S_T),\]

where \(\Psi\) is some real valued function. In particular, this definition describes a simple European contingent claim, since the payoff is only a function of the stock price \(S\) at maturity and not of the entire path. Specifically, this dissertation focuses on the most popular contingent claims, which have a fixed expiration date \(T\) and a payoff defined by

\[\Psi(S_T) = \max(\alpha(S_T - K), 0),\]

where \(\alpha = \pm 1\) indicates a call or a put option and \(K\) is a strike price. The pricing of the contingent claim in a Black Scholes model is described by a smooth function \(F\) with the terminal value corresponding to the payoff

\[F(T, S_T) = \Psi(S_T).\]  \(1.1\)

The innovation of Black and Scholes was to approach the problem of pricing these financial contracts by considering dynamic portfolio strategies that replicate (or hedge) the terminal payoff. In this dissertation we denote a portfolio strategy by
\[ \pi_t = (\vartheta_t, \eta_t), \] where \( \vartheta_t \) is the number of shares in the underlying asset (possibly fractional) and \( \eta_t \) is an amount in the bank account at time \( t \). For the purpose of the dissertation, time is discretised into fixed intervals unless mentioned otherwise. The value process for such a strategy is defined as

\[
V_t(\pi) = \vartheta_t S_t + \eta_t.
\]

A strategy is called self-financing if there is no exogenous infusion or withdrawal of money from the strategy. Mathematically, the self-financing condition in discrete time is

\[
\vartheta_t S_{t+1} + \eta_t (1 + r) = \vartheta_{t+1} S_{t+1} + \eta_{t+1},
\]

where \( r \) is the risk free rate for that period. This means that from inception until maturity, the strategy is self-supporting. The holding of the stock and the bank account can be reshuffled without changing the value of the strategy. The value only changes as a result of the changing stock price and the interest earned in the bank account. Returning to the literature, Black and Scholes [6] showed that an option payoff can be replicated exactly by a self-financing portfolio through the use of continuous trading. Assuming absence of arbitrage opportunities in the model means the price of the option is determined by the initial cost of the self-financing strategy, since the strategy and the option have equivalent payoffs. However, it was not clear that all contingent claims could be priced with the same principles.

Much rigorous mathematical formulation was developed in seminal papers by Harrison and Kreps [16] and Harrison and Pliska [17], where the link between martingale representation and contingent claim pricing was introduced. Furthermore, the concept of a complete market was formalised by Harrison and Pliska [18] by introducing the martingale representation property and proving uniqueness of the martingale measure. A market is said to be complete when every contingent claim has a unique price and can be replicated exactly with a self-financing strategy.

It is important to note that in a complete market all the risk and uncertainty can be hedged. However, the complete market model is only an approximation to reality. Discrete trading, transaction costs, portfolio constraints, stochastic volatility, jumps in the underlying process and other phenomena seen in reality introduce market incompleteness. When the market is incomplete, one can no longer replicate the option payoff exactly with a self-financing strategy. A unique preference free price and a unique hedging method no longer exist. As a result, the self-financing strategies produce residual risk which can be quantified as a hedging error. These strategies are the subject of this dissertation.
1.1 Strategies

To deal with conditions of market incompleteness, a subjective preference needs to be introduced which selects some criteria for minimising the residual risk. The possible criteria under examination are: minimising the variance of the hedging error, minimising the quadratic cost function or maximising investor utility. Associated with these criteria are optimal hedging strategies. The dissertation examines and focuses on the various methods of implementing such hedging strategies with market incompleteness introduced through discrete hedging and transaction costs. The strategies are divided into two groups: local-in-time and global-in-time. Local-in-time strategies attempt to minimise the residual risk for a predetermined hedging interval, while global-in-time strategies minimise the residual risk continuously, by hedging only when it is optimal to do so.

1.1.1 Wilmott Delta

Applying the Black Scholes strategy in discrete time is suboptimal, since the theory is designed to be used in a continuous setting. Boyle and Emanuel [7] determined the expression for the hedging error under the Black Scholes strategy when hedging takes place at discrete intervals. This hedging error is a function of the option gamma. A natural progression from this is to select a strategy which minimises the variance of the hedging portfolio. This is done by determining $\vartheta$ in the equation

$$\frac{\partial \text{Var}[\delta \Pi]}{\partial \vartheta} = 0,$$

where $\delta \Pi$ is the change of the hedging portfolio. The hedging portfolio $\Pi$ consists of a short option $F$ and an amount of $\vartheta$ in the underlying stock. Wilmott [43] provides a hedging strategy which is an adjustment to the Black Scholes delta. The adjustment term is a function of the stock drift term and the option gamma. Although in most cases the adjustment term is small, it can become significant in strong trending markets and near at-the-money positions. The Wilmott model is a local-in-time approach and is set in a discrete framework without transaction costs. The derivation of this hedge ratio was not fully provided in the article which led to the investigation of its derivation in this dissertation.
1.1 Strategies

1.1.2 Leland

In the paper by Leland [28], a hedging method is proposed under proportional transaction costs. The hedging error over one time step can be written as

\[
HE = \delta \Pi - r \delta t \Pi - TC.
\]

(1.3)

The Leland method chooses the hedging parameter such that it removes the transaction cost part, TC, from the hedging error. This is done by adjusting the volatility in the Black Scholes pricing equation. The Leland model is local-in-time as it attempts to minimise the hedging error for a predetermined hedging interval.

1.1.3 Quadratic Hedging

Another local-in-time approach is to consider the quadratic cost function of a hedging strategy given by

\[
E \left[ \left( C_T(\pi) - C_t(\pi) \right)^2 | \mathcal{F}_t \right],
\]

(1.4)

where \( C_t(\pi) \) is the cost process of the hedging strategy. The cost process of the hedging strategy is the accumulated difference between the intrinsic value of the contingent claim and the value process of a hedging strategy. This difference occurs in the incomplete market since hedging is no longer exact. By minimising this quadratic cost function it is possible to derive an optimal hedging strategy \( \pi \). This was first presented by Föllmer and Sondermann [15] where the underlying process was a martingale. Consequently Schweizer [38] extended it for the general semimartingale case. In his dissertation and consequently in [39], quadratic hedging was divided into local risk minimisation and mean-variance minimisation. The local risk minimisation involves deriving a strategy which is mean self-financing, which minimises the risk for each hedging interval. A numerical algorithm which implemented the local risk minimisation approach with transaction costs was provided by Mercurio and Vorst [30]. Furthermore, Lamberton, Pham and Schweizer [10] gave technical results for the case of quadratic hedging with transactions costs.

1.1.4 Utility Indifference Pricing

A utility indifference pricing approach uses a convex function denoted by \( U(x) \) which describes investor satisfaction with wealth \( x \). To determine a fair price of the option in this framework, one needs to consider two distinct problems. The first problem is represented by

\[
\sup_{\pi} \mathbb{E} [U(V_T(\pi))],
\]

(1.5)
where the investor wealth is invested in the market which consists of an underlying stock and a bank account. An investor attempts to maximise the expected utility of the wealth $V_T$ at some terminal time $T$. The second problem is represented by

$$\sup_\pi \mathbb{E}[U(V_T(\pi) - \Psi(S_T))],$$

where the investor sells an option, receives compensation for the option and invests the wealth in the market. Similarly, an investor wants to maximise his terminal expected utility while paying the option payoff. By considering these two problems, Hodges and Neuberger [20] pioneered the definition of the utility indifference price of the option. The utility indifference price of an option is determined when an investor is indifferent in utility terms between investing in the market with the option while receiving compensation and investing in the market without the option. The two proposed problems require the use of stochastic optimal control theory for their solution. Davis, Panas and Zariphopoulou [36] provided a global-in-time model inclusive of transaction costs by examining utility indifference pricing. Instead of minimising risk for a chosen hedging interval, a global-in-time model minimises risk at each time instant in a continuous time setting. The model hedges only when it is optimal to do so, giving it an advantage over local-in-time models where hedging is compulsory at fixed intervals. This was achieved by formulating partial differential equations (PDEs) using a Hamilton Jacobi Bellman equation which governs the option price and the hedging strategy. Monoyios [34] proposed an alternative algorithm which uses marginal utility to determine the hedging bounds. In both cases, the non-linear PDEs are discretised to form dynamic programming equations and solved using a Markov chain approximation method. Finally a different approach was taken by Whalley and Wilmott [41, 42], who performed an asymptotic analysis on the utility indifference pricing framework. The alternative hedging bounds for the strategy were approximated by a simple closed form expression based on a standard Black Scholes delta.

1.2 Comparison

This dissertation investigates these various hedging techniques for a European option in an incomplete market introduced by discrete hedging and transaction costs. Specifically, the parameter $\vartheta$ is derived for each strategy and its unique features are examined. To perform a fair comparison between the strategies examined, a hedging race was conducted to obtain the distribution of the total hedging errors (profit and loss). The distributions are obtained by running a Monte Carlo experiment which
hedges an option until maturity, using each of the discussed strategies. To obtain a reasonable distribution a large number of generated stock price paths were used. Running these strategies on a single point in the path is computationally expensive as some of the strategies use large binomial trees to compute the hedging parameters. The Monte Carlo experiment involves running these algorithms at multiple time points along multiple paths. To solve this computational problem in a reasonable amount of time, a cluster of computers was utilised. A problem of this nature is embarrassingly parallelisable, since each computer worker can be given a path to run the algorithms on. The results from this path are independent of the results of other workers. This allowed the usage of parallel computing techniques to complete the proposed Monte Carlo experiment. Notably without the use of the cluster, this experiment would take longer than a year to run and would not have been viable.

To make the comparisons fair, each strategy was given the same initial endowment to perform the hedging. The strategies were implemented as self-financing since no exogenous infusion or withdrawal of money was permitted, however money could be borrowed from the bank account to finance the required stock holding. Once the hedging race was complete, profit and loss was computed for each path and these distributions were compared. Furthermore, the parameters for the Monte Carlo experiment were chosen to simulate various trading regimes in the real world. Important care was taken to ensure the strategies were comparable since by definition global-in-time strategies are implemented continuously. For practical purposes, the global-in-time models were implemented discretely with hedging being performed at the same time points as the local-in-time models. Hedging was not, however, compulsory for the global-in-time models. Although this approach potentially penalised the utility indifference model, it was still expected to outperform the other approaches in the experiment.

For the comparison of hedging strategies, geometric Brownian motion was used to generate stock paths. It's been well documented that geometric Brownian motion has substantial drawbacks as it does not fully describe the market dynamics. To justify its use in this dissertation, it is important to note that the literature to date has primarily used geometric Brownian motion when dealing with transaction costs. The reason for this is that adjusting for transactions costs is difficult as it introduces market incompleteness. This issue is addressed further when discussing further research directions in Chapter 8.
1.3 Document Structure

The document is split by categorising each hedging technique into local-in-time and global-in-time approaches. Market incompleteness is introduced through the dissertation first by discrete hedging and later by proportional transaction costs. Finally the last chapter makes comparisons across the proposed models with transaction costs. This section now presents brief summaries of each of the chapters.

Chapter two introduces basic, preliminary results of the Black Scholes model used in continuous time. The Black Scholes model is then applied with discrete hedging intervals and an analytical expression for the hedging error is derived for a portfolio of options.

Chapter three continues with discrete hedging and derives the Wilmott optimal hedge ratio. Wilmott hedging method minimises the variance of the hedging portfolio by adjusting the Black Scholes delta. The significance of the adjustment term is examined numerically.

Chapter four introduces proportional transaction costs and examines the Leland method. The Leland method is shown to remove the transaction cost part of the hedging error by adjusting the Black Scholes volatility parameter. Furthermore the error in the Leland paper is examined and analysed numerically together with other features of the strategy.

Chapter five presents another local-in-time model by considering the quadratic hedging approach for European options with proportional transaction costs. A numerical method proposed by Mercurio and Vorst for local risk minimisation is demonstrated. A faster implementation of the method is provided.

Chapter six alters the direction of the dissertation and gives a broad overview of stochastic optimal control needed for the following chapter. A numerical approach, namely the Markov chain approximation method, for solving stochastic optimal control problems is also presented.

Chapter seven presents common properties of utility functions and reviews the utility indifference pricing framework of Davis, Panas and Zariphopoulou. The partial differential equations which govern the problem are derived and discretised to form dynamic programming equations. These equations are solved using a Markov chain approximation method. An outline of the Monoyios method is given which allows the derivation of the hedging bounds. The properties of the strategy are compared with the Whalley and Wilmott asymptotic analysis.

Chapter eight combines the results from previous chapters and numerically compares all hedging methods under proportional transaction costs. The technology used to compare the algorithms is discussed together with actual and estimated
times taken to run the algorithms on various machines. Furthermore, the advantages and disadvantages of each method are presented with the use of various market scenarios. Conclusions are drawn on the best strategy.

Lastly, it is important to note that this dissertation is aimed at market practitioners with strong quantitative backgrounds. The document provides a broad summary of each hedging technique and mostly ignores the subtleties of stochastic calculus. Although this review is neither comprehensive nor rigorous, this dissertation concentrates on practical issues regarding the implementation and characteristics of each hedging strategy.
Chapter 2

Discrete Hedging with Black Scholes

2.1 Introduction

The Black Scholes approach mitigates risk through the use of continuous hedging. While this leads to rigorous theoretical results, it is impossible to implement. In practice, hedging must be performed at discrete intervals of time. In this chapter, a basic overview of the Black Scholes strategy is given. It is then applied in a discrete setting and an analytical expression for the hedging error is derived. Finally the expression is investigated numerically.

2.2 Black Scholes Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space with filtration $\mathbb{F}$ generated by the Wiener process $W(t)$. In this space, the Black Scholes model uses a geometric Brownian motion for the underlying stock price given by

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

(2.1)

where $S(t)$ is a process which starts at $s_0$, $\mu$ is the drift and $\sigma$ is the volatility (diffusion). Note that for the purpose of this chapter a process is written in the given notation, however in subsequent chapters the notation $S_t$ is used.

Definition 2.1. A stochastic process $W(t)$ is called a Wiener process if the following conditions hold:

1. $W(0) = 0$.

2. The process $W(t)$ has independent increments.

3. For $s < t$ the stochastic variable $W(t) - W(s)$ has the Gaussian distribution with 0 mean and $\sqrt{t-s}$ variance.

4. $W(t)$ is continuous.
2.2 Black Scholes Model

Geometric Brownian motion is an important building block for modeling asset prices. Although in reality asset prices show many characteristics, known as stylised facts which are not explained by geometric Brownian motion, it is nonetheless a popular assumption. For the purpose of this dissertation, geometric Brownian motion will be used as a model for asset paths. The stochastic differential equation (2.1) can be solved directly using Itô calculus. The following proposition presents the result.

Proposition 2.2. The solution to the equation

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

is given by

\[ S(t) = s_0 \cdot \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W(t) \right). \]

The expected value is given by

\[ \mathbb{E}[S(t)] = s_0 \cdot \exp(\mu t). \]

Proof. See Chapter 5 in Björk [5].

To incorporate correlation structure in the underlying process, correlated geometric Brownian motion is used and given by

\[ dS_i(t) = \mu_i S_i(t) \, dt + \sigma_i S_i(t) \left( L \cdot d\overline{W}(t) \right)_i, \]

where the subscript \( i \) denotes the parameter applying to the \( i \)-th process, \( \overline{W} = (W^{(1)}(t) \cdots W^{(d)}(t))' \) is a vector of independent Wiener processes, \((\cdot)_i\) denotes the \( i \)-th element of a vector inside the brackets and \( L \) is the Cholesky decomposition of the correlation matrix \( \Sigma \) with \( L \cdot L^T = \Sigma \). For the purpose of the dissertation, in the multidimensional case the subscript notation is used to indicate the \( i \)-th stock, in the single stock case, the subscript will revert to indicate the time variable \( t \). The Black Scholes analysis uses a single geometric Brownian motion for the underlying stock price. The correlation structure does not impact the price of the portfolio of options as all the risk can be hedged in a continuous setting.

A hedging portfolio \( \Pi \) is examined which consists of a short option \( F \) and an amount of \( \frac{\partial F}{\partial S} \) in the underlying stock. Using Itô’s lemma to find a differential stochastic equation for the option and arranging terms, the portfolio dynamics can be written as

\[ d\Pi(t) = \left( -F_t - \frac{1}{2} S^2 \sigma^2 F_{SS} \right) dt, \]
where $F_t$ denotes a partial derivative of $F$ with respect to the variable $t$. Similarly $F_{SS}$ denotes a double partial derivative with respect to the variable $S$. This notation is used throughout the dissertation.

Through correct continuous rebalancing Black and Scholes showed that the stochastic element of the option and an underlying stock can be removed. By no arbitrage arguments, this results in the portfolio accumulating at a risk free rate and written as

$$d\Pi(t) = r\Pi dt. \tag{2.7}$$

Equations (2.6) and (2.7) are achieved if the hedger continuously holds $\vartheta_{BS} = \frac{\partial F}{\partial S}$ in the stock. The result implies that a continuously rebalanced delta hedge portfolio will replicate the value of the option at maturity. The strategy satisfies the self-financing condition as no additional cash flows are needed to replicate the option exactly. Although an European option is used for the purpose of this dissertation, it is noted that all contingent claims can be replicated exactly in this setup since it is a complete market.

Substituting (2.6) into (2.7) and using the definition of the portfolio, leads to the Black Scholes equation which gives the option price (2.10) when solved. This partial differential equation is a version of the heat equation and written as

$$F_t + \frac{1}{2}S^2\sigma^2F_{SS} + rSF_S - rF = 0. \tag{2.8}$$

The terminal boundary condition of the equation is the payoff of the option, which for a European calls ($\alpha = 1$) and puts ($\alpha = -1$) with strike $K$ is given by

$$F(S, K, 0, r, \sigma) = \Psi(S_T) = \max(\alpha(S_T - K), 0). \tag{2.9}$$

Black and Scholes solved equation (2.8) and showed that the corresponding price is determined by

$$F(\alpha, t, S, K, T, r, \sigma) = \alpha \left( SN \left( \alpha d^{(1)} \right) - Ke^{-r(T-t)} N \left( \alpha d^{(2)} \right) \right), \tag{2.10}$$

with

$$d^{(1)} = \frac{\ln \left( \frac{S}{K} \right) + (r + \frac{\sigma^2}{2}) (T - t)}{\sigma \sqrt{T - t}} \quad d^{(2)} = d^{(1)} - \sigma \sqrt{T - t},$$

where $t$ is the current time, $S$ is the current stock price, $K$ is the strike, $T$ is the maturity, $r$ is the risk free rate and $\sigma$ is the volatility.
2.3 Hedging Error

Although the Black Scholes model leads to powerful theoretical results, it disagrees with reality in a number of ways. One of which is the impractical assumption that hedging is possible in a continuous setting, as discussed by Derman and Taleb [13]. In practice an option hedger, who uses a Black Scholes model, can perform a discrete rebalanced delta hedge as follows:

- Sell one unit of the option at time $t = 0$.
- Compute $\vartheta^{BS} = \frac{\partial F}{\partial S}$ and buy this amount of the underlying stock. Borrow or invest the difference between the money received from the sale of the option and the cost of purchasing stock at the risk free rate, let this amount be $\eta^{BS}$.
- Wait a period of time. During this time, the underlying stock price may have changed, in which case the delta hedge is no longer correct.
- Compute the new value for $\vartheta^{BS}$ and adjust the stock holding accordingly. The difference is then placed in the new bank holding, $\eta^{BS}$.
- Repeat previous two steps until maturity of the option.
- At maturity, liquidate the stock holding and use the proceeds to pay the outstanding terminal amount of the option.

Performing this hedging strategy in a discrete setting results in residual risk as the hedger will have to finance the difference between the correct amounts in the stock and the bank account and the actual values at each hedging interval. This residual risk is called the hedging error while the total hedging error is simply the sum of all hedging errors over the lifetime of the option. The hedger has two choices, one is to subtract (or add) cash at each hedging update to cover the hedging error. Otherwise the hedger might be able to borrow (or invest) money from the bank account at each hedging update. Since the bank account is governed by the risk free rate, the hedging errors are different for both cases under a non-zero risk free rate. For the purpose of this dissertation, it is assumed that the hedger is able to fund (or invest) the hedging error using the bank account. In this case the total hedging error can be calculated as the difference between the payoff of the option and the liquidated value of the hedger’s portfolio at maturity.

Boyle and Emanuel [7] examine the hedging error under the Black Scholes model over one time step. The following extends their analysis to the case of $n$ options when the underlying processes are correlated geometric Brownian motions. Using
2.3 Hedging Error

the Euler scheme, equation (2.5) of the $i$-th process can be written as

\[
\frac{\delta S_i(t)}{S_i(t)} = \mu_i \delta t + \sigma_i (L\mathbf{\tau})_i \sqrt{\delta t},
\]

where $\delta(\cdot)$ denotes the discrete change in the variable, $(\cdot)_i$ denotes the $i$-th element of a vector inside the bracket, $\mathbf{\tau}$ is a vector of independent normal random variables and $L$ is the Cholesky decomposition of the correlation matrix $\Sigma$ with $L \cdot L^T = \Sigma$. For brevity in the following derivation the variable $t$ is dropped from $S_i(t)$ and instead written as $S_i$. The portfolio (\Pi) under consideration consists of $n$ call options ($F_i$) and amounts of $\partial F_i / \partial S_i$ in the respective underlying stocks. The portfolio dynamics are described by the following equations

\[
\Pi = \sum_{i=1}^{n} \left( F_i - \frac{\partial F_i}{\partial S_i} S_i \right),
\]

\[
\delta \Pi = \left( \sum_{i=1}^{n} \delta F_i \right) - \sum_{i=1}^{n} \left( \frac{\partial F_i}{\partial S_i} \delta S_i \right).
\]

Initially, the investment required to set up the portfolio is borrowed at the risk free rate $r$ and the hedging error is defined as

\[
\text{HE} = \Pi + \delta \Pi - \Pi e^{-rt}.
\]

Applying the exponential function series, (2.14) can be rewritten as

\[
\text{HE} = \delta \Pi - r \delta t \Pi + O(\delta t^2).
\]

Using the Black Scholes equation (2.10), Black Scholes delta for an option ($\partial F / \partial S = \alpha N(\alpha d_1)$) and substituting (2.12), (2.13) into (2.15) the hedging error expression becomes

\[
\text{HE} = \left( \sum_{i=1}^{n} \delta F_i - \sum_{i=1}^{n} \frac{\partial F_i}{\partial S_i} \delta S_i \right) + \alpha_r \delta t \sum_{i=1}^{n} \left( K_i e^{-rT_i} N(\alpha d_1^{(2)}) \right).
\]
Applying a Taylor series expansion the following expression for the sum of call options is
\[
\left( \sum_{i=1}^{n} \delta F_i \right) = \sum_{i=1}^{n} \left( \frac{\partial F_i}{\partial S_i} \delta S_i \right) + \sum_{i=1}^{n} \left( \frac{\partial F_i}{\partial t} \delta t \right) \\
+ \frac{1}{2} \sum_{i=1}^{n} \left( \frac{\partial^2 F_i}{\partial S_i^2} S_i^2 \sigma_i^2 (L\epsilon_i^2 \delta t) \right) + O(\delta t^{3/2}).
\] (2.17)

The Taylor series expansion uses the gamma of the option, which is defined by
\[
\Gamma = \frac{\partial^2 F}{\partial S^2}.
\] (2.18)

Using expression (2.17), the hedging error becomes
\[
\text{HE} = \sum_{i=1}^{n} \left( \frac{\partial F_i}{\partial t} \delta t \right) + \frac{1}{2} \sum_{i=1}^{n} \left( \frac{\partial^2 F_i}{\partial S_i^2} S_i^2 \sigma_i^2 (L\epsilon_i^2 \delta t) \right) \\
+ \alpha r \delta t \sum_{i=1}^{n} \left( K_i e^{-rT_i} N(\alpha d_i^{(2)}) \right) + O(\delta t^{3/2}).
\] (2.19)

Under the Black Scholes model, the rate of change in the option price with respect to time, theta, is written as
\[
\frac{\partial F}{\partial t} = -\frac{1}{2} \frac{\partial^2 F}{\partial S^2} S^2 \sigma^2 \alpha - \alpha r \left( K e^{-rT} N(\alpha d^{(2)}) \right).
\] (2.20)

Using this, the hedging error expression for \( n \) options is given by
\[
\text{HE} = \frac{1}{2} \sum_{i=1}^{n} \left( \frac{\partial^2 F_i}{\partial S_i^2} S_i^2 \sigma_i^2 (L\epsilon_i^2 - 1) \delta t \right).
\] (2.21)

Equation (2.21) is a general result and can be simplified for the case of non-correlated geometric Brownian motions since the Cholesky decomposition matrix \( L \) then becomes the identity matrix. The hedging error expression then becomes
\[
\text{HE} = \frac{1}{2} \sum_{i=1}^{n} \left( \frac{\partial^2 F_i}{\partial S_i^2} S_i^2 \sigma_i^2 (\epsilon_i^2 - 1) \delta t \right).
\] (2.22)

This is the result derived by Boyle and Emanuel [7]. In both cases the hedging error is proportional to the gamma of the option, the time increment, the square of the stock price and its volatility. The paper goes on to separate the hedging error into the stochastic component \( (L\epsilon_i^2 - 1) \) and a deterministic component \( \frac{\partial^2 F}{\partial S^2} S^2 \sigma^2 \delta t \).
the case of independent stocks, the stochastic part comprises of the squared normal random variable and the following proposition gives its distribution.

**Proposition 2.3.** Let \( \epsilon_1, \epsilon_2, \epsilon_3, \ldots, \epsilon_n \) have standard normal distributions, denoted \( N(0,1) \). If these variables are independent then \( Q = \sum_{i=1}^{n} \epsilon_i^2 \) has a chi-squared distribution with \( n \) degrees of freedom, denoted \( \chi^2_{(n)} \).

**Proof.** See Theorem 5.3-2 in Hogg and Tanis [21].

**Proposition 2.4.** The \( \chi^2_{(n)} \) distribution has the following probability distribution function

\[
    f(q) = \frac{1}{2^n \Gamma\left(\frac{n}{2}\right)} q^{\frac{n}{2}-1} e^{-\frac{q}{2}} \tag{2.23}
\]

and the following moments

\[
    \mathbb{E}[Q] = n \tag{2.24}
\]

\[
    \mathbb{E}[(Q - n)^2] = 2n \tag{2.25}
\]

\[
    \mathbb{E}[(Q - n)^3] = \sqrt{\frac{8}{n}} \tag{2.26}
\]

\[
    \mathbb{E}[(Q - n)^4] = \frac{12}{n} \tag{2.27}
\]

**Proof.** See Chapter 3 in Hogg and Tanis [21].

Although it seems that the hedging error cannot be manipulated, the hedger can choose the weights of the options carefully when shorting various options on the same underlying stock. The process of making the gamma of portfolio zero is known as gamma hedging.

### 2.4 Gamma Hedging

Delta neutrality provides protection from small movements in the underlying stock albeit with the already mentioned hedging error. The movement of an underlying stock results in changes in the delta thus necessitating rebalancing. The gamma of the option is the rate of change of the option delta with respect to the price of the underlying stock. This implies that if the gamma is high, rebalancing needs to be done more often and if the gamma is low, a delta hedge can be kept for a longer period. This is clearly reflected in the hedging error expression (2.21) in the gamma term. To manage the gamma risk, option hedgers can construct their portfolio to be gamma neutral. This protects the portfolio from large movements in the underlying stock.
stock price. Since the delta of underlying stock equals one and the gamma equals zero, hedgers cannot use the underlying stock to change the gamma of the portfolio. An option on the same underlying stock is used to achieve gamma neutrality with the following procedure:

- Acquire an amount of options on the same underlying stock such that the gamma of the portfolio is zero. This portfolio will generally not be delta neutral. Due to the positivity of the gamma for both call and put options, an opposite position is required.

- Now add the underlying stock in order to make the portfolio delta neutral.

Note that the procedure cannot be performed in a different order since acquiring a new position in the option will destroy the delta neutrality of the first step. Achieving gamma neutrality would minimise the hedging error expression (2.21) over one time step in relation to price changes. The hedging error will not be zero as the gamma terms could change over the hedging interval.

![Graph showing the distribution of hedging error](image-url)

**Fig. 2.1:** The distribution of the hedging error over one time step is given for one long option. The parameters are: $S(0) = 100$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $r = 0.04$ and $\delta t = \frac{0.25}{\sqrt{3}}$. 
2.5 Numerical Results

In this section, the distribution of hedging error over one time step is examined when hedging with Black Scholes delta. The shape of the hedging error for one option can be seen in Figure 2.1. The shape comes from a chi-squared distribution with one degree of freedom. It is positively skewed as the third moment around the mean for this distribution is $\sqrt{8}$. Note that the hedging error was derived for holding the option long. Generally hedgers would hold the options short and the distribution would be reversed giving hedgers a long negative tail. It is important to note that the hedging error is over one time step and the distribution of the total hedging error is symmetrical as will be seen in the numerical section of Chapter 8.

Figure 2.2 examines the distribution of hedging error for the case of a portfolio of one long and one short option. The correlation coefficient has an effect on the distribution of the two options, however it is not significant. These results can easily be expanded to any number of options in the portfolio using the hedging error expression.

![Figure 2.2](image)

**Fig. 2.2:** The distribution of the hedging error over one time step is given for a portfolio of a long and a short call option. The parameters are: $S_1(0) = 100, S_2(0) = 100, K_1 = 100, K_2 = 100, T = 0.25, \sigma_1 = 0.3, \sigma_2 = 0.35, r = 0.04$ and $\delta t = 0.25$.

In the discrete setting, using a Black Scholes strategy is not in any way optimal. Other strategies could produce hedging errors with better distribution statistics. The
next chapter examines a strategy which minimises the profit and loss distribution over one time step.
Chapter 3

Optimal Discrete Hedging

3.1 Introduction

The Black Scholes strategy in continuous time leads to an elegant solution. The strategy was examined in the previous chapter, however it may not be the most optimal way of performing hedging in a discrete setting. This section examines a hedging strategy introduced by Wilmott [43] set in discrete time, which minimises the variance of the hedging error. This chapter largely consists of a derivation of the Wilmott optimal hedge ratio and examines its properties numerically.

3.2 Outline

Using the definition for the hedging portfolio $\Pi$ for one option in (2.12), the following expression gives the variance of the hedge portfolio

$$\text{Var} [\delta \Pi] = \mathbb{E} [\delta \Pi^2] - \mathbb{E} [\delta \Pi]^2,$$

(3.1)

where the change in the hedge portfolio is

$$\delta \Pi = \delta F - \vartheta \delta S.$$

To find the optimal hedge ratio denoted by $\vartheta^W$ which minimises (3.1), we solve for $\vartheta^W$ in the expression

$$\frac{\partial \text{Var} [\delta \Pi]}{\partial \vartheta} = 0.$$

(3.2)

The bank holding for the Wilmott strategy $\eta^W$ is not given directly, but it can be determined by imposing the self-financing condition.
3.3 Derivation

The expression for the variance of the hedge portfolio is obtained using a strong Taylor approximation for the underlying process.

**Proposition 3.1.** Consider an Itô process \( X = X_t, t_0 \leq t \leq T \) satisfying the following scalar stochastic differential equation:

\[
\text{d}X_t = A(t, X_t) \text{d}t + B(t, X_t) \text{d}W_t
\]

with initial value \( X_{t_0} = X_0 \).

For a given discretisation \( t_0 = i_0 < i_1 < ... < i_n < ... < i_N = T \) of the time interval \([t_0, T]\), the Platen and Wagner approximation, denoted by \( Y \), is an order 1.5 strong Taylor scheme given by:

\[
Y_{n+1} = Y_n + A\Delta \tau + B \Delta W + \frac{1}{2} BB_X (\Delta W)^2 - \Delta \tau
\]

\[
+ A_X B \Delta Z + \frac{1}{2} \left( AA_X + \frac{1}{2} B^2 A_{XX} \right) (\Delta \tau)^2
\]

\[
+ \left( AB_X + \frac{1}{2} B^2 B_{XX} \right) [\Delta W \Delta \tau - \Delta Z]
\]

\[
+ \frac{1}{2} B \left( BB_{XX} + (B_X)^2 \right) \left[ \frac{1}{3} (\Delta W)^2 - \Delta \tau \right] \Delta W. \tag{3.3}
\]

where \( \Delta \tau = \tau_{n+1} - \tau_n \), \( \Delta W = W_{\tau_{n+1}} - W_{\tau_n} \) and \( \Delta Z_t = \int_{\tau_n}^{\tau_{n+1}} \int_{t}^{s_2} \text{d}W_{s_1} ds_2 \). The subscript in \( A_X \) denotes a partial derivative of \( A \) with respect to the variable \( X \).

Similarly a double partial derivative with respect to the variable \( X \) is denoted by \( A_{XX} \).

**Proof.** See Chapter 10 in Kloeden and Platen [24]. \( \blacksquare \)

Applying the strong Taylor approximation of order 1.5 to geometric Brownian
motion in (2.1), the following expressions for the change in $S$ are derived

$$\delta S = \sigma S \epsilon \delta t + \left( \mu + \frac{1}{2} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) S \delta t$$

$$+ \left( \mu + \frac{1}{6} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) \sigma S \epsilon \delta t^2 + O(\delta t^3),$$

$$(\delta S)^2 = \sigma^2 S^2 \epsilon^2 \delta t + 2 \sigma S^2 \epsilon \left( \mu + \frac{1}{2} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) \delta t^2$$

$$+ \left[ \left( \mu + \frac{1}{2} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right)^2 S^2 + 2 \left( \mu + \frac{1}{6} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) \sigma^2 S^2 \epsilon^2 \right] \delta t^2 + O(\delta t^2),$$

$$(\delta S)^3 = \sigma^3 S^3 \epsilon^3 \delta t^2 + 3 \sigma^2 S^3 \epsilon^2 \left( \mu + \frac{1}{2} \sigma^2 \epsilon - \frac{1}{2} \sigma^2 \right) \delta t^2 + O(\delta t^2),$$

The change in the derivative value $\delta F$ is expanded using a Taylor expansion with respect to two variables $\delta t$ and $\delta S$

$$\delta F = F_t \delta t + F_S \delta S_t + \frac{1}{2} (\delta S_t)^2 F_{SS} + (\delta S_t) F_{St} (\delta t) + \frac{1}{2} F_{tt} (\delta t)^2 + \frac{1}{6} F_{SSS} (\delta S_t)^3$$

$$= \sigma \epsilon S F_S (\delta t)^\frac{1}{2}$$

$$+ \left[ F_t + S \left( \mu + \frac{1}{2} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) \right] \sigma S \epsilon \left( \mu + \frac{1}{2} \sigma^2 \epsilon - \frac{1}{2} \sigma^2 \right) \right) F_S$$

$$+ \left[ \left( \mu + \frac{1}{6} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) \sigma S \epsilon + \sigma S^2 \left( \mu + \frac{1}{2} \sigma^2 \epsilon - \frac{1}{2} \sigma^2 \right) \right] F_{SS}$$

$$+ \frac{1}{2} F_{tt} + \frac{1}{2} S^2 \left( \mu + \frac{1}{2} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right)^2 F_{SS}$$

$$+ \left( \mu S + \frac{1}{2} \sigma^2 S \epsilon - \frac{1}{2} \sigma^2 \right) F_{St} + \left( \mu + \frac{1}{6} \sigma^2 \epsilon^2 - \frac{1}{2} \sigma^2 \right) \sigma^2 S^2 \epsilon^2 F_{SS}$$

$$+ \frac{3}{6} \sigma^2 S^3 \epsilon^2 \left( \mu + \frac{1}{2} \sigma^2 S \epsilon^2 - \frac{1}{2} \sigma^2 S \right) F_{SSS} \right] (\delta t)^2$$

$$+ O(\delta t)^2,$$

where $F_t$ denotes a partial derivative of $F$ with respect to the variable $t$. Similarly $F_{St}$ denotes a double partial derivative with respect to the variables $S$ and $t$. The expression for the derivative of the variance with respect to $\vartheta$ in (3.2) is expanded as

$$\frac{\partial \text{Var} [d\Pi]}{\partial \vartheta} = 2 \vartheta \mathbb{E} [\delta S^2] - 2 \vartheta \mathbb{E} [\delta S] \mathbb{E} [\delta F] + 2 \mathbb{E} [\delta F] \mathbb{E} [\delta S] - 2 \mathbb{E} [\delta F \delta S], \quad (3.4)$$
since the variance of the hedge portfolio is written fully as

\[
\text{Var}[\delta \Pi] = \mathbb{E} [\delta \Pi^2] - \mathbb{E} [\delta \Pi]^2 \\
= \mathbb{E} [(\delta F - \vartheta \delta S)^2] - \mathbb{E} [\delta F - \vartheta \delta S]^2 \\
= \mathbb{E} [\delta F^2] - 2 \mathbb{E} [\delta F \delta S] \vartheta + \mathbb{E} [\delta S^2] (\vartheta)^2 - \mathbb{E} [\delta F]^2 + 2 \mathbb{E} [\delta F] \mathbb{E} [\delta S] \vartheta \\
- \mathbb{E} [\delta S^2] (\vartheta)^2.
\]

Finally, the hedge ratio \( \vartheta_W \) which gives a minimum variance for the hedge portfolio is determined by

\[
\vartheta_W = \frac{\mathbb{E} [\delta F] \mathbb{E} [\delta S] - \mathbb{E} [\delta F \delta S]}{\mathbb{E} [\delta S^2] - \mathbb{E} [\delta S^2]}. \tag{3.5}
\]

Using the moments of normal random variable \( \epsilon \), the following expectations are evaluated

\[
\mathbb{E} [\delta S] = \mu_S \delta t + O(\delta t^2), \tag{3.6}
\]

\[
\mathbb{E} [\delta S^2] = \mu^2 S^2 \delta t^2 + O(\delta t^4), \tag{3.7}
\]

\[
\mathbb{E} [\delta S^2] = \sigma^2 S^2 \delta t + \mu^2 S^2 \delta t^2 + \frac{1}{2} \sigma^4 S^2 \delta t^2 + 2 \mu \sigma^2 S^2 \delta t^2 + O(\delta t^2), \tag{3.8}
\]

\[
\mathbb{E} [\delta F] = \left( F_t + \mu S F_S + \frac{1}{2} \sigma^2 S^2 F_{SS} \right) \delta t \\
+ \left( \frac{1}{2} F_{tt} + \frac{1}{2} \sigma^2 S^2 F_{SS} + \mu S F_{St} + \frac{1}{6} (\mu^2 S^3 + \sigma^4 S^3) F_{SSS} \right) \delta t^2 \\
+ O(\delta t^2), \tag{3.9}
\]

\[
\mathbb{E} [\delta F] \mathbb{E} [\delta S] = \left( \mu S F_t + \mu^2 S^2 F_S + \frac{1}{2} \mu \sigma^2 S^3 F_{SS} \right) \delta t^2 + O(\delta t^3), \tag{3.10}
\]

\[
\mathbb{E} [\delta F \delta S] = F_S \sigma^2 S^2 \delta t \\
+ \left( \mu S F_t + \mu^2 S^2 F_S + \frac{1}{2} \sigma^4 S^2 F_S + \frac{3}{2} \mu \sigma^2 S^3 F_{SS} + \frac{3}{2} \sigma^4 S^3 F_{SSS} \right) \delta t^2 + O(\delta t^3). \tag{3.11}
\]

After some algebra and ignoring terms of order \( O(\delta t^2) \), the expression for the hedge ratio becomes

\[
\vartheta_W \approx F_S + \left( \mu S F_{SS} + \frac{3}{2} \sigma^2 S F_{SSS} + F_{St} + \frac{1}{2} \sigma^2 S^2 F_{SSS} \right) \delta t \\
(1 + \frac{1}{2} \sigma^2 \delta t + 2 \mu \delta t)}.
\]
The $F_{St}$ term is a double derivative evaluated using the Black Scholes PDE (2.8)

\[ F_{St} = \frac{\partial}{\partial S} F_t = \frac{\partial}{\partial S} \left( rF - rSF_s - \frac{1}{2} \sigma^2 S^2 F_{SS} \right) = -rSF_{SS} - \sigma^2 SF_{SS} - \frac{1}{2} \sigma^2 S^2 F_{SSS}. \]

Finally the hedge ratio is written as the Black Scholes delta with an adjustment term

\[ \vartheta \approx F_S + \frac{\left( \frac{3}{2} \sigma^2 SF_{SS} + \mu SF_{SS} - (rSF_{SS} + \sigma^2 SF_{SS} + \frac{1}{2} \sigma^2 S^2 F_{SSS}) + \frac{1}{2} \sigma^2 S^2 F_{SSS} \right) \delta t}{(1 + \frac{1}{2} \sigma^2 \delta t + 2 \mu \delta t)} \]

\[ = F_S + \frac{\left( \mu - r + \frac{\sigma^2}{2} \right) SF_{SS} \delta t}{(1 + \frac{1}{2} \sigma^2 \delta t + 2 \mu \delta t)} \]

\[ = F_S + \left( \mu - r + \frac{\sigma^2}{2} \right) SF_{SS} \delta t \left( 1 - \left( \frac{1}{2} \sigma^2 \delta t + 2 \mu \delta t \right) + O(\delta t^2) \right) \]

\[ = F_S + \left( \mu - r + \frac{\sigma^2}{2} \right) SF_{SS} \delta t + O(\delta t^2). \]

Note the last result follows from using a Taylor series expansion on the denominator.

The adjustment term is a function of the drift of the underlying asset, the gamma of the option, the current stock price, the volatility of the underlying asset and the time increment. Wilmott argues that dependence on the drift is a result of the residual risk of a hedging portfolio in a discrete setting and consequently exposure to the underlying stock when writing the option.

### 3.4 Pricing Equation

Having derived the optimal hedge ratio, the option now needs to be priced. The Black Scholes pricing equation should not be used as we are now in a discrete setting with an adjusted hedge ratio. The pricing equation is derived using a discrete version of (2.7), which is written as

\[ E[\delta \Pi] = \left( r \delta t + \frac{1}{2} r^2 \delta t^2 + \cdots \right) \Pi. \]

The equation states that the expected return on the hedged portfolio is risk free and uses a power series approximation for the exponential function $e^{r \delta t} = 1 + r \delta t + \frac{1}{2} r^2 \delta t^2 + \cdots$. Substituting the expressions for $E[\delta \Pi]$ from (3.6), (3.9), (3.12) and
collecting $O(\delta t)$ terms, the option pricing equation becomes

$$F_t + \frac{1}{2}S^2\sigma^2 F_{SS} + rSF_S - rF + S^2(\mu - r)\left(r - \mu - \frac{\sigma^2}{2}\right)F_{SS}\delta t = 0. \quad (3.14)$$

The pricing equation also contains the drift of the process $\mu$. Risk neutral valuation is used when perfect hedging can be achieved. This is not possible in reality and the hedger is exposed to some risk which manifests itself in the drift term. The pricing equation derived is similar to the Black Scholes equation (2.8). This allows the use of a volatility adjustment $\sigma^*$ with the Black Scholes option pricing formula (2.10), where

$$\sigma^* = \sigma \sqrt{1 + \frac{2\delta t(\mu - r)}{\sigma^2}(r - \mu - \frac{\sigma^2}{2})}. \quad (3.15)$$

Notice that the new volatility adjustment $\sigma^*$ is less than the process volatility $\sigma$ when $\mu > r$.

### 3.5 Numerical Results

The new hedge ratio (3.12) is examined numerically in this section. In practice, it is difficult to estimate the drift of the underlying stock as there is only one realised path. Hence it is important to note when the adjustment becomes significant. In general the difference between the Black Scholes and Wilmott model is small, however in strong trending markets the drift term can increase the adjustment term. Furthermore the size of $\delta t$ also dictates the size of the correction term. By hedging more frequently, the correction term is diminished due to the small value of $\delta t$.

Figure 3.1 examines the differences between the derived hedge ratio and the Black Scholes delta. The greatest difference for the hedge ratio happens at-the-money of the option. The drift term also plays a more significant role as the size of the adjustment grows considerably when the drift term is increased.

Figure 3.2 examines the difference between the new adjusted price and the Black Scholes price with the same parameters. Similarly, the greatest price difference also occurs in the same region compounded with large values of the drift term. Wilmott argues that in trending markets the new adjustment in the hedge ratio will result in a better risk reduction since hedging is done with a view and at each step the variance of the hedging error is reduced. Although this is the case, in most cases there is a very small discrepancy between the hedge ratios and the prices of both models as seen by the scale in the figures.

To examine the reduction in variance, a Monte Carlo simulation was performed which hedges a call option using the vanilla Black Scholes strategy and Wilmott
strategy. The Monte Carlo simulation uses 10000 paths to estimate the variance of the hedging errors for both strategies.

Figures 3.3 and 3.4 plot the estimated variance against the drift and the number of hedging times, respectively. It is clear that the reduction in variance is significant only under a large drift or a small number of hedging times. Hence the adjustment term needs to be used under those scenarios.

Consequently, it has been shown that hedging using the Black Scholes delta is not optimal when performed in the discrete setting. The Wilmott adjustment was shown to improve the hedging in certain scenarios. The dissertation now shifts focus to investigate various trading strategies in the context of discrete trading with transaction costs.
Fig. 3.2: The difference between the Wilmott and the Black Scholes price is plotted against the stock price for various values of the drift. The parameters are: $K = 100$, $T = 0.25$, $\sigma = 0.3$, $r = 0.04$ and $\delta t = \frac{0.25}{90}$. 
Fig. 3.3: The difference between the Black Scholes variance and the Wilmott variance is plotted against the drift term of the price process. The parameters are: $S_0 = 95$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $r = 0.04$ and $\delta t = \frac{0.25}{360}$. 
Fig. 3.4: The difference between the Black Scholes variance and the Wilmott variance is plotted against the number of times hedging was performed during the life of the option. The parameters are: $S_0 = 100$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $\mu = 0.08$ and $r = 0.04$. 
Chapter 4

Discrete Hedging with Transaction Costs

4.1 Introduction

The modeling of transaction costs when hedging options is another deviation from the complete market model. Leland [28] argues that there are problems with using a naive Black Scholes strategy and hedging discretely with transaction costs. The error in replication will arise from two factors, hedging error and transaction costs. The hedger of the option will need to balance these two factors by choosing a revision interval. The hedging period can be decreased in the hope of minimising the hedging error, however the transaction cost can then dominate the replication cost. There is also uncertainty of the transaction costs payable by the hedger since these are path dependent. Leland proposed an option replicating strategy which modifies the volatility parameter according to the size of the transaction costs and the trading interval chosen. The proposed strategy is used with fixed discrete trading intervals and gives hedging errors which are uncorrelated to the market and tend to zero with frequent rebalancing. Furthermore, under arbitrarily small transaction costs the Leland strategy coincides with the Black Scholes strategy.

4.2 The Leland Method

Leland defines the proposed strategy that depends on the percent transaction cost \( \lambda \) and the fixed hedging period \( \delta t \). He then proves that this strategy will replicate the option payoff inclusive of transaction costs. Define the new modified volatility parameter in terms of the underlying stock \( S \) with drift \( \mu \) and volatility \( \sigma \) as

\[
\hat{\sigma}^2(\sigma^2, \lambda, \delta t) = \sigma^2 \left[ 1 + \frac{2\lambda}{\sigma^2 \delta t} \mathbb{E} \left[ \left| \frac{\delta S}{S} \right| \right] \right] \\
\approx \sigma^2 \left[ 1 + \sqrt{\frac{2}{\pi}} \frac{2\lambda}{\sigma \sqrt{\delta t}} \right],
\]

(4.1)
since

$$\mathbb{E} \left[ \frac{\delta S}{S} \right] \approx \mathbb{E} \left[ \mu \delta t + \sigma \sqrt{\delta t} \epsilon \right]$$

$$\leq |\mu| \delta t + \sigma \sqrt{\delta t} \mathbb{E} [ |\epsilon| ]$$

$$= |\mu| \delta t + \sigma \sqrt{\delta t} \sqrt{\frac{2}{\pi}}$$

$$\approx \sigma \sqrt{\delta t} \sqrt{\frac{2}{\pi}}, \quad (4.2)$$

where $|\epsilon|$ is an absolute standard normal random variable. The distribution of an absolute normal variable with parameters $\mu$ and $\sigma^2$ is a folded normal. The expectation is given by $\sigma \sqrt{\frac{2}{\pi}} \exp \left( -\frac{\mu^2}{2 \sigma^2} \right) + \mu \left( 1 - 2N \left( \frac{-\mu}{\sigma} \right) \right)$, where $N(\cdot)$ is the cumulative distribution function of a standard normal variable. Note that the drift term is removed since for small $\delta t$ the expression will be dominated by the second term.

The modified variance in (4.1) is used to price the option with the Black Scholes formula in (2.10). In a similar manner to the analysis in Chapter 2 on the hedging error, after the transaction costs, can be written for the new strategy as:

$$\text{HE} = \delta \Pi - r \delta t \Pi - TC$$

$$= \frac{\partial \hat{F}}{\partial S} \delta S - \delta \hat{F} - \left( \frac{\partial \hat{F}}{\partial S} S - \hat{F} \right) r \delta t - TC, \quad (4.3)$$

where TC is the term which arises from paying transaction costs

$$\text{TC} = \lambda \left| \delta \left( \frac{\partial \hat{F}}{\partial S} \right) (S + \delta S) \right|$$

$$= \lambda \left| \frac{\partial^2 \hat{F}}{\partial S^2} \delta S (S + \delta S) \right| + O(\delta t^{\frac{3}{2}})$$

$$= \lambda \frac{\partial^2 \hat{F}}{\partial S^2} S^2 \left| \delta S \right| + O(\delta t^{\frac{3}{2}}). \quad (4.4)$$

Using a Taylor series expansion for a function of two variables $S$ and $t$, we get the following approximation

$$\delta \hat{F} = \frac{\partial \hat{F}}{\partial S} \delta S + \frac{\partial \hat{F}}{\partial t} \delta t + \frac{1}{2} \frac{\partial^2 \hat{F}}{\partial S^2} (\frac{\delta S}{S})^2 + O(\delta t^{\frac{3}{2}}). \quad (4.5)$$
4.2 The Leland Method

Substituting (4.4) and (4.5) into (4.3) gives

\[
HE = \left( \hat{F} - \frac{\partial \hat{F}}{\partial S} S \right) r \delta t - \frac{\partial \hat{F}}{\partial t} \delta t - \frac{1}{2} \frac{\partial^2 \hat{F}}{\partial S^2} \left( \frac{\delta S}{S} \right)^2 + \frac{1}{2} \lambda \frac{\partial^2 \hat{F}}{\partial S^2} \left| \frac{\delta S}{S} \right| + O(\delta t^{3/2}). \tag{4.6}
\]

Using the expression for theta in (2.20)

\[
\left( \hat{F} - \frac{\partial \hat{F}}{\partial S} S \right) r = \frac{1}{2} \frac{\partial^2 \hat{F}}{\partial S^2} S^2 \sigma^2 + \frac{\partial \hat{F}}{\partial t}, \tag{4.7}
\]

and substituting it into (4.6), the equation becomes

\[
HE = \frac{1}{2} \frac{\partial^2 \hat{F}}{\partial S^2} S^2 \left[ \sigma^2 \delta t - \left( \frac{\delta S}{S} \right)^2 - 2 \lambda \left| \frac{\delta S}{S} \right| \right] + O(\delta t^{3/2}). \tag{4.8}
\]

Finally, substituting the expression for the modified variance (4.1), we get an expression for the hedging error inclusive of the transaction costs

\[
HE = \frac{1}{2} \frac{\partial^2 \hat{F}}{\partial S^2} S^2 \left[ \sigma^2 \delta t - \left( \frac{\delta S}{S} \right)^2 - 2 \lambda \left[ \mathbb{E} \left[ \left| \frac{\delta S}{S} \right| \right] - \left| \frac{\delta S}{S} \right| \right] \right] + O(\delta t^{3/2}). \tag{4.9}
\]

Under expectation the hedging error is shown to be \(O(\delta t^{3/2})\). Thus the total expected hedging error is \(O(\delta t^{1/2})\) and approaches zero as \(\delta t\) becomes small. Leland strategy has the following important properties.

- It replicates the option with transaction costs and the error is shown to tend to zero as the hedging period becomes shorter.
- Transaction costs become bounded as the hedging period is made shorter.
- The transaction costs can be calculated, given the hedging frequency and the option price can be determined.
- Leland shows that his strategy converges to the Black Scholes approach in the presence of arbitrarily small transaction costs.

The adjustment was chosen such that it removes the transaction cost part from the hedging error. Finally to price options using the Leland method, one simply uses the modified variance (4.1) with the Black Scholes option pricing formula (2.10). The hedger also needs to be compensated for the cost incurred when the hedge portfolio is set up. The hedge ratio for the Leland strategy is simply \(\vartheta_L = \frac{\partial \hat{F}}{\partial S}\).
4.3 Criticism of Leland Strategy

where the derivative equals to $\alpha N(\alpha \hat{d}(1))$. Similarly to the Wilmott strategy the bank holding $\eta$ can be determined by imposing the self-financing condition on the strategy. Depending on the initial position, further transaction costs are paid when purchasing or selling the underlying stock at inception. In the case of an initial all-cash holding, $\lambda S N(\alpha \hat{d}(1))$ is paid in transaction costs. The hedger needs to be compensated for this cost in the option price which is given by

$$
\hat{F}(\alpha, t, S_t, K, T, r, \hat{\sigma}, \lambda) = \alpha \left( S_t N(\alpha \hat{d}(1)) - Ke^{-r(T-t)} N(\alpha \hat{d}(2)) \right) + \lambda S_t N(\alpha \hat{d}(1)).
$$

(4.10)

In a case of all-stock holding, $\lambda S(1 - \alpha N(\alpha \hat{d}(1)))$ is paid in transaction costs. The hedger again needs to be compensated for this cost and the option price is then given by

$$
\hat{F}(\alpha, t, S_t, K, T, r, \hat{\sigma}, \lambda) = \alpha \left( S N(\alpha \hat{d}(1)) - Ke^{-r(T-t)} N(\alpha \hat{d}(2)) \right) + \lambda S(1 - \alpha N(\alpha \hat{d}(1))).
$$

(4.11)

4.3 Criticism of Leland Strategy

Zhao and Ziemba [45, 46] in correspondence with Leland [29] identified a mathematical flaw in the original Leland paper [28]. This mathematical flaw is related to the problem of uniform convergence and will be presented in this section. In the original paper, Leland showed that the hedging error over one time step is $O(\delta t^{3/2})$. This was done by closely examining the $\frac{\partial^2 \hat{F}}{\partial S^2} S^2$ term. Recall

$$
\frac{\partial^2 \hat{F}}{\partial S^2} S^2 = S N'(\hat{d}(1)) \frac{\hat{\sigma}}{\sqrt{T-t}}^{1/2},
$$

where

$$
N'(\hat{d}(1)) = \frac{\exp(-\frac{1}{2}(\hat{d}(1))^2)}{\sqrt{2\pi}},
$$

$$
\hat{d}(1) = \frac{\ln(S/K) + (r + \hat{\sigma}^2/2)(T-t)}{\hat{\sigma} \sqrt{T-t}},
$$

$$
\hat{\sigma}^2(\sigma^2, \lambda, \delta t) = \sigma^2 \left[ 1 + \frac{\sqrt{2} \cdot 2 \lambda}{\pi \sigma \sqrt{\delta t}} \right].
$$

In the original paper, Leland shows that

$$
\hat{d}(1) \to \frac{1}{2} \hat{\sigma}(T-t)^{1/2} \sim O(\delta t^{-5/4}) \text{ as } \hat{\sigma} \to \infty \text{ i.e. as } \delta t \to 0,
$$

(4.12)
which results in the term $\frac{\partial^2 \hat{F}}{\partial S^2} s^2$ being $O(\delta t^{\frac{3}{2}})$. Consequently this is required for the hedging error over one time step to be $O(\delta t^{\frac{3}{2}})$. However, the convergence in (4.12) only holds point wise. This means that it is true for a given time $t \in [0, T]$, but since the total hedging error is a sum of hedging errors over the subintervals, the convergence needs to be uniform over the entire interval $[0, T]$. Zhao and Ziemba show that just before maturity of the option, at time $t = T - \delta t$,

$$\hat{\sigma}(T - t)^{\frac{1}{2}} = \sigma \sqrt{\delta t + \frac{2\lambda \sqrt{\delta t}}{\sigma}}$$

(4.13)

and

$$\hat{\sigma}(T - t)^{\frac{1}{2}} \to 0 \text{ as } \delta t \to 0.$$  (4.14)

This shows that $\hat{d}(1) \sim O(\delta t^{-\frac{1}{2}})$ does not hold uniformly as required for the proof and the total hedging error may not always approach zero. In other words, the gamma $\frac{\partial^2 \hat{F}}{\partial S^2}$ can be chaotic when close to the maturity of the option since $N'(\hat{d}(1))$ cannot always be arbitrarily small. Leland [29] responded that the proposed method can still be used effectively for realistic trading costs and revision intervals as the average total errors are insignificantly different from zero.

### 4.4 Numerical Results

Zhao and Zemba [45, 46] investigate numerically whether the mean of the total hedging error is significantly different from zero. Their analysis caters for cases with and without transaction costs being paid when creating and liquidating the hedge portfolio. In this section the transaction costs at maturity are ignored as this is not accounted for in the Leland method. The histograms investigate the distribution of hedging errors with various levels of cost and hedging intervals. The research letters suggests that for small values of round trip cost $2\lambda$, the total hedging error is insignificantly different from zero. However when transaction costs are high, the hedging errors are significant, even when the transaction costs are ignored at inception and maturity. Interestingly the research letters also shows that the average of the hedging error tends closer to zero for higher strikes.

The following histograms are a realization of 10000 paths with the hedging error defined as the difference between the liquidated hedge portfolio and the call option value. The histograms show analogous results to Zhao and Zemba [46]. The following table contains the mean and standard deviation of the total hedging error. It is evident that for large costs the mean is different from zero, however ignoring liquidating costs at maturity, this discrepancy is trivial. Leland method remains an
effective method to hedge European options in the presence of proportional transaction costs. Comparisons to other strategies are performed in the last chapter.

<table>
<thead>
<tr>
<th>Transaction Cost</th>
<th>$\lambda = 0.03$</th>
<th>$\lambda = 0.02$</th>
<th>$\lambda = 0.01$</th>
<th>$\lambda = 0.005$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedging Interval $\delta t = 1/90$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.0520</td>
<td>-0.0225</td>
<td>-0.0050</td>
<td>-0.0023</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.9120</td>
<td>0.7883</td>
<td>0.6621</td>
<td>0.6001</td>
</tr>
<tr>
<td>Hedging Interval $\delta t = 1/360$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.0979</td>
<td>-0.0463</td>
<td>-0.0112</td>
<td>-0.0017</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.6179</td>
<td>0.5112</td>
<td>0.3961</td>
<td>0.3352</td>
</tr>
<tr>
<td>Hedging Interval $\delta t = 1/720$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.1135</td>
<td>-0.0543</td>
<td>-0.0137</td>
<td>-0.0025</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.5115</td>
<td>0.4168</td>
<td>0.3123</td>
<td>0.2545</td>
</tr>
</tbody>
</table>

**Tab. 4.1:** Total hedging error statistics for the Leland strategy. The parameters are: $S = 100$, $K = 100$, $T = 1$, $r = 0.05$ and $\sigma = 0.2$. 
Fig. 4.1: Total hedging error distributions for the Leland strategy. The parameters used are: $S = 100$, $K = 100$, $T = 1$, $r = 0.05$, $\sigma = 0.2$ and $\delta t = 1/360$. 
Chapter 5

Quadratic Hedging

5.1 Introduction

Quadratic hedging is a preference independent method of hedging contingent claims in an incomplete market. There are two quadratic hedging approaches: local and global risk minimisation. Local risk minimisation relaxes the self-financing constraint on the hedging portfolio and ensures that the payoff of the contingent claim is replicated exactly. Global risk minimisation (or mean-variance hedging) is a strategy which tries to reproduce the final payoff with smallest possible error, but insists on the self-financing constraint. One of the criticisms of quadratic hedging is that it equally penalises upside and downside risk. The advantages of this approach include mathematical tractability, optimal selection of the pricing measure and a hedging strategy that possesses relatively small replication errors with a reasonable initial price. The last claim is investigated fully in the final chapter. Mean-variance hedging is not presented here as its practical implementation remains unclear, while interested readers are directed to Schweizer [39, Section 4] for a rigorous overview.

This chapter looks at the local risk minimising strategy and provides results from Schweizer [38, 39], Lamberton, Pham and Schweizer [10] and Mercurio and Vorst [30].

5.2 Problem Formulation in Discrete Time

In this section a discrete time version of the local risk minimisation problem is defined. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space with filtration $\mathbb{F} = (\mathcal{F}_t)_{t=0,1,2,...,T}$ for some $T \in \mathbb{N}$. Leaving aside the issues regarding the choice of the numeraire, the stochastic process $\tilde{S} = (\tilde{S}_t)_{t=0,1,2,...,T}$ is a non-negative discounted stock process which is adapted to the filtration $\mathbb{F}$. There is also a riskless asset whose discounted price is 1 at all times. When dealing with quadratic hedging approaches a more stringent definition of a trading strategy than the one introduced in Chapter 1 is
5.2 Problem Formulation in Discrete Time

required.

**Definition 5.1.** A *trading strategy* \( \pi \) is a pair of processes \( \vartheta, \eta \) such that

- \( \vartheta = (\vartheta_t)_{t=0,...,T} \) is a square integrable predictable process\(^1\),
- \( \eta = (\eta_t)_{t=0,1,...,T} \) is adapted,

and the adapted value function defined by

\[
V_t(\pi) = \vartheta_t \tilde{S}_t + \eta_t \quad \text{for} \quad t = 0, 1, ..., T, \tag{5.1}
\]

is square integrable.

The square integrable condition is introduced to ensure that the risk process is well defined, while the predictability condition allows calculations of the value function at time step \( t \). The cost process for the strategy \( \pi \) is defined as

\[
C_t(\pi) = V_t(\pi) - \sum_{j=1}^{t} \vartheta_{j-1} \Delta \tilde{S}_j \quad \text{for} \quad t = 0, 1, ..., T, \tag{5.2}
\]

where \( \Delta U_t = U_t - U_{t-1} \) is the difference operator for any stochastic process \( U \). The cost process is interpreted as the deviation of the hedge portfolio from the value function up to time step \( t \) for some strategy \( \pi \). It is important to note that transaction costs are not taken into account here — the cost function with transaction costs is defined in the following section. The total risk process of a strategy \( \pi \) is defined as

\[
R_t(\pi) = \mathbb{E} \left[ (C_T(\pi) - C_t(\pi))^2 \mid \mathcal{F}_t \right] \quad \text{for} \quad t = 0, 1, ..., T, \tag{5.3}
\]

and is assumed to be square integrable. A formal definition of local risk minimising strategy is given.

**Definition 5.2.** Let \( \pi = (\vartheta, \eta) \) be a strategy and \( t \in \{0, 1, ..., T - 1\} \). A *local perturbation* of \( \pi \) at time step \( t \) is a strategy \( \pi^* = (\vartheta^*, \eta^*) \) with

\[
\vartheta^*_j = \vartheta_j \quad \text{for} \quad j \neq t \tag{5.4}
\]

and

\[
\eta^*_j = \eta_j \quad \text{for} \quad j \neq t. \tag{5.5}
\]

\(^1\) The predictability condition also implies that the value of \( \vartheta_t \) is \( \mathcal{F}_{t-1} \)-measurable.
\[ \pi^{QH} \text{ is called locally risk-minimising if } \]
\[ R_t(\pi^{QH}) \leq R_t(\pi^*) \quad \mathbb{P}-\text{a.s.} \quad (5.6) \]
for any time step \( t \in \{0, 1, ..., T-1\} \) and any local perturbation \( \pi^* \) of \( \pi^{QH} \) at time step \( t \).

The maturity time \( T \) is excluded in the definition of the perturbation strategy since the terminal values of \( \pi^{QH} \) and \( \pi^* \) are equal. In continuous time, the local risk minimising strategy is defined by minimising the continuous version of the risk process in (5.3). In discrete time, the risk can be minimised locally at each time step \( t \) by considering a strategy which minimises
\[ \mathbb{E} \left[ \left( C_{t+1}(\pi) - C_t(\pi) \right)^2 \mid \mathcal{F}_t \right] \]
instead of \( R_t(\pi) \). In fact, the two definitions are shown to be equivalent in Schweizer [38] and Lamberton, Pham and Schweizer [10] provided the underlying process is a general semimartingale. Intuitively, the risk process depends on the choice of the strategy via \( \vartheta_t, \vartheta_{t+1}, ..., \vartheta_T \) and \( \eta_t, \eta_{t+1}, ..., \eta_T \). However at time step \( t \), only the variables \( \vartheta_t \) and \( \eta_t \) are controlled, hence the minimisation takes place only with respect to these variables. Consider
\[ C_{t+1}(\pi) - C_t(\pi) = V_{t+1}(\pi) - V_t(\pi) - \vartheta_t(\tilde{S}_{t+1} - \tilde{S}_t) \]
\[ = \Delta V_{t+1}(\pi) - \vartheta_t \Delta \tilde{S}_{t+1}. \quad (5.7) \]

Using (5.7), the risk function for one time step can be written as:
\[ \mathcal{R}_t(\pi) = \mathbb{E} \left[ (C_{t+1}(\pi) - C_t(\pi))^2 \mid \mathcal{F}_t \right] = \mathbb{E} \left[ \left( \Delta V_{t+1}(\pi) - \vartheta_t \Delta \tilde{S}_{t+1} \right)^2 \mid \mathcal{F}_t \right] \]
\[ = \text{Var} \left[ V_{t+1}(\pi) - \vartheta_t \Delta \tilde{S}_{t+1} \mid \mathcal{F}_t \right] \]
\[ + \left( \mathbb{E} \left[ V_{t+1}(\pi) - \vartheta_t \Delta \tilde{S}_{t+1} \mid \mathcal{F}_t \right] - V_t(\pi) \right)^2, \quad (5.8) \]
since \( V_t(\pi) \) is \( \mathcal{F}_t \)-measurable. The \( \eta_t \) parameter appears only in the second term of (5.8), hence choosing \( \eta_t \) such that
\[ \mathbb{E} \left[ V_{t+1}(\pi) - \vartheta_t \Delta \tilde{S}_{t+1} \mid \mathcal{F}_t \right] = V_t(\pi), \quad (5.9) \]
makes the second term equal to zero. Thus minimising the contribution of this term. This also ensures the optimal strategy is mean self-financing, since the expected change in the cost process is zero,
\[ \mathbb{E} \left[ \Delta V_{t+1}(\pi) - \vartheta_t \Delta \tilde{S}_{t+1} \mid \mathcal{F}_t \right] = \mathbb{E} \left[ C_{t+1}(\pi) - C_t(\pi) \mid \mathcal{F}_t \right] = 0. \quad (5.10) \]
It follows from the mean self-financing property that the cost process under the local risk minimising strategy is a martingale, i.e.

$$E[C_{t+1}(\pi)|\mathcal{F}_t] = C_t(\pi).$$  \hfill (5.11)

Only the first term of (5.8) remains to be minimised with respect to the $\mathcal{F}_t$-measurable variable $\vartheta_t$. To achieve this Schweizer [38, Proposition 5, Chapter 1] proved that $C_t(\pi)\tilde{S}$ is a martingale when $\pi$ is a local risk minimising strategy. Using the definition of covariance and the fact that $C_t(\pi)$ and $\tilde{S}_t$ are martingales when $\pi$ is a risk minimising strategy, the martingale property can be written as

$$0 = E\left[C_{t+1}(\pi)\tilde{S}_{t+1} - C_t(\pi)\tilde{S}_t|\mathcal{F}_t\right] = E\left[C_{t+1}(\pi)\tilde{S}_{t+1}|\mathcal{F}_t\right] - E[C_{t+1}(\pi)|\mathcal{F}_t]E\left[\tilde{S}_{t+1}|\mathcal{F}_t\right] = \text{Cov}\left[C_{t+1}(\pi), \tilde{S}_{t+1}|\mathcal{F}_t\right].$$ \hfill (5.12)

Hence the term $\text{Var}\left[V_{t+1}(\pi) - \vartheta_t\Delta\tilde{S}_{t+1}|\mathcal{F}_t\right]$ is minimised if and only if the following property holds

$$\text{Cov}\left[V_{t+1}(\pi), \Delta\tilde{S}_{t+1}|\mathcal{F}_t\right] - \vartheta_t\text{Cov}\left[\Delta\tilde{S}_{t+1}, \Delta\tilde{S}_{t+1}|\mathcal{F}_t\right] = 0. \hfill (5.13)$$

Using the properties of covariance, (5.13) can be expanded as

$$\text{Cov}\left[V_{t+1}(\pi), \Delta\tilde{S}_{t+1}|\mathcal{F}_t\right] = \vartheta_t\text{Cov}\left[\Delta\tilde{S}_{t+1}, \Delta\tilde{S}_{t+1}|\mathcal{F}_t\right].$$ \hfill (5.14)

Consequently, the optimal stock holding $\vartheta_t^{QH}$ for a local risk minimising strategy $\pi^{QH}$ has the following form

$$\vartheta_t^{QH} = \frac{\text{Cov}\left[V_{t+1}(\pi), \Delta\tilde{S}_{t+1}|\mathcal{F}_t\right]}{\text{Var}\left[\Delta\tilde{S}_{t+1}|\mathcal{F}_t\right]}.$$ \hfill (5.15)

Using the Doob decomposition of the stochastic process $\tilde{S}$ into a martingale and a predictable process and the result in (5.13), Schweizer showed that, under the local risk minimising strategy, the martingale part of $\tilde{S}$ and the cost process $C_{t+1}(\pi)$ are strongly orthogonal\(^2\). Consequently, in discrete time a strategy is locally risk minimising if and only if

- The cost process $C(\pi)$ is a martingale, and

\(^2\) The martingales $X$ and $Y$ are said to be strongly orthogonal if the process $XY$ is also a martingale.
• The cost process $C(\pi)$ is orthogonal to the martingale part of $\tilde{S}$ under the Doob decomposition.

These properties are generalised for the continuous time case under specific conditions on $\tilde{S}$, but will not be explored further. For a complete account refer to Schweizer [38, 39].

5.3 Transaction Costs

In this section transaction costs are introduced into the local risk minimising strategy. The work closely follows Lamberton, Pham and Schweizer [10]. When transaction costs are incurred, different bid and ask prices are charged for the underlying asset. These prices are modeled using $\lambda \in [0, 1)$, with the bid and ask prices given by $(1 - \lambda)\tilde{S}_t$ and $(1 + \lambda)\tilde{S}_t$ respectively. The transaction cost parameter $\lambda$ makes the bid-ask spread symmetrical. The total outlay due to hedging and transaction cost at time step $t$ is

$$\eta_t - \eta_{t-1} + (\vartheta_t - \vartheta_{t-1})\tilde{S}_t(1 + \lambda \text{sgn}(\vartheta_t - \vartheta_{t-1}))$$

$$= V_t(\pi) - V_{t-1}(\pi) - \vartheta_{t-1} (\tilde{S}_t - \tilde{S}_{t-1}) + \lambda \tilde{S}_t |\vartheta_t - \vartheta_{t-1}|,$$

where $\text{sgn}$ is the signum function. In analogy to (5.2), the cost process is computed by considering the total outlay for all steps

$$C_t(\pi) = V_t(\pi) - \sum_{j=1}^{t} \vartheta_{j-1} \Delta \tilde{S}_j + \lambda \sum_{j=1}^{t} \tilde{S}_j |\Delta \vartheta_j|, \quad \text{for} \quad t = 0, 1, ..., T. \quad (5.17)$$

The definition of the risk process and local risk minimising strategy remains the same. The following theorem is analogous to the formulation in (5.8) with transaction costs.

**Theorem 5.3** (Properties of the local risk minimising strategy). A strategy $\pi = (\vartheta, \eta)$ is locally risk minimising (inclusive of transaction costs, with transaction cost parameter $\lambda$) if and only if it has the following properties:

1. $C_t(\pi)$ is a martingale.

2. For each $t \in \{0, 1, ..., T - 1\}$, $\vartheta_t$ minimises

$$\text{Var} \left[ V_{t+1}(\pi) - \vartheta_t^{*} \Delta \tilde{S}_{t+1} + \lambda \tilde{S}_t |\vartheta_{t+1} - \vartheta_t^{*}| |\mathcal{F}_t \right]$$

over all $\mathcal{F}_t$-measurable random variables $\vartheta_t^{*}$ such that $\vartheta_t^{*} \tilde{S}_{t+1}$ and $\vartheta_t^{*} \Delta \tilde{S}_{t+1}$ are square integrable.
5.3 Transaction Costs

Proof. The proof closely follows Lamberton, Pham and Schweizer [10, Proposition 2].

For the purpose of this proof, the time step \( t \in (0, \ldots, T - 1) \) is fixed and \( \pi^* = (\vartheta^*, \eta^*) \) is defined by setting \( \vartheta^* = \vartheta, \eta_j^* = \eta_j \) for \( j \neq t \) and

\[
\eta_t^* = \mathbb{E}[C_T(\pi) - C_t(\pi)|\mathcal{F}_t] + \eta_t.
\] (5.19)

Clearly \( \vartheta^* \) is adapted and \( \pi^* \) is a local perturbation of \( \pi \) at time step \( t \). Suppose now that \( \pi \) is a locally risk minimising strategy, Properties 1 and 2 need to be shown. From (5.19) and the definition of \( \pi^* \), it is possible to write

\[
V_t(\pi^*) = V_t(\pi) + \mathbb{E}[C_T(\pi) - C_t(\pi)|\mathcal{F}_t],
\]

\[
C_T(\pi^*) - C_t(\pi^*) = C_T(\pi) - C_t(\pi) + \mathbb{E}[C_T(\pi) - C_t(\pi)|\mathcal{F}_t].
\] (5.20)

Squaring both sides, taking expectations and using the definition of the variance, the following inequality is obtained

\[
R_t(\pi^*) = \mathbb{V} \text{ar}[C_T(\pi) - C_t(\pi)|\mathcal{F}_t] \leq \mathbb{E}[(C_T(\pi) - C_t(\pi))^2|\mathcal{F}_t] = R_t(\pi).
\] (5.21)

However, since \( \pi \) is a locally risk minimising strategy with \( R_t(\pi) \leq R_t(\pi^*) \), (5.21) should be an equality. This is only possible if \( \mathbb{E}[C_T(\pi) - C_t(\pi)|\mathcal{F}_t] = 0, \mathbb{P} \)-a.s., which shows that \( C_t(\pi) \) is a martingale. For the remainder of the proof, the expression (5.19) is relaxed and \( \pi^* \) remains a local perturbation of \( \pi \) at time step \( t \). To show the second property, it is necessary to compare the following expressions

\[
\mathbb{E}[R_{t+1}(\pi)|\mathcal{F}_t] + \mathbb{V} \text{ar}[C_{t+1}(\pi) - C_t(\pi)|\mathcal{F}_t] = \mathbb{E}[C_T(\pi)^2 - C_{t+1}(\pi)^2|\mathcal{F}_t] + \mathbb{E}[C_{t+1}(\pi)^2 - 2C_{t+1}(\pi)C_t(\pi) + C_t(\pi)^2|\mathcal{F}_t]
\]

\[
= \mathbb{E}[C_T(\pi)^2 - 2C_T(\pi)C_t(\pi) + C_t(\pi)^2|\mathcal{F}_t] = R_t(\pi) \quad \mathbb{P}\text{-a.s.},
\] (5.22)

and similarly

\[
\mathbb{E}[R_{t+1}(\pi^*)|\mathcal{F}_t] + \mathbb{E}[(C_{t+1}(\pi^*) - C_t(\pi^*))^2|\mathcal{F}_t] = \mathbb{E}[C_T(\pi)^2 - C_{t+1}(\pi^*)^2|\mathcal{F}_t] + \mathbb{E}[C_{t+1}(\pi^*)^2 - 2C_{t+1}(\pi)C_t(\pi^*) + C_t(\pi^*)^2|\mathcal{F}_t]
\]

\[
= \mathbb{E}[C_T(\pi)^2 - 2C_T(\pi)C_t(\pi^*) + C_t(\pi^*)^2|\mathcal{F}_t] = R_t(\pi^*) \quad \mathbb{P}\text{-a.s.}
\] (5.23)

The expressions (5.22) and (5.23) are derived using the martingale property of \( C_t(\pi) \).
The first expression divides the risk process of a local risk minimising strategy \( \pi \) into the expectation of the risk process one time step ahead for a strategy \( \pi \) and the variance of the change in the cost process over one time step for a strategy \( \pi \).

The second expression divides the risk process of a perturbation strategy \( \pi^* \) into the expectation of the risk process one time step ahead for a local risk minimising strategy \( \pi \) and an expectation of the squared difference of the cost process over one time step for a strategy \( \pi^* \). The choice of a stock holding \( \vartheta_t \) only influences the second term and as a result \( \vartheta_t \) minimises

\[
\text{Var} \left[ C_{t+1}(\pi^*) - C_t(\pi^*) \right] = \text{Var} \left[ V_{t+1}(\pi) - \vartheta_t \Delta S_{t+1} + \lambda S_t | \vartheta_{t+1} - \vartheta_t | \right].
\]

(5.24)

Using (5.22) and (5.23), the following inequality holds for any \( F_t \)-measurable choice of \( \vartheta^*_t \) and \( \eta^*_t \)

\[
E \left[ (C_{t+1}(\pi^*) - C_t(\pi^*))^2 | F_t \right] \geq \text{Var} \left[ C_{t+1}(\pi^*) - C_t(\pi^*) \right].
\]

(5.25)

In particular, the second property is obtained by fixing \( \vartheta^*_t \) together with choosing \( \eta^*_t \) such that

\[
E \left[ (C_{t+1}(\pi^*) - C_t(\pi^*))^2 | F_t \right] = 0
\]

and using the inequality (5.25).

Consider that property 1 and 2 hold, it remains to be shown that strategy \( \pi \) is locally risk minimising. Using (5.23), the following inequalities are derived

\[
R_t(\pi^*) = E \left[ R_{t+1}(\pi) | F_t \right] + E \left[ (C_{t+1}(\pi^*) - C_t(\pi^*))^2 | F_t \right]
\]

\[
\geq E \left[ R_{t+1}(\pi) | F_t \right] + \text{Var} \left[ C_{t+1}(\pi^*) - C_t(\pi^*) | F_t \right]
\]

\[
\geq E \left[ R_{t+1}(\pi) | F_t \right] + \text{Var} \left[ C_{t+1}(\pi) - C_t(\pi) | F_t \right]
\]

\[
= R_t(\pi),
\]

(5.26)

where the second inequality uses the second property and the last equality is from (5.23). This shows that strategy \( \pi \) is locally risk minimising.

This result implies that the local risk minimising strategy, inclusive of transaction costs, also minimises

\[
R_t(\vartheta, \eta) = E \left[ (C_{t+1}(\pi) - C_t(\pi))^2 | F_t \right]
\]

at each step, which is similar to the general formulation without transaction costs. This objective function with transaction costs can be written as

\[
R_t(\vartheta, \eta) = E \left[ (V_{t+1}(\pi) - V_t(\pi) - \vartheta_t(S_{t+1} - S_t) + \lambda S_{t+1} | \vartheta_{t+1} - \vartheta_t |)^2 | F_t \right].
\]

(5.27)

In a practical implementation, the risk minimising strategy can be found by using the properties of Theorem 5.3. The stock holding \( \vartheta_t \) is found by minimising the variance in (5.18) and then \( \eta_t \) is found by imposing the martingale property. Lamberton,
Pham and Schweizer [10] also prove the integrability, robustness and existence of such strategies. Interested readers are referred to their paper.

### 5.4 Total Hedging Error

Mercurio and Vorst present a useful proposition which describes the expected value and variance of the total hedging error under a local risk minimising strategy. The total hedging error is the sum of all the one time step hedging errors. The expression for the total hedging error with transaction costs can be written as

\[
\text{THE}(\vartheta, \eta, V_0) = \Psi(S_T) - V_0 - \sum_{t=1}^{T} \vartheta_{t-1}(S_t - S_{t-1}) + \lambda \sum_{t=1}^{T} S_t |\vartheta_t - \vartheta_{t-1}| , \quad (5.28)
\]

where \(\Psi(S_T)\) is a payoff of the contingent claim at maturity and \(V_0\) is the initial wealth used at inception of the replicating portfolio.

**Proposition 5.4.** The total hedging error for a local risk minimising strategy \(\pi^{QH}\), with respect to the initial wealth \(V_0\), has an expectation

\[
E \left[ \text{THE}(\vartheta^{QH}, \eta^{QH}, V_0) \right] = \vartheta_0^{QH} S_0 + \eta_0^{QH} - V_0 \quad (5.29)
\]

and variance

\[
\text{Var} \left[ \text{THE}(\vartheta^{QH}, \eta^{QH}, V_0) \right] = E \left[ \sum_{t=1}^{T} R_{t-1}(\pi^{QH}) \right] . \quad (5.30)
\]

**Proof.** The local risk minimising strategy is given by \(\pi^{QH} = (\vartheta^{QH}, \eta^{QH})\), \(V_T(\pi^{QH}) = \Psi(S_T)\) and \(V_0(\pi^{QH}) = \vartheta_0^{QH} S_0 + \eta_0^{QH}\). The expectation is determined as follows

\[
E \left[ \text{THE}(\vartheta^{QH}, \eta^{QH}, V_0) \right] = E \left[ \sum_{t=1}^{T} \left( C_t(\pi^{QH}) - C_{t-1}(\pi^{QH}) \right) \right] + \vartheta_0^{QH} S_0 + \eta_0^{QH} - V_0
\]

\[
= \vartheta_0^{QH} S_0 + \eta_0^{QH} - V_0
\]

since the cost process for a local risk minimising strategy is a martingale.
The variance is determined as follows

\[\text{Var} \left[ \text{THE}(\theta_Q^H, 1, V_0) \right] = E \left[ \sum_{t=1}^{T} \left( C_t(\pi_Q^H) - C_{t-1}(\pi_Q^H) \right)^2 \right] = \sum_{t=1}^{T} E \left[ R_{t-1}(\pi_Q^H) \right].\]

The proposition shows that the total hedging error under the local risk minimising strategy has a mean of zero if the initial wealth equals the value function at inception. It also gives an expression for the variance of the distribution for the total hedging error.

### 5.5 Numerical Implementation

So far the results have been presented in a general framework which has not specified the stochastic process \( \tilde{S} \) in detail. This section describes the numerical method of Mercurio and Vorst [30] which computes the local risk minimising strategy for a geometric Brownian motion as the underlying process.

The underlying process is formulated in discrete time by considering time steps \( t \in [0, 1, \ldots, N] \), where \( N \in \mathbb{N} \) (notice the variable \( T \) is swapped with \( N \) so as not to be confused with the maturity of the option). Hedging takes place at these time steps and the hedging times are given by \( \{0 = i_0 < i_1 < \cdots < i_t < \cdots < i_{N-1} \} \) with a constant interval \( \Delta t \) between them. The maturity time is omitted from the set and denoted by \( i_N = T \). Through the use of a binomial tree, the underlying process can be discretised further by choosing parameter \( M \) which determines the number of discretisation steps between two hedging times. The discretisation step is determined by \( \delta t = \Delta t / M \). The total number of discretisation steps is denoted by \( TS \) for the underlying process and determined by \( N \cdot M \). Figure 5.1 demonstrates the discretisation.

The aforementioned binomial tree has the following parameters for a geometric Brownian motion lattice

\[\hat{p} = \frac{e^{\mu \delta t} - w_d}{w_u - w_d}, \quad w_u = e^{\sigma \sqrt{\delta t}}, \quad w_d = e^{-\sigma \sqrt{\delta t}} = \frac{1}{w_u},\]

(5.31)

where \( \hat{p} \) and \( 1 - \hat{p} \) are the probabilities of an up move and a down move over one time step respectively, \( w_u \) and \( w_d \) are the factors by which the underlying asset moves up and down respectively and \( \delta t \) is the discretisation step. This parametrisation was
Fig. 5.1: The discretisation scheme is shown with parameters $N = 3$ and $M = 4$. The grey lines show the discretisation between the hedging times which is present throughout the tree.

first provided by Cox, Ross and Rubinstein [22] and is very tractable due to the recombinant property.

The extra discretisation between hedging times improves the approximation of the underlying process. The following multinomial model explains the price process evolution between one stock node at time $i_t$ and $M + 1$ stock nodes at time $i_{t+1}$. Let $j = 0, 1, ..., M$ be the variable that identifies each node at time $i_{t+1}$.

\[
S_{t+1} = \begin{cases} 
S_t w_u^M & \text{with probability } \hat{p}^M \\
\vdots & \vdots \\
S_t w_u^{M-2j} & \text{with probability } ^{M}j \hat{p}^{M-j}(1-\hat{p})^j \\
\vdots & \vdots \\
S_t w_u^{-M} & \text{with probability } (1-\hat{p})^M
\end{cases}
\quad \text{for } j = 0; \quad j = 1, \ldots, M - 1; \quad j = M.
\]

(5.32)

The probabilities in (5.32) apply to the future stock price at time $i_{t+1}$ given the current node.

Having defined the price process between inception and maturity with a binomial tree and the price process between the hedging times using a multinomial model, the terminal values of the strategy $\pi^{QH}$ can be determined for each stock price node.
For a long option these values are
\[ \vartheta_{QH} = \alpha 1_{\{\alpha(S_N - K) > 0\}} \]

\[ \mathcal{X}(S_N) = V_N(\pi_{QH}) = (\alpha(S_N - K))^+ \] \hspace{1cm} (5.33)

and for a short option
\[ \vartheta_{QH} = -\alpha 1_{\{\alpha(S_N - K) > 0\}} \]

\[ \mathcal{X}(S_N) = V_N(\pi_{QH}) = -(\alpha(S_N - K))^+ \] \hspace{1cm} (5.34)

where \( \alpha = \pm 1 \) indicates a call or a put option respectively. The following optimisation problem is solved for each stock node at each hedging time to find the local risk minimising strategy

\[ \min_{(\vartheta_{QH}, V_t) \in \mathbb{R}^2} \mathcal{R}_\ell(\vartheta, \eta) \] \hspace{1cm} (5.35)

subject to the given values of \( \vartheta_{QH} \) and \( V_T \) at maturity. The objective function with transaction costs is given in (5.27). The expectation is applied using the multinomial tree probabilities in (5.32) for each stock price node. The optimisation problems are solved backward in time starting at time \( i_{N-1} \) and the optimal values are then used recursively for the next set of optimisation problems.

## 5.6 Modified Algorithm

The proposed numerical procedure is computationally expensive especially for large values of \( N \) due to the large number of optimisation routines that are needed to be performed. For this reason, an algorithm which determines \( \vartheta_\ell \) and \( V_\ell \) for each node is provided in this section. Notably, the proposed method can be used in conjunction with any type of European option with a monotonic payoff.

Using the proposed multinomial model, the objective function in (5.27) can be written fully for some node at time \( i_\ell \) as

\[ \mathcal{R}_\ell(\vartheta, \eta) = \sum_{j=0}^{M} \binom{\mathcal{H}}{j} \hat{\vartheta}^{\mathcal{H}-j}(1 - \hat{\vartheta})^j \left( V_{\ell+1}(\pi) - V_{\ell}(\pi) - \vartheta_{\ell}(S_{\ell+1}^{(j)} - S_{\ell}) \right. \]

\[ + \lambda S_{\ell+1}^{(j)} \left( \vartheta_{\ell+1}^{(j)} - \vartheta_{\ell} \right) \right|^2 \] \hspace{1cm} (5.36)

where the superscript \( \cdot^{(j)}_{\ell+1} \) denotes the value of the ordered \( j \)-th node at time \( i_{\ell+1} \). The derivative of an absolute value function is a signum function which makes minimisation of the objective function problematic when done analytically. For
this reason, the function \(|\vartheta_{t+1}^{j} - \vartheta_{t}|\) can be considered for all possible values of \(\vartheta_{t}\). Consider the following \(M + 1\) stock holdings at some hedging time \(t_{t+1}\)

\[
\vartheta_{t+1}^{0} \geq \vartheta_{t+1}^{1} \geq \cdots \geq \vartheta_{t+1}^{(M-1)} \geq \vartheta_{t+1}^{M}.
\]  

(5.37)

The stock holdings are sorted in descending order. Using the sorted stock holdings it is possible to consider all possible relationships for the value of \(\vartheta_{t}\). Three possible cases are considered.

When

\[
\vartheta_{t} \geq \vartheta_{t+1}^{0} \geq \vartheta_{t+1}^{1} \geq \cdots \geq \vartheta_{t+1}^{(M-1)} \geq \vartheta_{t+1}^{M},
\]  

(5.38)

it is possible to evaluate the absolute value functions in (5.36) as follows

\[
|\vartheta_{t+1}^{j} - \vartheta_{t}| = -\left(\vartheta_{t+1}^{j} - \vartheta_{t}\right) \text{ since } \vartheta_{t+1}^{j} \leq \vartheta_{t} \text{ for } j = 0, 1, ..., M.
\]

When

\[
\vartheta_{t+1}^{0} \geq \vartheta_{t+1}^{1} \geq \cdots \geq \vartheta_{t+1}^{(n)} \geq \vartheta_{t} \geq \vartheta_{t+1}^{(n+1)} \geq \cdots \geq \vartheta_{t+1}^{(M-1)} \geq \vartheta_{t+1}^{M},
\]  

(5.39)

it is possible to find evaluate the absolute value functions in (5.36) as follows

\[
|\vartheta_{t+1}^{j} - \vartheta_{t}| = \left(\vartheta_{t+1}^{j} - \vartheta_{t}\right) \text{ since } \vartheta_{t+1}^{j} \geq \vartheta_{t} \text{ for } j = 0, 1, ..., n.
\]

and

\[
|\vartheta_{t+1}^{j} - \vartheta_{t}| = -\left(\vartheta_{t+1}^{j} - \vartheta_{t}\right) \text{ since } \vartheta_{t+1}^{j} \leq \vartheta_{t} \text{ for } j = n + 1, ..., M.
\]

Finally, when

\[
\vartheta_{t+1}^{0} \geq \vartheta_{t+1}^{1} \geq \cdots \geq \vartheta_{t+1}^{(M-1)} \geq \vartheta_{t+1}^{M} \geq \vartheta_{t}. 
\]  

(5.40)

it is possible to find evaluate the absolute value functions in (5.36) as follows

\[
|\vartheta_{t+1}^{j} - \vartheta_{t}| = \left(\vartheta_{t+1}^{j} - \vartheta_{t}\right) \text{ since } \vartheta_{t+1}^{j} \geq \vartheta_{t} \text{ for } j = 0, 1, ..., M.
\]

It is evident that \(M + 2\) cases can be considered. For each case, it is possible to substitute the determined expressions into the objective function (5.36). This allows one to take the derivatives in the following equations

\[
\frac{\partial \mathcal{R}_{t}(\vartheta, \eta)}{\partial \vartheta_{t}} = 0 \quad \text{and} \quad \frac{\partial \mathcal{R}_{t}(\vartheta, \eta)}{\partial V_{t}} = 0,
\]

(5.41)

and transform the results into a system of linear equations of the form \(Ax = b\).
which can easily be solved for $\vartheta_t$ and $V_t$.

To save on computation time a slightly different approach is taken to compute the correct values for $\vartheta_t$ and $V_t$. The first step is to check all the possible boundary values from (5.37). This is done in $M + 1$ cases by setting $\vartheta_t$ to each known value of $\vartheta^{(j)}_{t+1}$ and solving for $V_t$ in $\frac{\partial R_t(\vartheta, \eta)}{\partial V_t} = 0$. While evaluating $\frac{\partial R_t(\vartheta, \eta)}{\partial V_t}$, it is necessary to consider previously mentioned cases in (5.38), (5.39) and (5.40) and write down the function correctly for each boundary. The objective function (5.36) is then evaluated for each derived boundary value of $V_t$ and the boundary with the lowest value for the objective function is chosen. By choosing the smallest available value of the objective function, the optimal value for $\vartheta_t$ could lie either on the left or the right of that value. This directly corresponds to two possible cases which are considered when solving for the optimal values of $\vartheta_t$ and $V_t$.

As previously mentioned, instead of checking $M + 2$ possible cases, the equations in (5.41) only need to be solved twice to the left and right of the boundary with the smallest value. Having derived the optimal values for each case, it is necessary to check and disregard those values that do not correspond to the correct initial conditions in (5.38), (5.39) and (5.40). This is done since the optimisation, which needs to be performed is constrained, but solving a system of the form $Ax = b$ performs an unconstrained optimisation. Accordingly, it is possible for the optimal solution to lie outside the constrained region, but this will result in disregarding that solution. Once all the solutions are checked, the objective function is evaluated for all boundary values and the remaining optimal values. The optimal values of $\vartheta_t$ and $V_t$ are chosen which correspond to the smallest value of the objective function so that the strategy $\pi$ which minimises the risk function $R_t(\vartheta, \eta)$ is chosen. This procedure is repeated for each node at every hedging time.

This approach, while appearing to be computationally challenging, is much more efficient than using general purpose optimisation routines. It also takes care of all possible cases when minimising the risk function as long as the option is European and the payoff is monotonically increasing or decreasing. This modified algorithm made it possible to compute the results for large values of $M$. These are presented in the following section and in the last chapter of the dissertation.

### 5.7 Numerical Results

In this section the modified algorithm is implemented and the influence of the parameters is examined. Figures 5.2 and 5.3 examine the accuracy on the price and the hedge ratio with various tree sizes, while the number of hedging times was kept constant. The levels of transaction cost $\lambda$ were varied between 1%, 2% and 3%.
5.7 Numerical Results

The option price converges relatively quickly while the hedge ratio varies even with large tree sizes. Although the scale is relatively small for the hedging error in Figure 5.3, larger trees could be used in practice to achieve more accurate results. For the purpose of performing a large Monte Carlo simulation in the last chapter, a tree size between 60 to 180 and not below 20 was chosen when the time to maturity decreases.

Figure 5.4 examines the effect of different hedging frequencies on prices. It is evident that the prices increase with an increase in hedging times for all levels of transaction costs. This is due to the hedger paying transaction costs each time hedging is performed. For the same reason the curve is the steepest for the largest level of transaction cost. Although the figure for the variance of the hedging error is not presented in this section, the variance is expected to drop when performing more frequent hedging. This is shown in the final chapter.

The quadratic hedging algorithm gives a preference free method of hedging contingent claims in the incomplete market. Furthermore the numerical implementation of the method has proved to be stable which is indicative of the robustness of this approach.
5.7 Numerical Results

Fig. 5.3: The hedge ratios $\vartheta_0^{QH}$ with various levels of proportional transaction cost are plotted against the tree size. The parameters are: $S_0 = 100$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $\mu = 0.08$, $r = 0.04$ and $M = 90$.

Fig. 5.4: The prices $V_0(\pi^{QH})$ with various levels of proportional transaction cost are plotted against a number of hedging times. The parameters are: $S_0 = 100$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $\mu = 0.08$ and $r = 0.04$. 
Chapter 6

Introduction to Stochastic Optimal Control

6.1 Introduction

Control problems give a mathematical description of how to act optimally to maximise future reward. The controlled process may have a stochastic element in which case it is a stochastic optimal control problem. Central to stochastic optimal control is the notion of the dynamic programming principle which was originated by R. Bellman [3, 4] in 1950s. This principle was used to develop the Hamilton Jacobi Bellman equation (HJB) which if solved provides optimal control rules. This chapter sets up a type of a problem which can be solved using stochastic optimal control. A heuristic derivation of a general HJB equation, for the stochastic case, is presented which is largely based on Björk [5, Chapter 19]. The HJB equation is a non-linear second order partial differential equation, however solving it analytically proves to be difficult as the solutions need to be smooth, but this is not always the case even for simple problems. Viscosity solutions were introduced to address this issue by Crandall and Lions [9]. Furthermore, the Markov chain approximation method which numerically solves a type of stochastic optimal control problem is discussed. It uses the HJB equation for its intuitive meaning and breaks down the problem using approximations. The aim of this chapter is to give preliminary material required for the following chapter on utility indifference pricing and hedging of options. Note that no reference to financial problems is made in this chapter. All variables and functions are defined generally with no financial context.

6.2 Problem Formulation

This section presents a general class of stochastic optimal control problems with corresponding definitions and notation. Consider functions $\mu(t, x, u)$ and $\sigma(t, x, u)$ of the form

$$
\mu: \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n \quad \text{and} \quad \sigma: \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^{n \times d}.
$$
6.2 Problem Formulation

The \( n \)-dimensional state process \( X \) is written as

\[
\begin{align*}
    dX_t &= \mu(t, X_t, u_t) \, dt + \sigma(t, X_t, u_t) \, dW_t, \\
    X_0 &= x_0.
\end{align*}
\] (6.1)

The initial point is given by \( x_0 \in \mathbb{R}^n \) and the \( d \)-dimensional Weiner process is given by \( W \). The state process is governed by the \( k \)-dimensional control process \( u \) and \( u_t \) denotes its value at time \( t \). The control process \( u \) is adapted to the filtration generated by the Brownian motion of the state process. In particular, a feedback control law \( u(t, x_t) \) is required which is a function of time \( t \) and the current value of the multidimensional state process \( x_t \). Generally the control law needs to satisfy some control constraints which are modelled by taking a subset \( \mathcal{C} \subseteq \mathbb{R}^k \) and requiring that \( u_t \in \mathcal{C} \) for each \( t \).

In particular a special class of problem is considered, for which the control law exists and is admissible. A control law \( u \) is called admissible if the control constraints are satisfied (i.e. \( u(t, x) \in \mathcal{C} \) for all \( t \in \mathbb{R}_+ \) and all \( x \in \mathbb{R}^k \)) and for any given initial point \( x_t \), the stochastic differential equation

\[
\begin{align*}
    dX_s &= \mu(t, X_s, u(t, X_s)) \, dt + \sigma(t, X_s, u(t, X_s)) \, dW_t, \\
    X_t &= x_t.
\end{align*}
\] (6.3)

has a unique solution. The class of admissible control laws is denoted by \( \mathcal{A} \). In this setup it is assumed that the admissible control law exists. The existence assumption is a non-trivial one, since the considered optimal control law’s behaviour can be rapid. This can lead to irregular dynamics of the state process which results in a complicated existence problem. The existence problem is outside the scope of this dissertation and is not dealt with. To define the objective function for the proposed control problem, the following pair of functions is considered

\[
G : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R} \quad \text{and} \quad \Psi : \mathbb{R}^n \rightarrow \mathbb{R}.
\]

The function \( G \) is a running objective function and the function \( \Psi \) is a terminal objective function.\(^1\) The value function \( J_0 \) is

\[
J_0 : \mathcal{A} \rightarrow \mathbb{R},
\]

\(^1\) In Chapter 7 on utility indifference pricing, the running objective function is not used since consumption problems are not dealt with.
and defined by
\[
\mathcal{J}_0(u) = \mathbb{E}\left[ \int_0^T G(t, X^u_t, u_t) \, dt + \Psi(X^u_T) \right],
\]
(6.5)
where \(X^u\) is the solution to the state process given the initial point \(x_0\). The formal problem is to maximise the value function over all control laws. The optimal value is written as
\[
\hat{J}_0 = \sup_{u \in \mathcal{A}} \mathcal{J}_0(u).
\]
(6.6)
Hence an admissible control law \(\hat{u}\) is said to be optimal if
\[
\mathcal{J}_0(\hat{u}) = \hat{J}_0.
\]
The problem formulation is given for a general stochastic optimal control problem, and an optimal law may not always exist. The main objective is to find the optimal law provided it exists or derive some properties about the optimal law.

### 6.3 Derivation of Hamilton Jacobi Bellman equation

The Hamilton Jacobi Bellman (HJB) equation is central to the field of stochastic optimal control. It allows to represent the problem in the form of a non-linear second order PDE and finding the solution to the PDE is equivalent to obtaining an optimal feedback control law. This section is largely based on the heuristic derivation presented in Bjöörk [5, Chapter 19]. Bjöörk mentions that the derivation makes strong regularity assumptions and disregards a number of technical problems. Interested readers are directed to more rigorous books on this topic such as Fleming and Rishel [14]. The following assumptions are made for the derivation:

- There exists an optimal control law \(\hat{u}\).
- The optimal value function \(\hat{J}\) is regular, in the sense that it is continuous and twice differentiable.
- Limiting procedures are justified in the following argument.

The problem formulation above assumed an initial time 0, however to set up the derivation for the HJB equation a fixed time point \(t\) and fixed state point \(x_t\) are considered. The value function is then written as
\[
\mathcal{J}_t(t, x, u) = \mathbb{E}\left[ \int_t^T G(s, X^u_s, u_s) \, ds + \Psi(X^u_T) \right],
\]
6.3 Derivation of Hamilton Jacobi Bellman equation

Given previously mentioned dynamics and the control process constraints. The optimal value function at time $t$ is similarly given as

$$V(t, x) = \sup_{u \in A} J_t(t, x, u).$$

Note that the notation $V$ does not correspond to the value process used in the previous chapters, but corresponds to the notation used by Björk. The notation only applies for this chapter and reverts back in subsequent chapters. To derive the PDE, a small increment in time is considered and defined by $\delta t$ such that $t + \delta t < T$ and the control law $u^*$ is defined by

$$u^*(s, y) = \begin{cases} u(s, y), & (s, y) \in [t, t + \delta t] \times \mathbb{R}^n, \\ \hat{u}(s, y), & (s, y) \in (t + \delta t, T] \times \mathbb{R}^n. \end{cases}$$

The control law $u^*$ is split into an arbitrarily law $u$ for a short time interval and an optimal control law $\hat{u}$ for the rest of the interval. The expected utility functions are examined for two control laws, mainly $\hat{u}$ and $u^*$. Under the optimal law, the expected utility function is given by the optimal value function $J_t(t, x, \hat{u}) = V(t, x)$. Under the newly proposed law $u^*$, the expected utility for interval $[t, t + \delta t]$ is given by

$$\mathbb{E} \left[ \int_t^{t+\delta t} G(s, X^u_s, u_t) \, ds \right]$$

and for the remaining period it is

$$\mathbb{E} \left[ V(t + \delta t, X^\hat{u}_{t+\delta t}) \right].$$

Comparing the two strategies and using the fact that the first control law is optimal, the following inequality holds

$$V(t, x) \geq \mathbb{E} \left[ \int_t^{t+\delta t} G(s, X^u_s, u_s) \, ds + V(t + \delta t, X^u_{t+\delta t}) \right]. \quad (6.7)$$

Using the second assumption of a regular optimal value function, the following equation is derived using the multidimensional Itô formula

$$V(t + \delta t, X^u_{t+\delta t}) = V(t, x) + \int_t^{t+\delta t} \left( \frac{\delta V}{\delta t}(s, X^u_s) + A^u V(s, X^u_s) \right) \, ds$$

$$+ \int_t^{t+\delta t} \nabla_x V(s, X^u_s) \sigma^u \, dW_s, \quad (6.8)$$
where
\[ A^u = \sum_{i=1}^{n} \mu^u_i(t, x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \sigma^u_i(t, x) \sigma^u_j(t, x) \frac{\partial^2}{\partial x_i \partial x_j} \]
and
\[ \nabla_x = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}. \]

Taking the expectation and assuming integrability conditions are satisfied, the stochastic integral disappears. Substituting the expression (6.8) into the inequality (6.7), gives
\[ \mathbb{E} \left[ \int_t^{t+\delta t} \left( G(s, X^u_s, u_s) + \frac{\partial V}{\partial t}(s, X^u_s) + A^u V(s, X^u_s) \right) ds \right] \leq 0. \tag{6.9} \]
Dividing the previous expression by \( \delta t \), and letting \( \delta t \) tend to zero, the following expression is derived
\[ G(t, x, u) + \frac{\partial V}{\partial t}(t, x) + A^u V(t, x) \leq 0, \tag{6.10} \]
by the fundamental theorem of calculus. Note that (6.7) becomes an equality if and only if the control law \( u \) is an optimal control law. Using this and the definition of an optimal value function, a supremum is taken over all control laws that fulfill the constraints.
\[ \frac{\partial V}{\partial t}(t, x) + \sup_{u \in \mathcal{U}} (G(t, x, u) + A^u V(t, x)) = 0. \tag{6.11} \]
The derived PDE is known as a multidimensional Hamilton Jacobi Bellman equation.

By definition of the value function, the terminal boundary is
\[ V(T, x) = \Psi(x) \text{ for all } x \in \mathbb{R}^k. \tag{6.12} \]
The PDE holds for all points \((t, x) \in (0, T) \times \mathbb{R}^n\), since \((t, x)\) was an arbitrarily fixed point in the beginning of the discussion. The assumptions made for the derivation are specific to the general case described. However, in the academic literature these assumptions are presented in terms of initial data \( \mu, \sigma, G \) and \( \Psi \), coupled with greater technical complexity. Interestingly, the previous argument showed that the HJB equation is satisfied when \( V \) is the optimal value function and \( \hat{u} \) is the optimal control law. However the HJB equation is also a sufficient condition for the optimal control problem, meaning that the optimal value function follows the solution of the HJB equation. This result is known as the verification theorem and is now presented.
Theorem 6.1 (Verification theorem). Suppose that we have two functions $H(t, x)$ and $g(t, x)$, such that

- $H$ is sufficiently integrable and solves the HJB equation
  \[
  \frac{\partial H}{\partial t}(t, x) + \sup_{u \in \mathcal{C}} (G(t, x, u) + A^n H(t, x)) = 0, \quad \forall (t, x) \in (0, T) \times \mathbb{R}^n
  \]
  \[
  H(T, x) = \Psi(x) \quad \text{for all } x \in \mathbb{R}^k, \quad \forall x \in \mathbb{R}^n.
  \]
- The function $g$ is an admissible control law.
- For each fixed $(t, x)$, the supremum in the expression
  \[
  \sup_{u \in \mathcal{C}} (G(t, x, u) + A^n H(t, x))
  \]
  is obtained by the choice $u = g(t, x)$.

Then the following hold:

- The optimal value function $V$ to the control problem is given by
  \[
  V(t, x) = H(t, x).
  \]
- There exists an optimal control law $\hat{u}$ and it is determined by $\hat{u}(t, x) = g(t, x)$.

Proof. See Theorem 19.6, in Björk [5, Chapter 19].

Generally solving the HJB equation analytically or numerically is difficult as the solutions are not smooth enough to describe the derivatives involved in the equation. Viscosity solutions were introduced by Crandall and Lions which allow to obtain non-smooth solutions for the PDE by replacing the derivatives in the PDE with super-differentials and sub-differentials. Importantly, viscosity solutions maintain the uniqueness of the solution under some mild conditions and this leads to a powerful technique for solving partial differential equations including the HJB equation. It is beyond the scope of this dissertation to present an encompassing overview of viscosity solutions. The following section presents an alternative method of solving stochastic optimal control problems which will be used later when performing utility indifference pricing of options.

6.4 Markov Chain Approximation

The most powerful numerical method for stochastic control problems was developed by Kushner [25], namely a Markov chain approximation method. It is a numerical
technique which uses approximations to solve the proposed problem. In this section, an overview based Kushner [26] and Kushner and Dupuis [27] is presented.

The Markov chain approximation method breaks down the control problem into a simpler problem for which the computation is feasible. The computed values can be shown to converge to the original solution under some ‘local properties’. One of the major advantages of this method is that it does not require solving the HJB equation or dealing with the regularity issues seen in the previous section. However the HJB equation is still used implicitly to assist in constructing the relevant dynamic programming equations and algorithms. Another advantage of a Markov chain approximation is that the method is intuitive since the optimal value function for the approximation chain is the optimal value function for the proposed problem. In other words, the method does not rely on the analytical expressions (solutions of PDEs) for the functions that we wish to compute. It is rather based on approximating the controlled process by a simpler process for which the optimal control can be easily evaluated. Furthermore, the convergence results are determined using a probabilistic approach and are much simpler to show, then for example to show the existence of the viscosity solution. For this reason, it remains the standard numerical approach for stochastic optimal control problems.

The idea behind the method is to replace the original control problem with a suitable approximation using a locally consistent chain. A chain is a random process which undergoes transition between finite states. This allows easier calculations of the original problem. Specifically, the Markov property implies that the future states only depend on the current state and not on the previous history of the process. The approximating chain is a Markov chain which is chosen so that certain statistical properties are the same as the original process. In particular the mean and mean square change per step under any control need to be the same. One would expect the value from the approximating optimal value function over all controls to be close to the value of the optimal value function given by \((6.6)\) if the approximating controlled process is close to the original state process \((6.1)\) and the associated value function is close to the original value function \((6.5)\). The approximating controlled process is given by a finite state controlled Markov chain \(\xi_n^h\) where \(h > 0\) is the scalar approximation parameter and \(n < \infty\) indicates the discretised state space. The transition probability of a state process moving from state \(x_n\) to state \(x_{n+1}\) are defined by \(p^h(x_n, x_{n+1}|u)\) and the discretised control law is denoted by \(u_n^h\) with the value \(u\). To define the local consistency conditions for the Markov chain, it is necessary to define an interpolation interval \(\delta t_n^h = \delta t^h(\xi_n^h, u_n^h)\) and \(\delta t^h(x, u) > 0\) for values of state \(x\) and controlled law \(u\). To ensure a Markov chain approximates the state process in between the interpolated points, the interpolation interval must be
small and positive. The interpolation interval properties are written as
\[
\sup_{x,u} \delta t_n^h(x,u) \rightarrow 0 \quad \text{as} \quad h \rightarrow 0,
\]
\[
\inf_{x,u} \delta t_n^h(x,u) > 0 \quad \text{for each} \quad h > 0.
\]

The following local consistency conditions ensure that the Markov chain obeys the statistical properties of the original state process:
\[
\mathbb{E} \left( \delta \xi_{n+1}^h \mid \xi_n^h = x, u_n^h = u \right) = \mu^h(x,u) \delta t^h(x,u) + o(\delta t^h(x,u)),
\]
\[
\mathbb{V} \text{ar} \left( \delta \xi_{n+1}^h \mid \xi_n^h = x, u_n^h = u \right) = \sigma(x,u) \sigma(x,u)^t \delta t^h(x,u) + o(\delta t^h(x,u)),
\]
\[
\sup_n |\xi_{n+1}^h - \xi_n^h| \rightarrow 0 \quad \text{as} \quad h \rightarrow 0,
\]
(6.13)

where \( \delta \xi_{n+1}^h = \xi_{n+1}^h - \xi_n^h \). Notice (6.13) also defines functions \( \mu^h(\cdot) \) and \( \sigma^h(\cdot) \). The chain has the same properties as the state process in the sense that on the interval \([0, \delta t]\), with \( X_0 = x \) and \( u = u \) the state process has the following properties
\[
\mathbb{E} (X_{\delta t} - X_0 \mid F_0) = \mu(x,u) \delta t + o(\delta t),
\]
(6.14)
\[
\mathbb{V} \text{ar} (X_{\delta t} - X_0 \mid F_0) = \sigma(x,u) \sigma(x,u)^t \delta t + o(\delta t).
\]
(6.15)

The consistency properties given by (6.13) are all that is required for the approximating chain. Furthermore the control process \( u_n^h \) is called admissible if the chain has a Markov property:
\[
P(\xi_{n+1}^h = x \mid \xi_i^h, u_i^h \text{ for } i \leq n) = p(\xi_n^h, x \mid u_n^h).
\]
(6.16)

At this stage \( \xi_n^h \) is a discrete time parameter process. In order to approximate the continuous state process (6.1), a continuous time interpolation is needed for the Markov chain. There are two useful interpolation methods, the first being a continuous parameter Markov process which is usually used when constructing proofs and performing convergence analysis as it allows some simplifications. The second method uses interpolation intervals which are deterministic functions of the current state and control value. The second interpolation is the subject of this section.

The continuous parameter interpolations \( \xi^h(\cdot) \) and \( u^h(\cdot) \) are defined by
\[
\xi^h(t) = \xi_n^h, \quad u^h(t) = u_n^h, \quad t \in [t_n^h, t_{n+1}^h),
\]
(6.17)
where the interpolated time is given by

\[ t_h^n = \sum_{i=0}^{n-1} \delta t_h^i. \]  

(6.18)

Figure 6.1 shows the construction of the interpolated process \( \xi^h(t) \). It shows that the process is piecewise constant. A natural time scale of \( \delta t_h^i(x, u) \) can be used as a constant interval, although this could be restrictive. A more flexible approach of variable interpolation intervals can guarantee faster convergence in some cases. The interpolation intervals are obtained when specifying the transition probabilities for the Markov chain.

At this stage the sufficient conditions to approximate the state process for the Markov chain are defined in (6.13) and the Markov chain construction was examined. It is now possible to consider constructing approximations which solve the stochastic control problem. The proposed control problem falls in the class of finite time problems. In this class of problems, there are two types of methods for constructing such approximations, namely an explicit and implicit method. In fact, the Markov chain approximation method is closely related to finite difference methods for solving elliptic and parabolic PDEs. To illustrate this point and show how the Markov chain approximation method works, an example is provided which uses an explicit method to solve a one dimensional version of the problem proposed at the beginning of this chapter.

Consider a finite time problem where the value function is given by (6.5) and the state process (6.2) is driven by a single Weiner process. The HJB equation which satisfies the optimal value function of the example can be written as

\[
\frac{\partial V}{\partial t}(t, x) + \sup_{u \in \mathcal{C}} \left( G(t, x, u) \frac{\partial V}{\partial x}(t, x) + \frac{1}{2} (\sigma^u(t, x))^2 \frac{\partial^2 V}{\partial x^2}(t, x) \right) = 0.
\]

(6.19)
To approximate the derivatives the following finite difference forms are used:

\[
\begin{align*}
V_t(t, x) & \rightarrow \frac{V(t + \delta t, x) - V(t, x)}{\delta t} \\
V_x(t, x) & \rightarrow \frac{V(t + \delta t, x + h) - V(t + \delta t, x)}{h} \\
V_{xx}(t, x) & \rightarrow \frac{V(t + \delta t, x + h) + V(t + \delta t, x - h) - 2V(t + \delta t, x)}{h^2}
\end{align*}
\]  

These explicit forms are used if \( \mu^u(t, x) \geq 0 \). Substituting these expressions into (6.19) and multiplying all terms by \( \delta t \) the following finite difference equation is derived

\[
\tilde{V}(n\delta t, x) = \tilde{V}(n\delta t + \delta t, x) \left( 1 - \left( \sigma^u(n\delta t, x) \right)^2 \frac{\delta t}{h^2} - \mu^u(n\delta t, x) \frac{\delta t}{h} \right) \\
+ \tilde{V}(n\delta t + \delta t, x + h) \left( \frac{\left( \sigma^u(n\delta t, x) \right)^2 \delta t}{2h^2} + \mu^u(n\delta t, x) \frac{\delta t}{h} \right) \\
+ \tilde{V}(n\delta t + \delta t, x - h) \left( \frac{\left( \sigma^u(n\delta t, x) \right)^2 \delta t}{2h^2} \right) \\
+ G(n\delta t, x, u) \delta t.
\]  

(6.21)

where \( \tilde{V}(t, x) \) denotes the solution to the finite difference equation, \( x \) is an integral multiple of \( h \) and \( n\delta t < T \). The terminal boundary condition is

\[
\tilde{V}(T, x) = \Psi(x).
\]  

(6.22)

The method is explicit since (6.21) allows the solution at time \( n\delta t \) to be derived recursively from the solution at time \( n\delta t + \delta t \). This is due to the spatial derivatives being approximated at time \( n\delta t + \delta t \). Noticeably, the sum of \( V \) coefficients is unity. In fact, if the coefficients are all positive, the terms can be considered to be the transition probabilities and (6.21) can be re-written as

\[
\tilde{V}(n\delta t, x_n) = \sup_{u \in C} \left( \sum_{x_{n+1}} p^{h, \delta t}(x_n, x_{n+1} | u) \tilde{V}(n\delta t + \delta t, x_{n+1}) + G(n\delta t, x_n, u) \delta t \right).
\]  

(6.23)

Consider a Markov chain \( \xi^h_n \) that has locally consistent conditions and define the piecewise constant continuous parameter interpolation \( \xi^h(t) \) on the interval \( [n\delta t, n\delta t + \delta t] \). If \( \mu^u(t, x) < 0 \) then \( V_x(t, x) \rightarrow \frac{V(t + \delta t, x + h) - V(t + \delta t, x - h)}{h} \) is used. This ensures the finite difference equation has an interpretation in terms of a Markov chain.
The solution to (6.23) can be written as

\[
\tilde{V}(t, x) = \sup_{u \in \mathcal{U}} \mathbb{E} \left( \sum_n G(t, \xi^h_{i+1}, u) \delta t + \Psi(\xi^h_N) \right)
= \sup_{u \in \mathcal{U}} \mathbb{E} \left( \int_t^T G(s, \xi^h(s), u) \delta s + \Psi(\xi^h(T)) \right). \tag{6.24}
\]

The expression in (6.24) is an approximation to the optimal value process if the process \(\xi^h(t)\) is an approximation to the state process. In fact, (6.24) is a dynamic programming equation which concludes the Markov chain approximation method. The discrete equation (6.23) is easily solved by starting at time \(T\) with terminal condition (6.22) and then performing backward iterations all the way to time 0.

The example shows the link between using a finite difference method of solving the HJB equation and the Markov chain approximation method. As shown the Markov chain approximation is a powerful technique which can be applied for a large class of stochastic control problems. Without going into further detail, there also exists an implicit method of solving finite time problems, the time variable is discretised as a state variable and a different dynamic programming equation is derived. Interested readers are directed to a book by Kushner and Dupuis [27] for a rigorous overview on numerical methods for solving stochastic control problems.
Chapter 7

Utility Indifference Pricing with Transaction Costs

7.1 Introduction

This chapter introduces a global-in-time model which uses the utility indifference framework to price and hedge European options with proportional transaction costs. Previously, the strategies under examination had a fixed hedging interval. An option trader needs to choose this interval carefully as there is a trade-off between incurring transaction costs or a hedge slippage. Frequent rebalancing leads to high transaction costs while infrequent rebalancing leads to potentially large hedging errors and associate risk. The global in time mechanism chooses the trade off continuously by hedging only when it is optimal to do so. The utility indifference framework introduces a hedge rule which governs if the hedge portfolio needs to be rebalanced at each moment in time. This section introduces preliminary concepts on utility functions and discusses the characteristics of the approach taken from Henderson and Hobson [19]. It then follows work by Davis, Panas and Zariphopoulou [36] which is based on earlier work of Hodges and Neuberger [20]. Davis, Panas and Zariphopoulou derived the partial differential equations which give the hedging rule and the corresponding option price in the presence of transaction costs. Finally an algorithm proposed by Monoyios [34] is presented which numerically computes the solution to the partial differential equations and the option price, using a Markov chain approximation method.

7.2 Utility Functions

Utility functions help to relate the two important concepts of risk and wealth mathematically. This is needed to optimally choose between incurring transaction costs or taking on large hedging errors. A utility function, denoted by $U(x)$, quantitatively describes an investor’s satisfaction derived from wealth $x$ under an assumption
7.2 Utility Functions

of a rational investor. The utility function has the following properties:

- **Non-Satiation**: A utility function is monotonically increasing. This property ensures an investor always prefers more wealth to less.

- **Continuity**: A utility function is twice continuously differentiable. Changing investor wealth marginally should not lead to a sudden jump in the utility.

- **Law of diminishing returns**: The first derivative is monotonically decreasing. An increase in wealth diminishes the utility of a further increase in wealth.

- **Positive risk-aversion**: A utility function is strictly concave. An investor prefers a certain amount of wealth to a random outcome with the same expected value.

Practically each investor can have a different risk preference, and there are many classes of utility functions to choose from. Furthermore, each utility function is parametrised to fit an investor’s risk profile. In this section, a particular class of utility functions is examined namely HARA utilities. The utility function chosen to describe a rational investor using utility indifference pricing is selected from this class.

A useful quantity to describe utility functions is an absolute risk aversion defined by

\[ \text{ARA}(x) = -\frac{U_{xx}}{U_x}. \]  

(7.1)

where \( U_x \) denotes a derivative of \( U \) with respect to the variable \( x \). Similarly \( U_{xx} \) denotes a double derivative with respect to the variable \( x \). Intuitively, the higher the curvature of the utility function, the higher its risk aversion. This coefficient allows one to compare the intensity of risk aversion for various utility functions and sometimes used in the classification of utility functions. See Arrow [2] and Pratt [37] for a rigorous discussion of this coefficient. Furthermore a utility function belongs to the HARA (hyperbolic absolute risk aversion) class if the absolute risk aversion coefficient is of the following form

\[ \text{ARA}(x) = \frac{1}{A + Bx} \quad x \in I_D. \]  

(7.2)

It is important to define the interval \( I_D \) on which the utility function is defined and is consistent with the rational properties. In particular a utility function can be defined on a positive real line or over the whole real line. The denominator in (7.2) needs to be strictly positive to ensure a positive risk aversion (the fourth property of rational utility functions). The following conditions are placed on constants \( A \) and \( B \) to bound the interval \( I_D \) on which the utility function is defined. In a case where
$B = 0$ and $A > 0$ the interval $I_D = \mathbb{R}$. Alternatively, when $B > 0$ and $A$ takes some value, then the utility function is defined on $I_D = (-A/B, \infty)$, since $U(x) = \infty$ for $x < -A/B$. In this section an exponential utility function

$$U(x) = -e^{-\gamma x}$$

(7.3)

is selected, where $\gamma$ is the risk aversion parameter of the exponential utility function. The exponential utility function belongs to the HARA class since

$$\text{ARA}(x) = -\frac{-\gamma^2 e^{-\gamma x}}{\gamma e^{-\gamma x}} = \gamma,$$

(7.4)

with parameters $A = \frac{1}{\gamma}$ and $B = 0$. Clearly the interval $I_D$ is defined on the whole real line. The parameter $\gamma$ is also shown to be an absolute constant risk aversion parameter in (7.4). This means an investor’s attitude towards risk is the same irrelevant of how much initial wealth he has. This property is also known as invariant under any translation in wealth and can be written mathematically as

$$U(x + k) = f(k)U(x) + g(k),$$

(7.5)

where $k$ is some constant, $f(k)$ and $g(k)$ are some functions of $k$ and independent of $x$. For the exponential utility function this property holds as follows

$$U(x + k) = -e^{-\gamma(x+k)} = -e^{-\gamma k}(-e^{-\gamma x}) = -e^{-\gamma k}U(x).$$

(7.6)

As shown, $g(k) = 0$ which allows to separate the total investor’s wealth and the amount of money in the bank account when using an exponential utility function. This property is the reason this utility function is commonly used to reduce the dimensionality of the utility indifference pricing problem.

### 7.3 Utility Indifference Pricing

This section is largely based on work of Henderson and Hobson [19] and builds on the previous properties of utility functions. Consider a market which consists of three instruments: a risky asset, a riskless bank account and an option on the risky asset. The utility indifference price for the option is the price at which the investor is indifferent (in expected utility terms) between investing in the market with the long (short) option and paying (receiving) a suitable amount of compensation or investing in the market without the option. Mathematically this can be written as
follows:
\[
U^{(n)}(t, S, \eta - n \cdot p, \vartheta) = U^{(0)}(t, S, \eta, \vartheta),
\]
(7.7)
or
\[
U^{(-n)}(t, S, \eta + n \cdot p, \vartheta) = U^{(0)}(t, S, \eta, \vartheta),
\]
(7.8)
where
\[
U^{(n)}(t, S, \eta, \vartheta) = \sup_{\pi \in A_T} \mathbb{E}[U(V_T + n \cdot \Psi(S_T))].
\]
(7.9)
An investor uses an admissible trading strategy denoted by \( \pi \) which starts in a state \((t, S, \eta, \vartheta)\) and consists of holding amount \(\eta\) in cash and \(\vartheta\) shares of a stock \(S\). The supremum is taken over all terminal wealths \(V_T = \eta_T + \vartheta_T S_T\) which can be generated by the strategy \(\pi\). Note that the notation \(V\) no longer implies the optimal value function from Chapter 6 and corresponds to the wealth of the investor. The initial price of the option is given by \(p\), while the payoff at maturity of the option is described by \(\Psi(S_T)\). The number of options is denoted by \(n\).

These equations suggest that we need to solve two stochastic optimal control problems. In the first case, the investor does not buy the option and invests the initial wealth in the riskless and risky asset. This is the classic Merton problem [31]. The second problem involves the investor holding \(n\) options long (or short) for which the investor gets paid (or compensated). Both problems can be formulated as stochastic optimal control problems and described by the HJB equation. Alternatively, the problem can be formulated as a dual problem which involves selecting a price using expectations under a set of equivalent martingale measures. The dual problem will not be dealt with here.

The following are the characteristics of the utility indifference price methodology:

- **Non-linear Pricing**: As a result of concavity of the utility function, utility indifference pricing gives non-linear prices (i.e. the price of two options does not equal twice the price of a single option). An investor is expected to be compensated by a reduction in the price of options when taking additional non-diversifiable risk. This property is in contrast to other methodologies in the incomplete markets.

- **Bid-Ask Prices**: Utility indifference framework produces different prices for purchasing and selling an option.

- **Recovery of the complete market price**: When the market is made complete (e.g. when transaction costs are zero), the utility indifference option price produces the Black Scholes price.
7.4 Transaction Cost Formulation

- Monotonicity: A greater payoff between two options is reflected in the price of those options. Let \( p_i \) be the utility indifference price with one unit payoff of \( \Psi(S_T)^i \). If \( \Psi(S_T)^1 > \Psi(S_T)^2 \) then \( p_1 > p_2 \).

- Concavity: Let \( p_\lambda \) be the utility indifference buy price for the claim \( \lambda \Psi(S_T)^1 + (1 - \lambda) \Psi(S_T)^2 \) where \( \lambda \in [0, 1] \). Then

\[
p_\lambda \geq \lambda p_1 + (1 - \lambda)p_2.
\]

If considering the utility indifference sell prices then \( p_\lambda \) is convex.

7.4 Transaction Cost Formulation

Davis, Panas and Zariphopoulou [36] considered the problem of pricing European options with proportional transaction costs using utility indifference pricing. The option price presented here is a modification of the definition introduced by Hodges and Neuberger [20]. The derivation of the partial differential equation which governs the solution to the problem is presented in this section.

Consider a set of admissible trading strategies given by \( A(V_0) \) when starting with an initial wealth \( V_0 \). A trading strategy \( \pi \in A(V_0) \) describes an amount of money held in a bank account at time \( t \) denoted by \( \eta_t \) and a number of shares held at time \( t \) denoted by \( \vartheta_t \). The transaction costs are introduced by writing the liquidated cash value of the portfolio as

\[
c(\vartheta_t, S_t) = \begin{cases} 
(1 + \lambda)\vartheta_t S_t & \text{if } \vartheta_t < 0 \\
(1 - \lambda)\vartheta_t S_t & \text{if } \vartheta_t \geq 0,
\end{cases}
\]

where \( t \in [0, T] \) and \( \lambda \) is the proportional costs levelled when buying and selling the underlying asset. Suppose the option writer forms a portfolio to replicate the option and liquidates it at time \( T \). In the case of a call option, if \( S_T \leq K \) the option value is zero and the cash value of the portfolio is \( \eta_T + c(\vartheta_T, S_T) \) at time \( T \). Alternatively, if the option is in the money, \( S_T > K \), the writer receives \( K \) and pays the value of the underlying asset \( S_T \), hence the cash value of the portfolio is \( \eta_T + K + c(\vartheta_T - 1, S_T) \). Given the option writer’s utility function \( U(x) \) and taking into account the transaction costs the valuation function for \( n \) options is written as

\[
\Psi^{(n)}(t, S, \eta - n \cdot p, \vartheta) = \sup_{\pi^{(n)}} \mathbb{E}\left[U(\eta_T + 1_{\{S_T \leq K\}}[c(\vartheta_T, S_T)]
+ 1_{\{S_T > K\}}[n \cdot K + c(\vartheta_T - n, S_T)])\right] \quad (7.10)
\]
and a valuation function without the option is written as

$$
\mathcal{V}^{(0)}(t, S, \eta, \vartheta) = \sup_{\pi^{(0)}} \mathbb{E} \left[ U(\eta_T + c(\vartheta_T, S_T)) \right],
$$

(7.11)

where $1_{\{\cdot\}}$ is the indicator function for an event. Note the different valuation functions, where the subscript denotes the number of options the investor trades in the market. By definition, strategy $\pi^{(1)}$ deals with an investor trading optimally in the market with one option while strategy $\pi^{(0)}$ trades optimally in the market without the option. The definition of the hedging strategy which hedges one option is introduced.

**Definition 7.1.** The utility indifference hedging strategy $\pi^U$ for one option traded at price $p$ at time $t$ is defined by the holdings $\vartheta^U_k$ and $\eta^U_k$ at time $k \in [t, T]$ which satisfy

$$
\vartheta^U_k = \vartheta^{(1)}_k - \vartheta^{(0)}_k,
$$

(7.12)

$$
\eta^U_k = \eta^{(1)}_k - \eta^{(0)}_k,
$$

(7.13)

where $\vartheta^{(n)}$ and $\eta^{(n)}$ denotes the stock holding and bank holding of a trading strategy $\pi^{(n)}$ which invests in $n$ options and the market. The utility indifference hedging strategy for one option can be written as $\pi^U = \pi^{(1)} - \pi^{(0)}$.

We are now able to formulate the system as a control problem for both valuation functions with and without the option. Let $L(t)$ and $M(t)$ be the cumulative number of shares bought or sold at time $t$ over the life of the option given by

$$
L(t) = \int_0^t l(\tau) \, d\tau \quad \text{and} \quad M(t) = \int_0^t m(\tau) \, d\tau,
$$

(7.14)

where $l$ and $m$ are uniformly bounded by some finite amount $s$. The differential equations which govern the amount of cash, the number of shares and the stock price are

$$
d\vartheta_t = dL(t) - dM(t),
$$

$$
d\eta_t = r\eta_t \, dt - (1 + \lambda_b)S_t \, dL(t) + (1 - \lambda_s)S_t \, dM(t),
$$

$$
dS_t = \mu S_t \, dt + \sigma S_t \, dW_t,
$$

where $dL(t)$ and $dM(t)$ is a pair of right-continuous, non-decreasing processes which are adapted to the filtration of the stock process. Substituting (7.14) and re-
7.4 Transaction Cost Formulation

arranging terms we can write the equations as follows:

\[ d\vartheta_t = [l - m] \, dt, \]
\[ d\eta_t = [r\eta - l(1 + \lambda_b)S + m(1 - \lambda_s)S] \, dt, \]
\[ dS_t = \mu S \, dt + \sigma S \, dW. \]

Note a short hand notation is used and \( l, m \) are still functions of time \( t \). The system is described by two partial differential equations and one stochastic differential equation. A multi-variate HJB equation (6.11) is used to derive the control equations. The equations are derived for two cases: one where the trader invests in the market and one where the trader purchases the option and invests in the market. The two cases are denoted by the superscript \( j \in (0, 1) \). The control equations are written as

\[ -\mathcal{U}^{(j)} = \max_{0 \leq l, m \leq s} \left( (r\eta - l(1 + \lambda)S + m(1 - \lambda)S)\mathcal{U}^{(j)} + (l - m)\mathcal{U}^{(j)} \right. + (\mu S)\mathcal{U}^{(j)} + \frac{1}{2}(\sigma^2 S^2)\mathcal{U}^{(j)} \right), \quad (7.15) \]

where the subscript denotes a partial derivative with respect to that variable. To derive the optimal solution for the HJB equation we re-write (7.15) as

\[ \max_{0 \leq l, m \leq s} \left( (f_1)_l - (f_2)_m \right) + \mathcal{U}^{(j)} + (r\eta)\mathcal{U}^{(j)} + (\mu S)\mathcal{U}^{(j)} + \frac{1}{2}(\sigma^2 S^2)\mathcal{U}^{(j)} = 0 \quad (7.16) \]

where

\[ f_1 = \mathcal{U}^{(j)} - (1 + \lambda_b)S\mathcal{U}^{(j)}, \quad f_2 = \mathcal{U}^{(j)} - (1 - \lambda_s)S\mathcal{U}^{(j)}. \quad (7.17) \]

Examining (7.16) more closely we can obtain the optimal solution by the following three cases:

\[ f_1 \geq 0 \quad f_2 > 0 \quad \Rightarrow \quad l = s, \quad m = 0, \]
\[ f_1 < 0 \quad f_2 \leq 0 \quad \Rightarrow \quad l = 0, \quad m = s, \]
\[ f_1 \leq 0 \quad f_2 \geq 0 \quad \Rightarrow \quad l = 0, \quad m = 0. \]

All other combinations are inadmissible as a result of the valuation functions increasing in \( \vartheta \) and \( \eta \). It is conjectured that the state space region is divided into three distinct regions: the buy region, the sell region and the no transaction region. In those regions an investor’s optimal trading strategy is to buy stock, to sell stock and not to trade respectively. When the transaction region is reached with proportional costs, the investor needs to transact to the closest boundary of the no transaction region. Zakamouline [44] investigated the case with fixed and proportional costs in
which an investor transacts to special transact boundaries. In the proportional costs case, the transact boundaries are the boundaries between the no transaction region and the buy and sell regions which are denoted by \( \vartheta(b) \) and \( \vartheta(s) \) respectively.

In the buy region, the value function remains constant along the path of the state dictated by the optimal strategy and satisfies

\[
U(j)(t, S, \eta, \vartheta) = U(j)(t, S, \eta - S(1 + \lambda_b)\delta L, \vartheta + \delta L), \tag{7.18}
\]

where \( \delta L \) is the number of shares bought in order to take the investor’s holding to the boundary \( \vartheta(b) \). Allowing \( \delta L \downarrow 0 \), (7.18) becomes

\[
U_{\vartheta}(j) - (1 + \lambda_b)S\Upsilon(\eta) = 0. \tag{7.19}
\]

Similarly in the sell region, the value function satisfies

\[
U(j)(t, S, \eta, \vartheta) = U(j)(t, S, \eta + S(1 - \lambda_s)\delta M, \vartheta - \delta M), \tag{7.20}
\]

where \( \delta M \) is a number of shares sold in order to take the investor’s holding to the boundary \( \vartheta(s) \). Allowing \( \delta M \downarrow 0 \), (7.18) becomes

\[
U_{\vartheta}(j) - (1 - \lambda_s)S\Upsilon(\eta) = 0. \tag{7.21}
\]

Equations (7.19) and (7.21) govern the valuation functions \( U(j)(t, S, \eta, \vartheta) \) in the buy and sell regions and at the boundaries \( \vartheta(b) \) and \( \vartheta(s) \). In the no transaction region, it is sub-optimal to have any trades and the value function satisfies

\[
U(j)(t, S, \eta, \vartheta) \geq U(j)(t, S, \eta - S(1 + \lambda_b)\delta L, \vartheta + \delta L), \tag{7.22}
\]

and

\[
U(j)(t, S, \eta, \vartheta) \geq U(j)(t, S, \eta + S(1 - \lambda_s)\delta M, \vartheta - \delta M). \tag{7.23}
\]

The value function is evaluated using Bellman optimality principle which states that regardless of the decisions taken previously, the remaining decisions need to be optimal to acquire the optimal solution. This means that the optimal decisions can be traced back from maturity by taking an expectation over the time \( t + \delta t \)

\[
U(j)(t, S, \eta, \vartheta) = \mathbb{E}\left[U(j)(t + \delta t, S + \delta S, \eta + \delta \eta, \vartheta)\right]. \tag{7.24}
\]

In the no transaction region \( m = 0 \) and \( l = 0 \) since no trading takes places. Using
equation (7.16), the partial differential equation in the no transaction region is

\[ \mathcal{U}_t^{(j)} + (r\eta)\mathcal{U}_\eta^{(j)} + (\mu S)\mathcal{U}_S^{(j)} + \frac{1}{2}(\sigma^2 S^2)\mathcal{U}_{SS}^{(j)} = 0. \]  

(7.25)

It is sub-optimal to buy and sell at the same time as this will incur unnecessary transaction costs. To act optimally investor must act based on which region the state space is in. The complete, free boundary, non-linear PDE is written as:

\[
\max_{\pi} \left( \mathcal{U}_0^{(j)} - (1 + \lambda_b)S\mathcal{U}_S^{(j)} , -(\mathcal{U}_0^{(j)} - (1 - \lambda_s)S\mathcal{U}_S^{(j)}) , \right.

\[
\mathcal{U}_t^{(j)} + (rB)\mathcal{U}_B^{(j)} + (\mu S)\mathcal{U}_S^{(j)} + \frac{1}{2}(\sigma^2 S^2)\mathcal{U}_{SS}^{(j)} = 0. \]  

(7.26)

An analytical closed form solution for (7.26) is elusive and the valuation function must be solved numerically. The solution can be obtained using a Markov chain approximation method and evaluating the value function once the no transaction region and the boundaries \( \vartheta_b \) and \( \vartheta_s \) are evaluated using equations (7.18), (7.20) and (7.24).

In this setting the utility indifference option price \( p(t,S) \) is determined by the following equation

\[ p(t,S) = V_t^{(1)} - V_t^{(0)}, \]  

(7.27)

where

\[ V_t^{(j)} = \inf \left( V_t : \mathcal{U}^{(j)} \geq 0 \right) . \]  

(7.28)

The price of the option is the difference between the smallest wealth amount needed to ensure positive expected utility, when investing in the market with an option and the smallest wealth amount needed to ensure positive expected utility, when investing only in the market. The intuition for the derivation is the same as for the hedging strategy, the investor is indifferent between having an option and investing or just investing in the market. Furthermore Davis, Panas and Zariphopoulou [36, Theorem 1] showed that this definition of the option price reduces to the Black Scholes price in certain cases.

### 7.5 Transaction Cost under Exponential Utility

The valuation function \( \mathcal{U}^{(j)}(t,S,\eta,\vartheta) \) in the partial differential equation (7.26) is in four dimensions, namely time \( t \), asset price \( S \), amount of money held in cash \( \eta \) and the number of shares \( \vartheta \). By introducing an exponential utility function (7.3), an investor’s optimal trading strategy becomes independent of the amount \( \eta \) in the
7.5 Transaction Cost under Exponential Utility

This is due to the aforementioned constant risk aversion property as seen in (7.6). As a result the dimensionality of the problem can be reduced as follows. Define a new convex, non-increasing, continuous function

\[ Q^{(j)}(t, S, \vartheta) = \mathbb{U}^{(j)}(t, S, 0, \vartheta). \] (7.29)

By the constant risk aversion property (7.6) of the exponential utility function

\[ \mathbb{U}^{(j)}(t, S, \eta, \vartheta) = Q^{(j)}(t, S, \vartheta) \exp(-\gamma \eta e^{r(T-t)}). \] (7.30)

The partial differential equation (7.26) is transformed for the new valuation function \( Q^{(j)}(t, S, \vartheta) \) and written as

\[
\min_{\pi} \left[ Q^{(j)} \frac{\vartheta + \gamma (1 + \lambda_b)}{B(t)} S Q^{(j)}, - (Q^{(j)} - \frac{\gamma (1 - \lambda_s)}{B(t)} S Q^{(j)})) ,
Q_t^{(j)} + (\mu S) Q_S^{(j)} + \frac{1}{2} (\sigma^2 S^2) Q_{SS}^{(j)} \right] = 0,
\] (7.31)

where \( B(t) = e^{r(T-t)} \). Furthermore, (7.27) which defines the price of the option is now written as

\[ p(t, S) = \frac{\delta(T, t)}{\gamma} \log \left( \frac{Q^{(1)}(t, S, 0)}{Q^{(0)}(t, S, 0)} \right). \] (7.32)

The investor’s optimal trading strategy remains the same for each region. The optimal values for \( \delta L \) and \( \delta M \) are \( \delta L^* \) and \( \delta M^* \) respectively and satisfy

\[
\vartheta + \delta L^* = \vartheta_{(b)}(t, S) \quad \text{and} \quad \delta M^* = 0 \quad \text{if} \quad \vartheta < \vartheta_{(b)}(t, S),
\]
\[ \delta L^* = 0 \quad \text{and} \quad \delta M^* = 0 \quad \text{if} \quad \vartheta_{(b)}(t, S) \leq \vartheta \leq \vartheta_{(s)}(t, S), \]
\[ \vartheta - \delta M^* = \vartheta_{(s)}(t, S) \quad \text{and} \quad \delta L^* = 0 \quad \text{if} \quad \vartheta > \vartheta_{(s)}(t, S). \]

Equations (7.18), (7.20) and (7.24) can be written for the reduced valuation function in each region. In the buy region, the reduced valuation function follows

\[ Q^{(j)}(t, S, \vartheta) = Q^{(j)}(t, S, \vartheta_{(b)}(t, S)) \exp(\gamma S(1 + \lambda_b)(\vartheta_{(b)}(t, S) - \vartheta)e^{r(T-t)}) \]
for \( \vartheta < \vartheta_{(b)}(t, S). \) (7.33)

In the sell region, the reduced valuation function follows

\[ Q^{(j)}(t, S, \vartheta) = Q^{(j)}(t, S, \vartheta_{(s)}(t, S)) \exp(-\gamma S(1 - \lambda_s)(\vartheta - \vartheta_{(s)}(t, S))e^{r(T-t)}) \]
for \( \vartheta > \vartheta_{(s)}(t, S). \) (7.34)
Finally in the no transaction region, the reduced valuation function follows

\[ Q^{(j)}(t, S, \vartheta) = \mathbb{E} \left[ Q^{(j)}(t + \delta t, S + \delta S, \vartheta) \right] \quad \text{for} \quad \vartheta_{(b)}(t, S) \leq \vartheta \leq \vartheta_{(a)}(t, S). \quad (7.35) \]

It is clear that using the exponential utility function allows the reduction of the dimensionality of the problem by one. Having transformed the valuation function, it is possible to write down the dynamic programming equations in the following section.

### 7.6 Discrete Dynamic Programming Equations

The partial differential equations are discretised to form dynamic programming equations which are solved using a Markov chain approximation method. The discrete dynamic programming equation for the PDE in (7.26) is written as

\[
\tilde{U}^{(j)}(t, S, B, \vartheta) = \max_{\delta L, \delta M} \left[ \mathbb{E} \left[ \tilde{U}(t + \delta t, wS_t, (\eta - S_t(1 + \lambda_1)\delta L)R, \vartheta_t + \delta L) \right] \right. \\
\left. \mathbb{E} \left[ \tilde{U}(t + \delta t, wS_t, R\eta, \vartheta_t) \right] \right. \\
\left. \mathbb{E} \left[ \tilde{U}(t + \delta t, wS_t, R(\eta + S_t(1 - \lambda_2)\delta M, \vartheta_t - \delta M) \right] \right],
\]

where \( R = e^{r\delta t} \). Similarly the reduced form partial differential equation in (7.31) becomes

\[
\tilde{Q}^{(j)}(t, S, \vartheta) = \max_{\delta L, \delta M} \left[ \mathbb{E} \left[ \tilde{Q}(t + \delta t, wS_t, \vartheta_t + \delta L) \right] e^{\gamma S_t(1 + \lambda_1)\delta LB(t)} \right. \\
\left. \mathbb{E} \left[ \tilde{Q}(t + \delta t, wS_t, \vartheta_t) \right] \right. \\
\left. \mathbb{E} \left[ \tilde{Q}(t + \delta t, wS_t, \vartheta_t - \delta M) \right] e^{-\gamma S_t(1 - \lambda_2)\delta MB(t)} \right],
\]

where \( B(t) = e^{r(T-t)} \). These equations form the main focus for developing a numerical solution to the problem. The outline of the solution involves finding the hedging bounds for each state in a binomial tree of the underlying stock price. Once the no transaction region is identified, the valuation function is evaluated there. Finally the valuation function can be evaluated in other regions for all the values of \( \vartheta \). This is repeated for each state in the binomial tree for the underlying process. The exact details of the numerical scheme are identified in the next section.
7.7 Monoyios Algorithm

Monoyios [34] proposed an algorithm which prices European options in the presence of proportional transaction cost for the Davis, Panaz and Zariphopoulou framework. The algorithm uses a result proposed by Davis [11] in which an option is valued using marginal utility. This is done instead of solving a stochastic optimal control problem where an investor invests in the market and the option. This control problem is more complicated since it includes inserting the option payoff in the terminal value function.

Davis showed that the fair price of the option \( \hat{p} \) is determined when there is a neutral effect on the investor’s utility if an infinitesimal fraction of the initial wealth is diverted into the purchase or sale of the option at price \( \hat{p} \). Davis determined the fair price of the option to be

\[
\hat{p} = \frac{\mathbb{E} \left[ U'(\eta_T + c(\vartheta_T, S_T) + \Psi(S_T)) \right]}{\Psi^{(0)}(t, S, \eta, \vartheta)}, \tag{7.36}
\]

where \( U' \) is the derivative of the utility function and \( \Psi(S_T) \) is a payoff of a European option at maturity. The PDEs governing the solution remain the same as before, since both the numerator and denominator deal with problems investing in the market only. Using the exponential utility function the reduction in dimensionality is again applied to both valuation functions. Monoyios provides the following steps to implement equation (7.36) for pricing the option. Using the exponential utility function, the price of the option becomes

\[
\hat{p}(t, S, \vartheta) = e^{-r(T-t)} \frac{G(t, S, \vartheta)}{H(t, S, \vartheta)}, \tag{7.37}
\]

where \( G(t, S, \vartheta) \) and \( H(t, S, \vartheta) \) are the reduced form functions for the numerator and the denominator of (7.36). The following steps explain the implementation of the numerical algorithm:

- **Step 1:** Use the following equations to create a binomial tree for each \((S,t)\) node.

\[
S_{t+\delta t} = \begin{cases} 
S_tw_u & \text{with probability } \frac{1}{2} \\
S_tw_d & \text{with probability } \frac{1}{2}
\end{cases}
\]

and

\[
B_{t+\delta t} = Bte^{-\delta t},
\]
where

\[ w_u = e^{\mu \delta t + \sigma \sqrt{\delta t}} \quad w_d = e^{\mu \delta t - \sigma \sqrt{\delta t}}. \]

The number of time steps in the tree is given by parameter \( N \) and thus \( \delta t = \frac{T}{N} \).

- **Step 2:** Calculate the hedging boundaries for each node at time \( T - \delta t \) using the following formula

\[ \vartheta + \delta L^* = \vartheta_{(b)}(T - \delta t, S) = \frac{1}{\gamma S(w_u - w_d)} \log \left( \frac{q(1 - q_+)}{(1 - q)q_+} \right) \quad \text{BUY boundary} \]

\[ (7.38) \]

and

\[ \vartheta - \delta M^* = \vartheta_{(s)}(T - \delta t, S) = \frac{1}{\gamma S(w_u - w_d)} \log \left( \frac{q(1 - q_-)}{(1 - q)q_-} \right) \quad \text{SELL boundary} \]

\[ (7.39) \]

where

\[ q_+ = \frac{(R + \lambda) - w_d}{w_u - w_d} \quad q_- = \frac{(R - \lambda) - w_d}{w_u - w_d}, \]

\[ R = e^{r \delta t}, \quad q = \frac{1}{2}, \quad q_+, q_- \text{ are pseudo probabilities and } \delta L^*, \delta M^* \text{ are the optimal number of shares to be bought or sold to attain the optimal hedging boundaries } \vartheta_{(b)}, \vartheta_{(s)} \text{ respectively.} \]

Note that the analytical formula for the optimal boundary is derived by taking the expression in the transaction region, differentiating with respect to \( \delta L \) or \( \delta M \) and solving it for zero.

- **Step 3:** Create a vector of \( \vartheta \) values which ranges from the minimum to the maximum values from step 2 in increments of \( h_{\vartheta} \). The minimum and maximum vector values are chosen as

\[ \vartheta_{\text{min}} = \min_S \vartheta_{(b)}(T - \delta t, S) \quad \vartheta_{\text{max}} = \max_S \vartheta_{(s)}(T - \delta t, S). \]

Notice that the \( \vartheta \) vector needs to be large and include the whole no transaction region. The no transaction region is the largest at time \( T - \delta t \) and hence we will use this \( \vartheta \) vector throughout the algorithm at different time steps.

Using the newly derived \( \vartheta \) vector we can apply it to the terminal boundary condition for \( G \) and \( H \) functions at every \( S \) node at time \( T - \delta t \)

\[ G(T - \delta t, S, \vartheta) = -e^{-\gamma \vartheta S} \Psi(S_T) \]
and

\[ H(T - \delta t, S, \vartheta) = -e^{-\gamma \vartheta S}. \]

At this stage the values for \( G \) and \( H \) functions are determined at time \( T - \delta t \) for all values of \( S \) and \( \vartheta \). To compute the values for \( G \) and \( H \) functions at all other times, steps 4 and 5 must be performed until time 0 is reached.

- **Step 4:** Compare the terms in the following discrete equation and find the hedging boundaries using

\[
H(t, S, \vartheta) = \max_{\delta L, \delta M} \left[ \mathbb{E}[H(t + \delta t, wS, \vartheta + \delta L)] e^{\gamma S(1+\lambda_b)\delta LB(t)} \right. \\
\left. \mathbb{E}[H(t + \delta t, wS, \vartheta)] e^{-\gamma S(1-\lambda_s)\delta MB(t)} \right]
\]

where \( B(t) = e^{r(T-t)} \).

For each \( S \) node at time \( t \), we go through all values of the \( \vartheta \) vector in an increasing order and compare equations (7.40), (7.41) and (7.41), (7.42).

- The smallest \( \vartheta \) value for which (7.41) is greater than (7.40) is the buy boundary \( \vartheta(b)(t, S) \) for that particular node.
- Similarly, the largest \( \vartheta \) value for which (7.42) is greater than (7.41) is the sell boundary \( \vartheta(s)(t, S) \) for that particular node.

Note that \( \delta L \) and \( \delta M \) are set to \( h_\vartheta \) the discretisation step of \( \vartheta \) vector.

- **Step 5:** Having found the no transaction boundaries \( \vartheta(b) \) and \( \vartheta(s) \) for each \( S \) node at current time \( t \), the investor transacts towards the boundary when the portfolio is in the buy or sell region, using the following equations for the \( H \) function.

\[
H(t, S, \vartheta) = \begin{cases} 
\mathbb{E}[H(t + \delta t, wS, \vartheta)] & \text{if } \vartheta(b)(t, S) \leq \vartheta \leq \vartheta(s)(t, S) \\
H(t, S, \vartheta(b)(t, S)) e^{\gamma S(1+\lambda_b)(\vartheta(b)(t, S)-\vartheta)B(t)} & \text{if } \vartheta < \vartheta(b)(t, S) \\
H(t, S, \vartheta(s)(t, S)) e^{-\gamma S(1-\lambda_s)(\vartheta-\vartheta(s)(t, S))B(t)} & \text{if } \vartheta > \vartheta(s)(t, S).
\end{cases}
\]
Similarly the following equations are used for the $G$ function.

$$G(t, S, \vartheta) = \begin{cases} 
\mathbb{E}\left[ G(t + \delta t, wS, \vartheta) \right] & \text{if } \vartheta_{(b)}(t, S) \leq \vartheta \leq \vartheta_{(s)}(t, S) \\
G(t, S, \vartheta_{(b)}(t, S))e^{\gamma S(1 + \lambda_b)(\vartheta_{(b)}(t, S) - \vartheta)B(t)} & \text{if } \vartheta < \vartheta_{(b)}(t, S) \\
G(t, S, \vartheta_{(s)}(t, S))e^{-\gamma S(1 - \lambda_s)(\vartheta - \vartheta_{(s)}(t, S))B(t)} & \text{if } \vartheta > \vartheta_{(s)}(t, S). 
\end{cases}$$

Finally we are able to find $H(0, S, \vartheta)$ and $G(0, S, \vartheta)$ and solve for the option price at time 0

$$\tilde{p}(0, S, \vartheta) = e^{-r(T)} \frac{G(0, S, \vartheta)}{H(0, S, \vartheta)}.$$  

Notice the portfolio hedging bounds derived in the Monoyios algorithm are for maximising utility without holding the option. Hence the hedging bounds are not applicable for hedging of the option, but represent the optimal strategy to derive maximum terminal utility investing in the market without the option. To find the optimal hedging strategy for the option $\pi_1$, it is necessary to find the hedging strategy which maximises the terminal wealth with one option $\pi_1$ applicable to $\sup_{\pi_1} \mathbb{E}\left[ U(\eta_T + 1_{\{S_T \leq K\}}[c(\vartheta_T, S_T)] + 1_{\{S_T > K\}}[K + c(\vartheta_T - 1, S_T)]) \right]$.

Using definition 7.13, the option hedging bounds are written as

$$\vartheta_{U_{(b)}} = \vartheta_{(b)} - \vartheta_{(b)} \quad \vartheta_{U_{(s)}} = \vartheta_{(s)} - \vartheta_{(s)}.$$  

To calculate $\vartheta_{(b)}$ and $\vartheta_{(s)}$, we follow Monoyios algorithm with $H_{new}$ function which takes the following value at maturity

$$H_{new}(T, S, \vartheta) = e^{-\gamma(\vartheta S - \Psi(S_T))}.$$  

Furthermore the values of the $\vartheta_{(b)} = \vartheta + \delta L^{(1)}$ and $\vartheta_{(s)} = \vartheta - \delta M^{(1)}$ bounds at time $T - \delta t$ are

$$\vartheta_{(b)}(T - \delta t, S) = \frac{1}{\gamma S(w_u - w_d)} \left[ \log \left( \frac{q(1 - q_+)}{(1-q)q_+} \right) + \gamma(\Psi(S w_u) - \Psi(S w_d)) \right]$$  

and

$$\vartheta_{(s)}(T - \delta t, S) = \frac{1}{\gamma S(w_u - w_d)} \left[ \log \left( \frac{q(1 - q_-)}{(1-q)q_-} \right) + \gamma(\Psi(S w_u) - \Psi(S w_d)) \right].$$  

Consequently, step 4 and step 5 in the algorithm are adjusted, using $\vartheta_{(b)}$ and $\vartheta_{(s)}$. 

---
7.8 Issues with Implementation

In step 4 the hedging boundaries are found by comparing terms in the following equations

\[ H_{\text{new}}(t, S, \vartheta) = \max_{\delta L, \delta M} \left[ \mathbb{E}[H_{\text{new}}(t + \delta t, wS, \vartheta + \delta L)] e^{\gamma S(1+\lambda b)\delta LB(t)} - \mathbb{E}[H_{\text{new}}(t + \delta t, wS, \vartheta)] - \mathbb{E}[H_{\text{new}}(t + \delta t, wS, \vartheta - \delta M)] e^{-\gamma S(1-\lambda s)\delta MB(t)} \right]. \tag{7.45} \]

\[ \mathbb{E}[H_{\text{new}}(t + \delta t, wS, \vartheta)] \tag{7.46} \]

\[ \mathbb{E}[H_{\text{new}}(t + \delta t, wS, \vartheta - \delta M)] e^{-\gamma S(1-\lambda s)\delta MB(t)} \tag{7.47} \]

In step 5, the \( H_{\text{new}} \) function is evaluated using the boundaries \( \vartheta^{(1)}_{(b)} \) and \( \vartheta^{(1)}_{(s)} \).

\[ H_{\text{new}}(t, S, \vartheta) = \begin{cases} 
\mathbb{E}[H_{\text{new}}(t + \delta t, wS, \vartheta)] & \text{if } \vartheta^{(1)}_{(b)}(t, S) \leq \vartheta \leq \vartheta^{(1)}_{(s)}(t, S). \\
H_{\text{new}}(t, S, \vartheta^{(1)}_{(b)}(t, S)) e^{\gamma S(1+\lambda b)(\vartheta^{(1)}_{(b)}(t, S) - \vartheta)B(t)} & \text{if } \vartheta < \vartheta^{(1)}_{(b)}(t, S). \\
H_{\text{new}}(t, S, \vartheta^{(1)}_{(s)}(t, S)) e^{-\gamma S(1-\lambda s)(\vartheta - \vartheta^{(1)}_{(s)}(t, S))B(t)} & \text{if } \vartheta > \vartheta^{(1)}_{(s)}(t, S). 
\end{cases} \]

The Monoyios algorithm can be used to derive both the option price and the correct hedging bounds for the option. Furthermore it also gives the hedging bounds for an investor that wishes to invest optimally in the market with and without the option.

### 7.8 Issues with Implementation

Some issues were experienced when implementing the Monoyios algorithm. In particular issues with implementing equations (7.38), (7.39), (7.43) and (7.44) were experienced when running the algorithm along multiple paths. In particular, the expression in the log function needed to be positive but this might not always be the case due to the values of the chosen parameters. The following constraints need to hold so that the analytical solution for the boundary values at time \( T - \delta t \) can be found:

\[ \frac{q(w_u - (R + \lambda))}{(1 - q)(R + \lambda) - w_d} > 0, \tag{7.48} \]

\[ \frac{q(w_u - (R - \lambda))}{(1 - q)(R - \lambda) - w_d} > 0. \tag{7.49} \]

The only parameter that is controlled by the hedger is the \( \delta t \) term found in \( w_u, w_d \) and \( R \). Thus before evaluating the algorithm, the hedger needs to check that the chosen \( \delta t \) term satisfies (7.48) and (7.49). If one of the equations is not satisfied the number of time steps \( N \) in the algorithm need to be decreased which in turn will increase \( \delta t \). As a result, the algorithm uses a lesser number of time steps close to
maturity and is expected to have a larger numerical error in this region.

To overcome this problem, the analytical expression can be abandoned and the boundaries could be found by comparing terms as it is done in step 4. To do this a vector of $\vartheta$ values needs to be created, but it is unclear how the endpoints need to be chosen. Furthermore, some precision is lost since the solution is no longer analytical as it relies on the discretisation of the vector. This method was tested to see if the algorithm can use a large number of time steps close to maturity. However the values that needed to be compared increase significantly and the double machine precision is not big enough to store them. At this stage it remains unclear that using this method will allow to use a large number of time steps close to maturity and the original Monoyios algorithm was implemented instead in the last chapter.

The aforementioned implementation problem is more evident as Figure 7.1 shows the behaviour of the hedging bounds close to maturity. The hedging bounds intersect as time to maturity decreases although the final chapter shows this does not have a large negative impact on the results. Noticeably the number of time steps chosen for the Monoyios algorithm decreases substantially when close to maturity as the constraints (7.48) and (7.49) need to be satisfied. This can be seen as the cause of the intersecting boundaries.

![Graph showing hedging bounds](image)

**Fig. 7.1:** The hedging bounds for both utility indifference pricing algorithms are plotted time remaining to maturity of the option. The parameters are: $S_0 = 100, K = 100, T = 0.25, \sigma = 0.3, \mu = 0.08, r = 0.04, \lambda = 0.01$ and $\gamma = 2$ (for both algorithms).
7.9 Asymptotic Analysis

The analysis in the previous section indicates that the solution for the three dimensional problem is computationally intensive and complicated. Whalley and Wilmott [41, 42] used asymptotic analysis for the same problem which gives an expression for the hedging bounds, based on the option gamma. The advantage of the approach is that it gives an algorithm which is easy to implement and use. The solution was generalised for an abstract cost structure in [42], while the solution for fixed, proportional and proportional plus fixed transaction costs was summarised in [40].

Asymptotic analysis is a mathematical technique used to measure the magnitude of parameters in a complicated problem. In the proposed problem the following limiting behaviour is noted: if the costs are zero, the option price should be the Black Scholes price. In the presence of transaction costs, the option hedge should be the Black Scholes delta plus some adjustment. Whalley and Wilmott showed that in the limit of small transaction cost, the three dimensional problem becomes a simpler two dimensional inhomogeneous diffusion equation. The derived equation can be solved using a finite difference method to derive the option price. Interestingly, the hedging bandwidths are determined explicitly for the case of proportional transaction costs and given by

\[
\vartheta_{BS}^t - \left( \frac{3\lambda S_t e^{-r(T-t)} \Gamma^2}{2\gamma} \right)^{\frac{1}{3}} \leq \vartheta_{WWW}^t \leq \vartheta_{BS}^t + \left( \frac{3\lambda S_t e^{-r(T-t)} \Gamma^2}{2\gamma} \right)^{\frac{1}{3}},
\]

(7.50)

where \( \Gamma \) is the Black Scholes gamma given by

\[
\Gamma = \frac{\partial^2 F}{\partial S^2} = \frac{N'(d_1)}{S \sigma \sqrt{T-t}}.
\]

(7.51)

This approximation improves under an assumption of small transaction costs and under large values of the Black Scholes gamma. Similarly to the Monoyios algorithm, rebalancing takes place when the number of required stock holding moves outside the no transaction region and hedging is performed to the edge of the region. This is true for the case of proportional transaction costs only and the hedging boundary moves toward the center with the introduction of fixed costs. Finally, (7.50) is implemented for comparison purposes. Note that the method proposed by Whalley and Wilmott is not in any way related to the Wilmott delta derived in Chapter 3. The Whalley and Wilmott algorithm solely refers to the implementation of (7.50) and (7.51).
7.10 Numerical Results

In this section, numerical results on the utility indifference prices and hedging bounds are examined. Close attention is given to see how the hedging bounds compare to the Black Scholes delta. A calibration of the Monoyios algorithm together with Whalley and Wilmott algorithm is performed to choose an optimal risk aversion parameter for the Monte Carlo simulation in the last chapter. In particular the behaviour of the bounds for both utility indifference models is examined by changing the variables in the model. Certain implementation issues arose which are discussed in this section.

Figure 7.2 shows the bid and ask prices derived using Monoyios algorithm with three levels of transaction cost together with the Black Scholes price for various stock prices. The bounds are close to the Black Scholes price and widen with high levels of proportional transaction costs. This property is intuitive as the hedger is compensated by a bid-ask spread in the presence of transaction costs.

![Figure 7.2](image)

**Fig. 7.2:** Black Scholes price and Monoyios bid-ask spread are plotted for various levels of proportional transaction cost $\lambda$ against the stock price. The parameters are: $K = 15$, $T = 1$, $\sigma = 0.25$, $\mu = 0.15$, $r = 0.01$, $\gamma_{\text{Mon}} = 0.1$ and $\gamma_{\text{WW}} = 0.1$.

Figure 7.3 shows the buy and sell bounds of both algorithms together with the Black Scholes delta across various stock prices. The parameters used are taken from the Monoyios paper [34, Figure 5] and the Monoyios bounds appear to be identical to the paper. An interesting property which is not discussed in the paper is the issue of
the intersecting boundaries for the Monoyios algorithm. Under close investigation, the intersecting boundaries appear to be a function of the risk aversion parameter. This point is illustrated by Figures 7.4 and 7.5 when the risk aversion parameter is increased to $\gamma = 0.1$ and $\gamma = 10$ respectively, while keeping other parameters the same. The asymptotic bounds do not intersect each other, but appear to always be wider than the Monoyios bounds. It remains unclear why the Monoyios bounds intersect each other at low values of the risk aversion parameter and it is considered to be a property of the model. The three figures also illustrate that by increasing the risk aversion parameter, the bounds tighten around the Black Scholes delta. It is important to calibrate the risk aversion parameter for the strategy to perform optimally as very wide bounds will have high hedging errors, while tight bounds will incur large transaction costs.

Fig. 7.3: The hedging bounds for both utility indifference pricing algorithms are plotted against the stock price. The parameters are: $K = 15$, $T = 1$, $\sigma = 0.25$, $\mu = 0.15$, $r = 0.01$, $\gamma^{\text{Mon}} = 0.1$, $\gamma^{\text{WW}} = 0.1$ and $\lambda = 0.005$.

Figures 7.6 and 7.7 demonstrate the calibration which was performed for utility indifference algorithms. For each point on the graph, 1000 paths were generated and the algorithms were checked 180 times to see if hedging needed to be performed. The distribution of the hedging errors was worked out for each gamma parameter. Value at risk and expected shortfall were calculated for each distribution and plotted. The Black Scholes risk measures are constant since its hedging is independent of the risk
aversion parameter. The lowest expected shortfall was achieved at $\gamma = 2$ for the Monoyios algorithm and $\gamma = 3.5$ for the Whalley and Wilmott strategy. Due to the numerical error incurred when performing this type of calibration, it is impossible to find the exact optimal value for the risk aversion parameter, however the figures indicate that any value between 1.5 and 4.5 will be satisfactory for both models. It it important to note that by changing the number of times the hedging algorithms are checked, the results of the calibration could differ, however due to the computational constraints this calibration was only performed once for a given set of parameters. The performed calibration also indicates that the Monoyios algorithm outperforms the asymptotic approximation approach for any chosen risk aversion parameter. A more rigorous comparison is performed in the last chapter.

Figure 7.8 shows the behaviour of the bounds as the risk aversion parameter is changed. Both bounds are around the Black Scholes delta and tighten as the risk aversion parameter increases. This is as a result of the investor becoming more risk averse and not allowing his portfolio to incur hedge slippage.

Figure 7.9 shows the bounds while the proportional transaction costs are increased. The bounds widen as the costs increase, however Whalley and Wilmott bounds increase at a much faster rate. The bounds for the Monoyios algorithm
Fig. 7.5: The hedging bounds for both utility indifference pricing algorithms are plotted against the stock price. The parameters are: $K = 15$, $T = 1$, $\sigma = 0.25$, $\mu = 0.15$, $r = 0.01$, $\gamma_{\text{Mon}} = 10$, $\gamma_{\text{WW}} = 10$ and $\lambda = 0.005$.

... actually tighten for large values of the transaction costs. This could be the result of a numerical error since the time steps are decreased so that the constraints in (7.48) and (7.49) are satisfied for large values of transaction cost.

Figure 7.10 plots the algorithm bounds against the volatility of the option. The Black Scholes delta increases with an increase in the volatility while the asymptotic bounds tighten around it symmetrically. Interestingly the Monoyios bounds are not symmetrical around the delta especially for larger volatility values. In the figure the time steps are again decreased to satisfy the constraints for high values of volatility.

To investigate the effects of decreasing the time steps in the Monoyios model, Figure 7.11 shows the behaviour of the bounds for a given number of time steps. The algorithm gives stable bounds for the values of 25 time steps and above. For lower values the bounds become unstable as the time steps get closer to 1. This is attributed to the valuation of the needed expectations in the algorithm. Although as seen previously this problem arises as the time to maturity decreases, the result in the final chapter show that the algorithm can still be implemented successfully.

Conclusively, the figures also show that the asymptotic bounds lie on the outside of the Monoyios bounds for all the chosen parameters. In some instances the asymptotic bounds do not always follow them closely. As seen in Figures 7.5 and...
7.10 Numerical Results

Fig. 7.6: The expected shortfall 99% is calculated for each 1000 paths while hedging was checked 180 times. The parameters are: $S_0 = 100$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $\mu = 0.08$, $r = 0.04$ and $\lambda = 0.01$.

7.10 the Monoyios bounds are not even symmetrical around the Black Scholes delta. As a result one would expect the Monoyios algorithm to outperform the asymptotic approximation approach when hedging. This section also showed that the Monoyios algorithm is a stable, robust method to find prices and hedging bounds for the utility indifference pricing framework.
Fig. 7.7: The value at risk 99% is calculated for each 1000 paths while hedging was checked 180 times. The parameters are: \( S_0 = 100, \ K = 100, \ T = 0.25, \ \sigma = 0.3, \ \mu = 0.08, \ r = 0.04 \) and \( \lambda = 0.01 \).

Fig. 7.8: The hedging bounds for both utility indifference pricing algorithms are plotted against the risk aversion parameter. The parameters are: \( S_0 = 100, \ K = 100, \ T = 0.25, \ \sigma = 0.3, \ \mu = 0.08, \ r = 0.04 \) and \( \lambda = 0.01 \).
### 7.10 Numerical Results

**Fig. 7.9:** The hedging bounds for both utility indifference pricing algorithms are plotted against proportional transaction cost. The parameters are: $S_0 = 100$, $K = 100$, $T = 0.25$, $\sigma = 0.3$, $\mu = 0.08$, $r = 0.04$, $\gamma_{\text{Mon}} = 2$ and $\gamma_{\text{WW}} = 2$.

**Fig. 7.10:** The hedging bounds for both utility indifference pricing algorithms are plotted against volatility. The parameters are: $S_0 = 100$, $K = 100$, $T = 0.25$, $\mu = 0.08$, $r = 0.04$, $\gamma_{\text{Mon}} = 2$, $\gamma_{\text{WW}} = 2$ and $\lambda = 0.01$. 
Fig. 7.11: The hedging bounds for both utility indifference pricing algorithms are plotted against the number of time steps in the Monoyios algorithm. The parameters are: $S_0 = 100, K = 100, T = 0.25, \sigma = 0.3, \mu = 0.08, r = 0.04, \gamma_{\text{Mon}} = 2, \gamma_{\text{WW}} = 2$ and $\lambda = 0.01$. 
Chapter 8

Numerical Comparison

8.1 Introduction

This chapter provides a brief summary of the aforementioned hedging strategies and examines their properties numerically. The hedging strategies are examined by running Monte Carlo simulations with holdings of a long underlying stock and a short call option under various scenarios. Similar analysis has been performed by Mohamed [33] which used different parameters on a long term option. The analysis performed in this chapter differs from that of Mohamed, in that the term of the option is smaller and additional models are compared.

The strategies are examined to find the optimal time interval at which hedging should take place. This is important as the option hedger wants to be optimal in the trade off between hedge slippage and transaction costs. Furthermore, other scenarios are explored to show which hedging rules perform favourably under various trading regimes such as different levels of volatility and transaction costs. The models considered are divided into two groups, namely local-in-time and global-in-time models. Although by definition these strategies are implemented differently, careful consideration was taken to ensure fair comparisons. To evaluate the performance of each hedging rule, profit and loss distributions together with their risk measures, are compared. A quick summary of each model is presented, followed by comparisons and conclusion.

A parallel computing cluster was utilised due to computational complexity of calculating profit and loss distributions for each hedging algorithm. The computational difficulty arises due to the requirement of running the algorithms multiple times along each single path to get one realisation of a profit or loss. Furthermore, to observe the full distribution, thousands of paths are needed for each simulation. Note that each algorithm is viable for use in a real-time trading environment, with the Monoyios algorithm taking the longest amount of time to complete (of the order of 10 seconds to price and compute the hedge parameters). In this chapter, it
is shown that the calculations performed would not have been possible to do on a single desktop computer. A brief description of the technology used is provided.

8.2 Local-in-time Models

By definition, the local-in-time models attempt to minimise risk for a predetermined hedging interval. The hedging intervals are generally fixed and equally spaced through the lifetime of the option. Two models were considered, namely a hedging rule proposed by Leland [28] and a quadratic hedging approach which was originally proposed by Schweizer [38].

The Leland model adjusts the volatility by considering the given time step and the magnitude of proportional costs. The model uses the Black Scholes equation to price the option and as a result it is relatively easy to implement. The volatility adjustment is given by:

$$
\hat{\sigma}^2 = \sigma^2 \left[ 1 + \sqrt{\frac{2}{\pi}} \frac{2\lambda}{\sigma \sqrt{\delta t}} \right],
$$

(8.1)

where $\lambda$ is the proportional transaction cost and $\delta t$ is the length of the hedging interval. The option is priced using the Black Scholes pricing formula with the new modified volatility. The hedge ratio is simply the derivative of the modified option price ($\hat{F}$) with respect to the underlying stock

$$
\theta_L = \frac{\partial \hat{F}}{\partial S} = N\left(d_1\right).
$$

(8.2)

The alternative local-in-time model under consideration is quadratic hedging. Schweizer introduced the general framework where the strategy $\pi^{QH}$ minimises the following risk function:

$$
R_t(\pi) = \mathbb{E} \left[ (C_T(\pi) - C_t(\pi))^2 \mid \mathcal{F}_t \right],
$$

(8.3)

where $C_t(\pi)$ denotes the cost process for strategy $\pi$. It can be shown that in discrete time it is equivalent to minimise

$$
\mathcal{R}_t(\pi) = \mathbb{E} \left[ (C_{t+1}(\pi) - C_t(\pi))^2 \mid \mathcal{F}_t \right]
$$

(8.4)

The numerical procedure was proposed by Mercurio and Vorst [30] in which the quadratic hedging strategy is implemented in the presence of proportional transaction costs. The suggested algorithm discretises the stock process by implementing
Global-in-time models attempt to minimise risk over the lifetime of the option at each time instant. This is implemented as a choice between hedging (and incurring transaction costs) or not hedging (and incurring a hedge slippage). To implement this at every time instant, a global-in-time model produces hedging bounds which determine a region in which hedging is not performed. This is known as the no transaction region. As soon as the stock holding leaves this region, hedging is performed to get back into the no transaction region. Practically it is not possible to evaluate it at every instant of time as these models can be computationally expensive. To solve this issue, the model can be implemented discretely. It is expected that the model will approach its theoretical benchmarks as these discrete intervals become shorter. This is tested in a later section.

Davis, Panas and Zariphopoulou [36] proposed using utility functions to select optimally in the trade off between incurring hedge slippage or paying transaction costs. Their model uses utility indifference pricing to value the option and select an optimal hedging strategy. By definition, a utility indifference price for the option is the price at which the investor is indifferent (in expected utility terms) between investing in the market with the long (short) option and paying (receiving) a suitable amount of compensation or investing in the market without the option. To solve the proposed question, stochastic optimal control theory is applied and a Hamilton Jacobi Bellman equation is derived. Solving the PDEs analytically has not been possible and a numerical procedure was required. Monoyios [34] proposed a numerical algorithm using a Markov chain approximation method. Furthermore, the approach is different from the original paper since it uses a result by Davis [11] which determines the price of the option using marginal utility. The Monoyios algorithm is computationally intensive as it involves discretisation of a multidimensional problem. The parameter $N_{\text{Mon}}$ was chosen to control the number of discretised time steps for the stock process.
8.4 Simulation

Whalley and Wilmott [41, 42] give an alternative approach for solving the utility formulation of Davis, Panas and Zariphopoulou. The approach involves using asymptotic analysis on the derived Hamilton Jacobi Bellman equations. This gives an expression for the hedging bounds, based on the option gamma. The hedging bounds are given by

\[
\vartheta_{BS}^t - \left( \frac{3\lambda S_t e^{-r(T-t)} \Gamma^2}{2\gamma} \right)^{\frac{3}{2}} \leq \vartheta_{WW}^t \leq \vartheta_{BS}^t + \left( \frac{3\lambda S_t e^{-r(T-t)} \Gamma^2}{2\gamma} \right)^{\frac{3}{2}},
\]  

(8.5)

where \( \Gamma \) is the Black Scholes gamma given by

\[
\Gamma = \frac{\partial^2 F}{\partial S^2} = \frac{N'(d_1)}{S\sigma \sqrt{T-t}}.
\]  

(8.6)

The no transaction region for the model is given by (8.5). The advantage of this approach is that it gives an algorithm which is easy to implement and use, with the properties of a global-in-time model.

8.4 Simulation

The Monte Carlo scheme entails simulating a portfolio consisting of a long underlying stock and a short position of one call option through realistic trading with proportional transaction costs. This is done by generating geometric Brownian paths until maturity for a given set of parameters. The hedger’s initial position is cash only as it corresponds to an amount of cash received from the sale of the option. Each hedging strategy is implemented with the same initial starting point for comparison which is the Black Scholes price. The strategies only begin to differ as hedging begins to take place, since the strategies have different hedge ratios at each hedging point. The hedging points are uniformly spread through the lifetime of the option. Once each strategy is seeded with the same initial cash position the following takes place at each predetermined hedging point throughout the path:

- Using the current stock price and the time left to maturity, the new hedge ratios are calculated for each strategy.

- For global-in-time models, a check is made to see if hedging needs to take place. If the current holding is outside the no transaction region, the hedge is rebalanced and corresponding costs are paid.

- For local-in-time models, hedging is compulsory at every hedging point. If necessary, a new stock holding is acquired by paying the transaction cost and
transacting with the market and the bank account.

The hedger may borrow any amount from the bank account in which the lending rate is equal to the riskless rate $r$. At maturity of each path, the hedger liquidates the stock holding and pays the payoff of the option. It is important to note, that costs are charged initially to set up the hedge portfolio and to liquidate it at maturity. This is equivalent to assuming the option is settled in cash. For each path and each strategy, the hedging error is calculated to be the difference between the payoff of the option and the liquidated value of the portfolio.

As previously mentioned, global-in-time models are implemented differently to local-in-time models as hedging does not always occur. Due to computational constraints and comparison purposes, the global-in-time models were restricted to perform hedging at the same intervals as the local-in-time models. This allows one to compare the models for given hedge intervals. Furthermore, the performance of global-in-time model is expected to improve with an increase in the number of hedging times, as it is not compulsory to hedge at each interval.

8.5 Measuring Performance

In order to assess the performance of the hedging strategies, the cumulative hedging error was quantified for each path. This resulted in a profit and loss distribution for each hedging rule. For comparison purposes, the expected value and the variance of the distributions are examined. The variance incorporates both positive and negative deviations, however hedgers will find large positive deviations favourable. For this reason, several popular risk measures were chosen to examine the negative tail of the hedging error distribution. Value at risk (VaR) is one such measure and it is defined as

$$x^{(\alpha)}(X) = \sup \{x \mid P[X \leq x] \leq \alpha\},$$
$$\text{VaR}^{(\alpha)}(X) = -x^{(\alpha)}(X). \quad (8.7)$$

Practically, the value at risk is a chosen percentile, $\alpha$, of the loss distribution. Although VaR is very popular due to its simplicity and wide applicability, it is not a coherent risk measure as it is not subadditive. Expected shortfall (ES) was chosen as an alternative risk measure to compare the strategies. It is defined as

$$\text{ES}^{(\alpha)}(X) = -\frac{1}{\alpha} \left( E \left[ X 1_{\{X \leq x^{(\alpha)}\}} \right] - x^{(\alpha)} \left( P \left[ X \leq x^{(\alpha)} \right] - \alpha \right) \right). \quad (8.8)$$
Although the definition looks complex at first glance, it breaks down to a simple definition of a tail conditional expectation for continuous distributions

\[
\text{TCE}^{(\alpha)}(X) = \mathbb{E} \left[ -X \mid X \leq x^{(\alpha)}(X) \right].
\]  

(8.9)

The advantage of using expected shortfall is that it can also be easily calculated as it is simply the average of severe losses that occur \( \alpha \) percent of the time. Additionally, Acerbi and Tasche [1] show that expected shortfall is a coherent risk measure satisfying all four axioms: monotonicity, subadditivity, positive homogeneity and translation invariance. This is the reason expected shortfall was chosen to be the primary risk measure used in the comparisons.

### 8.6 Scenarios

The scenarios were selected to infer properties of the aforementioned hedging strategies. To do so, the number of possible hedging times (HT), levels of proportional transaction cost (\( \lambda \)) and stock price volatility (\( \sigma \)) were varied.

To implement the quadratic hedging approach, the initial number of time nodes where hedging takes place is given by \( N_{\text{QH}} \), while the tree is discretised further by choosing \( M_{\text{QH}} \) which determines the number of discretisation steps between two hedging times. The total number of binomial nodes is given by parameter \( TS \). As time to maturity decreases, the number of hedging times decreases and as a result, the total number of time steps also decreases. A parameter \( TS_{\text{min}} \) controls the minimum number of time steps in the quadratic hedging algorithm. To keep the remaining number of time steps in the model above \( TS_{\text{min}} \), parameter \( M_{\text{QH}} \) was manipulated to increase the size of the binomial tree. This was done without changing the remaining number of hedging times. The minimum total number of time steps was chosen to be 20 for all simulations.

As seen in Chapter 7 for utility indifference models, the calibration for the risk aversion parameter was performed by running simulations on 1000 paths with 180 hedging times. It was deemed sufficient to choose the risk aversion parameter value of 2 for both global-in-time models as it corresponded to the optimal region of lowest values for the expected shortfall and value at risk measures. Notably, this calibration was only performed for one set of parameters and one would expect improved results if calibration was performed for each simulation. The initial number of time steps for the Monoyios algorithm (\( N_{\text{Mon}} \)) was chosen to be 50 as this will give stable hedge ratios as seen in Figure 7.11.

One of the questions an option hedger faces is how often the option should be
hedged. To investigate this, the number of hedging times was varied. Table 8.1 is a summary of the parameters used for the first scenario. The number of hedging times is given by the first column, where the value varies from hedging biweekly to hedging twice daily.

### Table 8.1: Scenario 1: Varying the number of possible hedging times in the life of the option.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Stock Process</th>
<th>QH</th>
<th>Moneys</th>
<th>WW</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>HT</td>
<td>Paths</td>
<td>λ</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>6.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>8.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>12.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>18.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>24.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>6</td>
<td>30.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>7</td>
<td>45.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>8</td>
<td>60.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>9</td>
<td>90.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>10</td>
<td>120.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>11</td>
<td>150.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>12</td>
<td>180.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The size of proportional costs and volatility regimes could affect the models in different ways. In order to explore this, Table 8.2 and Table 8.3 show the parameters used to examine these two additional scenarios.

### Table 8.2: Scenario 2: Varying the size of the proportional transaction cost.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Stock Process</th>
<th>QH</th>
<th>Moneys</th>
<th>WW</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>HT</td>
<td>Paths</td>
<td>λ</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>90.00</td>
<td>10000</td>
<td>0.005</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>90.00</td>
<td>10000</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>90.00</td>
<td>10000</td>
<td>0.015</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>90.00</td>
<td>10000</td>
<td>0.02</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>90.00</td>
<td>10000</td>
<td>0.025</td>
<td>0.25</td>
</tr>
<tr>
<td>6</td>
<td>90.00</td>
<td>10000</td>
<td>0.03</td>
<td>0.25</td>
</tr>
</tbody>
</table>

For all scenarios, the chosen option was at-the-money with a maturity of 3 months. This was done for computational purposes as running 10000 paths and performing the chosen algorithms was not viable for a longer dated option. This period of time was deemed sufficient to make comparisons among the hedging strategies.
8.7 Technology Details

The scenarios are computationally intensive as each scenario involves running the aforementioned algorithms multiple times along a single path. This is then repeated for thousands of paths to determine the profit and loss distributions for each strategy. To run the discussed scenarios, a cluster of computers called Hydra, from the School of Computer Science at the University of Witwatersrand, was utilised. Hydra consists of 8 nodes, each installed with Intel Core i7 950 CPU @ 3.07GHz (four cores with hyper threading) and 6 GB of RAM. The nodes are connected using an ethernet hub and run 64-bit Ubuntu Linux, version 10.10. Furthermore, each node runs a 64-bit version of Matlab R2010b with a Parallel Computing Toolbox. In total 32 Matlab workers (one worker per core) were used in the cluster to perform the calculations.

Tab. 8.4: Time taken for Scenario 1 using 32 Matlab workers.

The problem was parallelised by sending price paths to the available workers in the cluster, through the use of `matlabpool` and `parfor` functions. Each worker then ran the hedging algorithms and returned a value for the hedging error for that particular path. Once a worker became free, it would be given the next generated price path. This process was repeated until the hedging error was calculated for each price path in the simulation. Table 8.4 shows the time taken for each simulation.
from the first scenario using 32 Matlab workers. For comparison purposes, Table 8.5 shows the time taken for the first simulation when only one Matlab worker was used.

<table>
<thead>
<tr>
<th>Simulation #</th>
<th>Time Taken</th>
<th>Simulation #</th>
<th>Time Taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 Days 0 Hours 28 Min 41 Sec</td>
<td>7</td>
<td>14 Days 14 Hours (estimated)</td>
</tr>
<tr>
<td>2</td>
<td>2 Days 19 Hours 49 Min 49 Sec</td>
<td>8</td>
<td>16 Days 21 Hours (estimated)</td>
</tr>
<tr>
<td>3</td>
<td>3 Days 15 Hours (estimated)</td>
<td>9</td>
<td>31 Days 11 Hours (estimated)</td>
</tr>
<tr>
<td>4</td>
<td>5 Days 5 Hours (estimated)</td>
<td>10</td>
<td>54 Days 10 Hours (estimated)</td>
</tr>
<tr>
<td>5</td>
<td>6 Days 21 Hours (estimated)</td>
<td>11</td>
<td>137 Days 5 Hours (estimated)</td>
</tr>
<tr>
<td>6</td>
<td>8 Days 1 Hours (estimated)</td>
<td>12</td>
<td>197 Days 7 Hours (estimated)</td>
</tr>
</tbody>
</table>

Tab. 8.5: Time taken for Scenario 1 using one Matlab worker.

By using the 32 Matlab workers in the cluster, the simulation was approximately sped up by 26.7 times. The total time taken to run the three scenarios on the cluster was 28 days, 6 hours, 44 minutes and 37 seconds. The estimated time to run the three scenarios with a single worker is 754 days. It is clear that this simulation would not have been possible on a desktop computer. This illustrates the importance of combining computer power to perform computational experiments of this nature.

8.8 Numerical Results

As previously mentioned, each simulation was seeded with an initial amount of cash equivalent to the Black Scholes price for the option. The same initial amount allows fair comparisons of the hedging distributions. This is also a realistic approach since in practice an option hedger would be given some amount (not necessarily the model’s price) to hedge the option. However, it is still important to compare the initial prices for each of the algorithms. The local-in-time prices determine an amount needed to set up a hedging portfolio with an expected value of zero. In a case of the Monoyios model, the bid-ask spread determine the prices with which an investor is willing (in utility terms) to buy or sell the option. In practice, a market maker would charge a risk premium over these prices. This is done to shift the hedging distribution to the right and attempt to make a profit on the trade. It follows that lower initial prices are favourable to the hedger as this will determine the size of the risk premium charged. A higher risk premium will also produce a higher mean and better risk measures for the hedging strategy. The option prices and their initial hedge ratios are given for the three scenarios in Tables 8.6, 8.7 and 8.8.

For the first scenario, only the local-in-time prices and hedge ratios change in
### 8.8 Numerical Results

#### Tab. 8.6: Scenario 1: The initial prices and hedge ratios for the implemented models.

<table>
<thead>
<tr>
<th>#</th>
<th>$F_0$</th>
<th>$\vartheta_0$</th>
<th>$F_0$</th>
<th>$\vartheta_0$</th>
<th>$F_0^{\text{Ask}}$</th>
<th>$F_0^{\text{Bid}}$</th>
<th>$\vartheta_0^{\text{Sell}}$</th>
<th>$\vartheta_0^{\text{Buy}}$</th>
<th>$\vartheta_0^{\text{Sell}}$</th>
<th>$\vartheta_0^{\text{Buy}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.4595</td>
<td>0.5563</td>
<td>7.7438</td>
<td>0.5571</td>
<td>6.7678</td>
<td>0.5430</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>2</td>
<td>6.4595</td>
<td>0.5563</td>
<td>7.8967</td>
<td>0.5573</td>
<td>6.9234</td>
<td>0.5454</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>3</td>
<td>6.4595</td>
<td>0.5563</td>
<td>8.0230</td>
<td>0.5575</td>
<td>7.0450</td>
<td>0.5375</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>4</td>
<td>6.4595</td>
<td>0.5563</td>
<td>8.2300</td>
<td>0.5579</td>
<td>7.2462</td>
<td>0.5357</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>5</td>
<td>6.4595</td>
<td>0.5563</td>
<td>8.3999</td>
<td>0.5583</td>
<td>7.5143</td>
<td>0.5395</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>6</td>
<td>6.4595</td>
<td>0.5563</td>
<td>8.5463</td>
<td>0.5586</td>
<td>7.6560</td>
<td>0.5339</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>7</td>
<td>6.4595</td>
<td>0.5563</td>
<td>8.8490</td>
<td>0.5593</td>
<td>7.9158</td>
<td>0.5460</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>8</td>
<td>6.4595</td>
<td>0.5563</td>
<td>9.0952</td>
<td>0.5599</td>
<td>8.2527</td>
<td>0.5428</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>9</td>
<td>6.4595</td>
<td>0.5563</td>
<td>9.4921</td>
<td>0.5609</td>
<td>8.9983</td>
<td>0.5448</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>10</td>
<td>6.4595</td>
<td>0.5563</td>
<td>9.8125</td>
<td>0.5619</td>
<td>9.3774</td>
<td>0.5469</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>11</td>
<td>6.4595</td>
<td>0.5563</td>
<td>10.0854</td>
<td>0.5627</td>
<td>9.6983</td>
<td>0.5485</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>12</td>
<td>6.4595</td>
<td>0.5563</td>
<td>10.3251</td>
<td>0.5634</td>
<td>9.9792</td>
<td>0.5499</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
</tbody>
</table>

#### Tab. 8.7: Scenario 2: The initial prices and hedge ratios for the implemented models.

<table>
<thead>
<tr>
<th>#</th>
<th>$F_0$</th>
<th>$\vartheta_0$</th>
<th>$F_0$</th>
<th>$\vartheta_0$</th>
<th>$F_0^{\text{Ask}}$</th>
<th>$F_0^{\text{Bid}}$</th>
<th>$\vartheta_0^{\text{Sell}}$</th>
<th>$\vartheta_0^{\text{Buy}}$</th>
<th>$\vartheta_0^{\text{Sell}}$</th>
<th>$\vartheta_0^{\text{Buy}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.4852</td>
<td>0.5596</td>
<td>7.3538</td>
<td>0.5566</td>
<td>6.7999</td>
<td>0.5337</td>
<td>5.0625</td>
<td>3.9183</td>
<td>0.6030</td>
<td>0.5114</td>
</tr>
<tr>
<td>2</td>
<td>6.4595</td>
<td>0.5563</td>
<td>9.4921</td>
<td>0.5609</td>
<td>8.9983</td>
<td>0.5448</td>
<td>6.9999</td>
<td>0.5936</td>
<td>0.5255</td>
<td>0.6365</td>
</tr>
<tr>
<td>3</td>
<td>8.4333</td>
<td>0.5596</td>
<td>11.6561</td>
<td>0.5676</td>
<td>11.1080</td>
<td>0.5553</td>
<td>9.0241</td>
<td>7.7640</td>
<td>0.5926</td>
<td>0.5390</td>
</tr>
<tr>
<td>4</td>
<td>10.4035</td>
<td>0.5655</td>
<td>13.6013</td>
<td>0.5753</td>
<td>13.1680</td>
<td>0.5659</td>
<td>11.0654</td>
<td>9.7546</td>
<td>0.5966</td>
<td>0.5510</td>
</tr>
<tr>
<td>5</td>
<td>12.3635</td>
<td>0.5727</td>
<td>15.6121</td>
<td>0.5836</td>
<td>15.1936</td>
<td>0.5758</td>
<td>12.9078</td>
<td>11.7513</td>
<td>0.6004</td>
<td>0.5650</td>
</tr>
<tr>
<td>6</td>
<td>14.3260</td>
<td>0.5807</td>
<td>17.6020</td>
<td>0.5922</td>
<td>17.1936</td>
<td>0.5858</td>
<td>14.7673</td>
<td>13.6749</td>
<td>0.6050</td>
<td>0.5776</td>
</tr>
</tbody>
</table>

#### Tab. 8.8: Scenario 3: The initial prices and hedge ratios for the implemented models.
value with each simulation as the Black Scholes model and global-in-time models are independent of the number of possible hedging times. The prices of the local-in-time models increase with more hedging times as the hedger requires compensation for the additional amount of transaction costs that will be paid at each of those additional hedging times.

In the second scenario, the prices of the models increase due to the increased transaction costs in each simulation. The hedging bounds of the Whalley and Wilmott model widen as it becomes more optimal to incur hedge slippage rather than to incur increased transaction costs. Although the Monoyios model would be expected to behave in a similar manner, the Monoyios model bounds widen initially and then tighten slightly. This is a result of numerical error since for large values of transaction cost the model needs to use a smaller number of time steps to fulfil the constraints of the model given by (7.48) and (7.49). Figure 7.9 shows this phenomenon in the previous chapter.

The volatility of the price process is increased gradually in the third scenario which results in all model prices and hedge ratios being effected. All prices increase as the increased volatility parameter suggests a higher level of risk. The hedger must be compensated for the additional risk by receiving a higher price for the option. The hedging bandwidth for each of the global-in-time models tighten for an increase in the volatility. With increased volatility it becomes optimal to incur more transaction costs rather than allow larger hedge slippage.

Disregarding the naive Black Scholes model, the ask price produced by the Monoyios model is always lower than the prices of other models for all given simulations. This is a favourable result for the global-in-time model. The Leland model produces the highest prices with the quadratic hedging model only being slightly behind.

Figure 8.1 shows the example of the observed distributions for one of the simulations. These hedging error distributions together with their statistics were determined for each simulation.

The influence of various hedging times is examined first by showing the risk measures for the profit and loss of each hedging strategy in Figures 8.2 and 8.3. The Black Scholes strategy is plotted as a benchmark, however this naive strategy is expected to perform poorly. At less frequent hedging intervals, all strategies are clustered together and the risk measures begin to decrease with more frequent hedging. The global-in-time risk measures decrease as the number of possible hedging times increase. The local-in-time models reach their minimum at between 20 and 40 hedging times and the risk measures start to increase with more frequent hedging times. For the chosen set of parameters, it is optimal to perform hedging around those levels
Fig. 8.1: The profit and loss distributions are plotted from the first scenario.

for the local-in-time models.

Figure 8.4 shows the variance of the profit and loss distributions for all the strategies. Apart from the Black Scholes model, this variance drops for all models as hedging intervals become more frequent. As the number of hedges increases, the mean of the distributions decreases as seen in Figure 8.5. The hedger might be interested in decreasing the variance of the profit and loss distribution. In this case, using a local-in-time model with a large number of hedging times would be preferable.

Since the proportional cost is levied every time rebalancing is performed, the local-in-time risk measures worsen as a number of hedges increase. The models compensate for this by charging a large initial price at these levels. However this increase in price is not taken into account here.

Clearly, the performance of global-in-time models with regard to the risk measures only improves with an increased number of hedging times. As previously mentioned this result is expected and now shown numerically in Figures 8.2 and 8.3. This is due to the fact that hedging is only performed if it is optimal at every hedge point and it is not compulsory. The global-in-time models will marginally improve further by having additional hedging times and will reach their minimums if the models are evaluated at every instant of time. This is of course not practically possible and the
8.8 Numerical Results

**Fig. 8.2:** Scenario 1: Value at Risk 95% of profit and loss distributions as a function of hedging times.

**Fig. 8.3:** Scenario 1: Expected Shortfall 95% of profit and loss distributions as a function of hedging times.
8.8 Numerical Results

Fig. 8.4: Scenario 1: Variance of profit and loss distributions as a function of hedging times.

Fig. 8.5: Scenario 1: Mean of profit and loss distributions as a function of hedging times.
8.8 Numerical Results

figures suggest that the risk measures flatten after 120 hedging times.

Having observed the properties of local and global models, it is possible to make
direct comparisons between individual models. Examining the risk measures, the
Monoyios algorithm outperforms all other models for each hedging time greater than
12. With less frequent hedging times, the Leland model performs well although all
models are clustered together with small differences between them. The Leland
model performs advantageously by having the lowest variance for all hedging times
with the Monoyios and quadratic hedging models competing closely for second place
in that regard. As discussed, this metric is not necessarily the best because a low
variance penalises both profits and losses.

Choosing the optimal hedging times for a hedger is not a simple matter as it
largely depends on what the hedger wants to achieve with the profit and loss distrib-
tion. If the hedger wants to minimise the variance, the Leland model seems to do
well, however it achieves this with the worst mean. As seen the Leland model also
produces highest prices out of all the models. This means the hedger would need to
sell the option at a high initial price together with an additional risk premium to
achieve a positive expectation on the trade. Of course, this might not be possible.
Furthermore, since the Leland model is a local-in-time approach the hedger would
need to select a hedging interval. This could be done by performing a similar ana-
lysis and choosing a hedging interval which corresponds to a level of a risk measure
which the hedger would deem acceptable.

Overall the Monoyios algorithm outperforms other strategies as it achieves best
risk measures of the profit and loss distributions while having the lowest initial
price. It does this at a variance which is not much higher than of the Leland model.
Moreover, the hedger does not need to worry about choosing a hedging interval as
it is a global-in-time model.

In the second scenario, the proportional transaction costs were varied for all
models and the total hedging error distribution properties were examined in Figures
8.6, 8.7, 8.8 and 8.9. As expected, the risk measures and variance increase with
higher costs while the means of the distribution fall. In the Figures 8.6 and 8.7, the
Monoyios algorithm performs favourably at lower end of the transaction costs with
the Whalley and Wilmott algorithm outperforming it marginally at higher levels of
costs. As in the first scenario the Leland algorithm achieves the lowest variance for
low transaction cost values while the Monoyios algorithm performs favourably at
high values of transaction costs.

In the third scenario, various volatility levels were examined in Figures 8.10,
8.11, 8.12 and 8.13. Similarly to the second scenario, the risk measures and variance
increase with higher levels of volatility while the means of the distributions decreased.
Fig. 8.6: Scenario 2: Value at Risk 95% of profit and loss distributions as a function of proportional transaction costs.

Fig. 8.7: Scenario 2: Expected Shortfall 95% of profit and loss distributions as a function of proportional transaction costs.
8.8 Numerical Results

Fig. 8.8: Scenario 2: Variance of profit and loss distributions as a function of proportional transaction costs.

Fig. 8.9: Scenario 2: Mean of profit and loss distributions as a function of proportional transaction costs.
In the figures, the Monoyios algorithm performs favourably again as it achieves the lowest values of VaR and expected shortfall. As seen previously, Figure 8.12, shows that the Leland model again achieves the lowest variance values for all levels of stock volatility.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{scenario_3_values_at_risk.png}
\caption{Scenario 3: Value at Risk 95\% of profit and loss distributions as a function of stock volatility.}
\end{figure}

\section*{8.9 Concluding Remarks}

After the analysis above, it is possible to draw some conclusions about the models presented. As expected prior to conducting this experiment, a global-in-time model outperformed the local-in-time approach. In particular, the Monoyios model achieved best risk measures in all three scenarios. The Monoyios algorithm also produced lowest ask prices. This is an important property since the hedger is likely to charge a risk premium when selling options. It can be argued that other models performed better than the Monoyios algorithm in other aspects. The Leland model achieved the lowest variance for all three scenarios. This is achieved with the worst mean and the highest initial price. Although variance is an important criteria, the option hedger should be more concerned with the negative deviations rather than making excess profit. For example, it is optimal for a hedger to choose a strategy with a large variance provided that a large part of the distribution is positive or
8.9 Concluding Remarks

Volatility

Expected Shortfall 95%

BS
QH
Leland
WW
Mon

Fig. 8.11: Scenario 3: Expected Shortfall 95% of profit and loss distributions as a function of stock volatility.

Variance of PnL

BS
QH
Leland
WW
Mon

Fig. 8.12: Scenario 3: Variance of profit and loss distributions as a function of stock volatility.
8.9 Concluding Remarks

Fig. 8.13: Scenario 3: Mean of profit and loss distributions as a function of stock volatility.

The distribution has a relatively low risk measure. It is for this reason that Whalley and Wilmott model also performed better than the local-in-time models. Whalley and Wilmott model achieved the worst variance of the hedging error distribution, however in all three scenarios it achieved second best results when comparing risk measures and the best mean. The quadratic hedging model performed poorly in almost all instances, as the Leland model outperformed it marginally in terms of risk measures. However the quadratic model produced a marginally better mean than the Leland model which could explain its slightly worse performance. Having performed this analysis, it remains up to the hedger to determine which model is best to use when dealing with proportional transaction costs as the preferences of the hedger need to be addressed. However, this dissertation concludes that the algorithm presented by Monoyios has given the most favourable results from all chosen models due to its low price, lowest risk measures, competitive variance and strong properties of global-in-time models.

As previously mentioned, geometric Brownian motion is not a great approximation of reality as it does not fully match the market dynamics. One area for further research would be to extend the presented transaction cost models to incorporate a more realistic model for the underlying asset.
Bibliography


