MODELLING APPROACHES FOR OPTIMAL LIQUIDATION UNDER A LIMIT-ORDER BOOK STRUCTURE

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2016

James Blair
School of Mathematics

Contents

Word count 60717

List of Tables

List of Figures

The University of Manchester

James Blair
Doctor of Philosophy
Modelling approaches for optimal liquidation under a limit-order book structure
April 21, 2016

This thesis introduces a selection of models for optimal execution of financial assets at the tactical level. As opposed to optimal scheduling, which defines a trading schedule for the trader, this thesis investigates how the trader should interact with the order book. If a trader is aggressive he will execute his order using market orders, which will negatively feedback on his execution price through market impact. Alternatively, the models we focus on consider a passive trader who places limit orders into the limit-order book and waits for these orders to be filled by market orders from other traders. We assume these models do not exhibit market impact. However, given we await market orders from other participants to fill our limit orders a new risk is borne: execution risk.

We begin with an extension of Guéant et al. (2012b) who through the use of an exponential utility, standard Brownian motion, and an absolute decay parameter were able to cleverly build symmetry into their model which significantly reduced the complexity. Our model consists of geometric Brownian motion (and mean-reverting processes) for the asset price, a proportional control parameter (the additional amount we ask for the asset), and a proportional decay parameter, implying that the symmetry found in Guéant et al. (2012b) no longer exists. This novel combination results in asset-dependent trading strategies, which to our knowledge is a unique concept in this framework of literature. Detailed asymptotic analyses, coupled with advanced numerical techniques (informing the asymptotics) are exploited to extract the relevant dynamics, before looking at further extensions using similar methods.

We examine our above mentioned framework, as well as that of Guéant et al. (2012b), for a trader who has a basket of correlated assets to liquidate. This leads to a higher-dimensional model which increases the complexity of both numerically solving the problem and asymptotically examining it. The solutions we present are of interest, and comparable with Markowitz portfolio theory. We return to our framework of a single underlying and consider four extensions: a stochastic volatility model which results in an added dimension to the problem, a constrained optimisation problem in which the control has an explicit lower bound, changing the exponential intensity to a power intensity which results in a reformulation as a singular stochastic control problem, and allowing the trader to trade using both market orders and limit orders resulting in a free-boundary problem.

We complete the study with an empirical analysis using limit-order book data which contains multiple levels of the book. This involves a novel calibration of the intensity functions which represent the limit-order book, before backtesting and analysing the performance of the strategies.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Prof. Peter Duck and Dr Paul Johnson, for their guidance, support, and encouragement throughout this period of research. Both shared their wealth of knowledge with me through insightful discussions, and I am extremely grateful for their invaluable assistance, without which this thesis may never have been completed.

Many thanks to other members of the Mathematical Finance group, including Prof. Syd Howell and Dr Geoff Evatt, both for their comments and advice among this journey. I would like to express my sincere appreciation to my examiners Dr Florian Theil and Prof. Sergei Fedotov for their insightful comments and suggestions.

This thesis would not have been the enjoyable experience it was if I did not share an office with an amazing group of friends. Many thanks to Andreas, Mishari, Mingliang, Wang, Hugo, Javier, Pui Wah, Angeliki, and Eduardo.

My thanks to the administrative and IT staff for their support. I would like to thank the School of Mathematics, University of Manchester for their financial support throughout my PhD, both on a monthly basis and for numerous conferences. Thanks to SIAM, IPAM and the Oxford-Man Institute for their support for various conferences.

Throughout my time in Manchester I have made some great friends both in and out of the department. There are too many to name but I'm sure you know who you. You have all made my time here in Manchester an enjoyable one. Thanks also to the friends back home and further abroad, through phone calls and messages you have helped with some long nights in the office.

Lastly, I would like to thank my family. They have always been there for me through many Skype conversations, phone calls, emails and text messages. Without their infinite support and love, I would have never accomplished this journey. To my parents, it seems only right that I dedicate this thesis to you.

Chapter 1

Introduction

1.1 Context of problem

This study focuses on one of the main services that a *sell-side* firm provides to investors, namely in providing *liquidity*. Market liquidity refers to the extent to which a market allows assets to be bought and sold at stable prices. Sell-side firms assist with helping investors access liquidity by two mediums:

- Brokers: Act on behalf of investors and consume liquidity in the market. The revenue generated is through commission.
- Market makers: Provide liquidity in the market by simultaneously buying and selling assets. Revenue is generated when assets are bought at a lower price and sold at a higher price.

Advances in technology have resulted in trading becoming essentially electronic through the use of automated strategies, known as algorithmic trading.

Algorithmic trading has exploded in recent years, with reports of 50-77% of trading volume in the US coming from computerised algorithms, see SEC (2010). Various algorithms have been investigated extensively by both academics and institutional traders; these algorithms focus on being the most profitable, the least risky, or a trade-off between the two. In this thesis we focus on algorithmic trading from the point of view of a sell-side firm, that is providing liquidity to an investor.

We shall now give a brief introduction to how equity markets work, and outline the need for quantitative methods as they progress towards a more electronic environment.

1.1.1 Market microstructure

Historically markets were centralised in primary exchanges (London Stock Exchange, New York Stock Exchange, etc.). With the advances in both technology and competition, markets have become fragmented. This means participants now have a choice of exchanges, each with its own matching rules, fee structure and transparency. The three main types of exchanges are *primary exchanges, multilateral trading facilities* and *dark pools*.

Primary exchanges provide pre-trade transparency and serve as a reference for prices for other exchanges. Multilateral trading facilities compete directly with primary exchanges, but have unique attributes characterised by their fee structure or latency regulations. One such fee structure is known as the *maker-taker* fee structure. Its basic structure gives a transaction rebate to market makers who provide liquidity (the makers); and charges a transaction fee to customers who take liquidity out of the market (the takers). Dark pool is an ominous-sounding term for private exchanges or forums for trading securities. Unlike public exchanges, dark pools are not accessible by public investors and do not provide pre-trade transparency. In dark pools, unlike other exchanges, investors can indicate their interest to trade with other investors in the dark pool, which is not publicly available prior to the trade occurring. This allows large institutions to offload positions without being gamed by high-frequency traders, among others. The trade is known only to the two parties involved until it is actually executed, thus the public is unaware of the shift in supply and demand of the large players. Primarily, the price at which these transactions occur is the mid-price of the primary exchange as this would be a reference for the 'fair' price at which the stock is currently trading.

In this thesis we consider a *lit* venue, which could be a primary exchange or multilateral trading facility. We do not consider dark pools as they are not available to all investors and thus inclusion would limit the applicability of our models to those with dark pools available to them.

1.1.2 Auction mechanisms

For primary exchanges, two main auction phases exist during the trading day: fixing auctions and continuous auctions. Fixing auctions have orders that are matched after being accumulated on a book during a certain period, which usually happens at the beginning and/or end of the day. This determines opening and closing prices. Continuous auctions have orders that are matched continuously as they appear on the order book. Most of the volume is traded during the continuous phase. In this study we will consider only continuous auctions.

During the continuous auction there is a ledger of unexecuted trades known as the limit-order book. There are two primary methods to interact with this book, that is through the use of market orders or limit orders.

- Market orders: Orders in which a quantity (to buy or sell) are specified to be immediately executed. In this case the trader consumes liquidity at the (best) available price.

- Limit orders: Orders in which a quantity and the price to trade are specified. These orders wait in the book until a market order arrives as a counterparty at this price, or until they are cancelled.

The order book is spilt into a bid side and an ask side. The bid side contains limit orders which are to buy at a specified price, and are filled by market orders looking to sell. Likewise, the ask side contains limit orders which are to sell at a specified price, and filled by market orders looking to buy. The highest proposed bid price is known as the best-bid price and the lowest proposed ask price is known as the best-ask price. The difference between the best bid and best ask is known as the spread, and the average of the best bid and best ask is known as the mid-price.

The limit-order book contains orders at the best-bid price, the best-ask price, and all orders lower than the best-bid and higher than the best-ask. All this information is available to the trader when he comes to the market. This is represented graphically in figure 1.1[1]. It can be seen that it not only contains the prices, but also the number of assets available at each level. This is known as the *market depth* and is an important concept when considering algorithmic strategies. In equity markets orders are filled

[1]Taken from www.maths.ox.ac.uk/groups/ociam/research/finance

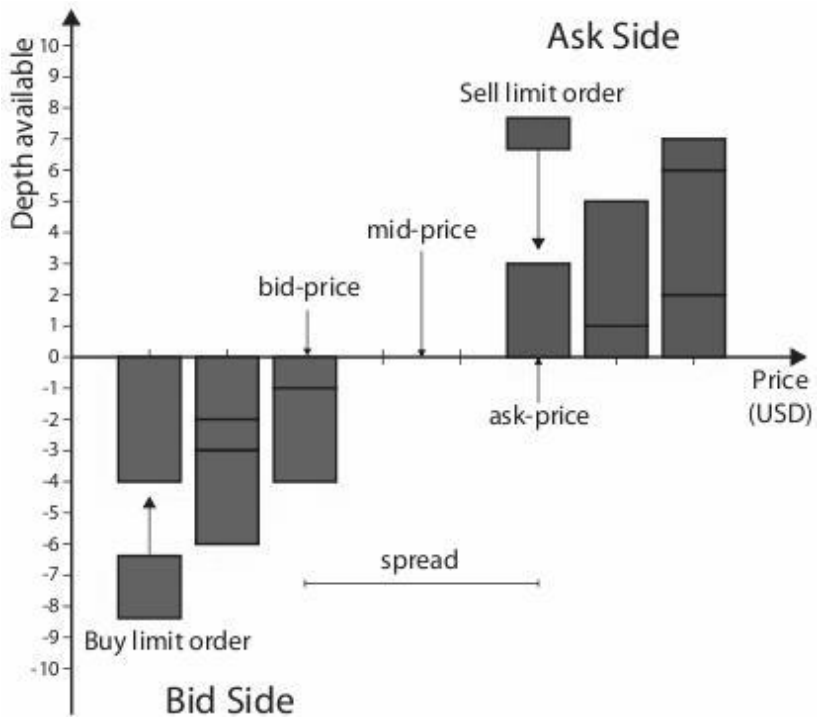


Figure 1.1: Graphical representation of a limit-order book.

by price priority, and for orders at the same price, by time priority. This is in contrast to some asset classes, for example interest rate products, which have a pro-rata limit-order book model. In this thesis we consider only the price/time priority structure.

Orders, limit and market, arrive to the book randomly and thus the best ask, best bid and mid-price change randomly. These changes are discrete as equity markets have a minimal unit of price, known as the tick size.

The use of market orders is known as aggressive trading, while the use of limit orders is known as passive trading. Aggressive trading exists when a trader buys or sells a large quantity of an asset which requires immediate execution, unbalancing supply and demand and thus having a feedback effect on the asset price, known as *market impact*. This is in contrast to passive trading where a dealer places his order in the limit-order book and waits for an opposite aggressive trade to fill his order.

The majority of this thesis focuses on a passive trader who liquidates using only limit orders. Although this is extended in chapter 8 to a trader who can trade using both passive and aggressive methods, i.e. using limit orders and market orders.

1.2 A random walk down Wall Street

Since Black and Scholes (1973) and Merton (1973) published their seminal papers on the pricing of options, the area of mathematical finance has grown enormously as their work showed that the models developed by mathematicians could be used on Wall Street. However, this image was first damaged with the downfall of Long Term Capital Management, a hedge fund in which Merton and Scholes were on the board of directors, and second with the credit crisis of 2008[2].

Currently the use of mathematics to develop trading strategies, especially in high-frequency trading, is under scrutiny. It is argued that trading has become a race with the winners being those with the fastest computers, or with servers closest to the exchange. It is also argued it causes market instability and, since the flash crash (see Easley et al., 2011), is a source of controversy. The *flash crash* was a stock market crash on May 6, 2010 which lasted approximately 36 minutes during which the Dow Jones Industrial Average had its biggest intra-day point drop, plunging 998.5 points (about 9%), mostly within minutes, only to recover a large part of the loss (see figure 1.2). However there are counter arguments for the use of algorithmic trading given they act as market makers and thus provide liquidity, even in times of a depressed economy (see Rijper et al., 2011). Empirical studies (e.g., Hendershott et al. (2011) and Menkveld (2013)) and a recent review (Grillet-Aubert, 2010) from AMF, the French regulator, point out this renewed competition among liquidity providers causes reduced market spreads, and therefore reduced indirect costs for final investors. For an interesting and easy read on high-frequency trading, including the flash crash, see Lewis (2014).

Since the financial crisis the number of banks focusing on derivatives pricing, and the size of the teams within those banks, has shrunk dramatically. Prior to the financial crisis, many exotic financial instruments were sold which were not completely hedged due to a lack of understanding of the full complexity of the models, which also led to mispricing. Post-crash, much stricter regulations are now in place for investment banks, e.g. Dodd-Frank law and Basel Accords, and as such there has been

[2]A quote from the Financial Times on March 20th, 2009 reads "Markets + Maths = mayhem". This is in relation to a report wrote by Lord Turner, former Chairman of the Financial Services Authority, in which he blamed "misplaced reliance on sophisticated math" easing money managers into a false sense of security about the risks they were taking (see Turner, 2009). A statement like this did not go without rebuttals and the author recommends Johnson (2009) and Thomas (2008) to the interested reader.

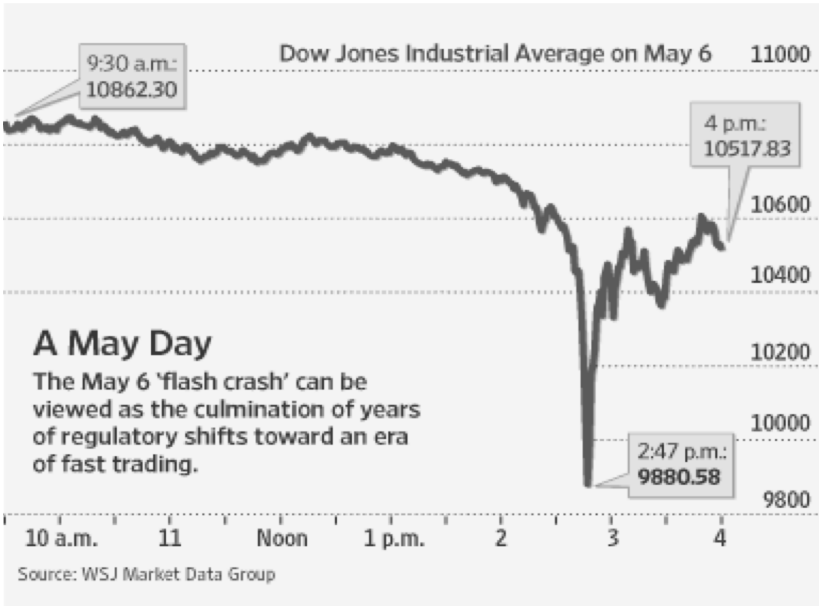


Figure 1.2: Dow Jones Industrial Average index on May 6, 2010. This was the day the *flash crash* occurred in which the market dropped over 9% in a few minutes only to recover a few minutes later.

a shift away from derivatives pricing and more towards either algorithmic trading and limit-order book modelling, or market-risk and credit-risk modelling. One of the main consequences of the evolution in trading automation is the increasing need for a quantitative approach to its analysis and optimisation. This thesis is thus using mathematical methods applied to an area which is currently very topical in mathematical finance, with academics and institutions alike.

1.3 Outline and contribution

This thesis uses a combined approach of asymptotic methods and accurate numerical techniques to examine the problem of optimal execution for a one-sided market maker. Using asymptotic and numerical methods we are able to contribute to knowledge and extend the current literature, notably Guéant et al. (2012b), which was originally derived from Avellaneda and Stoikov (2008), by introducing more general diffusion processes and modelling assumptions, which arguably result in more realistic solutions, as we will see in the coming chapters.

We begin in chapter 2 by giving an introduction to some of the relevant mathematics needed for an understanding of this thesis, including stochastic calculus, optimal

control theory, numerical methods and asymptotic analysis.

Chapter 3 is a detailed review of the optimal trading literature and a discussion of the contribution this work will make to it. We provide a detailed analysis of the framework of Guéant et al. (2012b), as it will be central in all of the following chapters in this thesis. We also contribute to the work of Guéant et al. (2012b) by showing an analytic solution can be derived from the differential-difference equation which they solve numerically.

Chapter 4 extends the current literature by introducing general diffusion processes for the asset price, as well as introducing a (more realistic) decay rate for the limit-order book which is proportional to the asset price, rather than constant for all asset prices which is the case in the base framework to which we are developing. We examine this problem both asymptotically and numerically, with each approach confirming and informing the other. The trading strategies are asset-price dependent, which to our knowledge is a unique concept in the passive optimal trading literature at a tactical level. In the steady-state limit we find a parameter constraint which is a necessary condition for a solution to exist.

The original framework of Guéant et al. (2012b) uses standard Brownian motion for the diffusion process for the asset price. In chapter 5 we extend the original framework to trading with multiple correlated underlying assets, examining it analytically in several limits, complementing the analysis with numerical methods, and discussing the resultant trading strategies. The resultant trading strategies relate to Modern Portfolio Theory (see Markowitz, 1959) due to the introduction of the correlation.

Chapter 6 develops chapter 4 by examining the case of multiple underlyings (which are driven by correlated geometric Brownian motions). This leads to a higher-dimensional model which increases the complexity of both numerically solving it and asymptotically examining it. We perform a thorough comparative analysis of the trading strategies found using numerical methods, before turning to analytic (asymptotic) solutions. In certain limits the multiple underlying problem is reduced to the single asset case and thus we can refer back to the single asset case for reassurance in our methods. We find the steady-state regime to be of particular interest from an analytic analysis perspective, and we perform a detailed study of the solution.

Chapter 7 extends the work of chapter 4 by replacing the constant volatility with

stochastic volatility, which is often seen as a more realistic assumption. We solve the system numerically, examine the financial implications of the results, before (again) asymptotically examining the system in several limits. The parameter constraint found in chapter 4, which was necessary for a solution to exist, was based on the (constant) volatility and thus examining the existence of a similar constraint under stochastic volatility is of interest.

Chapter 8 also extends the work of chapter 4 in three independent directions. We first consider an optimisation problem which constrains the admissible strategy space for the control, under the exponential intensity function originally proposed. Next we examine more general intensity functions for the filling process of orders, namely a power intensity function. Under a power intensity function a singular stochastic control problem results and we must reformulate the problem accordingly. Lastly, we follow a similar framework to chapter 4 but allow the trader to trade using both limit orders and market orders. The majority of the thesis examines a passive trader who can use only limit orders to trade. Allowing the trader to trade using both limit orders and market orders results in a free-boundary type problem, similar to that found with American options, as using a market order can be viewed in a similar light to early exercising an option.

In the chapters thus far we gain an understanding of the underlying problem using asymptotic and numerical methods. The parameter values used in the numerical solutions were exaggerated so as to give the reader insight into the properties and behaviour of the solutions and trading strategies. In chapter 9 we use empirical limit-order book data to calibrate and backtest some of the models discussed. Although the data is limited, due to the difficulty in obtaining limit-order book data (as financial institutions seek to protect their 'trade secrets'), this chapter gives an insight into how such modelling could be done with 'real-world' data, and an initial analysis of the performance of the models developed. We also discuss calibration methods used by other authors for a similar class of problems, and outline the novelty of the method we propose.

We conclude in chapter 10 and discuss some future potential work and research that may be of interest to undertake in this area.

Chapter 2

Background Material

In this chapter we discuss some of the relevant theory suggested for an understanding of this thesis. We begin by introducing some of the stochastic processes we use throughout and their application to financial mathematics. We follow this with a discussion of utility functions, what they are and why they are used. Next we discuss optimisation, in particular optimal control and some of the associated theory. We then introduce the methods we use in this thesis, introducing asymptotic methods and finishing with a discussion on the relevant numerical methods.

2.1 Stochastic processes

As we model in continuous time throughout this study, we shall focus on stochastic calculus in continuous time only, and follow Shreve (2004a); Zukerman (2013); Cont and Tankov (2004). For the interested reader we recommend Shreve (2004b) as a textbook on stochastic calculus in discrete time with applications in finance.

Let us first define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration, $(\mathcal{F}_t, t \in [0, T])$, where Ω denotes a continuous sample space. We define the σ-algebra $\mathcal{F}$ on Ω as a family of subsets with the following characteristics

- $\emptyset \in \mathcal{F}$,
- $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$ where $F^C = \Omega \setminus F$ is the complement of F in Ω,
- $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow A = \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that

- $\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = 1,$
- if $A_1, A_2, \cdots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ are disjoint subsets then

$$\mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Before continuing our discussion on stochastic processes, we will introduce two important definitions that are used throughout the financial mathematics literature.

Definition: Martingale *A stochastic process $X(t) : t \geq 0$ is a martingale if $X(t)$ is $\mathcal{F}_t$-measurable, $\mathbb{E}[|X(t)|] < \infty$ and*

$$\mathbb{E}[X(t)|\mathcal{F}_s] = X(s) \quad \text{for} \quad s < t.$$

Martingales are stochastic processes where the expectation of a future value given the current information is simply the current value. We define a semimartingale as a stochastic process that can be decomposed into the summation of martingales and adapted processes of finite variation. We next consider an important property used in mathematical finance which is the Markov property:

Definition: Markov Process *A $\mathcal{F}_t$-measurable process $X(t) : t \geq 0$ is Markov iff*

$$\mathbb{P}(X(t+h) \in B|\mathcal{F}_t) = \mathbb{P}(X(t+h) \in B|X(t)) \quad \forall \ t, h > 0$$

where $B \in \mathcal{B}(\mathbb{R})$ is a Borel set, which is the smallest σ-algebra containing all open sets.

Simply stated, a stochastic process has the Markov property if the conditional probability distribution of future states of the process depends only upon the present state, and as such is memoryless. The Markov property is one that is assumed in much of the mathematical finance literature. It relaxes a lot of the complexities that otherwise would be present and very difficult to work with.

2.1.1 Wiener process

We will now introduce a simple stochastic process known as a Wiener Process (also know as a Brownian motion). In 1827 the Scottish botanist Robert Brown (see Brown, 1828) gave his name to the random motion of small particles in a liquid. However, it was Wiener (1923) who developed a rigorous theory for Brownian motion, including proving existence.

Definition: Brownian motion *An adapted stochastic process* $(W(t) : t \geq 0)$ *is called a* $\mathbb{P}$*-Brownian motion, under the probability measure* $\mathbb{P}$*, if:*

- $W(0) = 0,\ \mathbb{P} - a.s.$
- W *has independent increments*
- *Increments are normally distributed*
- W *has continuous trajectories.*

A Brownian motion $W(t) : t \geq 0$ has mean zero and variance t. A stochastic process $X(t) = W(t)$ is Markovian and is a martingale. Brownian motion (either geometric or standard), along with a drift term, are the most common stochastic processes used for modelling asset values.

2.1.2 Point process

A point process is a sequence of events that occur at increasing random points in time t_i, where $i = 1, 2, \ldots,$ with $t_{i+1} > t_i$. A point process can be defined by its corresponding *counting process* $N(t), t \geq 0$ where $N(t)$ is the number of events within $[0, t)$. The following properties hold for the counting process $N(t)$:

1. $N(0) = 0$
2. $N(t) \in \mathbb{Z}$
3. If $t > s$ then $N(t) \geq N(s)$ and $N(t) - N(s)$ is the number of occurrences within the period $[s, t)$

4. Orderliness: The probability that two of more events happen at once is negligible;

$$\lim_{\Delta t \to 0} \mathbb{P}\left(N\left(t+\Delta t\right) - N\left(t\right) > 1 | N\left(t+\Delta t\right) - N\left(t\right) \geq 1\right) = 0$$

Two well-known processes that belong to the class of discrete-space point processes are the discrete-time *Bernoulli process* and the continuous-time *Poisson process*. In this thesis we use the latter process, particularly of the inhomogeneous (non-constant intensity) kind, and thus will discuss that next. The interested reader should see Zukerman (2013) for information on the Bernoulli process.

2.1.3 Poisson process

A Poisson process models jumps. It can be used to model processes such as queues or sales, or can be used in asset pricing to model jumps in price. In this thesis it is used to model sales of assets, also referred to as *trades*.

Definition: Poisson process *Let* $(\tau_i)_{i\geq 0}$ *be a sequence of independent exponentially distributed random variables with rate (intensity) parameter* $\lambda(t)$ *and* $T_n = \sum_{i=1}^{n} \tau_i$. *The process* $(N(t), t \geq 0)$ *defined by*

$$N(t) = \sum_{n\geq 1} \mathbb{1}_{t \geq T_n}$$

is a Poisson process.

An inhomogeneous Poisson process is a counting process and the following must hold true:

- $N(t)$ is finite $\mathbb{P}$-a.s.
- $N(t)$ follows a Poisson distribution with parameter $m(s,t) = \int_s^t \lambda(u)\, du$:

$$\mathbb{P}\left(N\left(t\right) - N\left(s\right) = n\right) = e^{-m(s,t)} \frac{m\left(s,t\right)^n}{n!},$$

- $N(t)$ has independent increments,
- $N(t)$ is Markovian.

It is inhomogeneous as the intensity, $\lambda(t)$, is non-constant. We have that $\mathbb{E}[N(t)] = m(0,t)$. We can define a compensating Poisson process as $\tilde{N}(t) = N(t) - m(0,t)$, which is a martingale with $\mathbb{E}[\tilde{N}(t)] = 0$.

The orderliness property of the Poisson process leads to the following small interval conditions:

1. $\mathbb{P}(N(t+\Delta t) - N(t) = 0) = 1 - \lambda(t)\Delta t + O(\Delta t^2)$,
2. $\mathbb{P}(N(t+\Delta t) - N(t) = 1) = \lambda(t)\Delta t + O(\Delta t^2)$,
3. $\mathbb{P}(N(t+\Delta t) - N(t) \geq 2) = O(\Delta t^2)$.

Alternatively one can express the Poisson process in the form

$$dN(t) = \lim_{\Delta t \to 0} N(t+\Delta t) - N(t)$$

which has dynamics

$$dN(t) = \begin{cases} 0 & \text{with probability } 1 - \lambda(t)\Delta t + O(\Delta t^2), \\ 1 & \text{with probability } \lambda(t)\Delta t + O(\Delta t^2). \end{cases}$$

Therefore, for the case of two uncorrelated Poisson processes, $N_1(t)$ and $N_2(t)$, the probability that a jump occurs in both at the same instant is $O(\Delta t^2)$ and is therefore negligible as $\Delta t \to 0$. In this thesis we consider uncorrelated jump processes and thus neglect the probability of a jump occurring in both processes at the same instantaneous time.

2.1.4 Itô's formula for jump-diffusion processes

We shall now extend our discussion of stochastic processes to Itô processes.

Definition: Itô process *A stochastic process* $(X(t) : t \in [0,T])$ *is called an Itô process, if it is of the form*

$$X(t) = X(0) + \int_0^t \gamma(X,s)\,ds + \int_0^t \phi(X,s)\,dW(s), \tag{2.1}$$

where $(\gamma(X,t) : t \in [0,T])$ *and* $(\phi(X,t) : t \in [0,T])$ *are adapted stochastic processes satisfying linear growth and Lipschitz conditions.*

Often for an Itô process of the form of (2.1) the shorthand notation:

$$dX(t) = \gamma(X,t)\,dt + \phi(X,t)\,dW(t),$$

is used. In this thesis Itô processes are primarily used for the modelling of asset prices.

Let $X(t)$ be an n-dimensional jump-diffusion process such that

$$X_i(t) = X_i(0) + \int_0^t \gamma_i(X_i,s)\,ds + \int_0^t \phi_i(X_i,s)\,dW_i(s) + \int_0^t \Delta X_i dN_i(t), \qquad (2.2)$$

for $i = 1,\ldots,n$. $(\gamma_i(X_i,t) : t \in [0,T])$ and $(\phi_i(X_i,t) : t \in [0,T])$ are adapted stochastic processes and $\Delta X_i = X_i(t) - X_i(t-)$. We assume a correlation between the Brownian motions with $\mathbb{E}[W_i(t)\,W_j(t)] = \rho_{ij}t$ with $\rho_{ij} = \rho_{ji}$ and $\rho_{ii} = 1$, and as discussed above we assume the jump components are uncorrelated, both with the other jump components and with the Brownian motions.

Definition: Itô formula (lemma) For any $C^{1,2}$ function $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}$, i.e. continuous first derivative in time and first and second derivative in space, the process $Y(t) = f(t, X(t))$ can be represented as:

$$\begin{aligned} f(t,X(t)) - f(0,X(0)) = &\int_0^t \frac{\partial f}{\partial s}(s,X(s))\,ds + \int_0^t \sum_{i=1}^n \frac{\partial f}{\partial X_i}(s,X(s-))\,dX_i^c(s) \\ &+ \frac{1}{2}\int_0^t \sum_{i,j=1}^n \frac{\partial f}{\partial X_{i,j}}(s,X(s-))\,d[X_i,X_j]^c(s) \qquad (2.3) \\ &+ \int_0^t \sum_{i=1}^n \left(f(s,X_i(s-)+\Delta X_i) - f(s,X_i(s-))\right) dN_i(s), \end{aligned}$$

where $[X_i,X_j]^c(s)$ is the quadratic (co)variation of the purely continuous part of X_i, X_i^c, which is defined as

$$[X_i,X_j]^c(t) = X_i^c(t)\,X_j^c(t) - X_i^c(0)\,X_j^c(0) - \int_0^t X_i^c(s)\,dX_j^c(s) - \int_0^t X_j^c(s)\,dX_i^c(s).$$

We have just discussed some of the fundamental elements in stochastic calculus which are needed for understanding of this thesis. If the reader would like a more thorough understanding the author recommends Shreve (2004a) for a general textbook on stochastic calculus in finance and Cont and Tankov (2004) for more of an insight into the applications of stochastic calculus in finance, with emphasis on jump processes.

2.2 Diffusion processes for asset prices

In this thesis we build on previous frameworks from the optimal trading literature by introducing diffusion processes that are used in other areas of mathematical finance, and are arguably more realistic from a modelling point of view, but, prior to this study, were yet to be implemented in the framework we examine. We shall introduce some of the more common diffusion processes and discuss their properties. Before beginning, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0, T])$. Unless specified, we assume that all stochastic processes in this section are $\mathcal{F}_t$ measurable.

2.2.1 Standard Brownian motion

Standard Brownian motion was the first stochastic process implemented in financial mathematics. The use of standard Brownian motion for asset pricing and option valuation was introduced by Bachelier (1900). His thesis, which discussed the use of Brownian motion to evaluate stock options, is historically the first paper to use advanced mathematics in the study of finance. As such, Bachelier is considered a pioneer in the study of financial mathematics and stochastic processes.

An asset $S(t)$ following standard Brownian motion with drift is defined as

$$dS(t) = \mu dt + \sigma dW(t), \tag{2.4}$$

where μ is the drift, σ is the volatility and $W(t)$ is a Wiener process. Equation (2.4) has solution

$$S(t) = S(0) + \mu t + \sigma W(t). \tag{2.5}$$

Equation (2.5) implies that $S(t)$ is normally distributed such that:

$$\mathbb{E}[S(t)] = S(0) + \mu t, \tag{2.6}$$

$$Var[S(t)] = \sigma^2 t. \tag{2.7}$$

It is clear from (2.5) that an asset following standard Brownian motion can obtain negative values, a property which is highly criticised from a financial perspective.

2.2.2 Geometric Brownian motion

Geometric Brownian Motion (GBM) gained popularity in mathematical finance literature following the seminal work of Black, Scholes and Merton (see Black and Scholes,

1973; Merton, 1973) and is now the dominant diffusion process in mathematical finance literature. A geometric Brownian motion process is defined for positive values, just like real stock prices and is therefore often favoured. Empirical evidence indicates that asset returns are normally distributed, rather than the asset prices themselves. This was first suggested by Osborne (1959) and first implemented in financial models by Samuelson (1965). Osborne (1959) argues a geometric Brownian motion process shows the same kind of 'roughness' in its paths as we see in real stock prices.

For an asset following geometric Brownian motion the diffusion process is defined as:

$$dS(t) = \mu S(t)\, dt + \sigma S(t)\, dW(t). \tag{2.8}$$

with μ the constant relative drift and σ the constant relative velocity. For a geometric Brownian motion process the returns are normally distributed and the asset price is log-normally distributed. This transformation is made using Itô's lemma. Under this transformation we find

$$S(t) = S(0)\, e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)} \tag{2.9}$$

which has the following properties:

$$\mathbb{E}[S(t)] = S(0)\, e^{\mu t} \tag{2.10}$$

$$Var[S(t)] = S(0)^2\, e^{2\mu t}\left(e^{\sigma^2 t} - 1\right) \tag{2.11}$$

Despite its general acceptance, geometric Brownian motion has received some criticism. In particular it falls short of reality because of the following: in real stock prices, volatility changes over time (possibly stochastically), but in geometric Brownian motion, volatility is assumed constant, and, returns are usually not normally distributed. Real stock returns have higher kurtosis, or 'fatter tails', which means there is a higher chance of large price changes. In addition, returns have negative skewness. The former criticism has led to several extensions of geometric Brownian motion. One assumes a non-constant, but deterministic function of stock price and time for the volatility. The second assumes a stochastic process for the volatility, the most well-known being the Heston (1993) model which assumes a mean-reverting process for the volatility, namely the Cox et al. (1985) process, which we discuss next.

2.2.3 Cox-Ingersoll-Ross (CIR) model

It has been argued that mean-reverting processes are more suitable for asset pricing, see for example Metcalf and Hassett (1995), than the alternative geometric Brownian motion that is typically used. Lund (1993) argues strongly against the use of geometric Brownian motion as a price process as it is unbounded above and as such it has the potential for infinite growth in time. Empirical studies have shown that mean-reverting processes are suitable for modelling prices of common stocks (see Poterba and Summers, 1988), futures (see Schwartz, 1997), currencies (see Cheung and Lai, 1994), interest rates (see Cox et al., 1985) and commodities (see Linetsky, 2004; Skorodumov, 2008), to name a few.

The CIR model was first suggested by Cox et al. (1985) as an extension to the Vasicek (1977) model. The attractiveness of the CIR model is its avoidance of the possibility of negative values, under certain regimes, which we shortly discuss. It was originally suggested for the modelling of interest rates but has since been suggested for other models such as commodities (see Linetsky, 2004). In our modelling we shall use it as the underlying process for the asset in section 4.4, as well as the process for the volatility when examining a stochastic volatility model in chapter 7.

An asset $S(t)$ following the CIR process solves the follow SDE

$$dS(t) = \kappa\left(\bar{S} - S(t)\right)dt + \sigma\sqrt{S(t)}dW(t), \tag{2.12}$$

with κ as the mean-reversion speed, $\bar{S}$ is the long-run mean asset price and $\sigma\sqrt{S(t)}$ is the absolute volatility. This process has an expected value and variance

$$\mathbb{E}[S(t)] = S(0)e^{-\kappa t} + \bar{S}\left(1 - e^{-\kappa t}\right), \tag{2.13}$$

$$Var[S(t)] = S(0)\frac{\sigma^2}{\kappa}\left(e^{-\kappa t} - e^{-2\kappa t}\right) + \frac{\bar{S}\sigma^2}{2\kappa}\left(1 - e^{-2\kappa t}\right)^2. \tag{2.14}$$

The CIR process has a conditional non-central chi-squared distribution. The boundary behaviour of (2.12) has been studied in great detail by Feller (1951) in which he found that (2.12) is defined on a positive domain if

$$2\kappa\bar{S} \geq \sigma^2,$$

else it is defined on a non-negative domain. There are a number of other mean-reverting processes used in mathematical finance, however they are not examined in

this thesis. These include the Uhlenbeck and Ornstein (1930) process

$$dS(t) = \kappa\left(\bar{S} - S(t)\right)dt + \sigma dW(t),$$

the Dixit and Pindyck (1994) process

$$dS(t) = \kappa\left(\bar{S} - S(t)\right)S(t)\,dt + \sigma S(t)\,dW(t),$$

and the Schwartz (1997) process

$$dS(t) = \kappa\left(\log\left(\bar{S}\right) - \log\left((t)\right)\right)dt + \sigma S(t)\,dW(t),$$

to name but a few. The author recommends the named references for further reading on these mean-reverting processes.

2.3 Utility functions

Utility functions represent preferences of some good or service in relation to another good or service. We define a utility function $U : \mathbb{R} \to \mathbb{R}$ as a concave, increasing, twice differentiable function. The utility function has the following properties: first, it has the property of non-satiation, which means the investor is never satisfied. This means the investor always wants more wealth, W. Mathematically, this property is defined as

$$\frac{\partial U(W)}{\partial W} > 0. \tag{2.15}$$

The second property is that the investor is risk-averse. This means the investor dislikes risk and will sometimes accept a lower expected payoff over a higher expected payoff if the lower expected payoff has less risk. Mathematically, this property is defined as

$$\frac{\partial^2 U(W)}{\partial W^2} < 0. \tag{2.16}$$

We also note that for a risk-averse investor

$$U(\mathbb{E}[W]) > \mathbb{E}[U(W)], \tag{2.17}$$

meaning he would rather except a guaranteed payoff then take a gamble; this will be discussed by example below.

Under the von Neumann-Morgenstern utility theorem, (see von Neumann and Morgenstern, 1947), four axioms of relativity must exist for an investor to have a utility

function. These four axioms are *completeness, transitivity, continuity*, and *independence*.

1. **Completeness:** Consider an investor who has two choices a or b, one of the following must hold: $a \prec b$ (b is preferred over a), $a \succ b$ (a is preferred over b) or $a \sim b$ (the investor is indifferent between a and b).

2. **Transitivity:** If the same investor has a third choice c with $a \prec b$ and $b \prec c$ then $a \prec c$.

3. **Continuity:** If $a \prec b \prec c$ then there exists a $p \in (0,1)$ such that $pa + (1-p)\, c \prec b$.

4. **Independence:** If $a \prec b$ then for any c there exists $p \in (0,1)$ such that $pa + (1-p)\, c \prec pb + (1-p)\, c$.

For a utility function $U(w)$, we define Arrow-Pratt absolute risk-aversion (see Arrow, 1965; Pratt, 1964), as:

$$A(w) = -\frac{U_{WW}(w)}{U_W(w)}, \tag{2.18}$$

where the subscript denotes the derivative. We can also define the relative risk-aversion as

$$R(w) = A(w)\, w = -\frac{wU_{WW}(w)}{U_W(w)}. \tag{2.19}$$

The utility functions we focus on in this thesis are exponential utility functions which take the form of the negative exponential such that

$$U(w) = a - e^{-\gamma w} \tag{2.20}$$

which has risk-aversion parameter γ. The constant a is chosen arbitrarily and is an upper bound for the utility function. Throughout this thesis, unless specified, $a = 0$. Exponential utility functions exhibit Constant Absolute Risk-Aversion (CARA) and are often referred to as CARA utility functions. This is the case as, from (2.18), we see that

$$A(w) = \gamma. \tag{2.21}$$

In option pricing, the exponential utility function has the property of initial wealth independence (see Cont and Tankov, 2004). This means that the 'fair' option price

is chosen independently of the initial wealth and as such two investors with different levels of initial wealth both find the same 'fair' option price if both investors utility is described by an exponential utility function with equal risk-aversion. The same property is present when finding the optimal trading strategy in limit-order books. The optimal sell/buy price is initial wealth independent and this property will be implemented throughout this thesis.

2.3.1 Certainty equivalent

Consider an investor who is risk-averse and consider the possible outcomes for one coin toss. This can be described by the sample space $\Omega = \{\text{heads}, \text{tails}\}$. A real-valued random variable $W : \Omega \to \{0, 100\}$ that models a \$100 payoff for a successful bet on heads is defined as

$$W(\omega) = \begin{cases} \$100, & \text{if } \omega = \text{heads}, \\ \$0, & \text{if } \omega = \text{tails}. \end{cases}$$

Assuming the coin is equally likely to land on either side, then W has a probability mass function $f : \mathbb{R} \to \mathbb{R}^+$ given by:

$$f(w) = \begin{cases} \frac{1}{2}, & \text{if } \omega = \text{heads}, \\ \frac{1}{2}, & \text{if } \omega = \text{tails}, \end{cases}$$

so $\mathbb{E}[W] = \$50$. We have three examples of investors: risk-averse, risk-neutral and risk-seeking. A risk-neutral investor is indifferent in tossing the coin or receiving \$50. A risk-seeking investor would prefer to toss the coin while the risk-averse investor would prefer the \$50 payoff.

Consider now the case where W is no longer modelling a coin toss but is instead modelling the value of a portfolio of assets driven by a stochastic process (which may be a vector of stochastic processes). For a utility function U we define the certainty equivalent, W_C, as the amount to which an investor is indifferent between having for certain or holding the portfolio with random outcome W, such that:

$$U(W_C) = \mathbb{E}[U(W)] \Rightarrow W_C = U^{-1}(\mathbb{E}[U(W)]).$$

W_C is interpreted as the compensation that a risk-averse investor with utility function U requires in order to assume the risk incurred by holding W. An investor who uses

the expected utility as a criterion is then indifferent between holding the random portfolio W or the lump sum W_C. Consider again the case of the coin toss. Under an exponential utility with risk-aversion parameter γ we have the equation for the certainty equivalent as

$$-e^{-\gamma W_C} = -\frac{1}{2}\left(e^{-\gamma 100} + e^{-0}\right) \Rightarrow W_C = -\frac{1}{\gamma}\ln\left(\frac{1}{2}\left(e^{-\gamma 100} + e^{-0}\right)\right),$$

where $\gamma > 0$ determines the degree of risk-aversion. It can be seen in table 2.1 how sensitive this simple example is to the choice of γ, with a relatively large γ corresponding to a high degree of risk-aversion and thus the willingness to accept a much lower compensation. It can also be seen how the risk-aversion constraint (2.17) is satisfied, given all elements in the fourth column are greater than the corresponding elements in the third column.

γ	W_C	$U(W_C) = \mathbb{E}[U(W)]$	$U(\mathbb{E}[W])$
0.01	37.98	-.68399	-.60653
0.05	13.73	-.50337	-.08208
0.10	6.93	-.50003	-.00674
0.20	3.47	-.50000	-4.54 $\times 10^{-5}$
0.50	1.38	-.50000	-1.39 $\times 10^{-11}$

Table 2.1: Certainty equivalent for various values of risk-aversion parameter, γ

2.3.2 St. Petersburg paradox

There have been some arguments for the use of expected utility theory in finance. One such argument arises when attempting to solve the *St. Petersburg paradox*. This paradox is named after Daniel Bernoulli's presentation of the problem and his solution, published in the Commentaries of the Imperial Academy of Science of Saint Petersburg (see Bernoulli, 1954). Consider a game in which a fair coin is tossed until the outcome of a tail occurs. The pot starts at £2 and is doubled every time a head appears. Letting X denote the payoff from this game, the expected value of the payoff is given as

$$\begin{aligned}\mathbb{E}[X] &= \frac{1}{2}\times 2 + \frac{1}{4}\times 4 + \frac{1}{8}\times 8 + \dots \\ &= 1 + 1 + 1 + \dots \\ &= \infty. \end{aligned} \tag{2.22}$$

Therefore, people should pay an infinite amount to play this game. However, the paradox is the discrepancy between what people seem willing to pay to enter the game and the infinite expected value.

A common method for solving the above problem is the use of expected utility theory. Bernoulli (1954) suggests the use of logarithmic utility, such that

$$U(w) = \ln(w),$$

for some positive wealth w. Defining $c > 0$ as the cost of playing the game, the expected utility of the game converges to a finite value and can be expressed as

$$\mathbb{E}[U(X)] = \sum_{k=1}^{\infty} \frac{\left(\ln\left(w + 2^{k-1} - c\right) - ln(w)\right)}{2^k} < \infty. \tag{2.23}$$

Equation (2.23) gives an implicit relationship between the wealth, w, and how much the player should be willing to pay to play, c. For example, with log utility a millionaire should be willing to pay up to £10.94, a person with £1000 should pay up to £5.94, and a person with £2 should pay up to £2.

For a more in-depth read on utility theory the author recommends Föllmer and Schied (2011) and Merton (1990).

2.4 Optimisation

In this section we shall discuss the optimisation techniques used in this thesis, as well as some of the complementary techniques that are used in continuous time optimisation. In this thesis we focus only on continuous time and as such we shall keep our discussion in this section to continuous time optimisation. We must distinguish between static optimisation problems and dynamic optimisation methods. The former consists of solving for a single optimal choice for each control variable, such as the optimal level of output per week or the optimal price to charge for a product, regardless of time. In contrast, dynamic optimisation poses the question of what is the optimal solution for a chosen variable at each point in time over a given time interval $[0, T]$ (with the infinite horizon case of $T = \infty$ being possible). In this thesis, and in this section, we deal with dynamic optimisation only. However, static optimisation problems have previously appeared in the literature of optimal trading, which we discuss in the next chapter.

In dynamic optimisation the objective is to maximise or minimise a functional, rather than a function. Consider the time paths $x : [0, T] \to \mathbb{R}$, which are functions in their own right. Let $J : x \to \mathbb{R}$ represent the associated path value, in which the general notation we use for this mapping is $J(x)$. It should be noted that this fundamentally differs from the composite-function $g(f(t))$ where g is a function of f which in turn is a function of t and thus in the final analysis g is a function of t. For the case of the functional, $x(t)$ represents the time path and as such we should not take J as a function of t, rather that it is a change in the position of the entire x path, as opposed to the change in time, hence the use of our notation $J(x)$.

A simple type of dynamic optimisation problem should contain:

1. A given initial point, $x(0)$, and terminal point, $x(T)$
2. A set of admissible paths $x : [0, T] \to \mathbb{R}$
3. A set of path values, $J(x)$, indicating some performance such as cost, profit etc
4. A specific objective: to choose the optimal path, $x^*(t)$, that either maximises or minimises $J(x^*)$.

There are three predominant approaches used to tackle the problem of dynamic optimisation: *the calculus of variations*, *optimal control theory*, and *dynamic programming*. Although it is dynamic programming we use predominately in this thesis, we shall briefly discuss the other methods and the interconnection between the three. In our discussion, we shall focus on the case of maximisation, as this is what is used throughout the thesis, although extending to minimisation is quite trivial. The notes below are based on Chiang (2000); Øksendal and Sulem (2005); Fleming and Soner (2006); Bardi and Capuzzo-Dolcetta (2008)

2.4.1 Calculus of variations

Consider the objective functional $J(x)$ that we seek to maximise, for which $x : [0, T] \to \mathbb{R}$ is an admissible trajectory (path). The fundamental problem can be stated as:

$$\max J(x) = \int_0^T F(t, x(t), x'(t))\, dt, \tag{2.24}$$

subject to

$$x(0) = x_0,$$
$$x(T) = x_T,$$

where $x_0, x_T \in \mathbb{R}$ are constants and are known *a priori*. The integrand function F depends on the time $t \in [0, T]$, the state $x(t)$ and the direction of the path $x'(t)$. The calculus of variations requires the use of the first and second derivatives and as such we shall restrict the set of admissible paths to those with continuous curves with continuous derivatives, and so assume that the integrand function F is twice differentiable. A smooth path x that is the maximum of $J(x)$ is called a maximal.

For the case of terminal time utility maximisation, the objective function often takes the form:

$$\max J(x) = G(T, x(T)), \tag{2.25}$$

subject to

$$x(0) = x_0,$$
$$x(T) = x_T,$$

where the function G is based on what happens at the terminal time only. It should be noted that it is possible to transform an objective function of the form of (2.25) to that of (2.24) using

$$z(t) \equiv G(t, x(t)) \qquad \text{with initial conditions } z(0) = 0, \tag{2.26}$$

as

$$\int_0^T z'(t)\, dt = z(t)\Big|_0^T = z(T) - z(0) = G(T, x(T)).$$

The basic first-order necessary condition in the calculus of variations is the Euler equation, i.e. it can be shown that the optimal trajectory $x^*(t)$ that satisfies (2.24) also satisfies

$$F_x - \frac{d}{dt}F_{x'} = 0, \tag{2.27}$$

which when expanded gives

$$F_{x'x'}x''(t) + F_{xx'}x'(t) + F_{tx'} - F_x = 0. \tag{2.28}$$

The expanded version shows the Euler equation is a non-linear second-order partial differential equation (PDE), in which we have an initial and terminal boundary condition as given above in (2.24).

Although the calculus of variations is the classical method for tackling problems of dynamic optimisation, it has two drawbacks. The first is that it requires continuously differentiable functions for its applicability. The second is that only interior solutions can be handled; a more modern development that overcomes the latter restriction is found in optimal control theory which we shall discuss next. For more on the calculus of variations the author recommends Chiang (2000).

2.4.2 Optimal control theory

The optimal control formulation of a dynamic optimisation problem focuses on finding the control variables that act as the instrument of optimisation. Whereas in the calculus of variations approach our goal was to determine the optimal path in time for a state variable x, in optimal control theory our goal is to determine the optimal path in time for a control variable, $u(t) \in \mathbb{U}$ for $t \in [0, T]$, where $\mathbb{U} \subset \mathbb{R}$ is some bounded control set, which, as an advantage over the calculus of variations approach, can be a closed bounded set.

The fundamental problem in optimal control theory reads:

$$\max_{u \in \mathbb{U}} J = \int_0^T F(t, x(t), u(t))\, dt + G(x(T)), \tag{2.29}$$

subject to

$$\begin{aligned} & x'(t) = f(t, x(t), u(t)), \\ & x(0) = x_0, \\ & u(t) \in \mathbb{U} \quad \text{for all } t \in [0, T], \end{aligned}$$

where $F : [0, T] \times \mathbb{R} \times \mathbb{U} \to \mathbb{R}$ is the running cost, $G : \mathbb{R} \to \mathbb{R}$ is the terminal time cost, $x_0 \in \mathbb{R}$ and $f : [0, T] \times \mathbb{R} \times \mathbb{U} \to \mathbb{R}$. Relating to the calculus of variations, we notice that $x'(t)$ has been replaced by $u(t)$ inside the cost functionals. Once the optimal control path $u^*(t)$ has been found, we can also find the optimal state path $x^*(t)$. The presence of the control variable u necessitates a linkage between u and x to tell us how u will specifically affect the course taken by the state variable x. This linkage is

given through the constraint equation $x'(t) = f(t, x(t), u(t))$, which is referred to as the equation of motion for the state variable. Note when we set $x'(t) = u(t)$, optimal control theory reduces precisely to the calculus of variations.

It should be noted that a feature of optimal control theory is that the control path need not be continuous in order to be admissible, it need only be piecewise continuous, and as such is allowed jump discontinuities. Also the state path need not be continuous but only piecewise differentiable, allowing for corners or sharp points. These two properties, along with the property of the control set having the ability to be a closed, bounded, convex set, make optimal control theory advantageous over the calculus of variations.

Optimal control and the Bellman principle

As stated in Bellman and Dreyfus (1962):

> *The Principle of Optimality: An optimal policy has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

Consider the fundamental problem (2.29). Let us examine this using the Bellman Principle approach. Let $t \in [0, T]$ and define $J(t, x(t), u(t))$ as the remaining reward between time $[t, T]$ when the control policy is $u : [t, T] \to \mathbb{R}$. This may be written as

$$J(t, x(t), u(t)) = \int_t^T F(s, x(s), u(s))\, ds + G(x(T)). \tag{2.30}$$

We can then define the value function

$$V(t, x(t)) = \max_{u \in \mathbb{U}} J(t, x(t), u(t)). \tag{2.31}$$

Consider the small increment Δt, then applying Bellman's Principle of Optimality to (2.31) results in:

$$V(t, x(t)) = \max_{u \in \mathbb{U}} \Big\{ \int_t^{t+\Delta t} F(\tau, x(\tau), u(\tau))\, d\tau + V(t + \Delta t, x(t) + \Delta x) \Big\}. \tag{2.32}$$

As F is continuous, we can approximate the integral in (2.32) to leading order as $F(t, x(t), u(t))\, \Delta t$ so (2.32) becomes

$$V(t, x(t)) = \max_{u(\tau) \in \mathbb{U}} \Big\{ F(t, x(t), u(t))\, \Delta t + V(t + \Delta t, x(t) + \Delta x) \Big\} + O\left(\Delta t^2\right). \tag{2.33}$$

As we assume that V is a continuously differentiable function, we use a Taylor series expansion of V with respect to Δt and obtain:

$$V\left(t+\Delta T, x\left(t+\Delta t\right)\right)=V\left(t, x\left(t\right)\right)+V_x\left(t, x\left(t\right)\right)x'\left(t\right)\Delta t+V_t\left(t, x\left(t\right)\right)\Delta t+O\left(\Delta t^2\right) \tag{2.34}$$

which when substituted into (2.33), dividing by Δt and taking the limit as $\Delta t \to 0$ we obtain the Hamilton-Jacobi-Bellman (HJB) equation

$$0=\max_{u\in\mathbb{U}}\left\{F\left(t, x\left(t\right), u\left(t\right)\right)+V_x\left(t, x\left(t\right)\right)f\left(t, x\left(t\right), u\left(t\right)\right)+V_t\left(t, x\left(t\right)\right)\right\}, \tag{2.35}$$

remembering from (2.29) that $x'\left(t\right)=f\left(t, x\left(t\right), u\left(t\right)\right)$. Equation (2.35) must be solved backwards in time with some known final condition

$$V\left(T, x\left(T\right)\right)=G\left(x\left(T\right)\right). \tag{2.36}$$

Equation (2.35) is generally non-linear with respect to the value function V. The solution obtained from the HJB equation is equivalent to the solution obtained using Pontryagin's Maximum Principle. Although it is not used directly in this thesis it is arguably the most important result in optimal control theory and therefore we shall discuss it briefly.

Pontryagin's maximum principle

Pontryagin's maximum principle, also known as the maximum principle, is a first-order necessary condition for optimality and was formulated in 1956 by Russian mathematician L.S. Pontryagin and his colleagues (see Pontryagin, 1987), although an American mathematician M.R. Hestenes produced comparable results in a RAND corporation report a few years previously (see Hestenes, 1950).

Consider the system (2.29). The objective is to find the optimal control $u^*\left(t\right)$ (and hence the optimal path $x^*\left(t\right)$) that maximises J subject to the given constraints. The constraints on the system can be implemented by introducing the time-dependent Lagrange multiplier $\lambda\left(t\right)$, with $\lambda\left(t\right)$ termed the *costates* of the system. The vehicle through which the costate variable gains entry into the optimal control problem is through the Hamiltonian H, defined as:

$$H\left(t, x\left(t\right), u\left(t\right), \lambda\left(t\right)\right)\equiv F\left(t, x\left(t\right), u\left(t\right)\right)+\lambda\left(t\right)f\left(t, x\left(t\right), u\left(t\right)\right) \tag{2.37}$$

The costates variable has the equation of motion

$$\frac{d\lambda}{dt} = -\frac{\partial H}{\partial x}, \tag{2.38}$$

while the equation of motion for x is given by

$$\frac{dx}{dt} = -\frac{\partial H}{\partial \lambda}. \tag{2.39}$$

Pontryagin's Maximum Principle states that if the optimal control trajectory $u^*(t)$ maximises J then it also maximises H. As such we have

$$\begin{aligned} u^*(t) &= \arg\max_u H(t, x(t), u(t), \lambda(t)), \\ \frac{d\lambda}{dt} &= -\frac{\partial H}{\partial x}, \\ \frac{dx}{dt} &= -\frac{\partial H}{\partial \lambda}, \end{aligned} \tag{2.40}$$

with H given by (2.37). Depending on the problem we attempt to solve, an additional condition might be required. This condition is known as the *transversality condition*, and it sets the value of λ at time T. This arises as the system (2.29) has a fixed terminal time T but free terminal-state condition $x(T)$. An example of a free terminal state condition could be a trader who is looking to profit as a market maker but with no constraint on the amount of inventory held, x, at the terminal time T. Hence we must set the value of $\lambda(T)$ as

$$\lambda(T) = 0.$$

If the terminal state was to be specified we would replace the transversality condition with the terminal condition

$$x(T) = x_T.$$

For varying terminal conditions and terminal states the value of $\lambda(T)$ has varying values and constraints and the reader is referred to Chiang (2000) for more insight.

Finally, in order to apply the maximum principle, we first must solve the equations of motion (2.40) for a general solution. Using the transversality condition we obtain particular solutions for $\lambda^*(t)$ and $x^*(t)$. We can then evaluate the optimal trajectory $u^*(t)$ that maximises H, and hence J.

2.4.3 Stochastic optimal control

So far we have discussed optimal control in the deterministic setting. In this thesis we deal with stochastic optimal control, as throughout we consider a stochastic setting, rather than a deterministic one.

Consider the problem outlined above in section 2.4.2. In a stochastic setting the state variable, previously x for the deterministic case, is now a stochastic process. We define the stochastic state variable as $X(t) : t \in [0, T]$ as a jump-diffusion of the form

$$X(t) = X(0) + \int_0^t \gamma(X, s)\, ds + \int_0^t \phi(X, s)\, dW(s) + \int_0^s \delta(X, s, u(s))\, dN(s), \quad (2.41)$$

which is a single dimension case of (2.2), and in shorthand notation can be expressed as

$$dX(t) = \gamma(X, t)\, dt + \phi(X, t)\, dW(t) + \delta(X, t, u(t))\, dN(t), \quad (2.42)$$

with $N(t)$ being Poisson distributed with rate $\lambda(u(t))$, where $u(t)$ is the Markov control process. We can conclude there exists a unique solution to (2.42) assuming it satisfies the following conditions

- At most linear growth: There exists a constant C_1 such that

$$|\gamma(x, t)| + |\phi(x, t)| + |\delta(x, t, u(t))\, \lambda(u(t))| \leq C_1 (1 + |x|).$$

- Lipschitz continuity: There exists a constant C_2 such that

$$|\gamma(x, t) - \gamma(y, s)| + |\phi(x, t) - \phi(y, s)| +$$
$$|\delta(x, t, u(t))\, \lambda(u(t)) - \delta(y, s, u(s))\, \lambda(u(s))| \leq C_2 (1 + |x - y| + |t - s|),$$

with

$$\mathbb{E}\Big[|X(t)|\Big] < \infty \quad \forall t.$$

For more on existence and uniqueness of stochastic differential equations see Øksendal (2003).

Given our state variable is now a stochastic process, our objective is to maximise the *expectation* of the reward functional. Define:

$$J(t, X(t), u(t)) = \mathbb{E}\Big[\int_t^T F(s, X(s), u(s))\, ds + G(X(T))\Big], \quad (2.43)$$

which is the total expected reward over time $[t, T]$. Our value function is the maximum of the total expected reward and takes the form:

$$V\left(t, X\left(t\right)\right) = \max_{u \in \mathbb{U}} J\left(t, X\left(t\right), u\left(t\right)\right), \tag{2.44}$$

to which we know a solution exists given the existence and uniqueness shown for the underlying stochastic differential equation. Through the use of Itô calculus and Bellman's Principle of Optimality we can derive an HJB equation for this problem, assuming the following conditions are satisfied

- The reward functional satisfies a polynomial growth condition, i.e. there exists constants $C \in \mathbb{R}^+$ and $k \in \mathbb{Z}^+$ such that

$$\left|J\left(t, x, u\right)\right| \leq C\left(1 + |x|^k\right), \quad \forall\left(t, X\right) \in [0, T] \times \mathbb{R}^+, \text{ and } u \in \mathbb{U}.$$

- The value function $V(t, x) \in C^{1,2}\left([0, T), \mathbb{R}^+\right)$.

Applying Bellman's Principle to (2.44) we have

$$V\left(t, X\left(t\right)\right) = \max_{u \in \mathbb{U}} \mathbb{E}\Big\{ \int_t^{t+h} F\left(\tau, X\left(\tau\right), u\left(\tau\right)\right) d\tau + V\left(t + h, X\left(t + h\right)\right) \Big\}. \tag{2.45}$$

Assume $u\left(s\right) = u$ for some $s \in [t, t + h]$ so

$$V\left(t, X\left(t\right)\right) \geq \mathbb{E}\Big\{ \int_t^{t+h} F\left(\tau, X\left(\tau\right), u\right) d\tau + V\left(t + h, X\left(t + h\right)\right) \Big\}, \tag{2.46}$$

and subtracting $V\left(t, X\left(t\right)\right)$ and dividing by h arrives at

$$0 \geq \frac{1}{h}\mathbb{E}\Big\{ \int_t^{t+h} F\left(\tau, X\left(\tau\right), u\right) d\tau + V\left(t + h, X\left(t + h\right)\right) - V\left(t, X\left(t\right)\right) \Big\}. \tag{2.47}$$

In the limit as $h \to 0$ we have

$$0 \geq \mathbb{E}\Big\{ F\left(t, X\left(t\right), u\right) dt + dV\left(t, X\left(t\right)\right) \Big\}. \tag{2.48}$$

Applying Itô calculus to dV, as discussed in section (2.1.4), we have

$$\begin{aligned} dV\left(t, X\left(t\right)\right) = {} & \frac{\partial V\left(t, X\right)}{\partial t} dt + \frac{\partial V\left(t, X\right)}{\partial X} dX^c\left(t\right) + \frac{1}{2}\frac{\partial^2 V\left(t, X\right)}{\partial X^2} d[X, X]^c\left(t\right) \\ & + \lambda\left(u\right)\left(V\left(t, X + \delta\left(X, t, u\right)\right) - V\left(t, X\right)\right) dt \\ & + \left(V\left(t, X + \delta\left(X, t, u\right)\right) - V\left(t, X\right)\right) d\tilde{N}\left(t\right), \end{aligned} \tag{2.49}$$

with $d[X,X]^c(t) = \phi(X,t)^2\,dt$ and $\tilde{N}(t)$ being a compensating Poisson process. Substituting (2.49) into (2.48) and dividing across by dt we arrive at

$$0 \geq \Big\{ \frac{\partial V(t,X)}{\partial t} + \gamma(X,t)\frac{\partial V(t,X)}{\partial X} + \frac{1}{2}\phi(X,t)^2\frac{\partial^2 V(t,X)}{\partial X^2} + F(t,X,u) + \lambda(u)\left(V(t,X+\delta(X,t,u)) - V(t,X)\right)\Big\}. \tag{2.50}$$

If we were to choose the control optimally we would have chosen $u(s) = u^*$ for $s \in [t, t+h]$ such that (2.50) would be strictly equal to zero, rather than less than or equal to. Therefore choosing the optimal control we have

$$0 = \max_{u\in\mathbb{U}} \Big\{ \frac{\partial V(t,X)}{\partial t} + \gamma(X,t)\frac{\partial V(t,X)}{\partial X} + \frac{1}{2}\phi(X,t)^2\frac{\partial^2 V(t,X)}{\partial X^2} + F(t,X,u) + \lambda(u)\left(V(t,X+\delta(X,t,u)) - V(t,X)\right)\Big\}, \tag{2.51}$$

which is the HJB equation that solves (2.44) with terminal condition:

$$V(T,X) = G(X(T)). \tag{2.52}$$

A verification theorem (see Bremaud, 1981) is used to prove that the solution we obtain from the HJB equation (2.51) is indeed the solution of the optimal control problem (2.44) we wish to solve. In this thesis we will derive the HJB equations we use which may give more of an insight for the reader into the derivation for a particular problem. We also provide verification theorems to show that the solutions we obtain from the HJB equations are equivalent to the solutions from the fundamental problems.

In the case where the smoothness assumptions (continuously differentiable) do not hold, it is not possible to find classical solutions to the HJB equation. In this case one may look towards non-smooth analysis such as viscosity solutions. Although this is outside the scope of this thesis, the author recommends Fleming and Soner (2006); Bardi and Capuzzo-Dolcetta (2008) for further reading. For a more in-depth read on stochastic optimal control the reader is referred Bertsekas et al. (1995) and Øksendal and Sulem (2005).

2.5 Asymptotic methods

When a PDE is to be solved, an analytic solution is generally preferred to a numerical solution; however, analytic solutions are not always available, particularly in high-dimensional non-linear PDEs. We can, however, use perturbation methods to examine

the limiting behaviour of a variable (or small/large parameter behaviour) which may lead to analytic solutions which are confined to specific regions. These methods have been rigorously implemented in the fluid mechanics literature (see Van Dyke, 1964) and are well suited for implementation in this thesis. Other areas of mathematical finance have relied on the use of asymptotic methods to gain a deeper understanding of the problem, for which an analytic solution is not available. These include: Duck et al. (2014) examining the limiting behaviour of a perpetual option with multiple underlyings, Widdicks et al. (2005) applying singular perturbation techniques to the Black-Scholes PDE used to price European, American, and barrier options, Whalley and Wilmott (1997) using asymptotic analysis for option pricing with transaction costs, and Schied and Schöneborn (2009) deriving asymptotic solutions for the optimal scheduling problem under varying utility functions, to name a few. This list is far from exhaustive, but a brief few examples to emphasis the range of topics (and power) in financial mathematics to which asymptotic analysis is used, especially on problems which are either non-linear and/or multi-dimensional. It is non-linear multi-dimensional problems that we examine in this thesis.

2.5.1 Perturbation theory

Perturbation theory comprises the use of mathematical methods for finding an approximate solution to a problem. Perturbation theory expands a solution in a power series involving the variable, the limit of which we are examining. This power series is substituted into the PDE and the coefficients of the series are solved for. The resulting solution becomes an approximation to the actual solution, valid in the limiting region. As defined by Bender and Orszag (1999), the power series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ is said to be asymptotic to the function $f(x)$ as $x \to x_0$ and we write $y(x) \sim \sum_{n=0}^{\infty} a_n (x - x_0)^n$ as $x \to x_0$ if $y(x) - \sum_{n=0}^{N} a_n (x - x_0)^n \ll (x - x_0)^N$ as $x \to x_0$ for every integer N. Thus a power series is asymptotic to a function if the remainder after N terms is much smaller than the last retained term as $x \to x_0$. By this definition a series may be asymptotic to a function without being convergent. Often an asymptotic approximation is required close to boundaries, for example either at zero or plus/minus infinity. For the former case we have $x_0 = 0$ so $y(x) \sim \sum_{n=0}^{\infty} a_n x^n$. The above definition for the latter case of $x \to \infty$ corresponds to $y(x) \sim \sum_{n=0}^{\infty} a_n x^{-n}$. We note that not all

functions can be expanded this way. For example, e^x has power series expansions $e^x \sim \sum_{n=0}^{\infty} x^n$ as $x \to 0$. However the form of $y(x) \sim \sum_{n=0}^{\infty} a_n x^{-n}$ does not hold for $y(x) = e^x$ as $x \to \infty$ because e^x grows more rapidly than any power of x as $x \to \infty$. If $y(x)$ can be expressed as $y(x) \sim \sum_{n=0}^{\infty} a_n (x - x_0)^n$ then the coefficients a_n can be determined uniquely as

$$a_i = \lim_{x \to x_0} \frac{y(x) - \sum_{n=0}^{i-1} a_n (x - x_0)^n}{(x - x_0)^i}.$$

2.5.2 Method of dominant balance

The method of dominant balance is often used to identify those terms in an equation that may be neglected. This is particularly helpful in the case of non-linear equations in which, for some limit, the non-linear term may be neglected such that a power series expansion may be found for the solution. For the method of dominant balance the technique we use consists of dropping all terms that appear small (this may be confirmed numerically) and we approximate the exact equation by an asymptotic relation, which is valid in the particular limit being examined. We should be able to verify that the neglected terms are in fact small by substituting the solution for the asymptotic relation into the neglected term. This method is used throughout this thesis and we refer to it as 'asymptotic balancing'.

For a more detailed discussion on asymptotic methods the author refers the reader to Bender and Orszag (1999); Van Dyke (1964), and for an analysis within the framework of PDEs the author refers the reader to Evans (2010).

2.6 Numerical methods

The vast majority of PDEs, along with many ordinary differential equations (ODEs), do not have analytic solutions and thus require a numerical approach to find solutions. In this section we discuss the main methods we use to solve the PDEs and ODEs discussed in this thesis. We begin by introducing finite-difference methods, in particular direct methods to solve PDEs and ODEs. This is followed by a discussion on iterative methods that are also used. Unless specified for a given method, in this section we follow Smith (1965).

2.6.1 Finite-difference methods

Physicists and applied mathematicians have for years been applying finite-difference methods to PDEs and that has strongly influenced the area of financial mathematics. It is well-known that the Black-Scholes PDE (see Black and Scholes, 1973) has a closed form solution for vanilla European call and put options; however this is not always the case. For more complex problems in financial mathematics, such as American options, exotic options or problems involving multiple underlyings, analytic solutions are generally not available. Merton et al. (1977) were the first authors to suggest applying finite-difference methods in financial mathematics, in this case to price an American put option in which they used the explicit method, and Brennan and Schwartz (1978) demonstrates its link to the trinomial lattice method (see Cox et al., 1979, for more on lattice methods).

There are various finite-difference methods, each with unique characteristics, in which we have a wealth of knowledge on given their previous use in other areas of science and engineering. Those available include the explicit, implicit, Crank-Nicolson (semi-implicit) and alternating-direction-implicit (ADI) schemes. For each we know the stability and convergence and shall discuss the first three below. The ADI scheme is not used in this thesis and we refer the interested reader to Peaceman and Rachford (1955).

Finite-differencing requires discretisation of the problem. A rectangular grid is formed which is usually equally spaced, although this need not always be the case. Let us consider a simple case in which we have one variable in space, S, and one in time, t. We are aiming to solve for the value function $V(\tau, S)$, with $\tau = T - t$, which is a solution of the general PDE which takes the form

$$-\frac{\partial V(\tau,S)}{\partial \tau} + \mu(\tau,S)\frac{\partial V(\tau,S)}{\partial S} + \sigma(\tau,S)\frac{\partial^2 V(\tau,S)}{\partial S^2} + g(\tau,S)V(\tau,S) = 0. \quad (2.53)$$

with some given initial condition.

We divide the S grid into $n+1$ points and the τ grid into $m+1$ points. We may then write (with $S_{min} = 0$ and $t = 0$)

$$\Delta\tau = \frac{T}{m} \qquad \tau_k = k\Delta\tau$$
$$\Delta S = \frac{S_{max}}{n} \qquad S_i = i\Delta S$$

in which S_{max} is an approximation for $S \to \infty$, which is implemented at the boundary. We can write the value function's numerical approximation v as

$$V(\tau_k, S_i) = v_i^k$$

Explicit method

The explicit method is popular in financial literature (Brennan and Schwartz, 1978; Hull and White, 1990) possibly due to its ease of implementation rather than its efficiency or accuracy. For the explicit method we approximate the derivatives as

$$\frac{\partial V}{\partial \tau}(\tau_k, S) = \frac{v_i^{k+1} - v_i^k}{\Delta \tau} + O(\Delta \tau), \tag{2.54}$$

$$\frac{\partial V}{\partial S}(\tau_k, S) = \frac{v_{i+1}^k - v_{i-1}^k}{2\Delta S} + O\left(\Delta S^2\right), \tag{2.55}$$

$$\frac{\partial^2 V}{\partial S^2}(\tau_k, S) = \frac{v_{i+1}^k - 2v_i^k + v_{i-1}^k}{\Delta S^2} + O\left(\Delta S^2\right), \tag{2.56}$$

which are found by a Taylor series expansion on V when small deviations in τ and S, of size $\Delta\tau$ and ΔS respectively, occur. Equation (2.54) is known as a forward difference while (2.55) and (2.56) are known as central differences. From (2.54)-(2.56) we expect the method to have $O(\Delta\tau, \Delta S^2)$ convergence. At time $k+1$, v_i^k is known for all i and as such the system we are solving is effectively a trinomial tree. One problem with this method is its stability. For the case of the general PDE (2.53) we have the constraint

$$\Delta\tau \leq \frac{\Delta S^2}{2\sigma(\tau, S)}$$

which requires the time interval to decrease with a decrease in the number of points in space or an increase in the $\sigma(\tau, S)$ coefficient. Although ΔS convergence is second order, reducing ΔS by half means we must reduce $\Delta\tau$ by a factor of four, thus for four times the accuracy computation time increases eight fold. As such we now turn to the implicit method.

Implicit method

The implicit method is unconditionally stable, although it has the same order of accuracy as the explicit method. For the implicit method we approximate V and the derivatives as

$$\frac{\partial V}{\partial \tau}(\tau_k + \Delta\tau, S) = \frac{v_i^{k+1} - v_i^k}{\Delta \tau} + O(\Delta \tau), \tag{2.57}$$

$$\frac{\partial V}{\partial S}(\tau_k + \Delta\tau, S) = \frac{v_{i+1}^{k+1} - v_{i-1}^{k+1}}{2\Delta S} + O\left(\Delta S^2\right), \tag{2.58}$$

$$\frac{\partial^2 V}{\partial S^2}(\tau_k + \Delta\tau, S) = \frac{v_{i+1}^{k+1} - 2v_i^{k+1} + v_{i-1}^{k+1}}{\Delta S^2} + O\left(\Delta S^2\right), \tag{2.59}$$

where we note that the space derivatives of (2.58) and (2.59) are calculated at time $k+1$ rather than at time k as is the case for the explicit scheme, and (2.57) is a backward difference. We now are not confined in choosing the size of the space step, although the scheme is only still $O(\Delta\tau, \Delta S^2)$ accurate. However using the implicit scheme does increase the computational expense as, for PDEs of the form (2.53), it requires the solving of a tridiagonal system of linear equations. However, there are methods such as Thomas algorithm or LU decomposition which exploit the tridiagonal structure of the system (see Smith, 1965).

Crank-Nicolson method

The following scheme is as simple to implement (for linear PDEs) as the implicit method and offers higher rates of convergence. It is thus often the method of choice for solving parabolic PDEs. The method was devised by Crank and Nicolson (1947) and has convergence $O(\Delta\tau^2, \Delta S^2)$, conditional on smoothness of the boundary conditions. The approximations for V and the derivatives are

$$V\left(\tau + \frac{1}{2}\Delta\tau_k, S\right) = \frac{v_i^{k+1} + v_i^k}{2} + O\left(\Delta\tau^2\right), \tag{2.60}$$

$$\frac{\partial V}{\partial \tau}\left(\tau_k + \frac{1}{2}\Delta\tau, S\right) = \frac{v_i^{k+1} - v_i^k}{\Delta\tau} + O\left(\Delta\tau^2\right), \tag{2.61}$$

$$\frac{\partial V}{\partial S}\left(\tau_k + \frac{1}{2}\Delta\tau, S\right) = \frac{1}{2}\left(\frac{v_{i+1}^{k+1} - v_{i-1}^{k+1}}{2\Delta S} + \frac{v_{i+1}^{k} - v_{i-1}^{k}}{2\Delta S}\right) + O\left(\Delta S^2\right), \tag{2.62}$$

$$\frac{\partial^2 V}{\partial S^2}\left(\tau_k + \frac{1}{2}\Delta\tau, S\right) = \frac{1}{2}\left(\frac{v_{i+1}^{k+1} - 2v_i^{k+1} + v_{i-1}^{k+1}}{\Delta S^2} + \frac{v_{i+1}^{k} - 2v_i^{k} + v_{i-1}^{k}}{\Delta S^2}\right) + O\left(\Delta S^2\right), \tag{2.63}$$

which when substituted into a PDE of the form of (2.53) results in a similar tridiagonal linear system as is the case of the implicit method, however the right-hand-side is now time-dependent.

2.6.2 Runge-Kutta method

When solving ODEs numerically we may require higher-order methods, over a standard Euler scheme (see Euler, 1768), which will reduce computational expense while increasing the convergence. For this we use the Runge-Kutta method (see Press et al., 2009). Consider the initial value problem

$$f' = g(t, f) \qquad f(t_0) = f_0. \tag{2.64}$$

Rather than having a grid we are only solving in a single domain, t, with $t \in [t_0, T]$. The time is discretised into N intervals of equal length (equality of intervals can be relaxed) such that we have $N + 1$ points. Let

$$h = \frac{T - t_0}{N}.$$

At time point t_n we define the order s Runge-Kutta method as

$$f_{n+1} = f_n + \sum_{i=1}^{s} b_i k_i,$$

with

$$t_{n+1} = t_n + h$$

where

$$\begin{aligned}
k_1 &= hg(t_n, y_n), \\
k_2 &= hg(t_n + c_2 h, f_n + a_{21} k_1), \\
k_3 &= hg(t_n + c_3 h, f_n + a_{31} k_1 + a_{32} k_2), \\
&\vdots \\
k_s &= hg(t_n + c_s h, f_n + a_{s1} k_1 + a_{s2} k_2 + \cdots + a_{s,s-1} k_{s-1}).
\end{aligned}$$

and $f_n = f(t_n)$. To specify a particular method, one needs to provide the integer s (the number of stages), and the coefficients a_{ij} (for $1 \leq j < i \leq s$), b_i (for $i = 1, 2, \ldots, s$) and c_i (for $i = 2, 3, \ldots, s$). The matrix $[a_{ij}]$ is called the Runge-Kutta matrix, while the b_i and c_i are known as the weights and the nodes. The Runge-Kutta method is consistent if

$$\sum_{j=1}^{i-1} a_{ij} = c_i \text{ for } i = 2, \ldots, s.$$

The method we use is the fourth-order explicit Runge-Kutta method which is $O(h^4)$ accurate. Consider again the system (2.64); the fourth-order explicit Runge-Kutta method, also known as the classical Runge-Kutta method, is defined as

$$f_{n+1} = f_n + \tfrac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{2.65}$$

with

$$k_1 = g(t_n, f_n), \tag{2.66}$$

$$k_2 = g\left(t_n + \tfrac{h}{2}, f_n + \tfrac{h}{2}k_1\right), \tag{2.67}$$

$$k_3 = g\left(t_n + \tfrac{h}{2}, f_n + \tfrac{h}{2}k_2\right), \tag{2.68}$$

$$k_4 = g(t_n + h, f_n + hk_3). \tag{2.69}$$

This is much more favourable than the Euler method, which is equivalent to the first order Runge-Kutta Scheme, and has convergence $O(h)$.

2.6.3 Iterative methods

In this thesis we often need to solve equations iteratively, whether they are non-linear ODEs, using Crank-Nicolson to solve non-linear PDEs, or solving free-boundary problems. We shall first discuss the method we use for solving a non-linear ODE as it is frequently used in this thesis, before examining the Successive-Over-Relaxation (SOR) method which is used to verify results when examining higher-order problems.

Newton's iterative method

Consider the second-order non-linear ODE for the value function V:

$$g(S, V, V', V'') = 0 \tag{2.70}$$

which has linear boundary conditions. We can approximate all derivatives using central differences which will result in the non-linear system of equations

$$\mathbf{G}(\mathbf{V}) = 0. \tag{2.71}$$

Given a starting value $\mathbf{V}^{(0)}$ we can approximate the $k^{th}+1$ iterative solution from the k^{th} iterative solution as

$$\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)} + \boldsymbol{\varepsilon}^{(k)} \tag{2.72}$$

in which $\boldsymbol{\varepsilon}^{(k)}$ is the correction at iteration k. The system is non-linear in both $\mathbf{V}^{(k)}$ and $\boldsymbol{\varepsilon}^{(k)}$. We can linearise in $\boldsymbol{\varepsilon}^{(k)}$ by taking the limit up to $O\left(\boldsymbol{\varepsilon}^{(k)}\right)$ as $\boldsymbol{\varepsilon}^{(k)} \to 0$. This results in a linear system of equations for $\boldsymbol{\varepsilon}^{(k)}$. To calculate $\boldsymbol{\varepsilon}^{(k)}$ we solve the system

$$\mathbf{J}\left(\mathbf{V}^{(k)}\right)\boldsymbol{\varepsilon}^{(k)} = -\mathbf{G}\left(\mathbf{V}\right), \tag{2.73}$$

where

$$\mathbf{J}\left(\mathbf{V}\right) = \frac{\partial \mathbf{G}}{\partial \mathbf{V}}\left(\mathbf{V}\right) = \left(\frac{\partial G_i}{\partial V_j}\left(\mathbf{V}\right)\right)_{0,\leq i,j\leq N} \tag{2.74}$$

is the Jacobi matrix of $\mathbf{G}$. As G_i is independent of V_j with $|i-j|>1$, the matrix $\mathbf{J}\left(\mathbf{V}^{(k)}\right)$ is tridiagonal for the class of problems we will solve in this thesis. Consequently, each step of the Newton iteration involves the solution of a system of linear equations with a tridiagonal matrix; this may be accomplished by using the Thomas algorithm. We repeat until some tolerance criteria is satisfied, namely that the sum of squares of the error are less than some tolerance, in which the tolerance should depend on the order of magnitude of $\mathbf{V}$.

Successive-over-relaxation method

We shall now discuss the SOR method (see Cryer, 1971). In this thesis we use this method for PDEs which have dimension higher than one space and one time variable. We shall thus examine it in a similar framework as (2.53) but with two space variables, S_1 and S_2

$$-\frac{\partial V}{\partial \tau} + \sum_{i=1}^{2}\mu_i\left(\tau, S_1, S_2\right)\frac{\partial V}{\partial S_i} + \sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{ij}\left(\tau, S_1, S_2\right)\frac{\partial^2 V}{\partial S_i \partial S_j} + g\left(\tau, S_1, S_2\right)V = 0, \tag{2.75}$$

with $V = V\left(\tau, S_1, S_2\right)$. With appropriate boundary conditions (2.75) would be used for the value of an option dependent on two underlyings.

Let $V_{i,j}^{k} = V\left(k\Delta\tau, i\Delta S_1, j\Delta S_2\right)$. Using implicit finite-differences we have a resulting linear system $\mathbf{AV} = \mathbf{b}$ at each time-step. The matrix $\mathbf{A}$ is an $(n\times m)\times(n\times m)$ sparse-banded matrix which consists of $m+1$ subdiagonals and $m+1$ superdiagonals, where n and m are the number of grid points in S_1 and S_2 respectively. It consists of a main diagonal of parameters, a main superdiagonal of parameters, a main subdiagonal of parameters, $m-3$ superdiagonals of zeros followed by three rows of parameters and $m-3$ subdiagonals of zeros followed by three rows of parameters.

For the system $\mathbf{AV} = \mathbf{b}$ we can use an iterative scheme. Let $V_{i,j}^{k,p}$ be the p^{th} iteration of $V_{i,j}^{k}$. For the SOR method we then have

$$y_{i,j} = \frac{1}{\beta_{i,j}}\Bigg(b_{i,j} - \alpha_{i,j}V_{i-1,j}^{k+1,p+1} - \gamma_{i,j}V_{i+1,j}^{k+1,p} - \delta_{i,j}V_{i,j-1}^{k+1,p+1} - \epsilon_{i,j}V_{i,j+1}^{k+1,p} - \nu_{i,j}\Big(V_{i-1,j-1}^{k+1,p+1} - V_{i+1,j-1}^{k+1,p} - V_{i-1,j+1}^{k+1,p+1} + V_{i+1,j+1}^{k+1,p}\Big)\Bigg), \tag{2.76}$$

$$V_{i,j}^{k+1,p+1} = V_{i,j}^{k,p} + \omega\left(y_{i,j} - V_{i,j}^{k+1,p}\right), \tag{2.77}$$

where β is the main diagonal of $\mathbf{A}$, α is the main subdiagonal, γ is the main superdiagonal, the lower band has main diagonal δ and superdiagonal and subdiagonal ν, and the upper band has main diagonal ϵ and superdiagonal and subdiagonal ν. Beginning by setting $V_{i,j}^{k+1,0} = V_{i,j}^{k}$, we iterate over these two equations until the difference between successive iterations is sufficiently small. There are various methods for deciding how small your tolerance between successive terms should be and these essentially relate to the order of magnitude of your solution (see Smith, 1965). The over-relaxation parameter ω has a range $0 < \omega < 2$ and for $0 < \omega < 1$ is known as under-relaxation. It has been shown, (see Young, 1971), that an optimal value of ω exists and this optimal value results in more rapid convergence than any other value of ω. There are means of calculating or estimating the optimal value of ω analytically. However, these typically take so many calculations that it is often quicker to adjust the value of ω at each time step so that it converges to an approximation of the optimal. This minimises the number of iterations it takes to solve the above equations. The method is outlined in pseudocode in algorithm 1. Given we are using implicit differences we expect the convergence to be $O\left(\Delta\tau, \Delta S_1^2, \Delta S_2^2\right)$.

Algorithm 1 Finding the approximate optimal ω for SOR

```
eps ← 1.e − 16
omega ← 1
domega ← 0.01
loops = sorsolver(A,b,f,omega,eps)
if loops > oldloops then
    domega ← domega × −1.
omega ← max(1.0,min(1.99,omega+domega))
oldomega ← omega
```

Projected successive-over-relaxation method

An extension to the above SOR method is the Projected Successive-Over-Relaxation (PSOR) method. This is often used in the case of free-boundary problems. Consider again (2.75) but with condition

$$V(\tau, S_1, S_2) \geq f(\tau, S_1, S_2). \tag{2.78}$$

This is a free-boundary that is present for the lifetime of the problem. Equation (2.75) can be solved with the free-boundary condition (2.78) by amending the SOR method (2.76, 2.77) to

$$\begin{aligned} y_{i,j} = \frac{1}{\beta_{i,j}}\Bigg(& b_{i,j} - \alpha_{i,j} V_{i-1,j}^{k+1,p+1} - \gamma_{i,j} V_{i+1,j}^{k+1,p} - \delta_{i,j} V_{i,j-1}^{k+1,p+1} - \epsilon_{i,j} V_{i,j+1}^{k+1,p} \\ & - \nu_{i,j}\Big(V_{i-1,j-1}^{k+1,p+1} - V_{i+1,j-1}^{k+1,p} - V_{i-1,j+1}^{k+1,p+1} + V_{i+1,j+1}^{k+1,p} \Big) \Bigg), \end{aligned} \tag{2.79}$$

$$V_{i,j}^{k+1,p+1} = \max\left(V_{i,j}^{k,p} + \omega\left(y_{i,j} - V_{i,j}^{k+1,p} \right), f_{i,j}^{k+1} \right), \tag{2.80}$$

where $f_{i,j}^{k} = f(k\Delta\tau, i\Delta S_1, j\Delta S_2)$.

Crank (1984) notes that although the PSOR method has been used heuristically as far back as Christopherson and Southwell (1938), convergence was not proved rigorously until Cryer (1971).

2.7 Summary

In this chapter we have discussed the prominent mathematical theory and techniques that are used in this thesis. When deriving our models we will first develop them in a stochastic framework. As trading is a search problem (a buyer searching for a seller and vice-versa), we use techniques from the theory of optimisation to formulate the problems. As most of the problems we encounter are highly non-linear, the probability of obtaining analytic solutions is small. Therefore, we use a combined approach of accurate numerical techniques and asymptotic approximation methods, all which have been discussed above, to find solutions to the problems.

Chapter 3

Literature Review of Optimal Trading

Optimal trading refers to the design of trading strategies and tactics by means of quantitative methods. There are two layers to this problem: the first is a *strategic* one, also known as a *macro-trader*, which is linked to time-scheduling, meaning, how to define the optimal trading rate. The second is a *tactical* one, also known as a *micro-trader*, which deals with the way we should interact with the market in order to obtain the desired number of shares, once a strategy from the macro-trader has been decided. In this thesis we focus on the tactical layer. However, we cannot develop an understanding of one without discussing the other[1].

Optimal scheduling of large orders to control trading costs was originally proposed for *aggressive* trading (trading using market orders) to find an optimal balance (and hence trading strategy) between reducing market impact (by trading slow) and reducing market risk (by trading fast).

Brokers face the problem of controlling all the transaction costs related to the liquidation of large orders. Among these costs, market impact deserves particular attention. Market impact is a specific type of liquidity risk which describes the risk of not being able to execute a trade at a specific price due to the feedback effect the execution of that trade has on the market. The rate of trading feedbacks negatively onto the price at which assets are sold, through some market impact function which depends

[1]It could be argued that a third layer exists; that layer being the *smart order routing* layer which focuses on which venue to send your order to. In this thesis we assume only one venue is available to the trader. For smart order routing we refer the interested reader to Lehalle and Laruelle (2013).

on the volume traded through the market. In order to reduce market impact, brokers split their parent orders into smaller child orders which are executed throughout a pre-allocated time. Their goal becomes to find the optimal rate at which to liquidate the order: if too fast, they face higher market-impact costs; if too slow, they risk market prices moving in an unfavourable direction, resulting in a worse-than-expected execution price at the end of the allocated period.

The first models were developed around the turn of the millennium by Bertsimas et al. (1999) and Almgren and Chriss (2001). These models focus on trading in the market book (i.e. using market orders only), in which some function of the trader's wealth, the objective function, is optimised. Both models use linear impact functions for discrete-time trading and static trading strategies are found. This means the strategies are known before trading begins and thus are independent of the asset price, which is driven by a standard Brownian motion with drift. This is further developed to continuous time by Almgren (2003).

Schied and Slynko (2011) argue that due to high-frequency trading, the models developed by Almgren and Chriss do not capture the full impact of trading when the time spans are short. We refer to the Almgren and Chriss framework as aggressive market-order models, given they do not consider the limit-order book structure. The framework of Almgren and Chriss (2001) is further developed for trading in the limit-order book, which is a better replication of real-life trading than trading in the market book as market depth and bid/ask prices are considered. Obizhaeva and Wang (2013), which has been a preprint paper since 2004, introduce aggressive trading in the limit-order book in which a block shaped density for the limit-order book was assumed. The temporary market impact function used in market-book models, in which only one trade was affected, was replaced by a transient market impact function, a decreasing function of time representing the limit-order book refilling post trades. These models shall be referred to as aggressive limit-order models.

Tactical trading on the other hand does not involve the trade-off between market impact and market risk. Market impact is replaced by a new type of risk: the risk of non-execution. In tactical trading, the trader interacts with the limit-order book by *passively* placing orders into the limit-order book which are only filled when met by an aggressive trader's order. The further into the limit-order book the trades are

placed, the higher the payoff for the trader but with a lower probability of the order being filled. Given this trading scenario, the filling of an order is not guaranteed prior to the terminal time, and hence there is a risk of non-execution. As the trader can continuously amend his asking price, the price at which the asset is sold is controlled by the trader. However, the trader cannot control the 'fair' price of the asset, which is driven by a stochastic process, and thus market risk is still present. This type of model was first suggested by Ho and Stoll (1981) but lay dormant in the literature for some time until Avellaneda and Stoikov (2008) revisited this problem with some modifications, with the task of a market maker in mind. More recently, strategies that use a combination of trading orders, both passive and aggressive, have been investigated so we will briefly discuss these also.

For the remainder of the chapter we shall give a detailed review of the optimal trading literature. We begin with a discussion on optimal scheduling problems using only aggressive market orders, in which the structure of the limit-order book is not considered. Next we discuss various methods for modelling the limit-order book, including dynamically modelling each level, and modelling the whole book statistically. The latter method is used in this thesis given that in terms of dimension it is more amenable to work it. This is followed by examining optimal scheduling in a limit-order book structure with aggressive market orders, which predominately assumes statistical properties of the limit-order book rather than modelling it directly. After this we discuss the literature which focuses on trading at a tactical level, which includes trading with passive limit orders or aggressive market orders, again assuming statistical properties of the limit-order book. We conclude by describing the framework and results of Guéant et al. (2012b), a similar framework to that of Avellaneda and Stoikov (2008) and Cartea and Jaimungal (2013b) but in a single-sided order book, which was a strong motivation for the work of this thesis.

3.1 Aggressive optimal scheduling with market orders

Bertsimas et al. (1999), working in discrete time, minimise the expected value of the mean cost of trading, in which they define the cost of trading as the difference between

the initial value of the portfolio prior to trading and the amount of money held post trading. They consider only a permanent price impact function, which is linear with respect to the rate of trading. The price at which assets are sold during a trade is the 'fair' price, driven by a standard Brownian motion, and some additional amount which is dependent on the rate of trading. This additional amount remains added to the 'fair' asset price for the remainder of the duration, indicating the indifference between supply and demand post trade.

Around the same time, Almgren and Chriss (2001) develop a similar model in which their objective is to minimise a mean-variance trade-off of the cost of trading while introducing a temporary impact function, as well as the permanent impact function introduced by Bertsimas et al. (1999). This temporary impact function is also linear with respect to the rate of trading but is different in that it only affects the current trade. Minimising the efficient frontier, a linear combination of the mean and variance of the cost of trading, Almgren and Chriss (2001) produce analytic expressions for the trading strategy, i.e. the amount of the portfolio sold at each discrete time step. These trading strategies are termed 'static' in the literature as they only consider the information available at the initial time step. The strategies found consisted of splitting the order into equally sized, equally time spaced trades.

There have been various extensions to this framework. Almgren (2003) further develops the Almgren and Chriss (2001) method to account for non-linear impact functions and examines the model in continuous time. Almgren and Lorenz (2007) allow for one update of the asset price at a fixed time, which is fixed before trading, and from this creates dynamic strategies based on this update. Lorenz and Almgren (2011) allow continuous updates of the share price to develop dynamic trading strategies. Almgren (2012) introduces stochastic liquidity and volatility. Labadie et al. (2012) derive explicit recursive formulas for Target Close (strategy that aims to execute a certain amount of shares as near as possible to the closing auction price) and Implementation Shortfall (strategy that aims to minimise the difference between the decision price and the final execution price) in the Almgren-Chriss framework. Cartea and Jaimungal (2015a) consider order-flow from other market participants (which is assumed to be stochastic) and how that affects the temporary and permanent impact.

Schied et al. (2010) take a different approach in which they wish to maximise the

expected utility of the revenues, rather a mean-variance trade-off, while using the same Almgren & Chriss approach as above for the dynamics of the asset price and impact functions. This is examined for a basket of correlated assets following standard Brownian motion. The problem is modelled as a finite fuel control problem, in which the control variable, the amount of assets held (fuel), is bounded. They assume the trader's utility is defined by CARA and as such the utility function has the form of a negative exponential function. Using this they derive an HJB equation and prove a unique solution exists. Asymptotic methods are discussed by Schied and Schöneborn (2009) for deriving solutions to the HJB equation for utility functions that do not follow the form of an exponential function. Haugh and Wang (2014) examine liquidating a fixed number of shares in multiple stocks while introducing stochasticity to the linear temporary and permanent impact. Bayraktar and Ludkovski (2011) consider a trader who submits orders discretely and solve a control problem corresponding to when to send market orders, and how large these should be.

Gatheral and Schied (2011) examine the risk-neutral framework of Bertsimas et al. (1999) but under geometric Brownian motion with zero drift driving the asset price. It was found that using the expected cost of trading as the objective function the optimal trading strategy is found to be static and is identically that of the trading strategy found by Bertsimas et al. (1999) when using standard Brownian motion. Forsyth (2011) considers numerical methods to solve the optimal trade execution problem for an asset following geometric Brownian motion under the Almgren and Chriss (2001) framework. The objective is to find the optimal trading strategy under a mean-variance objective function. In this model the asset exhibits both permanent and temporary price impact. Two PDEs are derived, one classical PDE for the expected value and one HJB equation for the variance. Using the semi-Lagrangian method Forsyth (2011) solves the HJB equation for the variance, and using the optimal trading strategy the classical PDE is solved giving the solution for the efficient frontier. In the numerical schemes Forsyth (2011) finds his solutions are similar to that developed in Almgren and Lorenz (2007), in which adaptive strategies are developed based on updated knowledge of the asset price, as opposed to stochastic dynamic programming. Forsyth et al. (2012) examine the optimal trading strategies of an asset under both standard Brownian motion and geometric Brownian motion, using a mean-variance cost function as the

risk criterion. They find that under the standard Brownian motion case the optimal strategies are the same as that under the Almgren and Chriss (2001) framework, and that the optimal trading strategy under standard Brownian motion is a good approximation for the optimal trading strategy under geometric Brownian motion, even under large volatility.

More recently, the concept of dark pools has been introduced into the literature. Kratz and Schöneborn (2013) consider a trader who can trade in the market book (with price impact) and in a dark pool (no price impact), with execution occurring randomly in the dark pool according to a Poisson process. Buti et al. (2011) examine a trader who can trade in a limit-order book and a dark pool, in which orders in the dark pool are randomly filled at the terminal time. They find that the introduction of a dark pool that competes with an illiquid order book is on average associated with trade creation, but also a deterioration of market quality and welfare.

Execution has also been studied from the point of view of complex products, such as derivatives and accelerated share repurchases (ASR). ASR occurs when a firm wants to repurchase their own shares. The firm often enters into such a contract with a bank. The bank buys the shares for the firm and is paid the average market price over the execution period, the length of the period being decided upon by the bank during the buying process. Guéant et al. (2015), working in discrete time, examine the problem of ASR when market impact is present. They determine the optimal stopping time and the optimal buying strategy for the bank. Guéant (2014) extends Guéant et al. (2015) by considering ASR for a fixed notional. Guéant and Pu (2013) examine the pricing and hedging of a call option when liquidity is considered. Rather than a frictionless market, as assumed under the original Black-Scholes model, buying and selling when hedging incurs permanent market impact and thus it becomes an optimal execution problem.

3.2 Dynamics of limit-order books

The behaviour of the limit-order book has been modelled both dynamically and statistically. It is the latter case that we consider in this thesis as it is more mathematically tractable due to its reduced dimension. However, we shall discuss some of the literature

of both frameworks.

3.2.1 Explicit representation of limit-order books

Given we focus on optimal trading strategies in limit-order books, it is necessary to consider some of the frameworks for which they are modelled. The stochastic dynamics of the limit-order book has been studied in some detail, supported by empirical evidence. Whereas most of the empirical analysis, e.g. Bouchaud et al. (2009), focuses on unconditional distributions of various features of the limit-order book, Cont and De Larrard (2013) focus on the conditional distributions and thus the idea of forecasting becomes logical. This is extended by Cont et al. (2013) who derive a model for the dynamics of the limit-order book, which we briefly describe.

In Cont et al. (2013) the state of the limit-order book is described by a continuous-time process:

$$X(t) = (s^b(t), q^b(t), q^a(t)), \tag{3.1}$$

which takes values on the discrete space $\delta\mathbb{Z} \times \mathbb{N}^2$, where δ is the size of one tick, with $s^b(t)$ as the best-bid price at time t, $q^b(t)$ as the size of the best-bid queue, and $q^a(t)$ as the size of the best-ask queue. Equation (3.1) is a reduced form of its original form, which was a four-dimensional process. This reduction is a result of assuming the bid-ask spread is equal to one tick, which was shown to be the case on empirical data over 98% of the time. This assumption reduces the process from four dimensions to three, as the best-ask price need not be modelled explicitly.

Considering orders, which may be cancellations, of size V_i^b and V_i^a for the bid and ask side respectively, the dynamics of the process are given as:

- If an order (limit or market) or cancellation, of size V_i^a, occurs on the ask side at time t:

$$\begin{aligned}(s^b(t), q^b(t), q^a(t)) &= (s^b(t-), q^b(t-), q^a(t-) - V_i^a)\mathbb{1}_{\{q^a(t-) > V_i^a\}} \\ &\quad +(s^b(t-) + \delta, R_i^b, R_i^a)\mathbb{1}_{\{q^a(t-) < V_i^a\}}.\end{aligned} \tag{3.2}$$

- If an order (limit or market) or cancellation, of size V_i^b, occurs on the bid size at time t:

$$\begin{aligned}(s^b(t), q^b(t), q^a(t)) &= (s^b(t-), q^b(t-) - V_i^b, q^a(t-))\mathbb{1}_{\{q_{t-}^b > V_i^b\}} \\ &\quad +(s_{t-}^b - \delta, \bar{R}_i^b, \bar{R}_i^a)\mathbb{1}_{\{q_{t-}^b < V_i^b\}},\end{aligned} \tag{3.3}$$

where (R_i^b, R_i^a) and $(\bar{R}_i^b, \bar{R}_i^a)$ are independent and identically distributed random variables with joint distribution functions f and $\bar{f}$ respectively. It is assumed all market events occur according to a Poisson process with market orders occurring at a rate μ, limit orders occurring at a rate λ, and cancellations occurring at a rate θ. Therefore the process has the following dynamics:

- $(T_i^a)_{i\geq 0}$, the durations between two consecutive orders at the best-ask queue, is a sequence of exponentially distributed random variables with rate $\lambda + \theta + \mu$.
- $(T_i^b)_{i\geq 0}$, the durations between two consecutive orders at the best-bid queue, is a sequence of exponentially distributed random variables with rate $\lambda + \theta + \mu$.
- $(V_i^a)_{i\geq 0}$ is a sequence of independent random variables with:

$$Pr[V_i^a = 1] = \frac{\lambda}{\lambda + \mu + \theta} \quad \text{and} \quad Pr[V_i^a = -1] = \frac{\mu + \theta}{\lambda + \mu + \theta}. \tag{3.4}$$

- $(V_i^b)_{i\geq 0}$ is a sequence of independent random variables with:

$$Pr[V_i^b = 1] = \frac{\lambda}{\lambda + \mu + \theta} \quad \text{and} \quad Pr[V_i^b = -1] = \frac{\mu + \theta}{\lambda + \mu + \theta}. \tag{3.5}$$

Using this model, Cont et al. (2013) derive analytic solutions based on conditional probabilities for the duration between price movements, and the probability of price changes. They also examine the distribution of price movements, their auto-correlations and the volatility of the price. The method they use to estimate the parameters μ, λ and θ is power law extrapolation and is described in their previous paper Cont et al. (2010). In Cont et al. (2010) the method keeps track of all the queues in the system. When one queue is depleted the value of the next is known, rather than generated randomly as above. Using this method the authors derive numerical solutions for the same quantities which are found analytically in Cont et al. (2013).

3.2.2 Statistical properties of limit-order books

One method that dominates the literature is to not explicitly consider the limit-order book, but instead to consider the statistical properties of its liquidity, reducing the dimension of the problem to a level that is more amenable to work with. In the literature, and throughout this thesis, asset sales are modelled as a Poisson process in

which the liquidity of the limit-order book is encapsulated by the inhomogeneous rate of the Poisson process. As we will see, the majority of the literature models the rate as a negative exponential function, decreasing the further into the limit-order book we place the order. However, power intensity functions are also examined, such as in Bayraktar and Ludkovski (2012). In this thesis we shall examine both, and the trading strategies that result.

These models assume limit orders are filled by significantly large incoming market orders of size Q. It is assumed that market orders to buy will fill all sell limit orders within $I(Q)$, and vice versa. These large orders cause a change of price, ΔS, proportional to Q. It is argued the distribution of the size of market orders has been found to obey a power law, i.e. the density, f_Q, of the order sizes, Q, obeys

$$f_Q(q) \propto q^{-\alpha-1}, \tag{3.6}$$

where the power α was found to be 1.53 by Gopikrishnan et al. (2000) for U.S. stocks, 1.4 by Maslov and Mills (2001) for the NASDAQ and 1.5 by Gabaix et al. (2006) for the Paris Bourse. There is less agreement on how price changes due to these market orders. Some authors find the change in price of an asset, ΔS, following a large market order, Q, is given by

$$\Delta S \propto Q^{\beta} \tag{3.7}$$

with Gabaix et al. (2006) finding $\beta = 0.5$ and Weber and Rosenow (2005) finding $\beta = 0.76$. Others, most noticeably Potters and Bouchaud (2003), find a better fit to be

$$\Delta S \propto \ln(Q). \tag{3.8}$$

Under (3.8) it is possible to derive a Poisson intensity, dependent on the distance from the reference price, as

$$\begin{aligned}
\Lambda(\delta) &= C \Pr(\Delta S > \delta) \\
&= C \Pr(\ln(Q) > K\delta) \\
&= C \Pr\left(Q > e^{K\delta}\right) \\
&= C \int_{e^{K\delta}}^{\infty} q^{-\alpha-1} dq \\
&= \lambda e^{-l\delta},
\end{aligned} \tag{3.9}$$

where $\lambda = C/\alpha$, $l = K\alpha$ and C and K are constants. The parameter l describes the exponential decay of the limit-order book, i.e. how quickly or slowly the demand changes as we move further into the limit-order book and thus how quickly or slowly the probability of execution decreases. Similarly, under (3.7) we derive

$$\begin{aligned}\Lambda(\delta) &= C\Pr(\Delta S > \delta) \\ &= C\Pr\left(Q^{\beta} > K\delta\right) \\ &= C\Pr\left(Q > (K\delta)^{1/\beta}\right) \\ &= C\int_{(K\delta)^{1/\beta}}^{\infty} q^{-\alpha-1}dq \\ &= \frac{\lambda}{\delta^{\alpha/\beta}},\end{aligned} \tag{3.10}$$

with $\lambda = \dfrac{CK^{\alpha/\beta}}{\alpha}$. This form of intensity function is examined in chapter 8, with the exponential intensity used elsewhere in this thesis.

3.3 Aggressive optimal scheduling with limit orders

Obizhaeva and Wang (2013) argue that on a finer time-scale the temporary market impact implemented first by Almgren and Chriss (2001) is in fact transient. For transient market impact an order creates some immediate price impact that subsequently decays over a certain, but short, period of time; this concept is known as resilience. A survey of empirical studies on the resilience of market microstructure is covered in Bouchaud (2010) and we refer to the references therein for more detail.

In the market-book models it was assumed that liquidity was encapsulated in the price. The introduction of limit-order book models by Obizhaeva and Wang (2013) coincided with a period when liquidity was examined through limit-order books. The original idea behind this is to examine liquidity as a dynamic process which represented supply and demand, rather than a static cost function that encapsulates spread and market depth. The novelty of the approach is to include the resilience of the limit-order book, which is the evolution of the limit-order book for future times under current trades. The optimal strategies found are dependent on the properties of the limit-order book and consists of large and small trades, which is in contrast to previous studies

by Bertsimas et al. (1999) and Almgren and Chriss (2001) that found the optimal strategy to be splitting the trade into equal sized, equally time spaced trades.

Obizhaeva and Wang (2013) use a similar framework as Bertsimas et al. (1999) and Almgren and Chriss (2001) in that their objective is to find a trading strategy that minimises the cost of trading. Rather than permanent and temporary market impact functions they include permanent and transient market impact functions. They assume that the limit-order book has a simple block-shaped density in which there is a constant amount of shares on offer at each price. When a trade occurs the order will 'eat into' the limit-order book by filling the cheapest orders first. This raises (lowers) the best-ask (best-bid) price for the case of buying (selling), which is unfavourable for the trader. It is assumed the trade has both a transient and permanent impact on the best-ask (best-bid) price. For the case of buying it permanently increases the best-ask price according to some linear (in order size) impact function. It also increases the price by some linear (in order size) impact function which is a decaying function of time, known as transient impact, which disappears as time increases. Obizhaeva and Wang (2013) allow for a combination of discrete and continuous trading, and find that the optimal strategy is a combination of both discrete and continuous trades, with a large trade at the initial and final time, and the rest of the trades spread out continuously between the initial and final time exclusively.

The work of Obizhaeva and Wang (2013) has since been developed. Alfonsi et al. (2008) introduce a time-dependant resilience parameter due to the daily U-shaped or W-shaped pattern in which trading volume typically occurs. The U-shaped (W-shaped) volume pattern indicates there is more volume at the beginning, end (and middle) of the day (see, for example, Wood et al., 1985; Harris, 1986; Jain and Joh, 1988). Alfonsi et al. (2010) consider general shape functions for the limit-order book and tackle the problem through the use of Lagrange multipliers, rather than through the use of dynamic programming as done by Obizhaeva and Wang (2013). Alfonsi et al. (2010) find a closed form solution for a broad class of limit-order books and show the suggested optimal strategies are similar to those derived by Obizhaeva and Wang (2013) for the block density function. Alfonsi and Schied (2010) continue the work of Alfonsi et al. (2008) and Alfonsi et al. (2010) by considering price manipulation of an asset which comes in the form of the exponential resilience. Price manipulation

is when a trader starts with nothing, and through a strategy of buying and selling shares, ends up with zero shares and a positive amount of wealth, i.e a profit. Alfonsi and Schied (2010) allow for trading times to be stopping times rather than equidistant discrete times. Alfonsi et al. (2012) combine three forms of linear impact which they suggest coincide better with empirical models than the other literature as they disallow price manipulation. The three impacts are temporary, transient and permanent. Using linear price impact functions they examine resilience functions that prelude price manipulation strategies. They find that for total exclusion of price manipulation it is necessary that price impact decay as a convex function of time. Kharroubi and Pham (2010) consider a continuous time setting but, to better replicate real life, allow only for discrete time trading by introducing a lag variable tracking the time interval between successive trades. This paper was followed by Guilbaud et al. (2010) which discusses numerical (finite-difference) methods used to solve this type of problem. Bouchard et al. (2011) examine the use of different algorithms, depending on the preferences of the trader. A novel approach is devised by Predoiu et al. (2011) in which the shape of a single-sided limit-order book is the determinant of price impact, rather than making assumptions about price impact directly. They only consider transient price impact but allow for the limit-order book to be discrete, continuous or to have gaps. In contrast to previous solutions such as Obizhaeva and Wang (2013) where the optimal trading strategy consists of a discrete lump trade at the beginning and ending and continuous equal sized trading within, Predoiu et al. (2011) find that the optimal trading strategy consists of an initial lump trade, then trading continuously at a rate which matches the resiliency of the limit-order book, then another lump trade, then trading continuously again matching the resiliency. This procedure is repeated until trading is complete, with a final lump trade at the end. Cartea et al. (2015b) examine the execution of a basket consisting of co-moving assets which exhibit temporary and permanent price impact.

Schied and Slynko (2011) and Cartea et al. (2015c) are recommended for a full review of the optimal scheduling literature. The latter text is also recommended for a more in-depth discussion on the material of the following section, that being optimal trading at the tactical level.

3.4 Optimal trading at the tactical level

The problem posed by Ho and Stoll (1981) considers a trader who has cash from selling assets, an inventory of an asset, and a base wealth which can be used to purchase assets, all following stochastic processes. The former two follow jump processes and the latter follows a Wiener process. The objective is to maximise the trader's expected terminal utility which is some transformation of his total wealth, by choosing the optimal trading strategy which is the additional amount the trader asks for the asset when placing your order in the limit-order book. This problem results in a multi-dimensional HJB equation in which the authors note that to the best of their knowledge there is no analytic solution to this problem.

The next time this problem is examined is 25 years later by Avellaneda and Stoikov (2008) in which they make some modifications as to be consistent with modelling a market maker. They assume the asset price is stochastic, rather than given exogenously as in Ho and Stoll (1981). Inventory takes unitary values and follows a jump process and the wealth changes accordingly each time a jump occurs. Under the objective of maximising the trader's utility, Avellaneda and Stoikov (2008) derive an HJB equation which they examine asymptotically, before numerically solving the problem (in its original form) using Monte Carlo methods. Guéant et al. (2012a), implementing an upper and lower bound on the inventory, reduce the HJB equation to a system of ODEs. Guéant et al. (2012b) examine this framework for a single-sided limit-order book, in which assets can only be sold, and a terminal time penalty is introduced. The framework for optimal liquidation with limit orders, as examined by Guéant et al. (2012b), is very close to market-making as we can consider that liquidation, at a tactical level, is a one-sided market maker. Guéant et al. (2012b) explicitly examine the case of liquidating a portfolio but extending it to filling a portfolio is quite straightforward. It is the framework of Guéant et al. (2012b) that this thesis develops and therefore we shall discuss the problem in more thorough detail below.

In addition to the above mentioned models, other contributions have been made to the literature. Bayraktar and Ludkovski (2012) develop a framework for portfolio liquidation for a risk-neutral trader in which the objective is to maximise the

expected revenue of sales. These authors consider a one-sided limit-order book, similar to Guéant et al. (2012b), but introduce a power-law intensity, as opposed to the negative exponential intensity which has dominated the literature. They derive an HJB equation which is reduced to an semi-analytic solution involving a recursive term. Guéant and Lehalle (2013) develop the framework of the previous paper Guéant et al. (2012b) by considering general intensity shapes for the Poisson process. In the appendix Guéant and Lehalle (2013) discuss the use of multiple correlated assets, deriving an HJB equation and providing a verification theorem. However they do not examine solving the problem and no asymptotic or numerical work is carried out. Cartea and Jaimungal (2013b) use a similar set-up to Guéant et al. (2012a) but under a high-frequency trading framework. Their objective is to maximise terminal wealth while penalising inventory deviations from zero and terminal inventory holdings, properties which comply with institutional high-frequency trading as outlined by Brunetti et al. (2011). This is one of the only papers to study the audit trail tape of a limit-order book. Brunetti et al. (2011) note that the net position of a high-frequency trader is zero at the beginning and the end of the day and his portfolio is always between -3000 and 3000 contracts, which is small compared to the average of -20000 to 30000 for ordinary traders. Using these two key points, that a high-frequency trader starts and ends with zero inventory and their inventory is relatively low during the day, Cartea and Jaimungal (2013b) derive trading strategies that are applicable for high-frequency traders. They consider a trader who wants to maximise his terminal wealth, while penalising and constraining inventories. Inventory feeds into the optimisation in three ways. Firstly, the amount a trader can hold is between an upper and lower bound. Secondly, deviations away from zero inventory are penalised throughout the trading horizon. Third, positions that are held at the terminal time are sold/bought at a discount as to guarantee immediate trade. The reference price of the asset is affected by both new information (standard Brownian motion) and the arrival of market orders (Poisson process) which has a permanent price impact on the asset. The authors derive an HJB equation which they reduce to a system of ODEs to obtain first-order solutions for the optimal trading strategies and value function, before examining various limiting regimes. Laruelle et al. (2011) examine the optimal splitting of an order across several trading venues, and especially dark pools, using an approach similar to

Avellaneda and Stoikov (2008) as a starting point. Cartea and Jaimungal (2015c) examine how a trader, using limit orders at the best bid/ask, adjusts his trading strategy under the assumption that the trader has the ability to observe order flow and knows how this affects the drift of the reference price.

Contrasting frameworks for obtaining optimal trading strategies have also been developed. To name a few, Cartea and Jaimungal (2013a) develop a hidden Markov model to understand the key behaviour of stock dynamics resulting in an optimal tick-by-tick trading strategy that a trader who uses limit orders to profit from the bid-ask spread should follow. Hult and Kiessling (2010) model the entire limit-order book as a high-dimensional Markov chain. Optimal trading strategies are discussed in the theory of Markov decision processes, and a value iteration procedure is presented that enables optimal strategies to be found numerically. Cartea et al. (2015a) build a measure of volume imbalance using data from the Nasdaq exchange and devise a trading strategy that demonstrates that including the volume imbalance process considerably boosts the strategy's progress.

3.4.1 Combining passive and aggressive trading

Some of the recent literature has turned to examining the use of various sources of liquidity, such as trading aggressively in the limit-order book (using market orders) while simultaneously passively trading in the limit-order book (using limit orders). Huitema (2013) derives a model where a trader, whose assets follow standard Brownian motion, can trade in both the market book (with the rate of trading impacting the price) and the limit-order book (with a Poisson process with controlled intensity driving the execution of orders), and solves the resulting partial integro-differential equation numerically. Guilbaud and Pham (2013b) develop a model where passive and aggressive trading is possible. Limit orders can be placed at the best quote or the best quote improved by one tick and are filled according to a Cox process, while aggressive trading can only occur at discrete times. This is further developed in Guilbaud and Pham (2013a) by examining a pro-rata micro-structure (which is found for certain securitised products), rather than the usual price/time priority structure (found in equities). Kühn and Stroh (2010) solve Merton's portfolio optimisation problem (see Merton, 1969, 1971) in the case where the trader can choose between market orders or

limit orders at the best bid or best ask. In Veraart (2011) the possibility to use market orders in addition to limit orders is also taken into account, in the context of market making in the foreign-exchange market. Cartea and Jaimungal (2015b) examine execution in a single-sided limit-order book framework, with the modelling assumptions similar to Cartea and Jaimungal (2013b), but where the trader can use both market orders and limit orders.

In chapter 8 we extend our framework to include market orders, which thus allows trading with limit orders and market orders. This work was done in parallel to Cartea and Jaimungal (2015b) in which it is treated as a free-boundary problem. Our novel contribution is (again) through the use of more general diffusion processes and modelling assumptions which require more complex numerical methods to obtain solutions.

3.5 Framework of Guéant et al. (2012b)

In this section we will give an outline of the framework of Guéant et al. (2012b), as it will be central in all of the following chapters in this thesis. This problem focuses on a trader wanting to liquidate an amount of an asset, who trades passively in the limit-order book. The framework is identical to that of a one-sided market maker.

3.5.1 Formulation of the problem

We consider a trader who wishes to maximise his expected utility, given a portfolio of assets to liquidate, before a specified terminal time, T. We assume at time $t = 0$ that the trader starts with an initial inventory of $q(0)$ assets, in which q takes positive integer values, and an initial wealth $X(0)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0, T])$. We assume the asset's reference price[2] $S(t)$ follows a standard Brownian motion, and so the diffusion process is defined as

$$dS(t) = \mu dt + \sigma dW(t) \tag{3.11}$$

[2]We use 'reference price' here as given the limit-order book structure the price at which assets are sold and bought is not the same, as there are various bid and ask prices, with the best bid and ask differing by the spread. The reference price may refer to the best-ask/bid price or mid-price.

where μ is the drift, σ is the volatility and $W(t)$ is a Wiener process which is $\mathcal{F}_t$ measurable. It is clear from (3.11) that an asset following standard Brownian motion can obtain negative values, a property that is often criticised from a financial aspect.

The trader will continuously post orders into the ask side of the limit-order book for price $S^a(t)$ which is $\delta = \delta(t, X, q, S)$ greater than the reference price $S(t)$,

$$S^a(t) = S(t) + \delta. \tag{3.12}$$

Asset sales follow a Poisson process, $N(t)$, with time-dependent intensity, which is $\mathcal{F}_t$ measurable and independent of $W(t)$

$$dq(t) = -dN(t). \tag{3.13}$$

For each occurrence of a jump (trade), the trader's wealth increases by the amount that asset was sold for, so the dynamics of the wealth is given by

$$dX(t) = (S(t) + \delta)\, dN(t), \tag{3.14}$$

where $N(t)$ is the same Poisson process as before. Therefore, when a trade occurs, the values of $q(t)$ and $X(t)$ change simultaneously, according to (3.13) and (3.14) respectively. It is thus assumed trades are of unitary size, which may refer to the size of one *lot* (bundle of 100) or one trade of the average trade size (ATS). This is because assets are primarily sold in bundles rather than individually. $N(t)$ has intensity $\Lambda(\delta)$ which takes the form:

$$\Lambda(\delta) = \lambda e^{-l(S^a - S)} = \lambda e^{-l\delta} \tag{3.15}$$

for some positive constants λ and l. The liquidity of the market is described by the intensity of the Poisson process. If no additional amount is added to the reference price then the rate at which the assets are sold is $\Lambda(0)$, which for the case of (3.15) is equal to λ. Given the negative exponential form of (3.15) there is less liquidity for assets sold for prices higher than their reference price and as such the probability of execution is lower. The parameter l can be interpreted as the exponential-decay factor for the fill rate of orders placed away from the reference price, i.e. how quickly or slowly the demand changes as we move further into the limit-order book and thus how quickly or slowly the probability of execution decreases.

The objective is to liquidate this portfolio before some final time T. Assets that are not liquidated before this time will be sold in the market at a discount, which

assumes immediate sale. The discount is such that the assets are sold for the current asset price $S(T)$ minus some penalty b per asset, for which b is a positive constant. The utility we seek to maximise takes the form of a negative exponential function and as such the trader has CARA defined by (2.18), which for the case of the exponential utility family is constant and equal to the risk-aversion parameter, $\gamma > 0$. We define our value function, $u(t, X, q, S)$, as the maximum expected utility at time t

$$u(t, X, q, S) = \sup_{\delta(t) \in \mathcal{A}} \mathbb{E}\left[-e^{-\gamma(X(T)+q(T)(S(T)-b))}\right], \tag{3.16}$$

where $\mathcal{A}$ is the set of admissible trading strategies.

Given the optimisation problem of (3.16), an HJB equation can be derived by applying the Bellman (1957) principle of optimality and using Itô's lemma:

$$\begin{aligned} &u_t(t, X, q, S) + \mu u_S(t, X, q, S) + \frac{1}{2}\sigma^2 u_{SS}(t, X, q, S) \\ &+\sup_{\delta}\left[\lambda e^{-l\delta}\left(u(t, X + S + \delta, q - 1, S) - u(t, X, q, S)\right)\right] = 0, \end{aligned} \tag{3.17}$$

with subscript denoting the partial derivative (which is used throughout this thesis), and conditions:

$$u(T, X, q, S) = -e^{-\gamma(X+q(S-b))}, \tag{3.18}$$

$$u(t, X, 0, S) = -e^{-\gamma X}. \tag{3.19}$$

The derivation of (3.17), along with a verification theorem, can be found in Guéant et al. (2012b).

Reduced form solution

Guéant et al. (2012b) derive a solution of (3.17) that can be written in the form

$$u\left(t, X, q, S\right) = -\exp\left(-\gamma\left(X + qS\right)\right) w(q, t)^{-\frac{\gamma}{l}} \tag{3.20}$$

where $w(q, t)$ solves

$$w_t(q, t) = (\frac{l}{2}\gamma\sigma^2 q^2 - l\mu q)w(q, t) - \lambda\left(1 + \frac{\gamma}{l}\right)^{-\left(1+\frac{l}{\gamma}\right)} w(q - 1, t), \tag{3.21}$$

with conditions

$$w(q, T) = e^{-lqb} \qquad \text{and} \qquad w(0, t) = 1, \tag{3.22}$$

and the optimal ask quote having solution

$$\delta^*(q,t) = \frac{1}{l}\ln\left(\frac{w(q,t)}{w(q-1,t)}\right) + \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{l}\right). \tag{3.23}$$

Equation (3.21) can be solved numerically using standard ODE solvers such as the Euler method or higher-order methods such as the Runge-Kutta method. We discussed the general Runge-Kutta method for an ODE in section 2.6.2. However we now have a system of q ODEs, rather than just one ODE, so we must adjust the numerical method accordingly. We can write the system (3.21) as

$$w_t(q,t) = g\left(t, w(q,t), w(q-1,t)\right) \quad \text{with} \quad w(q,T) = e^{-lqb} \quad w(0,t) = 1. \tag{3.24}$$

Discretising time into $N-1$ intervals of length h, such that we have N time points, we can solve using the fourth-order Runge-Kutta method at time n by solving

$$w_{n+1}(q) = w_n(q) + \tfrac{h}{6}\left(k_1(q) + 2k_2 l(q) + 2k_3(q) + k_4(q)\right) \tag{3.25}$$

with

$$k_1(q) = g\left(t_n, w_n(q), w_n(q-1)\right), \tag{3.26}$$

$$k_2(q) = g\left(t_n + \tfrac{h}{2}, w_n(q) + \tfrac{h}{2}k_1(q), w_n(q-1) + \tfrac{h}{2}k_1(q-1)\right), \tag{3.27}$$

$$k_3(q) = g\left(t_n + \tfrac{h}{2}, w_n(q) + \tfrac{h}{2}k_2(q), w_n(q-1) + \tfrac{h}{2}k_2(q-1)\right), \tag{3.28}$$

$$k_4(q) = g\left(t_n + h, w_n(q) + hk_3(q), w_n(q-1) + hk_4(q-1)\right). \tag{3.29}$$

with $k_i(q=0) = 0$ for $i = 1, \ldots, 4$. This method has $O(h^4)$ convergence, as can be seen in table 3.1.

Table 3.1: Convergence of Runge Kutta fourth-order scheme for underlying

N	$h\left(=\frac{T-t_0}{N-1}\right)$	$w(t=0,q=1)$	Difference (10^{-7})	ratio
11	0.10000	0.524295209007592	NA	NA
21	0.05000	0.524295151753517	0.572540743704053	NA
41	0.02500	0.524295148277688	0.034758290690107	16.47
81	0.01250	0.524295148063584	0.002141045118975	16.23
161	0.00625	0.524295148050299	0.000132847066681	16.12

N is the number of time points with h as the length of the intervals between respective points. We have set $T = 1, \mu = 0.05, \sigma = 3, \gamma = .5, \lambda = 0.3$ and $l = 0.3$. For a scheme of order p the ratio between successive differences should be 2^p, so for fourth-order convergence we expect a ratio of $2^4 = 16$.

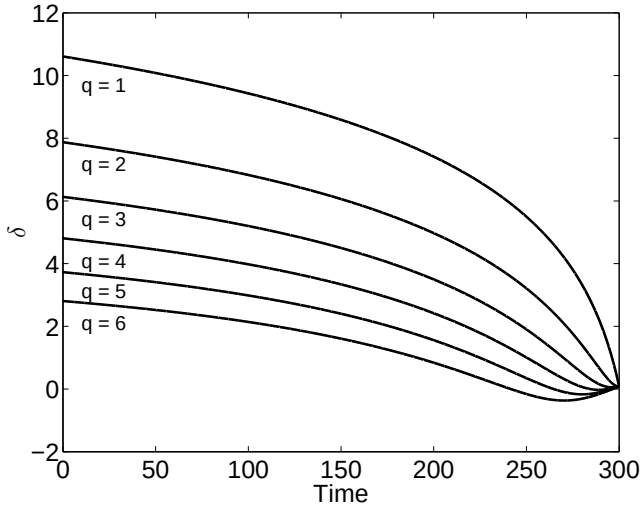


Figure 3.1: Optimal trading strategy under standard Brownian motion framework with $q(0) = 6$. These results were found using (3.23). The parameter values used are consistent with that of Guéant et al. (2012b) and take the values $T = 300, \mu = 0.0, \sigma = 0.3, \lambda = 0.1, l = 0.3, \gamma = 0.05$ and $b = 3$. We can see that as the number of shares increases the optimal strategy decreases.

It is noticeable here that $\delta^*(t, q)$, as plotted in figure 3.1, is independent of both the asset price S and wealth X. This was one of the significant results that tempted us to further examine this problem. Consider the results found in figure 3.1. With one asset remaining, $q = 1$, and $t = 0.5$, the optimal strategy is to sell the asset for an additional 7 price units, be it pounds, pence, dollars or euros. Under the model assumptions this will happen with equal probability regardless of the asset price. Realistically, for an asset priced at 10 units, selling it at 17 units seems less likely than an asset priced at 1000 units selling at 1007. However, the model treats this as having equal probability as the exponential-decay factor of the limit-order book, l, the risk-aversion parameter, γ, the volatility of the asset, σ, and the drift of the asset, μ, are all independent of the asset price. This is something we take into consideration when extending the model.

It is also worth mentioning, but not discussed in the original paper of Guéant et al. (2012b), that a solution of (3.21), and thus (3.17), can be solved analytically using recursion. Although this is tedious as q increases, it can be advantageous to gain further insight into the solution. For the case of $q = 1$ equation (3.21) has solution

$$w\left(q=1,t\right) = \left(e^{-lb} - \frac{\lambda\left(1+\frac{\gamma}{l}\right)^{-(1+l/\gamma)}}{\frac{1}{2}l\gamma\sigma^2 - l\mu}\right) e^{-\left(\frac{l}{2}\gamma\sigma^2 - l\mu\right)(T-t)} + \frac{\lambda\left(1+\frac{\gamma}{l}\right)^{-(1+l/\gamma)}}{\frac{l}{2}\gamma\sigma^2 - l\mu}, \quad (3.30)$$

which results in the solution for the optimal strategy

$$\delta^*(q=1,t) = \frac{1}{l}\ln\left(\left(e^{-lb} - \frac{\lambda\left(1+\frac{\gamma}{l}\right)^{-(1+l/\gamma)}}{\frac{l}{2}\gamma\sigma^2 - l\mu}\right)e^{-\left(\frac{l}{2}\gamma\sigma^2 - l\mu\right)(T-t)} + \frac{\lambda\left(1+\frac{\gamma}{l}\right)^{-(1+l/\gamma)}}{\frac{l}{2}\gamma\sigma^2 - l\mu}\right) + \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right). \tag{3.31}$$

This gives us analytic insight into the sensitivity of the solution with respect to the underlying parameters of the model.

3.5.2 Steady-state solution

It was also shown by Guéant et al. (2012b) that as time tends to infinity both the value function and thus the optimal trading strategy tend to a perpetual (steady-state) solution. This perpetual solution is found in terms of $w(q,t)$ by setting the time derivative to zero in (3.17), and is given by

$$\lim_{T\to\infty} w(q) = \frac{\eta^q}{q!}\prod_{j=1}^{q}\frac{1}{\alpha j - \beta} \tag{3.32}$$

with $w(q)$ now being t-independent, $\eta = \lambda\left(1+\frac{\gamma}{l}\right)^{-\left(1+\frac{l}{\gamma}\right)}$, $\alpha = \frac{l\gamma\sigma^2}{2}$, and $\beta = l\mu$. For a solution to exist the constraint

$$\beta < \alpha \tag{3.33}$$

must be satisfied which is equivalent to

$$\mu < \frac{\gamma\sigma^2}{2}. \tag{3.34}$$

This results in an optimal trading strategy of the form

$$\delta^*(q) = \frac{1}{l}\ln\left(\frac{\lambda}{l+\gamma}\times\frac{1}{\frac{1}{2}\gamma\sigma^2 q^2 - \mu q}\right) \tag{3.35}$$

which gives insight into the behaviour of the steady-state trading strategies, and the relationship between them and the parameters of the model.

3.5.3 Advantages of the model

The model of Guéant et al. (2012b), derived originally from Avellaneda and Stoikov (2008), is a successful attempt at a mathematically tractable model of a trading algorithm which liquidates orders passively in the order book. It encapsulates the two most

important aspects of an algorithm interacting with the order book seeking to liquid remaining inventory, these being controlling the short-term probability of execution, as well as the price risk.

Its mathematical tractability, which simultaneously integrates price and liquidity, is very appealing from a mathematical perspective. Firstly, liquidity appears as a statistical model with low dimensionality, with the parameters characterising intra-day features of the stock liquidity on a given exchange. Secondly, the stochastic behaviour of the price process is driven by a Brownian motion, which keeps consistent with classical models. Although the model is not an attempt to describe, in general, the processes arising in markets, it provides an insight into what concerns a trading algorithm's performance.

3.6 Summary

In this chapter we have given an overview of the literature related to optimal trading from the perspective of a sell-side firm. We paid particular attention to the framework of Guéant et al. (2012b) which models trading at the tactical level when liquidating assets, which can also be seen as modelling a market maker working in a single-side of the order book (which is the ask side in the case of liquidating). By choosing CARA utility functions, standard Brownian motion and an exponential intensity function with an absolute decay parameter, Guéant et al. (2012b) are able to cleverly build symmetry into the model from the start which allowed them to reduce the HJB equation to a system of ODEs.

In this thesis we expand on the framework of Guéant et al. (2012b) by introducing a different (and perhaps more realistic) concept to the optimal trading literature, that being the use of more general diffusion processes for the asset price, and more general intensity functions for the filling of orders in the limit-order book. Using more general diffusion processes and intensity functions will result in asset-dependent trading strategies, which to our knowledge are novel in the optimal trading literature at the tactical level (asset-dependent strategies have been found at the scheduling level, as discussed above, indicating they should also be present at the tactical level). We are also able to make a significant contribution by expanding the dimensionality of

the problem (multiple assets, stochastic volatility model), as well as examine a trader who can trade using both limit orders and market orders.

Chapter 4

Tactical Level Trading under General Diffusion Processes

The work in this chapter has recently been submitted for review:
Blair, J., Johnson, P., and Duck, P. (2015). Analysis of optimal liquidation in limit-order books. http://eprints.ma.man.ac.uk/2299.

In this chapter we extend the framework of Guéant et al. (2012b) by introducing more general diffusion processes for the asset price, and an exponential decay for the limit-order book which is relative to the asset price. It was over a decade after Almgren and Chriss (2001) published their seminal paper that Forsyth et al. (2012) examine the same problem under geometric Brownian motion. It is been argued through empirical evidence that geometric Brownian motion is favourable in modelling asset prices over standard Brownian motion (see Osborne, 1959). An important point here is to note that geometric Brownian motion can be used over any time scale, for both long and short trading horizons, and avoids the fallacy of negative price scenarios that can appear in standard Brownian motion (over longer time scales). Although it is true that under small-time horizons standard Brownian motion approximates geometric Brownian motion, many of the problems in optimal execution consider infinite horizon problems, for example Guéant et al. (2012b), Cartea and Jaimungal (2013b), Bayraktar and Ludkovski (2012) and Schied and Schöneborn (2009) to name a few, for which this approximation is no longer valid.

By choosing CARA utility functions, standard Brownian motion and an absolute decay parameter for the intensity function Guéant et al. (2012b) are able to cleverly

build symmetry into the model from the start which allows them to reduce the HJB equation to a system of ODEs. However, we attempt to build a model with geometric Brownian motion (and mean-reverting processes) for the asset price, a CARA utility function, a proportional control parameter (the additional amount we ask for the asset), and a proportional decay parameter, meaning that this symmetry no longer exists. This not only complicates the model but produces interesting (and arguably more realistic) results while doing so. Having an asset-dependent control parameter will induce some transparency in our results, given the optimal trading strategy will be expressed as a percentage of the asset. This novel combination will result in asset-dependent trading strategies, which to our knowledge is a unique concept in this framework of literature. This also results in not selling the asset for a negative price (effectively paying someone to take the asset), which can occur in the work of Guéant et al. (2012b), amongst others.

Although the equations are more complex to solve, which could be a reason geometric Brownian motion and more general diffusion processes were avoided in the initial framework, having more complex equations allows us to consider a richer set of solutions than has shown to be the case in the previous literature. We are able to carry out interesting analytic (asymptotic) analysis, finding analytic solutions in various limits, and the development of a singularity when considering a steady-state framework. We have thus taken a combined approach in solving the problem which involves both theoretically (analytically/asymptotically) examining the problem and using numerical methods to obtain solutions, with each approach informing and confirming the other. As we will see shortly, the methods we introduce are not constrained to the use of geometric Brownian motion as the driving process for the asset price, as we extend to a CIR process in section 4.4.

4.1 Problem formulation

We consider a trader who wishes to maximise his expected utility, given a portfolio of assets to liquidate, before a specified terminal time, T. We assume at time $t = 0$ that the trader starts with an initial inventory of $q(0)$ assets, with $q(t) \in \mathbb{Z}^+$, and an initial wealth $X(0)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0, T])$. We

assume the reference price (which may refer to the best-bid price, best-ask price or mid-price) of the asset, $S(t)$, follows a geometric Brownian motion, and so the diffusion process is defined as

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \tag{4.1}$$

with μ as the constant relative drift, σ as the constant relative volatility and $W(t)$ as a Wiener process which is $\mathcal{F}_t$ measurable.

The trader will continuously post orders into the ask side of the limit-order book for price $S^a(t)$ which is $\delta = \delta(t, X, q, S)$ percent greater than the asset value $S(t)$,

$$S^a(t) = S\left(t\right)\left(1 + \delta\right). \tag{4.2}$$

This is another distinct aspect to our approach as Guéant et al. (2012b) do not model the control as a percent of the asset price but instead have asking price $S^a(t) = S\left(t\right)+\delta$.

Asset sales follow a Poisson process, $N(t)$, with time-dependent intensity, which is $\mathcal{F}_t$ measurable and independent of $W(t)$

$$dq(t) = -dN(t), \tag{4.3}$$

for $q\left(t\right) > 0$, thus assuming the trader becomes inactive after liquidating and does not short sell. We assume trades are of unitary size, which are assumed to be lots or of the average trade size, consistent with that of previous literature as discussed in section 3.5. For each occurrence of a trade, the trader's wealth increases by the amount that asset was sold for, and thus the dynamics of the wealth is given by

$$dX(t) = S(t)\left(1 + \delta\right)dN(t), \tag{4.4}$$

where $N(t)$ is the same Poisson process as before. Therefore, the values of $q(t)$ and $X(t)$ change simultaneously, according to (4.3) and (4.4) respectively. $N(t)$ has intensity $\Lambda(\delta)$ which takes the form:

$$\Lambda(\delta) = \lambda e^{-l\left(\frac{S^a - S}{S}\right)} = \lambda e^{-l\delta} \tag{4.5}$$

for some positive constants λ and l. We notice (4.5) is contrasting in its derivation to (3.15) given the optimal control is proportional to the asset price and thus the exponential decay of the order book, l, is also proportional to the asset price. We discuss this comparison in more detail in section 4.1.2.

The objective is to liquidate this portfolio before some final time T. Assets that are not liquidated before this time will be sold in the market for their current value. In some of the literature, such as Guéant et al. (2012b), a terminal penalty is included such that the assets are sold at a discount of their actual price at the terminal price, which is a constant per asset and thus independent of the asset price. Others, such as Avellaneda and Stoikov (2008) and Guéant et al. (2012a), neglect inclusion of this terminal penalty. We believe the arguments of the latter studies are more profound. If the reference price was the best bid or mid-price then it could be argued that execution at the terminal time should be immediate. If the reference was the best ask, given the rate at which orders are filled at the best ask (as found in empirical results in chapter 9) it should be viable to liquidate the remaining position (which should be relatively small) almost immediately. We were considering adding a penalty, which we would make as a percent of the asset rather than a fixed amount which is independent of the asset price. We believe the former would be more realistic as the asset would never be sold for a negative price, effectively paying someone to take it. Inclusion of the penalty is quite trivial, and we outline its potential impact throughout the chapter. In terms of both solving this problem numerically and investigating it asymptotically, the terminal penalty would make little difference to both the difficulty of the methods used and the results obtained, given a realistic value for the (relative) penalty would be quite small. We have therefore chosen to neglect it.

The utility we seek to maximise takes the form of a negative exponential function, consistent with the previous literature of Avellaneda and Stoikov (2008) and Guéant et al. (2012b), the properties of which have been discussed in section 2.3. We define our value function, $u(t, X, q, S)$, as the maximum expected utility at time t

$$u(t, X, q, S) = \sup_{\delta(t) \in \mathcal{A}} \mathbb{E}\left[-e^{-\gamma(X(T)+q(T)S(T))}\right], \tag{4.6}$$

where $\gamma > 0$ is the risk-aversion characterising the trader and $\mathcal{A} \subseteq (-1, \infty)$ is the set of admissible trading strategies. Assuming the trader starts with some non-negative wealth and with some positive quantity of inventory the term $\gamma(X + qS)$ is strictly positive. Therefore the objective function, and by definition the value function, is bounded in the set $[-1, 0)$. We should note that the lower bound of the admissible strategy occurs naturally due to the trader never wanting to sell his asset for a negative

price, implying he would be paying someone to take the asset from him, and hence it does not have to be explicitly implemented. Using the form of the asking price $S^a = S + \delta$, under a standard Brownian motion, trading strategies are independent of the asset price and hence selling for a negative price can occur, which is a fundamental flaw of previous models.

Given the optimisation problem of (4.6), an HJB equation can be derived by applying the Bellman (1957) principle of optimality and using Itô's lemma:

$$u_t(t,X,q,S) + \mu S u_S(t,X,q,S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t,X,q,S)$$
$$+\sup_{\delta}\left[\lambda e^{-l\delta}\left(u(t,X+S\left(1+\delta\right),q-1,S) - u(t,X,q,S)\right)\right] = 0, \tag{4.7}$$

with conditions

$$u(T,X,q,S) = -e^{-\gamma(X+qS)}, \tag{4.8}$$

$$u(t,X,0,S) = -e^{-\gamma X}. \tag{4.9}$$

We will now discuss the derivation of the HJB equation (4.7), followed by a verification theorem which is used to prove that the solution we obtain from the HJB equation (4.7) is in fact the solution of the optimal control problem (4.6) we wish to solve.

4.1.1 Derivation of HJB equation and a verification theorem

Deriving HJB equation using Bellman principle

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0,T])$. Define the stochastic processes adapted to the filtration $\mathcal{F}_t$ as:

$$\begin{bmatrix} dS(t) \\ dq(t) \\ dX(t) \end{bmatrix} = \begin{bmatrix} \mu S\left(t\right) \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} \sigma S\left(t\right) dW(t) \\ -dN(t) \\ S(t)\left(1+\delta\right) dN(t) \end{bmatrix}, \tag{4.10}$$

which satisfies linear growth and Lipschitz continuity conditions and as such unique solutions exist, as discussed in section 2.4.3. We define the value function $u(t,X,q,S)$ as

$$u(t,X,q,S) = \sup_{\delta(t)\in\mathcal{A}} \mathbb{E}\left[\Phi(X(T),q(T),S(T))\right], \tag{4.11}$$

where $\mathcal{A} \subseteq (-1, \infty)$ is the set of admissible trading strategies and $\Phi(\cdot)$ is the form of the utility function. A unique solution of (4.11) exists given the existence and uniqueness of the underlying stochastic differential equations (4.10).

Bellman's Principle of Optimality reads:

$$u(t, X, q, S) = \sup_{\delta(t) \in \mathcal{A}} \mathbb{E}\left[u(t+h, X(t+h), q(t+h), S(t+h))\right]. \tag{4.12}$$

The intuition behind Bellman's Principle is that the supremum of the value function is found over $[t, T]$ if we find the supremum over $[t, t+h]$ and then continue optimally over $[t+h, T]$ with the state values at time $t+h$ as the initial values.

To derive the dynamic programming equation we will first assume $\delta(s) = \delta$ for $s \in [t, t+h]$, i.e. the control parameter is constant over the interval $[t, t+h]$. From (4.12) we have:

$$u(t, X, q, S) \geq \mathbb{E}\left[u(t+h, X(t+h), q(t+h), S(t+h))\right]. \tag{4.13}$$

Subtracting $u(t, X, q, S)$ from both sides gives

$$0 \geq \mathbb{E}\left[u(t+h, X(t+h), q(t+h), S(t+h)) - u(t, X, q, S)\right]. \tag{4.14}$$

Assuming $u \in C^{1,2}([0, T), \mathbb{R}^+)$ in t and S respectively, we now apply Itô's lemma to $u(t+h, X(t+h), q(t+h), S(t+h)) - u(t, X, q, S)$, such that

$$\begin{aligned}
u\left(t+h, X\left(t+h\right), q\left(t+h\right), S\left(t+h\right)\right) - u\left(t, X\left(t\right), q\left(t\right), S\left(t\right)\right) = \\
\int_t^{t+h} \Big(u_t\left(s, X\left(s\right), q\left(s\right), S\left(s\right)\right) + \mu S\left(s\right) u_S\left(s, X\left(s\right), q\left(s\right), S\left(s\right)\right) \\
+ \frac{1}{2}\sigma^2 S\left(s\right)^2 u_{SS}\left(s, X\left(s\right), q\left(s\right), S\left(s\right)\right) + \\
\left(u\left(s, X\left(s\right) + \Delta X\left(s\right), q\left(s\right) + \Delta q\left(s\right), S\left(s\right)\right) - u\left(s, X\left(s\right), q\left(s\right), S\left(s\right)\right)\right) \Lambda\left(\delta\right) \Big) ds \\
+ \int_t^{t+h} \sigma S\left(s\right) u_S\left(s, X\left(s\right), q\left(s\right), S\left(s\right)\right) dW\left(s\right) \\
+ \int_t^{t+h} \Big(u\left(s, X\left(s\right) + \Delta X\left(s\right), q\left(s\right) + \Delta q\left(s\right), S\left(s\right)\right) - u\left(s, X\left(s\right), q\left(s\right), S\left(s\right)\right) \Big) d\tilde{N}\left(s\right),
\end{aligned} \tag{4.15}$$

ignoring quadratic variation terms which equal zero, with $\tilde{N}(t)$ denoting a compensating Poisson process with intensity $\Lambda(\delta) = \lambda e^{-l\delta}$. Here we are assuming continuity in the value function, and not the underlying processes, and for further insight we refer the reader to Øksendal and Sulem (2005).

Substituting (4.15) into (4.14), taking the expectation, dividing by h and letting $h \to 0$ we obtain:

$$0 \geq \mathcal{L}u(t, X, q, S) + \Lambda(\delta) \left(u(t, X + S\left(1 + \delta\right), q - 1, S) - u(t, X, q, S) \right), \tag{4.16}$$

where we define the generator $\mathcal{L}$ of the geometric Brownian motion as

$$\mathcal{L}u(t, X, q, S) = u_t(t, X, q, S) + \mu S u_S(t, X, q, S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t, X, q, S). \tag{4.17}$$

Here we assume the final two integrals are martingales, which we will show to be true in the verification theorem below.

Equation (4.16) holds through for any $\delta \in \mathcal{A}$. Assume now that the optimal control is given by the Markov control policy $\delta^* = \delta^*(s, X^*(s), q^*(s), S(s))$. Here $X^*(s)$ and $q^*(s)$ denote the optimal state process controlled under δ^*. Using the optimal policy and making the same assessment as above we obtain

$$0 = \mathcal{L}u(t, X, q, S) + \Lambda(\delta^*) \left(u(t, X + S\left(1 + \delta^*\right), q - 1, S) - u(t, X, q, S) \right). \tag{4.18}$$

Combining (4.16) and (4.18) we obtain

$$0 = \sup_{\delta \in \mathcal{A}} [\mathcal{L}u(t, X, q, S) + \Lambda(\delta) \left(u(t, X + S\left(1 + \delta\right), q - 1, S) - u(t, X, q, S) \right)]. \tag{4.19}$$

As $u(t, X, q, S)$ is already defined as a supremum, defined by (4.11), we can take the generator process outside the supremum so (4.19) becomes

$$0 = \mathcal{L}u(t, X, q, S) + \sup_{\delta \in \mathcal{A}} [\Lambda(\delta) \left(u(t, X + S\left(1 + \delta\right), q - 1, S) - u(t, X, q, S) \right)] \tag{4.20}$$

with

$$u(T, X, q, S) = \Phi(X, q, S), \tag{4.21}$$

$$u(t, X, 0, S) = \Phi(X, 0, S). \tag{4.22}$$

Verification theorem

We shall now show that the solution we obtain from solving (4.20) is in fact the solution of (4.11). Define

$$\phi\left(t, X, q, S\right) := \mathbb{E}\left[\Phi(X(T), q(T), S(T))\right], \tag{4.23}$$

with $\phi \in C^{1,2}\left([0, T), \mathbb{R}^+\right)$ in t and S respectively. $\Phi(X, q, S) = -e^{-\gamma(X + qS)}$ for $q \in \mathbb{Z}^+$, $\gamma, X, S \in \mathbb{R}^+$ and hence Φ is bounded in $[-1, 0)$. By definition ϕ is also bounded in

$[-1, 0)$ and hence it satisfies a polynomial growth condition, i.e. there exists constants $C \in \mathbb{R}^+$ and $k \in \mathbb{Z}^+$ such that

$$|\phi(t, X, q, S)| \leq C\left(1 + |X + qS|^k\right), \quad \forall (t, X, q, S) \in [0, T] \times \mathbb{R}^+ \times \mathbb{Z}^+ \times \mathbb{R}^+,$$

Let $\phi^*(t, X, q, S)$ be the solution of

$$0 = \mathcal{L}\phi(t, X, q, S) + \sup_{\delta \in \mathcal{A}}[\Lambda(\delta)\left(\phi(t, X + S + \delta, q - 1, S) - \phi(t, X, q, S)\right)], \tag{4.24}$$

with $\mathcal{L}$ as the generator process defined by (4.17). We now want to show that $\phi^*(t, X, q, S)$ is identical to the solution obtained by solving (4.11). We begin by applying a Taylor's expansion to $\phi^*(T, X, q, S)$ which gives

$$\phi^*(T, X, q, S) = \phi^*(t, X, q, S) + \int_t^T d\phi^*(s, X, q, S). \tag{4.25}$$

Since $\phi \in C^{1,2}([0, T), \mathbb{R}^+)$ in t and S we have for all $(t, X, q, S) \in [0, T) \times \mathbb{R}^+ \times \mathbb{Z}^+ \times \mathbb{R}^+, r \in [t, T)$, and any stopping time τ valued in $[t, \infty)$, by Itô's formula

$$\begin{aligned}
\phi^*(r \wedge (t+h), X(r \wedge (t+h)), q(r \wedge (t+h)), S(r \wedge (t+h))) - \phi^*(t, X(t), q(t), S(t)) = \\
\int_t^{r \wedge (t+h)} \Big(\phi_t^*(s, X, q, S) + \mu S(s)\phi_S^*(s, X, q, S) + \frac{1}{2}\sigma^2 S(s)^2 \phi_{SS}^*(s, X, q, S) + \\
\left(\phi^*(s, X(s) + \Delta X(s), q(s) + \Delta q(s), S(s)) - \phi^*(s, X(s), q(s), S(s))\right)\Lambda(\delta)\Big) ds \\
+ \int_t^{r \wedge (t+h)} \sigma S(s)\phi_S^*(s, X(s), q(s), S(s)) dW(s) + \\
\int_t^{r \wedge (t+h)} \left(\phi^*(s, X(s) + \Delta X(s), q(s) + \Delta q(s), S(s)) - \phi^*(s, X(s), q(s), S(s))\right) d\tilde{N}(s),
\end{aligned}$$

with $x \wedge y = \max(x, y)$. We choose

$$\tau = \tau_n = \inf\left(r \geq t : \int_t^r |\sigma S(s)\phi_S^*(s, X(s), q(s), S(s))|^2 ds \geq n\right)$$

and notice that $\tau_n \to \infty$ when n goes to infinity. The stopped process

$$\int_t^{r \wedge \tau_n} \sigma S(s)\phi_S^*(s, X(s), q(s), S(s)) dW(s), \quad t \leq r \leq T,$$

is then a martingale with

$$\mathbb{E}\Big[\int_t^T \left(\sigma S(s)\phi_S^*(s, X(s), q(s), S(s))\right)^2 ds\Big] < \infty.$$

As the value function is bounded between -1 and 0, and noticing that δ is bounded from below and hence $\Lambda(\delta)$ is bounded we have that

$$\mathbb{E}\Big[\int_t^T |\phi^*(s, X(s) + \Delta X(s), q(s) + \Delta q(s), S(s))|\Lambda(\delta) ds\Big] < \infty,$$

and

$$\mathbb{E}\Big[\int_t^T |\phi^*(s, X(s), q(s), S(s))|\Lambda(\delta)\, ds\Big] < \infty,$$

so the final integral is a martingale with expectation zero.

We take the expected value of both sides of (4.25). The first integral goes to zero given it is identical to (4.24), as do the final two integrals as they have been shown to be martingales. This results in

$$\mathbb{E}\left[\Phi\left(X^*, q^*, S\right)\right] = \phi^*\left(t, X, q, S\right), \tag{4.26}$$

as $\phi^*\left(T, X, q, S\right) = \Phi\left(X^*, q^*, S\right)$. If δ was chosen arbitrarily we would have

$$\mathbb{E}\left[\Phi\left(X, q, S\right)\right] \leq \phi^*\left(t, X, q, S\right). \tag{4.27}$$

Thus we have

$$\sup_{\delta \in \mathcal{A}} \mathbb{E}\left[\Phi(X(T), q(T), S(T))\right] = \mathbb{E}\left[\Phi\left(X^*, q^*, S\right)\right] = \phi^*\left(t, X, q, S\right) = u(t, X, q, S). \tag{4.28}$$

We have thus shown that the solution found in the HJB equation (4.7) is the solution of the original optimisation problem (4.6). For further reading on the theory of optimal control the authors recommend Øksendal and Sulem (2005) and Pham (2009).

4.1.2 Reduction of the problem

The problem as stated in (4.7) is a four-dimensional HJB equation. Guéant et al. (2012b) and Cartea and Jaimungal (2013b) both reduce their (similar) problems to a system of ODEs by suggesting an ansatz form of the solution which assumes the solution can be separated into two functions. One function involves variables X, S and q, and the other involves variables q and t, with the PDE being reduced to a t-dependent ODE, thus arriving with trading strategies independent of X and S. Due to the use of geometric Brownian motion, as opposed to the standard Brownian motion used by Guéant et al. (2012b) and Cartea and Jaimungal (2013b), reducing the problem to a system of ODEs is no longer possible. However, we can still make use of the form of the utility function to make a significant reduction to the complexity of the problem.

We begin by assuming an ansatz solution

$$u\left(t,X,q,S\right)=e^{-\gamma X}f\left(\tau,q,S\right), \tag{4.29}$$

where we also use a change in the time variable $\tau=T-t$ so that we are now solving forward in τ rather than backward in t. Using this form of the solution we obtain:

$$\begin{aligned}&-e^{-\gamma X}f_{\tau}\left(\tau,q,S\right)+\mu Se^{-\gamma X}f_{S}\left(\tau,q,S\right)+\frac{1}{2}\sigma^{2}S^{2}e^{-\gamma X}f_{SS}\left(\tau,q,S\right)\\&+\sup_{\delta}\left[\lambda e^{-l\delta}\left(e^{-\gamma(X+S(1+\delta))}f\left(\tau,q-1,S\right)-e^{-\gamma X}f\left(\tau,q,S\right)\right)\right]=0.\end{aligned} \tag{4.30}$$

We can solve for the optimal control by differentiating the supremum with respect to the control, δ, and setting the result equal to zero which locates the stationary point. Solving this we obtain:

$$\delta^{*}\left(\tau,q,S\right)=\frac{1}{S\gamma}\ln\left(\frac{\left(\gamma S+l\right)f\left(\tau,q-1,S\right)}{lf\left(\tau,q,S\right)}\right)-1, \tag{4.31}$$

which we notice is independent of X, hence confirming the use of our ansatz solution. We did the second derivative test of the supremum to confirm this was in fact the maximum. We also notice the optimal strategy is a function of S which is a unique addition to the literature for this general class of problem. Using the form of (4.31) we find the asking price of the asset to be

$$S^{a*}\left(\tau\right)=\frac{1}{\gamma}\ln\left(\frac{\left(\gamma S+l\right)f\left(\tau,q-1,S\right)}{lf\left(\tau,q,S\right)}\right). \tag{4.32}$$

To confirm with previous literature, notably Guéant et al. (2012b), we will primarily keep our focus on the optimal control, δ^{*}, when examining our results.

Substituting (4.31) into (4.30) and cancelling common factors we obtain the following non-linear PDE

$$\begin{aligned}&-f_{\tau}\left(\tau,q,S\right)+\mu Sf_{S}\left(\tau,q,S\right)+\frac{1}{2}\sigma^{2}S^{2}f_{SS}\left(\tau,q,S\right)\\&-\frac{\lambda e^{l}\gamma Sf\left(\tau,q,S\right)}{S\gamma+l}\left(\frac{lf\left(\tau,q,S\right)}{\left(S\gamma+l\right)f\left(\tau,q-1,S\right)}\right)^{\frac{l}{S\gamma}}=0,\end{aligned} \tag{4.33}$$

with

$$f\left(0,q,S\right)=-e^{-\gamma qS}, \tag{4.34}$$

$$f\left(\tau,0,S\right)=-1. \tag{4.35}$$

When discussing results we shall now refer to $f\left(\tau,q,S\right)$ as the value function.

If we were to follow the framework of Guéant et al. (2012b) and use an absolute control parameter, so our asking price would be $S^a(t) = S(t) + \delta$, but still a proportional exponential decay parameter, which would take the form

$$\Lambda(\delta) = \lambda e^{-l\left(\frac{S^a - S}{S}\right)} = \lambda e^{-\frac{l\delta}{S}}, \tag{4.36}$$

we would still have an asset-dependent trading strategy, given by

$$\delta^*(\tau, q, S) = \frac{1}{\gamma} \ln\left(\frac{(\gamma S + l) f(\tau, q-1, S)}{l f(\tau, q, S)}\right) - S. \tag{4.37}$$

If we were to not use a proportional exponential-decay parameter, i.e.

$$\Lambda(\delta) = \lambda e^{-l(S^a - S)} = \lambda e^{-l\delta}, \tag{4.38}$$

the trading strategies we obtain would take the form

$$\delta^*(\tau, q, S) = \frac{1}{\gamma} \ln\left(\frac{(\gamma + l) f(\tau, q-1, S)}{l f(\tau, q, S)}\right) - S, \tag{4.39}$$

which would lead to a non-linear PDE analogous to (4.33) with trading strategies that are asset-price dependent. As discussed in section 3.5, Guéant et al. (2012b) use asset independent parameters and thus arrive with trading strategies which are asset-price independent, making them arguably unrealistic from a practical viewpoint. The inclusion of the asset proportional exponential-decay parameter provides a dynamic shape for the limit-order book, evolving as the asset price changes, becoming wider for higher asset prices and thinner for lower asset prices. This is emphasised in figure 4.1 which is a plot of (4.36) for various values of S, which portrays the wider limit-order book for increasing S.

The works of Avellaneda and Stoikov (2008), Cartea and Jaimungal (2013b) and Huitema (2013), to name a few, use the form of (4.38) for the intensity of the Poisson process, all allowing for negative δ, as do we in this thesis. Guéant and Lehalle (2013) have a twofold argument for the use of negative δ. The first being that it can be seen as indirectly introducing execution costs for liquidity taking orders, given we want execution to occur quicker than the average rate. Secondly, high-frequency traders have the ability to rapidly cancel orders and as such fill and/or kill orders, along with other limit orders, are used rather than market orders as these market orders may be retracted rapidly. It is also these two arguments they use against the use of a power

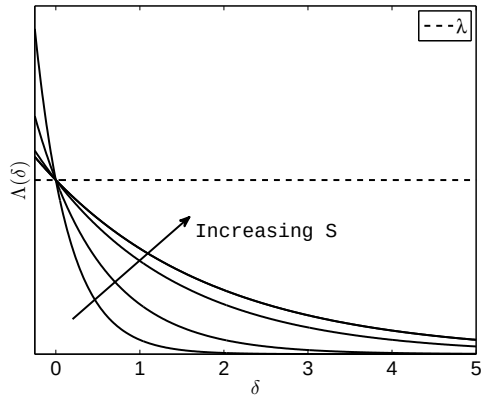


Figure 4.1: Limit-order book widens for proportional exponential-decay parameter as S increases, according to (4.36), with asking price taking the form $S^a(t) = S(t) + \delta$

intensity function, as examined by Bayraktar and Ludkovski (2012), which blows up at $\delta = 0$, indicating infinite liquidity. However, this is a two sided argument. If you are a liquidity consumer, then you should be guaranteed execution at some price, which is not the case under exponential intensity. It is for this reason that we extend the framework in chapter 8 to power intensity functions.

4.1.3 Rescaling of the PDE

We shall now perform a rescaling (effectively a non-dimensionalisation) of this PDE to eliminate several of the parameters. We use the following change of variables:

$$\tilde{\tau} = \lambda\tau, \quad \tilde{S} = S\gamma, \quad \tilde{\mu} = \frac{\mu}{\lambda}, \quad \tilde{\sigma} = \frac{\sigma}{\sqrt{\lambda}}. \tag{4.40}$$

Interestingly, the scaling of time with the rate λ came about when investigating the small-time solution, which we examine below in section 4.2. We noticed the length of time in which the small-time solution was valid was inversely proportional to the parameter λ, hence the above rescaling. Rescaling μ and σ we eliminate λ from the equation. When examining the numerical solutions, which we do below in section 4.1.5, we found the scaling which was present between S and γ, thus allowing us to rescale out the γ parameter. $\tilde{S}$ is now the risk-adjusted asset price, given that it is S scaled on γ. From now we will refer to $\tilde{S}$ as the asset price, dropping 'risk-adjusted'.

All the new variables of (4.40) are then dimensionless. This results in the non-linear

PDE:

$$-f_{\tilde{\tau}}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\mu}\tilde{S}f_{\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^{2}\tilde{S}^{2}f_{\tilde{S}\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)$$
$$-\frac{e^{l}\tilde{S}f\left(\tilde{\tau},q,\tilde{S}\right)}{\tilde{S}+l}\left(\frac{lf\left(\tilde{\tau},q,\tilde{S}\right)}{\left(\tilde{S}+l\right)f\left(\tilde{\tau},q-1,\tilde{S}\right)}\right)^{\frac{l}{\tilde{S}}}=0, \tag{4.41}$$

with

$$f\left(0,q,\tilde{S}\right)=-e^{-q\tilde{S}}, \tag{4.42}$$

$$f\left(\tilde{\tau},0,\tilde{S}\right)=-1. \tag{4.43}$$

Equation (4.41) is now dimensionless as are the conditions (4.42) and (4.43). The optimal trading strategy in dimensionless form is then

$$\delta^{*}\left(\tilde{\tau},q,\tilde{S}\right)=\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)f\left(\tilde{\tau},q-1,\tilde{S}\right)}{lf\left(\tilde{\tau},q,\tilde{S}\right)}\right)-1, \tag{4.44}$$

and the risk-adjusted asking price is

$$\tilde{S}^{a*}\left(\tilde{\tau}\right)=\ln\left(\frac{\left(\tilde{S}+l\right)f\left(\tilde{\tau},q-1,\tilde{S}\right)}{lf\left(\tilde{\tau},q,\tilde{S}\right)}\right). \tag{4.45}$$

The optimal strategy at $\tilde{\tau}=0$ can be confirmed by substituting (4.42) into (4.44) which results in:

$$\delta^{*}\left(\tilde{\tau}=0,q,\tilde{S}\right)=\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)}{l}\right). \tag{4.46}$$

As discussed above, if we were to include a penalty at $\tilde{\tau}=0$ we would make it a proportional penalty, b, as to neglect the possibility of selling for a negative value. The initial condition (4.42) would then take the form

$$f\left(0,q,\tilde{S}\right)=-e^{-q\tilde{S}(1-b)}, \tag{4.47}$$

with the optimal strategy at $\tilde{\tau}=0$ taking the form

$$\delta^{*}\left(\tilde{\tau}=0,q,\tilde{S}\right)=\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)}{l}\right)-b. \tag{4.48}$$

The same boundary conditions derived below would be suitable given a realistic value for b would be $b\ll O\left(1\right)$, and the trading strategies would still have the same form as they diffuse from the initial time, but would be shifted to slightly lower values.

Boundary conditions

To make this problem well-posed we must consider boundary conditions for $\tilde{S} = 0$ and $\tilde{S} \to \infty$. For $\tilde{S} = 0$ we take the limit of (4.41) which results in

$$\frac{\partial f}{\partial \tilde{\tau}} \left(\tilde{\tau}, q, \tilde{S} = 0 \right) = 0. \tag{4.49}$$

Given (4.49) and (4.42) the $\tilde{S} = 0$ boundary condition we arrive at is

$$f \left(\tilde{\tau}, q, \tilde{S} = 0 \right) = -1, \tag{4.50}$$

and so

$$\delta^*(\tilde{\tau}, q, \tilde{S} \to 0) = f_{\tilde{S}}(\tilde{\tau}, q, 0) - f_{\tilde{S}}(\tilde{\tau}, q - 1, 0) + \frac{1}{l} - 1. \tag{4.51}$$

For $\tilde{S} \to \infty$, $f \left(\tilde{\tau}, q, \tilde{S} \right) \to 0$ although for numerical purposes we assume a (softer) Neumann boundary condition

$$\frac{\partial f}{\partial \tilde{S}} \left(\tilde{\tau}, q, \tilde{S} \to \infty \right) \to 0, \tag{4.52}$$

which, from investigation, we found to be satisfactory. From a financial perspective we can interpret this as the trader not being able to become any more satisfied when he has an asset worth an infinite amount of money.

4.1.4 Numerical methods

To solve (4.41) we use a finite-difference scheme. We have tested both implementing implicit differences on the derivative terms while taking the non-linear term as an explicit term, and using an iterative Crank-Nicolson scheme, with each method confirming the other.

Using implicit differences and an explicit non-linear term we negate the need to use an iterative scheme as the non-linear terms remain on the right-hand-side of the linear system $\mathbf{Af} = \mathbf{b}$. We can thus use a direct solver such as the Thomas algorithm at each time step.

Solving using Crank-Nicolson is a bit trickier as we need to use an iterative scheme, given the non-linearity of (4.41). Assume we are calculating the solution, $\mathbf{f}^{n+1}$, at time-step $n + 1$. At the $k^{th} + 1$ iteration we have:

$$\mathbf{f}^{n+1,(k+1)} = \mathbf{f}^{n+1,(k)} + \boldsymbol{\varepsilon}^{(k)} \tag{4.53}$$

in which a reasonable guess for the initial solution $\mathbf{f}^{n+1,(0)}$ is the solution at time-step n, $\mathbf{f}^n$. Implementing Crank-Nicolson finite-differences in (4.41) results in

$$\begin{aligned}
&-\frac{f_{i,q}^{n+1}-f_{i,q}^{n}}{\Delta\tilde{\tau}}+\tilde{\mu}i\Delta\tilde{S}\left(\frac{f_{i+1,q}^{n+1}-f_{i-1,q}^{n+1}}{4\Delta\tilde{S}}+\frac{f_{i+1,q}^{n}-f_{i-1,q}^{n}}{4\Delta\tilde{S}}\right)\\
&+\frac{1}{2}\tilde{\sigma}^2i^2\Delta\tilde{S}^2\left(\frac{f_{i+1,q}^{n+1}-2f_{i,q}^{n+1}+f_{i-1,q}^{n+1}}{2\Delta\tilde{S}^2}+\frac{f_{i+1,q}^{n}-2f_{i,q}^{n}+f_{i-1,q}^{n}}{2\Delta\tilde{S}^2}\right)\\
&-\frac{e^li\Delta\tilde{S}f_{i,q}^{n+1}}{i\Delta\tilde{S}+l}\left(\frac{lf_{i,q}^{n+1}}{\left(i\Delta\tilde{S}+l\right)f_{i,q-1}^{n+1}}\right)^{\frac{l}{i\Delta\tilde{S}}}-\frac{e^li\Delta\tilde{S}f_{i,q}^{n}}{i\Delta\tilde{S}+l}\left(\frac{lf_{i,q}^{n}}{\left(i\Delta\tilde{S}+l\right)f_{i,q-1}^{n}}\right)^{\frac{l}{i\Delta\tilde{S}}}=0,
\end{aligned}\tag{4.54}$$

where $f_{i,q}^n = f\left(n\Delta\tilde{\tau}, q, i\Delta\tilde{S}\right)$. We substitute (4.53) into (4.54) for $\mathbf{f}^{n+1}$. Now the unknowns we are solving for is $\boldsymbol{\varepsilon}$. The system is still non-linear in $\mathbf{f}$, and non-linear in $\boldsymbol{\varepsilon}$. To linearise in $\boldsymbol{\varepsilon}$ we take the limit of the non-linear term up to $O(\boldsymbol{\varepsilon})$, using a Taylor expansion, as $\boldsymbol{\varepsilon}\to 0$. Rearranging we have a linear system of equations which can be wrote in the form $\mathbf{A}\boldsymbol{\varepsilon}=\mathbf{b}$ and solved directly for $\boldsymbol{\varepsilon}$. The right-hand-side $\mathbf{b}$ contains terms of $\mathbf{f}^n$ and $\mathbf{f}^{n+1,(k)}$ which are all known. At each iteration we update $\mathbf{f}^{n+1}$ according to (4.53) and repeat until some specified tolerance criteria is satisfied.

For the former we expected our method to exhibit $O\left(\Delta\tilde{\tau}, \Delta\tilde{S}^2\right)$ convergence which we found was the case; for the latter we expected $O\left(\Delta\tilde{\tau}^2, \Delta\tilde{S}^2\right)$ convergence which we also confirmed, where $\Delta\tilde{\tau}$ and $\Delta\tilde{S}$ are the grid sizes in $\tilde{\tau}$ and $\tilde{S}$ respectively. This can be seen in table 4.1. A number of calculations were performed on a transformed grid $Y=\ln\tilde{S}$, which did have some advantages for certain calculations, such as the steady-state problem, but also exhibited some disadvantages (including the need for two, rather than one, domain truncation parameters). However the results thus obtained did provide a useful check of the results from the untransformed ($\tilde{S}$) grid. We will return to the transformed grid when we examine the perpetual case in section 4.3.

4.1.5 Numerical examples

We shall focus mainly on the behaviour of the optimal trading strategy, δ^*, given that it is a transformation of the value function and, from a financial perspective, is more transparent than examining the value function per se. Unless specified, the parameter values used are: $\tilde{\mu}=0.04, \tilde{\sigma}=0.4, l=25$ and $\tilde{\tau}=2$. In chapter 9 we will calibrate

Table 4.1: Convergence of finite-difference scheme for single underlying

n	m	$f_{I-E}\left(\tilde{\tau}=2,q=1,\tilde{S}=5\right)$	ratio	$f_{C.N.}\left(\tilde{\tau}=2,q=1,\tilde{S}=5\right)$	ratio
501	2001	-0.012105020201	NA	-0.012104988415	NA
1001	2001	-0.012104875712	NA	-0.012104843964	NA
2001	2001	-0.012104839554	4	-0.012104807874	4
4001	2001	-0.012104830584	4	-0.012104798766	4
8001	2001	-0.012104828267	4	-0.012104796598	4
8001	501	-0.012104922072	NA	-0.012104796606	NA
8001	1001	-0.012104859753	NA	-0.012104796514	NA
8001	2001	-0.012104828237	2	-0.012104796591	4
8001	4001	-0.012104812371	2	-0.012104796485	4
8001	8001	-0.012104804444	2	-0.012104796484	4

n and m are the number of space and time points respectively. We have set $\tilde{S}_{max} = 20, \tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$. f_{I-E} is the value corresponding to the implicit derivatives-explicit non-linear term finite-difference scheme and $f_{C.N.}$ is the value corresponding to the Crank-Nicolson scheme. The ratio gives the ratio of the change in errors for successive grids.

the model using real world data. However the values used in this chapter are used to highlight the properties and behaviour of the solutions.

The results can be seen in figure 4.2 for $q = 1$. Looking first at the optimal strategy over time, figure 4.2(a), for various values of $\tilde{S}$ we notice that the general behaviour of the optimal strategy is decreasing for increasing $\tilde{S}$. This is due to the CARA characteristic of the utility function. Given the trader is risk-averse, the trader will dislike the higher absolute volatility present as $\tilde{S}$ increases. He will thus opt to sell the asset quicker, so as to lock in the current price and avoid the risk of the asset price moving against him. What is particularly vivid for larger values of $\tilde{S}$ but also true for smaller values, is that as $\tilde{\tau}$ increases the optimal trading strategy tends to a perpetual (time-independent) type solution. This can be seen in the dot-dash line tending to a constant value around $\tilde{\tau} \sim 0.3$. It can also be seen in figure 4.2(b) that the solution is tending to a perpetual-type solution as we can see the values for $\tilde{\tau} = 1$ are close to those at $\tilde{\tau} = 2$, with the two lines are almost identical for $\tilde{S} > 10$. We notice in both figure 4.2(a) and figure 4.2(b) that as we perturb away from the terminal time solution there is interesting behaviour and as such a small-$\tilde{\tau}$ solution is also of some interest. Near the terminal time, the optimal strategy is increasing as $\tilde{\tau}$ increases in the small $\tilde{S}$ regime while decreasing as $\tilde{\tau}$ increases for larger values of $\tilde{S}$. This is due

to the drift prospect outweighing price-movement risk for small $\tilde{S}$ while for larger $\tilde{S}$ the volatility risk outweighs the drift prospect.

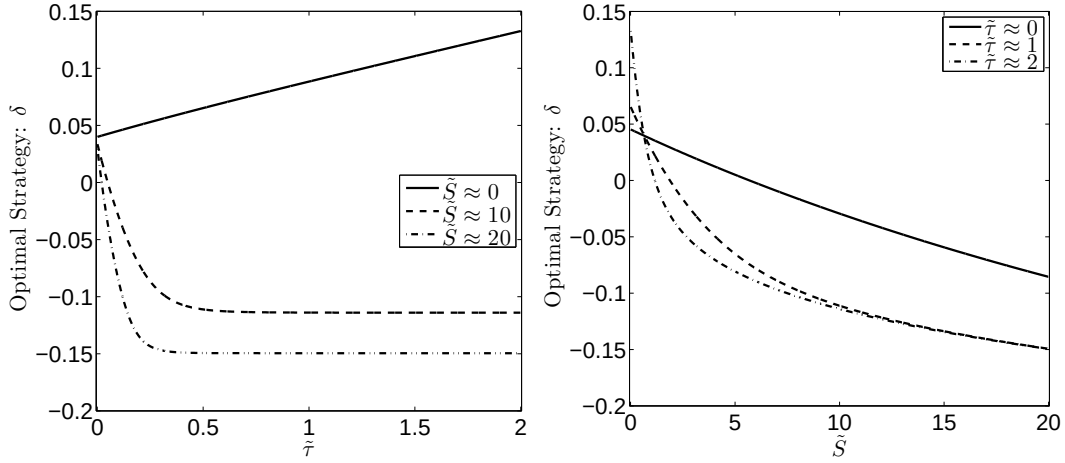
Given the optimal strategy, δ^*, is a transformation of $f\left(\tilde{\tau}, q, \tilde{S}\right)$ the function f exhibits similar behaviour as the optimal strategy, δ^*, does in figure 4.2. Therefore in the following sections we shall investigate both the small-$\tilde{\tau}$ solution and perpetual solution for f.

We have also included a plot of the asking price, $\tilde{S}^{a*}(\tilde{\tau})$, given by (4.45). This can be seen in figure 4.3, and the results are corresponding to the optimal trading strategies in figure 4.2. In figure 4.3(a) we see the solution at $\tilde{S} = 0$ is graphically indistinguishable from the horizontal axis. This is because $\tilde{S}^{a*} \to 0$ as $\tilde{S} \to 0$, as can be seen in figure 4.3(b). We also notice that although the optimal trading strategy at $\tilde{\tau} = 1$ and $\tilde{\tau} = 2$ are distinguishable from each other (see figure 4.2(b)), when examining the asking price we notice they are relatively indistinguishable from each other (see figure 4.3(b)), given in monetary terms the difference is actually relatively small.

The remainder of this section consists of examination of the trading strategies for varying amounts of inventory, q, and varying parameter values. We examine these at a time close to the terminal time, $\tilde{\tau} = 0.01$, and at (an earlier time), $\tilde{\tau} = 2$. The reason for doing so is that the solution varies rapidly as $\tilde{\tau}$ initially increases before tending to a perpetual solution. Therefore, showing the solutions at both times give a good insight into the overall behaviour of the solution.

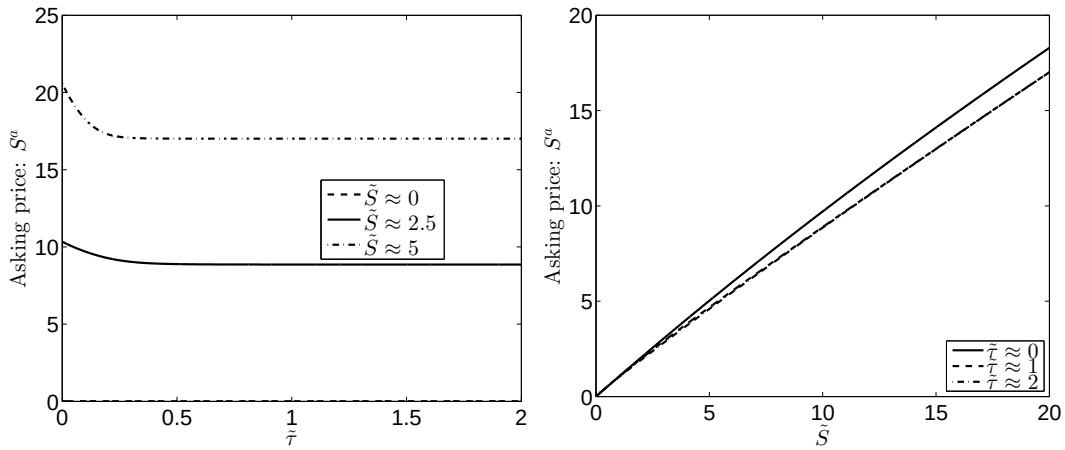
The properties of the optimal trading strategy for $q > 1$ are similar to those of $q = 1$. Figure 4.4 gives an indication of how the optimal trading strategy behaves for various values of q for a single time step around the terminal time ($\tilde{\tau} = 0.01$, figure 4.4(a)) and at an earlier time ($\tilde{\tau} = 2$, figure 4.4(b)). The optimal trading strategy is not dependent on q at $\tilde{\tau} = 0$, as shown in (4.46). We have included a plot of (4.46) in figure 4.4(a) to highlight the significant difference in solutions from a small deviation in time, and as q is increased.

Figure 4.5 shows the value function corresponding to the optimal trading strategies in figure 4.4. It can be seen that for the value function there is almost an indistinguishable difference (graphically) between the three curves for $\tilde{\tau} = 0.01$ and $\tilde{\tau} = 2$. However, the difference between the results of the optimal strategy at the two times



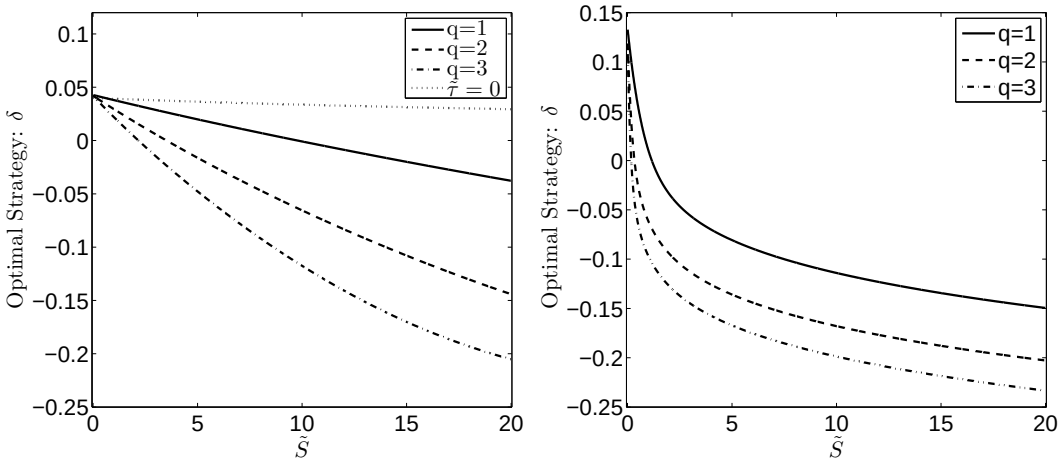
(a) Optimal strategy for various values of $\tilde{S}$ (b) Optimal strategy for various values of $\tilde{\tau}$

Figure 4.2: Optimal strategy for trader with one asset remaining, with $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.



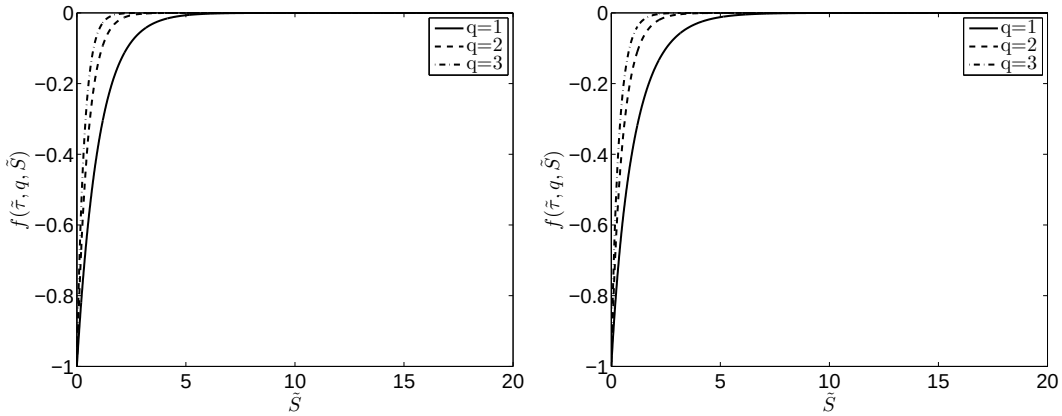
(a) Asking price for various values of $\tilde{S}$ (b) Asking price for various values of $\tilde{\tau}$

Figure 4.3: Asking price for trader with one asset remaining, with $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.



(a) Optimal strategy at $\tilde{\tau} = 0.01$ (b) Optimal strategy at $\tilde{\tau} = 2$

Figure 4.4: Optimal strategy at various times for numerous assets remaining. The parameter values used are $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.



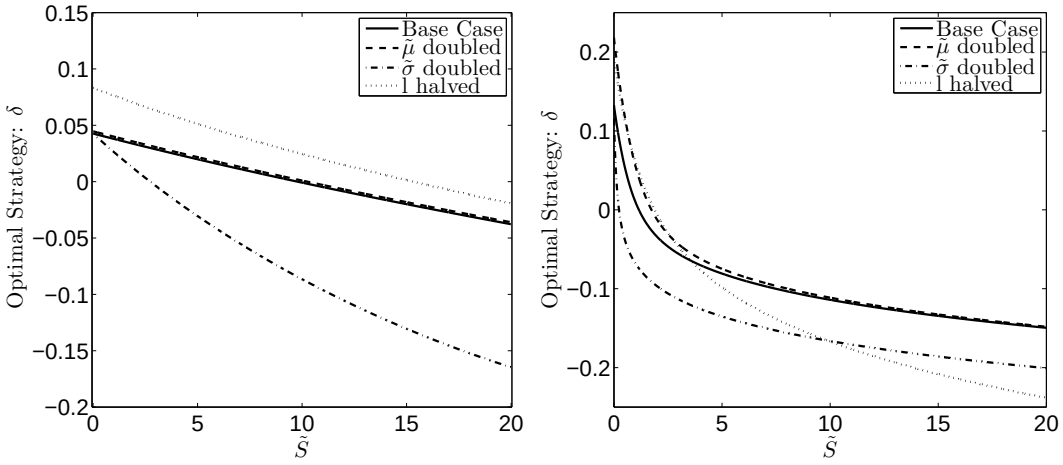
(a) Value function at $\tilde{\tau} = 0.01$ (b) Value function at $\tilde{\tau} = 2$

Figure 4.5: Value function at various times for numerous assets remaining. The parameter values used are $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.

is quite transparent. It is for this reason that we focus our attention on the optimal trading strategy when examining results, rather than the value function.

To conclude the discussion on the numerical examples we shall briefly describe how the optimal trading strategy changes in relation to changes to the various parameters in the model. In figure 4.6 the optimal strategy for various times has been plotted with each parameter, $\tilde{\mu}, \tilde{\sigma}$ and l being varied, as well as a 'base' case which is calculated using the same parameters as stated above. As can be seen, the properties which we now discuss hold at both ends of the time spectrum for a given parameter variation.

As the drift, $\tilde{\mu}$, increases the trader will be more happy to hold on to the asset as he



(a) Optimal strategies at $\tilde{\tau} = 0.01$ (b) Optimal strategies at $\tilde{\tau} = 2$

Figure 4.6: Optimal strategy under varying parameters. The base case has parameter values: $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.

expects the price to rise and thus his wealth to rise. He therefore asks for a higher price for the asset over assets with lower drift. As the volatility, $\tilde{\sigma}$, increases the trader's asking price decreases. Understandably, a trader who is risk-averse dislikes higher volatility as it brings a greater level of risk in stock price movement and thus will wish to sell quicker than if volatility was lower. As we decrease l, the exponential-decay parameter for the limit-order book, we see the trader asks for a higher price when his asset price is low and a lower price when his asset price is high. This switch is centred around whether the trader is selling the asset above par or at a discount, i.e. δ^* greater than zero or less than zero, which occurs as $\tilde{S}$ is increased. When selling above par a lower l signifies that the trader can place his asset further into the limit-order book without significantly reducing his probability of sale. When the trader is eager to sell he will sell the asset for a discount. In order to do so under lower l the trader must sell at a high discount in order to increase his probability of sale by a significant amount. The opposite is true for larger l; when selling above par the probability of selling significantly decreases while the probability of selling significantly increases when selling at a discount; this can be seen in figure 4.6(b). However if we were to plot our graph at $\tilde{\tau} = 0.01$ for higher values of $\tilde{S}$ we would see the curve representing lower l cross the curve representing the 'base' case eventually for large enough $\tilde{S}$.

It is interesting to observe the preferences of the trader for constant wealth but with varying asset quantities and respective asset price, i.e. does the trader prefer one

asset worth £100 or 10 assets worth £10. We shall consider this at a time $\tilde{\tau} \gg 0$ as close to the terminal time it is optimal to have as least assets as possible in order to liquidate before the terminal time[1]. It was found from numerical investigate that in certain parameter regimes the trader prefers more assets with less value per asset than less assets with greater value while in other parameter regimes the trader prefers less assets with higher value per asset than more assets with less value (assuming constant wealth). This has to do with his risk-aversion. In the case of larger drift, $\tilde{\mu}$, and lower volatility $\tilde{\sigma}$, the trader prefers a greater number of assets over a lower number as the optimal strategy is to ask for an above par value for the assets as their growth prospects outweigh the risk. However, for lower values of $\tilde{\mu}$ and higher $\tilde{\sigma}$, in which risk outweighs the growth prospects, the trader will ask for a below par value for the assets and prefers to have a lower number of assets, as to have more control over the inventory. This relates to the parameter constraint found when examining the perpetual solution, as discussed in section 4.3.

4.2 Small-time-to-termination solution

In the light of the results from section 4.1.5 we now visit the concept of a small-$\tilde{\tau}$ solution. From the numerical results we see that there is a lot of 'activity' (which mirrors financial interest) taking place around $\tilde{\tau} = 0$. Indeed, it is well-known that options priced using the Black-Scholes-Merton framework close to expiry (see Evans et al., 2002) can also exhibit rapid changes in value.

In this regime the trader is more worried about not being able to sell the asset for any price higher than its current price or, when the asset is already at a high price, the trader is worried that volatility could cause a decrease in his asset price before the terminal time. This is a property of the trader's risk-aversion.

To examine the small-time-to-termination solution we shall expand $f\left(\tilde{\tau}, q, \tilde{S}\right)$ as follows

$$f\left(\tilde{\tau}, q, \tilde{S}\right) = f_0\left(q, \tilde{S}\right) + \tilde{\tau} f_1\left(q, \tilde{S}\right) + \tilde{\tau}^2 f_2\left(q, \tilde{S}\right) + O\left(\tilde{\tau}^3\right), \tag{4.55}$$

which is a power-series expansion in $\tilde{\tau}$ up to $O\left(\tilde{\tau}^3\right)$. The $O\left(\tilde{\tau}^0\right)$ term, $f_0\left(q, \tilde{S}\right)$, is

[1]This was found to be the case for all the parameter regimes tested and intuitively is what we would expect.

merely equal to the terminal condition given by (4.42). By substituting (4.55) into (4.41) we can find an analytic solution for $f_1\left(q,\tilde{S}\right)$ by collecting the $O\left(\tilde{\tau}^0\right)$ terms

$$f_1\left(q,\tilde{S}\right)=\left(-\tilde{\mu}\tilde{S}q+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2q^2-\frac{\tilde{S}}{\tilde{S}+l}\left(\frac{l}{\tilde{S}+l}\right)^{\frac{l}{\tilde{S}}}\right)f_0\left(q,\tilde{S}\right), \tag{4.56}$$

and from the $O\left(\tilde{\tau}^1\right)$ terms

$$\begin{aligned} f_2\left(q,\tilde{S}\right)=\frac{1}{2}\Bigg(&\tilde{\mu}\tilde{S}f_{1\tilde{S}}\left(q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f_{1\tilde{S}\tilde{S}}\left(q,\tilde{S}\right)-\frac{1}{\tilde{S}+l}\left(\frac{l}{\tilde{S}+l}\right)^{\frac{l}{\tilde{S}}} \\ &\times\left(\left(l+\tilde{S}\right)f_1\left(q,\tilde{S}\right)-lf_1\left(q-1,\tilde{S}\right)e^{-\tilde{S}}\right)\Bigg). \end{aligned} \tag{4.57}$$

$f_{1\tilde{S}}$ and $f_{1\tilde{S}\tilde{S}}$ are the first and second derivative of f_1 respectively which can both be easily calculated analytically, as given by

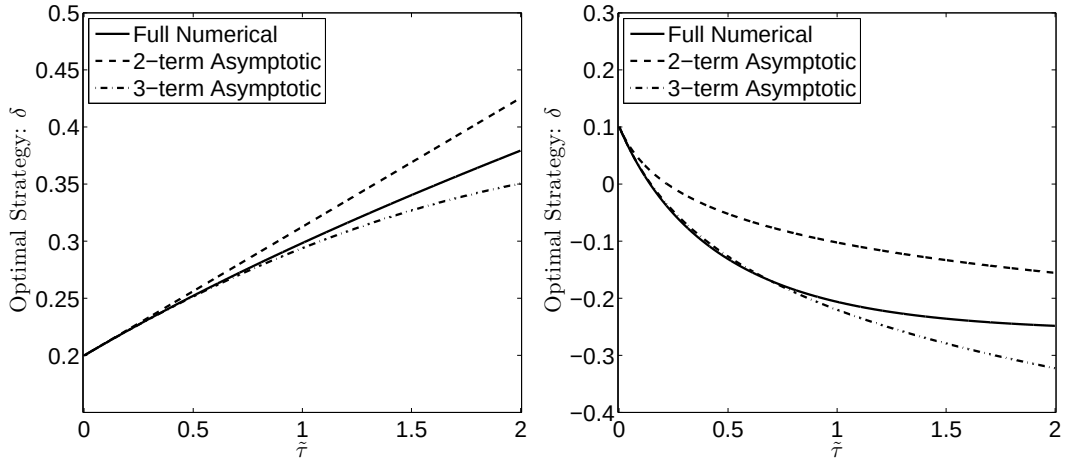
$$\begin{aligned} f_{1\tilde{S}}\left(q,\tilde{S}\right)=&-qf_1\left(q,\tilde{S}\right) \\ &+\left(-\tilde{\mu}q+\tilde{\sigma}^2q^2\tilde{S}+\left(\frac{l}{\tilde{S}+l}\right)^{\left(\frac{l}{\tilde{S}}+1\right)}\left(\frac{\log\left(\frac{l}{l+\tilde{S}}\right)}{\tilde{S}}\right)\right)f_0\left(q\tilde{S}\right), \end{aligned} \tag{4.58}$$

and

$$\begin{aligned} f_{1\tilde{S}\tilde{S}}\left(q,\tilde{S}\right)=&-qf_{1\tilde{S}}\left(q,\tilde{S}\right) \\ &-q^2\left(\tilde{\sigma}^2\left(1-q\tilde{S}\right)-\tilde{\mu}\right)f_0\left(q,\tilde{S}\right) \\ &-\frac{1}{\tilde{S}^2}\left(\frac{l}{l+\tilde{S}}\right)^{\frac{l}{\tilde{S}}+1}\left(\frac{\tilde{S}}{l+\tilde{S}}+\left(q\tilde{S}-2\right)\log\left(\frac{l}{l+\tilde{S}}\right)\right. \\ &\left.+\frac{l}{\tilde{S}}\log^2\left(\frac{l}{l+\tilde{S}}\right)\right)f_0\left(q,\tilde{S}\right). \end{aligned} \tag{4.59}$$

Therefore (4.56) and (4.57) have fully analytic solutions. In the limit as $\tilde{S}\to 0$, both f_1 and f_2 go smoothly to a constant and hence are finite. This characteristic will guide us when examining the steady-state solution.

We shall now discuss the accuracy of this asymptotic expansion, along with its limitations. The values we use for our parameters are the same as that of section 4.1.5 but with $l=5$, for reasons that will be explained later. In figure 4.7 we have plotted the solution of the optimal trading strategy along with its asymptotic approximations for two values of $\tilde{S}$ which was calculated by substituting (4.55) into (4.44). We can see that in both figures the three-term approximation is a good approximation for



(a) $\tilde{S} = 0$ (b) $\tilde{S} = 10$

Figure 4.7: Comparison of full-numerical optimal trading strategy and asymptotic expansion for $\tilde{S}$ near zero and $\tilde{S}$ near 10 using a two-term and three-term model, with parameters $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 5$.

the solution for $\tilde{\tau}$ of $O(1)$ while the two-term approximation is valid for smaller $\tilde{\tau}$. We have tested this for various q values and these observations still hold true. We mentioned in section 4.1 that the rescaling of λ and τ came about when examining the small-time solution. In the original dimensions, for $\lambda = 1$ we arrive with similar plots as that in figure 4.7. If we increased λ by a factor of two, the time for which the small-time solution was valid would decrease by a factor of two and so forth. Thus the rescaling of λ against τ was deemed appropriate.

This approximation is surprisingly accurate for a range of parameter values. However, there are restrictions to how well the model holds depending on the parameter values used. For higher values of $\tilde{S}$ and l the solution of $f\left(\tilde{\tau}, q, \tilde{S}\right)$ diverges faster from the small-time-to-termination solution due to the stronger presence of the non-linear term, hence the use of the smaller l in our numerical example, compared to the l used in the non-steady state. In figure 4.8 we have plotted the optimal strategy against $\tilde{S}$ for increasing $\tilde{\tau}$. We can see that as $\tilde{\tau}$ increases the range of $\tilde{S}$ for which the approximation is accurate decreases. It can also be seen that for $\tilde{\tau} = 0.1$ the difference between the full numerical and three-term asymptotic approximation is negligible. If l were to be reduced even further the asymptotic approximation remains valid for a larger range of $\tilde{\tau}$ values.

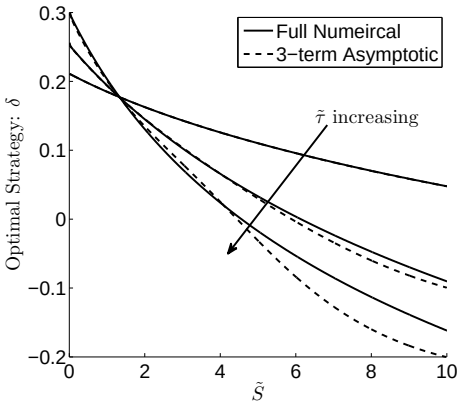


Figure 4.8: Comparison of full-numerical optimal trading strategy and three-term asymptotic expansion for increasing $\tilde{\tau}$, with $\tilde{\tau} = \{0.1, 0.5, 1\}$ and parameter values $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 5$.

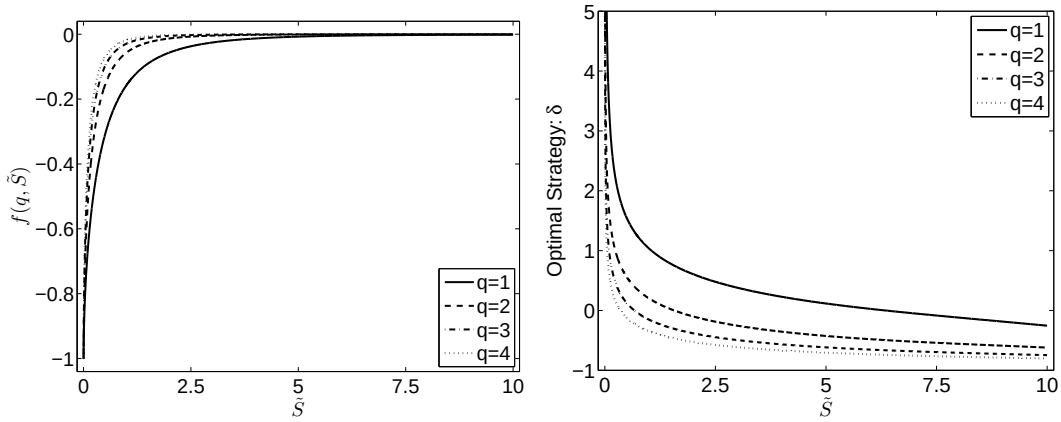
It is worth noting the original investigation of this problem was done in the original dimensional space, before transforming to the non-dimensional space, which was performed in section 4.1.3. It was through examining the small-time-to-termination solution that we found the relationship between time remaining, τ, and the rate at which assets are sold at their reference price, λ, and as such scaled $\tilde{\tau} = \lambda\tau$. Examining the accuracy for various parameters and values of S also led us to the scaling $\tilde{S} = \gamma S$. Thus this is evidence that asymptotic analysis can add insight into a problem.

We shall now move on to examine what happens when $\tilde{\tau} \to \infty$, i.e. the "perpetual" state.

4.3 Perpetual solution

We saw from figure 4.2 that it is possible that perpetual (time-independent) solutions exist. Indeed, Guéant et al. (2012b) also find such solutions (in certain parameter regimes). In this section we shall study the perpetual solution in some detail. For the perpetual solution we require $f_{\tilde{\tau}}\left(\tilde{\tau}, q, \tilde{S}\right) \to 0$ as $\tilde{\tau} \to \infty$. In making this assumption, we are implicitly assuming that the terminal time in which the assets must be sold is sufficiently distant for us to neglect the time variation. Then (4.41) becomes

$$\tilde{\mu}\tilde{S}f_{\tilde{S}}\left(q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f_{\tilde{S}\tilde{S}}\left(q,\tilde{S}\right)-\frac{e^l\tilde{S}f\left(q,\tilde{S}\right)}{\tilde{S}+l}\left(\frac{lf\left(q,\tilde{S}\right)}{\left(\tilde{S}+l\right)f\left(q-1,\tilde{S}\right)}\right)^{\frac{l}{\tilde{S}}}=0, \quad (4.60)$$



(a) Value function (b) Optimal strategy

Figure 4.9: Value function and optimal strategy for perpetual solution for various q, with parameter values: $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.

with

$$f\left(0, \tilde{S}\right) = -1. \tag{4.61}$$

To solve this non-linear ODE (4.60) we must consider two boundary conditions, one for $\tilde{S} = 0$ and one for $\tilde{S} \to \infty$. We know from the non-steady-state problem that $f\left(\forall\tilde{\tau}, q, \tilde{S} = 0\right) = -1$ and thus $f\left(q, \tilde{S} = 0\right) = -1$ for the perpetual case. For $\tilde{S} \to \infty$ we use the same boundary condition as used in section 4.1.3 for the full problem, that being a Neumann boundary condition as given by (4.52). We solve (4.60) using a Newton iterative method, see section 2.6.3. The optimal trading strategy takes the form

$$\delta^*\left(q, \tilde{S}\right) = \frac{1}{\tilde{S}} \ln\left(\frac{\left(\tilde{S} + l\right) f\left(q-1, \tilde{S}\right)}{l f\left(q, \tilde{S}\right)}\right) - 1. \tag{4.62}$$

We plot the value function (figure 4.9(a)) and optimal trading strategy (figure 4.9(b)) for $q \in \{1, 2, 3, 4\}$ with the same parameter values as used in section 4.1.5. This solution was tested against the full numerical solution for the non-perpetual (time-varying) case. We notice that the optimal strategy for the perpetual case is larger for small $\tilde{S}$ than for the full problem. This is for two reasons. The first being that $\delta^*_{\tilde{\tau}} > 0$ for small $\tilde{S}$ (as can be seen in figure 4.7(a)) which is due to the prospect of the drift increasing the value of the asset dominating over the prospect of the volatility decreasing the value. The second is due to the presence of a singularity about $\tilde{S} = 0$, which we will now discuss.

4.3.1 Analytic analysis

We found that there are convergence issues for certain parameter values. Examining $f_{\tilde{S}}$ for small $\tilde{S}$ it appeared that a singularity was present for $\tilde{S} \ll 1$. On closer inspection using asymptotic balancing suggests that in this limit, the solution behaves as

$$f\left(q, \tilde{S}\right) \approx -1 + c(q)\tilde{S}^{\beta} \tag{4.63}$$

for some positive constant β and $c(q) \geq 0$, and so the solution is bounded, provided

$$\beta = 1 - \frac{2\tilde{\mu}}{\tilde{\sigma}^2} > 0, \tag{4.64}$$

which gives us the requirement that

$$\tilde{\mu} < \frac{\tilde{\sigma}^2}{2} \tag{4.65}$$

is necessary for a solution to exist, given $\beta > 0$. We have verified this constraint numerically (later in this section we address the issue of parameters outside this region). This constraint is analogous to that found by Guéant et al. (2012b) in which they derived (in their notation)

$$\mu < \frac{\gamma\sigma^2}{2} \tag{4.66}$$

to be a necessary condition for a solution to exist for the perpetual case under standard Brownian motion. We notice there is a lack of the γ parameter in our constraint in comparison to Guéant et al. (2012b). Although we have scaled out γ in our model, even without this rescaling we still arrive at a constraint that is independent of γ.

As can be seen in figure 4.9(b) we notice a singularity is apparently present in the optimal trading strategy. We can make an asymptotic approximation for the optimal trading strategy for $\tilde{S} \ll 1$ given as:

$$\delta^*\left(q, \tilde{S}\right) \approx K(q)\tilde{S}^{\beta-1} \tag{4.67}$$

which is found by taking the limit of (4.62) as $\tilde{S} \to 0^+$, using the form (4.63) for the value function, and noting that $c(q) > c(q-1)$. Given the latter point we expect $K(q)$ to be a decreasing function of q, tending to zero as $q \to \infty$, which can be seen in figure 4.10.

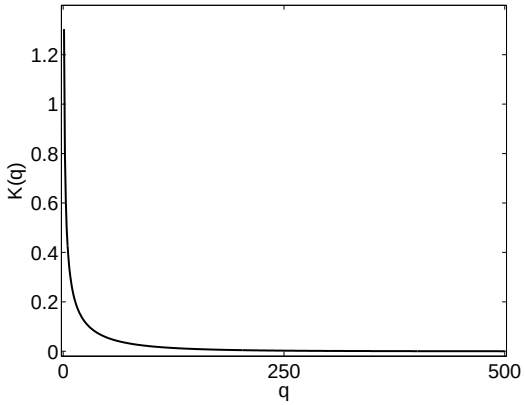


Figure 4.10: $K(q)$ variation with q for $S \to 0$ with parameter values: $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 1$. We set $Y_{min} = -20$.

Rescaling to log-space

Examining the small-$\tilde{S}$ region can be computationally expensive as we need to have a very fine grid around small $\tilde{S}$ and given we are using uniform grids this means that the whole grid will be quite fine. A way of working around this is to transform the variable to the log of $\tilde{S}$, such that we set $Y = \ln\left(\tilde{S}\right)$. Y has a double infinite domain and given $\tilde{S} = e^Y$ we can choose Y_{min} such that $|Y_{min}|$ is large and thus a good approximation for small $\tilde{S}$. The grid need not be as fine as that of the $\tilde{S}$ grid and thus relieves us of some numerical expense. The log transformation is often used in finance mathematics under the Black-Scholes framework when transforming the Black-Scholes PDE to the heat equation (see Wilmott, 1995).

Taking $Y = \ln\left(\tilde{S}\right)$ we transform (4.60) to

$$\left(\tilde{\mu} - \frac{1}{2}\sigma^2\right) f_Y(q, Y) + \frac{1}{2}\tilde{\sigma}^2 f_{YY}(q, Y) - \frac{e^{l+Y}(q, Y)}{e^Y + l}\left(\frac{l f(q, Y)}{(e^Y + l) f(q-1, Y)}\right)^{le^{-Y}} = 0, \tag{4.68}$$

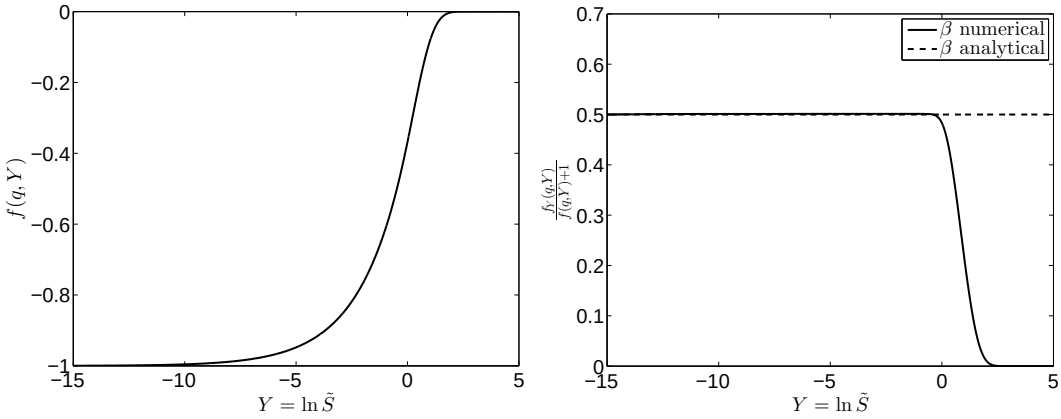
with

$$f(0, Y) = -1. \tag{4.69}$$

For the large Y boundary condition, we transform the Neumann in $\tilde{S}$-space given by (4.52) and arrive with a similar Neumann in log-space, given as

$$\frac{\partial f}{\partial \tilde{Y}}\left(q, \tilde{Y} \to \infty\right) \to 0, \tag{4.70}$$

For the negative infinity boundary condition we transform (4.63) to log-space and



(a) Value function (b) Numerical and analytic solution β

Figure 4.11: Examining the solution of the value function under log-transform and numerically confirming the boundary condition behaviour assumed, with parameter values: $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.

arrive at

$$f(q, Y) \approx -1 + c(q)e^{\beta Y} \tag{4.71}$$

to which we have the Robin condition

$$\frac{\partial f(q, Y)}{\partial Y} - \beta f(q, Y) = \beta \quad \text{as} \quad Y \to -\infty. \tag{4.72}$$

To numerically validate our assumption of the behaviour of the small-$\tilde{S}$ regime we have examined the numerical value of β, against its actual as depicted by (4.64). We can numerical examine β as

$$\beta = \frac{f_Y(q, Y)}{f(q, Y) + 1} \tag{4.73}$$

which is simply a rearrangement of (4.72).

Figure 4.11 shows a plot of the value function (figure 4.11(a)) as well as a plot of β (figure 4.11(b)), which is calculated from (4.73). In figure 4.11(b) we see the numerical value of β, found from (4.73), is identical to that of the analytic value (4.64) for $Y \sim O(1)$, confirming our solution assumption (4.63) for the small-$\tilde{S}$ limit. $K(q)$ from equation (4.67), was also verified using the log-transform method, as seen in figure 4.10.

Development of singularity

In the following we return to the $\tilde{S}$ space and have set $l = 1$; larger values of l lead to a general dominance of the non-linear term, and rapid valuation variations close to

$\tilde{S} = 0$, both of which conceal valuation behaviours, and thereby make insights thereof more difficult to discern.

It is of some interest to inspect how (quickly) the time-varying solution approaches the perpetual state (assuming (4.64) is satisfied). Figure 4.12(a) shows sample results for the difference between the former and latter states, indicating a quite rapid asymptotic approach in this case (which, indeed, generally was a canonical observation found in most calculations). However, although this difference is diminishing, quite rapidly, it is clear that there is a maximum deviation at decreasing values of $\tilde{S}$ as $\tilde{\tau}$ increases. It is clear that the $O(\tilde{S}^{\beta})$ in the perpetual state as $\tilde{S} \to 0$ leads to singular behaviour, whilst the numerical results of the time-varying solution very strongly point to the non-linear term in (4.41) being negligible in this regime. Therefore, we examined the linearised PDE

$$-f_{\tilde{\tau}}\left(\tilde{\tau}, q, \tilde{S}\right) + \tilde{\mu}\tilde{S}f_{\tilde{S}}\left(\tilde{\tau}, q, \tilde{S}\right) + \frac{1}{2}\tilde{\sigma}^2\tilde{S}^2 f_{\tilde{S}\tilde{S}}\left(\tilde{\tau}, q, \tilde{S}\right) = 0, \tag{4.74}$$

which is valid in the small-$\tilde{S}$ regime, and has initial conditions given by (4.42) and (4.43). Assuming a power series expansion of $f\left(\tilde{\tau}, q, \tilde{S}\right)$ about $\tilde{S}$ as $\tilde{S} \to 0$ we obtain a series solution of the linearised form of this PDE which takes the form

$$f\left(\tilde{\tau}, q, \tilde{S}\right) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{n!}\left(q\tilde{S}\right)^n e^{\left(n\tilde{\mu} + \frac{1}{2}n(n-1)\tilde{\sigma}^2\right)\tilde{\tau}}. \tag{4.75}$$

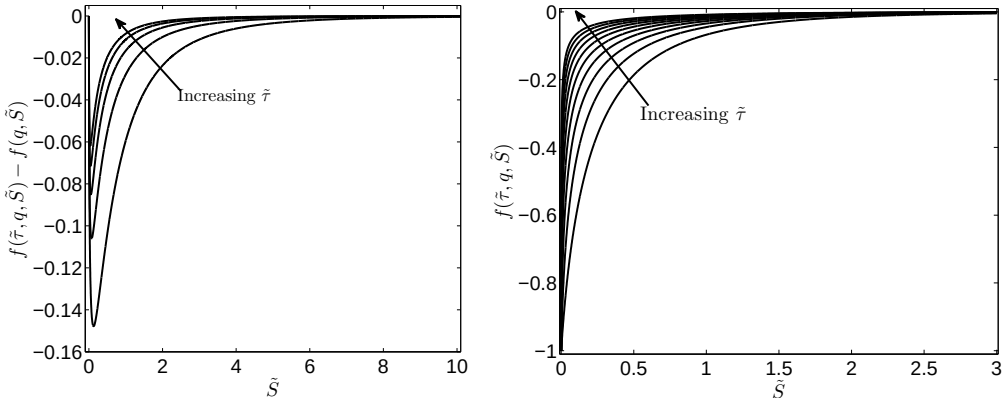
Although this is clearly a divergent series (for sufficiently large n), nonetheless it does highlight the likely occurrence of (very) short $\tilde{S}$ scales as $\tilde{\tau}$ increases (and suitable truncations of this series does lead to results consistent with the full numerical system). The $n = 1$ term leads to the result that

$$f_{\tilde{S}}(\tilde{\tau}, q, 0) = qe^{\tilde{\mu}\tilde{\tau}}, \tag{4.76}$$

which clearly suggests the development of an (exponentially thin) region close to $\tilde{S} = 0$, together with an increasing magnitude for this derivative (as opposed to a diminishing value for $|f|$), which we surmise leads ultimately with a connection to (4.63). Outside of this region, the valuation is effectively that predicted by the perpetual solution.

It is possible to find an exact solution to this linear system (4.74), satisfying the appropriate conditions at $\tilde{\tau} = 0$. Returning to the log-transformation, we obtain a linearised PDE

$$-f\left(\tilde{\tau}, q, Y\right) + \left(\tilde{\mu} - \frac{1}{2}\sigma^2\right) f_Y\left(q, Y\right) + \frac{1}{2}\tilde{\sigma}^2 f_{YY}\left(q, Y\right) = 0, \tag{4.77}$$



(a) $f\left(\tilde{\tau}, q, \tilde{S}\right) - f\left(q, \tilde{S}\right)$ (b) $f\left(\tilde{\tau}, q, \tilde{S}\right)$, for $\tilde{\tau} = \{20, 40, \ldots, 200\}$

Figure 4.12: Figure 4.12(a) is the solution of (4.41), $f\left(\tilde{\tau}, q, \tilde{S}\right)$, minus the solution of the (4.60), $f\left(q, \tilde{S}\right)$, for $\tilde{\tau} = \{5, 10, 15, 20, 25\}$ with $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 1$. Figure 4.12(b) is the solution of (4.41), $f\left(\tilde{\tau}, q, \tilde{S}\right)$, for $\tilde{\tau} = \{20, 40, \ldots, 200\}$ with $q = 1$, $\tilde{\mu} = 0.06$, $\tilde{\sigma} = 0.25$ ($\beta = -.92$) and $l = 1$.

with

$$f\left(0, q, Y\right) = -e^{-qe^{Y}}, \tag{4.78}$$

$$f\left(\tilde{\tau}, 0, Y\right) = -1. \tag{4.79}$$

Equation (4.77) is essential the heat equation with an advection term and can be solved using a Fourier transformation. This yields the solution

$$f\left(\tilde{\tau}, q, Y\right) = \frac{-1}{\sigma\sqrt{2\pi\tilde{\tau}}} \int_{-\infty}^{\infty} \exp\left(-qe^{(Y-p)} - \frac{\left(\tilde{\mu} - \frac{1}{2}\tilde{\sigma}^2\right)\tilde{\tau}}{2\tilde{\sigma}^2} + p(\frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2}) - \frac{p^2}{2\tilde{\sigma}^2\tilde{\tau}} \right) dp. \tag{4.80}$$

We have not been able to find an analytic expression for the integral in (4.80) but we can solve it numerically using a quadrature method (see Press et al., 2009). Up to some large $\tilde{\tau}$ the solutions of (4.41), (4.74) and (4.77) agree in the $\tilde{S} \to 0$ ($Y \to -\infty$) limit.

A natural question that arises is what is the situation if the values of $\tilde{\mu}$ and $\tilde{\sigma}$ do not comply with (4.65)? Figure 4.12(b) shows such a case (corresponding to $\beta = -.92$), and it is very clear (particularly when compared with previous results), that a time-independent state is not being approached, rather that the non-negligible valuation is confined to a diminishingly small regime, close to $\tilde{S} = 0$.

4.3.2 Perpetual solution for large q

To conclude our examination of the perpetual problem we shall look at the case of large q. Examination of figure 4.9 suggests that as q is increased the solutions tend to collapse on to a single distribution (a conclusion supported by a number of numerical experiments conducted). We have seen in figure 4.10 that $K(q) \to 0$ as $q \to \infty$. Further to that, examining the non-linear term we find

$$\lim_{q\to\infty}\left(\frac{f\left(q,\tilde{S}\right)}{f\left(q-1,\tilde{S}\right)}\right)\to 1, \tag{4.81}$$

which was backed up numerically. Using (4.81) we can reduce (4.60) to a linear ODE which is now independent of q,

$$\tilde{\mu}\tilde{S}f_{\tilde{S}}\left(\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f_{\tilde{S}\tilde{S}}\left(\tilde{S}\right)-\frac{e^l\tilde{S}f\left(\tilde{S}\right)}{\tilde{S}+l}\left(\frac{l}{\tilde{S}+l}\right)^{\frac{l}{\tilde{S}}}=0, \tag{4.82}$$

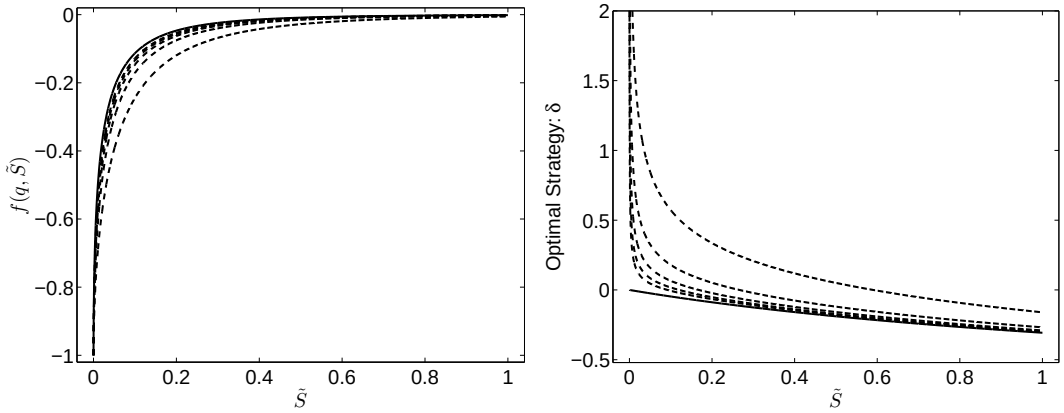
in which we shall refer $f\left(\tilde{S}\right)$ as the $q\to\infty$ solution, with boundary conditions $f\left(\tilde{S}=0\right)=-1$ and $f_{\tilde{S}}\left(\tilde{S}\to\infty\right)\to 0$, as used on the non-linear ODE previously. We have been unable to find an analytic solution for (4.82) but it is straightforward to solve numerically using finite-difference methods (for example). Remembering the explicit form of the optimal control given by (4.62), we can see that under the assumption of (4.81) the optimal trading strategy is given by

$$\delta^*\left(\tilde{S}\right)=\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)}{l}\right)-1. \tag{4.83}$$

The solution is shown in figure 4.13. In figure 4.13(a) we can see that as q is increased the value function solutions tend to the $q\to\infty$ solution. The optimal trading strategy is given by (4.83) and can be seen in figure 4.13(b). We notice how, as with the value function, the solutions of the optimal trading strategies converge to the $q\to\infty$ optimal trading strategy as q is increased. It can be seen that the solution for δ^* at $S\to 0^+$ is near zero for the $q\to\infty$ solution while singular for the solutions for finite q. This is expected given

$$\delta^*\left(\tilde{S}\to 0\right)=\lim_{\tilde{S}\to 0^+}\left(\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)}{l}\right)-1\right)=\frac{1}{l}-1. \tag{4.84}$$

If we were to increase q significantly, the coefficient of the singularity (4.67) in this limit decreases towards zero, in line with our observations above regarding $K(q\to\infty)$.



(a) Value function (b) Optimal trading strategy

Figure 4.13: Perpetual solution for large q. In figure 4.13(a) the solid line represents the solution of (4.82) while the broken lines represent the solutions of (4.60) for $q = \{10, 20, 30, 40, 50\}$, beginning with $q = 10$ which is furthest from the $q \to \infty$ solution. We can see that as q is increased the solutions converge to the $q \to \infty$ solution. The same can be seen for the optimal strategy in figure 4.13(b). The parameter values used: $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 1$.

4.4 Optimal liquidation of mean-reverting assets

In this section we consider using a mean-reverting diffusion process for the asset price, so as to emphasise the adaptability and advantages of using numerical methods to solve this class of problem. The process we shall use is the CIR process, which was first suggested in 1985 by Cox et al. (1985) as an extension to the Vasicek (1977) model. The attractiveness of the CIR model is its avoidance of the possibility of negative values, assuming the Feller (1951) condition is satisfied (which we discuss below). It was originally suggested for the modelling of interest rates but has since been suggested for other models such as stochastic volatility models and, of particular interest to us for this framework, commodities (see Linetsky, 2004) and futures (see Schwartz, 1997).

4.4.1 Formulating model for asset driven by CIR process

An asset $S(t)$ following the CIR process solves the following stochastic differential equation

$$dS(t) = \kappa\left(\zeta - S(t)\right)dt + \sigma\sqrt{S(t)}dW(t) \tag{4.85}$$

with κ as the mean-reversion speed, ζ is the long-term mean asset price and $\sigma\sqrt{S(t)}$ is the absolute volatility. Following the same problem formulation as we have in section 4.1 but with the CIR process for the asset price, we can derive a similar HJB equation

$$\begin{aligned} &u_t(t,X,q,S)+\kappa\left(\zeta-S\right)u_S(t,X,q,S)+\frac{1}{2}\sigma^2 S u_{SS}(t,X,q,S)\\ &+\sup_{\delta}\left[\lambda e^{-l\delta}\left(u(t,X+S\left(1+\delta\right),q-1,S)-u(t,X,q,S)\right)\right]=0, \end{aligned} \tag{4.86}$$

with initial conditions (4.8) and (4.9).

Using the same ansatz form of the solution as used for geometric Brownian motion, given by (4.29), and the change of variables

$$\tilde{\tau}=\lambda\left(T-t\right),\quad \tilde{S}=S\gamma,\quad \tilde{\zeta}=\zeta\gamma,\quad \tilde{\kappa}=\frac{\kappa}{\lambda},\quad \tilde{\sigma}=\sigma\sqrt{\frac{\gamma}{\lambda}} \tag{4.87}$$

we can find the optimal strategy in terms of the value function, and thus reduce (4.86) to the dimensionless non-linear PDE:

$$\begin{aligned} &-f_{\tilde{\tau}}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\kappa}\left(\tilde{\zeta}-\tilde{S}\right)f_{\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}f_{\tilde{S}\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)\\ &\qquad-\frac{e^l\tilde{S}f\left(\tilde{\tau},q,\tilde{S}\right)}{\tilde{S}+l}\left(\frac{lf\left(\tilde{\tau},q,\tilde{S}\right)}{\left(\tilde{S}+l\right)f\left(\tilde{\tau},q-1,\tilde{S}\right)}\right)^{\frac{l}{\tilde{S}}}=0, \end{aligned} \tag{4.88}$$

with initial conditions (4.42) and (4.43). Note that the optimal trading strategy takes the same form as (4.44).

For the CIR model it is necessary to be careful when establishing boundary conditions as the concept of positivity and non-negativity come into play, the difference being the inclusion of zero to the asset-price domain. Under certain conditions the CIR model is defined on a strictly positive domain, while when these conditions are not satisfied the CIR model has a positive probability of reaching zero. The boundary behaviour of (4.85) has been studied in great detail by Feller (1951) in which he found that the CIR model is defined on a positive domain if (in our notation)

$$2\tilde{\kappa}\tilde{\zeta}\geq\tilde{\sigma}^2. \tag{4.89}$$

Equation (4.89) is satisfied for most realistic parameter values applicable to finance, and as such we shall focus only on this regime in this thesis. However, it has been found in previous studies, namely Aquan-Assee (2009), that when the Feller condition

is not satisfied the boundary conditions for zero and infinite boundary conditions are still appropriate and provide consistent and convergent solutions.

For the boundary condition at $\tilde{S} = 0$ we implement the degenerate form of the PDE (4.88) resulting in

$$- f_{\tilde{\tau}}(\tilde{\tau}, q, 0) + \tilde{\kappa}\tilde{\zeta} f_{\tilde{S}}(\tilde{\tau}, q, 0) = 0, \tag{4.90}$$

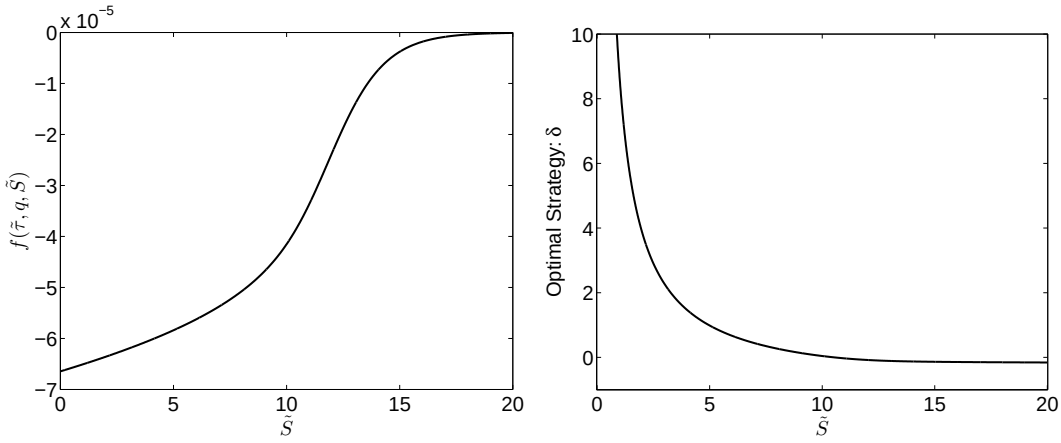
which could be seen as the boundary behaviour, as $\tilde{S}$ never reaches zero. Indeed, (4.90) implies that $f(\tilde{\tau}, q > 0, \tilde{S} = 0)$ approaches zero as $\tilde{\tau}$ increases since $f_{\tilde{S}}(\tilde{\tau} > 0, q, 0) > 0$, then as $\tilde{S} \to 0$, $\delta^* = O(\frac{1}{\tilde{S}})$, which we will see shortly is consistent with the numerics.

For the case of $\tilde{S} \to \infty$ we found from numerical investigation that the same boundary condition as used for the geometric Brownian motion case, that being a Neumann condition given by (4.52), is consistent and satisfactory.

4.4.2 Numerical results

Several of the properties found in the results for the geometric Brownian motion model are mirrored in the CIR model and thus we shall not repeat them. However a unique property for the mean-reverting process not found in the geometric Brownian motion case was that of a trader becoming risk-seeking. If we examine the graph of the value function given in figure 4.14(a) we see that the function is not strictly concave. Interestingly this occurs when the asset price increases past the long-term mean, $\tilde{\zeta}$. In this case the asset price is expected to return back to $\tilde{\zeta}$, with some volatility of course. As the asset price is due to decrease the trader will try to sell the asset before it does so. His trading strategy is thus to change the asking price drastically in comparison to the geometric Brownian motion model where the asset value was always expected to increase in the long run. This results in a negative δ^* which implies selling the asset at a discount, but at a price higher than $\tilde{\zeta}$. This change in asking price results in the risk-seeking behaviour we observe around the long-term mean; this trading strategy can be seen in figure 4.14(b).

We can observe that the optimal trading strategy grows significantly as $\tilde{S}$ approaches zero (in line with our comments above). This is due to the mean-reverting characteristic of the asset price. When $\tilde{S}$ is near zero the trader expects the future value of the asset to revert to the long-term mean, $\tilde{\zeta}$. He will thus ask for a large



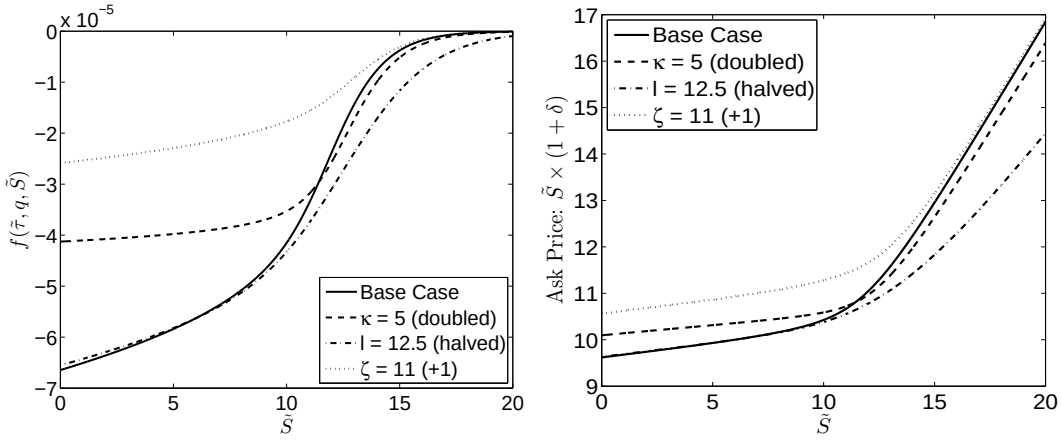
(a) Value function (b) Optimal trading strategy

Figure 4.14: Value function and optimal trading strategy under CIR process at time $\tilde{\tau} = 2$. The parameters take the values: $q = 1$, $\tilde{\zeta} = 10, \tilde{\kappa} = 2.5, \tilde{\sigma} = 0.8$ and $l = 25$. Note the value function is not concave over the whole domain and as such the trader switches from being risk-averse to risk-seeking.

multiple of the current asset price. Figure 4.16, which is the solution of (4.88) for $\tilde{S} = 0$, reinforces the explanation of why the optimal trading strategy increases in this region and also as $\tilde{\tau}$ increases.

The risk-seeking behaviour observed in figure 4.14(a) is more pronounced for higher values of the exponential-decay parameter, l, and the speed of reversion, $\tilde{\kappa}$. This is due to the trader decreasing his asking price more predominantly as the probability of execution decreases more rapidly for larger l and the asset price decreases faster to the long-term mean $\tilde{\zeta}$ for large values of $\tilde{\kappa}$. This can be seen in figure 4.15 which compares the results of varying the parameters $\tilde{\zeta}$, $\tilde{\kappa}$ and l, against a base case of $\tilde{\zeta} = 10, \tilde{\kappa} = 2.5,$ and $l = 25$, for the value function (figure 4.15(a)) and asking price (figure 4.15(b)). As noted above we see the sharp change in asking price about the long-term mean, $\tilde{\zeta}$. We also see the asking price as $\tilde{S} \to 0$ is near the long-term mean and as such the optimal strategy, δ^*, is asymptotically large, as discussed previously. The risk-seeking behaviour is also more pronounced for larger values of the time-to-expiry, $\tilde{\tau}$, as the asset has more time to revert back to the long-term mean, $\tilde{\kappa}$.

Similar to the case under geometric Brownian motion, the CIR model has interesting characteristics close to the terminal time which we shall investigate in the next section. This is of particular interest as Bhave and Libertin (2013) empirically find



(a) Value function (b) Optimal asking price

Figure 4.15: Value function and asking price under varying parameters for CIR process at time $\tilde{\tau} = 2$. The base case parameters take the values: $q = 1$, $\tilde{\zeta} = 10, \tilde{\kappa} = 2.5, \tilde{\sigma} = 0.8$ and $l = 25$.

mean-reverting patterns in equities over short-time scales. However, unlike the geometric Brownian motion model, the CIR model does not tend to a perpetual-type solution as $\tilde{\tau} \to \infty$, but rather the value diminishes towards zero, as evidenced by figure 4.16. Therefore we cannot construct a non-trivial perpetual solution in the same sense as we could for the geometric Brownian motion case. From a financial perspective, when $\tilde{S}$ is small (and less than $\tilde{\zeta}$) the trader asks for a relatively high price for the asset given it is expected to increase and revert back to the its long-term mean, $\tilde{\zeta}$. The trader's value function thus approaches zero as $\tilde{\tau}$ increases.

4.4.3 Small-time-to-termination solution

To examine the small-$\tilde{\tau}$ solution we use the same perturbation method as used in section 4.2. Substituting (4.55) into (4.88) we collect the $O\left(\tilde{\tau}^0\right)$ and $O\left(\tilde{\tau}^1\right)$ terms in which we find the values for $f_1\left(q, \tilde{S}\right)$ and $f_2\left(q, \tilde{S}\right)$ to be

$$f_1\left(q, \tilde{S}\right) = \left(-\tilde{\kappa}\left(\tilde{\zeta} - \tilde{S}\right) q + \frac{1}{2}\tilde{\sigma}^2 \tilde{S} q^2 - \frac{\tilde{S}}{\tilde{S} + l}\left(\frac{l}{\tilde{S} + l}\right)^{\frac{l}{\tilde{S}}}\right) f_0\left(q, \tilde{S}\right) \tag{4.91}$$

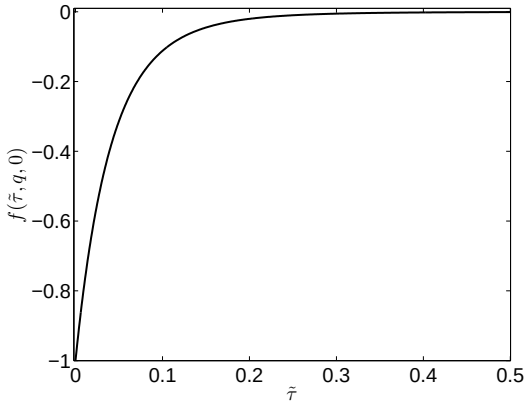
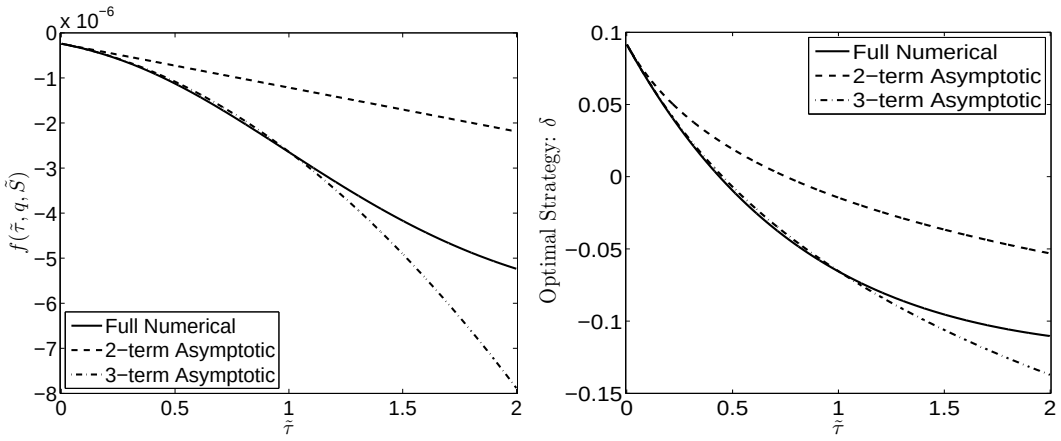


Figure 4.16: Value function at $\tilde{S} = 0$. The parameters take the values: $q = 1$, $\tilde{\zeta} = 10, \tilde{\kappa} = 2.5, \tilde{\sigma} = 0.8$ and $l = 25$. Notice the approach of $f\left(\tilde{\tau}, q, \tilde{S} = 0\right)$ to zero as $\tilde{\tau}$ increases, which reinforces the growth in δ^* as $\tilde{S} \to 0$ as seen in figure 4.14(b).

and

$$f_2\left(q, \tilde{S}\right) = \frac{1}{2}\Bigg(\tilde{\kappa}\left(\tilde{\zeta} - \tilde{S}\right) f_{1\tilde{S}}\left(q, \tilde{S}\right) + \frac{1}{2}\tilde{\sigma}^2 \tilde{S} f_{1\tilde{S}\tilde{S}}\left(q, \tilde{S}\right) - \frac{1}{\tilde{S} + l}\left(\frac{l}{\tilde{S} + l}\right)^{\frac{l}{\tilde{S}}} \times \left(\left(l + \tilde{S}\right) f_1\left(q, \tilde{S}\right) - l f_1\left(q - 1, \tilde{S}\right) e^{-\tilde{S}}\right)\Bigg) \tag{4.92}$$

As for the case of geometric Brownian motion, $f_{1\tilde{S}}$ and $f_{1\tilde{S}\tilde{S}}$ are the first and second derivative of f_1 respectively which can both be easily calculated analytically; therefore (4.91) and (4.92) have fully analytic solutions. In the limit as $\tilde{S} \to 0$, both f_1 and f_2 go smoothly to a constant and hence are finite. A comparison of the full numerical solution against a two-term and three-term asymptotic can be found in figure 4.17. We see the asymptotic approximation is accurate for $\tilde{\tau} \sim O(1)$. Note that in this case, both $|f_1|$ and $|f_2|$ are numerically large, which explains the initial rapid variation in values close to the terminal time, especially as $\tilde{S} \to 0$. It is also found (verified numerically) that as $\tilde{S}$ increases the signs of f_1 and f_2 change, as do the signs of the first derivatives with respect to $\tilde{S}$, which explains the risk-seeking behaviour observed previously (given a risk-averse trader's value function should have strictly negative first derivative, as discussed in section 2.3). This does not occur under geometric Brownian motion.



(a) Value function (b) Optimal trading strategy

Figure 4.17: Comparison of full-numerical value function and optimal trading strategy against a two-term and three-term asymptotic expansion for $\tilde{S} = \tilde{\zeta}$ for the CIR model. The parameters take the values: $\tilde{\zeta} = 10, \tilde{\kappa} = 2.5, \tilde{\sigma} = 0.8$ and $l = 5$.

4.5 Summary

In this chapter we have developed a novel approach to the problem involving optimal trading at the tactical level. We suggest using geometric Brownian motion as the driving process for the asset along with an asset proportional control parameter (trading strategy) and asset proportional exponential-decay parameter. This is distinct from the standard Brownian motion and non-proportional control and exponential-decay parameter used by Guéant et al. (2012b). The trading strategies we found were variable with respect to the asset price, as opposed to the trading strategies found by Guéant et al. (2012b) which were found to be constant for all asset values. The former is a novel characteristic we feel would be more realistic in the finance industry.

Focusing on the problem, we reduced the four-dimensional HJB equation to a three-dimensional non-linear PDE with explicit non-linear term by finding an explicit form of the optimal control in terms of the reduced variable function. We also used a change of variables (non-dimensionalisation) which eliminated two parameters from the model. We used numerical methods to solve this problem, noting that the three-dimensional PDE approach is much less computationally expensive than solving the full HJB equation and is thus much more attractive from an algorithmic trading perspective.

We investigated both the small-time-to-termination solution and the perpetual solution, which provided us with a deeper understanding of the solution topology.

The small-$\tilde{\tau}$ solution provided trading strategies which could be calculated extremely quickly and thus would be especially useful for a high-frequency trading framework in which little time is left to liquidate the portfolio and a quick solution is needed. The perpetual solution on the other hand provided trading strategies which could be implemented for a trader with a long time remaining before expiry. We found a constraint for the perpetual case similar (but different from) that found by Guéant et al. (2012b) under standard Brownian motion.

Solving this problem numerically has led to the development of a methodology which can easily be generalised. This was highlighted in section 4.4 by changing the diffusion process for the asset price from a geometric Brownian motion to a mean-reverting process, that being the CIR process. Implementing other such mean-reverting processes such as the Uhlenbeck and Ornstein (1930) process, the Dixit and Pindyck (1994) process or the Schwartz (1997) process would also be straightforward using our numerical methods. As well as changing the stochastic process driving the asset price, numerical methods allow us to expand on this problem in many forms to make it more mathematically interesting and financially realistic. This will be seen in the coming chapters.

Chapter 5

Extension of Guéant et al. (2012b) to Multiple Underlyings

In this chapter we shall extend the framework of Guéant et al. (2012b), who consider liquidation for a single underlying, to a trader who has a basket of correlated assets to liquidate. Considering multiple underlyings, in which the correlation is accounted for, is critical in optimal trading from a risk-management perspective. Many algorithmic trading strategies prefer to have little or no risk overnight or over weekends. With correlated assets the risk in portfolios can sometimes be diversified or theoretically eliminated completely, resulting in reduced risk even if positions are held outside of market hours.

Guéant and Lehalle (2013) suggest the use of multiple correlated assets in the appendix of their paper, simply deriving an HJB equation, leaving much room for analytic and numerical analysis. Our contribution lies in reducing an HJB equation to a system of non-linear ODEs before further reducing it to a system of linear ODEs by introducing a novel parameter constraint. We solve these numerically and discuss the resultant solutions. After solving a small-time solution, we examine the perpetual case in which the non-linear ODE is simply a non-linear difference equation, while the linear ODE is a simple linear difference equation. Albeit tedious and time consuming, the latter difference equation can be solved analytically by recursion. We can thus verify our numerics with the analytic counterpart.

5.1 Problem formulation for N underlyings

We follow the same problem formulation as discussed in section 3.5 but developed for multiple assets. We consider a trader who wishes to maximise his expected utility, given a portfolio of assets to liquidate, which contains various assets, before a specified terminal time, T. We assume at time $t = 0$ that the trader starts with an initial inventory of $q_i(0)$ assets of asset i, in which q_i takes positive integer values, and an initial wealth $X(0)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0, T])$. We assume the asset's reference price, $S_i(t)$, for asset i follows a standard Brownian motion, and so the diffusion process is defined as

$$dS_i(t) = \mu_i dt + \sigma_i dW_i(t) \tag{5.1}$$

where μ_i is the drift of asset i, σ_i is the volatility of asset i, and $W_i(t)$ is a Wiener process which is $\mathcal{F}_t$ measurable with

$$\mathbb{E}[W_i(t) W_j(t)] = \rho_{ij} t, \tag{5.2}$$

in which $\rho_{ij} = \rho_{ji}$, $-1 \leq \rho_{ij} \leq 1$, and $\rho_{ii} = 1$, implying the various assets are correlated for $\rho_{ij} \neq 0$.

Correlated assets are seen throughout finance and are used every day by investors, traders and hedge funds for risk-management. We have included figure 5.1 for a reference of typical correlation between bonds and stocks[1]. We can see a strong positive correlation of 0.96 between LT.NS (Larsen and Toubro Ltd which is an Indian company that specializes in the engineering, I.T. and shipbuilding industries) and SBIN.NS (State Bank of India). This strong correlation means that these assets will likely move in a similar direction. A strong negative correlation of -0.80 can be seen between SUNPHARMA.NS (Sun Pharmaceutical Ltd which is a manufacturer of sun care products in India) and ONGC.NS (Oil and Natural Gas Corporation Ltd which is an Indian state-owned oil and gas company). We expect these assets to have opposite movements. A very weak (almost negligible) correlation of 0.02 can be seen between MARUTI.NS (Maruti Suzuki India Ltd which is India's largest passenger car company) and ONGC.NS (Oil and Natural Gas Corporation Ltd). These companies

[1] Obtained from http://www.marketcalls.in/trading-lessons/correlation-study-on-nifty-stocks.html on 30/09/14

	GRASIM.NS	BHEL.NS	SBIN.NS	HEROHONDA.NS	LT.NS	MARUTI.NS	SUNPHARMA.NS	ONGC.NS	RELIANCE.NS	ICICIBANK.NS	TCS.NS	SIEMENS.NS	GAIL.NS	ITC.NS	NTPC.NS	SAIL.NS
GRASIM.NS		0.41	0.46	-0.65	0.34	0.83	-0.05	0.24	-0.08	0.61	0.63	0.43	0.24	-0.42	-0.65	0.3
BHEL.NS	0.41		0.77	-0.47	0.78	0.22	-0.47	0.78	0.75	0.81	0.26	0.79	0.48	0.45	0.18	0.94
SBIN.NS	0.46	0.77		-0.78	0.96	0.32	-0.66	0.8	0.64	0.95	0.33	0.95	0.6	0.29	-0.07	0.82
HEROHONDA.NS	-0.65	-0.47	-0.78		-0.73	-0.44	0.63	-0.68	-0.19	-0.79	-0.32	-0.82	-0.41	0.26	0.47	-0.54
LT.NS	0.34	0.78	0.96	-0.73		0.23	-0.76	0.88	0.69	0.9	0.13	0.94	0.69	0.37	-0.05	0.87
MARUTI.NS	0.83	0.22	0.32	-0.44	0.23		0.09	0.02	-0.22	0.37	0.62	0.19	0.29	-0.45	-0.67	0.1
SUNPHARMA.NS	-0.05	-0.47	-0.66	0.63	-0.76	0.09		-0.8	-0.47	-0.57	0.24	-0.77	-0.37	-0.21	0.1	-0.64
ONGC.NS	0.24	0.78	0.8	-0.68	0.88	0.02	-0.8		0.7	0.8	0.05	0.87	0.65	0.35	0.05	0.91
RELIANCE.NS	-0.08	0.75	0.64	-0.19	0.69	-0.22	-0.47	0.7		0.64	-0.06	0.64	0.42	0.76	0.51	0.84
ICICIBANK.NS	0.61	0.81	0.95	-0.79	0.9	0.37	-0.57	0.8	0.64		0.39	0.93	0.58	0.23	-0.15	0.85
TCS.NS	0.63	0.26	0.33	-0.32	0.13	0.62	0.24	0.05	-0.06	0.39		0.17	0.12	-0.17	-0.29	0.14
SIEMENS.NS	0.43	0.79	0.95	-0.82	0.94	0.19	-0.77	0.87	0.64	0.93	0.17		0.49	0.24	-0.06	0.86
GAIL.NS	0.24	0.48	0.6	-0.41	0.69	0.29	-0.37	0.65	0.42	0.58	0.12	0.49		0.3	-0.09	0.57
ITC.NS	-0.42	0.45	0.29	0.26	0.37	-0.45	-0.21	0.35	0.76	0.23	-0.17	0.24	0.3		0.7	0.51
NTPC.NS	-0.65	0.18	-0.07	0.47	-0.05	-0.67	0.1	0.05	0.51	-0.15	-0.29	-0.06	-0.09	0.7		0.16
SAIL.NS	0.3	0.94	0.82	-0.54	0.87	0.1	-0.64	0.91	0.84	0.85	0.14	0.86	0.57	0.51	0.16	

Figure 5.1: Examples of correlation between securities and bonds.

would be expected to move independently of each other and a movement in one would be almost useless in forecasting a movement in the other.

The trader will continuously post orders into the ask side of the limit-order book. For asset i he will post orders for price $S_i^a(t)$ which is $\delta_i = \delta_i(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N)$ greater than the asset value $S_i(t)$, such that

$$S_i^a(t) = S_i\left(t\right) + \delta_i, \tag{5.3}$$

which is similar to that as used previously by Guéant et al. (2012b), rather than $S_i^a(t) = S_i\left(t\right)\left(1 + \delta_i\right)$ as used previously in chapter 4.

Asset sales follow a Poisson process, $N_i(t)$, for asset i, with time-dependent intensity, which is $\mathcal{F}_t$ measurable and independent of both $W_j(t)$ and $N_j\left(t\right)$. The inventory process is thus

$$dq_i(t) = -dN_i(t). \tag{5.4}$$

The Poisson processes are uncorrelated and, as discussed in section 2.1, the probability of more than one jump occurring in any infinitesimal small instant Δt is $O\left(\Delta t^2\right)$ and thus negligible to leading order. The dynamics of the wealth is given by

$$dX_i(t) = \sum_{i=1}^{N}\left(S_i(t) + \delta_i\right) dN_i(t), \tag{5.5}$$

where $N_i(t)$ is the same Poisson process as before. $N_i(t)$ has intensity $\Lambda_i(\delta_i)$ which takes the form:

$$\Lambda_i(\delta_i) = \lambda_i e^{-l_i\left(S_i^a - S_i\right)} = \lambda_i e^{-l_i\delta_i} \tag{5.6}$$

for some positive constants λ_i and l_i. This is consistent with that previously used by Guéant et al. (2012b) but varies from the intensity implemented in chapter 4.

The objective is to liquidate this portfolio before some final time T while maximising the trader's utility. The utility we seek to maximise takes the form of a negative exponential function and as such the trader has CARA. We define our value function

$$u(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N) = \sup_{\underline{\delta}(t) \in \mathcal{A}} \mathbb{E}\left[-e^{-\gamma\left(X(T)+\sum_{i=1}^{N} q_i(T)S_i(T)\right)}\right], \tag{5.7}$$

with $\underline{\delta} = [\delta_1, \ldots, \delta_N]^\top$.

Given the optimisation problem of (5.7), an HJB equation can be derived by applying the Bellman (1957) principle of optimality and using Itô's lemma:

$$u_t + \sum_{i=1}^{N} \mu_i u_{S_i} + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{2}\rho_{ij}\sigma_i\sigma_j u_{S_iS_j} + \sum_{\substack{i=1 \\ q_i>0}}^{N} \sup_{\delta_i}\left[\lambda_i e^{-l_i\delta}\left(u(t, X + S_i + \delta_i, q_1 \ldots, q_i - 1, \ldots, q_N, S_1, \ldots, S_N) - u\right)\right] = 0, \tag{5.8}$$

with $u = u(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N)$ and conditions

$$u(T, X, q_1, \ldots, q_N, S_1, \ldots, S_N) = -e^{-\gamma\left(X(T)+\sum_{i=1}^{N} q_i(T)S_i(T)\right)}, \tag{5.9}$$

$$u(t, X, q_i = 0, S_1, \ldots, S_N) = -e^{-\gamma X(T)} \ \forall\, i. \tag{5.10}$$

Guéant and Lehalle (2013) provide a derivation of (5.8) and a verification theorem.

We cannot implement the same methodology as used in Guéant et al. (2012b) for a single underlying, as discussed in section 3.5, to reduce the HJB equation (5.8) to a linear ODE unless we impose stricter conditions. Therefore, we shall first investigate the general case before introducing a case involving a novel parameter constraint.

For the general case we assume an ansatz form of the solution

$$u\left(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N)\right) = -e^{-\gamma\left(X(T)+\sum_{i=1}^{N} q_i(T)S_i(T)\right)} f\left(t, q_1, \ldots, q_N\right) \tag{5.11}$$

Writing $f = f\left(t, q_1, \ldots, q_N\right)$ we use this ansatz solution to reduce (5.8) to:

$$f_t + \left(\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{2}\rho_{ij}\sigma_i\sigma_j q_i q_j \gamma^2 - \sum_{i=1}^{N} \mu_i q_i \gamma\right) f - \sum_{\substack{i=1 \\ q_i>0}}^{N} \frac{\lambda_i \gamma f}{\gamma + l_i}\left(\frac{l_i f\left(t, q_1, \ldots, q_N\right)}{(\gamma + l_i) f\left(t, q_1, \ldots, q_i - 1, \ldots, q_N\right)}\right)^{\frac{l_i}{\gamma}} = 0 \tag{5.12}$$

by finding the form of the optimal strategy

$$\delta_i^* = \frac{1}{\gamma} \ln\left(\frac{(\gamma + l_i) f\left(t, q_1, \ldots, q_i - 1, \ldots, q_N\right)}{l_i f\left(t, q_1, \ldots, q_N\right)}\right), \tag{5.13}$$

and substituting it into the HJB equation (5.8). Equation (5.12) has initial conditions:

$$f\left(t=T,q_1,\ldots,q_N\right)=1, \tag{5.14}$$

$$f\left(t,q_i=0\right)=1\ \ \forall\, i. \tag{5.15}$$

If we restrict the exponential-decay factor l_i such that $l_i = l\ \ \forall\, i$ we can further reduce the HJB equation (5.8) to a linear ODE. In the single asset case, studied by Guéant et al. (2012b), there is one limit-order book and thus one exponential-decay parameter. In making the restriction $l_i = l\ \ \forall\, i$ we assume the decay of each order book is equal, which could be the case for similar stocks, such as constituents of an index.

Let us assume $l_i = l\ \ \forall\, i$. We can then introduce the ansatz solution

$$u\left(t,X,q_1,\ldots,q_N,S_1,\ldots,S_N\right))=-e^{-\gamma\left(X(T)+\sum_{i=1}^{N}q_i(T)S_i(T)\right)}g\left(t,q_1,\ldots,q_N\right)^{-\frac{\gamma}{l}}. \tag{5.16}$$

Using this ansatz solution we can reduce (5.8) to:

$$\begin{aligned}-g_t+\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{2}\rho_{ij}\sigma_i\sigma_j q_i q_j\gamma l-\sum_{i=1}^{N}\mu_i q_i l\right)g\left(t,q_1,\ldots,q_N\right)\\ -\sum_{\substack{i=1\\ q_i>0}}^{N}\lambda_i\left(\frac{l}{(\gamma+l)}\right)^{1+\frac{l}{\gamma}}g\left(t,q_1,\ldots,q_i-1,\ldots,q_N\right)=0,\end{aligned} \tag{5.17}$$

by finding the form of the optimal strategy

$$\delta_i^*=\frac{1}{l}\ln\left(\frac{g\left(t,q_1,\ldots,q_N\right)}{g\left(t,q_1,\ldots,q_i-1,\ldots,q_N\right)}\right)+\frac{1}{\gamma}\ln\left(\frac{(\gamma+l)}{l}\right), \tag{5.18}$$

with

$$g\left(t=T,q_1,\ldots,q_N\right)=1, \tag{5.19}$$

$$g\left(t,q_i=0\right)=1\ \ \forall\, i. \tag{5.20}$$

For the case of the non-linear ODE (5.12), if we were to set $l_i = l\ \ \forall\, i$ and introduce a function such that

$$g\left(t,q_1,\ldots,q_N\right)=f\left(t,q_1,\ldots,q_N\right)^{-\frac{\gamma}{l}} \tag{5.21}$$

we could reduce the non-linear ODE (5.12) in f to a linear ODE (5.17) in g. Thus (5.21) is the link between the functions f and g. The focus of the remainder of this chapter is on the two-dimensional, $N = 2$, case.

5.1.1 Problem formulation for two underlyings

For $N = 2$ and $\tau = T - t$, so we are now working forwards in τ rather than backwards in t, equation (5.12) and (5.17) reduce to

$$-f_\tau(\tau, q_1, q_2) + \frac{\gamma}{l}\beta(q_1, q_2) f(\tau, q_1, q_2) - \frac{\lambda_1 \gamma f(\tau, q_1, q_2)}{\gamma + l_1}\left(\frac{l_1 f(\tau, q_1, q_2)}{(\gamma + l_1) f(\tau, q_1 - 1, q_2)}\right)^{\frac{l_1}{\gamma}} - \frac{\lambda_2 \gamma f(\tau, q_1, q_2)}{\gamma + l_2}\left(\frac{l_2 f(\tau, q_1, q_2)}{(\gamma + l_2) f(\tau, q_1, q_2 - 1)}\right)^{\frac{l_2}{\gamma}} = 0 \tag{5.22}$$

and

$$g_\tau(\tau, q_1, q_2) + \beta(q_1, q_2) g(\tau, q_1, q_2) - \lambda_1\left(\frac{l}{(\gamma + l)}\right)^{1+\frac{l}{\gamma}} g(\tau, q_1 - 1, q_2) - \lambda_2\left(\frac{l}{(\gamma + l)}\right)^{1+\frac{l}{\gamma}} g(\tau, q_1, q_2 - 1) = 0 \tag{5.23}$$

respectively, with the restriction of $l_1 = l_2 = l$ needed for (5.23) to exist, with

$$\beta(q_1, q_2) = \frac{1}{2}\sigma_1^2 q_1^2 \gamma l + \frac{1}{2}\sigma_2^2 q_2^2 \gamma l + \rho\sigma_1\sigma_2 q_1 q_2 \gamma l - \mu_1 q_1 l - \mu_2 q_2 l \tag{5.24}$$

and boundary conditions

$$f(\tau = 0, q_1, q_2) = g(\tau = 0, q_1, q_2) = 1, \tag{5.25}$$

$$f(\tau, q_1 = 0, q_2 = 0) = g(\tau, q_1 = 0, q_2 = 0) = 1, \tag{5.26}$$

for both (5.22) and (5.23). If either $q_1 = 0$ or $q_2 = 0$ the respective $q_i - 1$ term in the ODE is neglected. When solving the non-linear ODE (5.22) the optimal trading strategies take the form

$$\delta_1^*(\tau, q_1, q_2) = \frac{1}{\gamma}\ln\left(\frac{(\gamma + l_1) f(\tau, q_1 - 1, q_2)}{l_1 f(\tau, q_1, q_2)}\right), \tag{5.27}$$

$$\delta_2^*(\tau, q_1, q_2) = \frac{1}{\gamma}\ln\left(\frac{(\gamma + l_2) f(\tau, q_1, q_2 - 1)}{l_2 f(\tau, q_1, q_2)}\right), \tag{5.28}$$

while for the linear ODE (5.23) the optimal strategies take the form

$$\delta_1^*(\tau, q_1, q_2) = \frac{1}{l}\ln\left(\frac{g(\tau, q_1, q_2)}{g(\tau, q_1 - 1, q_2)}\right) + \frac{1}{\gamma}\ln\left(\frac{(\gamma + l)}{l}\right), \tag{5.29}$$

$$\delta_2^*(\tau, q_1, q_2) = \frac{1}{l}\ln\left(\frac{g(\tau, q_1, q_2)}{g(\tau, q_1, q_2 - 1)}\right) + \frac{1}{\gamma}\ln\left(\frac{(\gamma + l)}{l}\right). \tag{5.30}$$

As far as we are aware analytic solutions for these ODEs do not exist and thus they must be solved numerically. For (5.22) we used the iterative procedure as outlined in section 2.6.3 for each $\{q_1, q_2\}$. This iterative method is constrained to $O(\Delta\tau)$ which is why it can be unfavourable. For (5.23) we used the classic fourth-order explicit Runge-Kutta method, as outlined in section 2.6.2 and discussed further in section 3.5.1 for systems of ODEs. This method has an error $O(\Delta\tau^4)$ which is more favourable over the iterative method, however we are constrained to $l_1 = l_2 = l$ for this method. Both $O(\Delta\tau)$ for the iterative method and $O(\Delta\tau^4)$ for the Runge-Kutta method were found for (5.22) and (5.23) respectively, as can be seen in table 5.1.

Table 5.1: Convergence of iterative scheme and Runge-Kutta fourth-order scheme

N	$\Delta\tau$	$f(\tau = 1, q_1 = 1, q_2 = 1)^{-\gamma/l}$	ratio	$g(\tau = 1, q_1 = 1, q_2 = 1)$	ratio
11	0.100000	10.4662419795	NA	8.1917520759	NA
21	0.050000	9.17979422942	NA	8.1917946177	NA
41	0.025000	8.65506463578	2.45	8.1917971298	16.94
81	0.012500	8.41641215446	2.20	8.1917972824	16.46
161	0.006250	8.30242508412	2.09	8.1917972918	16.23
321	0.003125	8.24670021424	2.05	8.1917972924	16.11

N is the number of time points with $\Delta\tau$ as the length of the intervals between respective points. We have set $\tau = 1, \mu_1 = \mu_2 = 0.05, \sigma_1 = \sigma_2 = 3, \gamma = 0.5, \lambda_1 = \lambda_2 = 0.3$ and $l_1 = l_2 = l = 0.3$. We can see the iterative scheme exhibits linear convergence while the Runge-Kutta scheme exhibits quartic convergence.

5.1.2 Numerical results

In the case of $q_1 > 0, q_2 = 0$ or $q_1 = 0, q_2 > 0$ the system simplifies down to the single asset problem in $q_1, \tilde{S}_1$ and $q_2, \tilde{S}_2$ respectively, in which the results of Guéant et al. (2012b), discussed in section 3.5, hold. Thus we shall numerically examine the case of $q_1, q_2 > 0$. In the next chapter we extend the framework of chapter 4 to multiple underlying assets. We do a thorough comparative analysis when parameter values are asymmetric. Many of the results in the standard Brownian motion case are analogous to that of the geometric Brownian motion case (in terms of how they behave for varying parameters) as discussed in the next chapter and thus we will not examine them here. Here we shall focus on the effect of the correlation.

Having a portfolio of correlated assets to reduce risk was first examined by Markowitz

(1959) in what became known as Modern Portfolio Theory (MPT). Markowitz introduced the concept of 'efficiency' and 'market portfolios' in which he showed that a combination of correlated assets could provide a higher expected return with less volatility than holding individual assets. From MPT we expect a trader with negatively correlated assets to want a balance of assets in his portfolio, such that $q_1 = q_2$, while for positively correlated assets the trader would seek an unbalanced portfolio.

We will focus our attention to the optimal trading strategy for the various regimes of the correlation parameter, ρ. For symmetric parameter regimes

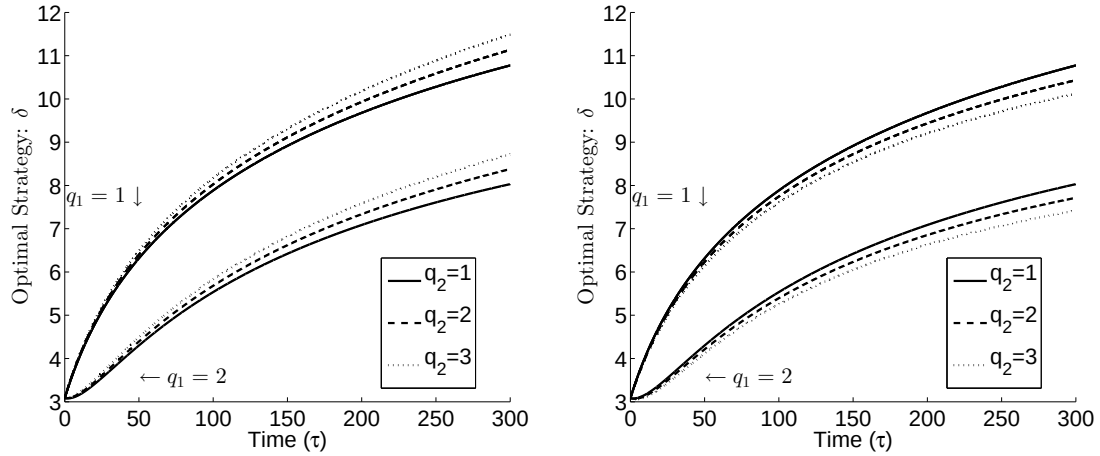
$$\delta_1^*(\tau, q_1, q_2) = \delta_2^*(\tau, q_1, q_2)$$

and so we only need to focus on either δ_1^* or δ_2^*. The main interest is how the optimal strategy varies for various values of the inventory, $\{q_1, q_2\}$, under different correlation, ρ, regimes. The parameters we choose are analogous to those of the single asset case of Guéant et al. (2012b), but used for both assets (given we examine a symmetric regime).

For the case of $\rho = 0$, numerical simulations indicated conclusively that there is no difference in the optimal trading strategies for asset one, δ_1^*, for various q_2 (and vice-versa). This was pre-supposed given the randomness driving the asset prices, as well as the jump process driving asset sales, are independent for asset one and two. We shall return to this below with analytic analysis when examining the steady-state.

For non-zero ρ, the portfolio is more risky for positive ρ and less risky for negative ρ. We thus expect lower optimal trading strategies for $\rho > 0$ and higher optimal trading strategies of $\rho < 0$. This can be seen in figure 5.2. Here we have plotted curves of δ_1^* against time for various $\{q_1, q_2\}$. The curves that initially group together correspond to constant q_1, with q_2 being varied which corresponds to their divergence. If ρ was zero than we would see only two curves, one corresponding to $q_1 = 1$ for various q_2 (which lie on top of each other) while the other would correspond to $q_1 = 2$, again for various q_2.

We see in figure 5.2(a) how the optimal trading strategies of asset one, δ_1^*, increase in q_2 for constant $q_1 = 1$ (top three curves) as well as increase in q_2 for constant $q_1 = 2$ (bottom three curves) when ρ is negative. This is because the trader can now hedge his risk so a more balanced portfolio is preferred. In the case of $\rho > 0$, as in figure



(a) δ_1^* against time for $\rho < 0$ (b) δ_1^* against time for $\rho > 0$

Figure 5.2: Optimal trading strategy for asset one, δ_1^*, against time for various combinations of assets for various correlations. Negative ρ results in higher asking prices while positive ρ results in lower asking prices, with the differences between asking prices diverging as time remaining increases. The parameter values used are symmetric, taking the values $\mu_1 = \mu_2 = 0.0, \sigma_1 = \sigma_2 = 0.3, \lambda_1 = \lambda_2 = 0.1, l = 0.3, \gamma = 0.05$ and $\rho = \pm 0.75$.

5.2(b), the opposite is true. The trader now wants to sell his inventory of asset one quicker if he holds more of asset two as to reduce the risk in the portfolio given the positive correlation of the portfolio. We see that the solutions are similar near the terminal time and diverge as the time remaining increases. Similar results occur if we were to examine δ_2^*.

As with the single underlying case, discussed in section 3.5, we see the trading strategies are asset-price independent. This means the additional amount we add to the asset does not dependent on its current price, which was one of the motivations for extending the framework discussed in section 3.5, that of Guéant et al. (2012b).

5.2 Small-time-to-termination solution

In this section we shall examine the small-time-to-termination solution of (5.23). We choose the $l_1 = l_2 = l$ case as the $l_1 \neq l_2$ case leads to the non-linear ODE (5.22) and small-time-to-termination investigation of this problem is similar (in terms of the non-linear term) to that we will perform in section 6.3, which is a multiple asset extension of chapter 4 (geometric Brownian motion driving the asset price). Given what we have

learnt from previous small-time-to-termination examination of this kind of problem, see section 4.2, we decide to perform a non-dimensionalisation. We did not perform a non-dimensionalisation previously as we wanted our solutions to be comparable with Guéant et al. (2012b) who do not work with dimensionless variables.

The non-dimensionalisation we use is

$$\tilde{\tau} = \lambda_1 \tau, \quad \tilde{\mu_i} = \frac{\mu_i}{S_{i,0}\lambda_1}, \quad \tilde{\sigma_i} = \frac{\sigma_i}{S_{i,0}\sqrt{\lambda_1}}, \quad \lambda = \frac{\lambda_2}{\lambda_1}, \quad \tilde{S}_i = S_i\gamma, \quad \tilde{l} = \frac{l}{\gamma} \tag{5.31}$$

with $S_i(0) = S_{i,0}$ for $i \in \{1,2\}$, which is assumed known. Using (5.31) with (5.23) we arrive at

$$g_{\tilde{\tau}}(\tilde{\tau}, q_1, q_2) + \tilde{\beta}(q_1, q_2)\, g(\tilde{\tau}, q_1, q_2) - \left(\frac{\tilde{l}}{\left(1+\tilde{l}\right)}\right)^{1+\tilde{l}} g(\tilde{\tau}, q_1 - 1, q_2)$$
$$-\lambda \left(\frac{\tilde{l}}{\left(1+\tilde{l}\right)}\right)^{1+\tilde{l}} g(\tilde{\tau}, q_1, q_2 - 1) = 0 \tag{5.32}$$

with

$$\tilde{\beta}(q_1, q_2) = \frac{1}{2}\tilde{S}_{1,0}^2\tilde{\sigma}_1^2 q_1^2\tilde{l} + \frac{1}{2}\tilde{S}_{1,0}^2\tilde{\sigma}_2^2 q_2^2\tilde{l} + \rho\tilde{S}_{1,0}\tilde{S}_{2,0}\tilde{\sigma}_1\tilde{\sigma}_2 q_1 q_2\tilde{l} - \tilde{S}_{1,0}\tilde{\mu}_1 q_1\tilde{l} - \tilde{S}_{2,0}\tilde{\mu}_2 q_2\tilde{l}, \tag{5.33}$$

in which we have eliminated the parameters γ and λ_1 from the equations.

We perform a small-time perturbation in $\tilde{\tau}$ which is a power series in $\tilde{\tau}$ and takes the form

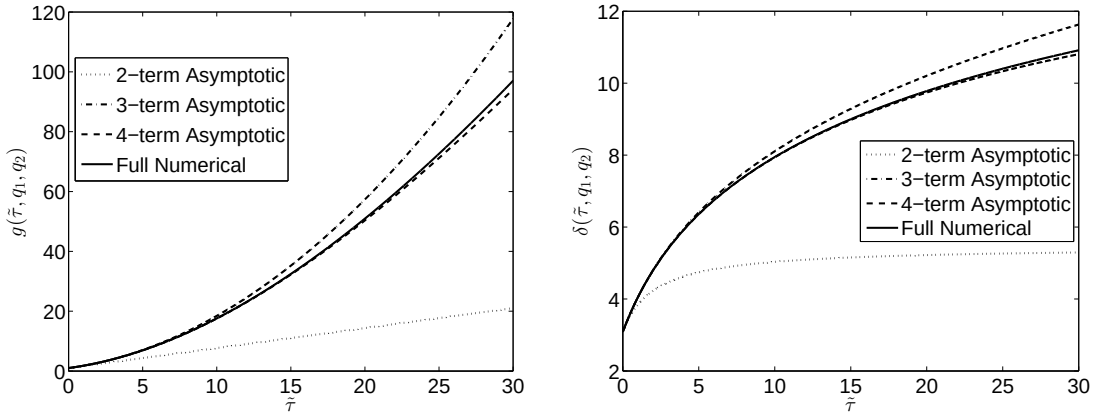
$$g(\tilde{\tau}, q_1, q_2) = g_0(q_1, q_2) + \tilde{\tau} g_1(q_1, q_2) + \tilde{\tau}^2 g_2(q_1, q_2) + O\left(\tilde{\tau}^3\right) \tag{5.34}$$

with $g_0(q_1, q_2) = g(\tilde{\tau} = 0, q_1, q_2)$ which is given by (5.25). Substituting (5.34) into (5.32) we arrive with a recursive system for $g_i(q_1, q_2)$ which takes the form

$$g_i(q_1, q_2) = -\frac{1}{i}\left(\tilde{\beta}(q_1, q_2)\, g_{i-1}(q_1, q_2) - \left(\frac{\tilde{l}}{\left(1+\tilde{l}\right)}\right)^{1+\tilde{l}} g_{i-1}(q_1 - 1, q_2)\right.$$
$$\left. -\lambda \left(\frac{\tilde{l}}{\left(1+\tilde{l}\right)}\right)^{1+\tilde{l}} g_{i-1}(q_1, q_2 - 1)\right) \tag{5.35}$$

with $g_i(q_1 > 0, q_2 < 0) = g_i(q_1 < 0, q_2 > 0) = 0$ for $i \geq 1$.

Figure 5.3 shows a comparison of the accuracy of the small-time solution, for various numbers of terms, against the full-numerical solution of (5.23) solved using the Runge-Kutta method. This is done for both $g(\tilde{\tau}, q_1, q_2)$ (figure 5.3(a)) and $\delta_1^*(\tilde{\tau}, q_1, q_2)$ (figure



(a) Asymptotic approximation for $g(\tilde{\tau}, q_1, q_2)$ (b) Asymptotic approximation for $\delta_1^*(\tilde{\tau}, q_1, q_2)$

Figure 5.3: Examination of asymptotic approximation for small-time solution verses the full-numerical solution for $g(\tilde{\tau}, q_1, q_2)$ and $\delta_1^*(\tilde{\tau}, q_1, q_2)$. We compare a two-term, three-term and four-term approximation against the numerical solution. The parameters we use are symmetric and correspond to those of figure 5.2, taking the values $q_1 = q_2 = 1, \tilde{\mu}_1 = \tilde{\mu}_2 = 0, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 1.27, \tilde{l} = 0.6, \lambda = 1, \tilde{S}_{1,0} = \tilde{S}_{2,0} = 1$ and $\tilde{\tau} = 30$.

5.3(b)). We can see this is a very good approximation for $\tilde{\tau}$ up to $O(10)$. However, this is parameter dependent which we now discuss.

Let us refer back to the single asset case which was solved analytically for $q = 1$. This has solution (3.30) as discussed in section 3.5.1. It can be seen that this solution grows exponentially in time, and the rate of its growth depends on the combination

$$\frac{1}{2} l\gamma\sigma^2 - l\mu, \tag{5.36}$$

which is equal to $\beta(1,0) = \beta(0,1)$, assuming symmetric parameters. Given the ODE (5.23) has a similar form as the single asset case, as given by (3.21), but with a second difference term, we would expect the solution of (5.23) to be of similar form (but more complicated) to that of (3.31), with a similar exponential-growth term. Assuming this true, the larger the value of (5.36) the worse we expect the accuracy of the power series expansion in $\tilde{\tau}$. This is what we are finding (numerically) is the case and hence the reason for the accurate solution found in figure 5.3(a) and figure 5.3(b) for $\tilde{\tau} \approx 30$.

5.3 Steady-state case

The steady-state solution has previously been examined for the single asset case by Guéant et al. (2012b). In this section we shall focus on the steady-state case but with

the novel constraint $l_i = l \ \forall i$ as so we can relate it back to the single asset case of Guéant et al. (2012b), and for that reason we also return to the dimensional variables of section 5.1.1, rather than dimensionless that were just discussed in section 5.2. We did consider the case of $l_1 \neq l_2$ which results in a non-linear algebraic equation. As neither an analytic solution can be found, nor interesting asymptotic analysis be performed, it is of little benefit or interest to us.

The steady state of (5.23) is given by

$$\beta\left(q_1, q_2\right) g\left(q_1, q_2\right) = \lambda_1 \left(\frac{l}{(\gamma + l)}\right)^{1+\frac{l}{\gamma}} g\left(q_1 - 1, q_2\right) + \lambda_2 \left(\frac{l}{(\gamma + l)}\right)^{1+\frac{l}{\gamma}} g\left(q_1, q_2 - 1\right) \tag{5.37}$$

with

$$g\left(q_1 = 0, q_2 = 0\right) = 1 \tag{5.38}$$

and $\beta\left(q_1, q_2\right)$ given by (5.24). As with the non-steady case, if either $q_1 = 0$ or $q_2 = 0$ the respective $q_i - 1$ term in the ODE is neglected. For the single asset case Guéant et al. (2012b) find analytic solutions for the difference and as such we can implement that to the case when either $q_1 = 0$ or $q_2 = 0$. If $q_1 > 0$ and $q_2 = 0$, equation (5.37) has solution

$$g\left(q_1, 0\right) = \frac{\eta_1^{q_1}}{q_1!} \prod_{i=1}^{q_1} \frac{1}{\frac{l\gamma\sigma_1^2}{2} i - l\mu_1}, \tag{5.39}$$

while if $q_1 = 0$ and $q_2 > 0$, equation (5.37) has solution to

$$g\left(0, q_2\right) = \frac{\eta_2^{q_2}}{q_2!} \prod_{i=1}^{q_2} \frac{1}{\frac{l\gamma\sigma_2^2}{2} i - l\mu_2}, \tag{5.40}$$

with

$$\eta_i = \lambda_i \left(\frac{l}{(\gamma + l)}\right)^{1+\frac{l}{\gamma}}. \tag{5.41}$$

It was found by Guéant et al. (2012b) for the single asset case, which is analogous to our finding in section 4.3, that for a steady-state solution to exist

$$\mu < \frac{1}{2}\sigma^2\gamma \tag{5.42}$$

must be satisfied. We find an analogous constraint for the multiple asset case, that being

$$\mu_i < \frac{1}{2}\sigma_i^2\gamma \quad \text{for } i \in \{1, 2\}. \tag{5.43}$$

For the multiple asset case another constraint was found, notably

$$\beta(q_1, q_2) = \frac{1}{2}\sigma_1^2 q_1^2 \gamma l + \frac{l}{2}\sigma_2^2 q_2^2 \gamma l + \rho\sigma_1\sigma_2 q_1 q_2 \gamma l - \mu_1 q_1 l - \mu_2 q_2 l \neq 0. \tag{5.44}$$

From (5.39) and (5.40) we can find an analytic expression for $\delta_1^*(q_1, 0)$ and $\delta_2^*(0, q_2)$ which take the form

$$\delta_1^*(q_1, 0) = \frac{1}{l}\ln\left(\frac{\eta_1}{\frac{l}{2}\gamma\sigma_1^2 q_1^2 - l\mu_1 q_1}\right) + \frac{1}{\gamma}\ln\left(\frac{\gamma + l}{l}\right), \tag{5.45}$$

$$\delta_2^*(0, q_2) = \frac{1}{l}\ln\left(\frac{\eta_2}{\frac{l}{2}\gamma\sigma_2^2 q_2^2 - l\mu_2 q_2}\right) + \frac{1}{\gamma}\ln\left(\frac{\gamma + l}{l}\right), \tag{5.46}$$

respectively. From (5.45) and (5.46) we see that the optimal strategy is increasing in λ_i and μ_i, decreasing in γ, q and σ_i and can be increasing or decreasing in l, depending on whether δ_i^* is positive or negative, as described in chapter 4.

We are yet to find a fully analytic solution for (5.37), however we can use recursive methods to calculate analytic solutions at each $\{q_1, q_2\}$, albeit this can be quite tedious and become quite complicated as $\{q_1, q_2\}$ grow. We shall consider the case of $\{q_1 = 1, q_2 = 1\}$. From (5.37) we find

$$g(1,1) = \left(\frac{\eta_1\eta_2}{\frac{l}{2}\gamma\sigma_1^2 - l\mu_1 + \frac{l}{2}\gamma\sigma_2^2 - l\mu_2 + \rho\sigma_1\sigma_2\gamma l}\right)\left(\frac{1}{\frac{l}{2}\gamma\sigma_1^2 - l\mu_1} + \frac{1}{\frac{l}{2}\gamma\sigma_2^2 - l\mu_2}\right) \tag{5.47}$$

which results in

$$\begin{aligned}\delta_i^*(1,1) = \frac{1}{l}\ln\left(\frac{\eta_i}{\frac{l}{2}\gamma\sigma_i^2 - l\mu_i} \times \frac{\frac{1}{2}\gamma\sigma_1^2 - \mu_1 + \frac{1}{2}\gamma\sigma_2^2 - \mu_2}{\frac{1}{2}\gamma\sigma_1^2 - \mu_1 + \frac{1}{2}\gamma\sigma_2^2 - \mu_2 + \rho\sigma_1\sigma_2\gamma}\right) \\ + \frac{1}{\gamma}\ln\left(\frac{\gamma + l}{l}\right)\end{aligned} \tag{5.48}$$

for $i \in \{1, 2\}$. Notice that if $\rho = 0$

$$\delta_i^*(1,1) = \frac{1}{l}\ln\left(\frac{\eta_i}{\frac{l}{2}\gamma\sigma_i^2 - l\mu_i}\right) + \frac{1}{\gamma}\ln\left(\frac{\gamma + l}{l}\right) = \frac{1}{l}\ln\left(\frac{\lambda_i}{l + \gamma} \times \frac{1}{\frac{1}{2}\gamma\sigma_i^2 - \mu_i}\right) \tag{5.49}$$

which is identical to the single asset case for $q = 1$, i.e. (5.45) with $q_1 = 1$. This indicates that $\delta_1^*(1, 0) = \delta_1^*(1, 1)$ when $\rho = 0$ and thus when selling asset one the trader asks for the same price regardless of whether he has one or no assets of asset two remaining. Numerical simulations indicated conclusively that δ_1^* is independent of q_2, and δ_2^* is independent of q_1, when $\rho = 0$. Considering again (5.48) but now for the

case of $\rho \neq 0$. When ρ is positive this results in higher risk as there is less ability to hedge, and from (5.48) we can see this lowers the value of $\delta_i^*(1,1)$. Contrastingly, for $\rho < 0$ the trader can hedge his risk and as such (as seen in (5.48)) the trader asks for a higher price for the asset. This is similar to the unsteady case, as seen in figure 5.2.

5.4 Summary

In this chapter we developed the work of Guéant et al. (2012b) and examined the optimal trading strategy of a trader with various amounts of various assets (rather than an amount of one asset) in which the assets are driven by correlated standard Brownian motions. We began by deriving a general model for N correlated assets before examining in detail the case of $N = 2$.

We reduced the resulting HJB equation to a non-linear ODE, and, with a novel constraint that the exponential-decay factor of both limit-order books are equal, were able to further reduce it to a linear ODE with a difference term. To solve the non-linear ODE we used a standard iterative solver, while for the linear ODE we used a higher-order Runge-Kutta method which results in $O(\Delta t^4)$ convergence (the standard iterative solver had $O(\Delta t)$ convergence). We found the correlation term generates some interest, with results relating back to Modern Portfolio Theory (see Markowitz, 1959), in which a trader with positively correlated assets wants an unbalanced portfolio in order to reduce risk while a trader with negatively correlated assets wants a balanced portfolio so as to diversify the risk.

We preformed some interesting asymptotic analysis. We first examined a small-time solution (which was not considered in the single asset case by Guéant et al. (2012b)). Performing a non-dimensionalisation, in order to reduce the number of parameters by two, we found a small-time-to-termination solution which was fully analytic, easily calculated, and, for parameter regimes analogous to those examined in Guéant et al. (2012b) for a single asset, resulted in highly accurate solutions. We next examined a perpetual solution which can be solved analytically using recursion. This solution provided analytic insight into the relationship between the strategy of one asset and the amount of inventory of the other asset being held.

Chapter 6

Tactical Level Trading with Multiple Underlyings under Geometric Brownian Motion

The work in this chapter is in preprint format and ready for submission for review:
Blair, J., Johnson, P., and Duck, P. (2015). Analysis of optimal liquidation in limit-order books for portfolios of correlated assets with stochastic volatility.
http://eprints.ma.man.ac.uk/2325.

In this chapter we expand on the work of chapter 4 by considering the case of multiple correlated underlyings. Although the case of a single underlying asset is dominant in the optimal trading literature, multiple underlyings were briefly discussed in the appendix of Guéant and Lehalle (2013) in which they derived the HJB equation for the case of standard Brownian motion for the asset price, as examined in the previous chapter. Furthermore, in the optimal scheduling literature, Schied et al. (2010) and Cartea et al. (2015b) both consider liquidation for a basket of assets with multiple underlyings, as discussed in chapter 3.

Aside from the optimal trading literature, multiple underlyings have been examined in some detail in other areas of financial mathematics. For example, the problem of multiple underlyings under the Black and Scholes (1973) option pricing framework is now well established (see for example Wilmott, 1998). Duck et al. (2014) extend this further by examining perpetual options with multiple underlyings, using both numerical and asymptotic analysis, and it is a similar framework which we use to

examine the perpetual multi-dimensional case.

We shall thus extend our framework of chapter 4, firstly for the time-dependent problem before examining the asymptotic characteristics of the steady-state problem.

6.1 Problem formulation for N underlyings

We follow a similar framework as discussed in chapter 4 but for the case of multiple underlyings, as derived under standard Brownian motion in the previous chapter.

We consider a trader who wishes to maximise his expected terminal time utility, given a portfolio of assets to liquidate, before a specified terminal time, T. We assume at time $t = 0$ the trader starts with an initial inventory of $q_i(0)$ assets of asset i, in which q_i takes non-negative integer values implying we cannot short sell, and an initial cash holding $X(0)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0, T])$. We assume the reference price $S_i(t)$ of asset i follows a geometric Brownian motion, and so the diffusion process is defined as

$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t)dW_i(t), \tag{6.1}$$

with μ_i as the constant relative drift, σ_i as the constant relative volatility and $W_i(t)$ as a Wiener process which is $\mathcal{F}_t$ measurable, with

$$\mathbb{E}[W_i(t) W_j(t)] = \rho_{ij} t, \tag{6.2}$$

in which $\rho_{ij} = \rho_{ji}$, $-1 \leq \rho_{ij} \leq 1$, and $\rho_{ii} = 1$, implying the various assets are correlated for $\rho_{ij} \neq 0$.

As in chapter 4, the trader will continuously post orders into the ask side of the limit-order book. For asset i he will post orders for price $S_i^a(t)$ which is $\delta_i = \delta_i(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N)$ percent greater than the reference price $S_i(t)$, such that

$$S_i^a(t) = S_i(t)(1 + \delta_i). \tag{6.3}$$

Here we are using the same distinct approach as used in chapter 4 of modelling the control as a percent of the asset price rather than an absolute amount, as implemented in chapter 5.

Asset sales follow a Poisson process, $N_i(t)$, for asset i, with time-dependent intensity, which is $\mathcal{F}_t$ measurable and independent of $W_j(t)$, for all j. The inventory process

for asset i is thus

$$dq_i(t) = -dN_i(t). \tag{6.4}$$

The Poisson processes are uncorrelated and, as discussed in section 2.1, the probability of more than one jump occurring in any infinitesimal small instant Δt is $O\left(\Delta t^2\right)$ and thus negligible to leading order.

Each time an asset is sold, the trader's cash increases by the amount that asset was sold for, such that the dynamics of the cash is given by

$$dX(t) = \sum_{i=1}^{N} S_i(t)\left(1+\delta_i\right) dN_i(t), \tag{6.5}$$

where $N_i(t)$ is the same Poisson process as before. Therefore, when a trade occurs, the values of $q_i(t)$ and $X(t)$ change simultaneously, according to (6.4) and (6.5) respectively. $N_i(t)$ has intensity $\Lambda_i(\delta_i)$ which takes the form:

$$\Lambda_i(\delta_i) = \lambda_i e^{-l_i\left(\frac{S_i^a - S_i}{S_i}\right)} = \lambda_i e^{-l_i\delta_i}, \tag{6.6}$$

for some positive constants λ_i and l_i. This is consistent with that of chapter 4 but distinct from that of chapter 5.

The objective is to liquidate this portfolio before some final time T while maximising the trader's utility, which takes the form of a CARA (negative exponential) utility function. We define our value function

$$u(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N) = \sup_{\underline{\delta}(t)\in\mathcal{A}} \mathbb{E}\left[-e^{-\gamma\left(X(T)+\sum_{i=1}^{N} q_i(T)S_i(T)\right)}\right], \tag{6.7}$$

where $\underline{\delta} = [\delta_1, \ldots, \delta_N]^\top$, $\gamma > 0$ is the risk-aversion characterising the trader and $\mathcal{A} \subseteq (-1, \infty)$ is the set of admissible trading strategies. By the same argument as discussed in section 4.1 the objective function, and by definition the value function, is bounded in the set $[-1, 0)$.

Given the optimisation problem of (6.7), an HJB equation can be derived by applying the Bellman (1957) principle of optimality and using Itô's lemma:

$$\begin{aligned} &u_t + \sum_{i=1}^{N} \mu_i S_i u_{S_i} + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{2}\rho_{ij}\sigma_i\sigma_j S_i S_j u_{S_iS_j} \\ &+ \sum_{\substack{i=1 \\ q_i>0}}^{N} \sup_{\delta_i}\left[\lambda_i e^{-l_i\delta}\left(u(t, X + S_i\left(1+\delta_i\right), q_1 \ldots, q_i - 1, \ldots, q_N, S_1, \ldots, S_N) - u\right)\right] = 0, \end{aligned} \tag{6.8}$$

with $u = u(t, X, q_1, \ldots, q_N, S_1, \ldots, S_N)$ and conditions:

$$u(T, X, q_1, \ldots, q_N, S_1, \ldots, S_N) = -e^{-\gamma\left(X(T)+\sum_{i=1}^{N} q_i(T)S_i(T)\right)}, \tag{6.9}$$

$$u(t, X, q_i = 0, S_1, \ldots, S_N) = -e^{-\gamma X(T)} \quad \forall\, i. \tag{6.10}$$

The derivation of (6.8) is analogous to the derivation of the single asset case which is outlined in section 4.1.1, in which a verification theorem is shown, and thus shall not be reiterated.

The above problem has dimension $2N + 1$. The focus of the remainder of this chapter will be the $N = 2$ case.

6.2 Problem formulation for two underlyings

For $N = 2$, let $u = u\left(t, X, q_1, q_2, S_1, S_2\right)$. The HJB equation given by (6.8) is reduced to

$$\begin{aligned}
&u_t + \mu_1 S_1 u_{S_1} + \frac{1}{2}\sigma_1^2 S_1^2 u_{S_1S_1} + \mu_2 S_2 u_{S_2} + \frac{1}{2}\sigma_2^2 S_2^2 u_{S_2S_2} + \rho\sigma_1\sigma_2 S_1 S_2 u_{S_1S_2} \\
&+ \sup_{\delta_1}\left[\lambda_1 e^{-l_1\delta_1}\left(u\left(t, X + S_1\left(1+\delta_1\right), q_1 - 1, q_2, S_1, S_2\right) - u\left(t, X, q_1, q_2, S_1, S_2\right)\right)\right] \\
&+ \sup_{\delta_2}\left[\lambda_2 e^{-l_2\delta_2}\left(u\left(t, X + S_2\left(1+\delta_2\right), q_1, q_2 - 1, S_1, S_2\right) - u\left(t, X, q_1, q_2, S_1, S_2\right)\right)\right] = 0,
\end{aligned} \tag{6.11}$$

with

$$u\left(t = T, q_1, q_2, S_1, S_2\right) = -e^{-\gamma(X+q_1S_1+q_2S_2)}, \tag{6.12}$$

$$u\left(t, q_1 = 0, q_2 = 0, S_2, S_2\right) = -e^{-\gamma X}. \tag{6.13}$$

If $q_i = 0$ and $q_j > 0$ equation (6.11) reduces to

$$\begin{aligned}
&u_t + \mu_j S_j u_{S_j} + \frac{1}{2}\sigma_j^2 S_j^2 u_{S_jS_j} \\
&+ \sup_{\delta_j}\left[\lambda_2 e^{-l_j\delta_j}\left(u\left(t, X + S_j\left(1+\delta_j\right), q_i, q_j - 1, S_1, S_2\right) - u\left(t, X, q_i, q_j, S_1, S_2\right)\right)\right] = 0,
\end{aligned} \tag{6.14}$$

as we cannot short sell and thus the derivative terms with respect to S_i are zero as S_i has no impact on the solution. Equation (6.14) is identical to the single asset case examined in chapter 4.

6.2.1 Reduction and rescaling

We can now make a similar reduction as was performed for the single asset case in section 4.1.2[1]. We introduce the ansatz

$$u(t, X, q_1, q_2, S_1, S_2) = e^{-\gamma X} f(t, q_1, q_2, S_1, S_2), \tag{6.15}$$

which factors out the cash of the trader, and we then use the following change of variables

$$\tilde{\tau} = \lambda_1 (T - t), \quad \tilde{S}_i = \gamma S_i, \quad \tilde{\mu}_i = \frac{\mu_i}{\lambda_1}, \quad \tilde{\sigma}_i = \frac{\sigma_i}{\sqrt{\lambda_1}}, \quad \tilde{\lambda} = \frac{\lambda_2}{\lambda_1}, \tag{6.16}$$

noting that $\tilde{S}_i$ is now the risk-adjusted asset price for asset i, given that it is S_i scaled on γ. From now we will refer to $\tilde{S}_i$ as the asset price of asset i, dropping 'risk-adjusted'. Substituting (6.15) and (6.16) into (6.11) we can solve for the optimal controls, δ_i^*, by differentiating the supremum with respect to the controls and setting the result equal to zero, which locates the stationary point. Solving this we obtain

$$\delta_i^* \left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = \frac{1}{\tilde{S}_i} \ln \left(\frac{\left(\tilde{S}_i + l_i\right) f\left(\tilde{\tau}, q_i - 1, q_j, \tilde{S}_1, \tilde{S}_2\right)}{l_i f\left(\tilde{\tau}, q_i, q_j, \tilde{S}_1, \tilde{S}_2\right)} \right) - 1, \tag{6.17}$$

which we notice is independent of X, hence confirming the use of our ansatz solution. We perform the second derivative test to ensure the stationary point we find is indeed the maximum.

Substituting (6.17), and using (6.15) and (6.16), we can transform (6.11) to

$$\begin{aligned} &- f_{\tilde{\tau}} + \tilde{\mu}_1 \tilde{S}_1 f_{\tilde{S}_1} + \tilde{\mu}_2 \tilde{S}_2 f_{\tilde{S}_2} + \frac{1}{2}\tilde{\sigma}_1^2 \tilde{S}_1^2 f_{\tilde{S}_1\tilde{S}_1} + \frac{1}{2}\tilde{\sigma}_2^2 \tilde{S}_2^2 f_{\tilde{S}_2\tilde{S}_2} + \rho\tilde{\sigma}_1\tilde{\sigma}_2\tilde{S}_1\tilde{S}_2 f_{\tilde{S}_1\tilde{S}_2} \\ &- \frac{e^{l_1}\tilde{S}_1 f}{\tilde{S}_1 + l_1} \left(\frac{l_1 f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)}{\left(\tilde{S}_1 + l_1\right) f\left(\tilde{\tau}, q_1 - 1, q_2, \tilde{S}_1, \tilde{S}_2\right)} \right)^{\frac{l_1}{\tilde{S}_1}} \\ &- \frac{\tilde{\lambda} e^{l_2}\tilde{S}_2 f}{\tilde{S}_2 + l_2} \left(\frac{l_2 f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)}{\left(\tilde{S}_2 + l_2\right) f\left(\tilde{\tau}, q_1, q_2 - 1, \tilde{S}_1, \tilde{S}_2\right)} \right)^{\frac{l_2}{\tilde{S}_2}} = 0, \end{aligned} \tag{6.18}$$

with $f = f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)$ and

$$f\left(\tilde{\tau} = 0, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = -e^{-q_1\tilde{S}_1 - q_2\tilde{S}_2}, \tag{6.19}$$

[1]Note this can also be performed for the general N assets case

$$f\left(\tilde{\tau}, q_1 = 0, q_2 = 0, \tilde{S}_1, \tilde{S}_2\right) = -1. \tag{6.20}$$

For $q_i = 0$ and $q_j > 0$ equation (6.14) can similarly be transformed to

$$-f_{\tilde{\tau}} + \tilde{\mu}_j \tilde{S}_j f_{\tilde{S}_j} + \frac{1}{2}\tilde{\sigma}_j^2 \tilde{S}_j^2 f_{\tilde{S}_j \tilde{S}_j} - \frac{\lambda e^{l_j} \tilde{S}_j f}{\tilde{S}_j + l_j} \left(\frac{l_j f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)}{\left(\tilde{S}_j + l_j\right) f\left(\tilde{\tau}, q_i, q_j - 1, \tilde{S}_1, \tilde{S}_2\right)} \right)^{\frac{l_j}{\tilde{S}_j}} = 0, \tag{6.21}$$

with $\lambda = 1$ if $j = 1$ or $\lambda = \tilde{\lambda}$ if $j = 2$. We note here not only have we reduced the number of input parameters by two but also factored out one of the variables. When discussing results we shall now refer to $f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)$ as the value function.

6.2.2 Boundary conditions

To solve (6.18) we need boundary conditions in the asset prices, $\tilde{S}_1$ and $\tilde{S}_2$, which must be imposed along lines (rather than points given the extra dimension).

For $\tilde{S}_1 = \tilde{S}_2 = 0$ equation (6.18) reduces to

$$\frac{\partial f}{\partial \tilde{\tau}}\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = 0. \tag{6.22}$$

Using this, and (6.19), we have

$$f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1 = 0, \tilde{S}_2 = 0\right) = -1. \tag{6.23}$$

Next we look at the case of $\tilde{S}_i = 0$ while $\tilde{S}_j \neq 0$, for which (6.18) reduces to the single asset case in $\tilde{S}_j$ given by (6.21). Boundary conditions and numerical methods for (6.21) were discussed in chapter 4. Equation (6.21) has the $\tilde{S}_j = 0$ boundary condition given by (6.23). For large $\tilde{S}_j$ we use a Neumann boundary condition of the form

$$\frac{\partial f}{\partial \tilde{S}_j}\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) \to 0 \quad \text{as} \quad \tilde{S}_j \to \infty, \tag{6.24}$$

which is used for both (6.18), for all values of $\tilde{S}_i$, and hence also (6.21).

We examined using a cross-derivative boundary condition as both asset prices tended to infinity, $\tilde{S}_1, \tilde{S}_2 \to \infty$, of the form

$$\frac{\partial^2 f}{\partial \tilde{S}_1 \partial \tilde{S}_2}\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1 \to \infty, \tilde{S}_2 \to \infty\right) \to 0. \tag{6.25}$$

However this gave similar results to using a Neumann in both $\tilde{S}_1$ and $\tilde{S}_2$ for the infinite limits given the smooth, exponential-decay form of the solution as $\tilde{S}_1, \tilde{S}_2 \to \infty$. Therefore, for large $\tilde{S}_1$ and $\tilde{S}_2$ the boundary condition we use is a Neumann condition.

6.2.3 Numerical method

To solve the system (6.18) we use a finite-difference scheme using implicit differences for the derivatives in $\tilde{S}_1$ and $\tilde{S}_2$ and explicit differencing for the non-linear terms. This results in a linear system $\mathbf{Af} = \mathbf{b}$ for each combination of variables q_1, q_2 as well as each time-step, in which all the non-linear terms are on the right-hand-side, i.e. in $\mathbf{b}$. The matrix $\mathbf{A}$ is an $(n \times m) \times (n \times m)$ sparse-banded matrix, where n and m are the number of grid points in $\tilde{S}_1$ and $\tilde{S}_2$ respectively. It consists of a main diagonal of parameters, a main superdiagonal of parameters, a main subdiagonal of parameters, $m-3$ superdiagonals of zeros followed by three rows of parameters and $m-3$ subdiagonals of zeros followed by three rows of parameters. The matrix $\mathbf{A}$ is $\tilde{\tau}$ and $\{q_1, q_2\}$ independent. We exploit this by using LU decomposition, in which the matrices $\mathbf{L}$ and $\mathbf{U}$ only need to be calculated once, and, having been stored, can be used repeatedly for all $\tilde{\tau}$ and $\{q_1, q_2\}$. We further exploit the sparseness of the matrix by using a NAG (see NAG, 2014) routine *nag_dgbtrf* which takes $O\left((n \times m)^2\right)$ floating point operations which is an order of magnitude smaller than standard Gauss-Elimination which takes $O\left((n \times m)^3\right)$ floating point iterations. The NAG routine is also storage efficient as it neglects the zero subdiagonals below the $m^{th}+1$ subdiagonal, as these remain zero after LU decomposition. After using *nag_dgbtrf* to calculate $\mathbf{A} = \mathbf{PLU}$, the NAG routine *nag_dgbtrs* is then used to solve the system $\mathbf{Af} = \mathbf{b}$ which solves $\mathbf{PLY} = \mathbf{B}$ and then $\mathbf{Uf} = \mathbf{Y}$. For further reading on LU decomposition the reader is referred to Smith (1965). We expect this routine to be $O\left(\Delta\tilde{\tau}, \Delta\tilde{S}_1^2, \Delta\tilde{S}_2^2\right)$ convergent which was found to be the case, as can be seen in table 6.1. We confirmed our numerics by solving this problem using the successive-over-relaxation (SOR) method, as discussed in section 2.6. However, implicit differences with explicit non-linear terms was the scheme we used primarily.

6.2.4 Numerical results

In the case of $q_1 > 0, q_2 = 0$ or $q_1 = 0, q_2 > 0$ equation (6.18) simplifies down to the single asset problem (6.21) in $q_1, \tilde{S}_1$ and $q_2, \tilde{S}_2$ respectively. It is the case, which is confirmed numerically, that $\partial/\partial\tilde{S}_i \equiv 0$ when $q_i = 0$. Therefore the results of chapter 4 then hold. Thus we shall numerically examine $q_1, q_2 > 0$ which is governed by (6.18).

Table 6.1: Convergence of finite-difference scheme for multiple underlyings

$n = m$	k	f_{LU}	ratio	f_{SOR}	ratio
51	1001	-0.0002609054341644	NA	-0.0002609054342111	NA
101	1001	-0.0002593726011118	NA	-0.0002593726011122	NA
201	1001	-0.0002589897564976	4.00	-0.0002589897564976	4.00
401	1001	-0.0002588935201294	3.98	-0.0002588935201297	3.98
801	1001	-0.0002588691542179	3.95	-0.0002588691542196	3.95
401	251	-0.0002592484453679	NA	-0.0002592484453682	NA
401	501	-0.0002590119089189	NA	-0.0002590119089192	NA
401	1001	-0.0002588935201294	2.00	-0.0002588935201297	2.00
401	2001	-0.0002588342956051	2.00	-0.0002588342956053	2.00
401	4001	-0.0002588046758120	2.00	-0.0002588046758121	2.00

$n = m$ and k are the number of space and time points respectively. f_{LU} and f_{SOR} are the solution of (6.18) at $\tilde{\tau} = 1, q_1 = 1, q_2 = 1, \tilde{S}_1 = 5, \tilde{S}_2 = 5$. We have set $\tilde{S}_{1,max} = \tilde{S}_{2,max} = 10, \tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4, \rho = 0.5, \lambda = 1$ and $l_1 = l_2 = 10$. The ratio is the ratio of the change in errors for successive grids.

We shall first examine the effect of the correlation before examining the case of asymmetric parameters. As discussed in chapter 5, from Modern Portfolio Theory (see Markowitz, 1959) we expect a trader with negatively correlated assets to want a balance of assets in his portfolio, such that $q_1 = q_2$, while for positively correlated assets the trader would seek an unbalanced portfolio. Figure 6.1 shows the value function for various values of the correlation parameter, ρ, with otherwise constant parameters, for $q_1 = q_2 = 1$ which is a slice of the main diagonal, corresponding to $\tilde{S}_1 = \tilde{S}_2$. It can be seen how the value is monotonically decreasing as the correlation, ρ, increases, indicating a preference for a negatively correlated portfolio.

We will now shift our attention to the optimal trading strategy for the various regimes of the correlation parameter, ρ. We investigate the optimal trading strategies rather than the value function as these are simply transformations of the value function and, from a financial perspective, are more transparent than examining the value function per se. This can be quite difficult to show graphically given we have $\delta_1^*\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)$ and $\delta_2^*\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)$, which is a function of five variables, with q_1 and q_2 taking discrete values. To investigate the optimal trading strategies we will take a slice of the optimal trading strategies δ_1^*, δ_2^* for a single value of time-to-expiry, $\tilde{\tau}$, the price of asset one, $\tilde{S}_1$, and the price of asset two, $\tilde{S}_2$, with $\tilde{\tau} = 1$ and $\tilde{S}_1 = \tilde{S}_2 = 6$. We shall then plot the optimal strategy for asset one, δ_1^*, and the optimal strategy for

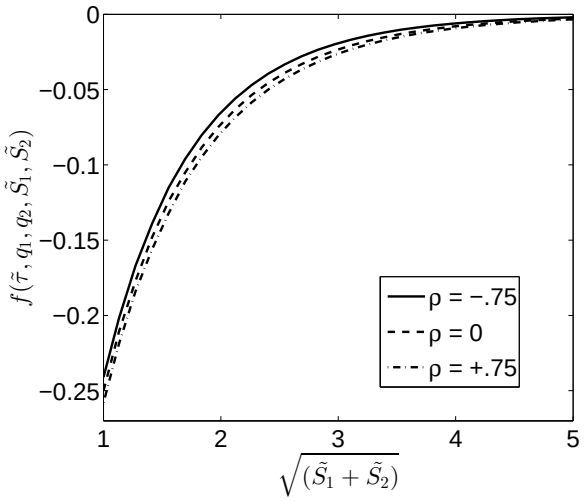
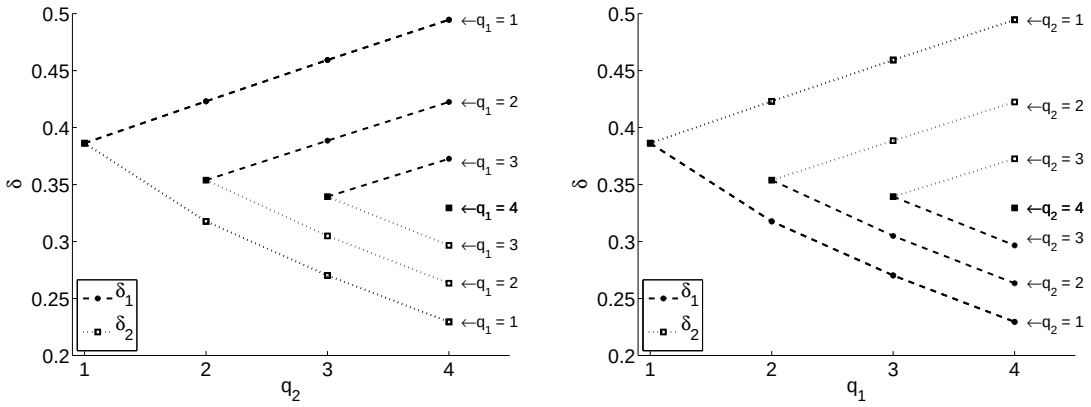


Figure 6.1: This is a plot of the value function for various values of ρ, to which we can see the value function is a decreasing function for increasing ρ. The parameters are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.03, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.6, l_1 = l_2 = 10, \tilde{\lambda} = 1$ and varying ρ.

asset two, δ_2^*, on the vertical axis, with the number of asset two we have remaining, q_2, on the horizontal axis, for various values of inventory for asset one, q_1, with $q_1 \geq q_2$ for ease of graphically interpretation. In a corresponding figure, we also plot the optimal strategy of asset one, δ_1^*, and asset two, δ_2^*, on the vertical axis, with the inventory for asset one, q_1, on the horizontal axis, for various values of q_2, with $q_2 \geq q_1$. Although we are plotting discrete points for improved presentation we connect these with lines to show lines of constant q_1 for the former case and q_2 for the latter case, as can be seen in figure 6.2.

We begin by investigating the case of negative ρ, in which the optimal strategies can be seen in figure 6.2. In figure 6.2(a) the point at which the lines of δ_1^* and δ_2^* meet corresponds to $q_1 = q_2$. For each point corresponding to $q_1 = q_2$ we can see that the adjacent line corresponding to the values of the optimal trading strategy for asset one, δ_1^*, increase for increasing inventory of asset two, q_2, while the lines corresponding to the optimal trading strategy of asset two, δ_2^*, decrease for increasing inventory of asset two, q_2. This is due to having more of asset two than asset one as when this is the case the trader seeks to sell asset two faster than asset one so as to balance the risk in his portfolio, and hence $\delta_2^* > \delta_1^*$. The same pattern, but with δ_1^* and δ_2^* switched, due to the symmetry, can be seen in figure 6.2(b).

For positively correlated assets, $\rho > 0$, the trading strategies take lower values than



(a) δ_1^*, δ_2^* against q_2 (b) δ_1^*, δ_2^* against q_1

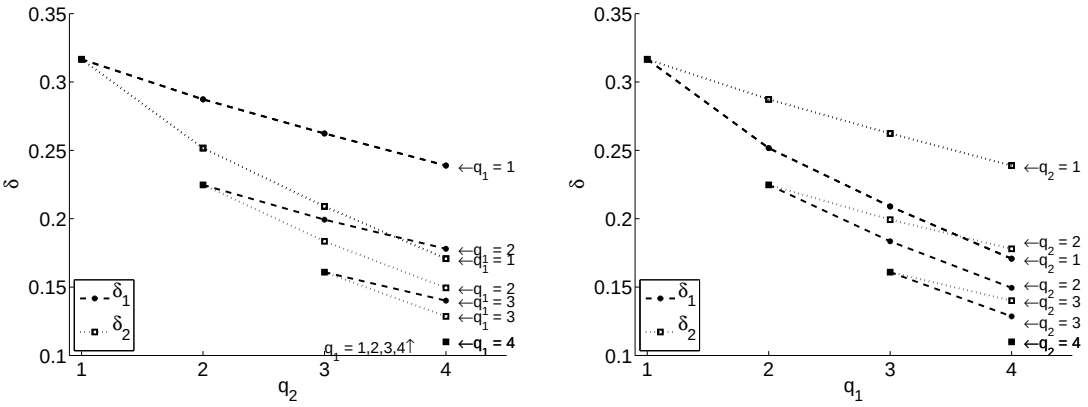
Figure 6.2: Optimal trading strategy δ_1^*, δ_2^* against q_2 (figure 6.2(a)) and q_1 (figure 6.2(b)) at $\tilde{\tau} = 1, \tilde{S}_1 = \tilde{S}_2 = 6$ for negative correlation ($\rho = -0.75$). The parameters used are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.03, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.1, l_1 = l_2 = 10, \tilde{\lambda} = 1$ and $\rho = -0.75$.

that seen in figure 6.2 as the trader would like to sell quicker to de-risk his portfolio. Positive ρ increases the risk for a risk-averse trader and as such the more assets he holds the more he would like to sell them, thus he decreases his asking price. This can be seen in figure 6.3. Having an unbalanced portfolio is also evident in these results. Looking at figure 6.3(a) we can see that the optimal trading strategy of asset two, δ_2^*, is much lower for $q_1 = q_2 = 2$ than it is for $q_1 = 1, q_2 = 2$. This indicates the trader would rather hold an unbalanced portfolio as it is not as risky. The same can be seen for other values of the inventory, q_1, q_2, as well as for the optimal strategy in asset one, δ_1^* (in figure 6.3(b)).

In chapter 5, through analytic and numerical verification, we found the trading strategies of one asset was independent of the other asset for zero correlation, $\rho = 0$. Although we cannot perform the same analytic analysis in this framework, numerical computations indicated conclusively that the same property is present.

We shall now discuss varying parameters such that the symmetry is lost. We shall do this by modifying each parameter (or parameter pair) individually such that we see its specific impact. The base case for comparison is seen in figure 6.2.

If the ratio of intensities at the best ask, $\tilde{\lambda}$, is increased the values of the optimal strategy of asset two, δ_2^*, increase, while that of asset one, δ_1^*, decrease. Increasing $\tilde{\lambda}$ implies the rate at which asset two is sold, assuming all else equal, is a multiple of that of asset one. This can be seen in figure 6.4, in which the intensity of asset two



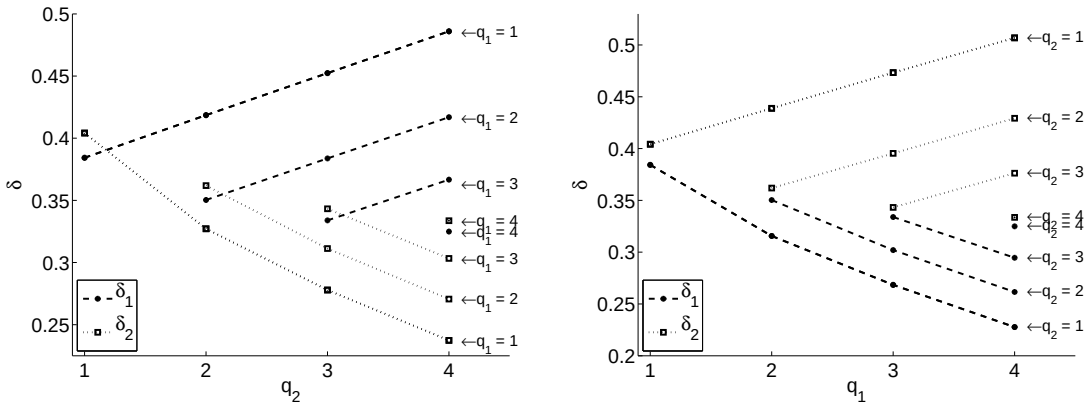
(a) δ_1^*, δ_2^* against q_2 (b) δ_1^*, δ_2^* against q_1

Figure 6.3: Optimal trading strategy δ_1^*, δ_2^* against q_2 (figure 6.3(a)) and q_1 (figure 6.3(b)) at $\tilde{\tau} = 1, \tilde{S}_1 = \tilde{S}_2 = 6$ for positive correlation ($\rho = 0.75$). The parameters used are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.03, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.1, l_1 = l_2 = 10, \tilde{\lambda} = 1$ and $\rho = 0.75$.

has been doubled so $\tilde{\lambda} = 2$.

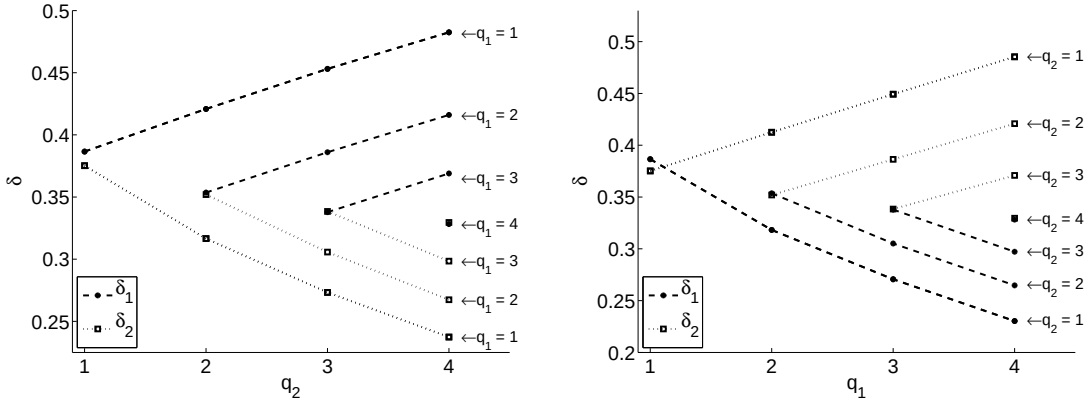
When the exponential decay of the limit-order book, l_i, is decreased the optimal trading strategy of asset i, δ_i^*, increases when the asking price is above par value and decreases when the asking price is below par value. When selling above par a lower l_i signifies that the trader can place his asset further into the limit-order book without significantly reducing his probability of sale. When the trader is eager to sell, he will sell the asset at a discount. In order to do so under lower l_i the trader must sell at a higher discount in order to increase his probability of sale by a significant amount. The opposite is true for larger l_i. Figure 6.5 shows the case of altering the exponential-decay parameter l_2 such that $l_2 > l_1$ with $l_2 = 20$ and $l_1 = 10$. Given the optimal strategies are positive, it can be seen that the optimal strategy of asset two, δ_2^*, has increased. This also leads to a decreases in the optimal strategy of asset one, δ_1^*.

Figure 6.6 shows the case of increasing the drift in one asset while the rest of the parameters remain constant, such that $\tilde{\mu}_2 > \tilde{\mu}_1$ with $\tilde{\mu}_2 = 0.1$ and $\tilde{\mu}_1 = 0.03$. We see in figure 6.6(a) and 6.6(b) that the values of the optimal strategy of asset two, δ_2^*, and asset one, δ_1^*, are greater in comparison to the case of equal drifts, $\tilde{\mu}_1 = \tilde{\mu}_2$, as seen in figure 6.2. This is expected given an asset with a higher expected growth is more valuable than an asset with lower expected growth. The trader is thus not as keen to selling in comparison to the asset with lower growth as he wants to relinquish the higher growth prospect of the asset while he still holds it.



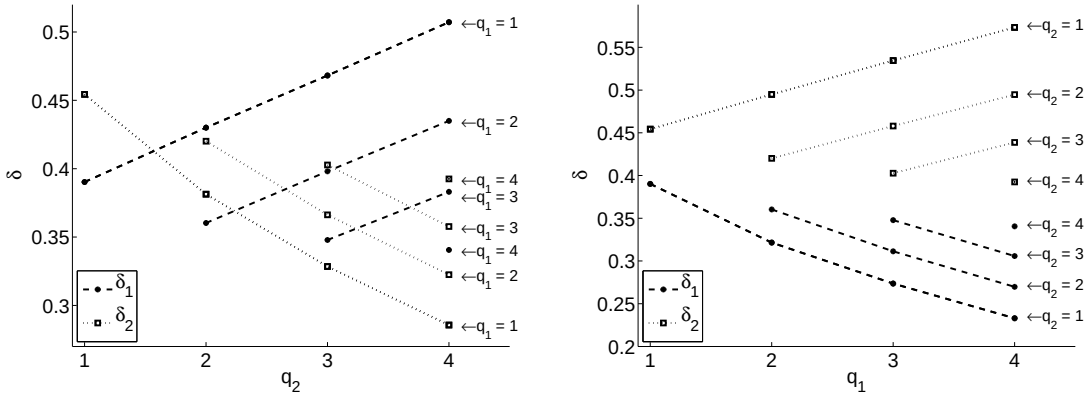
(a) δ_1^*, δ_2^* against q_2 (b) δ_1^*, δ_2^* against q_1

Figure 6.4: Optimal trading strategy δ_1^*, δ_2^* against q_2 (figure 6.4(a)) and q_1 (figure 6.4(b)) at $\tilde{\tau} = 1, \tilde{S}_1 = \tilde{S}_2 = 6$ when one asset has a higher base intensity than the other ($\tilde{\lambda} = 2$). The parameters used are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.03, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.1, l_1 = l_2 = 10, \tilde{\lambda} = 2$ and $\rho = -0.75$.



(a) δ_1^*, δ_2^* against q_2 (b) δ_1^*, δ_2^* against q_1

Figure 6.5: Optimal trading strategy δ_1^*, δ_2^* against q_2 (figure 6.5(a)) and q_1 (figure 6.5(b)) at $\tilde{\tau} = 1, \tilde{S}_1 = \tilde{S}_2 = 6$ when one of the asset's limit-order book has a higher exponential-decay than the other ($l_2 > l_1$). The parameters used are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.03, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.1, l_1 = 10, l_2 = 20, \tilde{\lambda} = 1$ and $\rho = -0.75$.



(a) δ_1^*, δ_2^* against q_2 (b) δ_1^*, δ_2^* against q_1

Figure 6.6: Optimal trading strategy δ_1^*, δ_2^* against q_2 (figure 6.6(a)) and q_1 (figure 6.6(b)) at $\tilde{\tau} = 1, \tilde{S}_1 = \tilde{S}_2 = 6$ when one asset has a higher drift rate than the other ($\mu_2 > \mu_1$). The parameters used are $\tilde{\mu}_1 = 0.03, \tilde{\mu}_2 = 0.1, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.1, l_1 = l_2 = 10, \tilde{\lambda} = 1$ and $\rho = -0.75$.

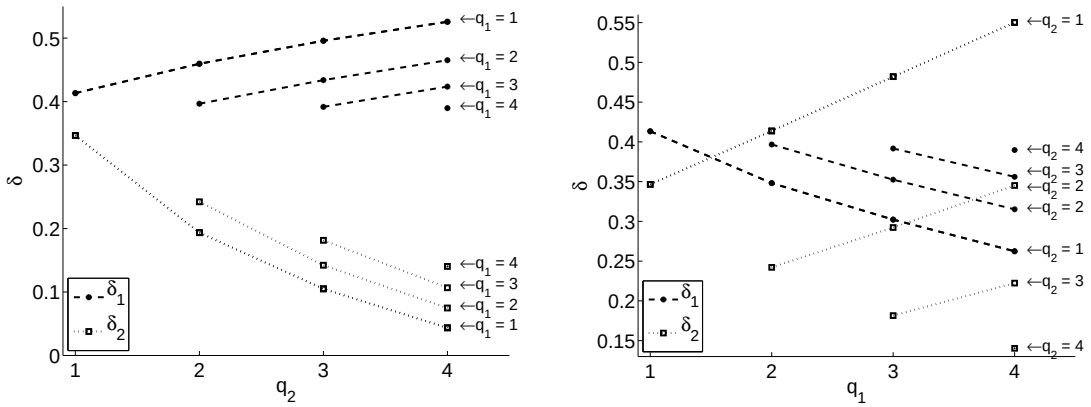
Unlike the drift, if the volatility, $\tilde{\sigma}_i$, of asset i is increased it will decrease the asking price of that asset. An asset with a higher volatility is more riskier than an asset with lower volatility. The trader thus wants to sell the riskier assets faster than the less risky assets. Figure 6.7 shows the case of increasing the volatility in asset two, $\tilde{\sigma}_2$, while keeping the volatility in asset one, $\tilde{\sigma}_1$, the same as previous, such that $\tilde{\sigma}_2 > \tilde{\sigma}_1$ with $\tilde{\sigma}_2 = 0.2$ and $\tilde{\sigma}_1 = 0.1$. We see in figure 6.7(a) and 6.7(b) that the values of the optimal strategy of asset one, δ_1^*, are greater while that of asset two, δ_2^*, are lower, in comparison to the $\tilde{\sigma}_1 = \tilde{\sigma}_2$ case as seen in figure 6.2.

Similar to the single asset case explored in chapter 4, there is interesting behaviour in the small-time-to-termination regime and thus we examine this next.

6.3 Small-time-to-termination solution

Using asymptotic analysis we can find a fully analytic solution which is valid in the regime as $\tilde{\tau} \to 0$. This could be especially relevant for a high-frequency trading framework given the time remaining would be small and thus analytic expressions would be preferred over numerical solutions.

To examine the small-time-to-termination solution we anticipate a perturbation



(a) δ_1^*, δ_2^* against q_2 (b) δ_1^*, δ_2^* against q_1

Figure 6.7: Optimal trading strategy δ_1^*, δ_2^* against q_2 (figure 6.7(a)) and q_1 (figure 6.7(b)) at $\tilde{\tau} = 1, \tilde{S}_1 = \tilde{S}_2 = 6$ when one asset has a higher volatility than the other ($\sigma_2 > \sigma_1$). The parameters used are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.03, \tilde{\sigma}_1 = 0.1, \tilde{\sigma}_2 = 0.2, l_1 = l_2 = 10, \tilde{\lambda} = 1$ and $\rho = -0.75$.

series about the parameter $\tilde{\tau}$, for $\tilde{\tau} \ll 1$, of the form

$$f\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = f_0\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) + \tilde{\tau} f_1\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) + \tilde{\tau}^2 f_2\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) + O\left(\tilde{\tau}^3\right), \tag{6.26}$$

which is an extension to the perturbation series we used in chapter 4 when examining trading with a single underlying (see (4.55)), which gave very satisfactory results. The $O\left(\tilde{\tau}^0\right)$ term, $f_0\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)$, is merely equal to the terminal condition given by (6.19). By substituting (6.26) into (6.18) we can find an analytic solution for $f_1\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)$ by collecting the $O\left(\tilde{\tau}^0\right)$ terms

$$f_1\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = \Bigg(-\sum_{i=1}^{2} \tilde{\mu}_i \tilde{S}_i q_i + \frac{1}{2} \sum_{i=1}^{2} \tilde{\sigma}_i^2 \tilde{S}_i^2 q_i^2 + \rho \tilde{\sigma}_1 \tilde{\sigma}_2 \tilde{S}_1 \tilde{S}_2 q_1 q_2 - \sum_{i=1}^{2} \frac{\tilde{S}_i}{\tilde{S}_i + l_i} \left(\frac{l_i}{\tilde{S}_i + l_i}\right)^{\frac{l_i}{\tilde{S}_i}} \Bigg) f_0\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right), \tag{6.27}$$

and from the $O\left(\tilde{\tau}^1\right)$ terms

$$f_2\left(q_1,q_2,\tilde{S}_1,\tilde{S}_2\right)=\frac{1}{2}\Bigg(\sum_{i-1}^{2}\tilde{\mu}_i\tilde{S}_i f_{1\tilde{S}_i}+\frac{1}{2}\sum_{i=1}^{2}\tilde{\sigma}_i^2\tilde{S}_i^2 f_{1\tilde{S}_i\tilde{S}_i}+\rho\tilde{\sigma}_1\tilde{\sigma}_2\tilde{S}_1\tilde{S}_2 f_{1\tilde{S}_1\tilde{S}_2}$$
$$-\frac{1}{\tilde{S}_1+l_1}\left(\frac{l_1}{\tilde{S}_1+l_1}\right)^{\frac{l_1}{\tilde{S}_1}}\left(\left(l_1+\tilde{S}_1\right)f_1\left(q_1,q_2,\tilde{S}_1,\tilde{S}_2\right)-l_1 f_1\left(q_1-1,q_2,\tilde{S}_1,\tilde{S}_2\right)e^{-\tilde{S}_1}\right)$$
$$-\frac{1}{\tilde{S}_2+l_2}\left(\frac{l_2}{\tilde{S}_2+l_2}\right)^{\frac{l_2}{\tilde{S}_2}}\left(\left(l_2+\tilde{S}_2\right)f_1\left(q_1,q_2,\tilde{S}_1,\tilde{S}_2\right)-l_2 f_1\left(q_1,q_2-1,\tilde{S}_1,\tilde{S}_2\right)e^{-\tilde{S}_2}\right)\Bigg), \tag{6.28}$$

with

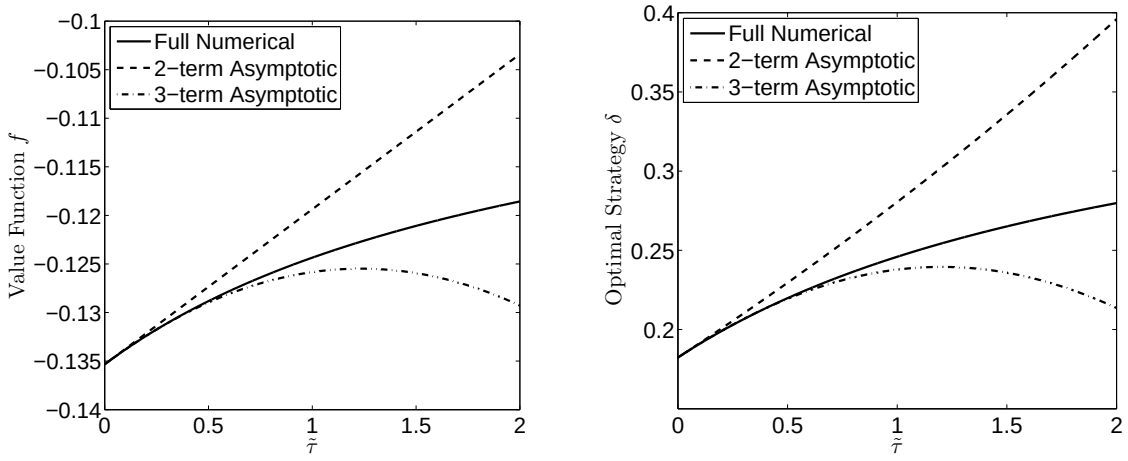
$$f_1\left(q_1=0,q_2,\tilde{S}_1,\tilde{S}_2\right)=f_1\left(q_1,q_2=0,\tilde{S}_1,\tilde{S}_2\right)=0.$$

In (6.28) $f_{1\tilde{S}_i}$ and $f_{1\tilde{S}_i\tilde{S}_j}$ are the first and second derivative of f_1, which as we see from (6.27), can be calculated analytically and as such we have a fully analytic expression for (6.26). We neglect the analytic solutions of the derivatives due to the complexity of the equations but they are simple extensions to (4.58) and (4.59) which were found in chapter 4 for the case of a single underlying asset. We could include higher-order terms, although this would become increasingly complex. As analogous to the PDE, if $q_i-1=0$ the corresponding f_1 term in (6.28) is zero.

The approximation (6.26) is quite accurate even for $\tilde{\tau}$ of $O\left(1\right)$ for a range of parameter values, an example of which can be seen in figure 6.8. This figure shows the full-numerical solution compared to the two-term and three-term asymptotic approximation, for both the value function (figure 6.8(a)) and optimal trading strategy (figure 6.8(b)). As in the single asset case, for which the reader is referred to section 4.2, there are restrictions to how well the expression holds depending on the parameter values used. For higher values of $\tilde{S}_i$ and l_i the solution of $f\left(\tilde{\tau},q_1,q_2,\tilde{S}_1\tilde{S}_2\right)$ diverges faster from the small-time-to-termination solution due to the stronger presence of the non-linear terms.

6.4 Perpetual problem

Numerical simulations positively indicated the possibility that perpetual (time-independent) solutions exist, as expected given the analysis of a single underlying in chapter 4. It



(a) Value function (b) Optimal strategy

Figure 6.8: The two above figures show the small-time approximation (using a first-order and second-order expansion in time) against the solution obtained using the full-numerical scheme. This can be seen for both the value function (figure 6.8(a)) and optimal trading strategy (figure 6.8(b)). The parameter values used are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4, \rho = -0.4, l_1 = l_2 = 5$ and $\lambda = 1$, at a point along the main diagonal of the solution, with $q_1 = q_2 = 1$ and $\tilde{S}_1 = \tilde{S}_2 = 1$.

is thus of interest to examine these for the multi-dimensional case, by requiring

$$\frac{\partial f}{\partial \tilde{\tau}}\left(\tilde{\tau}, q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = 0 \quad \text{as} \quad \tilde{\tau} \to \infty.$$

In making this assumption, we are implicitly assuming that the terminal time in which the assets must be sold is sufficiently distant for us to neglect the time variation.

6.4.1 Problem formulation for steady-state

For non-zero $\{q_1, q_2\}$ the steady-state PDE we examine is

$$\begin{aligned}
&\tilde{\mu}_1 \tilde{S}_1 f_{\tilde{S}_1} + \tilde{\mu}_2 \tilde{S}_2 f_{\tilde{S}_2} + \frac{1}{2}\tilde{\sigma}_1^2 \tilde{S}_1^2 f_{\tilde{S}_1\tilde{S}_1} + \frac{1}{2}\tilde{\sigma}_2^2 \tilde{S}_2^2 f_{\tilde{S}_2\tilde{S}_2} + \rho\tilde{\sigma}_1\tilde{\sigma}_2\tilde{S}_1\tilde{S}_2 f_{\tilde{S}_1\tilde{S}_2} \\
&- \frac{e^{l_1}\tilde{S}_1 f}{\tilde{S}_1 + l_1}\left(\frac{l_1 f\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)}{\left(\tilde{S}_1 + l_1\right) f\left(q_1 - 1, q_2, \tilde{S}_1, \tilde{S}_2\right)}\right)^{\frac{l_1}{\tilde{S}_1}} \\
&- \frac{\tilde{\lambda} e^{l_2}\tilde{S}_2 f}{\tilde{S}_2 + l_2}\left(\frac{l_2 f\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right)}{\left(\tilde{S}_2 + l_2\right) f\left(q_1, q_2 - 1, \tilde{S}_1, \tilde{S}_2\right)}\right)^{\frac{l_2}{\tilde{S}_2}} = 0,
\end{aligned} \tag{6.29}$$

with

$$f\left(q_1 = 0, q_2 = 0, \tilde{S}_1, \tilde{S}_2\right) = -1. \tag{6.30}$$

We use the same boundary conditions for the perpetual problem as discussed in section 6.2.2 for the non-steady-state problem, these being (6.21) (with time-derivative equal to zero), (6.23) and (6.24). When the inventory for asset i is zero, $q_i = 0$, the derivatives with respect to the asset price of asset i, $\tilde{S}_i$, are zero everywhere as $\tilde{S}_i$ has no impact on the solution. As such the problem reduces to the single asset case which was examined thoroughly in chapter 4. Thus we an interested in the case of $q_1, q_2 > 0$.

Numerical method

To solve the steady-state problem we use the Newton-iteration method described for an ODE in section 2.6.3, but now for a PDE, while using the same NAG routine as described in section 6.2.3. Consider the non-linear PDE for fixed $\{q_1, q_2\}$

$$G\left(S_1, S_2, f, f_{\tilde{S}_1}, f_{\tilde{S}_2}, f_{\tilde{S}_1\tilde{S}_1}, f_{\tilde{S}_2\tilde{S}_2}, f_{\tilde{S}_1\tilde{S}_2}\right) = 0, \tag{6.31}$$

for value function f which has linear boundary conditions. Using central differences for the derivatives, we can solve the algebraic system by iteratively solving

$$\mathbf{J}\left(\mathbf{f}^{(k)}\right)\boldsymbol{\varepsilon}^{(k)} = -\mathbf{G}\left(\mathbf{f}\right), \tag{6.32}$$

with

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \boldsymbol{\varepsilon}^{(k)}, \tag{6.33}$$

where $\mathbf{f}^{(k)}$ and $\boldsymbol{\varepsilon}^{(k)}$ are the value function and correction at iteration k respectively. Here we have taken a first-order approximation of $\boldsymbol{\varepsilon}$, as detailed in section 2.6.3. The Jacobi matrix of $\mathbf{G}$

$$\mathbf{J}\left(\mathbf{f}\right) = \frac{\partial \mathbf{G}}{\partial \mathbf{f}}\left(\mathbf{f}\right) = \left(\frac{\partial G_i}{\partial f_j}\left(\mathbf{f}\right)\right)_{0,\leq i,j\leq N}, \tag{6.34}$$

has the same banded structure as the matrix $\mathbf{A}$ discussed in section 6.2.3. We therefore have a similar system to $\mathbf{Af} = \mathbf{b}$ and as such we use the same NAG routine, *nag_dgbtrf* to calculate the LU decomposition before using *nag_dgbtrs* to solve the system.

Unlike the non-steady-state case, which has a matrix $\mathbf{A}$ that is $\tilde{\tau}$ and q_i independent for all i, the Jacobian matrix $\mathbf{J}$ is a function of $\mathbf{f}$ given the non-linear structure of the PDE and as such is updated at each iteration. Therefore we must calculate $\mathbf{A} = \mathbf{PLU}$ using the NAG routine *nag_dgbtrf* at each iteration which can be quite computationally expensive for fine grids.

Numerical results

Many of the results for the non-steady-state case are quantitatively similar to the perpetual case and thus we will not discuss them in detail. We will however examine two solutions which are included for use of reference later in this chapter.

Figure 6.9(a) shows contours of constant value of the value function. The symmetry in the asset prices of asset one, $\tilde{S}_1$, and asset two, $\tilde{S}_2$, is clear to see, and is on account of the symmetry of the chosen parameters. A second example presented in figure 6.9(b) shows an analogous contour plot to that of figure 6.9(a) but with the asset price, $\tilde{S}_1, \tilde{S}_2$, symmetry broken by an asymmetric parameter choice.

It was found in chapter 4 that there was quite interesting behaviour around the small $\tilde{S}$ region for the perpetual problem. This is also (unsurprisingly) the case for the multiple underlying problem. For zero inventory in asset j, $q_j = 0$, the system is reduced to the single asset case in asset i and thus the same singular behaviour as found in chapter 4 is present. It is also of interest to examine the case of positive inventory for both assets, $q_1, q_2 > 0$. We shall therefore next examine the small asset price, $\tilde{S}_1, \tilde{S}_2$, solutions.

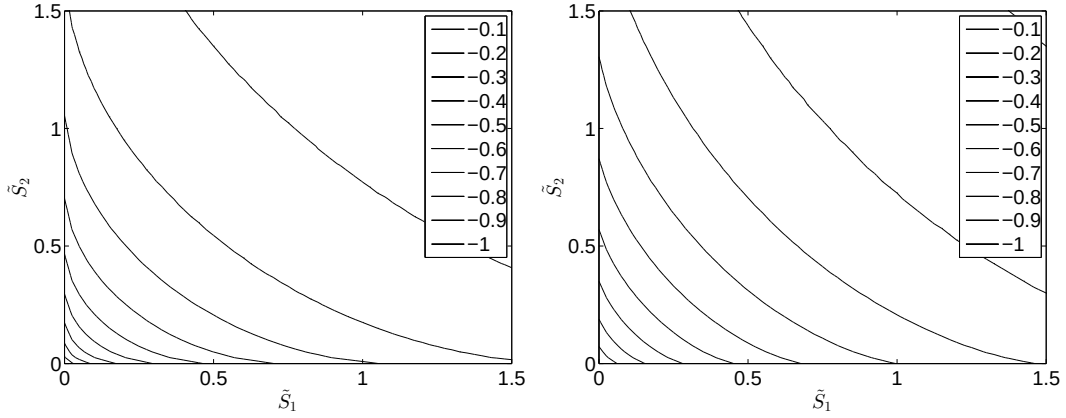
6.4.2 Asymptotic analysis of the asset prices

Since the PDE (6.29) appears to have no analytic solutions, we find it useful to use the complementary approach of asymptotic and numerical methods. Examining figures 6.9(a) and 6.9(b) suggests the use of polar coordinates will be advantageous in gaining analytic insight. A similar examination is performed by Duck et al. (2014) for the case of American options with multiple underlyings in which converting to polars proved quite insightful and more efficient.

To transform to polar conditions we assign

$$\tilde{S}_1 = R\cos\theta, \qquad \tilde{S}_2 = R\sin\theta, \tag{6.35}$$

with $R = \left(\tilde{S}_1^2 + \tilde{S}_2^2\right)^{\frac{1}{2}}$. Under this transformation $f\left(q_1, q_2, \tilde{S}_1, \tilde{S}_2\right) = f\left(q_1, q_2, R, \theta\right)$



(a) Symmetric case (b) Asymmetric case

Figure 6.9: Contours of constant value for the $f\left(q_1 = 1, q_2 = 1, \tilde{S}_1, \tilde{S}_2\right)$ for the symmetric case ($\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4, l_1 = l_2 = 1, \tilde{\lambda} = 1$ and $\rho = -0.4$) and asymmetric case ($\tilde{\mu}_1 = 0.03, \tilde{\mu}_2 = 0.06, \tilde{\sigma}_1 = 0.5, \tilde{\sigma}_2 = 0.7, l_1 = 1, l_2 = 2, \tilde{\lambda} = 2$ and $\rho = -0.4$)

and as such (6.29) is transformed to

$$
\begin{aligned}
&R^2\left(\frac{1}{2}\tilde{\sigma}_1^2\cos^4\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\cos^2\theta\sin^2\theta + \frac{1}{2}\tilde{\sigma}_2^2\sin^4\theta\right) f_{RR} \\
&+ R\cos\theta\sin\theta\left(-\tilde{\sigma}_1^2\cos^2\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) + \tilde{\sigma}_2^2\sin^2\theta\right) f_{R\theta} \\
&+ \cos^2\theta\sin^2\theta\left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\sigma_2^2\right) f_{\theta\theta} \\
&+ R\left(\cos^2\theta\sin^2\theta\left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\sigma_2^2\right) + \tilde{\mu}_1\cos^2\theta + \tilde{\mu}_2\sin^2\theta\right) f_R \\
&+ \cos\theta\sin\theta\left(\tilde{\sigma}_1^2\cos^2\theta - \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) - \tilde{\sigma}_2^2\sin^2\theta - \tilde{\mu}_1 + \tilde{\mu}_2\right) f_\theta \\
&- \frac{e^{l_1} f R\cos\theta}{R\cos\theta + l_1}\left(\frac{l_1 f\left(q_1, q_2, R, \theta\right)}{\left(R\cos\theta + l_1\right) f\left(q_1 - 1, q_2, R, \theta\right)}\right)^{\frac{l_1}{R\cos\theta}} \\
&- \frac{\tilde{\lambda} e^{l_2} f R\sin\theta}{R\sin\theta + l_2}\left(\frac{l_2 f\left(q_1, q_2, R, \theta\right)}{\left(R\sin\theta + l_2\right) f\left(q_1, q_2 - 1, R, \theta\right)}\right)^{\frac{l_2}{R\sin\theta}} = 0,
\end{aligned}
\tag{6.36}
$$

with

$$
f\left(q_1 = 0, q_2 = 0, R, \theta\right) = -1. \tag{6.37}
$$

When $\theta = 0$ the PDE reverts to the single asset case in q_1 and $\tilde{S}_1$. Similarly when $\theta = \frac{\pi}{2}$ the PDE reverts to the single asset case in q_2 and $\tilde{S}_2$. We are therefore interested in the behaviour of $\theta \neq \{0, \frac{\pi}{2}\}$, as well as the asymptotic properties as $R \to 0$. To solve numerically we need to impose boundary conditions as $R \to 0, \infty$. As $R \to 0, \infty$ we have from the boundary conditions in Cartesian coordinates (6.23) and (6.24) (with

time-derivative equal to zero)

$$f\left(q_1, q_2, R = 0, \theta\right) = -1, \tag{6.38}$$

and

$$\frac{\partial f}{\partial R}\left(q_1, q_2, R \to \infty, \theta\right) \to 0. \tag{6.39}$$

In the case of the single asset problem we found quite interesting behaviour around the small asset price, $\tilde{S} \to 0$, regime, as discussed in section 4.3. We will thus begin by examining this region for the multiple underlying case, first in polar coordinates, and then transforming back to Cartesian coordinates in $\tilde{S}_1$ and $\tilde{S}_2$ space to confirm our asymptotic approximations.

The limit of $R \to 0$

For the single asset case in asset j, i.e. (6.21) with the time-derivative equal to zero, we found in the limit of the small asset price, $\tilde{S}_j \to 0$, the non-linear term was negligible and the solution behaved like

$$f\left(q, \tilde{S}_j\right) \approx -1 + c_j(q)\tilde{S}_j^{\beta}, \tag{6.40}$$

in which

$$\beta = 1 - \frac{2\tilde{\mu}_j}{\tilde{\sigma}_j^2} > 0 \Rightarrow \tilde{\mu} < \frac{\tilde{\sigma}^2}{2}, \tag{6.41}$$

with the value of $c_j\left(q\right)$ found numerically, see section 4.3.1. Asymptotic balancing for the multiple asset case, confirmed with numerical computations, indicates a similar behaviour in the limit as $R \to 0$ and thus (6.36) is reduced to a linear PDE of the form

$$\begin{aligned}
&R^2\left(\frac{1}{2}\tilde{\sigma}_1^2\cos^4\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\cos^2\theta\sin^2\theta + \frac{1}{2}\tilde{\sigma}_2^2\sin^4\theta\right) f_{RR} \\
&+ R\cos\theta\sin\theta\left(-\tilde{\sigma}_1^2\cos^2\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) + \tilde{\sigma}_2^2\sin^2\theta\right) f_{R\theta} \\
&+ \cos^2\theta\sin^2\theta\left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\sigma_2^2\right) f_{\theta\theta} \\
&+ R\left(\cos^2\theta\sin^2\theta\left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\sigma_2^2\right) + \tilde{\mu}_1\cos^2\theta + \tilde{\mu}_2\sin^2\theta\right) f_R \\
&+ \cos\theta\sin\theta\left(\tilde{\sigma}_1^2\cos^2\theta - \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) - \tilde{\sigma}_2^2\sin^2\theta - \tilde{\mu}_1 + \tilde{\mu}_2\right) f_{\theta} = 0.
\end{aligned} \tag{6.42}$$

In the limit as $R \to 0$ it is found that (6.42) emits two families of solutions, both a homogeneous family of solutions and an inhomogeneous family of solutions, both of which we will investigate.

Inhomogeneous family of solutions

We shall consider first the inhomogeneous family of solutions, both for the symmetric and asymmetric scenarios. For asymmetric parameters, asymptotic balancing suggests that (6.42) admits solutions of the form:

$$f = -1 + R^{\hat{\alpha}_1} \hat{f}_1(\theta) + R^{\hat{\alpha}_2} \hat{f}_2(\theta), \tag{6.43}$$

assuming $\hat{\alpha}_1, \hat{\alpha}_2 > 0$. This results in the linear family of ODEs

$$\begin{aligned}
&\cos^2\theta \sin^2\theta \left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\sigma_2^2\right)\hat{f}_{i\theta\theta} \\
&+ \Bigg(\cos\theta \sin\theta \left(\tilde{\sigma}_1^2 \cos^2\theta - \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) - \tilde{\sigma}_2^2 \sin^2\theta - \tilde{\mu}_1 + \tilde{\mu}_2\right) \\
&+ \hat{\alpha}_i \cos\theta \sin\theta \left(-\tilde{\sigma}_1^2 \cos^2\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) + \tilde{\sigma}_2^2 \sin^2\theta\right) \Bigg)\hat{f}_{i\theta} \\
&+ \Bigg(\hat{\alpha}_i\left(\hat{\alpha}_i - 1\right)\left(\frac{1}{2}\tilde{\sigma}_1^2 \cos^4\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2 \cos^2\theta \sin^2\theta + \frac{1}{2}\tilde{\sigma}_2^2 \sin^4\theta\right) \\
&+ \hat{\alpha}_i \left(\cos^2\theta \sin^2\theta \left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\tilde{\sigma}_2^2\right) + \tilde{\mu}_1 \cos^2\theta + \tilde{\mu}_2 \sin^2\theta\right)\Bigg)\hat{f}_i = 0,
\end{aligned} \tag{6.44}$$

for $i \in \{1, 2\}$. The next step is to consider the limit of $\hat{f}_1$ and $\hat{f}_2$ as $\theta \to 0$ and $\theta \to \frac{\pi}{2}$.

Consider first the solutions of $\hat{f}_1$, examining (6.44) at $\theta = 0$ we see it reduces to

$$\left(\hat{\alpha}_1\left(\hat{\alpha}_1 - 1\right)\frac{1}{2}\tilde{\sigma}_1^2 + \hat{\alpha}_1\tilde{\mu}_1\right)\hat{f}_1 = 0, \tag{6.45}$$

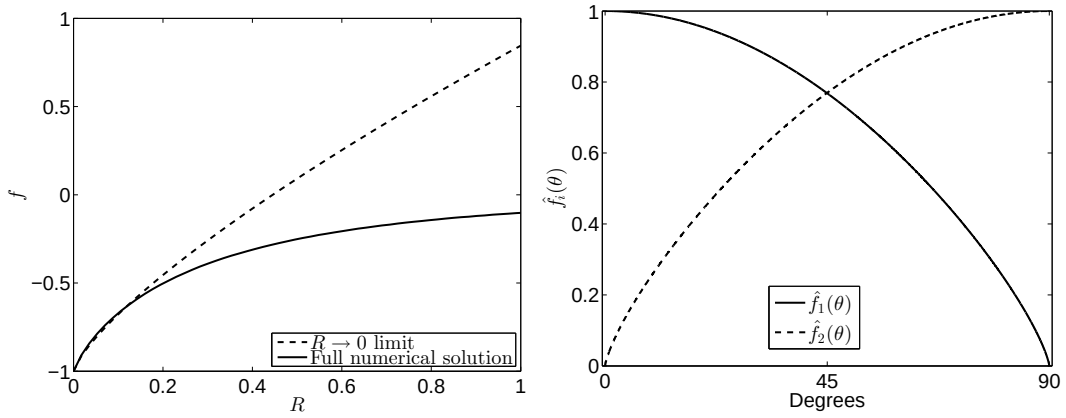
which for non-zero $\hat{f}_1$ and non-zero $\hat{\alpha}_1$ results in

$$\hat{\alpha}_1 = 1 - \frac{2\tilde{\mu}_1}{\tilde{\sigma}_1^2} > 0 \Rightarrow \tilde{\mu}_1 < \frac{\tilde{\sigma}_1^2}{2}, \tag{6.46}$$

which is analogous with that of the single asset case (see (6.40),(6.41)), in which a parameter constraint is also present. With $\theta = 0$, which is equivalent to the limit of the small asset price of asset one, $\tilde{S}_1 \to 0$, with $\tilde{S}_2 = 0$, the solution of (6.43) should match with the single asset case in $\tilde{S}_1$ space and be independent of $\tilde{S}_2$. For this we need that $\hat{f}_1(\theta = 0) = c_1(q_1)$ and $\hat{f}_1\left(\theta = \frac{\pi}{2}\right) = 0$.

Turning now to (6.44) for $i = 2$ we can apply similar analysis. As above, examining (6.44) at $\theta = \frac{\pi}{2}$ we find

$$\hat{\alpha}_2 = 1 - \frac{2\tilde{\mu}_2}{\tilde{\sigma}_2^2} > 0 \Rightarrow \tilde{\mu}_2 < \frac{\tilde{\sigma}_2^2}{2}. \tag{6.47}$$



(a) f as $R \to 0$ (b) $\hat{f}(\theta)$

Figure 6.10: Figure 6.10(a) shows the solution of f given by (6.43) as $R \to 0$ alongside the full-numerical solution, along the line $\theta = \frac{\pi}{4}$, with parameter values corresponding to that of figure 6.9(b). Figure 6.10(b) shows the solution of $\hat{f}_1(\theta)$ and $\hat{f}_2(\theta)$, which is the solution of the ODEs (6.44) for $i = 1$ and $i = 2$ respectively, again with parameter values corresponding to that of figure 6.9(b). Although figure 6.10(b) looks symmetric we stress it is not.

Referring to the single asset case, at $\theta = \frac{\pi}{2}$, which is equivalent to the limit of the small asset price of asset two, $\tilde{S}_2 \to 0$, with $\tilde{S}_1 = 0$, the solution of (6.43) should match with the single asset case in $\tilde{S}_2$ space and be independent of $\tilde{S}_1$. Therefore we need that $\hat{f}_2(\theta = 0) = 0$ and $\hat{f}_2\left(\theta = \frac{\pi}{2}\right) = c_2(q_2)$.

We can solve (6.44) using a central finite-difference approach. In (6.43) it is clearly the least positive of $\hat{\alpha}_1, \hat{\alpha}_2$ that will dominate in the limit as $R \to 0$. The solution f, given by (6.43), for $q_1 = q_2 = 1$ along the line $\theta = \frac{\pi}{4}$ can be seen in figure 6.10(a) along with the solution of (6.44) which can be seen in figure 6.10(b).

If we consider the symmetric parameter case, with $\tilde{\sigma}_1 = \tilde{\sigma}_2 = \tilde{\sigma}$ and $\tilde{\mu}_1 = \tilde{\mu}_2 = \tilde{\mu}$, (6.43) can be reduced to

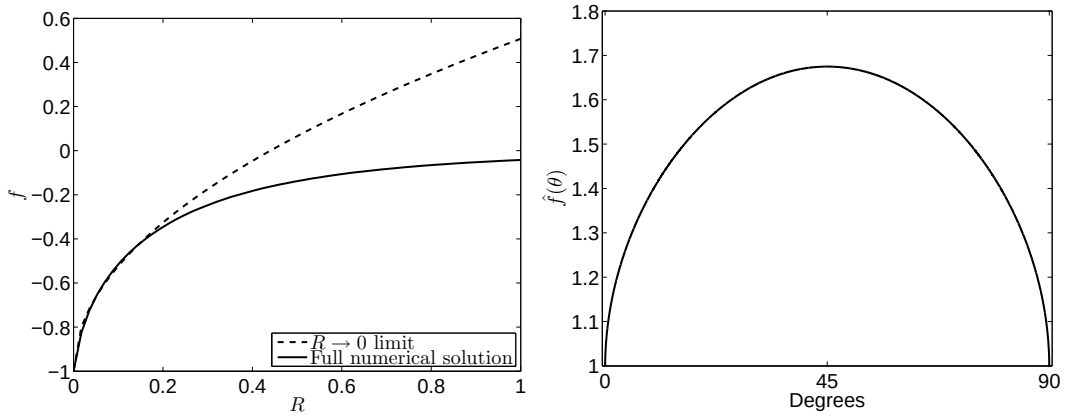
$$f = -1 + R^{\hat{\alpha}} \hat{f}(\theta), \tag{6.48}$$

and rather than a family of ODEs (6.44) reduces to a single ODE, with boundary conditions

$$\hat{f}(\theta = 0) = \hat{f}(\theta = \pi/2) = c(q_1, q_2). \tag{6.49}$$

The value of $\hat{\alpha}$ and the boundary conditions of $\hat{f}(\theta)$ can be deduced in a similar fashion as above, with

$$\hat{\alpha} = 1 - \frac{2\tilde{\mu}}{\tilde{\sigma}^2} > 0 \Rightarrow \tilde{\mu} < \frac{\tilde{\sigma}^2}{2}, \tag{6.50}$$



(a) f as $R \to 0$ (b) $\hat{f}(\theta)$

Figure 6.11: Figure 6.11(a) shows the solution of f given by (6.48) as $R \to 0$ alongside the full-numerical solution, along the line $\theta = \frac{\pi}{4}$, with parameter values corresponding to that of figure 6.9(a). Figure 6.11(b) shows the solution of $\hat{f}(\theta)$, again with parameter values corresponding to that of figure 6.9(a).

in which we see the same parameter constraint as in the single asset case (6.41) is present. The solution f of (6.48) for $q_1 = q_2 = 1$ along the line $\theta = \frac{\pi}{4}$ can be seen in figure 6.11(a) along with the solution of $\hat{f}(\theta)$ which can be seen in figure 6.11(b).

Homogeneous family of solutions

Let us now consider the family of solutions which correspond to the homogeneous case, i.e. with boundary conditions

$$f(\theta = 0) = f(\theta = \pi/2) = 0, \tag{6.51}$$

for the PDE (6.42). This PDE emits solutions of the form

$$f = R^{\tilde{\alpha}} \tilde{f}(\theta), \tag{6.52}$$

which results in the same ODE as (6.44), with $\hat{\alpha}_i$ replaced with $\tilde{\alpha}$, with homogeneous boundary conditions of the form

$$\tilde{f}(\theta = 0) = \tilde{f}(\theta = \pi/2) = 0. \tag{6.53}$$

This implies the problem is a non-linear (quadratic) eigenvalue problem for $\tilde{\alpha}$. In order to address this non-linearity we can introduce a new variable (in the standard way)

$$f^* = \tilde{\alpha}\tilde{f}, \tag{6.54}$$

and this renders the eigenvalue problem in the system $\left(\tilde{f}, f^*\right)$ a linear one (see Tisseur and Meerbergen, 2001). Equation (6.44) can thus be wrote in the form

$$\begin{aligned}
&\cos^2\theta\sin^2\theta\left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\sigma_2^2\right)\tilde{f}_{\theta\theta} \\
&+ \cos\theta\sin\theta\left(\tilde{\sigma}_1^2\cos^2\theta - \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) - \tilde{\sigma}_2^2\sin^2\theta - \tilde{\mu}_1 + \tilde{\mu}_2\right)\tilde{f}_\theta \\
&+ \Bigg(\left(\cos^2\theta\sin^2\theta\left(\frac{1}{2}\tilde{\sigma}_1^2 - \rho\tilde{\sigma}_1\tilde{\sigma}_2 + \frac{1}{2}\tilde{\sigma}_2^2\right) + \tilde{\mu}_1\cos^2\theta + \tilde{\mu}_2\sin^2\theta\right) \\
&- \left(\frac{1}{2}\tilde{\sigma}_1^2\cos^4\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\cos^2\theta\sin^2\theta + \frac{1}{2}\tilde{\sigma}_2^2\sin^4\theta\right)\Bigg) f^* \\
&= -\tilde{\alpha}\cos\theta\sin\theta\left(-\tilde{\sigma}_1^2\cos^2\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\left(2\cos^2\theta - 1\right) + \tilde{\sigma}_2^2\sin^2\theta\right)\tilde{f}_\theta \\
&- \tilde{\alpha}\left(\frac{1}{2}\tilde{\sigma}_1^2\cos^4\theta + \rho\tilde{\sigma}_1\tilde{\sigma}_2\cos^2\theta\sin^2\theta + \frac{1}{2}\tilde{\sigma}_2^2\sin^4\theta\right) f^*. \qquad (6.55)
\end{aligned}$$

Using second-order differencing on (6.55), with boundary conditions given by (6.53), coupled with (6.54) we can write the system as

$$\mathbf{AX} = \tilde{\alpha}\mathbf{BX}, \qquad (6.56)$$

with $\tilde{\alpha}$ as the eigenvalues and the corresponding eigenvectors being X, which contains both $\tilde{f}$ and f^*. The numerical set-up is such that if we discretise $\tilde{f}$ into n grid points, the matrices $\mathbf{A}$ and $\mathbf{B}$ will be size $2n \times 2n$, with corresponding elements alternating between the parameter values of $\tilde{f}$ and f^*. This can be seen in (6.57), (6.58) and (6.59), with the parameters corresponding to $\tilde{f}$ labelled $\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}, \tilde{e}$ and $\tilde{h}$ while the parameters corresponding to f^* are labelled b^* and e^*. Superscript $\top$ denotes the transpose.

$$\mathbf{A} = \begin{bmatrix}
\tilde{b}_0 & b_0^* & \tilde{c}_0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\tilde{a}_1 & 0 & \tilde{b}_1 & b_1^* & \tilde{c}_1 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & 0 \\
\vdots & \ddots & 0 & \tilde{a}_{k_{max}-1} & 0 & \tilde{b}_{k_{max}-1} & b^*_{k_{max}-1} & \tilde{c}_{k_{max}-1} & 0 \\
0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & \dots & 0 & 0 & 0 & \tilde{a}_{k_{max}} & 0 & \tilde{b}_{k_{max}} & b^*_{k_{max}} \\
0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \qquad (6.57)$$

$$\mathbf{B} = \begin{bmatrix} \tilde{e}_0 & e_0^* & \tilde{h}_0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \tilde{d}_1 & 0 & \tilde{e}_1 & e_1^* & \tilde{h}_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & 0 \\ \vdots & \ddots & 0 & \tilde{d}_{k_{max}-1} & 0 & \tilde{e}_{k_{max}-1} & e^*_{k_{max}-1} & \tilde{h}_{k_{max}-1} & 0 \\ 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \tilde{d}_{k_{max}} & 0 & \tilde{e}_{k_{max}} & e^*_{k_{max}} \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \tag{6.58}$$

$$\mathbf{X} = \begin{bmatrix} \tilde{f}_0 & f_0^* & \tilde{f}_1 & f_1^* & \dots & \dots & \tilde{f}_{k_{max}} & f^*_{k_{max}} \end{bmatrix}^\top. \tag{6.59}$$

Given this problem formulation we can treat the problem as an algebraic generalised eigenvalue problem, using a QZ algorithm. We use the NAG QZ algorithm *f02bjc*. The QZ algorithm takes two asymmetric matrices $\mathbf{A}$ and $\mathbf{B}$ and finds the generalized eigenvalues and eigenvectors such that (6.56) is satisfied and $\tilde{\alpha}$ obeys $\det(\mathbf{A} - \tilde{\alpha}\mathbf{B}) = 0$. The QZ decomposition factorises both matrices as

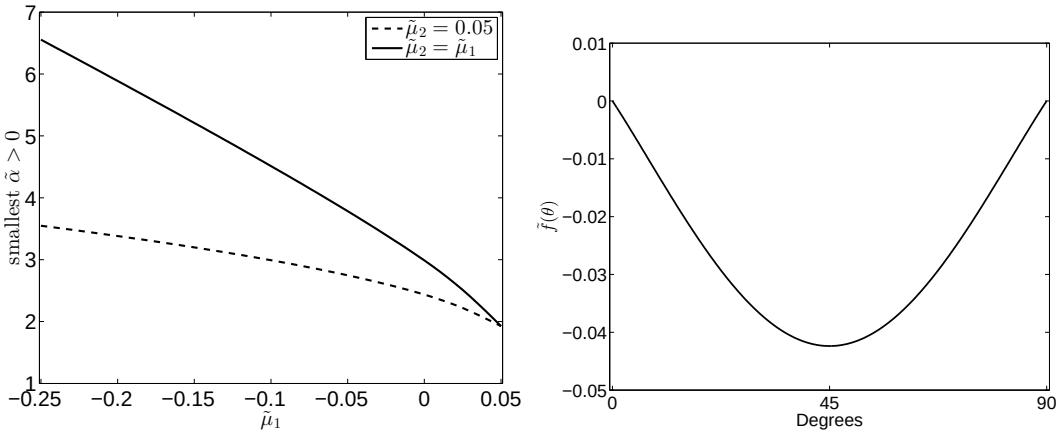
$$\mathbf{A} = \mathbf{QSZ}^\top, \qquad \mathbf{B} = \mathbf{QTZ}^\top, \tag{6.60}$$

where $\mathbf{Q}$ and $\mathbf{Z}$ are unitary and $\mathbf{S}$ and $\mathbf{T}$ are upper triangular. The generalised eigenvalues can then be computed as

$$\tilde{\alpha}_i = \mathbf{S}_{ii}/\mathbf{T}_{ii}. \tag{6.61}$$

For more on QZ decomposition the author recommends Golub and Van Loan (2012).

Computations indicated conclusively that no complex eigenvalue $\tilde{\alpha}$ exist, although both positive and negative values were found. For this problem it is the positive eigenvalues that are of interest for $R \to 0$, particularly the smallest positive eigenvalue as this will exhibit the slowest decay in this region. Figure 6.12(a) shows results for the smallest positive value of $\tilde{\alpha}$ for a range of values of the drift of asset one, $\tilde{\mu}_1$, one curve for $\tilde{\mu}_2 = \tilde{\mu}_1$, the other for $\tilde{\mu}_2 = 0.05$. Other parameters are those used previously in figure 6.9(a), namely volatility $\tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4$ and correlation $\rho = -0.4$. Numerical results indicate that there are a (likely infinite) discrete set of eigenvalues $\tilde{\alpha}$. A plot



(a) Variation of smallest positive $\tilde{\alpha}$

(b) $\tilde{f}(\theta)$ for smallest positive $\tilde{\alpha}$

Figure 6.12: Variation of smallest positive $\tilde{\alpha}$ with $\tilde{\mu}_1$ (figure 6.12(a)), and $\tilde{f}(\theta)$ for smallest positive $\tilde{\alpha} = 2.17$ (figure 6.12(b)). The parameters match that of figure 6.9(a), that being $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4$ and $\rho = -0.4$.

of the eigenvector $\tilde{f}(\theta)$ corresponding to the smallest positive $\tilde{\alpha}$ can be seen in figure 6.12(b).

Given the inhomogeneous and homogeneous solutions we have discussed we conclude in the limit as $R \to 0$ the solution can be wrote as an infinite series of the form

$$f = -1 + R^{\hat{\alpha}_1}\hat{f}_1(\theta) + R^{\hat{\alpha}_2}\hat{f}_2(\theta) + \sum_{i=0}^{\infty} R^{\tilde{\alpha}_i}\tilde{f}_i(\theta), \tag{6.62}$$

with the *hat* terms being solutions of the inhomogeneous problem and the *tilde* terms being the solutions of the homogeneous (eigenvalue) problem. As $R \to 0$ it is the smallest positive value of $\hat{\alpha}_1, \hat{\alpha}_2, \tilde{\alpha}_i$ that will have the corresponding dominant behaviour. For the parameter values used in the example shown in figure 6.9(a) we have $\hat{\alpha}_1 = \hat{\alpha}_2 = 0.5$ and the smallest positive eigenvalue $\tilde{\alpha}_i$ takes the value $\tilde{\alpha}^+_{\min} = 2.17$. For this example it is thus the inhomogeneous family of solutions which dominate.

The limit as $\tilde{S}_1, \tilde{S}_2 \to 0$

Given the above analytic analysis in polar coordinates we shall confirm these by examining the zero limit of the asset price in Cartesian coordinates, $\tilde{S}_1$ and $\tilde{S}_2$. In Cartesian coordinates we are examining (6.29) in the limit as $\tilde{S}_1, \tilde{S}_2 \to 0$. In this limit, asymptotic balancing suggests the non-linearity is negligible and as such we examine the

linear PDE

$$\tilde{\mu}_1\tilde{S}_1 f_{\tilde{S}_1} + \tilde{\mu}_2\tilde{S}_2 f_{\tilde{S}_2} + \frac{1}{2}\tilde{\sigma}_1^2\tilde{S}_1^2 f_{\tilde{S}_1\tilde{S}_1} + \frac{1}{2}\tilde{\sigma}_2^2\tilde{S}_2^2 f_{\tilde{S}_2\tilde{S}_2} + \rho\tilde{\sigma}_1\tilde{\sigma}_2\tilde{S}_1\tilde{S}_2 f_{\tilde{S}_1\tilde{S}_2} = 0. \tag{6.63}$$

Firstly, we shall examine the inhomogeneous solution. It was found (6.63) emits solution of the form

$$f = -1 + \tilde{S}_1^{\hat{\alpha}_1}\hat{f}_1\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right) + \tilde{S}_2^{\hat{\alpha}_2}\hat{f}_2\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right), \tag{6.64}$$

which is analogous to the polar coordinate form given by (6.43). Substituting the form of (6.64) into (6.63) results in the linear family of ODEs

$$\begin{aligned}
&\zeta^2\left(\frac{1}{2}\tilde{\sigma}_1^2 + \frac{1}{2}\tilde{\sigma}_2^2 - \tilde{\sigma}_1\tilde{\sigma}_2\rho\right)\hat{f}_1'' \\
&+ \zeta\left(-\tilde{\mu}_1 + \tilde{\mu}_2 - \tilde{\sigma}_1^2\left(\hat{\alpha}_1 - 1\right) - \tilde{\sigma}_1\tilde{\sigma}_2\rho\left(\hat{\alpha}_1 - 1\right)\right)\hat{f}_1' \\
&+ \left(\tilde{\mu}_1\hat{\alpha}_1 + \frac{1}{2}\tilde{\sigma}_1^2\left(\hat{\alpha}_1 - 1\right)\hat{\alpha}_1\right)\hat{f}_1 = 0,
\end{aligned} \tag{6.65}$$

and

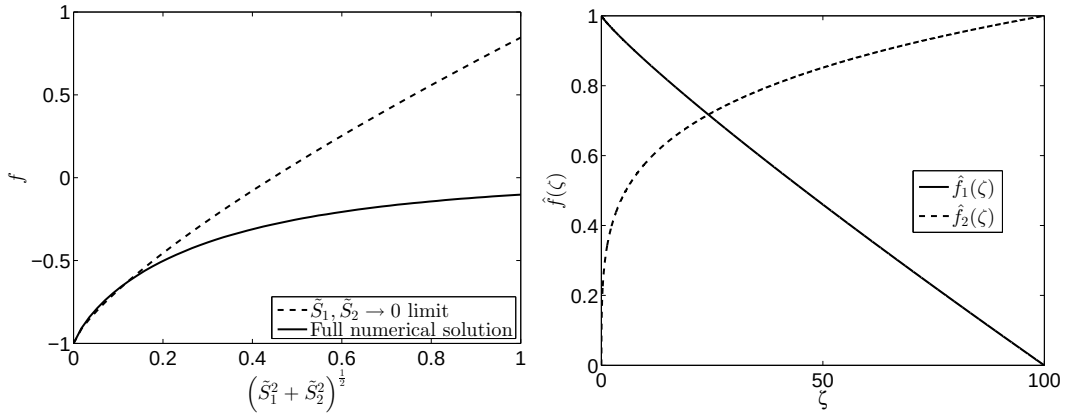
$$\begin{aligned}
&\zeta^2\left(\frac{1}{2}\tilde{\sigma}_1^2 + \frac{1}{2}\tilde{\sigma}_2^2 - \tilde{\sigma}_1\tilde{\sigma}_2\rho\right)\hat{f}_2'' \\
&+ \zeta\left(-\tilde{\mu}_1 + \tilde{\mu}_2 + \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2\hat{\alpha}_2 + \tilde{\sigma}_1\tilde{\sigma}_2\rho\left(\hat{\alpha}_2 + 1\right)\right)\hat{f}_2' \\
&+ \left(\tilde{\mu}_1\hat{\alpha}_2 + \frac{1}{2}\tilde{\sigma}_2^2\left(\hat{\alpha}_2 - 1\right)\hat{\alpha}_1\right)\hat{f}_2 = 0,
\end{aligned} \tag{6.66}$$

with $\zeta = \tilde{S}_2/\tilde{S}_1$, which is the ratio of the price of asset two to asset one. We use similar methods as used for the polar coordinates, by referring to the single asset case, to derive boundary conditions and solve for values of $\hat{\alpha}_1, \hat{\alpha}_2$. Two boundary conditions are needed for both $\hat{f}_1(\zeta)$ and $\hat{f}_2(\zeta)$, as $\zeta \to 0$ and $\zeta \to \infty$. For $\zeta = 0$, which is equivalent to the limit of the small asset price of asset one, $\tilde{S}_1 \to 0$, with $\tilde{S}_2 = 0$, the solution of (6.64) should match with the single asset case in $\tilde{S}_1$ space and be independent of $\tilde{S}_2$. For this we need that $\hat{f}_1 = c_1(q_1)$ and $\hat{f}_2 = 0$ at $\zeta = 0$. In this limit we can obtain values of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ which are given by:

$$\hat{\alpha}_1 = 1 - \frac{2\tilde{\mu}_1}{\tilde{\sigma}_1^2}, \qquad \hat{\alpha}_2 = 1 - \frac{2\tilde{\mu}_2}{\tilde{\sigma}_2^2}, \tag{6.67}$$

which are identical to those found in polar coordinates, (6.46) and (6.47).

Similarly, as $\zeta \to \infty$, which is equivalent to the limit of the small asset price of asset two, $\tilde{S}_2 \to 0$, with $\tilde{S}_1 = 0$, the solution of (6.64) should be independent of $\tilde{S}_1$, for which we need that $\hat{f}_1 = 0$ and $\hat{f}_2 = c_2(q_2)$ as $\zeta \to \infty$.



(a) f as $\tilde{S}_1, \tilde{S}_2 \to 0$ along $\zeta = 1$ (b) $\hat{f}(\zeta)$

Figure 6.13: Figure 6.13(a) shows the solution of f from (6.64) along $\zeta = 1$ as $\tilde{S}_1, \tilde{S}_2 \to 0$ alongside the full-numerical solution, with parameters values corresponding to that of figure 6.9(b). Figure 6.13(b) shows the solution of $\hat{f}_1(\zeta)$ and $\hat{f}_2(\zeta)$, which is the solution of the ODEs (6.65) and (6.66) respectively, again with parameter values corresponding to that of figure 6.9(b).

The solution f, given by (6.64), for $q_1 = q_2 = 1$ along the line $\zeta = 1$ can be seen in figure 6.13(a), and the solutions of (6.65) and (6.66) can be seen in figure 6.9(b).

Similar to when using polar coordinates form, when symmetry is present due to a symmetrical choice in parameters, (6.64) can be reduced to

$$f = -1 + \left(\tilde{S}_1^{\hat{\alpha}} + \tilde{S}_2^{\hat{\alpha}}\right) \hat{f}\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right), \tag{6.68}$$

in which $\hat{f}$ solves the ODE

$$\begin{aligned} &\zeta^2\left(1+\zeta^{\hat{\alpha}}\right)\tilde{\sigma}^2(1-\rho)\hat{f}'' + \zeta\left((\hat{\alpha}+1)\zeta^{\hat{\alpha}} - (\hat{\alpha}-1)\right)\tilde{\sigma}^2(1-\rho)\hat{f}' \\ &+ \left(1+\zeta^{\hat{\alpha}}\right)\left(\tilde{\mu}\hat{\alpha} + \tilde{\sigma}^2(\hat{\alpha}-1)\hat{\alpha}\right)\hat{f} = 0, \end{aligned} \tag{6.69}$$

with boundary conditions

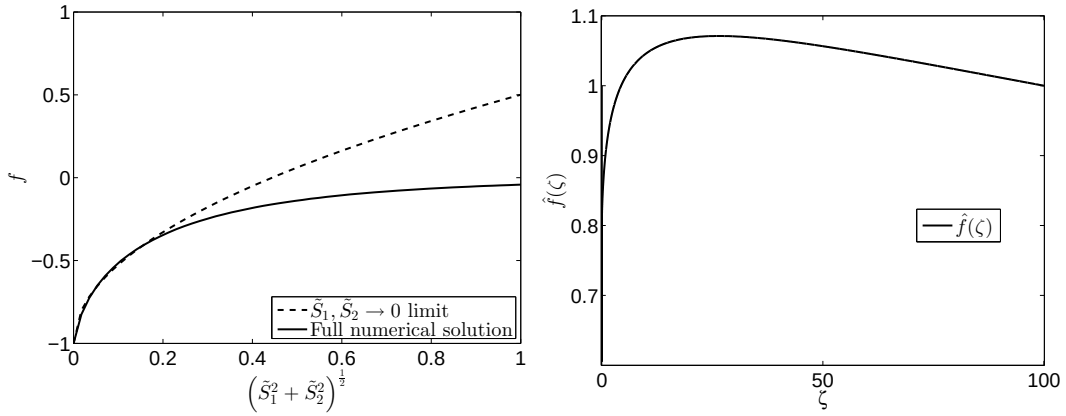
$$\hat{f}(\zeta = 0) = \hat{f}(\zeta \to \infty) = c(q_1, q_2) \tag{6.70}$$

Examining (6.69) at $\zeta = 0$ we find

$$\hat{\alpha} = 1 - \frac{2\tilde{\mu}}{\tilde{\sigma}^2} > 0 \Rightarrow \tilde{\mu} < \frac{\tilde{\sigma}^2}{2}, \tag{6.71}$$

which is consistent with that of (6.50) as found when using polar coordinates.

The solution (6.68) for $q_1 = q_2 = 1$ along the line $\zeta = 1$ can be seen in figure 6.13(a), as well as the solutions of (6.69) which can be seen in figure 6.9(a).



(a) f as $\tilde{S}_1, \tilde{S}_2 \to 0$ along $\zeta = 1$ (b) $\hat{f}(\zeta)$

Figure 6.14: Figure 6.14(a) shows the solution of f from (6.68) along $\zeta = 1$ as $\tilde{S}_1, \tilde{S}_2 \to 0$ alongside the full-numerical solution, with parameters values corresponding to that of figure 6.9(a). Figure 6.14(b) shows the solution of $\hat{f}(\zeta)$, which is the solution of the ODE (6.69), again with parameter values corresponding to that of figure 6.9(a).

Next we examine the solutions found when examining the homogeneous PDE, i.e. (6.63) with homogeneous boundary conditions. This PDE emits solutions of the form

$$f = \left(\tilde{S}_1\tilde{S}_2\right)^{\tilde{\alpha}} \tilde{f}\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right), \tag{6.72}$$

which results in an ODE for $\tilde{f}$ of the form

$$\begin{aligned}&\zeta^2\left(\frac{1}{2}\tilde{\sigma}_1^2+\frac{1}{2}\tilde{\sigma}_2^2-\tilde{\sigma}_1\tilde{\sigma}_2\rho\right)\tilde{f}''+\zeta\left(\tilde{\mu}_2-\tilde{\mu}_1-\tilde{\sigma}_1^2\left(\tilde{\alpha}-1\right)+\tilde{\sigma}_2^2\tilde{\alpha}-\tilde{\sigma}_1\tilde{\sigma}_2\rho\right)\tilde{f}'\\&+\left(\tilde{\mu}_1\tilde{\alpha}+\tilde{\mu}_2\tilde{\alpha}+\frac{1}{2}\tilde{\sigma}_1^2\left(\tilde{\alpha}_1-1\right)\tilde{\alpha}+\frac{1}{2}\tilde{\sigma}_2^2\left(\tilde{\alpha}-1\right)\tilde{\alpha}+\tilde{\sigma}_1\tilde{\sigma}_2\rho\tilde{\alpha}^2\right)\tilde{f}=0,\end{aligned} \tag{6.73}$$

with

$$\tilde{f}\left(\zeta=0\right)=\tilde{f}\left(\zeta\to\infty\right)=0. \tag{6.74}$$

Equation (6.73) is similar to (6.55) in that it is a non-linear (quadratic) eigenvalue problem $\tilde{\alpha}$. We thus use a similar method for solving as discussed above which involves introducing the new function

$$f^* = \tilde{\alpha}\tilde{f}, \tag{6.75}$$

so we now have an eigenvalue problem for the system $\left(\tilde{f}, f^*\right)$ with

$$\begin{aligned}&\zeta^2(\frac{1}{2}\tilde{\sigma}_1^2+\frac{1}{2}\tilde{\sigma}_2^2-\tilde{\sigma}_1\tilde{\sigma}_2\rho)\tilde{f}''+\zeta\left(\tilde{\mu}_2-\tilde{\mu}_1+\tilde{\sigma}_1^2-\tilde{\sigma}_1\tilde{\sigma}_2\rho\right)\tilde{f}'+(\tilde{\mu}_1+\mu_2-\frac{1}{2}\tilde{\sigma}_1^2-\frac{1}{2}\tilde{\sigma}_2^2)f^*\\&=\tilde{\alpha}\left(\left(\tilde{\sigma}_1^2-\tilde{\sigma}_2^2\right)\tilde{f}'-\left(\frac{1}{2}\tilde{\sigma}_1^2+\frac{1}{2}\tilde{\sigma}_2^2+\tilde{\sigma}_1\tilde{\sigma}_2\rho\right)f^*\right).\end{aligned} \tag{6.76}$$

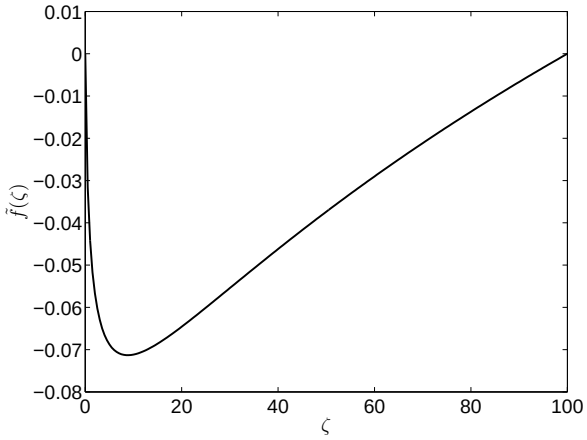


Figure 6.15: $\tilde{f}(\theta)$ for smallest positive $\tilde{\alpha}$. The parameters match that of figure 6.9(a), that being $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4$ and $\rho = -0.4$.

Using second-order differencing (6.76) can be discretised and, along with (6.74), can be wrote as:

$$\mathbf{AX} = \tilde{\alpha}\mathbf{BX}$$

and solved using the QZ method discussed above. As with the polar coordinate case, it is the smallest positive eigenvalue $\tilde{\alpha}$ we are interested in as this will be dominant over the larger $\tilde{\alpha}$'s for $\tilde{S}_1, \tilde{S}_2 \to 0$. In figure 6.15 we have plotted the eigenvector $\tilde{f}(\zeta)$ corresponding to the smallest positive $\tilde{\alpha}$ for $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4$ and $\rho = -0.4$. Numerical computations similarly suggested a (likely infinite) set of discrete eigenvalues.

Given the homogeneous and inhomogeneous solutions examined we conclude in the limit $\tilde{S}_1, \tilde{S}_2 \to 0$ the solution takes the form

$$f = -1 + \tilde{S}_1^{\hat{\alpha}_1} \hat{f}_1\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right) + \tilde{S}_2^{\hat{\alpha}_2} \hat{f}_2\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right) + \sum_{i=0}^{\infty} \left(\tilde{S}_1 \tilde{S}_2\right)^{\tilde{\alpha}_i} \tilde{f}_i\left(\frac{\tilde{S}_2}{\tilde{S}_1}\right), \tag{6.77}$$

with the *hat* terms being solutions of the inhomogeneous problem and the *tilde* terms being the solutions of the homogeneous (eigenvalue) problem. This solution is the same as that of the polar coordinates as given by (6.62).

6.4.3 Large inventory solution

We can investigate a large inventory solution for the multiple asset case such that q_1 and/or q_2 become large (both together and individually).

Similar to the single asset case, numerical simulations and asymptotic analysis lead us to conclude

$$\lim_{q_i \to \infty} \left(\frac{f\left(q_1, \ldots, q_i, \ldots, q_N, \tilde{S}_1, \ldots, \tilde{S}_N\right)}{f\left(q_1, \ldots, q_i - 1, \ldots, q_N, \tilde{S}_1, \ldots, \tilde{S}_N\right)} \right) \to 1. \tag{6.78}$$

For the case of only one amount of inventory tending to infinity, say $q_2 \to \infty$ and $q_1 \sim O(1)$ equation (6.29) reduces to

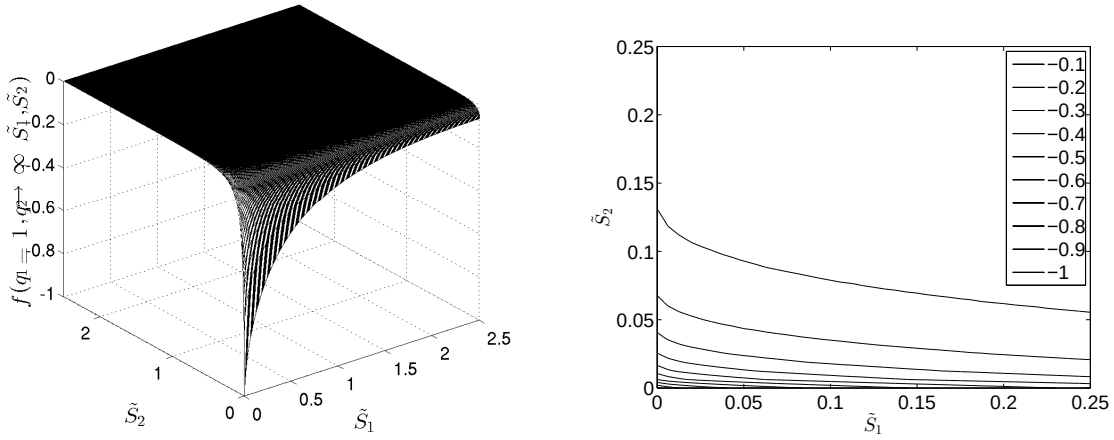
$$\tilde{\mu}_1 \tilde{S}_1 f_{\tilde{S}_1} + \tilde{\mu}_2 \tilde{S}_2 f_{\tilde{S}_2} + \frac{1}{2}\tilde{\sigma}_1^2 \tilde{S}_1^2 f_{\tilde{S}_1\tilde{S}_1} + \frac{1}{2}\tilde{\sigma}_2^2 \tilde{S}_2^2 f_{\tilde{S}_2\tilde{S}_2} + \rho\tilde{\sigma}_1\tilde{\sigma}_2\tilde{S}_1\tilde{S}_2 f_{\tilde{S}_1\tilde{S}_2}$$
$$- \frac{e^{l_1}\tilde{S}_1 f}{\tilde{S}_1 + l_1} \left(\frac{l_1 f\left(q_1, \tilde{S}_1, \tilde{S}_2\right)}{\left(\tilde{S}_1 + l_1\right) f\left(q_1 - 1, \tilde{S}_1, \tilde{S}_2\right)} \right)^{\frac{l_1}{\tilde{S}_1}} - \frac{\tilde{\lambda} e^{l_2}\tilde{S}_2 f}{\tilde{S}_2 + l_2} \left(\frac{l_2}{\left(\tilde{S}_2 + l_2\right)} \right)^{\frac{l_2}{\tilde{S}_2}} = 0, \tag{6.79}$$

with $f = f\left(q_1, \tilde{S}_1, \tilde{S}_2\right)$ and boundary conditions analogous to those discussed in section 6.2.2. For $q_2 \to \infty$, we only consider the case of the inventory of asset one being positive, $q_1 > 0$, as $q_1 = 0$ is identical to the single asset case examined in chapter 4. Equation (6.79) is still a non-linear PDE and thus an iterative method must be used with appropriate boundary conditions in the asset prices, $\tilde{S}_1$ and $\tilde{S}_2$. To solve (6.79) we use a similar iterative method as described above for the perpetual case when the inventory was of order one, $q_1, q_2 \sim O(1)$, and use the same boundary conditions in the asset prices, $\tilde{S}_1$ and $\tilde{S}_2$. A solution for (6.79) can be seen in figure 6.16, in which figure 6.16(a) is a surface plot of $f(q_1 = 1, q_2 \to \infty, \tilde{S}_1, \tilde{S}_1)$, along with a contour plot in figure 6.16(b). Figure 6.16(a) shows the rapid decay to zero as the price of asset two, $\tilde{S}_2$, increases, while at $\tilde{S}_2 = 0$ we see the more gentle decay, corresponding to $q_1 = 1$. This is also portrayed in figure 6.16(b), which we can compare to figure 6.9(a) which is the case of $q_1 = q_2 = 1$.

We can also examine the case of when both the inventory of asset one and asset two are large, $q_1, q_2 \to \infty$, in which the non-linear PDE (6.79) reduces to the *linear* PDE:

$$\tilde{\mu}_1 \tilde{S}_1 f_{\tilde{S}_1} + \tilde{\mu}_2 \tilde{S}_2 f_{\tilde{S}_2} + \frac{1}{2}\tilde{\sigma}_1^2 \tilde{S}_1^2 f_{\tilde{S}_1\tilde{S}_1} + \frac{1}{2}\tilde{\sigma}_2^2 \tilde{S}_2^2 f_{\tilde{S}_2\tilde{S}_2} + \rho\tilde{\sigma}_1\tilde{\sigma}_2\tilde{S}_1\tilde{S}_2 f_{\tilde{S}_1\tilde{S}_2}$$
$$- \frac{e^{l_1}\tilde{S}_1 f}{\tilde{S}_1 + l_1} \left(\frac{l_1}{\left(\tilde{S}_1 + l_1\right)} \right)^{\frac{l_1}{\tilde{S}_1}} - \frac{\tilde{\lambda} e^{l_2}\tilde{S}_2 f}{\tilde{S}_2 + l_2} \left(\frac{l_2}{\left(\tilde{S}_2 + l_2\right)} \right)^{\frac{l_2}{\tilde{S}_2}} = 0, \tag{6.80}$$

with $f = f(\tilde{S}_1, \tilde{S}_2)$ with the same boundary conditions to those above for inventory of order one, $q_1, q_2 \sim O(1)$. As this is a linear PDE we need not use an iterative



(a) $f(q_1 = 1, q_2 \to \infty, \tilde{S}_1, \tilde{S}_1)$ mesh plot (b) $f(q_1 = 1, q_2 \to \infty, \tilde{S}_1, \tilde{S}_1)$ contour plot

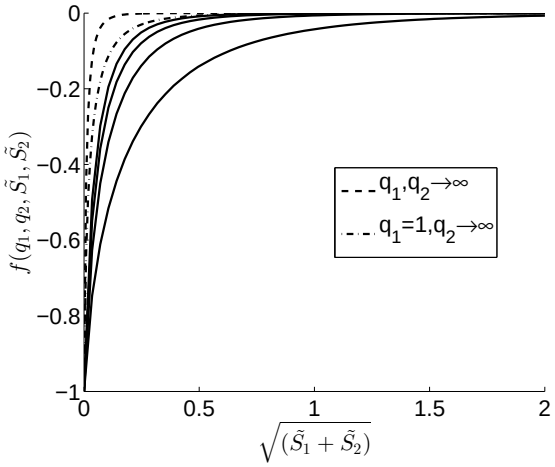
Figure 6.16: Plot of $q_1 = 1$ with $q_2 \to \infty$ as a surface plot and contour plot with parameter values $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4, l_1 = l_2 = 1, \tilde{\lambda} = 1$ and $\rho = -0.4$

method. A solution of (6.80) can be seen in figure 6.17. In figure 6.17(a) we show the convergence of increasing q_1, q_2. This figure shows curves along the main diagonal of the asset prices, i.e. $\tilde{S}_1 = \tilde{S}_2$, for $f(q_1 = q_2, \tilde{S}_1, \tilde{S}_2)$ (solid line), $f(q_1 = 1, q_2 \to \infty, \tilde{S}_1, \tilde{S}_2)$ (dot-dash line) and $f(q_1 \to \infty, q_2 \to \infty, \tilde{S}_1\tilde{S}_2)$ (broken line). We see the solution of $f(q_1, q_2, \tilde{S}_1\tilde{S}_2)$, for $q_1, q_2 \sim O(1)$, converges to $f(q_1 = 1, q_2 \to \infty, \tilde{S}_1\tilde{S}_2)$ which converges to $f(q_1 \to \infty, q_2 \to \infty, \tilde{S}_1\tilde{S}_2)$. Figure 6.17(b) is a contour plot of $f(q_1 \to \infty, q_2 \to \infty, \tilde{S}_1, \tilde{S}_2)$, which is comparable to figure 6.16(b) and figure 6.9(a).

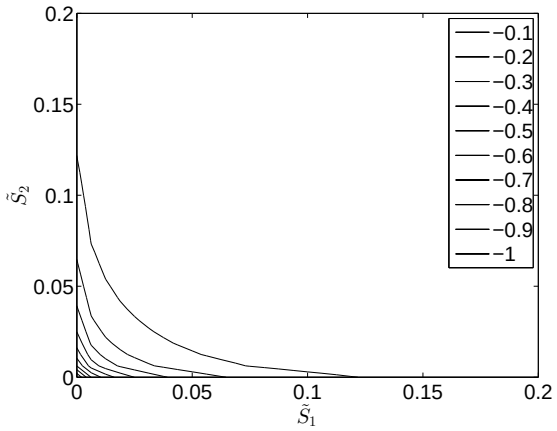
6.5 Summary

In this chapter we examined the trading strategy of a trader with various amounts of various assets, rather than just one asset as examined in chapter 4. We derived a general model for N correlated assets before examining in detail the case of $N = 2$.

We first examined the case of $N = 2$ numerically before examining the solution in several limits. We used a tactical LU decomposition scheme in which the LU matrices only had to be calculated once, greatly reducing computational time. We examined the optimal trading strategies under various parameter regimes and found the correlation term generates some interest, with results relating back to Modern Portfolio Theory (see Markowitz, 1959), in which a trader with positively correlated assets wants an unbalanced portfolio in order to reduce risk while a trader with negatively correlated assets wants a balanced portfolio as to diversify the risk.



(a) $f(q_1, q_2, \tilde{S}_1, \tilde{S}_2)$ for various limits



(b) $f(q_1 = q_2 \to \infty, \tilde{S}_1, \tilde{S}_1)$ contour plot

Figure 6.17: Figure 6.17(a) shows curves of the main diagonal of the asset price for $f(q_1 = q_2, \tilde{S}_1, \tilde{S}_2)$ (solid line), $f(q_1 = 1, q_2 \to \infty, \tilde{S}_1, \tilde{S}_2)$ (dot-dash line) and $f(q_1 \to \infty, q_2 \to \infty, \tilde{S}_1, \tilde{S}_2)$ (broken line). Figure 6.17(b) shows the solution of (6.80) as a contour plot. The parameter values are $\tilde{\mu}_1 = \tilde{\mu}_2 = 0.04, \tilde{\sigma}_1 = \tilde{\sigma}_2 = 0.4, l_1 = l_2 = 1, \tilde{\lambda} = 1$ and $\rho = -0.4$

After producing some interesting numerical solutions we let the numerics, and indeed the single asset case, guide us to which limits are of interest. We examined a small-time-to-termination solution which in the small-time limit has a fully analytic solution. This may be of use in a high-frequency trading framework where the time remaining is small and analytic solutions are preferred over numerical solutions.

In contrast, we examined a perpetual-type solution in which the time-to-termination tends to infinity. We investigated the small $\tilde{S}_1, \tilde{S}_2$ regime in terms of the original coordinates as well as in polar coordinate form, as guided by the numerics. We found the existence of an inhomogeneous and homogeneous family of solutions, the former of which resulted in solving a family of ODEs, while the latter resulted in solving a non-linear (quadratic) eigenvalue problem. It is clear the polar coordinate case is advantageous given the semi-infinite domain for the $\tilde{S}_2/\tilde{S}_1$ case whereas the polar coordinate case has a finite domain. When using numerical boundary conditions we can use more precise Dirclet boundary conditions rather than the need to approximate as $\tilde{S}_2/\tilde{S}_1 \to \infty$.

Finally we examined the limit as the number of assets held, $\{q_1, q_2\}$, went to infinity, both together and separately. This is beneficial as it reduces the problem from a non-linear PDE to a linear PDE which is numerically easier and quicker to solve.

Chapter 7

Extension of Single Underlying: A Stochastic Volatility Model

The work in this chapter is in preprint format and ready for submission for review:
Blair, J., Johnson, P., and Duck, P. (2015). Analysis of optimal liquidation in limit-order books for portfolios of correlated assets with stochastic volatility.
http://eprints.ma.man.ac.uk/2325.

In the previous chapters we have considered volatility to be constant. We shall now extend that work to include stochastic volatility. Stochastic volatility has been implemented rigorously in other areas of mathematical finance, most notably option pricing, but, to the best of our knowledge, has yet to be consider in the optimal trading literature for a trader at the tactical level. It has however been examined under the framework of optimal scheduling by Almgren (2012). Thus it seems appropriate that we extend the tactical level framework in a similar light.

One of the strongest arguments for the use of stochastic volatility is the implied volatility surface which is found to have a smile or skew shape and thus is non-constant[1]. Originally this problem was investigated in Merton (1973) for deterministic functions of time in an option pricing framework. This was further investigated by Derman and Kani (1994) and Rubinstein (1994) who introduce a local volatility model which depends on time and the underlying state variable. It was finally suggested that volatility should itself be stochastic, with the pioneering work including that of Hull

[1]Implied volatility is the volatility which if used in the Black-Scholes option pricing formula will provide the value at which that contract is traded in the market (see Wilmott, 2007).

and White (1987) and Heston (1993), among others. While Hull and White (1987) suggest the volatility squared (variance) should follow a geometric Brownian motion correlated with the asset price, Heston (1993) assumes a mean-reverting process for the variance with a correlation between the variance and the underlying asset. The Heston (1993) model stands out as a dominant model as not only does it explain the mean-reverting nature of the volatility smile, it also allows for only positive values, assuming the Feller (1951) condition is satisfied (the Schöbel and Zhu, 1999 model was similar to that of Heston, 1993 but with an Uhlenbeck and Ornstein, 1930 process for the variance, and as such allowed for negative values). We shall therefore contribute to knowledge by examining the previous framework of chapter 4 under a Heston (1993) stochastic volatility model.

7.1 Problem formulation

We shall now examine when the trader has a quantity $q(0)$ of a single underlying asset to liquidate, but for which the volatility of this underlying is assumed stochastic. We shall consider a similar framework as previous in which a trader wishes to maximise his expected terminal time utility. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0,T])$. We follow Heston (1993) by assuming the asset follows

$$dS(t) = \mu S(t)\, dt + \sqrt{\nu(t)} S(t)\, dW(t), \tag{7.1}$$

which is a geometric Brownian motion, with the square-root of the variance, $\sqrt{\nu(t)}$, (which is the volatility) following an Uhlenbeck and Ornstein (1930) process

$$d\sqrt{\nu(t)} = -\beta\sqrt{\nu(t)}dt + \alpha dB(t). \tag{7.2}$$

It can be shown through Itô's lemma that the variance $\nu(t)$ follows the process

$$d\nu(t) = \left(\alpha^2 - 2\beta\nu(t)\right) dt + 2\alpha\sqrt{\nu(t)}dB(t), \tag{7.3}$$

which can be represented as a mean-reverting square-root process

$$d\nu(t) = \kappa(\theta - \nu(t))\, dt + \sigma\sqrt{\nu(t)}dB(t), \tag{7.4}$$

with $\mathbb{E}[W(t) B(t)] = \rho t$, such that the asset price and variance are correlated. Heston's model has the idea that both high and low values of volatility (variance) are temporary

and will tend to lead to an average specified value over time. The average value of the variance is denoted θ and the speed that the variance reverts to the average is denoted κ. When the variance is higher than the average it will decrease towards it. Similarly when the variance is lower than the average it will increase towards it.

The value function we define is now a function of the variance, ν, and is defined as

$$u(t,X,q,S,\nu) = \sup_{\delta(t)\in\mathcal{A}} \mathbb{E}\left[-e^{-\gamma(X(T)+q(T)S(T))}\right]. \tag{7.5}$$

In a similar framework to previous chapters we have the inventory and wealth following

$$dq(t) = -dN(t), \tag{7.6}$$

and

$$dX(t) = S(t)\left(1+\delta\right)dN(t), \tag{7.7}$$

respectively, where $N\left(t\right)$ is a Poisson process, independent of $W\left(t\right)$ and $B\left(t\right)$, and has intensity

$$\Lambda(\delta) = \lambda e^{-l\left(\frac{S^a-S}{S}\right)} = \lambda e^{-l\delta}. \tag{7.8}$$

Using a similar formulation we can define the HJB equation as

$$\begin{aligned} &u_t + \mu S u_S + \frac{1}{2}\nu S^2 u_{SS} + \kappa\left(\theta-\nu\right)u_\nu + \frac{1}{2}\sigma^2\nu u_{\nu\nu} + \rho\sigma\nu S u_{S\nu} \\ &+\sup_\delta\left[\lambda e^{-l\delta}\left(u\left(t,X+S\left(1+\delta\right),q-1,S,\nu\right)-u\left(t,X,q,S,\nu\right)\right)\right] = 0, \end{aligned} \tag{7.9}$$

with $u = u(t,X,q,S,\nu)$ and initial conditions

$$u(T,X,q,S,\nu) = -e^{-\gamma(X+qS)}, \tag{7.10}$$

$$u(t,X,0,S,\nu) = -e^{-\gamma X}. \tag{7.11}$$

The derivation of this HJB equation is a relatively straightforward extension to that of the constant volatility case and the reader is referred to the description in section 4.1.1 for further details. Extending the verification theorem from the constant volatility case is also straightforward and is thus not detailed here.

As similar to that performed in section 4.1, let us choose an ansatz of the form $u\left(t,X,q,S,\nu\right) = e^{-\gamma X} f\left(\tau,q,S,\nu\right)$, with $\tau = T-t$. Performing a change of variables

$$\tilde{\tau} = \lambda\tau, \quad \tilde{S} = S\gamma, \quad \tilde{\nu} = \frac{\nu}{\lambda}, \quad \tilde{\theta} = \frac{\theta}{\lambda}, \quad \tilde{\mu} = \frac{\mu}{\lambda}, \tag{7.12}$$

we can non-dimensionalise (7.9), such that all the *tilde* variables are dimensionless. Examining the supremum term of (7.9) and using the ansatz we find the optimal trading strategy is given by

$$\delta^* \left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right) = \frac{1}{\tilde{S}} \ln \left(\frac{\left(\tilde{S} + l\right) f \left(\tilde{\tau}, q-1, \tilde{S}, \tilde{\nu}\right)}{l f \left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right)} \right) - 1. \tag{7.13}$$

This results in the non-linear dimensionless PDE

$$\begin{aligned} &- f_{\tilde{\tau}} + \tilde{\mu}\tilde{S} f_{\tilde{S}} + \frac{1}{2}\tilde{\nu}\tilde{S}^2 f_{\tilde{S}\tilde{S}} + \kappa\left(\tilde{\theta} - \tilde{\nu}\right) f_{\tilde{\nu}} + \frac{1}{2}\sigma^2 \tilde{\nu} f_{\tilde{\nu}\tilde{\nu}} + \rho\sigma\tilde{\nu}\tilde{S} f_{\tilde{S}\tilde{\nu}} \\ &- \frac{e^l \tilde{S} f}{\tilde{S} + l} \left(\frac{l f \left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right)}{\left(\tilde{S} + l\right) f \left(\tilde{\tau}, q-1, \tilde{S}, \tilde{\nu}\right)} \right)^{\frac{l}{\tilde{S}}} = 0, \end{aligned} \tag{7.14}$$

with $f = f\left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right)$ and initial conditions

$$f\left(0, q, \tilde{S}, \tilde{\nu}\right) = -e^{-q\tilde{S}}, \tag{7.15}$$

$$f\left(\tau, 0, \tilde{S}, \tilde{\nu}\right) = -1, \tag{7.16}$$

in which we have factored out the variable X, as well as the γ and λ parameter.

To solve (7.14) we require two boundary conditions in each of the two semi-infinite domains of the asset price, $\tilde{S}$, and variance, $\tilde{\nu}$. As $\tilde{S} \to 0$, equation (7.14) reduces to

$$-f_{\tilde{\tau}} + \kappa\left(\tilde{\theta} - \tilde{\nu}\right) f_{\tilde{\nu}} + \frac{1}{2}\sigma^2 \tilde{\nu} f_{\tilde{\nu}\tilde{\nu}} = 0. \tag{7.17}$$

Similarly, as $\tilde{\nu} \to 0$, equation (7.14) reduces to

$$-f_{\tilde{\tau}} + \tilde{\mu}\tilde{S} f_{\tilde{S}} + \kappa\tilde{\theta} f_{\tilde{\nu}} - \frac{e^l \tilde{S} f}{\tilde{S} + l} \left(\frac{l f \left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right)}{\left(\tilde{S} + l\right) f \left(\tilde{\tau}, q-1, \tilde{S}, \tilde{\nu}\right)} \right)^{\frac{l}{\tilde{S}}} = 0. \tag{7.18}$$

As both $\tilde{S} \to 0$ and $\tilde{\nu} \to 0$, equation (7.17) and (7.18) reduce to

$$\frac{\partial f}{\partial \tilde{\tau}} \left(\tilde{\tau}, q, \tilde{S} \to 0, \tilde{\nu} \to 0\right) = 0 \quad \Rightarrow \quad f\left(\tilde{\tau}, q, \tilde{S} \to 0, \tilde{\nu} \to 0\right) = -1, \tag{7.19}$$

which arises from the initial condition (7.15). Similar to previous cases, as $\tilde{S} \to \infty$ we use a Neumann condition of the form

$$\frac{\partial f}{\partial \tilde{S}} \left(\tilde{\tau}, q, \tilde{S} \to \infty, \tilde{\nu}\right) = 0. \tag{7.20}$$

We use a similar condition for the case of the variance tending to infinity which takes the form

$$\frac{\partial f}{\partial \tilde{\nu}}\left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu} \to \infty\right) = 0. \tag{7.21}$$

Intuitively, the higher the volatility the lower the trader's value function as he is risk-averse. Thus the derivative of the value function with respect to $\tilde{\nu}$ is decreasing which can be approximated as zero when $\tilde{\nu} \to \infty$.

7.1.1 Numerical results

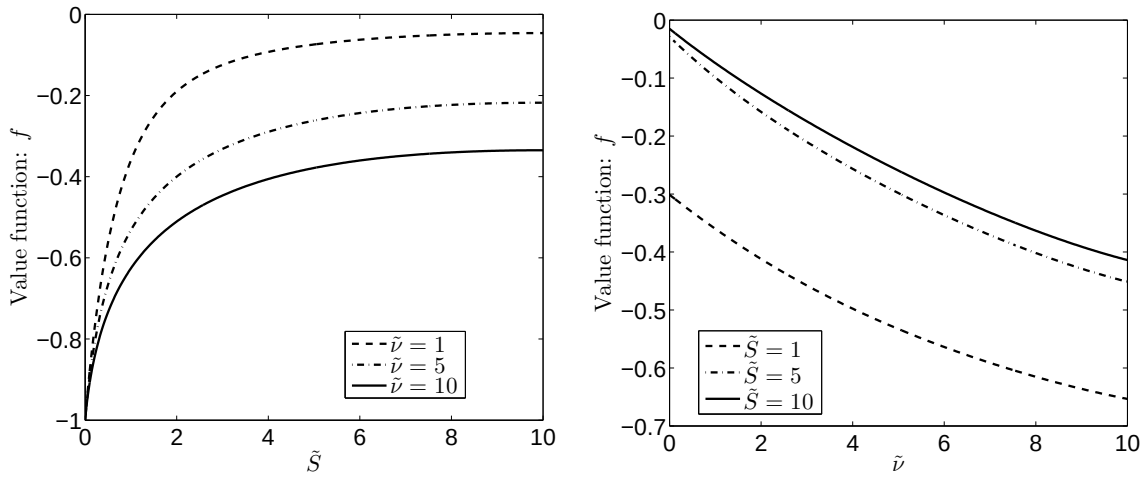
To solve this problem numerically we use a similar implicit-explicit finite-difference scheme to that as used in the multiple underlying case, described in section 6.2.3. This involves taking the derivative terms in the asset price, $\tilde{S}$, and variance, $\tilde{\nu}$, as implicit finite-differences and taking the non-linear term as explicit, allowing us to use LU decomposition in which the LU matrices are $\tilde{\tau}$ and q independent, thus only needing to perform the LU decomposition once. We expect $O\left(\Delta\tilde{\tau}, \Delta\tilde{S}^2\right)$ convergence, which we see in table 7.1 is the case.

Table 7.1: Convergence of finite-difference scheme for stochastic volatility model

$n = m$	k	$f_{I-E}\left(\tilde{\tau} = 1, q = 1, \tilde{S} = 5, \tilde{\nu} = 5\right)$	ratio
51	1001	-0.2097362454	NA
101	1001	-0.2076657046	NA
201	1001	-0.2070685831	3.47
401	1001	-0.2069097475	3.76
801	1001	-0.2068696183	3.96
401	251	-0.2068153647	NA
401	501	-0.2068782465	NA
401	1001	-0.2069097475	2
401	2001	-0.206925513	2
401	4001	-0.2069333995	2

$n = m$ and k are the number of space and time points respectively. We have set $\tilde{S}_{max} = \tilde{\nu}_{max} = 10, \tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$. f_{I-E} is the value corresponding to the implicit derivatives-explicit non-linear term finite-difference scheme. The ratio is the ratio of the change in errors for successive grids.

The solution of the value function, f, from (7.14) can be seen in figure 7.1. It can be seen that the value function, f, is increasing in asset price, $\tilde{S}$, as is always the



(a) f against $\tilde{S}$, varying $\tilde{\nu}$

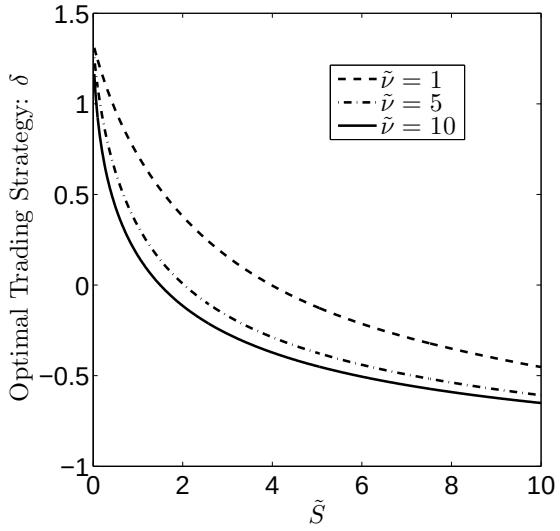
(b) f against $\tilde{\nu}$, varying $\tilde{S}$

Figure 7.1: The above figures shows the value function, f, plotted against $\tilde{S}$ (figure 7.1(a)) and against $\tilde{\nu}$ (figure 7.1(b)). The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

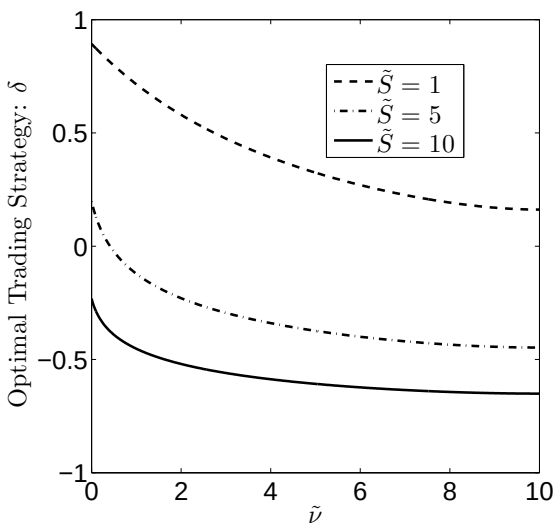
case, and decreasing in the variance, $\tilde{\nu}$. The corresponding optimal trading strategies can be in found in figure 7.2. The properties of the optimal trading strategy mirror that of the value function. As the risk-adjusted asset price increases the trader would like to sell quicker to lock in profit. Similarly, a higher variance represents a riskier portfolio for the trader and given the trader is risk-averse he would have preference over a lower variance. The trader thus opts to sell quicker under higher volatility. This was also observed in chapter 4 for constant volatility when we examined how the solutions varied with respect to the parameters. We found that as we increased the volatility of the geometric Brownian motion the optimal strategy decreased due to the risk-aversion of the trader.

To gain further insight into how the solutions depend on the parameters of the stochastic volatility process we shall consider the optimal trading strategies while keeping all but one parameter constant. We exclude varying $\tilde{\mu}$ and l as this has been done in the constant volatility model, examined in chapter 4, and the results found are similar to the stochastic volatility case.

We first investigate altering the speed of reversion, κ. For larger values of κ the speed of reversion to the long-term mean, $\tilde{\theta}$, is faster. It can be seen in figure 7.3, for $\kappa = \{0.05, 0.75, 2.0\}$, that when $\tilde{\nu} < \tilde{\theta}$ the variance is expected to increase and as such a lower value of κ is preferred so it increases slowly. In contrast, when $\tilde{\nu} > \tilde{\theta}$

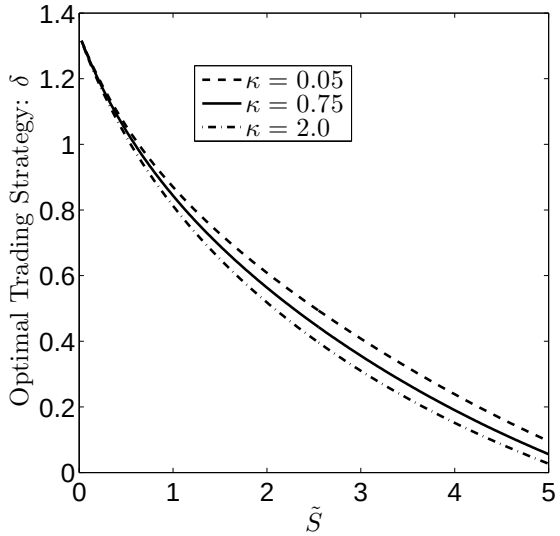


(a) δ^* against $\tilde{S}$, varying $\tilde{\nu}$

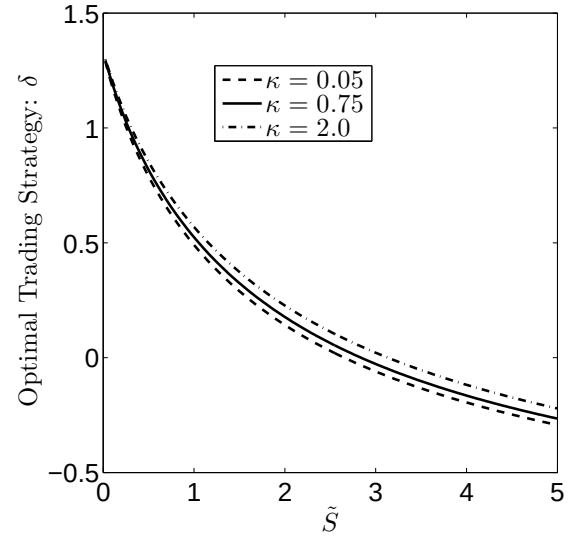


(b) δ^* against $\tilde{\nu}$, varying $\tilde{S}$

Figure 7.2: The above figures shows the optimal trading strategy, δ^*, plotted against $\tilde{S}$ (figure 7.2(a)) and against $\tilde{\nu}$ (figure 7.2(b)). The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.



(a) δ^* against $\tilde{S}$, varying κ with $\tilde{\nu} < \tilde{\theta}$



(b) δ^* against $\tilde{S}$, varying κ with $\tilde{\nu} > \tilde{\theta}$

Figure 7.3: The above figures shows the optimal trading strategy, δ^*, plotted against $\tilde{S}$ for various values of κ with $\tilde{\nu} = 0.25 < \tilde{\theta}$ (figure 7.3(a)) and $\tilde{\nu} = 2.5 > \tilde{\theta}$ (figure 7.3(b)). The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = \{0.05, 0.75, 2.0\}, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

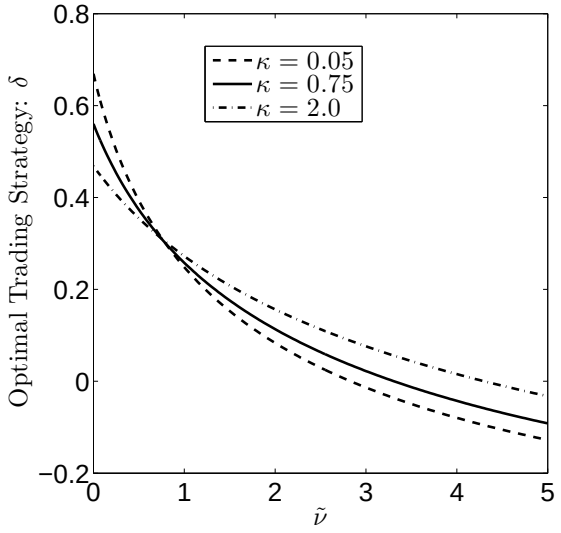
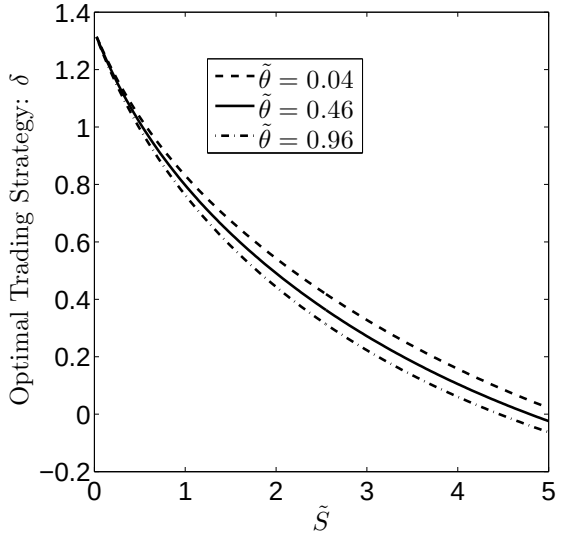


Figure 7.4: The above figure shows the optimal trading strategy, δ^*, plotted against $\tilde{\nu}$ for various values of κ, with $\tilde{S} = 2.5$. The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = \{0.05, 0.75, 2.0\}, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

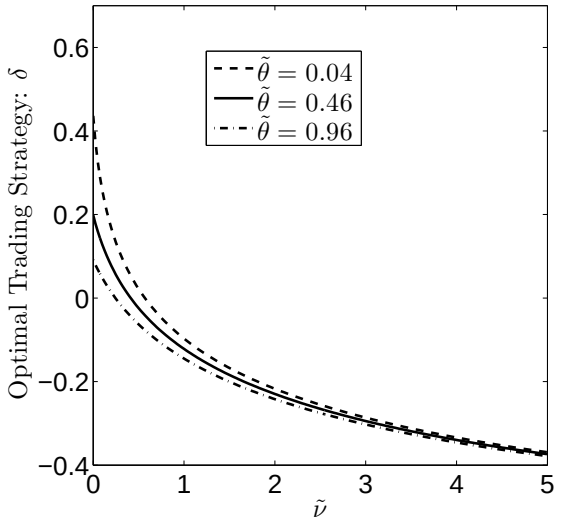
the variance is expected to decrease and as such a higher value of κ is preferred so it decreases quickly. It is also portrayed in figure 7.4 that as $\tilde{\nu}$ increases past $\tilde{\theta}$ the curves for various speeds of reversion, κ, intersect, indicating the change in preference.

Next we investigate the change in the long-term mean level of variance, $\tilde{\theta}$. This can be seen in figure 7.5, with each figure having three curves for the long-term mean, corresponding to $\tilde{\theta} = \{0.04, 0.46, 0.92\}$. It can be seen that for all values of the asset price, $\tilde{S}$, and variance, $\tilde{\nu}$, it is favourable to have the smallest value of the long-term mean, $\tilde{\theta}$. This is an expected result as smaller volatility is preferred over higher volatility for a risk-averse trader.

Next we investigate the change in correlation parameter, ρ, between the asset price, $\tilde{S}$, and the variance, $\tilde{\nu}$. Figures 7.6(a) and 7.6(b) each have three curves corresponding to different values of the correlation parameter, $\rho = \{-0.64, 0.00, +0.64\}$. In both we can see there is a switch in preference of whether a positive or negative correlation is preferred. A positive correlation is preferred when the variance, $\tilde{\nu}$, is small and expected to increase back to its long-term mean. If the variance, $\tilde{\nu}$, is expected to increase then it is preferable to have the asset price, $\tilde{S}$, correlated positively to the variance as so it is also expected that the asset will increase in value. The opposite is true for large variance, $\tilde{\nu}$. A negative correlation is preferred when $\tilde{\nu}$ is large and expected to decrease back to its long-term mean. If the correlation, ρ, is negative

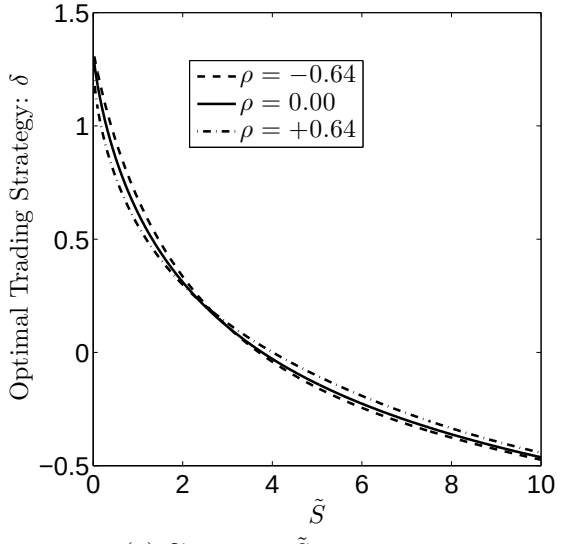


(a) δ^* against $\tilde{S}$, varying θ

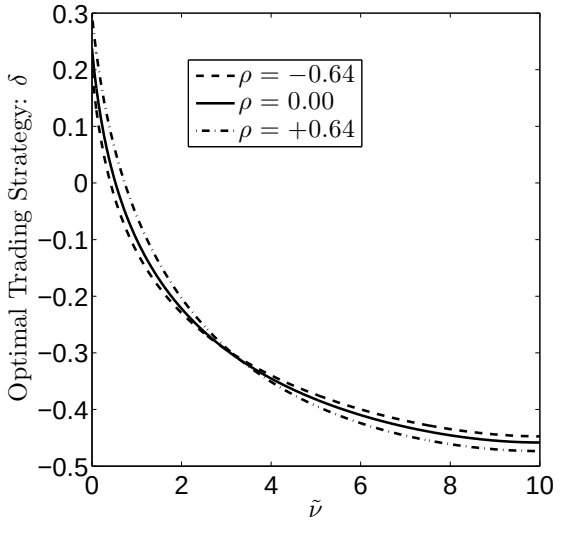


(b) δ^* against $\tilde{\nu}$, varying θ

Figure 7.5: The above figures shows the optimal trading strategy, δ^*, for various values of $\tilde{\theta}$ plotted against $\tilde{S}$ (figure 7.5(a), with $\tilde{\nu} = 0.26$) and against $\tilde{\nu}$ (figure 7.5(b), with $\tilde{S} = 5$.). The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = \{0.04, 0.46, 0.96\}, \sigma = 2.78$ and $\rho = -0.64$.

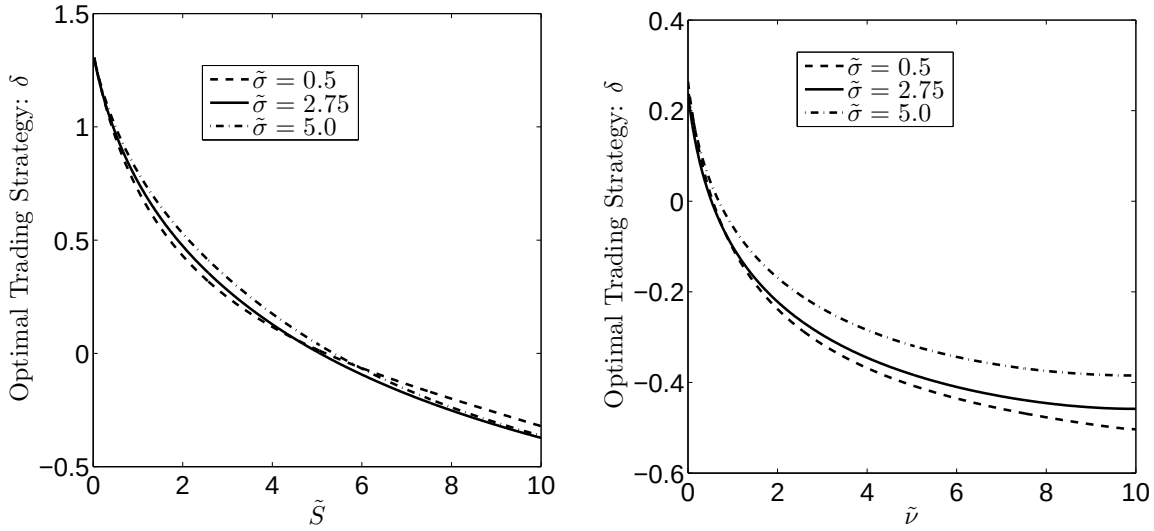


(a) δ^* against $\tilde{S}$, varying ρ



(b) δ^* against $\tilde{\nu}$, varying ρ

Figure 7.6: The above figures shows the optimal trading strategy, δ^*, for various values of ρ plotted against $\tilde{S}$ (figure 7.6(a), with $\tilde{\nu} = 1.25$) and against $\tilde{\nu}$ (figure 7.6(b), with $\tilde{\nu} = 5.00$). The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = \{-0.64, 0.0, +0.64\}$.



(a) δ^* against $\tilde{S}$, varying σ

(b) δ^* against $\tilde{\nu}$, varying σ

Figure 7.7: The above figures shows the optimal trading strategy, δ^*, for various values of σ plotted against $\tilde{S}$ (figure 7.7(a), with $\tilde{\nu} = 0.5$) and against $\tilde{\nu}$ (figure 7.7(b), with $\tilde{S} = 5$). The parameter values used are $\tilde{\tau} = 1, l = 10, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = \{0.5, 2.78, 5.0\}$ and $\rho = 0.0$.

and if the variance, $\tilde{\nu}$, is expected to decrease then it is expected the asset price, $\tilde{S}$, will increase and this is favourable to the holder of that asset. If the correlation was positive then it would be expected that the asset would decrease in value and this is obviously not favourable to the trader. It is generally accepted that asset prices and volatility have a negative correlation. As asset prices fall a company becomes riskier, and hence more volatile, as the relative value of debt to equity rises. This is known as the leverage effect (see Black, 1976; Christie, 1987). However, research has also shown that declines in stock prices are accompanied by larger increases in volatility than the decline in volatility that accompanies rising stock markets (see Nelson, 1991; Engle and Ng, 1993).

Investigation of the effect of σ, the volatility of the variance, proved quite interesting. We investigate three values, $\sigma = \{0.5, 2.94, 5.0\}$, which can be seen in figure 7.7. It was found that in certain regimes a smaller σ was preferred, such as when the variance was below the long-term mean, $\tilde{\theta}$, as being risk-averse the trader prefers more certainty. However, when the variance, $\tilde{\nu}$, was above the long-term mean, $\tilde{\theta}$, a higher value of σ resulted in a higher value function for the trader as the variance would change more rapidly. These results are consistent with those found in other areas of finance where stochastic volatility is concerned and the reader is referred to

Yang (2010) for an analysis of option pricing under a Heston (1993) framework. There is little literature regarding the noise of volatility, despite the vast amount of literature on volatility itself. Wilmott (2007) suggests "the volatility of volatility is large" without further development or explanation. There is a volatility index that is traded on the Chicago Board of Options called the CBOE Volatility Index, or VIX for short. It a popular measure of the implied volatility of the S&P 500 index options. It represents one measure of the annualised market's expectation of stock market volatility over the next 30 day period. Table 7.2 represents some of the statistical properties of the volatility of variance as calculated using the VIX for 22 years of data, from Feb. 9, 1990 through Feb. 3, 2012 (see Moran, 2012). This data represents the statistical properties of α, see (7.2). α is the volatility of $\sqrt{\nu(t)}$ whereas σ is the volatility of $\nu(t)$. Using Itô's lemma it is found that $\sigma = 2\alpha$.

Table 7.2: Properties for the VIX

	α	$\sigma\ (=2\alpha)$
Maximum	2.45	4.90
Minimum	0.344	0.688
Average	0.931	1.862
Median	0.862	1.724

7.2 Small-time-to-termination solution

We found from the numerics that there was a lot of interesting behaviour occurring around $\tilde{\tau} = 0$, which is the termination time. This is similar to the case of constant volatility, as discussed in section 4.2. For that reason we shall investigate a small-time-to-termination solution.

From the previous analysis of similar models we expect to be able to derive a small-time-to-termination solution which is fully analytic. To examine the small-$\tilde{\tau}$ solution we use a perturbation series of the form

$$f\left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right) = f_0\left(q, \tilde{S}, \tilde{\nu}\right) + \tilde{\tau} f_1\left(q, \tilde{S}, \tilde{\nu}\right) + \tilde{\tau}^2 f_2\left(q, \tilde{S}, \tilde{\nu}\right) + O\left(\tilde{\tau}^3\right), \tag{7.22}$$

which is a power series expansion in $\tilde{\tau}$ of $f\left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right)$ up to $O(\tilde{\tau}^3)$. We found using a solution of this form successful in previous chapters, such as the constant volatility case in chapter 4 (see (4.55)) and with multiple underlyings in chapter 6 (see (6.26))

We can find a fully analytic expression for each term of (7.22) by substituting (7.22) into (7.14) and matching the terms of similar order in $\tilde{\tau}$. Before doing this, we examine (7.22) at $\tilde{\tau} = 0$ to which we have

$$f\left(\tilde{\tau} = 0, q, \tilde{S}, \tilde{\nu}\right) = f_0\left(q, \tilde{S}, \tilde{\nu}\right) = -e^{-q\tilde{S}}, \tag{7.23}$$

which is from the initial condition (7.15). Upon substituting (7.22) into (7.14), we find

$$f_1\left(q, \tilde{S}, \tilde{\nu}\right) = \left(-\tilde{\mu}\tilde{S}q + \frac{1}{2}\tilde{\nu}\tilde{S}^2q^2 - \frac{\tilde{S}}{\tilde{S}+l}\left(\frac{l}{\tilde{S}+l}\right)^{\frac{l}{\tilde{S}}}\right) f_0\left(q, \tilde{S}\right), \tag{7.24}$$

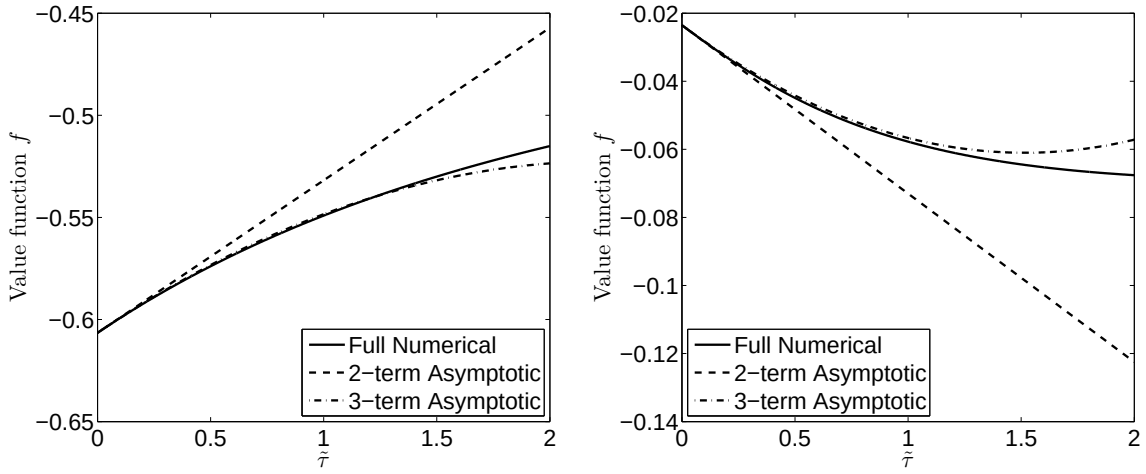
by collecting the $O\left(\tilde{\tau}^0\right)$ terms and

$$\begin{aligned} f_2\left(q, \tilde{S}, \tilde{\nu}\right) = \frac{1}{2}\Bigg(&\tilde{\mu}\tilde{S}f_{1\tilde{S}} + \kappa\left(\tilde{\theta} - \tilde{\nu}\right) f_{1\tilde{\nu}} + \frac{1}{2}\tilde{\nu}\tilde{S}^2 f_{1\tilde{S}\tilde{S}} + \frac{1}{2}\sigma\tilde{\nu}^2 f_{1\tilde{\nu}\tilde{\nu}} + \rho\tilde{\nu}\sigma\tilde{S}f_{1\tilde{S}\tilde{\nu}} \\ &- \frac{1}{\tilde{S}+l}\left(\frac{l}{\tilde{S}+l}\right)^{\frac{l}{\tilde{S}}}\left(\left(l+\tilde{S}\right) f_1\left(q, \tilde{S}, \tilde{\nu}\right) - l f_1\left(q-1, \tilde{S}, \tilde{\nu}\right) e^{-\tilde{S}}\right)\Bigg), \end{aligned} \tag{7.25}$$

by collecting the $O\left(\tilde{\tau}^1\right)$ terms. Given the expression (7.24) of $f_1\left(q, \tilde{S}, \tilde{\nu}\right)$, the derivatives of $f_1\left(q, \tilde{S}, \tilde{\nu}\right)$ can be calculated analytically and hence (7.25) can be expressed analytically. We neglect the analytic solutions of the derivatives due to the complexity of the equations but they are simple extensions to (4.58) and (4.59) which were found in chapter 4 for the case of a single underlying asset with constant volatility.

A comparison of the accuracy for the two-term and three-term asymptotic expansion can be seen in figure 7.8, in which we have shown a comparison for two values of the asset price, small $\tilde{S}$ (figure 7.8(a)) and larger $\tilde{S}$ (figure 7.8(b)). We examined these at a value of the variance, $\tilde{\nu}$, near the long term mean, $\tilde{\theta}$. It can be seen that the three-term asymptotic expansion is a strong approximation to the full-numerical solution for $\tilde{\tau} \sim O\left(1\right)$. It is of interest to observe that, as was the case under constant volatility, the value function is increasing in $\tilde{\tau}$ in the small $\tilde{S}$ regime (figure 7.8(a)), while decreasing in $\tilde{\tau}$ for larger values of $\tilde{S}$ (figure 7.8(b)).

It is also worth noting that there are restrictions to how well the expansion approximates the full-numerical solution. The larger we make the asset price, $\tilde{S}$, the faster the expansion deviates as an approximation for the full-numerical solution. This is also the case as the variance, $\tilde{\nu}$, is increased. Figure 7.9 shows the same figures as figure 7.8 but for a larger value of the variance, with $\tilde{\nu}$ doubling from $\tilde{\nu} = 0.4$ to $\tilde{\nu} = 0.8$. It can be seen that for both small $\tilde{S}$ (figure 7.9(a), $\tilde{S} = 0.5$) and larger $\tilde{S}$

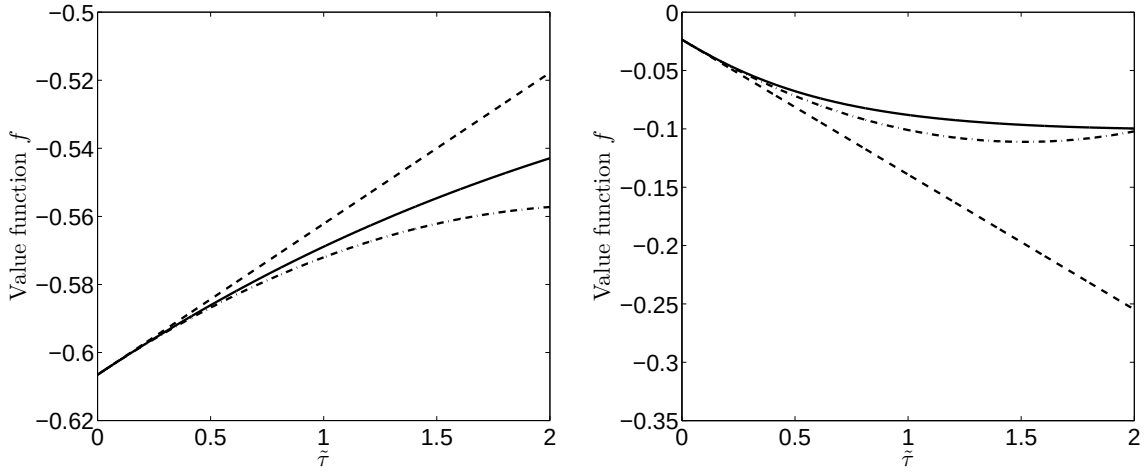


(a) $\tilde{S} = 0.5, \tilde{\nu} = 0.4$ (b) $\tilde{S} = 4, \tilde{\nu} = 0.4$

Figure 7.8: The above figures shows the comparison of the full-numerical solution and the two-term and three-term asymptotic expansion of the value function, for $\tilde{S} = 0.5$ (figure 7.8(a)) and $\tilde{S} = 4$ (figure 7.8(b)). The parameter values used are $l = 1, \tilde{\mu} = 0.05, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 0.5$ and $\rho = -0.64$.

(figure 7.9(b), $\tilde{S} = 4$) the asymptotic solution deviates faster from the full-numerical solution in comparison to the smaller value of the variance, $\tilde{\nu}$. Finally, relatively large parameter values can lead to a quicker deviation of the asymptotic expansion solution from the full-numerical solution, hence the use of the smaller σ in figure 7.8.

This section has investigated when the time remaining until termination is small. We shall now change from examining a small-time-to-termination solution to examining the case in which the time horizon is so large that we can approximate it as infinite and thus the change in time in negligible. This is known as the perpetual (steady-state) solution which was investigated thoroughly under constant volatility (see section 4.3) and with multiple underlyings (see section 6.4). Both investigations gave in-depth insight into the solution topology and provided parameter constraints for the existence of solutions. Some of these constraints involved the (constant) volatility (see, for example, (4.65)) and it is thus of interest to see if similar constraints exist under a stochastic volatility framework.



(a) $\tilde{S} = 0.5, \tilde{\nu} = 0.8$ (b) $\tilde{S} = 4, \tilde{\nu} = 0.8$

Figure 7.9: The above figures shows the comparison of the full-numerical solution and the two-term and three-term asymptotic expansion of the value function, for $\tilde{S} = 0.5$ (figure 7.9(a)) and $\tilde{S} = 4$ (figure 7.9(b)), for larger $\tilde{\nu}$ in relation to figure 7.8. The parameter values used are $l = 1, \tilde{\mu} = 0.05, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 0.5$ and $\rho = -0.64$.

7.3 Perpetual case

Let us now examine a perpetual-type solution which was shown to be of asymptotic interest in previous chapters. Assuming a steady-state, and thus setting

$$\frac{\partial f}{\partial \tilde{\tau}}\left(\tilde{\tau}, q, \tilde{S}, \tilde{\nu}\right) = 0 \quad \text{as} \quad \tilde{\tau} \to \infty,$$

the non-linear PDE (7.14) reduces to

$$\tilde{\mu}\tilde{S}f_{\tilde{S}} + \frac{1}{2}\tilde{\nu}\tilde{S}^2 f_{\tilde{S}\tilde{S}} + \kappa\left(\tilde{\theta} - \tilde{\nu}\right) f_{\tilde{\nu}} + \frac{1}{2}\sigma^2\tilde{\nu}f_{\tilde{\nu}\tilde{\nu}} + \rho\sigma\tilde{\nu}\tilde{S}f_{\tilde{S}\tilde{\nu}}$$
$$- \frac{e^l\tilde{S}f}{\tilde{S}+l}\left(\frac{lf\left(q,\tilde{S},\tilde{\nu}\right)}{\left(\tilde{S}+l\right)f\left(q-1,\tilde{S},\tilde{\nu}\right)}\right)^{\frac{l}{\tilde{S}}} = 0, \tag{7.26}$$

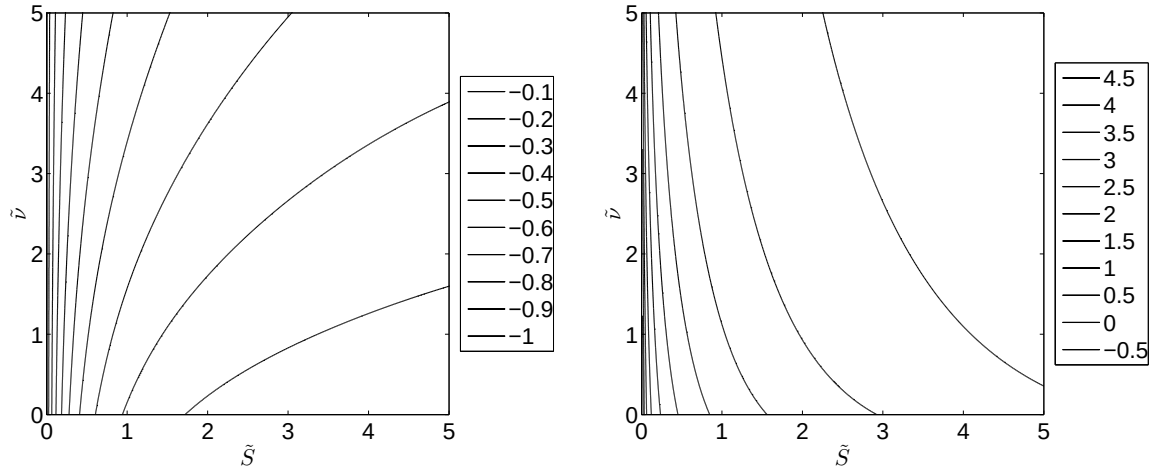
with $f = f\left(q, \tilde{S}, \tilde{\nu}\right)$ and initial condition

$$f\left(q = 0, \tilde{S}, \tilde{\nu}\right) = -1. \tag{7.27}$$

The optimal strategy for the steady-state solution takes the form

$$\delta^*\left(q, \tilde{S}, \tilde{\nu}\right) = \frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)f\left(q-1,\tilde{S},\tilde{\nu}\right)}{lf\left(q,\tilde{S},\tilde{\nu}\right)}\right) - 1. \tag{7.28}$$

To solve (7.26) we require two boundary conditions in each of the two semi-infinite domains of the asset price, $\tilde{S}$, and variance, $\tilde{\nu}$. These are analogous to the time-dependent case above, given by (7.17),(7.18),(7.20) and (7.21), but with the time



(a) Contour of f (b) Contour of δ^*

Figure 7.10: Contour plot of stochastic volatility steady-state solution for value function (figure 7.10(a)) and optimal trading strategy (figure 7.10(b)). The parameter values used are $l = 1, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

derivative set to zero. We solve (7.26) using an iterative numerical method similar to that described in section 6.4. Figure 7.10 shows contour plots of the value function (figure 7.10(a)) and the optimal trading strategy (figure 7.10(b)). In figure 7.10(a) we can see the contours curve as the variance, $\tilde{\nu}$, increases indicating the value function is a decreasing function of $\tilde{\nu}$. This is confirmed in figure 7.10(b), as we see the optimal trading strategy is a decreasing function of the variance, $\tilde{\nu}$.

7.3.1 Limit of small asset price

Let us now consider the limit of (7.26) when the asset price is small, $\tilde{S} \to 0$, which previously has shown to give us insight into the parameter regimes to which solutions exist. In the single asset case with constant volatility it was found that for a steady-state to exist the constraint

$$\tilde{\mu} < \frac{\tilde{\sigma}^2}{2}, \tag{7.29}$$

had to be satisfied, where $\tilde{\sigma}$ is the volatility of the geometric Brownian motion process driving the asset price, i.e. $\tilde{\nu}(\tau) = \tilde{\sigma}^2 \ \forall \ \tau$. This is not to be confused with σ, which in this chapter is the volatility of the variance, $\tilde{\nu}(\tau)$. Given now we have stochastic volatility it is of interest to examine if similar constraints exist.

In the limit as $\tilde{S} \to 0$ asymptotic balancing suggests the non-linear term is negligible and thus a solution of the form

$$f\left(q, \tilde{S}, \tilde{\nu}\right) = -1 + c\left(q\right) \tilde{S}^{\beta} e^{\alpha \tilde{\nu}}, \tag{7.30}$$

solves the linear PDE

$$\tilde{\mu}\tilde{S} f_{\tilde{S}} + \frac{1}{2}\tilde{\nu}\tilde{S}^2 f_{\tilde{S}\tilde{S}} + \kappa\left(\tilde{\theta} - \tilde{\nu}\right) f_{\tilde{\nu}} + \frac{1}{2}\sigma^2 \tilde{\nu} f_{\tilde{\nu}\tilde{\nu}} + \rho\sigma\tilde{\nu}\tilde{S} f_{\tilde{S}\tilde{\nu}} = 0, \tag{7.31}$$

with initial condition (7.16). Substituting (7.30) into (7.31) we obtain

$$\alpha = -\frac{\tilde{\mu}}{\kappa\tilde{\theta}}\beta \tag{7.32}$$

and

$$\frac{1}{2}\beta\left(\beta - 1\right) + \rho\sigma\beta\alpha + \frac{1}{2}\sigma^2\alpha^2 - \kappa\alpha = 0, \tag{7.33}$$

by collection the $O\left(\tilde{\nu}^0\right)$ and $O\left(\tilde{\nu}^1\right)$ terms respectively. For the solution to be bounded we need $\beta > 0$ and $\alpha < 0$. Solving (7.33) gives

$$\beta = \frac{\left(1 - \frac{2\tilde{\mu}}{\tilde{\theta}}\right)}{\left(1 - \frac{2\rho\sigma\tilde{\mu}}{\kappa\tilde{\theta}} + \left(\frac{\sigma\tilde{\mu}}{\kappa\tilde{\theta}}\right)^2\right)}, \tag{7.34}$$

which can be substituted into (7.32) to obtain

$$\alpha = -\frac{\frac{\tilde{\mu}}{\kappa\tilde{\theta}}\left(1 - \frac{2\tilde{\mu}}{\tilde{\theta}}\right)}{\left(1 - \frac{2\rho\sigma\tilde{\mu}}{\kappa\tilde{\theta}} + \left(\frac{\sigma\tilde{\mu}}{\kappa\tilde{\theta}}\right)^2\right)}. \tag{7.35}$$

Examining β we see that the denominator of (7.34) is always positive; this is clear when $\rho = 0$. At the extremes, when $\rho = \pm 1$ the denominator of (7.34) becomes

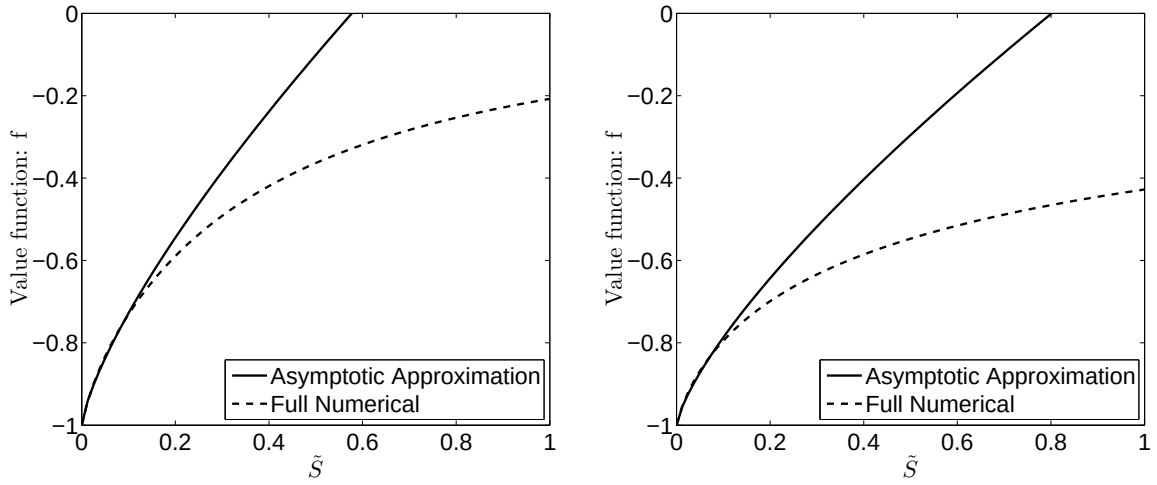
$$\left(1 \mp \left(\frac{\sigma\tilde{\mu}}{\kappa\tilde{\theta}}\right)\right)^2 > 0.$$

Therefore, from the numerator of (7.34), we have the constraint

$$1 - \frac{2\tilde{\mu}}{\tilde{\theta}} > 0 \Rightarrow \tilde{\mu} < \frac{\tilde{\theta}}{2}, \tag{7.36}$$

which is analogous to the constraint of the constant volatility case (7.29) given $\tilde{\theta}$ is the long-term mean of the variance. For $\alpha < 0$ we need $\frac{\tilde{\mu}}{\kappa\tilde{\theta}} > 0$, given $\beta > 0$, which sets the restriction

$$\tilde{\mu} > 0, \tag{7.37}$$



(a) $\tilde{\nu} = 0.25$ (b) $\tilde{\nu} = 4.00$

Figure 7.11: Above we compare the small $\tilde{S}$ asymptotic solution against the full-numerical solution. We do this for small $\tilde{\nu}$ (figure 7.11(a)) and relatively large $\tilde{\nu}$ (figure 7.11(b)). The parameter values used are $l = 1, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

which is another parameter constraint, that was not present in the constant volatility framework. Figure 7.11 shows the comparison of the asymptotic solution given by (7.30), for small variance (figure 7.11(a)) and larger variance (figure 7.11(b)). We can see the solution is approximated well in the small asset price, $\tilde{S}$, regime, before diverging away from the solution, for both small and larger variance, $\tilde{\nu}$.

Given β mainly takes non-integer values, singular behaviour is present. In the small asset price limit, $\tilde{S} \to 0$, the optimal trading strategy is found, to leading order, to take the form

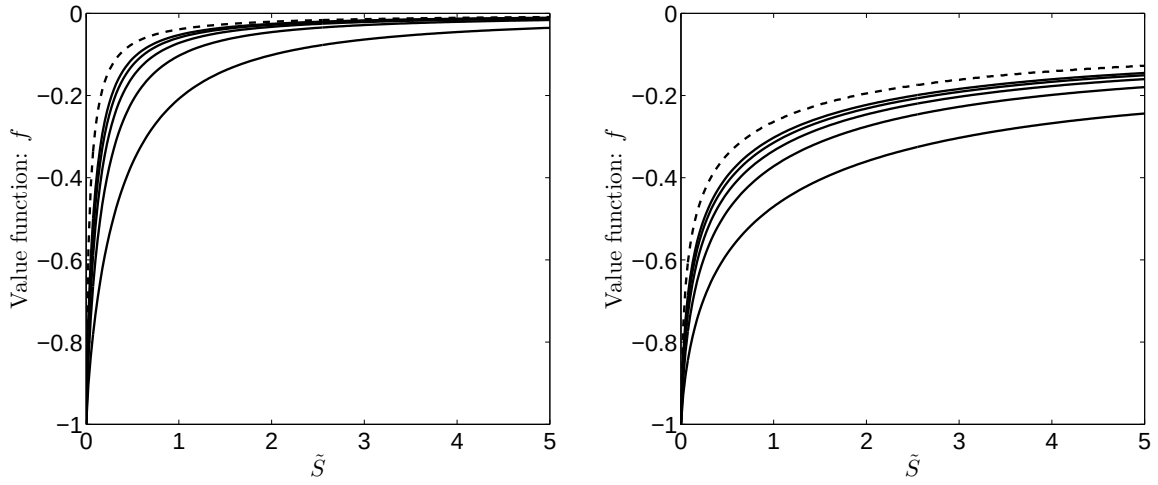
$$\delta^* \approx K(q)\, \tilde{S}^{\beta-1} e^{\alpha\tilde{\nu}}, \tag{7.38}$$

in which $K(q)$ is a decreasing function of q. This is similar to that found under constant volatility (4.67), discussed in section 4.3.1.

7.3.2 Large number of assets

We shall now investigate (briefly) the limit as the number of assets held becomes large, i.e. $q \to \infty$. It is found (confirmed by numerical investigation) that as $q \to \infty$

$$\lim_{q\to\infty} \frac{f\left(q, \tilde{S}, \tilde{\nu}\right)}{f\left(q-1, \tilde{S}, \tilde{\nu}\right)} \to 1, \tag{7.39}$$



(a) $\tilde{\nu} = 0.25$ (b) $\tilde{\nu} = 5.00$

Figure 7.12: Above we compare the $q \to \infty$ solution against the solutions of increasing values of q. We plot $q = \{1, \ldots, 5\}$ which are represented by the increasing solid lines and converging to the $q \to \infty$ solution, represented by the broken line. We do this for small $\tilde{\nu}$ (figure 7.12(a)) and large $\tilde{\nu}$ (figure 7.12(b)). The parameter values used are $l = 1, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

which when substituted into (7.26) yields

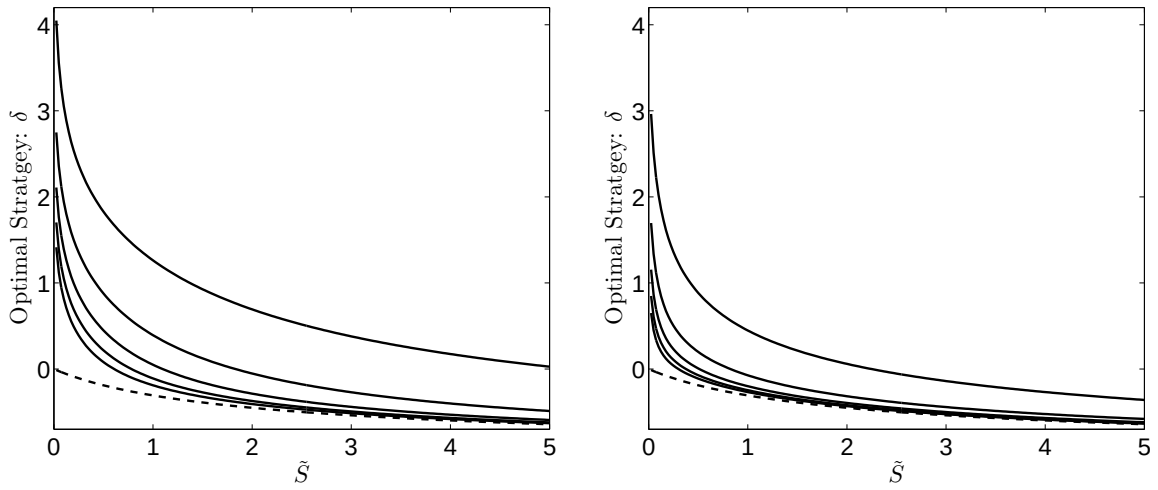
$$\tilde{\mu}\tilde{S}f_{\tilde{S}} + \frac{1}{2}\tilde{\nu}\tilde{S}^2 f_{\tilde{S}\tilde{S}} + \kappa\left(\tilde{\theta} - \tilde{\nu}\right) f_{\tilde{\nu}} + \frac{1}{2}\sigma^2\tilde{\nu}f_{\tilde{\nu}\tilde{\nu}} + \rho\sigma\tilde{\nu}\tilde{S}f_{\tilde{S}\tilde{\nu}} - \frac{e^l\tilde{S}f}{\tilde{S}+l}\left(\frac{l}{\left(\tilde{S}+l\right)}\right)^{\frac{l}{\tilde{S}}} = 0, \tag{7.40}$$

with $f = f(\tilde{S}, \tilde{\nu})$. This results in an optimal trading strategy which takes the form

$$\delta^*\left(\tilde{S}, \tilde{\nu}\right) = \frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)}{l}\right) - 1, \tag{7.41}$$

which we notice is independent of the variance, $\tilde{\nu}$. Intuitively this makes sense, given that a trader with a large number of assets will receive a large amount of cash regardless of the asset price. The variance of the asset is thus not of concern to him.

To solve (7.40) we use a second-order finite-difference scheme, implementing the same boundary conditions as used for the $q \sim O(1)$ case as discussed above in section 7.3. We examine the convergence of increasing the inventory, q, towards the large inventory, $q \to \infty$, solution. This can be seen in figure 7.12 for the value function and figure 7.13 for the optimal trading strategies. We have included two figures in each, one for small variance, $\tilde{\nu}$, and one for relatively large $\tilde{\nu}$. In each figure, we have curves



(a) $\tilde{\nu} = 0.25$ (b) $\tilde{\nu} = 5.00$

Figure 7.13: Above we compare the $q \to \infty$ solution against the solutions of increasing values of q for the optimal trading strategy, corresponding to those of figure 7.12. We plot $q = \{1, \ldots, 5\}$ which are represented by the increasing solid lines and converging towards the $q \to \infty$ solution, represented by the broken line. We do this for small $\tilde{\nu}$ (figure 7.13(a)) and relatively large $\tilde{\nu}$ (figure 7.13(b)). The parameter values used are $l = 1, \tilde{\mu} = 0.02, \kappa = 0.75, \tilde{\theta} = 0.46, \sigma = 2.78$ and $\rho = -0.64$.

representing increasing values of inventory, q, (solid curves) which converge towards the broken curve representing the large inventory, $q \to \infty$, solution. The values of the inventory we use are $q = \{1, \ldots, 5\}$. This convergence corroborates with our approximation of the ratio of the value function for subsequent inventory, q, values, given by (7.39). The singular behaviour (7.38) can be observed in figure 7.13 for the optimal trading strategies for finite q as $\tilde{S} \to 0$. As q increases it tends to the $q \to \infty$ solution which is not singular as $\tilde{S} \to 0$, as from (7.41) we have

$$\delta^* \left(\tilde{S} \to 0, \tilde{\nu} \right) = \lim_{\tilde{S} \to 0^+} \left(\frac{1}{\tilde{S}} \ln \left(\frac{\left(\tilde{S} + l \right)}{l} \right) - 1 \right) = \frac{1}{l} - 1. \tag{7.42}$$

7.4 Summary

In this chapter we returned to investigating the optimal trading strategy of a trader who holds an amount of a single asset, but now the volatility of the asset is driven by a stochastic process, correlated to the stochastic process driving the asset price. This is the Heston (1993) framework in which the asset follows geometric Brownian motion and the variance is driven by a stochastic mean-reverting process. We begin by deriving an HJB equation and providing an ansatz and non-dimensionalisation which

reduces this five-dimensional HJB equation to a four-dimensional non-linear PDE with explicit non-linear term in which the number of input parameters reduced by two.

We first examined the PDE numerically before examining the solution in several limits. We used a similar LU decomposition scheme as used for the multiple asset case, see chapter 6, in which the LU matrices only had to be calculated once, greatly reducing computational time. We examined the optimal trading strategies under various parameter regimes to get a full understanding of the effect of the mean-reverting stochastic volatility on the trading strategies.

After producing some interesting numerical solutions we let the numerics, and indeed the single asset case, guide us to which limits are of interest. We examined a perpetual-type solution in which the time-to-termination tends to infinity. This led to an interesting small $\tilde{S}$ solution which explained the singular behaviour that was developing as $\tilde{S} \to 0$, and also provided us with parameter constraints necessary for a steady-state solution to exist. We had already found an analogous constraint for the constant volatility case (section 4.3) so it was of interest to see if a similar constraint existed when the volatility was a variable rather than a parameter.

Finally we examined the limit as the number of assets held, q, went to infinity. This is beneficial as it reduces the problem from a non-linear PDE to a linear PDE which is numerically easier and quicker to solve.

Chapter 8

Extension of Single Underlying: Various Intensity Functions and Trading with Market Orders

In this chapter we focus on three extensions to chapter 4. The first is imposing a constraint on the admissible strategy space that the control may take for an exponential intensity. The second is to replace the negative exponential intensity function with a power intensity function. The third is to return to the original setting of chapter 4 but allow for trading using both limit orders and market orders. Doing such analysis exemplifies the significance (and usefulness) of solving the underlying problem using asymptotic and numerical methods, the latter of which is prominent in this chapter.

In previous chapters we have focused on the intensity function of the Poisson process having a negative exponential form, with the values taken by the optimal strategy naturally bounded from below, with $\delta \in (-1, \infty)$. Having $\delta < 0$ implies we are selling at a discount. It can be argued that in crossing the reference price we are being charged a commission. This could correspond to the maker-taker fee (see section 1.1) which occurs in practice on some exchanges, including the NYSE and NASDAQ. It could also correspond to temporary market impact (see chapter 3) as introduced by Almgren and Chriss (2001). Guéant and Lehalle (2013), who extend their original model in Guéant et al. (2012b) by constraining δ, argue that setting

$\delta < 0$ is equivalent to using marketable limit orders[1]. To avoid marketable limit orders/crossing the reference price, we introduce a constrained optimisation problem, using a negative exponential intensity function with $\delta > \delta_{min}$ (with $\delta_{min} = 0$ being the most obvious case).

An alternative model, implemented by Bayraktar and Ludkovski (2012) in a risk-neutral setting and backed empirically by the work of Gopikrishnan et al. (2000), Maslov and Mills (2001) and Gabaix et al. (2006) to name a few, is to use a power intensity function. Bayraktar and Ludkovski (2012) argue that if selling at the reference price we should be able to execute almost certainly, and this is not the case using the exponential intensity function. A power intensity function, with the power taking a negative value, would tend to infinity as $\delta \to 0$, and they argue that this is a more realistic assumption. Not only does this result in interesting and contrasting trading strategies, given the instantaneous liquidity available as $\delta \to 0$, but also results in an interesting problem from a mathematical perspective. As $\delta \to 0$ the Hamiltonian may tend to infinity resulting in a singular stochastic control problem. The problem can thus not be solved in a similar manner as the negative exponential intensity case.

Thus far in this thesis we have considered a trader who can trade explicitly using only limit orders[2]. Posting limit orders has the usual trade-off between price risk and non-execution risk. Posting market orders eliminates the latter risk but the trader must pay for this privilege. The price the trader pays comes in the form of selling the asset for a discount which in practice could represent market impact. Although we are viewing immediate execution as a market order, it could also be interpreted as selling to a central-risk desk in which a premium must be paid for this service. We thus study the strategy of a trader who can post both limit orders and market orders throughout the trading period. In our framework, given the trader has the ability to execute immediately, the problem will be formulated as a free-boundary problem and can be viewed similar to that as an American option which allows for early exercise. This problem is of particular interest given the trader is risk-averse and thus before examining the results we expect the trader will take advantage of having the ability

[1]Harris (2002) defines a marketable limit order as a buy order with a price at or above the lowest offer in the market or a sell order with a price at or below the highest bid in the market

[2]Using the best bid as a reference price, liquidation at the terminal time, or under power intensity as $\delta \to 0$, could be interpreted as trading with market orders.

to early execute (using market orders), resulting in solutions contrasting to that of chapter 4.

8.1 Constraining the trading strategy

In previous chapters δ was allowed to take any value (however naturally it was bounded by $\delta \in (-1, \infty)$). In this section we shall impose a constraint $\delta \geq \delta_{min} \geq 0$, which avoids selling for a price lower than the reference price (although any value could be set for $\delta_{\min}$). Firstly, we will briefly recap the model.

8.1.1 Formulating the problem

We begin with an initial formulation as derived in chapter 4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration, $(\mathcal{F}_t, t \in [0, T])$. We assume the asset reference price $S(t)$ follows a geometric Brownian motion

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \tag{8.1}$$

with μ as the constant relative drift, σ as the constant relative volatility and $W(t)$ as a Wiener process which is $\mathcal{F}_t$ measurable. The trader posts orders into the ask side of the limit-order book for price $S^a(t)$ which is $\delta = \delta(t, X, q, S) \geq \delta_{min}$ percent greater than the reference price $S(t)$

$$S^a(t) = S\left(t\right)\left(1 + \delta\right). \tag{8.2}$$

Asset sales and the trader's wealth follow the same Poisson process, $N(t)$, with time-dependent intensity, which is $\mathcal{F}_t$ measurable and independent of $W(t)$:

$$dq(t) = -dN(t), \tag{8.3}$$

and

$$dX(t) = S(t)\left(1 + \delta\right)dN(t), \tag{8.4}$$

accordingly. Therefore, when a jump (trade) occurs, the values of $q(t)$ and $X(t)$ change simultaneously, according to (8.3) and (8.4) respectively. As previous, $N(t)$ has intensity $\Lambda(\delta)$ which takes the form:

$$\Lambda(\delta) = \lambda e^{-l\delta}, \tag{8.5}$$

for some positive constants λ and l.

The objective is to liquidate this portfolio before some final time T while maximising the terminal time utility which takes the form of a negative exponential function. We define our value function

$$u^{min}(t,X,q,S) = \sup_{\delta(t)\in\mathcal{A}^{min}} \mathbb{E}\left[-e^{-\gamma(X(T)+q(T)S(T))}\right], \tag{8.6}$$

where $\gamma > 0$ is the risk-aversion characterising the trader, $\mathcal{A}^{min}$ is the set of admissible trading strategies bounded below by δ^{min}, and the value function's superscript min referring to the constraint $\delta \geq \delta^{min}$ being imposed.

Given the optimisation problem of (8.6), an HJB equation can be derived by applying the Bellman (1957) principle of optimality and using Itô's lemma:

$$\begin{aligned} &u_t^{min}(t,X,q,S) + \mu S u_S^{min}(t,X,q,S) + \frac{1}{2}\sigma^2 S^2 u_{SS}^{min}(t,X,q,S) \\ &+ \sup_{\delta\geq\delta^{min}} \left[\lambda e^{-l\delta}\left(u^{min}(t,X+S\left(1+\delta\right),q-1,S) - u^{min}(t,X,q,S)\right)\right] = 0, \end{aligned} \tag{8.7}$$

with conditions:

$$u^{min}(T,X,q,S) = -e^{-\gamma(X+qS)}, \tag{8.8}$$

$$u^{min}(t,X,0,S) = -e^{-\gamma X}. \tag{8.9}$$

The derivation of (8.7) is similar to that of (4.7) which is shown in section 4.1.1, and a similar verification theorem (also described in section 4.1.1) can then be used to show that the solution obtained from solving (8.7) is in fact the solution of (8.6).

8.1.2 Reduction of problem

Using the same techniques as in previous chapters (see section 4.1.2), we assume an ansatz solution (4.29)

$$u^{min}\left(t,X,q,S\right) = e^{-\gamma X} f^{min}\left(\tau,q,S\right), \tag{8.10}$$

and use the change of variables (4.40)

$$\tilde{\tau} = \lambda\tau, \quad \tilde{S} = S\gamma, \quad \tilde{\mu} = \frac{\mu}{\lambda}, \quad \tilde{\sigma} = \frac{\sigma}{\sqrt{\lambda}}, \tag{8.11}$$

where we also use a change in the time variable $\tau = T - t$, so that we are now solving forward in $\tilde{\tau}$ rather than backward in t. Using this form of the solution we obtain:

$$-f^{min}_{\tilde{\tau}}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\mu}\tilde{S}f^{min}_{\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f^{min}_{\tilde{S}\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)$$
$$+\sup_{\delta\geq\delta^{min}}\left[e^{-l\delta}\left(e^{-\tilde{S}(1+\delta)}f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)-f^{min}\left(\tilde{\tau},q,\tilde{S}\right)\right)\right]=0, \tag{8.12}$$

with

$$f^{min}\left(0,q,\tilde{S}\right)=-e^{-q\tilde{S}}, \tag{8.13}$$
$$f^{min}\left(\tilde{\tau},0,\tilde{S}\right)=-1. \tag{8.14}$$

Let us define the Hamiltonian

$$\begin{aligned}H^{min}\left(\tilde{\tau},q,\tilde{S},f^{min}\right)&=\sup_{\delta\geq\delta^{min}}\left[e^{-l\delta}\left(e^{-\tilde{S}(1+\delta)}f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)-f^{min}\left(\tilde{\tau},q,\tilde{S}\right)\right)\right]\\&=\sup_{\delta\geq\delta^{min}}L^{min}\left(f,\tilde{\tau},q,\tilde{S},\delta\right).\end{aligned} \tag{8.15}$$

For fixed $\left(\tilde{\tau},q,\tilde{S}\right)$, a unique maximiser for $L^{min}\left(f,\tilde{\tau},q,\tilde{S},\delta\right)$ was found in chapter 4 for the unconstrained problem. This means that $L^{min}\left(f,\tilde{\tau},q,\tilde{S},\delta\right)$ is strictly increasing on $(-1,\delta^*]$ and strictly decreasing on $[\delta^*,\infty)$. Hence for the constrained problem the unique maximiser over $\delta\geq\delta^{min}$ is $\max\left(\delta^{min},\delta^*\right)$. We note that this leads to sub-optimal strategies when $\delta^*<\delta^{min}$. We thus treat two cases of the Hamiltonian (8.15), one when $\delta=\delta^*$ and one when $\delta=\delta^{min}$. This results in

$$H^{min}\left(\tilde{\tau},q,\tilde{S},f^{min}\right)=\begin{cases}\sup\limits_{\delta}\left[e^{-l\delta}\left(e^{-\tilde{S}(1+\delta)}f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)-f^{min}\left(\tilde{\tau},q,\tilde{S}\right)\right)\right]\\ \qquad\qquad\text{if } \delta\geq\delta^{min},\\ e^{-l\delta^{min}}\left(e^{-\tilde{S}\left(1+\delta^{min}\right)}f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)-f^{min}\left(\tilde{\tau},q,\tilde{S}\right)\right)\\ \qquad\qquad\text{if } \delta\leq\delta^{min}.\end{cases} \tag{8.16}$$

We know from chapter 4 that in the unconstrained case we can explicitly find the optimal strategy in terms of the value function, which is given by

$$\delta^*=\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)}{lf^{min}\left(\tilde{\tau},q,\tilde{S}\right)}\right)-1. \tag{8.17}$$

This holds for $\delta^*\geq\delta^{min}$ and hence

$$\delta^{min*}=\max\left(\delta^{min},\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)}{lf^{min}\left(\tilde{\tau},q,\tilde{S}\right)}\right)-1\right). \tag{8.18}$$

From (8.18) and (8.16) we have

$$H^{min}\left(\tilde{\tau},q,\tilde{S},f^{min}\right)=\begin{cases}-\dfrac{e^{l}\tilde{S}f^{min}\left(\tilde{\tau},q,\tilde{S}\right)}{\tilde{S}+l}\left(\dfrac{lf^{min}\left(\tilde{\tau},q,\tilde{S}\right)}{\left(\tilde{S}+l\right)f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)}\right)^{\frac{l}{\tilde{S}}} \\ \qquad\qquad \text{if } \delta^{*}\geq\delta^{min}, \\ e^{-l\delta^{min}}\left(e^{-\tilde{S}\left(1+\delta^{min}\right)}f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)-f^{min}\left(\tilde{\tau},q,\tilde{S}\right)\right) \\ \qquad\qquad \text{if } \delta^{*}\leq\delta^{min}.\end{cases} \tag{8.19}$$

The problem is thus to solve

$$-f_{\tilde{\tau}}^{min}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\mu}\tilde{S}f_{\tilde{S}}^{min}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^{2}\tilde{S}^{2}f_{\tilde{S}\tilde{S}}^{min}\left(\tilde{\tau},q,\tilde{S}\right) +H^{min}\left(\tilde{\tau},q,\tilde{S},f^{min}\right)=0, \tag{8.20}$$

with initial conditions (8.13) and (8.14), and H^{min} given by (8.19). The boundary conditions used are the same as those used in chapter 4, namely

$$f\left(\tilde{\tau},q,\tilde{S}=0\right)=-1, \tag{8.21}$$

and

$$\frac{\partial f}{\partial\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\to\infty\right)\to 0. \tag{8.22}$$

8.1.3 Numerical method

To solve (8.20) we use a similar method as used for the unconstrained problem in section 4.1.5, but with a slight modification to allow for the constraint to be implemented.

We use finite-differences to solve (8.20), implementing implicit differences for the derivatives while taking the non-linear term, $H^{min}\left(\tilde{\tau},q,\tilde{S},f^{min}\right)$, explicitly. Dropping the *min* superscript, and letting $f^{min}\left(j\Delta\tilde{\tau},q,i\Delta\tilde{S}\right)=f_{i,j}^{q}$ we can approximate (8.20) as

$$-\frac{f_{i,j+1}^{q}-f_{i,j}^{q}}{\Delta\tilde{\tau}}+\tilde{\mu}i\Delta\tilde{S}\frac{f_{i+1,j+1}^{q}-f_{i-1,j+1}^{q}}{2\Delta\tilde{S}}+\frac{1}{2}\tilde{\sigma}^{2}\left(i\Delta\tilde{S}\right)^{2}\frac{f_{i+1,j+1}^{q}-2f_{i,j+1}^{q}+f_{i-1,j+1}^{q}}{\Delta\tilde{S}^{2}} +H\left(j\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right)=0, \tag{8.23}$$

Table 8.1: Convergence of finite-difference scheme for constrained optimisation

n	m	$f\left(\tilde{\tau}=1,q=1,\tilde{S}=2.5\right)$	ratio
1001	1001	-0.064119345463	NA
2001	1001	-0.064119247917	NA
4001	1001	-0.064119223433	3.984002
8001	1001	-0.064119217263	3.968233
16001	1001	-0.064119215696	3.937878
1001	501	-0.064120717188	NA
1001	1001	-0.064119345463	NA
1001	2001	-0.064118660074	2.001382
1001	4001	-0.064118317498	2.000691
1001	8001	-0.064118146240	2.000346

n and m are the number of space and time points respectively. We have set $\tilde{S}_{max} = 5, \tilde{\tau} = 1, l = 25, \tilde{\mu} = 0.04$ and $\tilde{\sigma} = 0.4$. The ratio is the ratio of the change in errors.

with

$$H\left(j\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right)=\begin{cases}-\dfrac{e^{l}i\Delta\tilde{S}f_{i,j}^{q}}{i\Delta\tilde{S}+l}\left(\dfrac{lf_{i,j}^{q}}{\left(i\Delta\tilde{S}+l\right)f_{i,j}^{q-1}}\right)^{\frac{l}{i\Delta\tilde{S}}} & \text{if } \delta^{*}\geq\delta^{min}\\ e^{-l\delta^{min}}\left(e^{-i\Delta\tilde{S}\left(1+\delta^{min}\right)}f_{i,j}^{q-1}-f_{i,j}^{q}\right) & \text{if } \delta^{*}\leq\delta^{min}.\end{cases} \tag{8.24}$$

and conditions (8.13) and (8.14).

At each $\left(\tilde{\tau},q,\tilde{S}\right)$ we must examine whether the constraint is satisfied or not, and hence which case of (8.19) we use. For the former case we need (from (8.18))

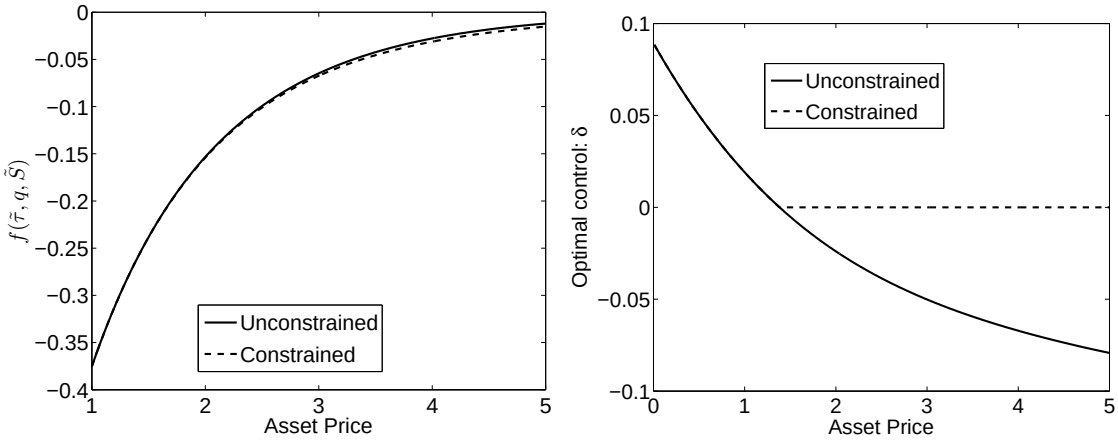
$$\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)}{lf^{min}\left(\tilde{\tau},q,\tilde{S}\right)}\right)-1\geq\delta^{min} \tag{8.25}$$

which is equivalent to the constraint

$$f^{min}\left(\tilde{\tau},q,\tilde{S}\right)\geq e^{-\tilde{S}\left(1+\delta^{min}\right)}\frac{\left(\tilde{S}+l\right)f^{min}\left(\tilde{\tau},q-1,\tilde{S}\right)}{l}. \tag{8.26}$$

Therefore, if (8.26) is satisfied, we use the former case of (8.24), otherwise we use the latter. As we are employing explicit differences for the non-linear term we use the explicit terms in checking the constraint (8.26), thus eliminating the need of iteration.

The above method should exhibit $O\left(\Delta\tilde{\tau},\Delta\tilde{S}^{2}\right)$ convergence, which we see in table 8.1 is the case. In the next section, in which we replace the exponential intensity with a power intensity, we shall discuss a novel numerical method (see section 8.2.4) in



(a) Value function against $\tilde{S}$

(b) Optimal strategy against $\tilde{S}$

Figure 8.1: Comparing the value function and optimal strategy of the constrained and unconstrained model against values of $\tilde{S}$ for a trader with one asset remaining, at $\tilde{\tau} = 1$. The parameter values take $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.

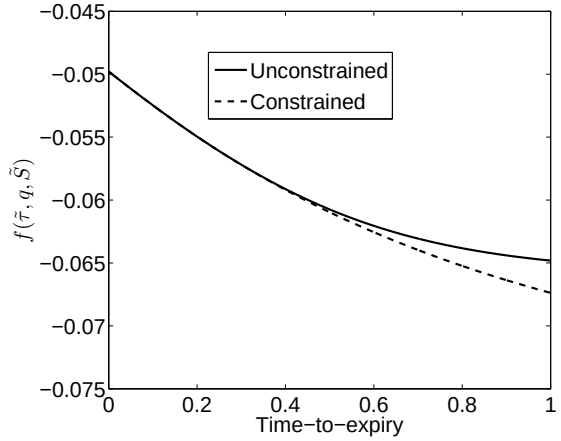
which the PDE, including the non-linear term, is taken implicitly, and the control is found numerically. We will use this implicit method to verify the above method.

8.1.4 Numerical examples

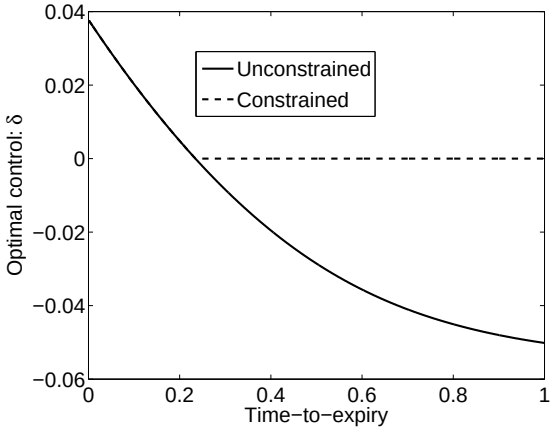
We shall briefly discuss how the solutions of the constrained model vary from those of the unconstrained model of chapter 4.

As mentioned, solutions of the constrained optimisation problem are suboptimal compared to those of the unconstrained optimisation problem. We therefore expect the value function of the constrained problem to be less than that of the unconstrained value function. In figure 8.1 we have plotted the value function (figure 8.1(a)) and optimal strategy (figure 8.1(b)) against $\tilde{S}$, at the initial time, $\tilde{\tau} = 1$. In figure 8.1(a) we can observe the value function of the unconstrained problem being greater than that of the constrained problem. Although it is only graphically observable for larger $\tilde{S}$ values, it should be noted that the unconstrained value function is greater than the constrained value function for all values of $\tilde{S}$ due to the diffusive nature of the PDE. Figure 8.1(b) shows the difference in the optimal strategy for the constrained and unconstrained problem. It is clear where the constraint kicks in and the two solutions diverge. However, like the value function, the optimal strategies differ for all $\tilde{S}$ values due to the diffusion.

The transition to the constraint can be seen in figure 8.2 which shows the value

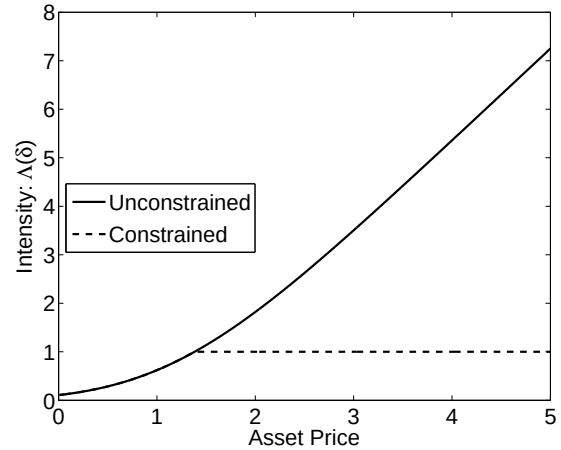


(a) Value function against $\tilde{\tau}$

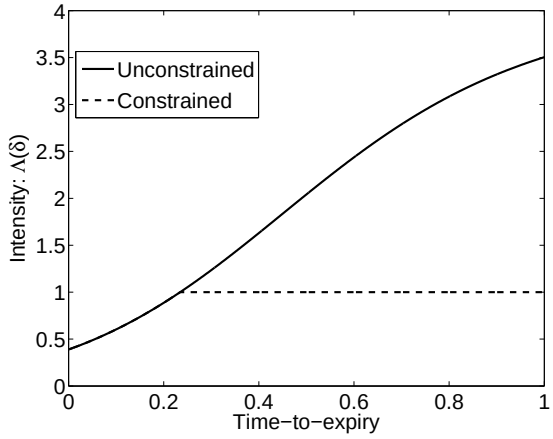


(b) Optimal strategy against $\tilde{\tau}$

Figure 8.2: Comparing the value function and optimal strategy of the constrained and unconstrained model against values of $\tilde{\tau}$ for a trader with one asset remaining, at $\tilde{S} = 3$. The parameter values take $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.



(a) Intensity function against $\tilde{S}$



(b) Intensity function against $\tilde{\tau}$

Figure 8.3: Comparing the intensity function of the Poisson process for the constrained and unconstrained model values of $\tilde{S}$ (figure 8.3(a)) and $\tilde{\tau}$ (figure 8.3(b)). This is for a trader with one asset remaining. The parameter values take $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $l = 25$.

function (figure 8.2(a)) and optimal strategy (figure 8.2(b)) against $\tilde{\tau}$ at a particular $\tilde{S}$ value; the divergence of the solutions as $\tilde{\tau}$ increases is vivid.

It is also of interest to see how the constraint affects the intensity of the Poisson process, which is given by (8.5). Figure 8.3 is a plot of (8.5) using the constrained and unconstrained optimal control, with figure 8.3(a) showing the intensity against $\tilde{S}$ and figure 8.3(b) showing the intensity against $\tilde{\tau}$.

We shall now turn our attention to a similar framework but under power intensity, which naturally encapsulates a lower bound of zero for the optimal trading strategy.

8.2 Introducing a power function intensity

In this section we examine changing the form of the intensity function from a negative exponential function to a power function with the power taking a negative value. This self imposes the constraint $\delta \in (0, \infty)$.

Implementing this change highlights the power of using a combined approach of asymptotic and numerical methods. We have shown in chapter 4 that we are not constrained by the underlying process driving the asset while in chapters 6 and 7 we highlighted that our methods are still applicable for higher-dimensional problems. In this section we demonstrate that our (numerical) methods can still be used even when we are unable to reduce the HJB equation to a non-linear PDE with explicit non-linear term and indeed need to solve the optimal solution numerically. Bayraktar and Ludkovski (2012) use a combination of a linear utility function, and a power intensity function, to reduce the problem to a recurrence relation. However in our formulation we cannot reduce the resulting HJB equation to a non-linear PDE with explicit non-linear term (as in previous cases, e.g. chapter 4) and thus must rely completely on numerical methods to find both the optimal strategy and solve the PDE.

8.2.1 Formulating the problem

The formulation of the problem is similar to that of section 8.1. We assume the asset price $S(t)$ follows a geometric Brownian motion given by (8.1). The trader will continuously post orders into the ask side of the limit-order book for price $S^a(t)$ given by (8.2). Asset sales and the trader's wealth follow a Poisson process, $N(t)$, with

time-dependent intensity, given by (8.3) and (8.4) respectively. The intensity $\Lambda(\delta)$ takes the form:

$$\Lambda(\delta) = \frac{\lambda}{\delta^{\alpha}}, \quad \alpha > 1 \tag{8.27}$$

which is distinct from (8.5) which has the negative exponential form. The derivation of (8.27) was discussed in section 3.2.2. The constraint $\alpha > 1$ is necessary for a unique solution to exist, which we discuss further in section 8.2.4.

The objective is to liquidate this portfolio before some final time T. We define our value function

$$u(t, X, q, S) = \sup_{\delta(t) \in \mathcal{A}} \mathbb{E}\left[-e^{-\gamma(X(T)+q(T)S(T))}\right], \tag{8.28}$$

where $\mathcal{A}$ is the set of admissible trading strategies bounded below by zero.

Given the optimisation problem of (8.28), an HJB equation can be derived by applying the Bellman (1957) principle of optimality and using Itô's lemma:

$$\begin{aligned} u_t(t, X, q, S) + \mu S u_S(t, X, q, S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t, X, q, S) \\ + \sup_{\delta \in \mathcal{A}} \left[\frac{\lambda}{\delta^{\alpha}} \left(u(t, X + S\left(1 + \delta\right), q - 1, S) - u(t, X, q, S)\right)\right] = 0, \end{aligned} \tag{8.29}$$

with initial conditions (8.8) and (8.9). However, unlike the case of the exponential intensity, the Hamiltonian

$$\begin{aligned} H\left(t, X, q, S, u\right) &= \sup_{\delta \in \mathcal{A}} \left[\frac{\lambda}{\delta^{\alpha}} \left(u(t, X + S\left(1 + \delta\right), q - 1, S) - u(t, X, q, S)\right)\right] \\ &= \sup_{\delta \in \mathcal{A}} L\left(\delta, t, X, q, S, u\right) \end{aligned} \tag{8.30}$$

may tend to infinity in some domain of (t, X, q, S, u) as $\delta \to 0$. In such a case, as discussed in Pham (2009), from the arguments leading to the derivation of the HJB equation (see section 8.2.2), we introduce a continuous function $G\left(t, X, q, S, u\right)$ such that

$$H\left(t, X, q, S, u\right) < \infty \iff G\left(t, X, q, S, u\right) \leq 0. \tag{8.31}$$

This results in

$$u_t(t, X, q, S) + \mu S u_S(t, X, q, S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t, X, q, S) + H\left(t, X, q, S, u\right) \leq 0, \tag{8.32}$$

$$G\left(t, X, q, S, u\right) \leq 0. \tag{8.33}$$

For the problem stated here we have

$$G\left(t, X, q, S, u\right) = u\left(t, X + S, q - 1, S\right) - u\left(t, X, q, S\right), \tag{8.34}$$

which represents when the asset is sold at its reference price, which we assume occurs instantaneously given the intensity of the Poisson process tends to infinity as $\delta \to 0$. If (8.33) has a strict inequality for some (t, X, q, S), then there exists a neighbourhood of (t, X, q, S, u) for which H is finite. Then from a similar argument that led to the derivation of the HJB equation we should have equality in (8.32). This results in a variational inequality for the dynamic programming equation

$$\max \Big[u_t(t,X,q,S) + \mu S u_S(t,X,q,S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t,X,q,S) + H\left(t,X,q,S,u\right), \\ u\left(t,X+S,q-1,S\right) - u\left(t,X,q,S\right) \Big] = 0, \tag{8.35}$$

with conditions (8.8) and (8.9), and $H\left(t,X,q,S,u\right)$ given by (8.30).

8.2.2 Derivation of HJB equation and verification theorem

This derivation is an extension to that shown for the case of the negative exponential intensity as described in section 4.1.1 and the reader is referred to that section for further insight.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration , $(\mathcal{F}_t, t \in [0,T])$. We define the stochastic processes adapted to the filtration $\mathcal{F}_t$ as:

$$\begin{bmatrix} dS(t) \\ dq(t) \\ dX(t) \end{bmatrix} = \begin{bmatrix} \mu S\left(t\right) \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} \sigma S\left(t\right) dW(t) \\ -dN(t) \\ S(t)\left(1+\delta\right) dN(t) \end{bmatrix}, \tag{8.36}$$

and the value function $u(t,X,q,S)$ as

$$u(t,X,q,S) = \sup_{\delta(t)\in\mathcal{A}} \mathbb{E}\left[\Phi(X(T),q(T),S(T))\right], \tag{8.37}$$

where $\mathcal{A} \subseteq \mathbb{R}^+$ is the set of admissible trading strategies and $\Phi\left(\cdot\right)$ is the form of the utility function. Bellman's Principle of Optimality reads:

$$u(t,X,q,S) = \sup_{\delta(t)\in\mathcal{A}} \mathbb{E}\left[u(t+h,X(t+h),q(t+h),S(t+h))\right]. \tag{8.38}$$

To derive the dynamic programming equation we will first assume $\delta(s) = \delta$ for $s \in [t, t+h]$, i.e. the control parameter is constant over the interval $[t, t+h]$. From (8.38) we have:

$$u(t,X,q,S) \geq \mathbb{E}\left[u(t+h,X(t+h),q(t+h),S(t+h))\right]. \tag{8.39}$$

Subtracting $u(t, X, q, S)$ from both sides gives

$$0 \geq \mathbb{E}\left[u(t+h, X(t+h), q(t+h), S(t+h)) - u(t, X, q, S)\right]. \tag{8.40}$$

Assuming $u \in C^{1,2}\left([0,T), \mathbb{R}^+\right)$ in t and S respectively, we now apply Itô's lemma to $u(t+h, X(t+h), q(t+h), S(t+h)) - u(t, X, q, S)$, which when divided by h and letting $h \to 0$ results in (this is shown in more detail in section 4.1.1)

$$0 \geq \mathcal{L}u(t, X, q, S) + \frac{\lambda}{\delta^\alpha}\left(u(t, X + S\left(1+\delta\right), q-1, S) - u(t, X, q, S)\right) \tag{8.41}$$

where we define the generator $\mathcal{L}$ of the geometric Brownian motion as

$$\mathcal{L}u(t, X, q, S) = u_t(t, X, q, S) + \mu S u_S(t, X, q, S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t, X, q, S). \tag{8.42}$$

We define the Hamiltonian as

$$H(t, X, q, S) = \sup_{\delta(t) \in \mathcal{A}} \frac{\lambda}{\delta^\alpha}\left(u(t, X + S\left(1+\delta\right), q-1, S) - u(t, X, q, S)\right). \tag{8.43}$$

If

$$\frac{\lambda}{\delta^\alpha}\left(u(t, X + S\left(1+\delta\right), q-1, S) - u(t, X, q, S)\right) > 0 \text{ as } \delta \to 0 \Rightarrow H(t, X, q, S) \to \infty.$$

Similarly, if

$$\frac{\lambda}{\delta^\alpha}\left(u(t, X + S\left(1+\delta\right), q-1, S) - u(t, X, q, S)\right) \leq 0 \text{ as } \delta \to 0 \Rightarrow H(t, X, q, S) < \infty.$$

Therefore

$$H\left(t, X, q, S, u\right) < \infty \Longleftrightarrow G\left(t, X, q, S, u\right) = u\left(t, X + S, q-1, S\right) - u\left(t, X, q, S\right) \leq 0. \tag{8.44}$$

If the inequality of (8.44) is strictly unequal then the optimal control is given by the non-zero Markov control policy $\delta^* = \delta^*(s, X^*(s), q^*(s), S(s))$. Using the optimal policy and making the same assessment as above we obtain

$$0 = \mathcal{L}u(t, X, q, S) + \frac{\lambda}{\left(\delta^*\right)^\alpha}\left(u(t, X + S\left(1+\delta^*\right), q-1, S) - u(t, X, q, S)\right), \tag{8.45}$$

and so we can conclude

$$0 = \mathcal{L}u(t, X, q, S) + \sup_{\delta \in \mathcal{A}}[\Lambda(\delta)\left(u(t, X + S\left(1+\delta\right), q-1, S) - u(t, X, q, S)\right)]. \tag{8.46}$$

If the inequality of (8.44) is strictly equal we have

$$u\left(t,X,q,S\right)=u\left(t,X+S,q-1,S\right), \tag{8.47}$$

and hence we can conclude

$$\begin{aligned}\max\Big[u_t(t,X,q,S)+\mu Su_S(t,X,q,S)+\frac{1}{2}\sigma^2S^2u_{SS}(t,X,q,S)+H\left(t,X,q,S,u\right),\\ u(t,X+S,q-1,S)-u(t,X,q,S)\Big]=0,\end{aligned} \tag{8.48}$$

with initial conditions (8.8) and (8.9), and $H\left(t,X,q,S,u\right)$ given by (8.43). This problem is known as a singular stochastic control problem and the reader is referred to Pham (2009) and Øksendal and Sulem (2005) for further insight.

Using the same methods as section 4.1.1 a verification theorem can be derived to show that the solution from the variational inequality (8.48) is indeed the solution of (8.37).

8.2.3 Reduction of the problem

As with the negative exponential intensity case (see section 8.1) we are able to perform both a reduction of variables and a reduction of parameters on the power intensity case. Using the ansatz solution (8.10) with $\tau=T-t$ and the non-dimensionalisation (8.11) we can rewrite (8.29) as

$$\begin{aligned}-f_{\tilde{\tau}}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\mu}\tilde{S}f_{\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f_{\tilde{S}\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)\\ +\sup_{\delta\in\mathcal{A}}\left[\frac{1}{\delta^\alpha}\left(e^{-\tilde{S}(1+\delta)}f\left(\tilde{\tau},q-1,\tilde{S}\right)-f\left(\tilde{\tau},q,\tilde{S}\right)\right)\right]=0,\end{aligned} \tag{8.49}$$

with initial conditions (8.13) and (8.14). The Hamilton, H, previously defined by (8.30) can be expressed as

$$\begin{aligned}H\left(\tilde{\tau},q,\tilde{S},f\right)&=\sup_{\delta\in\mathcal{A}}\left[\frac{1}{\delta^\alpha}\left(e^{-\tilde{S}(1+\delta)}f\left(\tilde{\tau},q-1,\tilde{S}\right)-f\left(\tilde{\tau},q,\tilde{S}\right)\right)\right]\\ &=\sup_{\delta\in\mathcal{A}}L\left(\delta,\tilde{\tau},q,\tilde{S},f\right),\end{aligned} \tag{8.50}$$

and the function G, given by (8.34) can be expressed as

$$G\left(\tilde{\tau},q,\tilde{S},f\right)=e^{-\tilde{S}}f\left(\tilde{\tau},q-1,\tilde{S}\right)-f\left(\tilde{\tau},q,\tilde{S}\right). \tag{8.51}$$

We can therefore rewrite the variational inequality (8.35) as

$$\max\Big[-f_{\tilde{\tau}}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\mu}\tilde{S}f_{\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f_{\tilde{S}\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+H\left(\tilde{\tau},q,\tilde{S},f\right),\\ e^{-\tilde{S}}f\left(\tilde{\tau},q-1,\tilde{S}\right)-f\left(\tilde{\tau},q,\tilde{S}\right)\Big]=0, \quad (8.52)$$

along with initial conditions (8.13) and (8.14), boundary conditions (8.21) and (8.22), and $H\left(\tilde{\tau},q,\tilde{S},f\right)$ given by (8.50).

8.2.4 Numerical method

When examining the case of an exponential intensity function we were able to reduce the HJB equation to a non-linear PDE with explicit non-linear term by finding the optimal control in terms of the value function, given by (8.17), and substituting it back into the HJB equation. For the case of the power intensity function we are unable to use similar methods and must approach the problem with a different numerical method.

We use a method based on the SOR scheme (the SOR scheme is discussed in section 2.6.3) as now we must also use numerical methods to find the optimal control, δ^*, by numerically finding the supremum of the (8.50).

Letting $f\left(j\Delta\tilde{\tau},q,i\Delta\tilde{S}\right)=f^q_{i,j}$ we can approximate (8.49) using implicit differencing as

$$-\frac{f^q_{i,j+1}-f^q_{i,j}}{\Delta\tilde{\tau}}+\mu i\Delta\tilde{S}\frac{f^q_{i+1,j+1}-f^q_{i-1,j+1}}{2\Delta\tilde{S}}+\frac{1}{2}\sigma^2\left(i\Delta\tilde{S}\right)^2\frac{f^q_{i+1,j+1}-2f^q_{i,j+1}+f^q_{i-1,j+1}}{\Delta\tilde{S}^2}\\ +H\left((j+1)\,\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right)=0, \quad (8.53)$$

with

$$H\left((j+1)\,\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right)=\sup_{\delta\in\mathcal{A}}\left[\frac{1}{\delta^\alpha}\left(e^{-i\Delta\tilde{S}(1+\delta)}f^{q-1}_{i,j+1}-f^q_{i,j+1}\right)\right]\\ =\sup_{\delta\in\mathcal{A}}L\left(\delta,(j+1)\,\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right). \quad (8.54)$$

At each iteration we must recalculate the optimal strategy by finding

$$\delta^*=\arg H\left((j+1)\,\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right). \quad (8.55)$$

Once this is calculated we substitute it into

$$L^*\left(\delta^*,(j+1)\,\Delta\tilde{\tau},q,i\Delta\tilde{S},f\right)=\frac{1}{(\delta^*)^\alpha}\left(e^{-i\Delta\tilde{S}(1+\delta^*)}f^{q-1}_{i,j+1}-f^q_{i,j+1}\right), \quad (8.56)$$

so now we have a PDE to solve and then perform another iteration. To find the supremum of (8.54), and thus the value of (8.55), we must specify a discrete range of δ to search over. This range cannot contain zero as this will result in computational error (given there is division by some positive power of δ). We thus set a range for δ which has a lower bound, $\underline{\delta}$, near to zero (say 10^{-12}). We then search over the range of δ, the method of which is described next. If we find the optimal control, δ^*, is equal to the lower bound of the discrete space of δ, then it is assumed it should in fact be zero in which case the Hamiltonian would be infinite and thus the second term in the variation inequality (8.52) becomes active. We thus modify the numerical (SOR) scheme appropriately.

When we are solving the HJB equation part of the variational inequality (8.53) we can write it in the form of $\mathbf{Af} = \mathbf{b}$, where $\mathbf{A}$ is a tridiagonal matrix. Defining $f_{i,j}^{q,p}$ as the p^{th} iteration of $f_{i,j}^{q}$ we can formulate a scheme based on the SOR method as

$$y = \frac{1}{\beta_i}\left(b_i - \alpha_i f_{i-1,j+1}^{q,p+1} - \gamma_i f_{i+1,j+1}^{q,p}\right) \tag{8.57}$$

$$f_{i,j+1}^{q,p+1} = \begin{cases} f_{i,j+1}^{q,p} + \omega\left(y - f_{i,j+1}^{q,p}\right) & \text{if } \delta^* \neq \underline{\delta} \\ e^{-i\Delta S} f_{i,j+1}^{q-1} & \text{if } \delta^* = \underline{\delta} \end{cases} \tag{8.58}$$

where β is a vector containing the main diagonal of $\mathbf{A}$, α is a vector containing the subdiagonal of $\mathbf{A}$ and γ is a vector containing the superdiagonal of $\mathbf{A}$. We iterate over these two equations until the difference between successive iterations is sufficiently small, beginning by setting $f_{i,j+1}^{q,0} = f_{i,j}^{q}$, and recalculating (8.55) during each iteration. When δ^* is recalculated the values of $\mathbf{b}$ and β must be recalculated as they contain the q and $q-1$ components of (8.56) respectively. Convergence of the above method is discussed by Huang et al. (2012), and the reader is referred to this article for more insight.

Finding the supremum

In each of the SOR iterations we must find the optimal strategy by solving (8.55) for each $\left(\tilde{\tau}, q, \tilde{S}\right)$. We must first show that $L\left(\delta, \tilde{\tau}, q, \tilde{S}, f\right)$ is a unimodal function in δ. A unimodal function is defined as follows: a function $g(x)$ is a unimodal function if for some value m, it is monotonically increasing for $x \leq m$ and monotonically decreasing for $x \geq m$. In that case, the maximum value of $g(x)$ is $g(m)$ and there

are no other local maxima. It was shown that under a negative exponential intensity function $L\left(\delta, \tilde{\tau}, q, \tilde{S}, f\right)$ is a unimodal function, given we were able to find unique maximum explicitly in terms of the value function. We cannot explicitly find the unique maximum under a power intensity function however we can still show that one exists.

Over a domain $\delta \in (0, \infty)$ we have

$$\begin{aligned}\frac{\partial L}{\partial \delta}\left(\delta, \tilde{\tau}, q, \tilde{S}, f\right) = \Lambda'(\delta)\left(e^{-\tilde{S}(1+\delta)} f\left(\tilde{\tau}, q-1, \tilde{S}\right) - f\left(\tilde{\tau}, q, \tilde{S}\right)\right) \\ - \tilde{S}\Lambda(\delta)\left(e^{-\tilde{S}(1+\delta)} f\left(\tilde{\tau}, q-1, \tilde{S}\right)\right)\end{aligned} \tag{8.59}$$

where $\Lambda(\delta)$ is defined by (8.27), and $\Lambda'(\delta)$ denotes its derivative. A maximum of $L\left(\delta, \tilde{\tau}, q, \tilde{S}, f\right)$ exists at

$$\delta^* - \frac{1}{\tilde{S}} \ln\left(1 - \frac{\tilde{S}\Lambda(\delta^*)}{\Lambda'(\delta^*)}\right) = \frac{1}{\tilde{S}} \ln\left(\frac{f\left(\tilde{\tau}, q-1, \tilde{S}\right)}{f\left(\tilde{\tau}, q, \tilde{S}\right)}\right) + 1. \tag{8.60}$$

We define the strictly increasing function

$$h(x) = x - \frac{1}{\tilde{S}} \ln\left(1 - \frac{\tilde{S}\Lambda(x)}{\Lambda'(x)}\right), \tag{8.61}$$

which we know is strictly increasing given

$$h'(x) = \frac{-\tilde{S}\Lambda(x)\Lambda'(x)}{\Lambda'(x)^2 - \tilde{S}\Lambda(x)\Lambda'(x)} + \frac{2\Lambda'(x)^2 - \Lambda(x)\Lambda''(x)}{\Lambda'(x)^2 - \tilde{S}\Lambda(x)\Lambda'(x)} \tag{8.62}$$

is strictly positive under the condition $\Lambda(x)\Lambda''(x) \leq 2\Lambda'(x)^2$. This is also a condition found by Bayraktar and Ludkovski (2012), and to which (8.27) for $\alpha > 1$, and the exponential intensity (8.5), satisfy. Given (8.61) is strictly increasing, δ^* satisfying (8.60) for fixed $\left(\tilde{\tau}, q, \tilde{S}\right)$ is unique and a unique maximum of $L\left(\delta, \tilde{\tau}, q, \tilde{S}, f\right)$ exists.

The algorithm to solve for a coordinate m, corresponding to the maximum $g(m)$, of a unimodal function $g(x)$, in a domain $x \in [x_{min}, x_{max}]$ is as follows:

1. Choose a tolerance level ε, and from this calculate the number of iterations m (see below). Set $counter = 0$.
2. Set $\Delta x = \dfrac{x_{max} - x_{min}}{n}$.
3. Discretise the domain into $n+1$ values $\{x_0, \ldots, x_n\}$ with $x_i = x_{min} + i\Delta x$.

4. Calculate $g(x_i)$ for $i \in \{0, \ldots, n\}$ and find $x^* = \arg\max\{g(x_i) : i \in \{0, \ldots, n\}\}$.

5. If $x^* = x_{min}$, $x_{max} = x_{min} + \Delta x$, else if $x^* = x_{max}$, $x_{min} = x_{max} - \Delta x$, else $x_{min} = x^* - \Delta x$, $x_{max} = x^* + \Delta x$.

6. If $counter < m$ increment counter and return to step 2. Else finish.

Optimal number of sections

A question arises now as to what value n should take. Using the optimal bisection algorithm requires a different approach than when using a standard root finding algorithm. In the root finding algorithm we can use one division to go from the interval $[x_{min}, x_{max}]$ to either $[x_{min}, (x_{min} + x_{max})/2]$ or $[(x_{min} + x_{max})/2, x_{max}]$. However when searching for a maximum we do not know if the maximum is contained in the left or right interval. If we divide the interval into n sections, in each iteration we can only narrow down to the interval $[x_i - (x_{max} - x_{min})/n, x_i + (x_{max} - x_{min})/n]$ where x_i was the result that gave the value of the maximum.

If we seek to reduce a starting interval $[x_{min}, x_{max}]$ by some relative tolerance ϵ, such that the length of the interval is reduced to less than $(x_{max} - x_{min}) \times \epsilon$, then the number of iterations m required for fixed n is

$$m = \frac{\log(\epsilon)}{\log(2/n)}.$$

So if the work done in the algorithm at each stage is proportional to n, we can write the cost function to achieve an accuracy of ϵ as

$$f(n) = An\frac{\log(\epsilon)}{\log(2/n)} + B.$$

Differentiating with respect to n gives

$$\frac{df}{dn}(n) = A\log(\epsilon)\frac{\log(2/n) + 1}{\log^2(2/n)},$$

so the minimum cost is achieved at

$$n = 2e.$$

Obviously we must choose an integer, and it appears as though $n = 5$ gives a slightly lower cost than $n = 6$. To find the number of iterations m in the algorithm for a required tolerance we choose

$$m = \frac{\log(\epsilon)}{\log(0.4)}.$$

For example, consider the interval [0,1]. If we want to reduce it to 5% of its original space, we find the number of iterations to be

$$m = \frac{\log(0.05)}{\log(0.4)} = 3.27$$

which we round to 4 as it must be an integer. In the first search we reduce the interval to $[x_i^1 - 0.2, x_i^1 + 0.2]$, where the superscript on x represents it is the optimal value found in that search. In the next search we reduce the interval to $[x_i^2 - 0.08, x_i^2 + 0.08]$. The third and fourth search reduce the interval to $[x_i^3 - 0.032, x_i^3 + 0.032]$ and $[x_i^4 - 0.0128, x_i^4 + 0.0.0128]$ respectively, in which the interval is now of size 0.0256 which is within the absolute tolerance.

Other methods for finding the maximum of a unimodal function include the Golden Section Search algorithm developed by Kiefer (1953) which derives its name from the fact that the algorithm maintains the function values for triplets of points whose distances form a golden ratio; the above numerical techniques are outlined in Press et al. (2009).

The above numerical scheme should exhibit $O\left(\Delta\tilde{\tau}, \Delta\tilde{S}^2\right)$ convergence which we can see from table 8.2 is the case. To further confirm our numerical method we returned to the problem examined in chapter 4, for which we can solve using a direct solver. We solved (4.30), which requires the supremum to be found numerically (rather than directly), and compared the results to those found when the supremum is calculated as an explicit function of the value function, as in (4.33). We also verified the numerics by solving the constrained optimisation problem discussed in section 8.1 using the above method.

8.2.5 Numerical examples

We shall briefly discuss how the solutions under a power intensity function vary from that under an exponential intensity (as was previously discussed in chapter 4).

Figure 8.4 shows the optimal strategy as a function of $\tilde{S}$ for various values of the inventory, q. This can be seen at times-to-expiry $\tilde{\tau} = 0.05$ (figure 8.4(a)) and $\tilde{\tau} = 1$ (figure 8.4(b)). We notice these plots have similar characteristics to those of the exponential model (see figures 4.2 and 4.4). The optimal strategies are decreasing in q and decreasing in $\tilde{S}$. However, there is a contrast to the exponential model.

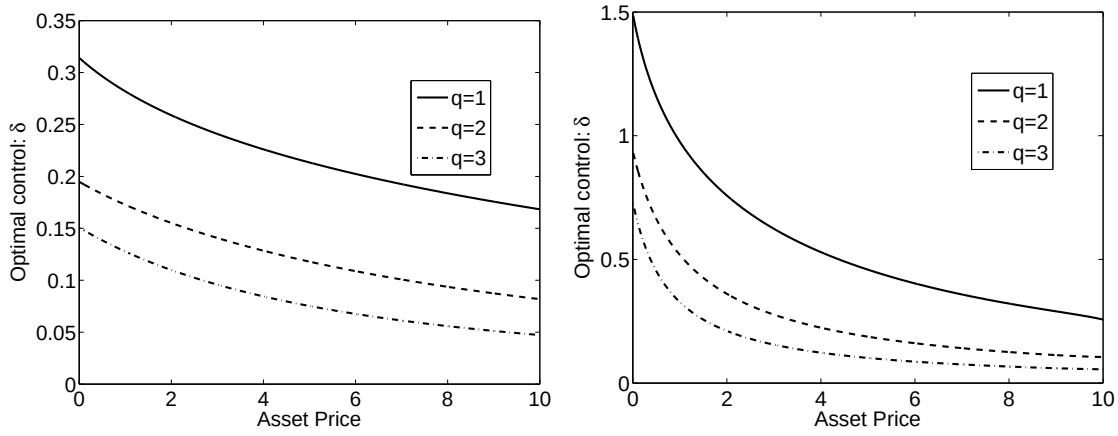
Table 8.2: Convergence of SOR scheme under power intensity

n	m	$f\left(\tilde{\tau}=1, q=1, \tilde{S}=5\right)$	ratio
251	1001	-0.001469455352	NA
501	1001	-0.001469320629	NA
1001	1001	-0.001469286931	3.997988
2001	1001	-0.001469278515	4.003909
4001	1001	-0.001469276376	3.934555
1001	251	-0.001503341513	NA
1001	501	-0.001480462120	NA
1001	1001	-0.001469286931	2.047338
1001	2001	-0.001464036210	2.128315
1001	4001	-0.001461615175	2.168791

n and m are the number of space and time points respectively. We have set $\tilde{S}_{max} = 10, \tilde{\tau} = 1, \alpha = 2, \tilde{\mu} = 0.04$ and $\tilde{\sigma} = 0.4$. The ratio is the ratio of the change in errors.

The exponential model had increasing optimal strategy as time-to-expiry increases for certain $\tilde{S}$ regimes (small $\tilde{S}$), and decreasing optimal strategy as time-to-expiry increases for other $\tilde{S}$ regimes. On the other hand, numerical calculations indicated conclusively that the optimal strategy increases for all $\tilde{S}$ regimes as time-to-expiry increases. This can be seen in figure 8.5. As $\tilde{\tau} \to 0$ the optimal trading strategy $\delta^* \to 0$. However, it will always stay greater than zero as the trader will try making some revenue from selling above the reference price. We believe this is due to the effect of having the option of selling immediately by sending $\delta \to 0$ if the trader so chooses, and as such has less liquidation risk than that of the exponential model. Bayraktar and Ludkovski (2012) also find (analytically) that δ remains strictly positive, under their framework.

Finally, we discuss the effect of varying parameters. Changing $\tilde{\mu}$ and $\tilde{\sigma}$ has the same properties as the exponential intensity model, which was discussed in chapter 4, so we will not reiterate . We will thus focus on the different regimes of the parameter α. Firstly, let us consider the properties of the intensity function (8.27) for varying α. For smaller α, equation (8.27) will tend to infinity slower as $\delta \to 0$, while tending to zero faster as $\delta \to \infty$. We see similar behaviour in figure 8.6. In figure 8.6(a) we see the optimal strategy decreasing in α for larger values of δ^* while increasing in α for smaller values of δ^*. This indicates that for larger values of δ^* the trader has preference over smaller α while when he wants to liquidate quicker (and thus have smaller δ^*) he



(a) Optimal strategy against $\tilde{S}$ at $\tau = 0.05$ (b) Optimal strategy against $\tilde{S}$ at $\tau = 1$

Figure 8.4: Comparing the optimal strategy of the power intensity model against values of $\tilde{S}$ for various assets remaining, at two different times-to-expiry. The parameter values take $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.4$ and $\alpha = 2$.

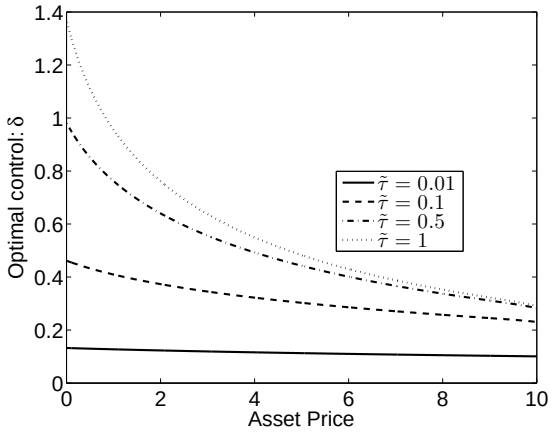
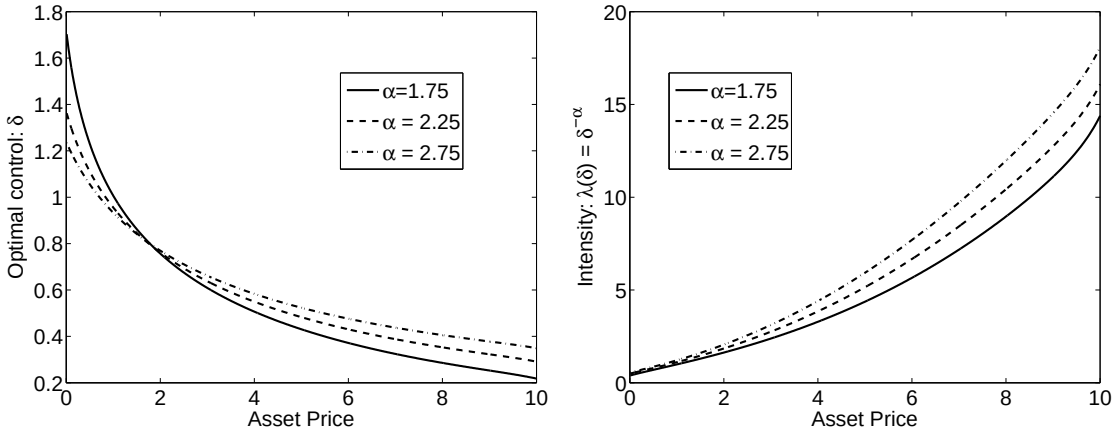


Figure 8.5: Comparing the optimal strategy for various $\tilde{S}$ values at different time-to-expiry for a trader with one asset remaining. The parameter values take $\tilde{\mu} = 0.04$ and $\tilde{\sigma} = 0.4$ and $\alpha = 2$.



(a) Optimal Strategy against $\tilde{S}$ for various α (b) Intensity against $\tilde{S}$ for various α

Figure 8.6: Comparing the optimal strategy and intensity function for various α against values of $\tilde{S}$ for a trader with one asset remaining, at $\tilde{\tau} = 1$. The parameter values take $\tilde{\mu} = 0.04$ and $\tilde{\sigma} = 0.4$.

prefers a higher value of α. The corresponding intensities can be seen in figure 8.6(b).

We shall now turn our attention to the exponential intensity, under a framework which allows the trader to trade with both market orders and limit orders, thus also allowing for immediate execution.

8.3 Including market orders

In this section we shall examine the strategy of a trader who can post both limit orders and market orders throughout the trading period.

This analysis is in parallel with the recent work of Cartea and Jaimungal (2015b), who through standard Brownian motion, exponential intensity function with absolute decay parameter and linear utility function were able to cleverly build symmetry into the model from the start which allowed them to reduce the HJB equation to an ODE within the variational inequality equations. Our novel contribution is the use of more general diffusion processes and intensity functions, which lead to more complex equations that require accurate numerical schemes to solve. The solutions are significant as they are asset-price dependent, unlike the independent strategies found by Cartea and Jaimungal (2015b). A more in-depth discussion on studies that focus on trading using both limit and market orders can be found in section 3.4.1.

8.3.1 Formulating the problem

We begin with a similar set-up to that of the opening section of this chapter, section 8.1. We have a trader with wealth X and inventory q of an asset with price S which follow the stochastic processes (8.4), (8.3) and (8.1) respectively. The intensity of the Poisson process, $N(t)$, driving (8.4) and (8.3) is a negative exponential intensity of the form (8.5). The objective is to maximise the expected terminal time utility

$$u(t,X,q,S) = \sup_{\delta(t)\in\mathcal{A}} \mathbb{E}\left[-e^{-\gamma(X(T)+q(T)S(T))}\right], \tag{8.63}$$

which results in the HJB equation (8.7) with initial conditions (8.8) and (8.9)

Let us now assume the trader can place market orders of unitary size, at a discount of p relative to the asset price, so the assets are sold at $S(1-p)$. Here p could be seen as the temporary market impact that only affects the current transaction, or could be seen as the commission, spread, maker-taker fee, or some combination, again only affecting the current transaction.

We can model execution using market orders as a free-boundary problem. When the trader feels he can profit using limit orders his value function is greater than that if he were to liquidate using a market order, which equates to

$$u(t,X,q,S) \geq u(t,X+S(1-p),q-1,S). \tag{8.64}$$

When (8.64) is not satisfied, it is optimal for the trader to use a market order to liquidate. Let (τ_n) represent a series of stopping times representing the market-order decision times of a trader. At times τ_n the cash holdings and inventory follow

$$q(\tau_n) = q\left(\tau_n^-\right) - 1, \tag{8.65}$$

and

$$X(\tau_n) = X\left(\tau_n^-\right) + S(\tau_n^-)(1-p), \tag{8.66}$$

with τ_n^- representing the instantaneous time prior to τ_n.

From (8.64) and the HJB equation (8.7) we have

$$\begin{aligned} &u_t(t,X,q,S) + \mu S u_S(t,X,q,S) + \frac{1}{2}\sigma^2 S^2 u_{SS}(t,X,q,S) \\ &+\sup_{\delta}\left[\lambda e^{-l\delta}\left(u(t,X+S(1+\delta),q-1,S) - u(t,X,q,S)\right)\right] \leq 0, \end{aligned} \tag{8.67}$$

$$u(t,X+S(1-p),q-1,S) - u(t,X,q,S) \leq 0. \tag{8.68}$$

Thus the revised dynamic programming equation associated with the free-boundary problem can be formulated as a variational inequality

$$\max\Bigg(u_t(t,X,q,S)+\mu S u_S(t,X,q,S)+\frac{1}{2}\sigma^2S^2u_{SS}(t,X,q,S) \\ +\sup_{\delta}\left[\lambda e^{-l\delta}\left(u(t,X+S\left(1+\delta\right),q-1,S)-u(t,X,q,S)\right)\right], \\ u\left(t,X+S\left(1-p\right),q-1,S\right)-u\left(t,X,q,S\right)\Bigg)=0, \tag{8.69}$$

with initial conditions (8.8) and (8.9). The derivation of (8.69) is analogous to that of (8.35) which was discussed in section 8.2.2. Using the same methods as section 4.1.1 a verification theorem can be derived to show the solution from the variational inequality (8.69) is indeed the solution of (8.63) subject to the free-boundary (8.64).

For the interested reader, Jaillet et al. (1990) examine theoretical properties of variational inequality problems related to finance, in particular that of the American option, which is analogous to (8.69).

Reduction and transformation

We can use the same transformation as used in chapter 4 to non-dimensionalise the problem and reduce the number of variables and parameters. Using the ansatz solution (8.10) with $\tau = T-t$ and the non-dimensionalisation (8.11) we can reformulate (8.69) as

$$\max\Bigg(-f_{\tilde{\tau}}\left(\tilde{\tau},q,\tilde{S}\right)+\tilde{\mu}\tilde{S}f_{\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right)+\frac{1}{2}\tilde{\sigma}^2\tilde{S}^2f_{\tilde{S}\tilde{S}}\left(\tilde{\tau},q,\tilde{S}\right) \\ -\frac{e^l\tilde{S}f\left(\tilde{\tau},q,\tilde{S}\right)}{\tilde{S}+l}\left(\frac{lf\left(\tilde{\tau},q,\tilde{S}\right)}{\left(\tilde{S}+l\right)f\left(\tilde{\tau},q-1,\tilde{S}\right)}\right)^{\frac{l}{\tilde{S}}}, \\ e^{-\tilde{S}(1-p)}f\left(\tilde{\tau},q-1,\tilde{S}\right)-f\left(\tilde{\tau},q,\tilde{S}\right)\Bigg)=0, \tag{8.70}$$

with initial conditions (8.13) and (8.14), by finding the optimal strategy when early execution is not optimal as

$$\delta^*=\frac{1}{\tilde{S}}\ln\left(\frac{\left(\tilde{S}+l\right)f\left(\tilde{\tau},q-1,\tilde{S}\right)}{lf\left(\tilde{\tau},q,\tilde{S}\right)}\right)-1. \tag{8.71}$$

Table 8.3: Convergence of PSOR scheme for trader using market orders

n	m	$f\left(\tilde{\tau}=1, q=1, \tilde{S}=5\right)$	ratio
1001	1001	-0.078174679282	NA
2001	1001	-0.078174709111	NA
4001	1001	-0.078174717568	4.000032
8001	1001	-0.0781747181432	3.999966
16001	1001	-0.078174718898	4.00017
1001	501	-0.078174052374	NA
1001	1001	-0.078174679282	NA
1001	2001	-0.078174992632	2.000666
1001	4001	-0.078175149281	2.000333
1001	8001	-0.0781752287599	2.000166

n and m are the number of space and time points respectively. We have set $\tilde{S}_{max} = 10, \tilde{\tau} = 1, l = 5, \tilde{\mu} = 0.04$ and $\tilde{\sigma} = 0.2$. The ratio is the ratio of the change in errors.

To solve (8.70) we use the PSOR method as described in section 2.6.3, with boundary conditions (8.21) and (8.22). The PSOR scheme should exhibit $O\left(\Delta\tilde{\tau}, \Delta\tilde{S}^2\right)$ convergence, which we see in table 8.3 is the case.

8.3.2 Interpreting trading strategies with numerical examples

In this section we shall use the numerical results to examine some properties of the trading strategies when the trader can liquidate with market orders, and how these strategies compare to those devised in chapter 4.

We can examine analytically when it is optimal to execute early, through the optimal trading strategy, δ^*. When early execution is optimal

$$f\left(\tilde{\tau}, q, \tilde{S}\right) = e^{-\tilde{S}(1-p)} f\left(\tilde{\tau}, q-1, \tilde{S}\right). \tag{8.72}$$

Substituting (8.72) into (8.71) we have

$$\delta^{ee} = \frac{1}{\tilde{S}} \ln\left(\frac{\left(\tilde{S}+l\right)}{l}\right) - p, \tag{8.73}$$

where the superscript ee indicates early execution. Thus when

$$\delta^* < \delta^{ee} \tag{8.74}$$

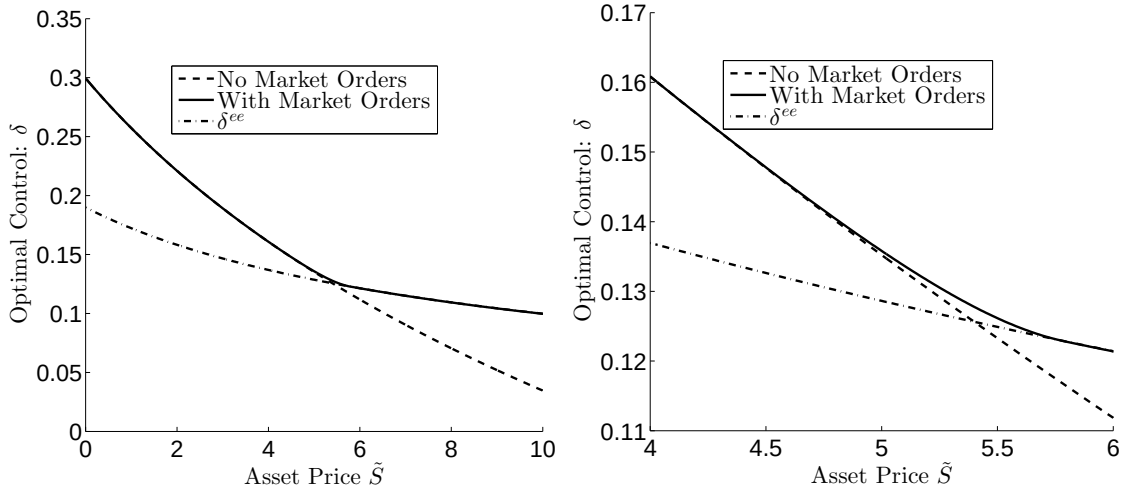
it is optimal to sell using a market order rather than placing a limit order.

Equation (8.73) has some interesting properties. Firstly, we notice it is both q and $\tilde{\tau}$ independent. It depends only on the risk-adjusted asset price, $\tilde{S}$, the decay parameter of the limit-order book, l, and the temporary market impact, p. This means that the lower bound to place a limit order at, before switching to a market order, is independent of the number of assets the trader has remaining and the time remaining. However, the value of the optimal control, δ^*, that we compare this lower bound against will be a function of q and $\tilde{\tau}$, as well as the previously mentioned variables and parameters.

Figure 8.7 indicates how the trading strategy differs when the trader can liquidate using both limit and market orders. In figure 8.7(a) we see the comparison of trading strategies when market orders can be used, compared to when they cannot be used (as discussed in chapter 4). For small $\tilde{S}$ (before it is optimal to liquidate using a market order) both strategies are approximately the same (but not identical, which we will discuss momentarily). However, as $\tilde{S}$ increases, and the risk increases, the trader will liquidate using a market order. This is portrayed when the strategy hits, and sticks, to the cut-off δ^{ee}, indicating the trader has liquidated with a marker order.

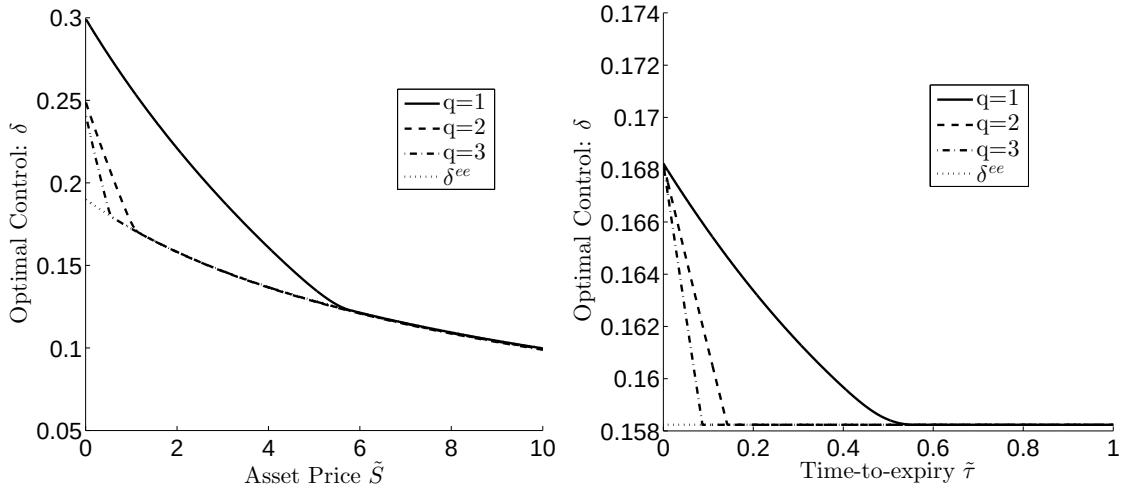
We mentioned that before the trader chooses to liquidate with a marker order his strategy is similar to a strategy if he was unable to liquidate with market orders. It is indeed similar, but it is not exact. Given the trader has the ability to liquidate with market orders, his value function will be higher than if he was unable to liquidate with marker orders. This is not only the case in the space where he chooses to liquidate, but also throughout the whole space given the diffusive nature of the PDE. Not only is his value function higher, but having the option to liquidate using market orders will also lead to higher values of the optimal trading strategy in areas that he chooses not to liquidate with market orders (given the optimal trading strategy is a functional of the value function). This can be seen in figure 8.7(b) which is a magnification of figure 8.7(a). Interestingly, and as expected, this is the opposite of what we saw in section 8.1, in which the constraint $\delta > \delta^{\min}$ resulted in suboptimal strategies and a value function less than the unconstrained problem. Here, the introduction of marker orders has led to trading strategies that are more favourable over those in which market orders cannot be used.

The behaviour of the trading strategy for various values of the inventory q can be



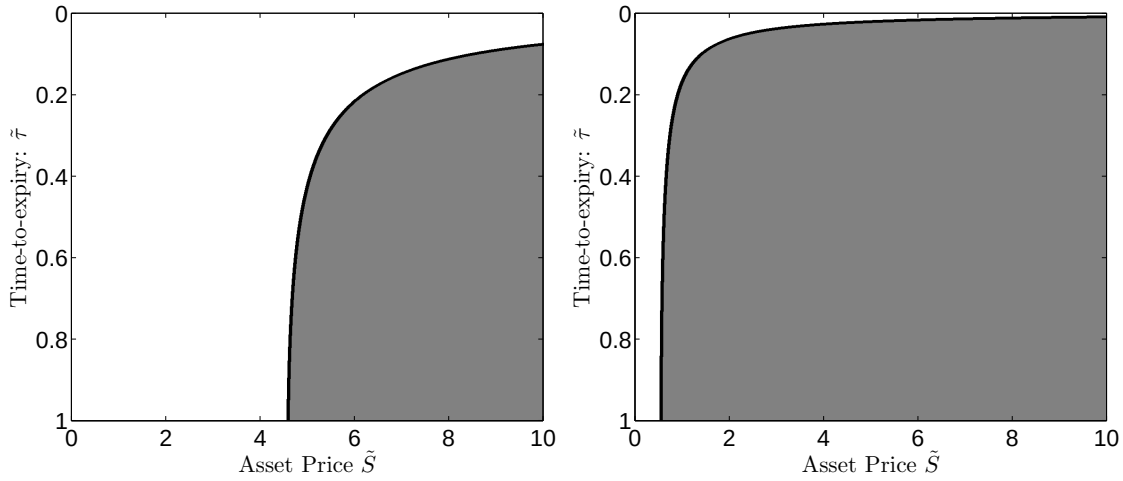
(a) Optimal strategy for various values of $\tilde{S}$ (b) Optimal strategy for zoomed values of $\tilde{S}$

Figure 8.7: Comparing optimal strategy for trader with one asset remaining who can or can't liquidate with market orders at $\tilde{\tau} = 1$, with $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.2, l = 5$ and $p = 0.01$,.



(a) Optimal strategy for various values of $\tilde{S}$ (b) Optimal strategy for various values of $\tilde{\tau}$

Figure 8.8: Optimal strategy for trader with varying amounts of inventory remaining who can liquidate using market orders, with $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.2, l = 5$ and $p = 0.01$. Figure 8.8(a) has $\tilde{\tau} = 1$ while figure 8.8(b) has $\tilde{S} = 2$

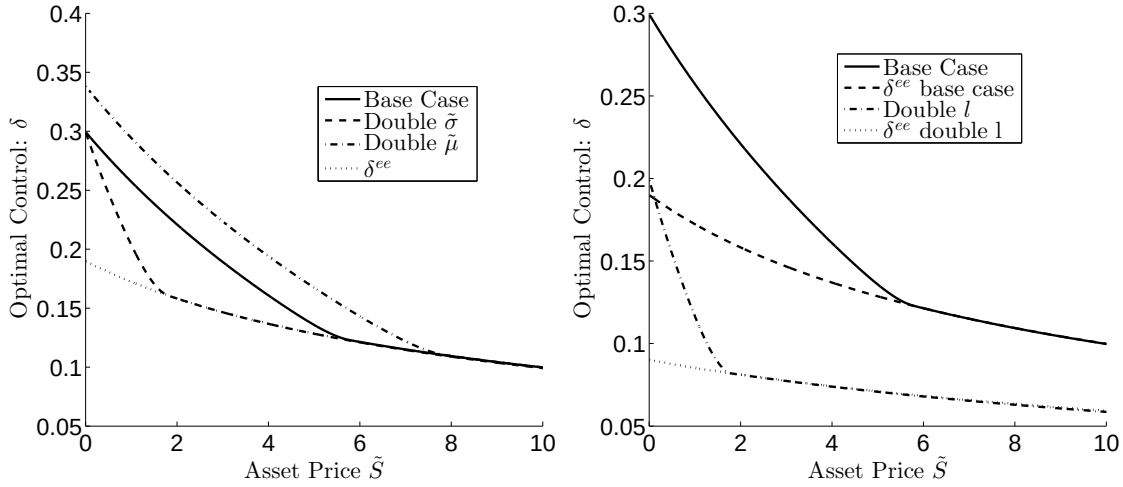


(a) Optimal strategy for $q = 1$ (b) Optimal strategy for $q = 3$

Figure 8.9: Heat map of early execution region, with $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.2, l = 5$ and $p = 0.01$. The area shaded white corresponds to using limit orders while the area shaded grey corresponds to market orders, with the free-boundary in black.

seen in figure 8.8. This is plotted against the risk-adjusted asset price, $\tilde{S}$, in figure 8.8(a), and against the time-to-expiry, $\tilde{\tau}$, in figure 8.8(b). In both we notice the trader chooses to use market orders if he has a larger amount of inventory, or the asset price is larger. This is due to him being characterised by a risk-averse utility function. We also notice that as the time-to-expiry tends to zero, the trader chooses to use limit orders over market orders given over shorter time frames he is not as concerned about the riskiness of his portfolio and is looking to profit. This is also portrayed in figure 8.9 which shows a heatmap for when it is optimal to use market orders over limit orders, plotted against the risk-adjusted asset price, $\tilde{S}$, and time-to-expiry, $\tilde{\tau}$, for $q = 1$ (figure 8.9(a)) and $q = 3$ (figure 8.9(b)).

Lastly we shall examine how the behaviour of the trading strategies change in different parameter regimes. We previously examined this in chapter 4 for the case of a trader who can liquidate using limit orders only. Given the behaviour found, and that the trader is risk-averse, we fully expect that as the riskiness increases the trader will choose to use market orders more often than in a less riskier scenario. Figure 8.10 shows the varying trading strategies for varying parameter regimes. In figure 8.10(a) we examine how the strategy changes if the drift, $\tilde{\mu}$, and the volatility, $\tilde{\sigma}$, are doubled. For the former case the trader uses market orders less often compared to the base case, as he will hold on to the asset in the hope that he can enjoy its higher expected



(a) Optimal strategy with adjusted $\tilde{\mu}$ and $\tilde{\sigma}$ (b) Optimal strategy with adjusted l

Figure 8.10: Comparing parameter sensitivity of optimal strategy for trader with one asset remaining who can liquidate with market orders. Base Case: $\tilde{\mu} = 0.04, \tilde{\sigma} = 0.2, l = 5$ and $p = 0.01$.

rate of growth. In the latter case we see that when the volatility is doubled the trader will use market orders more frequently, given his aversion to risk. We notice that the lower bound to which a limit order should be placed, δ^{ee}, is independent of both these parameter changes; it is however dependent on the decay factor of the limit-order book, l. Hence we have examined the sensitivity of this parameter in a separate figure, figure 8.10(b). It can be seen that as l is doubled, the lower bound δ^{ee} decreases, and the trader will use market orders more frequently in comparison to the base case. This is because as l increases, the intensity of filling orders using a limit order decreases more rapidly for positive values of δ^*. Thus the trader decreases his value of δ^* respectively, which in doing so results in hitting the early execution level, δ^{ee}.

8.4 Summary

In this chapter we extended the work of chapter 4 in three aspects. Firstly, we investigated a constrained optimisation model. Secondly, we investigated replacing the negative exponential intensity function with a power intensity function. Finally, we introduced the concept of trading explicitly with both limit orders and market orders.

Having the constraint $\delta > 0$ can be seen as more realistic than the unconstrained model (however there are also arguments for a negative δ). In either case implementing

the constraint allows us to examine how the trading strategies change, and indeed how this constraint alters the trader's value function. It was shown (as expected) the constraint led to a suboptimal value function compared to the unconstrained problem.

The model under a power intensity had a self-contained constraint as the problem was defined only for positive δ. We had to tackle the derivation of this problem differently, as well as modify the numerical methods used. The previous HJB equation derived under the exponential intensity became a variational inequality problem, given the fact that the trader could liquidate immediately by sending $\delta \to 0$. With this modified form, we could no longer perform the transformation from an HJB equation to a non-linear PDE with explicit non-linear term. We thus had to use a numerical method to find the optimal, and then solve the PDE, as well as checking the variational inequality constraint. The numerical solutions both had similarities and differences to those found with the negative exponential intensity.

Introducing the concept of trading explicitly with both limit orders and market orders allows the trader to execute his order immediately but at a discount to the current price, if he so chooses. Mathematically this led to an interesting problem in the form of a free-boundary problem, which can be related back to other areas of mathematical finance such as in American options. It also led to the need of using a PSOR method. From a financial perspective, the solutions obtained were intriguing. In the introduction of this chapter we hypothesised that the trader would take advantage of having the ability to 'early exercise' given he is risk-averse. This hypothesis was based on the observations of the trading strategies we found in previous chapters. In our numerical examination we found this hypothesis to be correct. The trader would execute with a market order when his position was becoming more risky, such as when the asset price was larger (and thus the absolute volatility was larger), his inventory holdings was larger, or when the time-to-expiry was not near zero, as the asset would have a higher probability of moving against him. Finally, we note that the problem of including market orders is of interest to us under negative exponential intensity but not under power intensity. Under a power intensity, as the optimal control approaches zero the rate of trading approaches infinity and thus immediate execution occurs at the reference price. Thus the trader would never choose to sell at a discount when he can obtain a sale at the reference price.

Chapter 9

Calibration Techniques and Backtesting

In the thesis thus far we took as a baseline the model of Guéant et al. (2012b) and derived various extensions, which were examined theoretically using a combined approach of asymptotic and numerical techniques. In this chapter we shall first look at calibration techniques for the model proposed in chapter 4, as well as for the extension of the intensity function to a power function as discussed in section 8.2. In the second part of this chapter we will backtest the strategies. This will consist of three tasks: firstly, we use limit-order book data to estimate parameters for the intensity of the Poisson process. Secondly, we use the estimated parameters in the PDE and solve for the optimal strategies. Finally, we backtest the optimal strategies on market data to analyse their performance.

Given the statistical nature of the problem, interpreting the different variables is far from trivial. This is due to the model having a continuous intensity for each side of the order book, as well as the price process being continuous, as opposed to a tick-by-tick representation which is present in practice.

Calibration of the model of Guéant et al. (2012b) is discussed in Fernandez-Tapia (2015), which is the PhD thesis of one of the authors of Guéant et al. (2012b), in which methods for parameter estimation using level-one order-book data are examined. Level-one order-book data refers to the data at the best-bid/ask level. The improvement in our analysis is the use of order-book data which includes levels higher than level one. Thus the data we have consists of the best bid/ask, the second best

bid/ask and so on up to the tenth queue (level) at each side. Having more levels gives us more quantitative insight into the behaviour of the limit-order book. However, as mentioned in chapter 1, obtaining such data is difficult and thus in this study we only have a limited amount to work with. As such this chapter outlines the methods that can be used when the whole book is considered. An area of future work could be to revisit this problem when more data is available.

9.1 Calibration of parameters

In the modelling thus far we have considered two different aspects to model, the price process and the liquidity. The former is driven by a geometric Brownian motion (in this analysis), which is continuous in space, and the latter by a Poisson Process, which is discrete in space. Both are continuous in time; this is an issue we shall first address.

In practice trading is both discrete in space (due to tick size) and discrete in time. Orders sit in a queue at a particular level and changing position too often reduces the actual chance to be reached by a market order as we would be moved to the back of the queue each time an order is changed. Therefore we must consider our model in a discrete way. In terms of prices, our asking price must not be between two ticks; we shall round the optimal quotes to the nearest tick. In terms of time, an order is sent to the market and is not cancelled nor modified for a given period of time Δt, unless a market order fills the order. If a trade occurs and changes the inventory or when an order stays in the order book for longer than Δt, then the optimal quote is updated and, if necessary, a new order is inserted.

Let us first focus on the calibration of the intensity function of the Poisson process, and later in this chapter revisit the discrepancy between the continuous modelling of the geometric Brownian motion and the discrete world in which trading occurs.

9.1.1 Calibration of λ, l and α

In our modelling we assume assets in the limit-order book, at a price $S \times (1 + \delta)$, are filled at a rate $\Lambda(\delta)$. S refers to the reference price of the asset and we take it as the best-ask price. It could also be taken as the best-bid or the mid-price. The intensity

$\Lambda(\delta)$ is defined by (4.5) for exponential intensity, which takes the form

$$\Lambda(\delta) = \lambda e^{-l\delta}, \quad l > 0,$$

and (8.27) for power intensity, which takes the form

$$\Lambda(\delta) = \frac{\lambda}{\delta^{\alpha}}, \quad \alpha > 1.$$

In our calibration we assume the intensity function to be constant over an interval ΔT for fixed δ. This indicates that the change in price over this interval ΔT, and thus the change in instantaneous probability of being executed as the price moves towards or away from our limit order, is encapsulated within. This is in contrast to the model of calibration derived by Fernandez-Tapia (2015) who accounts for the change in price over the interval ΔT, and thus the change in intensity over the interval ΔT, which we shall discuss later in the chapter.

For the case of exponential intensity we have

$$\Lambda(\delta, t, t+\Delta t) = \int_{t}^{t+\Delta T} \lambda e^{-l\delta} dt \approx \lambda e^{-l\delta} \Delta T, \tag{9.1}$$

from which we can deduce

$$\log(\Lambda(\delta, t, t+\Delta t)) = \log(\lambda) - l\delta + \log(\Delta T). \tag{9.2}$$

Given Λ has units per unit time, we can estimate Λ for a given unit of time and scale it accordingly to another unit of time. Thus, setting the unit of time equal to ΔT, we can reduce (9.2) to

$$\log(\Lambda(\delta, t, t+\Delta t)) = \log(\lambda) - l\delta. \tag{9.3}$$

Denoting $\hat{\Lambda}(\delta)$ as an estimate for $\Lambda(\delta, t, t+\Delta t)$, we can formulate the problem as a least squares problem (see Friedman et al., 2001) such that

$$r(\lambda, l) = \sum_{\delta_i} \left(\log\left(\hat{\Lambda}(\delta_i)\right) - \log(\lambda) + l\delta_i \right)^2, \tag{9.4}$$

where the objective is to choose λ and l such that $r(\lambda, l)$ is minimised. A similar approach can be taken for the power intensity which results in minimising

$$r(\lambda, l) = \sum_{\delta_i} \left(\log\left(\hat{\Lambda}(\delta_i)\right) - \log(\lambda) + \alpha \log(\delta_i) \right)^2, \tag{9.5}$$

with $\hat{\Lambda}(\delta)$ as an estimate for

$$\Lambda(\delta, t, t+\Delta t) \approx \frac{\lambda}{\delta^{\alpha}}. \tag{9.6}$$

9.1.2 Estimate of $\hat{\Lambda}(\delta)$

We now compute the intensity of the Poisson process defining the number of market orders arriving at a price $S \times (1 + \delta)$. Under the hypothesis that the events taking place in each interval are independent from the events taking place on the next one, the problem reduces to estimating the intensity of a Poisson process. As suggested by Fernandez-Tapia (2015), there are two methods to calibrate this problem, as an external observer, or as a participant:

- As an external observer we do not consider our trades in the market. We use purely historical data from the limit-order book to examine the frequency that other participant's trades were filled at given distances from the reference price.

- As a participant we can observe if our trades are executed or not, under certain assumptions. We ignore other participant's orders and assume their trading is encapsulated in the movement of the asset price. We can then place orders as we please in the limit-order book and observe the *waiting time* until they are filled, which occurs when the reference price hits our asking price. We assume the waiting time for orders that are not filled is equal to the duration of the time remaining since placed.

We thus have two methods to calculate $\hat{\Lambda}(\delta)$. The first counts the number of events, the second is based on waiting times.

For both approaches Fernandez-Tapia (2015) uses level-one order-book data. For the second approach only the reference price is needed and thus level-one data is sufficient. However, we can extend the first approach by including more levels of the limit-order book.

Estimation by counting trades

This corresponds to being an external observer. Over an interval ΔT, we can count the number of trades that occur which were entered into the limit-order book at price $S \times (1 + \delta)$ for a fixed δ, and stayed in the book until filled or cancelled. Note we only count filled trades and not cancelled orders. Thus let X_k^δ represent the total number of shares traded that occurred over the interval $[k\Delta T, (k-1)\Delta T]$ which were entered

at a price δ percent higher than the reference price at the time the order was placed. The estimator $\hat{\Lambda}(\delta)$ is given by

$$\hat{\Lambda}(\delta) = \frac{1}{n}\sum_{k=1}^{n} X_k^{\delta}. \tag{9.7}$$

By the strong law of large numbers, it converges almost surely towards $\Lambda(\delta)$, and it was shown in Fernandez-Tapia (2015) to be unbiased.

Estimation by waiting times

In this approach level-one limit-order book data is used to estimate the waiting time τ_n^{δ} for a trade to occur which was entered into the limit-order book at price $S \times (1+\delta)$ for a fixed δ. We can enter the trade into the limit-order book and observe whether the order is executed by following the path of the reference (best-ask) price, which is by definition the level-one data we have. Let T be the size of the time window for which we have data for, in units corresponding to those we want to estimate $\hat{\Lambda}(\delta)$. Let $\tau_1^{\delta}, \ldots, \tau_n^{\delta}$ be a sequence of stopping times which correspond to orders being filled. We define $X_n^{\delta} = \min\left(\tau_n^{\delta}, \Delta T\right)$. The maximum likelihood estimator for $\hat{\Lambda}(\delta)$ is given by

$$\hat{\Lambda}(\delta) = \frac{\sum_{i=1}^{n} 1_{X_i^{\delta} < \Delta T} \times ATS}{\sum_{i=1}^{n} X_i^{\delta}}, \tag{9.8}$$

where ATS is the Average Trade Size. As we only know when orders are filled, and do not know the trade size, we assume the trades are the ATS. Shares are rarely sold individually and normally in bundles of 100, thus we can assume each time a sell occurs it is a bundle of 100, or whatever the ATS is. Using this, (9.7) and (9.8) should be approximately equal. It was shown in Fernandez-Tapia (2015) that (9.8) is the maximum likelihood of $\Lambda(\delta)$, converges towards $\Lambda(\delta)$ as $n \to \infty$, and is a biased estimator.

Discussion

The key issues we will discuss here include how our model compares to that in Fernandez-Tapia (2015), and the assumptions we make on the filling of orders when calculating $\hat{\Lambda}(\delta)$.

It was mentioned above that Fernandez-Tapia (2015) accounts for the change in price during the interval ΔT which affects the probability of the order being filled,

while we assume this is encapsulated in the intensity. For the model of Fernandez-Tapia (2015) the intensity over a period ΔT (under the case of exponential intensity) is given as

$$\Lambda\left(\delta, t, t+\Delta t\right) = \int_{t}^{t+\Delta T} \lambda e^{-l\delta + l(S_u - S_t)} dt, \tag{9.9}$$

whereas we have it defined by (9.1). The expectation of the stochastic term in (9.9), which is not present in (9.1), can be found explicitly for standard Brownian motion (which is what is used in the model of Fernandez-Tapia, 2015). In terms of doing the regression (see (9.4)) the expectation of the stochastic term acts as an additional intercept and thus does not change the remainder of the calibration methodology. If we were to use the formula (9.9), adjusted for relative δ and geometric Brownian motion, we would also be able to formulate an explicit solution for the expectation of the stochastic term. However under power intensity, the expectation is undefined for $\alpha > 1$. As discussed in section 8.2, $\alpha > 1$ are the values this problem is defined for and, as we will shortly see, are the values that we find empirically. Therefore, given we cannot use the same method as Fernandez-Tapia (2015) for both the exponential intensity and power intensity, we make the assumption of constant intensity over the interval ΔT so we can examine both intensity functions together, on a comparative level.

For the estimation by counting trades method, Fernandez-Tapia (2015) estimates $\hat{\Lambda}\left(\delta\right)$ by counting all trades within the interval ΔT that are filled at a distance δ from the initial reference price. However, it is not necessarily true that all those orders were entered a distance δ from the initial price as they may have been entered within the interval ΔT in which the price is not constant and may have moved. In our approach we include higher levels of the limit-order book which allows us to see exactly the distance from the reference price that the trade was entered, and thus count exactly how many orders were filled over an interval ΔT that were placed a (relative) distance δ from the reference price.

For the estimation by waiting times method, we use the same method as discussed in Fernandez-Tapia (2015). The flaw in this method is that it does not consider position in the queue, and thus we need to make an assumption on when the order is filled. We consider two methods, an optimistic approach and a pessimistic approach. The optimistic approach is to assume the order is filled once the reference price hits

the asking price, indicating our order is in the front of the queue. The pessimistic approach is to assume our order is at the end of the queue and is only filled when the entire queue is filled, which corresponds to the reference price hitting an order in the next queue, which has a price higher than our asking price. We found the former approach is more in line with the results found in the estimation by counting trades method, as the latter method effectively does not hit any orders at the best ask which is where the majority of orders are filled. Thus we use the optimistic approach, as also used by Fernandez-Tapia (2015).

9.1.3 Calibration of price process parameters

So far we have considered the calibration of the intensity function. In this section we discuss the calibration of the price process parameters.

Firstly, given the short time-scales we work over, we set $\mu = 0$. If we were to find the annual drift rate of a stock and rescale it to the time-frame we were examining we would find it be to negligible in most cases. This is also the case implemented by Fernandez-Tapia (2015).

Secondly, the problem of estimating volatility using limit-order tick data is far from trivial and there is still ongoing research into the best approach. The main issue arising when estimating the volatility using limit-order data is that the standard realised volatility estimator,

$$\hat{\sigma}^2 = \frac{1}{t_N - t_0} \sum_{i=1}^{N} \left(S\left(t_i\right) - S\left(t_{i-1}\right) \right)^2, \tag{9.10}$$

is affected by microstructure noise and does not converge to an accurate proxy for risk (see Bacry et al., 2013). A popular approach for dealing with this is to model microstructure noise using the concept of *latent price*. Rather than being able to observe the efficient price, $S\left(t\right)$, driven by some sort of Brownian motion, the trader is able to observe a noisy version, $\tilde{S}\left(t\right)$. Gloter and Jacod (2001a,b) suggest an additive microstructure noise model which has been examined in the context of financial data by Aït-Sahalia et al. (2005); Zhang et al. (2005, 2006) to name a few. The goal is to separate the noise from the true signal, as the volatility of the true signal can be calculated using (9.10).

An alternative method, implemented by Fernandez-Tapia (2015), is to extend the work of Garman and Klass (1980) who develop an estimation of the volatility by using daily high, low, open and close data. The extension is adapting it to the frequencies of several minute intervals, where the high, low, open and close is found in each interval.

Given our lack of limit-order book data, we use daily data (which is widely available online, e.g. https://uk.finance.yahoo.com/) to calculate the realised volatility using (9.10). We adjust this to our given time frame by using the square-root-of-time rule, which assumes normality.

Lastly, we come to setting a value for γ, the risk-aversion parameter. This is a mathematically derived parameter and in practice cannot be calibrated using standard methods. As such, deciding its value is subjective. We set $\gamma = 0.05$ which is the value chosen by Guéant et al. (2012b).

We stress that the significance of this chapter is to calibrate the intensity functions using higher-level limit-order book data. For a more in-depth discussion on calibrating the price process parameters using tick data we refer to the references above.

9.1.4 Data

The data we obtained is limit-order book data with ten levels on each side of the book, i.e. ten levels in the ask side and ten levels in the bid side. However, it is only the ask side we focus on in this analysis as we are only considering liquidation. Repeating for the case of acquiring an order (i.e. buying) would be trivial.

The data was obtained from LOBSTER (2015) which provides limit-order book data on Nasdaq listed stocks. The format of the data can be seen in figure 9.1, with an explanation of variables given in table 9.1. The data we have is one trading day worth of data for five different names, those being Amazon Inc. (AMZN), Google Inc. (GOOG), Microsoft Corporation (MSFT), Intel Corporation (INTC) and Apple Inc. (AAPL). Although it is only one day's data for each name, the number of trades are in the order of magnitude of $O\left(10^6\right)$ for each name. However, given we only have one day for each name we must assume stationarity of the intensity throughout the trading day. An interesting extension, which would require more data, would be to examine various trading times such as morning, midday and evening, and examine the U-shape or W-shape intra-day volume curve that has been found in empirical literature (see,

Time	Type	Ref	Size	Price	Buy/Sell	Best Ask	Size	Best Bid	Size
34210.706474037	1	17017091	29	2240300	-1	2240300	29	2239200	364
34210.910418382	3	17017091	29	2240300	-1	2240700	100	2239200	364
34210.910418382	1	17030256	29	2240400	-1	2240400	29	2239200	364
34210.966978058	4	17030256	29	2240400	-1	2240700	100	2239200	364
34210.966978058	5	0	71	2240400	-1	2240700	100	2239200	364
34211.618610908	1	17063006	400	2240700	-1	2240700	500	2239200	364
34211.943114597	4	16728917	100	2240700	-1	2240700	400	2239200	364
34211.943114597	4	17063006	300	2240700	-1	2240700	100	2239200	364
34211.967039643	4	17063006	100	2240700	-1	2241000	1100	2239200	364
34211.967968333	4	16681306	100	2241000	-1	2241000	1000	2239200	364
34211.970165989	4	16681306	100	2241000	-1	2241000	900	2239200	364
34212.161685865	1	17084297	100	2239300	1	2241000	900	2239300	100
34212.165104681	3	16681306	900	2241000	-1	2242400	20	2239300	100
34212.165315424	5	0	317	2239500	1	2242400	20	2239300	100
34212.165315424	5	0	100	2239400	1	2242400	20	2239300	100

Figure 9.1: Order-book data for AMZN on 21/6/2012.

for example, Wood et al., 1985; Harris, 1986; Jain and Joh, 1988).

Data manipulation

As we are considering the sell side only we first removed all buy orders, i.e. those corresponding to +1. A buy market order filling a sell limit order is labelled as -1. As we are only interested in orders that are filled, we remove cancelled orders (both fully and partial). As we cannot tell when hidden orders were inserted into the book we remove these also. Given the lack of data, we did not encounter any trading halt indicators, which occur when the exchange suspends trading.

We calculate δ by looking at the difference in price to which a sell limit order was placed and the current best ask (and divide this by the current best ask, given it is relative). We then need to find the orders that were filled; this is doable given each order has a reference number. By matching reference numbers, we can observe whether the order was filled or cancelled, or if it remained in the book until the end of the trading day (which for our analysis can be consider unfilled/cancelled).

For the ATS we instead use the median which is equivalent to 100, which is the same for all names and so makes the results comparable. The mean of all names is close to 100, but not exact.

Variable	Meaning
Time:	Seconds after midnight
Type:	
1:	Submission of a new limit order
2:	Cancellation (partial deletion of a limit order)
3:	Deletion (total deletion of a limit order)
4:	Execution of a visible limit order
5:	Execution of a hidden limit order
7:	Trading halt indicator
Reference	Unique order reference number
Size	Number of shares
Price	Price corresponding to order, times 10000
Buy/Sell	+1 buy limit order, -1 sell limit order
Best Ask	Best-ask price, times 10000
Size	Size of best-ask queue
Best Bid	Best-bid price, times 10000
Size	Size of best-bid queue

Table 9.1: Variable explanation for Lobster Data (figure 9.1)

Calibration

Using the methods discussed above we fit exponential and power intensity functions to the data we have available. We consider an hourly time unit for this analysis so the unit for the intensity is per hour. The intensity is also normalised to the ATS (which as we discussed above is in fact the median trade size). We do this for both the method of a participant (part) in the market, and that of an external observer (obs) in the market.

The calibrated parameters can be seen in table 9.2. We see that the parameters found using each of the methods are similar for a given name, but not exactly equal. This is because using the observation method (9.7) we find the exact amount of shares that were executed for each trade, while using the participation method (9.8) we assume the volume of each trade is equal to the ATS. We can see from the results of table 9.2 that this assumption is plausible.

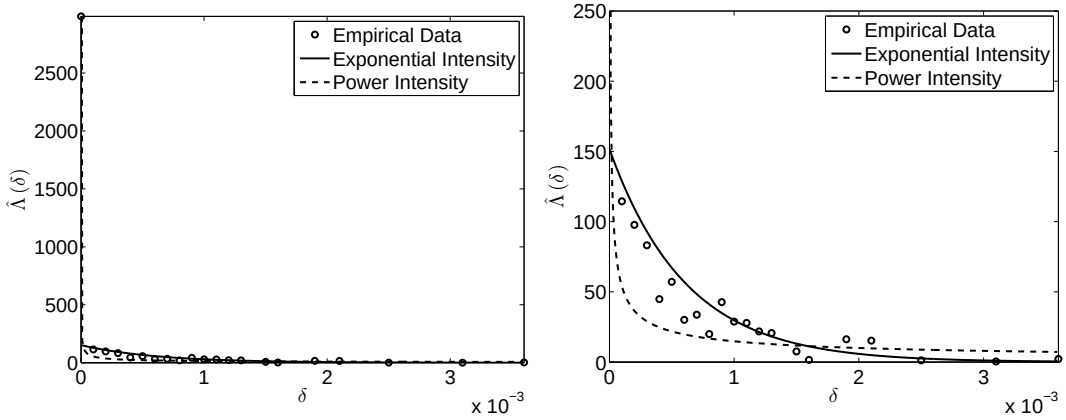
We include a plot of the intensity functions fitted to the data, using least squares as discussed above. We do this for the parameters obtained from the observation method, given we find it of more interest as it requires higher levels of limit-order book data. However both methods give comparable results, as seen in table 9.2. The plots are shown in figure 9.2 for Amazon Inc (AMZN). The original plot is shown in

	AMZN		GOOG		INTC		AAPL		MSFT	
Method	Part	Obs	Part	Obs	Part	Obs	Part	Obs	Part	Obs
λ_{exp}	152	119	355	306	1226	1837	313	494	3173	4219
l_{exp}	1631	2428	5978	6311	3331	1565	2723	4551	1725	1857
R^2_{exp}	0.71	0.72	0.74	0.85	0.73	0.54	0.64	0.66	0.66	0.65
λ_{pow}	0.30	0.02	0.39	0.09	0.04	1.41	0.33	0.04	3.34	3.42
α_{pow}	1.56	1.76	1.49	1.59	1.78	1.59	1.59	1.76	1.55	1.55
R^2_{pow}	0.62	0.52	0.85	0.75	0.9	0.66	0.84	0.53	0.71	0.62

Table 9.2: Calibrated parameters for various names using both the participation (part) method, and external observer (obs) method

figure 9.2(a) with a close-up, excluding the first point, in figure 9.2(b). We see that the first point is an order of magnitude greater than the second, with the first point value approximately 3000 and second approximately 125. This is because the majority of limit orders are placed at the best ask as they have a much higher probability of being executed, given anyone trading with a market order will fill a trade at the best-ask queue until the queue is depleted.

Given this property we come to a discussion on which intensity function is more suitable. Given the evidence in figure 9.2 there are arguments for and against both. The R-squared of the regression on AMZN for the exponential intensity is 0.7095 while the R-squared for the power intensity is 0.6184. The higher R-squared of the exponential intensity is due to it having a better fit to points excluding the first (best-ask) point; it does not approximate the best-ask intensity well. In contrast, although it has a lower R-squared, and does not fit the majority of the data very well in comparison to the exponential intensity, the power intensity function fits the best-ask point, in which most trading occurs, very well. It tends to infinity as $\delta \to 0$ which is a reasonable approximation given the change in orders of magnitude between the first point, and the remaining points as we look to trade further in the book. Thus, one conclusion is that the power intensity is a better fit for a passive trader who trades a lot at the best ask while the exponential intensity could be better suited to a passive trader who trades using queues other than the best ask. However, as can be seen in table 9.2, it is not always the case that the exponential intensity R^2 is higher than the R^2 of the power intensity. We believe more data is required before a significant conclusion can be reached.



(a) Estimated Intensity Function (b) Exlcuding First Point

Figure 9.2: Intensity functions fitted to AMZN data, with parameters in table 9.2

Given the difference in orders of magnitude between the best ask and the remaining points, one future development could be to look at relaxing the least-squares assumption of equal importance on all points and use weighted least-squares. More weight would be given to the best-ask point and the weight would decrease as we move further down the book. For a discussion on weighted least-squares see Friedman et al. (2001).

9.2 Historical simulation

In this section we look at backtesting our model. We mentioned in the introduction of this chapter that this consists of three tasks

1. Calibrate the necessary parameters,
2. Find the optimal trading strategies,
3. Backtest the strategies on historical data.

The calibration of these parameters was discussed in the previous section. Once these parameters are found we substitute them into the code for solving the model developed in chapter 4 (with the extension to power intensity discussed in section 8.2). This process will return to us the optimal trading strategies which are functions of time, asset price and inventory. With these we can use the methods discussed in this section to backtest the model.

We shall begin by examining some of the properties of the expected inventory, before concluding with a study of a trading simulation.

9.2.1 Expected inventory and trading rate

It is of interest to investigate the expected inventory and trading rate of our calibrated models, as this will give us insight into the contrasting behaviour between them. If we first calculate the expected inventory, we can calculate the expected trading rate as its derivative. We use Monte Carlo simulations to calculate this expectation. Given the trading rate is taken as the derivative of a simulation, it will not be smooth due to the numerics, but should be smooth as the number of simulations, N, becomes large.

From section 2.1.3 we have that our process follows a Poisson distribution

$$\mathbb{P}\left(N\left(t\right)-N\left(s\right)=n\right)=e^{-m(s,t)}\frac{m\left(s,t\right)^{n}}{n!}, \tag{9.11}$$

with $m\left(s,t\right)=\int_{s}^{t}\lambda\left(u\right)du$. We previously noted that trading is a discrete time problem. In this case the intensity is constant over an interval Δt, with $m\left(t,t+\Delta t\right)=\lambda\left(t\right)\Delta t$. Therefore (9.11) reduces to

$$\mathbb{P}\left(N\left(t+\Delta t\right)-N\left(t\right)=1\right)=e^{-\lambda(t)\Delta t}\lambda\left(t\right)\Delta t=\lambda\left(t\right)\Delta t+O\left(\Delta t^{2}\right)\approx\lambda\left(t\right)\Delta t \tag{9.12}$$

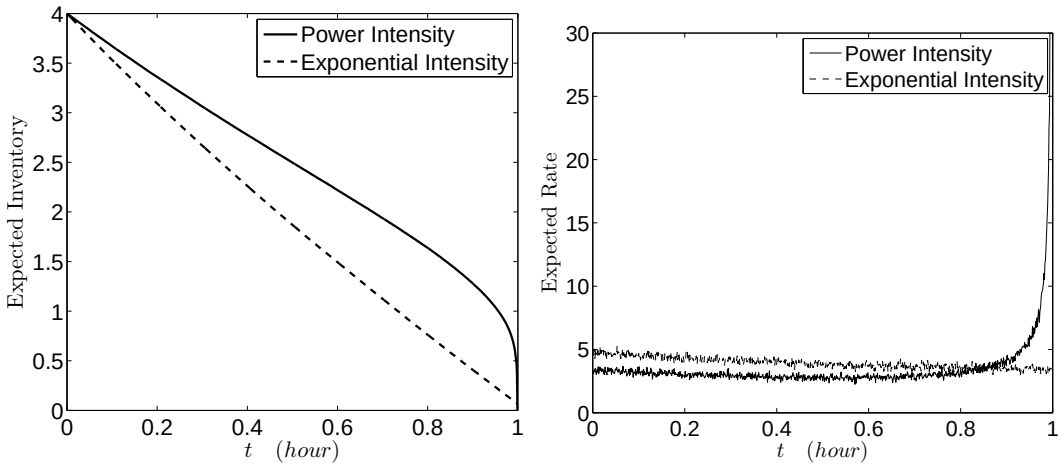
ignoring $O\left(\Delta t^{2}\right)$ terms. Assuming the control, δ, is constant over an interval Δt, given we cannot update our orders continuously, we can use Monte Carlo simulation to calculate the expected inventory.

Assuming our inventory follows a Poisson process with intensity $\hat{\Lambda}\left(\delta\right)$ we can calculate the probability of a jump occurring over the interval Δt as

$$\mathbb{P}\left(N\left(t+\Delta t\right)-N\left(t\right)=1\right)\approx\hat{\Lambda}\left(\delta\right)\Delta t, \tag{9.13}$$

where δ is fixed in the interval Δt. We thus calculate the expected inventory as follows:

- Repeat N times:
  - While $q(t)>0$ and $t<T$
    1. Generate uniform random variable $X\sim U\left(0,1\right)$
    2. If $X<\hat{\Lambda}\left(\delta\left(t,q,S\right)\right)\Delta t$ increment $q(t)=q(t)-1$
    3. Increment time: $t=t+\Delta t$
- Average $q\left(t\right)$ over N simulations resulting in expected inventory $\hat{q}\left(t\right)$



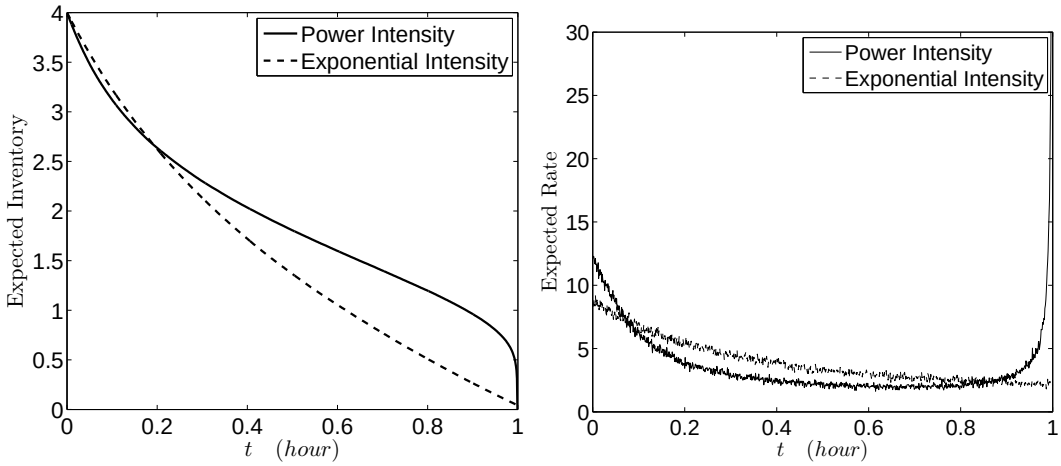
(a) Expected Inventory (b) Expected Trading Rate

Figure 9.3: Expected Inventory and Trading Rate fitted to AMZN data, with parameters found in table 9.2, for $S = 5$

As the optimal strategy is a function of S, we will assume S is constant over the time frame we examine. We can thus examine the expected inventory and trading rate for various values of S. We outline how the solutions vary for various S, which was indeed one of the motivations for carrying out this study, given the previous work of Guéant et al. (2012b) has asset-independent trading strategies.

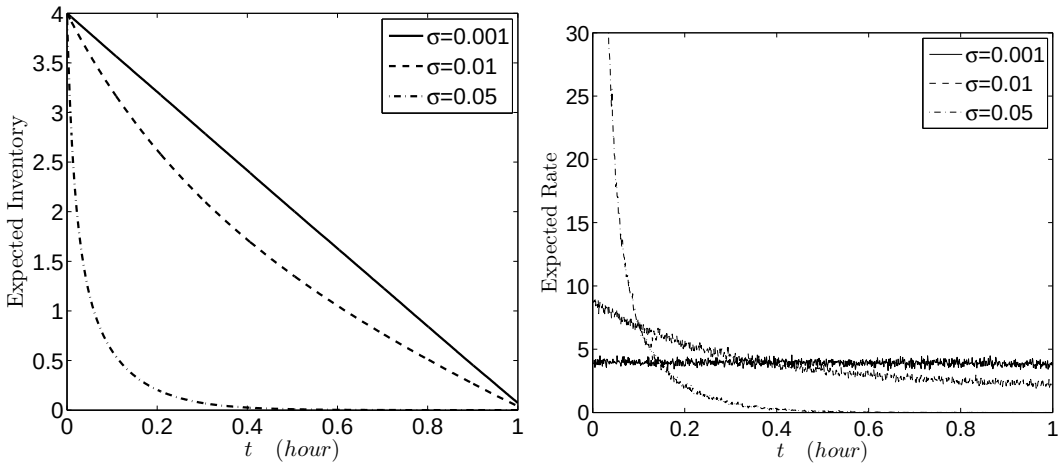
Figure 9.3 is a plot of the expected inventory (figure 9.3(a)) and expected trading rate (figure 9.3(b)) for $S = 5$. It can be seen that the expected inventory under exponential intensity decreases almost linearly, while under power intensity the trader initially trades slower as he has the ability to liquidate quickly at the terminal time, in which indeed he does. Turning to the expected trading rates, we see the exponential intensity trading rate is approximately linear with a slight negative slope. For the power intensity we see the rate is approximately linear up to some time when the trader decides he should start liquidating more quickly and his trading rate increases exponentially fast.

We can see contrasting behaviour for a larger value of S, remembering the trader is risk-averse. Figure 9.4 is a plot of the expected inventory (figure 9.4(a)) and expected trading rate (figure 9.4(b)) for $S = 100$. Under the exponential intensity we see the trader starts trading fast, and his rate of liquidation decreases as he approaches the terminal time, and as his inventory decreases. Under the power intensity we see a comparable and contrasting strategy. Initially the trader trades fast as in the



(a) Expected Inventory (b) Expected Trading Rate

Figure 9.4: Expected Inventory and Trading Rate fitted to AMZN data, with parameters found in table 9.2, for $S = 100$



(a) Expected Inventory (b) Expected Trading Rate

Figure 9.5: Expected Inventory and Trading Rate fitted to AMZN data, with parameters found in table 9.2, for varying σ

exponential intensity case. He then slows down to a rate which is slower than that under the exponential intensity case. However, towards the end of the trading period, as was the case for $S = 5$, the trader rapidly increases his rate of liquidation to ensure he has completely unwound his position by the terminal time.

Finally we will examine how the expected inventory and trading rate differ for various volatility regimes, with σ taking one of the three values $\sigma = 0.001$, 0.01, 0.05 *per hour*$^{1/2}$. We do this for the exponential intensity only, as a similar analysis under power intensity yields analogous conclusions. The resultant expected inventory and trading rate can be seen in figure 9.5. As expected, increasing the volatility has comparable results to examining the differences between increasing values of S, given the trader is risk-averse. For higher volatility levels the trader liquidates at a much quicker rate, compared to lower volatility levels.

9.2.2 Trading simulation

In this trading simulation we use the calibrated parameters as found in section 9.1 to calculate the optimal trading strategies. Once these are found we can use them to decide where to place orders in the limit-order book, examine if these trades get filled, and analyse the performance of the strategy.

First we must make some assumptions about the trading algorithm. Some of the same assumptions that were used in the calibration must hold for the trading simulation. In terms of prices, our asking price must not be between two ticks and we shall round the optimal quotes to the nearest tick. In terms of time, an order is sent to the market and is not cancelled nor modified for a given period of time Δt, unless a market order fills our order. If a trade occurs and changes the inventory, or when an order stays in the order book for a period Δt, the optimal quote is updated and, if necessary, a new order is inserted. We assume that the entire order is filled when a trade occurs at or above the quoted price [1]. We must also note that given the limitation of data we have there is a bias in our simulation. This is because we use the same data for calibration as we do for testing the simulation. Ideally we would

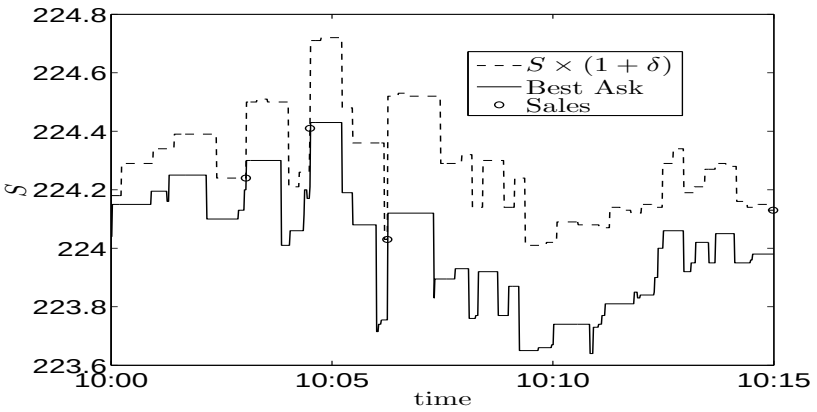
[1] Changing this assumption to strictly above the quoted price (so we assume the order is at the back of the queue) was something we examined. It is trivial to change in the simulation if we so please. Another extension to be considered would be to assign a probability to an order being filled at the quote price, such as a Bernoulli random variable to establish the outcome. We deem this outside the scope of this study but may be considered as future work.

prefer to perform an out-of-sample test in which the in-sample is used for calibration and the out-sample is used for testing.

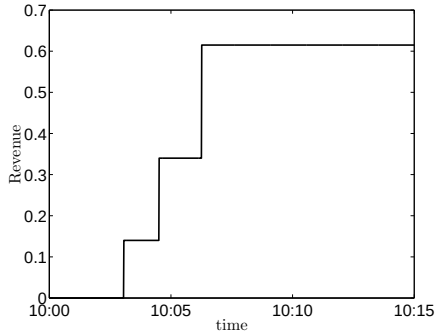
Given the above assumptions we can discuss the framework of this simulation. The data we need is the best-ask (reference) price for the asset we are considering, which will be Amazon Inc (AMZN). We consider two different time frames for the simulation. The first is liquidating a portfolio of 4 units of the asset, with each unit representing the ATS, thus $q(0) = 4$. We seek to liquidate this position before the terminal time of 15 minutes, so $T = 0.25$. We examine this under the exponential intensity and power intensity. Next we look at liquidating a larger portfolio consisting of 20 units of the asset, thus $q(0) = 20$. We seek to liquidate this position before the terminal time of two hours, so $T = 2$. The length of time Δt discussed above is equal to 60 seconds, so $\Delta t = \frac{1}{60}$ given the unit of time is an hour; consistent with Fernandez-Tapia (2015).

For each simulation we show a plot containing the evolution of the best-ask price, the evolution of our asking price and the execution of our trades. We also show a plot of the inventory process in time (with the curve of the expected inventory superimposed), as well as a plot of the revenue accumulated from implementing the strategy. This is seen in figure 9.6 under exponential intensity and figure 9.7 under power intensity for liquidation for $q(0) = 4$ assets in 15 minutes. We see that the trading curves are approximately similar, with the exponential intensity model generating more revenue.

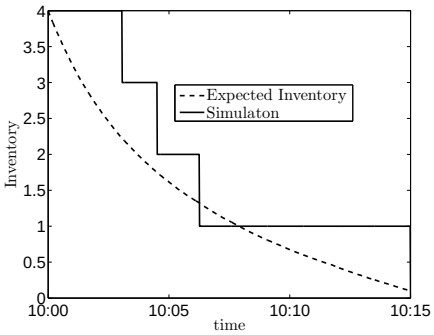
Figure 9.8 shows the plots for the larger simulation in which the trader is aiming to liquid 20 units of an asset over a period of two hours. We see the trader begins by selling rapidly, with his inventory curve following closely the curve of his expected inventory. On a per asset basis this is not the most profitable part of his strategy but as he has a large market inventory, and is risk-averse, he wants to liquidate quickly. Approximately half his inventory is unwound in the first 15 minutes. After this period the rate of trading decreases substantially. He places limit orders further from the best ask as to try profit from the strategy given the remaining time he has. During the bullish (market rally) period between 11:15am and 11:30am we see he is able to sell at prices significantly higher than the best ask given the price is increasing and hitting his orders. He generates approximately the same revenue in this period from three orders as he did in the first 15 minutes from ten orders. During the bearish (market decline) period that follows he is unable to liquidate. However, given he has



(a) Evolution of price processes



(b) Evolution of revenue



(c) Evolution of inventory

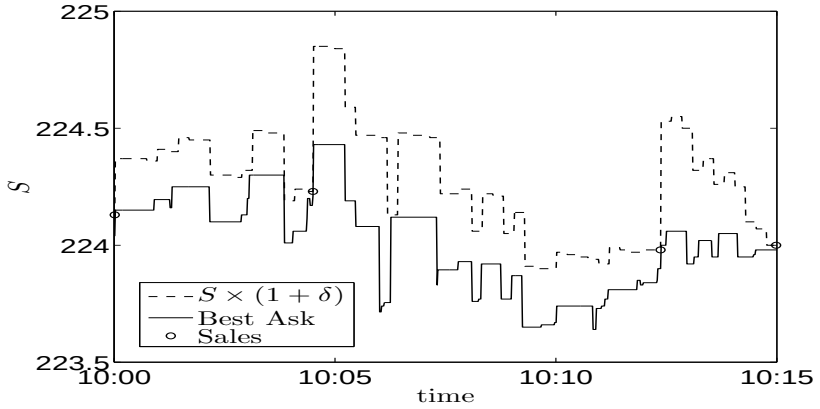
Figure 9.6: Backtest example with $q(0) = 4$ under exponential intensity

only a few units remaining, his asking price would be substantially above the current best ask. As time remaining decays he is able to unwind a few more positions, given he drops his asking price as he approaches the terminal time, and is left with only one asset remaining at the terminal time.

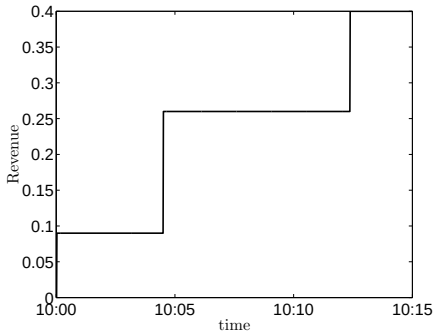
We performed the same analysis on the other names and obtained similar results. This is expected given they are all highly liquid, large capitalisation names, which are all constitutions of the S & P 500 index. What would be of interest in a future study (if the data was available) would be to consider the performance of the strategy for less liquid names.

9.3 Summary

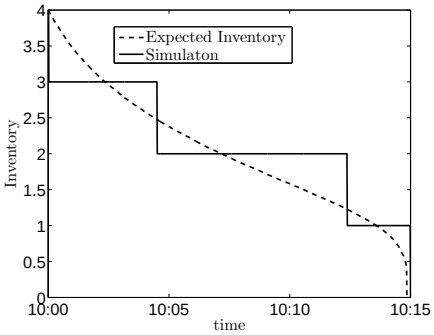
In this chapter we studied the model developed in chapter 4, and extended in chapter 8, from the point of view of its statistical calibration. As opposed to using only level-one order-book data as used in Guéant et al. (2012b), discussed in Fernandez-Tapia



(a) Evolution of price processes



(b) Evolution of revenue



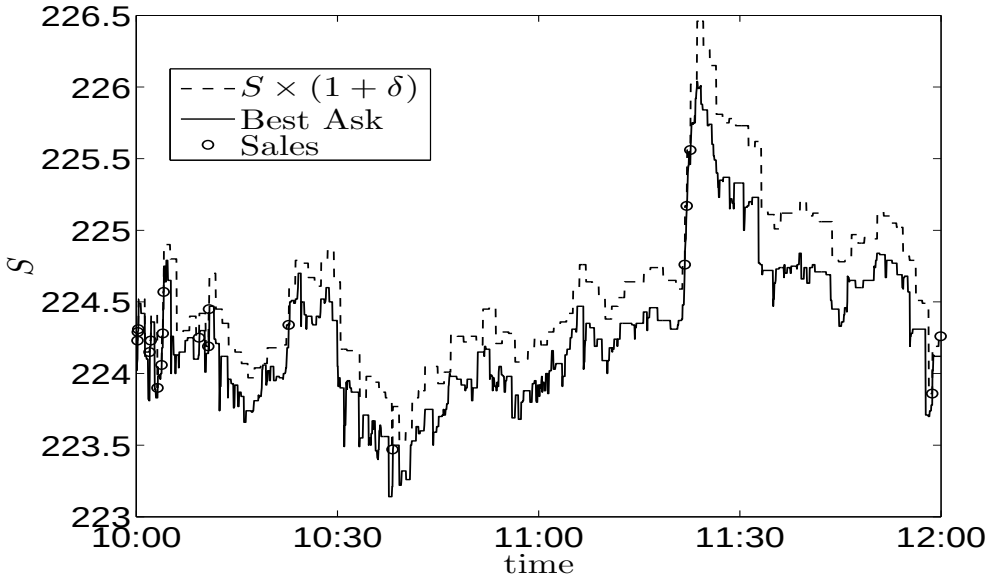
(c) Evolution of inventory

Figure 9.7: Backtest example with $q(0) = 4$ under power intensity

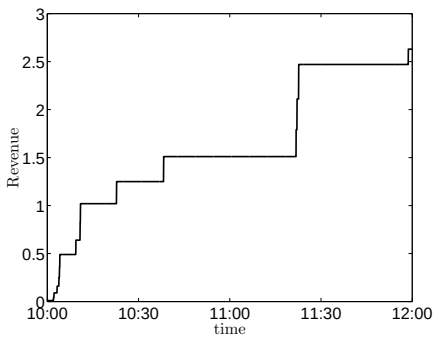
(2015), we take the new approach of calibrating the intensities using ten levels of the order book at each side.

Several issues were discussed in the chapter. Firstly, we addressed the issues of examining our model, which is continuous in time and space, in a discrete way. Secondly, we studied various methods for calibration of the intensity functions, including counting trades and estimating waiting times. The former involved counting trades obtained from historical limit-order book data, while the latter involved placing orders into the book and observing whether or not these orders were executed by observing the empirical price process of the reference price. We also briefly discussed calibration of the price process parameters which is not trivial when considering tick data. Finally, we presented the results of backtesting the calibrated model, which portrayed successful liquidation strategies.

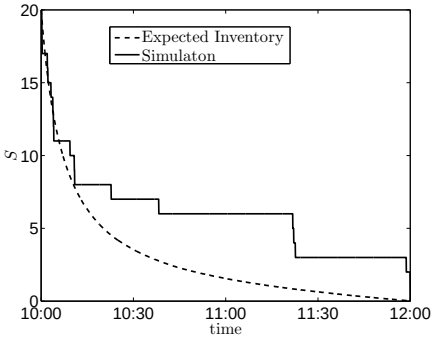
As mentioned previously, the data needed for the calibration outlined in this chapter is difficult to obtain for academics. Financial institutes are hesitant in providing it and data service providers often charge considerable amounts given the demand of



(a) Evolution of price processes



(b) Evolution of Revenue



(c) Evolution of Inventory

Figure 9.8: Backtest example with $q(0) = 20$ under exponential intensity

financial institutions in developing their own models. The lack of access to data was a topical subject at the *Market Microstructure: Confronting Many Viewpoints* conference held in Paris, December 2014. This conference was attended by both academics and practitioners. A key talking point was the need from practitioners for solutions to problems in optimal execution but the unwillingness to provide academics with the data needed to solve these problems.

We thus conclude the chapter by outlining that although the statistical methods within this chapter have been implemented on data, the amount of the data we used is far too less to deem the results statistically significant. An area of future work is to do a more in-depth statistical analysis of the problem, if a larger data set becomes available to allow us to do so.

Chapter 10

Conclusions and Future Work

In this chapter we discuss the main ideas that this work was intended to convey, highlighting the scientific contributions presented, before suggesting some possible directions for future research.

10.1 Conclusions

This thesis has developed the framework posed by Guéant et al. (2012b), who extend the original framework of Avellaneda and Stoikov (2008), by introducing more general diffusion processes and modelling assumptions. By choosing CARA utility functions, standard Brownian motion and an absolute decay parameter for the intensity function Guéant et al. (2012b) are able to cleverly build symmetry into the model from the start, which allow them to reduce the HJB equation to a system of ODEs. In this thesis we used a combined approach of asymptotic analyses and accurate numerical techniques that allowed us to gain insight into solutions when this symmetry was no longer present.

The reduction from an HJB equation to an ODE by Guéant et al. (2012b) allow the authors to solve the resultant differential-difference equation using numerical methods. When we introduced the framework of Guéant et al. (2012b) in chapter 3 we were able to contribute to their original framework by finding an analytic solution for the case of one asset remaining, $q = 1$, which provided insight into the solution that was not realised by Guéant et al. (2012b) in their original study of the problem.

To first extend the framework of Guéant et al. (2012b) we built a model with

geometric Brownian motion (and mean-reverting processes) for the asset price, a proportional control parameter (the additional amount we ask for the asset), and a proportional decay parameter, meaning the symmetry found in Guéant et al. (2012b) no longer existed. This novel combination resulted in asset-dependent trading strategies, which to our knowledge is a unique concept in this framework of literature. This also resulted in not selling the asset for a negative price, which can occur in the work of Guéant et al. (2012b), amongst others. We reduced the resulting four-dimensional HJB equation to a novel three-dimensional non-linear PDE with explicit non-linear term, as well as rescaling the variables to reduce the number of input parameters by two. We carried out interesting analytic (asymptotic) analyses, finding analytic solutions in various limits, and the development of a singularity when considering a steady-state, as well as a complementary parameter constraint necessary for a solution to exist.

We extended the framework of both Guéant et al. (2012b) and that discussed above (chapter 4) to a trader who has a basket of correlated assets to liquidate. This is studied both under standard Brownian motion (chapter 5) and geometric Brownian motion (chapter 6). Under both frameworks we solve the problem numerically and asymptotically, and discuss the resultant trading strategies. We found the correlation term generates some interest, with results relating back to Modern Portfolio Theory (see Markowitz, 1959). For negatively correlated assets the trader aimed for a balanced portfolio while for a positively correlated portfolio the trader aimed to unwind faster (in comparison to negatively or uncorrelated assets) as the positive correlation led to increased market risk. Furthermore, under geometric Brownian motion in the perpetual limit, a change of variables to polar coordinates gave insight into a parameter constraint necessary for a solution to exist, as well as confirming singular behaviour which the numerical results were indicating.

In chapter 7 we extended the work of chapter 4 by replacing the constant volatility with stochastic volatility. Solving numerically, we examined the optimal trading strategies under various parameter regimes to get a full understanding of the effect of the mean-reverting stochastic volatility on the trading strategies. The parameter constraint of chapter 4 contained the (constant) volatility; therefore, we were interested in investigating a similar constraint now that the volatility was stochastic. Using asymptotic analysis we indeed found an analogous parameter constraint.

Chapter 8 examined three extensions of the original model derived in chapter 4. We first introduced a constraint in the value the optimal strategy may take, with $\delta^{min} = 0$, which does not allow the trader to trade at a discount. This resulted in a reformulation of the problem, with interesting (and contrasting) trading strategies. These strategies were suboptimal to those found for an unconstrained strategy. Secondly, we replaced the exponential intensity function with a power intensity function. We had to tackle the derivation of this problem differently, given it is a singular stochastic control problem, as well as modifying the numerical methods used. We compared the resultant trading strategies with those of the exponential intensity model. The strategies were contrasting given, under power intensity, a trader can instantaneously liquidate by sending $\delta \to 0$. Finally, we studied the strategy of a trader who can post both limit orders and market orders throughout the trading period. In this framework, given the trader has the ability to execute immediately, the problem was formulated as a free-boundary problem and could be viewed similar to that of an American option which allows for early exercise. From a financial perspective, the solutions obtained were intriguing. We hypothesised that the trader would take advantage of having the ability to execute with market orders given he is risk-averse, which was the case.

In chapter 9 we calibrated the model under geometric Brownian motion, for both exponential and power intensity functions. The improvement in our analysis was the use of order-book data which includes levels higher than level one, as having more levels gives us more quantitative insight into the behaviour of the limit-order book. We first addressed the issues of examining our model in a discrete way, given it is continuous in time and space. Secondly, we studied various methods for calibrating the intensity functions, discussing the methods already proposed by Fernandez-Tapia (2015), and the contribution we made through the use of multiple levels of the limit-order book. Finally, we presented the results of backtesting the calibrated model, which portrayed successful liquidation strategies.

10.2 Future work

In this thesis we have considered various extensions that could be examined further in a future study.

The first relates to changing the intensity function such that the Poisson process is instead a Hawkes (1971) process. A Hawkes process is a self-exciting point process with intensity function

$$\lambda(t) = \mu(t) + \sum_{t_i<t} \nu(t - t_i),$$

where $\mu(t)$ is a deterministic base intensity and $\nu(t - t_i)$ expresses the positive influence of the past events t_i on the current value of the intensity process. The original study of Hawkes (1971) proposes an exponential kernel of the form

$$\nu(t) = \sum_{j=1}^{p} \alpha_j e^{-\beta_j t}.$$

Hawkes processes have been implemented previously in the optimal scheduling literature for the price process (see Alfonsi and Blanc, 2014) and the order flow (see Bacry and Muzy, 2014), to name a few examples, and would make for an interesting extension to the tactical trading literature. The Hawkes process could model bursts of volume, such that the probability of being executed increases directly after a trade occurs. This could be seen as a more accurate representation of reality, rather than having the independence assumption of a Poisson process.

The second relates to the framework of section 8.3 in which the trader can trade using both market orders and limit orders. In section 8.3 we included an instantaneous penalty for trading using market orders which could be seen as the spread, temporary impact, or both. An interesting extension would be to include permanent and/or transient market impact which would result in a jump in the asset price every time an order is executed using a market order, similar to the market impact assumptions of Almgren and Chriss (2001) and Obizhaeva and Wang (2013).

Finally, if data were to become available, a more extensive calibration study would be of interest. This includes studying whether it is more advantageous for the trader to trade in periods of the day when there is more volume, such as morning or evening, verses periods when there is less volume, such as in the middle of the day, as found empirically by the U-shaped volume curve (see Wood et al., 1985; Harris, 1986; Jain and Joh, 1988). If this was found to be the case, it could be adapted in the model by introducing a local, or indeed stochastic, intensity function to represent this non-stationary intra-day volume.

Bibliography

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial studies*, 18(2):351–416.

Alfonsi, A. and Blanc, P. (2014). Dynamic optimal execution in a mixed-market-impact hawkes price model. *arXiv preprint arXiv:1404.0648*.

Alfonsi, A., Fruth, A., and Schied, A. (2008). Constrained portfolio liquidation in a limit-order book model. *Banach Center Publ*, 83:9–25.

Alfonsi, A., Fruth, A., and Schied, A. (2010). Optimal execution strategies in limit-order books with general shape functions. *Quantitative Finance*, 10(2):143–157.

Alfonsi, A. and Schied, A. (2010). Optimal trade execution and absence of price manipulations in limit-order book models. *SIAM Journal on Financial Mathematics*, 1(1):490–522.

Alfonsi, A., Schied, A., and Slynko, A. (2012). Order book resilience, price manipulation, and the positive portfolio problem. *SIAM Journal on Financial Mathematics*, 3(1):511–533.

Almgren, R. (2012). Optimal trading with stochastic liquidity and volatility. *SIAM Journal on Financial Mathematics*, 3(1):163–181.

Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40.

Almgren, R. and Lorenz, J. (2007). Adaptive arrival price. *Trading*, 2007(1):59–66.

Almgren, R. F. (2003). Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10(1):1–18.

Aquan-Assee, J. (2009). Boundary conditions for mean-reverting square root process. *University of Waterloo*.

Arrow, K. J. (1965). *Aspects of the theory of risk-bearing*. Yrjö Jahnssonin Säätiö.

Avellaneda, M. and Stoikov, S. (2008). High-frequency trading in a limit-order book. *Quantitative Finance*, 8(3):217–224.

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13(1):65–77.

Bacry, E. and Muzy, J.-F. (2014). Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, 14(7):1147–1166.

Bardi, M. and Capuzzo-Dolcetta, I. (2008). *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media.

Bayraktar, E. and Ludkovski, M. (2011). Optimal trade execution in illiquid markets. *Mathematical Finance*, 21(4):681–701.

Bayraktar, E. and Ludkovski, M. (2012). Liquidation in limit-order books with controlled intensity. *Mathematical Finance*.

Bellman, R. (1957). *Dynamical Programing*. Princton University Press.

Bellman, R. E. and Dreyfus, S. E. (1962). Applied dynamic programming. *Rand Corporation*.

Bender, C. and Orszag, S. (1999). *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*. Springer.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, pages 23–36.

Bertsekas, D. P., Bertsekas, D. P., Bertsekas, D. P., and Bertsekas, D. P. (1995). *Dynamic programming and optimal control*, volume 1. Athena Scientific Belmont, MA.

Bertsimas, D., Lo, A. W., and Hummel, P. (1999). Optimal control of execution costs for portfolios. *Computing in Science and Engineering*, 1(6):40–53.

Bhave, A. and Libertin, N. (2013). A study of short-term mean-reversion in equities. *361 Capital Investment Team*.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association*.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654.

Bouchard, B., Dang, N.-M., and Lehalle, C.-A. (2011). Optimal control of trading algorithms: a general impulse control approach. *SIAM Journal on Financial Mathematics*, 2(1):404–438.

Bouchaud, J.-P. (2010). Price impact. *Encyclopedia of quantitative finance*.

Bouchaud, J.-P., Farmer, J. D., and Lillo, F. (2009). How markets slowly digest changes in supply and demand. *Handbook of financial markets: dynamics and evolution*, 1:57.

Bremaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer.

Brennan, M. J. and Schwartz, E. S. (1978). Finite difference methods and jump processes arising in the pricing of contingent claims: A synthesis. *Journal of Financial and Quantitative Analysis*, 13(03):461–474.

Brown, R. (1828). A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine, or Annals of Chemistry, Mathematics, Astronomy, Natural History and General Science*, 4(21):161–173.

Brunetti, C., Kirilenko, A., and Mankad, S. (2011). Identifying high-frequency traders in electronic markets: Properties and forecasting. *Tech. rep., Johns Hopkins University, 2011.*

Buti, S., Rindi, B., and Werner, I. M. (2011). Dark pool trading strategies. *Charles A. Dice Center Working Paper*, (2010-6).

Cartea, Á., Donnelly, R. F., and Jaimungal, S. (2015a). Enhancing trading strategies with order book signals. *Available at SSRN 2668277*.

Cartea, Á., Gan, L., and Jaimungal, S. (2015b). Liquidating baskets of co-moving assets. *Available at SSRN*.

Cartea, Á. and Jaimungal, S. (2013a). Modelling asset prices for algorithmic and high-frequency trading. *Applied Mathematical Finance*, (ahead-of-print):1–36.

Cartea, Á. and Jaimungal, S. (2013b). Risk metrics and fine tuning of high-frequency trading strategies. *Mathematical Finance*.

Cartea, Á. and Jaimungal, S. (2015a). Incorporating order-flow into optimal execution. *Available at SSRN 2557457*.

Cartea, Á. and Jaimungal, S. (2015b). Optimal execution with limit and market orders. *Quantitative Finance*, (ahead-of-print):1–13.

Cartea, Á. and Jaimungal, S. (2015c). Order-flow and liquidity provision. *Available at SSRN 2553154*.

Cartea, Á., Jaimungal, S., and Penalva, J. (2015c). *Algorithmic and High-Frequency Trading*. Cambridge University Press.

Cheung, Y.-W. and Lai, K. S. (1994). Mean reversion in real exchange rates. *Economics Letters*, 46(3):251–256.

Chiang, A. (2000). Elements of dynamic optimization. *Illinois: Waveland Press Inc.*

Christie, A. A. (1987). On cross-sectional analysis in accounting research. *Journal of Accounting and Economics*, 9(3):231–258.

Christopherson, D. G. and Southwell, R. (1938). Relaxation methods applied to engineering problems. III. problems involving two independent variables. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 317–350.

Cont, R. and De Larrard, A. (2013). Price dynamics in a markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1):1–25.

Cont, R., Kukanov, A., and Stoikov, S. (2013). The price impact of order book events. *Journal of financial econometrics*, 12(1):47–88.

Cont, R., Stoikov, S., and Talreja, R. (2010). A stochastic model for order book dynamics. *Operations research*, 58(3):549–563.

Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes*. CRC press.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263.

Crank, J. (1984). *Free and moving boundary problems*. Clarendon press Oxford.

Crank, J. and Nicolson, P. (1947). A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 43, pages 50–67. Cambridge Univ Press.

Cryer, C. W. (1971). The solution of a quadratic programming problem using systematic overrelaxation. *SIAM Journal on Control*, 9(3):385–392.

Derman, E. and Kani, I. (1994). Riding on a smile. *Risk*, 7(2):32–39.

Dixit, A. and Pindyck, R. (1994). *Investment Under Uncertainty*. Princton University Press.

Duck, P. W., Evatt, G. W., and Johnson, P. V. (2014). Perpetual options on multiple underlyings. *Applied Mathematical Finance*, 21(2):174–200.

Easley, D., Lopez de Prado, M., and O'Hara, M. (2011). The microstructure of the flash crash: Flow toxicity, liquidity crashes and the probability of informed trading. *The Journal of Portfolio Management*, 37(2):118–128.

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778.

Euler, L. (1768). *Institutionum calculi integralis*, volume 1. imp. Acad. imp. Saènt.

Evans, J., Kuske, R., and Keller, J. B. (2002). American options on assets with dividends near expiry. *Mathematical Finance*, 12(3):219–237.

Evans, L. (2010). *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society.

Feller, W. (1951). Two singular diffusion problems. *The Annals of Mathematics*, 54(1):173–182.

Fernandez-Tapia, J. (2015). *Modeling, optimization and estimation for the on-line control of trading algorithms in limit-order markets*. PhD thesis, Université Pierre et Marie Curie (Paris VI).

Fleming, W. H. and Soner, H. M. (2006). *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media.

Föllmer, H. and Schied, A. (2011). *Stochastic finance: an introduction in discrete time*. Walter de Gruyter.

Forsyth, P. A. (2011). A Hamilton–Jacobi–Bellman approach to optimal trade execution. *Applied numerical mathematics*, 61(2):241–265.

Forsyth, P. A., Kennedy, J. S., Tse, S., and Windcliff, H. (2012). Optimal trade execution: a mean quadratic variation approach. *Journal of Economic Dynamics and Control*, 36(12):1971–1991.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin.

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2006). Institutional investors and stock market volatility. *The Quarterly Journal of Economics*, 121(2):461–504.

Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of business*, pages 67–78.

Gatheral, J. and Schied, A. (2011). Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework. *International Journal of Theoretical and Applied Finance*, 14(03):353–368.

Gloter, A. and Jacod, J. (2001a). Diffusions with measurement errors. I. local asymptotic normality. *ESAIM: Probability and Statistics*, 5:225–242.

Gloter, A. and Jacod, J. (2001b). Diffusions with measurement errors. II. optimal estimators. *ESAIM: Probability and Statistics*, 5:243–260.

Golub, G. H. and Van Loan, C. F. (2012). *Matrix computations*, volume 3. JHU Press.

Gopikrishnan, P., Plerou, V., Gabaix, X., and Stanley, H. E. (2000). Statistical properties of share volume traded in financial markets. *Physical Review E*, 62(4):R4493.

Grillet-Aubert, L. (2010). Négociation dactions: une revue de la littératurea lusage des régulateurs de marché. *AMF, available at: http://www. amffrance. org/documents/general/9530*, 1.

Guéant, O. (2014). Optimal execution of ASR contracts with fixed notional. *arXiv preprint arXiv:1410.1481*.

Guéant, O. and Lehalle, C.-A. (2013). General intensity shapes in optimal liquidation. *Mathematical Finance*.

Guéant, O., Lehalle, C.-A., and Fernandez-Tapia, J. (2012a). Dealing with the inventory risk: a solution to the market making problem. *Mathematics and Financial Economics*, pages 1–31.

Guéant, O., Lehalle, C.-A., and Fernandez-Tapia, J. (2012b). Optimal portfolio liquidation with limit orders. *SIAM Journal on Financial Mathematics*, 3(1):740–764.

Guéant, O. and Pu, J. (2013). Option pricing and hedging with execution costs and market impact. *arXiv preprint arXiv:1311.4342*.

Guéant, O., Pu, J., and Royer, G. (2015). Accelerated share repurchase: pricing and execution strategy. *International Journal of Theoretical and Applied Finance*, page 1550019.

Guilbaud, F., Mnif, M., and Pham, H. (2010). Numerical methods for an optimal order execution problem. *arXiv preprint arXiv:1006.0768*.

Guilbaud, F. and Pham, H. (2013a). Optimal high-frequency trading in a pro rata microstructure with predictive information. *Mathematical Finance*.

Guilbaud, F. and Pham, H. (2013b). Optimal high-frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79–94.

Harris, L. (1986). A transaction data study of weekly and intradaily patterns in stock returns. *Journal of financial economics*, 16(1):99–117.

Harris, L. (2002). *Trading and exchanges: Market microstructure for practitioners*. Oxford University Press.

Haugh, M. and Wang, C. (2014). Dynamic portfolio execution and information relaxations. *SIAM Journal on Financial Mathematics*, 5(1):316–359.

Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 438–443.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, 66(1):1–33.

Hestenes, M. R. (1950). A general problem in the calculus of variations with applications to paths of least time. *DTIC Document*.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343.

Ho, T. and Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial economics*, 9(1):47–73.

Huang, Y., Forsyth, P. A., and Labahn, G. (2012). Combined fixed point and policy iteration for Hamilton–Jacobi–Bellman equations in finance. *SIAM Journal on Numerical Analysis*, 50(4):1861–1882.

Huitema, R. (2013). Optimal portfolio execution using market and limit orders. *Available at SSRN 1977553*.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300.

Hull, J. and White, A. (1990). Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis*, 25(01):87–100.

Hult, H. and Kiessling, J. (2010). *Algorithmic trading with Markov chains*. PhD thesis, Stockholm University, Sweden.

Jaillet, P., Lamberton, D., and Lapeyre, B. (1990). Variational inequalities and the pricing of american options. *Acta Applicandae Mathematica*, 21(3):263–289.

Jain, P. C. and Joh, G.-H. (1988). The dependence between hourly prices and trading volume. *Journal of Financial and Quantitative Analysis*, 23(03):269–283.

Johnson, T. (2009). Paying the price. *https://plus.maths.org/content/paying-price*.

Kharroubi, I. and Pham, H. (2010). Optimal portfolio liquidation with execution cost and risk. *SIAM Journal on Financial Mathematics*, 1(1):897–931.

Kiefer, J. (1953). Sequential minimax search for a maximum. *Proceedings of the American Mathematical Society*, 4(3):502–506.

Kratz, P. and Schöneborn, T. (2013). Portfolio liquidation in dark pools in continuous time. *Mathematical Finance*.

Kühn, C. and Stroh, M. (2010). Optimal portfolios of a small investor in a limit order market: a shadow price approach. *Mathematics and Financial Economics*, 3(2):45–72.

Labadie, M., Lehalle, C.-A., et al. (2012). Optimal starting times, stopping times and risk measures for algorithmic trading. *No. hal-00705056. 2012.*

Laruelle, S., Lehalle, C.-A., and Pages, G. (2011). Optimal split of orders across liquidity pools: a stochastic algorithm approach. *SIAM Journal on Financial Mathematics*, 2(1):1042–1076.

Lehalle, C.-A. and Laruelle, S. (2013). *Market Microstructure in Practice*. World Scientific.

Lewis, M. (2014). *Flash boys: a Wall Street revolt*. WW Norton & Company.

Linetsky, V. (2004). Computing hitting time densities for CIR and OU diffusions: Applications to mean-reverting models. *Journal of Computational Finance*, 7:1–22.

LOBSTER (2015). Limit-order book system - the efficient reconstructor made by frischedaten UG (haftungsbeschränkt), Germany. *www.lobsterdata.com*.

Lorenz, J. and Almgren, R. (2011). Mean–variance optimal adaptive execution. *Applied Mathematical Finance*, 18(5):395–422.

Lund, D. (1993). The lognormal diffusion is hardly an equilibrium price process for exhaustible resources. *Journal of Environmental Economics and Management*, 25(3):235–241.

Markowitz, H. (1959). Portfolio selection: efficient diversification of investments. *New York*.

Maslov, S. and Mills, M. (2001). Price fluctuations from the order book perspective - empirical facts and a simple model. *Physica A: Statistical Mechanics and its Applications*, 299(1):234–246.

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712–740.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of economic theory*, 3(4):373–413.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of economics and management science*, pages 141–183.

Merton, R. C. (1990). *Continuous-time finance*. Blackwell Boston.

Merton, R. C., Brennan, M. J., and Schwartz, E. S. (1977). The valuation of american put options. *The Journal of Finance*, 32(2):449–462.

Metcalf, G. E. and Hassett, K. A. (1995). Investment under alternative return assumptions comparing random walks and mean reversion. *Journal of Economic Dynamics and Control*, 19(8):1471–1488.

Moran, M. (2012). VIX Views. http://www.spvixviews.com/2012/03/16/new-vvix-index-measures-the-volatility-of-volatility. Accessed: 2014-09-22.

NAG (2014). *The NAG Library, The Numerical Algorithms Group (NAG), Oxford, United Kingdom*. www.nag.com.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.

Obizhaeva, A. A. and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32.

Øksendal, B. (2003). *Stochastic differential equations*. Springer.

Øksendal, B. K. and Sulem, A. (2005). *Applied stochastic control of jump diffusions*, volume 498. Springer.

Osborne, M. F. (1959). Brownian motion in the stock market. *Operations research*, 7(2):145–173.

Peaceman, D. W. and Rachford, Jr, H. H. (1955). The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3(1):28–41.

Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media.

Pontryagin, L. S. (1987). *Mathematical theory of optimal processes*. CRC Press.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1):27–59.

Potters, M. and Bouchaud, J.-P. (2003). More statistical properties of order books and price impact. *Physica A: Statistical Mechanics and its Applications*, 324(1):133–140.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 122–136.

Predoiu, S., Shaikhet, G., and Shreve, S. (2011). Optimal execution in a general one-sided limit-order book. *SIAM Journal on Financial Mathematics*, 2(1):183–212.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (2009). *Numerical recipes in C++: the art of scientific computing*. Cambridge University Press.

Rijper, T., Sprenkeler, W., and Kip, S. (2011). High frequency trading: Position paper. *Unpublished*.

Rubinstein, M. (1994). Implied binomial trees. *The Journal of Finance*, 49(3):771–818.

Samuelson, P. A. (1965). Rational theory of warrant pricing. *Industrial management review*, 6:13–31.

Schied, A. and Schöneborn, T. (2009). Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastics*, 13(2):181–204.

Schied, A., Schöneborn, T., and Tehranchi, M. (2010). Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance*, 17(6):471–489.

Schied, A. and Slynko, A. (2011). Some mathematical aspects of market impact modeling. *Available at SSRN 1735465*.

Schöbel, R. and Zhu, J. (1999). Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *European Finance Review*, 3(1):23–46.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973.

SEC (2010). Concept realease on equity market structure. *Concept release No. 34-61358; File No. S7-02-10, SEC. 17 CFR Part 242.*

Shreve, S. E. (2004a). *Stochastic calculus for finance II: Continuous-time models*, volume 11. Springer Science & Business Media.

Shreve, S. E. (2004b). *Stochastic Calculus for Finance: The Binomial Asset Pricing Model. I.* Springer.

Skorodumov, B. (2008). Estimation of mean reversion in oil and gas markets. *MISUI & CO. Energy Risk Management Ltd., Technical Report, MITSUI/2008-10-14.*

Smith, G. D. (1965). *Numerical solution of partial differential equations.* Oxford university press London.

Thomas, R. (2008). Is maths to blame? *https://plus.maths.org/content/os/latestnews/sep-dec08/financecrisis/index.*

Tisseur, F. and Meerbergen, K. (2001). The quadratic eigenvalue problem. *SIAM review*, 43(2):235–286.

Turner, J. (2009). The Turner review: A regulatory response to the global banking crisis. *http://www.fsa.gov.uk/pubs/other/turner_review.pdf.*

Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical review*, 36(5):823.

Van Dyke, M. (1964). *Perturbation methods in fluid mechanics*, volume 964. Academic Press New York.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188.

Veraart, L. A. (2011). Optimal investment in the foreign exchange market with proportional transaction costs. *Quantitative Finance*, 11(4):631–640.

von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior*, volume 60. Princeton university press Princeton, NJ.

Weber, P. and Rosenow, B. (2005). Order book approach to price impact. *Quantitative Finance*, 5(4):357–364.

Whalley, A. E. and Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7(3):307–324.

Widdicks, M., Duck, P. W., Andricopoulos, A. D., and Newton, D. P. (2005). The Black-Scholes equation revisited: Asymptotic expansions and singular perturbations. *Mathematical Finance*, 15(2):373–391.

Wiener, N. (1923). Differential space. *Journal of Mathematical Physics*, 2:131–174.

Wilmott, P. (1995). *The mathematics of financial derivatives: a student introduction*. Cambridge University Press.

Wilmott, P. (1998). *Derivatives*. John Willey & Sons, New York.

Wilmott, P. (2007). *Paul Wilmott on Quantitative Finance, 3 Volume Set*. John Wiley & Sons.

Wood, R. A., McInish, T. H., and Ord, J. K. (1985). An investigation of transactions data for nyse stocks. *The Journal of Finance*, 40(3):723–739.

Yang, C. (2010). *Multi-asset and stochastic volatility option and bond pricing models: valuations and calibrations*. PhD thesis, School of Mathematics, University of Manchester, UK.

Young, D. (1971). *Iterative solution of large linear systems*. Elsevier.

Zhang, L. et al. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12(6):1019–1043.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales. *Journal of the American Statistical Association*, 100(472).

Zukerman, M. (2013). Introduction to queueing theory and stochastic teletraffic models. *arXiv preprint arXiv:1307.2968*.