Department of Mathematics

MPhil in Mathematics

January 2021 - August 2023

Optimal Execution for a Risk-Averse Trader

Adam Nii Armah Hesse
Email: adam.nad.hesse@outlook.com
Brunel email: adamniiarmah.hesse@brunel.ac.uk
Tel: +44 (0)7580 632 063

Originally submitted: Thursday 31st August 2023
Corrected: Sunday 6st October 2024

Brunel University London
Department of Mathematics
Uxbridge
Middlesex
UB8 3PH
United Kingdom
T: +44 1895 203397
F: +44 (0) 1895 251686

Abstract

We solve optimal stochastic control problems for a risk-averse trader who uses market orders and/or limit orders to liquidate a large position in a risky asset. In each case we aim to maximise terminal wealth, while managing the loss due to the price impact of our own trader's trading activity. We solve the problems using various utility functions for the trader's risk-aversion and penalty functions for the trader's urgency to liquidate the position and reduce market risk. We compare and contrast the performance of the strategies, and compare them to industry benchmarks such as TWAP and VWAP.

Acknowledgements

I would like to thank Dr Elena Boguslavskaya for her supervision on this project. It has been a joy to work with her. I have greatly appreciated her understanding and her guidance in my time at Brunel University.

I also thank Dr Paresh Date and Dr Dalia Chakrabarty for their helpful advice in my progress reviews and in other meetings.

I would like to thank the EPSRC for my funding in 2021 and 2022.

I would also like to thank my colleagues at the university, particularly Kam, with whom I have shared many fascinating and insightful conversations on stochastic analysis, mathematics, trading, and options. Mentions too to Sanna, Jugraj, Kane, Michael, Alexander, and Arianna.

Thanks to Xiaochuan and Lashi for the great conversations on mathematics and life in London. Lashi's philosophical conversations have been fantastic and I've greatly enjoyed them.

Thanks to the staff at Brunel Gym. The gym was a highlight of my time at Brunel and has helped me keep confident and strong both physically and mentally during this project.

I thank my mother and my two brothers, the constants in my life. My mother most of all for her belief and for always pushing me on to achieve. There is more to come.

Contents

Chapter 1

Algorithmic Trading

An algorithmic trading strategy is a set of rules that determine which securities to invest in, the size of the position, whether to go long or short, at which times to enter and execute the trade, and the time and the type of order used to execute the trade. Algorithmic trading is of great importance in the financial industry. It accounts for a very large proportion of the volume in stock markets, and is increasing its share of volume in fixed-income markets. Since the turn of the twenty-first century, the rise of electronic trading over more traditional pit trading has led to the proliferation of computer-based algorithmic trading. Regulation NMS in the United States, and the Market in Financial Instruments Directive (MiFID) II in the European Union have also played a role in the increasing dominance of algorithmic trading in the financial markets in both regions.

1.1 Types of algorithmic trading strategies

Algorithmic trading strategies can be divided into several different types. Optimal execution strategies try to minimise the price impact incurred when executing large buy or sell orders. Market-making involves a trader simultaneously posting buy and sell limit orders for a particular asset into the market to provide liquidity, matching buyers and sellers of the asset. Statistical arbitrage techniques try to capitalise on anomalies in the statistical characteristics of prices of one or several assets, for example the discrepancies in the co-integration of two assets in pairs trading.

1.2 A brief history of algorithmic trading

Algorithmic trading has developed over the past 50 to 60 years as the use and the availability of computers for performing financial calculations, building electronic exchanges, and executing trades has increased. Before the 1960s and 1970s, trading was performed manually. Hull [Hul12] and Guo et al [GLSW17] describe how "open-outcry" trading was performed on trading floors or in the trading pits of exchanges. Trade in a particular asset was announced. Traders would communicate orders verbally or using a complex system of hand gestures to indicate the amount of the asset they wanted to trade, whether they wanted to buy or sell, and at what price they wished to transact. Stock tickers, tiny printers which printed out abbreviated company names followed by the most recent transacted price, were used to transmit price updates to the public. Traded volume was added in the 1930s [GLSW17]. These trading floors were rendered obsolete by the development and eventual mass adoption of electronic trading since the 1970s. In banks and funds, the trading floor was replaced by systems of computers running algorithms to execute trades electronically. In many cases, these computers are monitored by human traders. Despite the almost universal replacement of physical trading floors with computers and servers, some open-outcry pits still remain, for example the London Metal Exchange.

1.2.1 Trading pre-1970s: The Manual Era

Trading decisions were made by various participants on the "buy-side" and facilitated by other participants on the "sell-side", as detailed in Irene Aldridge's book on "High Frequency Trading" [Ald13].

Buy-side Market Participants pre-1970s

- **Discretionary asset managers**: hedge funds, pension funds, mutual funds
- **Retail flow**: smaller investors and businesses.
- **Speculators**: individual traders betting on the direction of asset prices.

Sell-side Market Participants pre-1970s

- **Market-makers**: brokers at stockbrokers or dealing firms who provided quotes for buy-side traders to buy and sell, held short-term inventory risk, and profited from fees for providing these services.
- **Exchange**: Exchanges in each asset class, often one per asset class. Examples include the Chicago Board Options Exchange CBOE, the New York Stock Exchange, the London Stock Exchange, etc.
- **Stockbrokers**
- **Bucket Shops**: unregulated trading houses where individual traders would speculate on prices by placing bets. (described in great detail in Edwin Lefevre's "Reminiscences of a Stock Operator" [Lef23])

Aldridge [Ald13] describes how these market participants traded prior to the developments of the 1970s. Brokers would contact their buy-side clients with trade ideas and tip-offs, and those clients would in return trade through the broker. The clients would give a verbal order in person or over the telephone to the broker to buy or sell a particular quantity of a given security.

The broker would then execute the trade. If the order was for a large number of units of the security, the broker would either execute the order immediately, or would divide the client's order into a series of smaller orders and execute trades sequentially in the market. If the order was smaller than the minimum order size executable on an exchange (a "round lot"), the broker would accumulate orders from other clients into a single "round lot".

The broker would route the order to the exchange for the asset class of the security desired. Representatives of the exchanges, known as specialists, would match the order, selling to the broker if the order was a buy order, and buying from the broker if the order was a sell order.

Now, specialists at exchanges would often provide preferential terms for some of their clients. Aldridge describes how this practice "resulted in Wall Street cliques capable of significant price discrimination for in-group verses out-of-group customers" [Ald13] (page 4).

The broker would finally inform the client of the execution of the trade. Aldridge argues that brokers were the main profiters from this older system, because they were able to influence prices and receive compensation for favourable treatment of particular clients. High transaction costs were a feature of these "highly manual and labour-intensive" markets [Ald13] (page 3), which Aldridge argues led to relatively low turnover and a high degree of error associated with manual processing of orders. Guo et al [GLSW17] (page 3) corroborate this, stating that "out trades (orders unmatched due to human error in manually entering two supposedly matching trades to buy and sell a given security) were considered a cost of trading".

1.2.2 Trading since 1970s, the Electronic Trading era

In 1969, Instinet, an electronic communication network (ECN), started electronic trading among institutional clients. Guo et al [GLSW17] describe how Instinet "grew rapidly in the 1980s and became the dominant ECN by the ... late 1990s". NASDAQ was founded in 1971 as the first "electronic stock market". It provided automatic quotations as a computer bulletin board. Guo et al [GLSW17] state that most trades on NASDAQ took place via telephone until 1987. In the 1980s and 1990s many stock exchanges around the world began to offer electronic trading services: the London Stock Exchange in 1986, the Chicago Mercantile Exchange in 1992, the Toronto Stock Exchange in 1997, and the New York Stock Exchange in 2006.

As computing advances led to electronification of the financial markets, trading itself was also becoming automated. Quantitative analysis of asset prices, derivatives, volatility, and other market phenomena led to greater profits for institutions and were adopted widely on both the sell-side and the buy-side. Since the 1990s, improvements in computing technology led to the introduction of high frequency trading (HFT). This industry boomed in the latter half of the first decade of the 2000s. Alridge lists the new participants to the markets since the 1970s.

New Market Participants, post 1970s

- **Quantitative Funds**: mutual funds, hedge funds, proprietary trading firms using mathematical and statistical models of asset prices and volatility to make trading decisions and also using analysis of market microstructure to design execution strategies (the focus of this work).
- **Automated Market Makers**: market-makers using automated execution programs and mathematical and statistical analysis of market microstructure to provide competititve spreads for quoting limit orders to buy and sell securities and provide liquidity to the markets.
- **Automated arbitrageurs**: Statistical arbitrage funds looking to identify and profit from discrepancies in price movement, correlation, and co-integration.
- **Alternative trading venues**: Dark pools, new exchanges (such as BATS, Chi-X [Men13], IEX [Lew15])

Aldridge argues that the introduction of these participants led to a shift in the balance of power from brokers to their customers in funds and banks. Quantitative methods introduced at banks and funds gave them an edge over brokers in forecasting price movements. Brokers' expertise was reduced in scope from "the all-encompassing sell-side research into securities behaviour to ... algorithmic execution strategies" [Ald13] (page 4).

Electronic trading networks facilitated more accurate order recording and allowed the customer to control entry and recording of their own order directly to the exchange or broker almost instantaneously.

Following receipt of the order, the broker selects the appropriate optimal execution algorithm, designed to minimise the customer's exposure to unfavourable price direction during execution (price impact), and if requested via the order type used, to reduce the visibility of the size of the order to the market (say, via an "iceberg order" [CJP15]).

The algorithm slices the up the customer's full order into a smaller child orders and executes them in the market via various exchanges. "Smart order routing" [GLSW17] can be used to find the optimal exchanges to send each child order to at given intervals in time.

The trading venues match the respective child orders and provide notice of execution to the broker, which then accumulates and sends the confirmation of the full order back to the customer. This process typically occurs within microseconds or seconds.

Menkveld [Men13], Guo et al [GLSW17] and Aldridge [Ald13] all argue that the introduction of electronic trading and high frequency trading since the late 1990s has led to much faster trade execution, reduced transaction costs, and increased liquidity. Guo et al mention increased transparency of the market, since bid and ask prices are transmitted electronically via the Internet all over the world. A consequence of the increased competition between market-makers is reduced bid-ask spreads [Men13], [GLSW17].

The book “Flash Boys” by Martin Lewis [Lew15] presents the view, that HFT firms are predatory and target slower market participants (such as mutual funds, hedge funds) using various techniques to manipulate markets. An example is “front-running” to take advantage of order flow provided to them by banks and executing proprietary trades at better prices, forcing customer orders to be filled at less favourable prices. Aldridge does describe some of the more exploitative strategies used by certain HFT firms, such as “quote stuffing”, producing “fake” liquidity by flooding exchanges with limit orders which are quickly cancelled [Ald13] (page 201) and spoofing, again flooding exchanges with limit orders at various prices which are not cancelled, in order to produce “fake” liquidity and distort the limit-order book (page 202). Spoofing was made illegal in the USA under the Dodd-Frank Act [Ald13] (page 202). Despite acknowledging some of these exploitative tactics used by some HFT firms, Aldridge is overall positive that the benefits provided by high-frequency trading in particular outweigh the disadvantages.

1.3 Price

We can define and model the price of an asset in several ways. Given a list of limit orders for transacting in a given asset in the limit order book, the highest limit order to buy is known as the best bid, and the highest order to sell is known as the best ask. Let us say one is trading over a finite time horizon $[0, T]$. At a given time $t \in [0, T]$, taking the average of the best bid S_t^{bid} and the best ask S_t^{ask} yields us the **midprice** S_t^{mid} of the asset.

$$S_t^{\text{mid}} = \frac{S_t^{\text{ask}} + S_t^{\text{bid}}}{2}. \tag{1.1}$$

The asset midprice can be modelled mathematically in a variety of ways. Almgren and Chriss [AC00] divide the trading period $[0, T]$ into N intervals of length τ. Define the times $t_k = k\tau$ for each $k = 0, 1, ..., T$. Then they model the midprice as an arithmetic random walk in discrete time. Between time t_{k-1} and time t_k, the asset midprice takes the value:

$$S_k = S_{k-1} + \mu\tau + \sigma\sqrt{\tau}\epsilon, \tag{1.2}$$

where μ models the expected growth rate of the asset, σ models the volatility of the asset, and $\epsilon \sim \mathcal{N}(0, 1)$ is a standard Normal random variable. This model does not incorporate price impact as Almgren and Chriss do themselves. We will discuss price impact later.

In much of the financial literature, the price has been modelled in continuous time using a Wiener process $W = \{W_t\}_{0 \le t \le T}$, which has the properties described in the Appendix A.3.1.

A common model for asset prices in the short term is to model the price as Brownian motion with an associated scaling factor σ, which represents the volatility of the asset at each time. The volatility can be defined numerous ways, but one is to model it as the standard deviation of the asset price over a given time interval. Thus the midprice can be modelled as:

$$S_t = \sigma W_t, \quad 0 \leq t \leq T. \tag{1.3}$$

This model is used by many authors in dealing with optimal execution problems, for example Cartea et al [CJP15], and also market-making problems as in Gueant et al [GLFT13], Carmona and Webster [CW12], and Avellaneda and Stoikov [AS08]. Shreve [Shr08] shows that a Wiener Process can be seen to be a continuous-time analogue to the discrete-time random walk, obtained in the limit as the number of partitions N of the trading period $[0, T]$ tends to infinity.

1.4 Market Orders, Limit Orders, and the Limit Order Book (LOB)

Electronic exchanges offer a wide variety of order types for traders to buy or sell securities. There are two main order types offered by most electronic exchanges: market orders and limit orders.

A **limit order (LO)** is an order to buy or to sell a specified quantity of an asset at a price equal to or better than a specified limit price.

Limit orders posted to the exchange are entered into a **limit order book** (LOB). In her book “High Frequency Trading”, Irene Aldridge [Ald13] describes the LOB as a “table with columns corresponding to sequential price increments, and rows recording sizes of limit orders posted at each price increment” [Ald13]. The LOB records incoming limit orders and accumulates them into bins based on their price, measured to the “tick”, the smallest possible increment. Aldridge [Ald13] defines **liquidity** to be the “cumulative trade size of all limit orders available to meet incoming market orders at any given time on a specific trading venue”.

Orders resting in the limit order book are categorised by whether they are bids (limit orders to buy a particular security) or offers (limit orders to sell a security, also known as “asks”). The highest bid is known as the “**best bid**” and the lowest ask is known as the “**best ask**” or “**best offer**”. The difference between the best ask and the best bid is called the **bid-ask spread**.

When the bid-ask spread is 0, the best bid matches the best ask, and so resting limit orders execute against each other at that price until all the liquidity at that price is finished, i.e. all of the trades possible at the common best bid and best ask are executed. When the best ask is greater than the best bid, the limit orders rest in the limit order book awaiting a matching market order to arrive at the exchange and trigger against either the best bid or best ask.

A **market order (MO)** is an order to buy or to sell a specified quantity of an asset at the best available price. As Cartea et al explain [CJP15], this often results in an immediate execution. Trade does not always occur immediately, however. Aldridge [Ald13] explains that most exchanges use a price-time priority execution system, also known as the "first-in, first-out" schedule, for limit orders in the limit order book. In this system, at each price level, the limit order which arrived first is the first to execute against an incoming market order which matches the order.

1.5 Trade Execution and Price Impact

When a client gives an order to a broker or an execution trader to execute, the trader must decide how to execute that order to maximize the revenue of the sale for the client. We can equivalently model this as minimizing the execution costs for the client. The trader must also manage the risks of adverse price movements in the form of price impact, and also the risk that the order will not be filled. As described in Cartea *et al* [CJP15], the trader will carve the full "parent" order into smaller "child" orders, and try to execute these child orders over the trading horizon. Cartea *et al* describe how the trader must simultaneously address the issues of price impact and price risk.

Price impact occurs when a trader's activity causes the price of the asset being traded to move adversely against him due to the size of his order. Price impact can be divided into temporary and permanent price impact. As explained by Almgren and Chriss [AC00], temporary impact reflects changes in supply and demand caused by the trader's activity which change the price his trades are executed at. It is caused when a large market order "walks the limit order book" [CJP15]. When an market order "walks the book", it executes against the best limit orders that it can attain based on "price-time priority" [CJP15]. A market order first executes as much of the quantity at the best available price. If there is sufficient volume quoted at the best available price in the market, then the market order will execute in full against the oldest limit order posted at that price. If there is not sufficient volume available at the best price, the market order will execute as much as possible at the best price, and then attempt to execute the rest of the order at limit orders posted deeper in the limit order book, with the earliest limit orders given priority. This results in at least part of a large order executing against prices that are less than optimal for the trader. Permanent impact is a change in the actual market price of the security, the effect of which lasts throughout the trading period.

A large order to buy a security causes market prices to rise, thus increasing the trader's execution price so he pays more for the purchase. A large order to sell a security causes market prices to fall, decreasing the trader's execution price so he receives less for the sale.

1.5.1 TWAP and VWAP

Two commonly used execution algorithms are **Time-Weighted Average Price (TWAP)** and **Volume-Weighted Average Price (VWAP)**.

TWAP

A TWAP order is specified as follows: "Buy / Sell quantity Q of asset Y over time T with a limit price of S". TWAP aims to buy or sell a given quantity of a given asset over a time period by dividing the time period and the quantity into equal partitions, and executing a set fraction of the order (a "slice" [Ald13] or "child order" [CJP15]) at each interval, as long as the price at the time of execution of each child order remains better than the limit price specified in the original TWAP order.

Given an initial position of Q_0 units of an asset to liquidate over a time horizon $[0, T]$, at each time $t \in [0, T]$, a trader using TWAP will hold a quantity

$$Q_t^{\mathrm{TWAP}} = \left(1 - \frac{t}{T}\right) Q_0 \tag{1.4}$$

and will execute trades at the average trading speed

$$v_t^{\mathrm{TWAP}} = \frac{Q_0}{T}. \tag{1.5}$$

The trader will transact on average at the "TWAP price". Assuming the trader transacts each trade at the midprice, we can write TWAP as the arithmetic mean of the prices sampled at each time t:

$$\mathrm{TWAP} = \frac{1}{T} \sum_{t=1}^{T} S_t. \tag{1.6}$$

In our continuous-time models for each time $t \in [0, T]$, we use an integral:

$$\mathrm{TWAP} = \frac{1}{T} \int_0^T S_t dt. \tag{1.7}$$

As Brent Donnelly explains [Don19], the limit price referred to in the order request is optional. When no price is specified, the TWAP algorithm simply splits the full order into child orders according to the specified time interval and begins to execute them, executing the first child order immediately, at the best available price. When the TWAP order request specifies a limit price, each child order is executed only if the price at the time specified by the TWAP order is at the limit price or better. Thus it is possible for TWAP orders to execute only partially. Donnelly advises that TWAP orders are best utilised when the market is either not moving much, or when one expects the price to move favourably over the specified time horizon for the trade. For example, a trader looking to buy a given asset could use TWAP when he expects price to fall steadily over the trading period. That way, he would execute against increasingly favourable prices. But when he expects the price to rise over the time horizon, it would be better to use a market order and liquidate immediately rather than incur increasingly worse prices for child orders, not as a result of price impact but instead by timing the market incorrectly.

Irene Aldridge [Ald13] explains that the total number N of slices or child orders (i.e. the partition into which we divide the time horizon) is "best determined using characteristics specific to the traded security" (page 254), such as historical volume variation during the trading day, or market depth at the beginning of execution. More sophisticated algorithms which incorporate market depth have been developed by Obizhaeva and Wang [OW13] among other authors. Aldridge [Ald13] states that the objective when placing a TWAP order is to "select slices small enough so each child order does not significantly move the market", but that will be large enough or frequent enough so that the order is executed within the time horizon. In this case, the trader is trying to balance the two risks of market risk (retaining a position in the asset whole price becomes unfavourable) and price impact (incurring increasingly unfavourable prices due to size of order).

In order to better manage these two risks, more sophisticated algorithms have been designed. One of these is VWAP.

VWAP

The Volume-Weighted Average Price, VWAP, described both by Cartea et al [CJP15] and Aldridge [Ald13] as one of the most popular execution algorithms, aims to divide the parent order into child orders such that the size of a child order is larger when the trading volume at the time of execution is larger, and smaller when the trading volume at the time of execution is smaller. As Aldridge explains, trading during periods of higher volume is likely to provide a larger pool of matching orders and thus lead to faster and less costly execution. The point of VWAP is to take advantage of better liquidity when it is present, and reduce risk of bad execution in periods of reduced liquidity.

Volume V_t is measured over a given time interval $[T1, T2]$. At each time $t \in [T1, T2]$, the Volume-Weighted Average Price (VWAP) is given by

$$\text{VWAP} = \frac{\int_{T1}^{T2} S_t dV_t}{\int_{T1}^{T2} dV_t}, \tag{1.8}$$

where S_t is the midprice.

Cartea and Jaimungal [CJ16a] explain the rationale for an execution trader to aim to divide his full order into child orders whose size is not large relative to the liquidity available in the best quotes of the limit order book. An execution trader using market orders to execute his child orders must be aware that costs can be incurred from several features of high frequency markets. There is a delay or **latency** between the time an order is sent and the execution of that order. This leads to **slippage**, defined by Aldridge [Ald13] as "adverse changes in the market price" during this time. The actions of other market participants also has an effect on markets in the short term. They might cancel their limit orders or add more, and so the quantity of prices available at the best bid or best ask might change, or the best bid or best ask might change completely to a new quote if the liquidity at that level is depleted. Cartea and Jaimungal

[CJ16a] design optimal strategies which aim to optimise the trading speed such that each child order aims to take (1) a percentage of volume (PoV), or (2) a percentage of cumulative volume (PoCV) observed over the trading horizon up to that point.

1.5.2 Optimal Execution: The Literature

A mathematical approach to solving optimal execution problems was introduced in two seminal papers towards the end of the 1990s, first in discrete-time by Bertsimas and Lo [BL98], and then in continuous time by Almgren and Chriss [AC00]. In their paper of 2000, "Optimal Execution of Portfolio Transactions", Almgren and Chriss created a model to minimize volatility risk and transactions costs incurred from both permanent and temporary price impact for a trader managing a portfolio of risky assets [AC00]. Their model extends the discrete-time optimal execution model introduced earlier by Bertsimas and Lo. Rather than stochastic control to solve their optimization problem, the authors use a more direct approach, solving the constrained optimization problem using Lagrange multipliers.

Cartea, Jaimungal, and Penalva (2015) [CJP15] expand the Almgren-Chriss model to include several phenomena seen in real markets: in Chapter 7 they allow the stock midprice process to include order of market participants, and in Chapter 12 they use volume imbalance as an indicator of future order flow. The original Almgren-Chriss model allowed only for market orders, which in real trading incurs expenses because of the spread between the bid and ask prices, and also slippage due. to latency or the market order walking the limit order book. Cartea et al experiment first with limit orders only for execution that does not incur these costs, at the expense of having a fill probability less than 1. They then use a model that allows for both market orders and limit orders, and solve the resulting stochastic control problems numerically. In Chapter 9 they examine trade execution strategies which target volume.

In this text we will focus on optimal execution using market orders and limit orders. We model a risk-averse trader seeking to liquidate a number of shares in a finite time period. We model risk-aversion using utility functions. First we model a trader using market orders to liquidate the entire position, assuming his trading can be executed continuously. Afterwards, we allow the trader to use limit orders, given incoming market orders measured by a Poisson process.

Chapter 2

Methodology: Stochastic Control and Reinforcement Learning

2.1 Stochastic Control

2.1.1 The procedure

Given a stochastic control problem, we attempt to solve the problem using the following procedure:

1. Create a **stochastic model** for the controlled process.
2. Express the optimization problem as a **value function**.
3. Use the **Dynamic Programming Principle (DPP)** to simplify the optimization problem.
4. Use the DPP to get the **Hamilton-Jacobi-Bellman partial differential equation (HJB PDE)**.
5. **Solve the HJB PDE** to obtain optimal control parameters for the value function.

2.1.2 A stochastic control problem

Assume that we work in a finite time horizon $[0, T]$ on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(0 \leq t \leq T)}, \mathbb{P})$, with the filtration $(\mathcal{F}_t)_{(0 \leq t \leq T)}$ satisfying the "usual conditions" of being complete: each subfiltration $\mathcal{F}_t$ contains all of the "null" sets of probability zero of the full filtration $\mathcal{F}$, and being right-continuous, i.e. $\mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s = \mathcal{F}_t$.

Suppose we have a process, in our case, a trader's total wealth from a trading scheme, that we want to optimize using one or several control parameters. Due to the random fluctuations in asset prices in financial markets, we model the trader's wealth process as a system of stochastic differential equations (SDE), indexed over the amount of admissible controls. Using the notation of Cartea and Jaimungal [CJP15]:

$$dX_t^\alpha = \mu\left(t, X_t^\alpha, \alpha_t\right) dt + \sigma\left(t, X_t^\alpha, \alpha_t\right) dW_t, \quad X_0^\alpha = x, \tag{2.1}$$

where the stochastic process $W = (W_t^\alpha)_{0 \le t \le T}$ is a d-dimensional vector of independent Brownian motions on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(0 \le t \le T)}, \mathbb{P})$. The p-dimensional vector $\alpha = (\alpha_t)_{0 \le t \le T}$ represents the control processes, which are progressively measurable with respect to $(\mathcal{F})_{0 \le t \le T}$, and are valued in a subset $A \subset R^p$. The n-dimensional vector process $X = (X_t^\alpha)_{0 \le t \le T}$ represents the controlled variables. The drift $\mu : \mathbb{R}^n \times A \to \mathbb{R}$ and volatility $\sigma : \mathbb{R}^n \times A \to \mathbb{R}^{n \times d}$ can potentially depend on the time t, the value α_t of the particular control being used, and the value of the controlled process X_t at time t (hence there can be a feedback effect from the choice of control onto the controlled processes). They are Lipschitz continuous functions, i.e. there exists a number $K \ge 0$, for each $x_1, x_2 \in \mathbb{R}^n$, and for each control $a \in A$ with the property:

$$|\mu(x_1, a) - \mu(x_2, a)| + |\sigma(x_1, a) - \sigma(x_2, a)| \le K|x_1 - x_2|. \tag{2.2}$$

The admissible set $\mathcal{A}$ is a set of controls which satisfy certain properties making them realistic and regular. For example, admissible trading strategies must be square-integrable (so as not to be infinite and therefore impossible to implement in a world of finite resources and speed), and depend only on the information available up to a given time τ in the trading horizon $[t, T]$ (so a trader cannot look into the future in order to design a perfect strategy).

Using Huyên Pham's [Pha09] formulation for a finite-horizon problem, the admissible set $\mathcal{A}$ is the set of control processes $\alpha \in A$ such that:

$$\mathbb{E}\left[\int_0^T |\mu(0, \alpha_t)|^2 + |\sigma(0, \alpha_t)|^2 \, dt\right] < \infty. \tag{2.3}$$

That is, the control processes α are such that the drift and the diffusion coefficients of the controlled processes X in the underlying model are both square-integrable, and so is their sum, i.e. they satisfy a quadratic growth condition. This prevents all of the controlled processes from "blowing up" to infinity. These admissible controls ensure that the controlled processes have the properties which allow there to be a strong solution to the SDE 2.1 above, for all controls $\alpha \in \mathcal{A}$ and any initial condition $(t, x) \in [0, T] \times \mathbb{R}^n$, when the controlled process begins at a particular value $X_t = x$. Pham denotes this solution by $\{X_s^{t,x}, t \le s \le T\}$. The conditions of Lipschitz continuity (2.2) and quadratic growth (2.3) imply that the strong solution $X_s^{t,x}$ satisfies these properties:

$$\mathbb{E}\left[\sup_{t\le s\le T}|X_s^{t,x}|^2\right] < \infty \tag{2.4}$$

$$\lim_{h\downarrow 0^+}\mathbb{E}\left[\sup_{s\in[t,t+h]}|X_s^{t,x}-x|^2\right] = 0. \tag{2.5}$$

2.1.3 Value function and performance criteria

We now express the optimization problem as a value function. We find the optimal value of of this function over all admissible controls $u \in \mathcal{A}$.

We first introduce two measurable functions $f : [0,T]$ x $\mathbb{R}^n$ x $A \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. The first function f represents what Cartea et al [CJP15] call a running penalty/reward. This is gain or loss received by the agent at a given time, depending on the values attained by the (controlled) state variables given his or her use of a particular control. The second is the terminal reward attained at the end of the time horizon, depending on the final value of the (controlled) state variables.

Following Pham [Pha09], we suppose that g has a lower bound, or that g satisfies a quadratic growth condition: for each $x \in \mathbb{R}^n$, there exists some constant C, which is independent of x, for which:

$$|g(x)| \le C(1+|x|^2). \tag{2.6}$$

Now for $(t,x) \in [0,T]\times\mathbb{R}^n$, we denote by $\mathcal{A}_{(t,x)}$ the subset of controls in $\mathcal{A}$ such that the function f has finite expected value on the interval [t,T], given that the state variables have attained particular values at time t:

$$\mathbb{E}\left[\int_t^T f(s,X_s^\alpha,\alpha_s)ds \;\middle|\; X_t^\alpha = x\right] < \infty. \tag{2.7}$$

We define the **performance criteria** (Pham [Pha09] calls this the **gain function**), which expresses our problem mathematically as the expectation at time t of the sum of the running penalty or reward, and the terminal penalty or reward:

Definition 2.1.1 (Performance Criteria / Gain Function)**.** *For all $(t,x) \in [0,T]\times\mathbb{R}^n$ and all admissible controls $\alpha \in \mathcal{A}_{(t,x)}$, we define the **performance criteria** as:*

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^T f\left(s,X_s^\alpha,\alpha_s\right)ds + g(X_T^\alpha) \;\middle|\; X_t^\alpha = x\right]. \tag{2.8}$$

Our optimization problem can then be expressed mathematically by the **value function**. We maximize the total reward (or minimize the total penalty) over all admissible strategies $u \in \mathcal{A}$:

Definition 2.1.2 (Value Function)**.** *For all $(t,x) \in [0,T]\times\mathbb{R}^n$ and all admissible controls $\alpha \in \mathcal{A}_{(t,x)}$, we define the **value function** as:*

$$V(t,x) = \sup_{\alpha\in\mathcal{A}_{(t,x)}} \mathbb{E}\left[\int_t^T f\left(s, X_s^\alpha, \alpha_s\right) ds + g(X_T^\alpha) \;\middle|\; X_t^\alpha = x\right]. \tag{2.9}$$

Thus the value function is the supremum of the performance criteria.

$$V(t,x) = \sup_{\alpha\in\mathcal{A}_{(t,x)}} J(t,x,\alpha). \tag{2.10}$$

Definition 2.1.3 (Optimal Control)**.** *Given an initial condition $(t,x) \in [0,T]\times\mathbb{R}^n$, we say that an admissible control $\alpha^* \in \mathcal{A}_{(t,x)}$ is an **optimal control** if*

$$V(t,x) = J(t,x,\alpha^*). \tag{2.11}$$

2.1.4 The Dynamic Programming Principle (DPP)

It is extremely computationally expensive to solve the optimization problem over all possible scenarios for our controlled process X_t over each time instant in the time interval $[t,T]$. We address this problem by first flowing the process to a stopping time τ between t and T, and then optimize over the rest of the time horizon $[\tau,T]$.

Theorem 2.1.1 (The Dynamic Programming Principle (DPP) for Diffusion Processes)**.** *Given the stochastic control problem modelled by (2.1), the value function (2.9) and performance criteria (2.8) satisfy the Dynamic Programming Principle: For all $(t,x) \in [0,T]\times\mathbb{R}^n$, and all stopping times $\tau \le T$,*

$$V(t,x) = \sup_{\alpha\in\mathcal{A}_{(t,x)}} \mathbb{E}\left[\int_t^\tau f\left(s, X_s^\alpha, \alpha_s\right) ds + V(\tau, X_\tau^\alpha) \;\middle|\; X_t^\alpha = x\right]. \tag{2.12}$$

Proof. We choose an arbitrary admissible control $\alpha \in \mathcal{A}$ and start with the performance criteria:

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^T f\left(s, X_s^\alpha, \alpha_s\right) ds + g(X_T^\alpha) \;\middle|\; X_t^\alpha = x\right]. \tag{2.13}$$

We introduce an arbitrary stopping time $\tau \in [t,T]$ and split the integral of the running penalty/reward function.

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^\tau f\left(s, X_s^\alpha, \alpha_s\right) ds + \int_\tau^T f\left(s, X_s^\alpha, \alpha_s\right) ds + g(X_T^\alpha) \;\middle|\; X_t^\alpha = x\right]. \tag{2.14}$$

Introduce the notation that $\mathbb{E}_{t,x}\left[\bullet\right] = \mathbb{E}\left[\bullet \mid X_t^\alpha = x\right]$.

By iterated expectations, and the definition of the performance criteria (2.8)

$$\begin{aligned}
&\mathbb{E}_{t,x}\left[\int_{\tau}^{T} f\left(s, X_s^{\alpha}, \alpha_s\right) ds + g(X_T^{\alpha})\right] \\
&= \mathbb{E}_{t,x}\left[\mathbb{E}_{\tau, X_{\tau}^{\alpha}}\left[\int_{\tau}^{T} f\left(s, X_s^{\alpha}, \alpha_s\right) ds + g(X_T^{\alpha})\right]\right] \\
&= \mathbb{E}_{t,x}\left[J(\tau, X_{\tau}^{\alpha}, \alpha)\right]. \qquad (2.15)
\end{aligned}$$

Therefore we can define the performance criteria at any time $t \in [0, T]$ as the expectation, conditional on the process X_t attaining the value x, of the sum of the performance criteria at a stopping time $\tau \in [t, T]$ and the running penalty or reward attained between our starting point t and that stopping time τ:

$$J(t, x, \alpha) = \mathbb{E}_{t,x}\left[J(\tau, X_{\tau}^{\alpha}, \alpha) + \int_{t}^{\tau} f\left(s, X_s^{\alpha}, \alpha_s\right) ds\right]. \qquad (2.16)$$

By definition of the value function (2.9) at time t as the supremum taken over all admissible controls of the performance criteria (2.8), i.e. since $V(t, x) = \sup_{\alpha \in \mathcal{A}} J(t, x, \alpha)$:

$$\begin{aligned}
J(t, x, \alpha) &= \mathbb{E}_{t,x}\left[J(\tau, X_{\tau}^{\alpha}, \alpha) + \int_{t}^{\tau} f\left(s, X_s^{\alpha}, \alpha_s\right) ds\right] \\
&\le \mathbb{E}_{t,x}\left[V(\tau, X_{\tau}^{\alpha}) + \int_{t}^{\tau} f\left(s, X_s^{\alpha}, \alpha_s\right) ds\right]. \qquad (2.17)
\end{aligned}$$

Taking supremum over all admissible controls on both sides gives us an upper bound for the value function, and one part of the Dynamic Programming Principle.

$$\begin{aligned}
\sup_{\alpha \in \mathcal{A}_{(t,x)}} J(t, x, \alpha) &\le \sup_{\alpha \in \mathcal{A}_{(t,x)}} \mathbb{E}_{t,x}\left[V(\tau, X_{\tau}^{\alpha}) + \int_{t}^{\tau} f\left(s, X_s^{\alpha}, \alpha_s\right) ds\right] \\
\implies V(t, x) &\le \sup_{\alpha \in \mathcal{A}_{(t,x)}} \mathbb{E}_{t,x}\left[V(\tau, X_{\tau}^{\alpha}) + \int_{t}^{\tau} f\left(s, X_s^{\alpha}, \alpha_s\right) ds\right]. \qquad (2.18)
\end{aligned}$$

We now derive the lower bound for the value function, and show that it equals the upper bound. We choose an arbitrary admissible control $\alpha \in \mathcal{A}_(t, x)$ and a stopping time $\tau \in [t, T]$.

There exists for each $\epsilon > 0$ and each outcome $\omega \in \Omega$ a control $\alpha^{\epsilon,\omega} \in \mathcal{A}_{(\tau(\omega), X_{\tau(\omega)}(\omega))}$ which is an ϵ-optimal control, that is

$$V(\tau(\omega), X_{\tau(\omega)}) - \epsilon \le J(\tau(\omega), X_{\tau(\omega)}, \alpha^{\epsilon,\omega}) \le V(\tau(\omega), X_{\tau(\omega)}). \qquad (2.19)$$

We define a modification of the ϵ-optimal control $\tilde{\alpha}_s(\omega)$ as

$$\tilde{\alpha}_s(\omega) = \begin{cases} \alpha_s(\omega), \; s \in [0, \tau(\omega)] \\ \alpha_s^{\epsilon,\omega}(\omega), \; s \in [\tau(\omega), T]. \end{cases}$$

Pham states that this control $\tilde{\alpha}_s(\omega)$ can be shown to be progressively measurable, and therefore $\tilde{\alpha}_s(\omega) \in \mathcal{A}_{(t,x)}$. Cartea et al formulate the same process as

$$\tilde{\alpha}_s(\omega) = \alpha_s(\omega)\mathbb{I}_{t\leq\tau} + \alpha_s^{\epsilon,\omega}(\omega)\mathbb{I}_{t>\tau}. \tag{2.20}$$

This modification $\tilde{\alpha}_s(\omega)$ is ϵ-optimal after the stopping time τ, but could be sub-optimal before time τ. Comparing the value function to the performance criteria achieved using this control, we have:

$$\begin{aligned}
V(t,x) &\geq J(t,x,\tilde{\alpha}_s(\omega)) \\
&= \mathbb{E}_{t,x}\left[J\left(\tau, X_\tau^{\tilde{\alpha}_s(\omega)}, \tilde{\alpha}_s(\omega)\right) + \int_t^\tau f\left(s, X_s^{\tilde{\alpha}_s(\omega))}, \tilde{\alpha}_s(\omega))\right) ds\right] \\
&= \mathbb{E}_{t,x}\left[J\left(\tau, X_\tau^{\alpha}, \alpha_s^{\epsilon}\right) + \int_t^\tau f\left(s, X_s^{\alpha}, \alpha_s\right)\right) ds\right] \\
&\geq \mathbb{E}_{t,x}\left[V\left(\tau, X_\tau^{\alpha}\right) + \int_t^\tau f\left(s, X_s^{\alpha}, \alpha_s\right)\right) ds\right] - \epsilon,
\end{aligned} \tag{2.21}$$

where the third line follows from our definition of the modification (2.20) and the fourth from the definition of the ϵ-optimal control (2.19) and the linearity of expectation.

Taking limits as $\epsilon \downarrow 0$ eliminates the ϵ:

$$V(t,x) \geq \mathbb{E}_{t,x}\left[V\left(\tau, X_\tau^{\alpha}\right) + \int_t^\tau f\left(s, X_s^{\alpha}, \alpha_s\right)\right) ds\right]. \tag{2.22}$$

Since the inequality holds for an arbitrary admissible control $\alpha \in \mathcal{A}_(t,x)$, it also holds for the admissible control that achieves the supremum:

$$V(t,x) \geq \sup_{\alpha\in\mathcal{A}_{(t,x)}} \mathbb{E}_{t,x}\left[V\left(\tau, X_\tau^{\alpha}\right) + \int_t^\tau f\left(s, X_s^{\alpha}, \alpha_s\right)\right) ds\right]. \tag{2.23}$$

Thus we have the lower bound for the value function, and it is equal to the upper bound. Therefore for all $(t,x) \in [0,T] \times \mathbb{R}^n$, and all stopping times $\tau \leq T$,

$$V(t,x) = \sup_{\alpha\in\mathcal{A}_{(t,x)}} \mathbb{E}_{t,x}\left[V\left(\tau, X_\tau^{\alpha}\right) + \int_t^\tau f\left(s, X_s^{\alpha}, \alpha_s\right)\right) ds\right]. \tag{2.24}$$

$\square$

The Dynamic Programming Principle states that the value function calculated at time $t \in [0,T]$ equals the expected value, conditional on X_t attaining value x, of the sum of the value function evaluated at a future stopping time $\tau \in [t,T]$ plus the running penalty/reward obtained between time t and τ, provided we are using the control $\alpha \in \mathcal{A}$ which maximises that expected value over the time interval $[t,T]$.

2.1.5 The Hamilton-Jacobi-Bellman Partial Differential Equation

The Hamilton-Jacobi-Bellman (HJB) equation is an infinitesimal continuous-time version of the Dynamic Programming Principle. By differentiating the HJB equation with respect to each of the control variables, we can find the optimal controls in "feedback form", that is, in the form of partial derivatives of the value function. Feeding these optimal control representations back into the HJB equation leads to a partial differential equation (PDE) for the value function.

This partial differential equation is usually non-linear and deterministic: it removes the randomness from the stochastic control problem, making it much more tractable. Some examples can be solved explicitly, for example, Almgren and Chriss' optimal liquidation model. Others must be solved numerically using techniques such as finite difference methods.

Let us define some items which appear in the HJB equation. Here we follow similar notation to that of Cartea *et al* [CJP15]

Lemma 2.1.2 (Itô's Lemma)**.** *Let $W = \{W_t\}_{t\geq 0}$ denote an n-dimensional column vector of independent Brownian motions. Let $X = \{X_t\}_{t\geq 0}$ be an Itô process, i.e. let X be an m-dimensional column vector of stochastic processes which satisfies the SDE:*

$$dX_t = \mu\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dW_t, \quad X_0 = x, \tag{2.25}$$

where $\mu(t, X_t)$ is an m-dimensional vector of drifts and $\sigma(t, X_t)$ is an $(m \times n)$-dimensional matrix of diffusion coefficients (or volatilities, in the financial context).

We introduce a new stochastic process $Y = f(t, X_t)$ where $f(t, x) \in \mathcal{C}^{(1,2)}$, i,e. $f(t, x)$ is differentiable with respect to t and twice-differentiable with respect to each x_i, with continuous derivatives.

Then Y is another Itô process which satisfies the SDE:

$$\begin{aligned} dY_t &= df(t, X_t) \\ &= \left(\partial_t f(t, X_t) + \mu\left(t, X_t\right)' \mathcal{D} f(t, X_t) + \frac{1}{2}\mathbf{Tr}\left(\sigma(t, X_t)\sigma(t, X_t)' \mathcal{D}^2 f(t, X_t)\right)\right) dt \\ &\quad + \mathcal{D} f(t, X_t)' \sigma(t, X_t) dW_t, \end{aligned} \tag{2.26}$$

where for any vector or matrix A, A' denotes its transpose, $\mathcal{D} f(t, X_t)$ denotes the n-dimensional column vector of first partial derivatives of $f(t, X_t)$, and $\mathcal{D}^2 f(t, X_t)$ denotes the $(m \times n)$-dimensional matrix of second partial derivatives of $f(t, X_t)$.

Definition 2.1.4 (Infinitesimal Generator)**.** *The infinitesimal generator, denoted $\mathcal{L}_t$, of a process X_t is a differential operator that acts on twice-differentiable functions $\phi(x) \in \mathcal{C}^2$ as follows:*

$$\mathcal{L}_t = \lim_{h \searrow 0} \frac{\mathbb{E}\left[\phi(X_{t+h}) \mid X_t = x\right] - \phi(x)}{h}. \tag{2.27}$$

Via Itô's lemma, we can express the infinitesimal generator of a twice differentiable function $\phi(x)$ of the Itô process (2.1) in the following form:

$$\mathcal{L}_t\phi(x) = \mu(t,x)^{'}\ \mathcal{D}\phi(x) + \frac{1}{2}\mathbf{Tr}\left(\sigma(t,x)\sigma(t,x)^{'}\ \mathcal{D}^2\phi(x)\right), \tag{2.28}$$

where for any vector or matrix A, $A^{'}$ denotes its transpose, $\mathcal{D}\phi(x)$ denotes the n-dimensional column vector of first partial derivatives of $\phi(x)$, and $\mathcal{D}^2\phi(x)$ denotes the $(m \times n)$-dimensional matrix of second partial derivatives of $\phi(x)$.

Note that with the definition (2.28), we can write Itô's formula for a function $Y = f(t, X_t)$ of an Itô process X_t as:

$$dY_t = (\partial_t + \mathcal{L}_t)\, f(t,X_t)\ dt + \mathcal{D}f(t,X_t)^{'}\sigma(t,X_t)\ dW_t, \tag{2.29}$$

or in integral form: Given an Itô process $\{X_t\}_{t\geq 0}$, and a new stochastic process $Y = f(t, X_t)$ where $f(t,x) \in \mathcal{C}^{(1,2)}$, then for all $t \in [0,T]$ and all $T \geq 0$,

$$Y_t = Y_0 + \int_0^t (\partial_s + \mathcal{L}_s)\, f(s,X_s)\ ds + \int_0^t \mathcal{D}f(s,X_s)^{'}\sigma(s,X_s)\ dW_s. \tag{2.30}$$

Theorem 2.1.3 (The Hamilton-Jacobi-Bellman Partial Differential Equation)**.** *If the value function (2.9) satisfies the Dynamic Programming Principle (2.1.1), then it satisfies the HJB equation: For each $(t,x) \in [0,T]\times\mathbb{R}^n$, and each control $\alpha \in \mathcal{A}$,*

$$\partial_t V(t,x) + \sup_{\alpha\in\mathcal{A}} \left(\mathcal{L}_t^\alpha V(t,x) + f(t,x,\alpha)\right) = 0, \tag{2.31}$$

$$V(T,x) = g(x). \tag{2.32}$$

Proof. We follow Cartea et al's exposition [CJP15] and we choose the stopping time $\tau \leq T$ in the DPP (2.1.1) such that it is the minimum between the time it takes the process X_t^α to exit a ball of size ϵ around its starting point x at time t, and a fixed time h.

$$\tau = \min\left\{T, \inf\left\{s > t\ :\ \Big(\ (s-t)\ ,\ |\ X_s^\alpha - x\ |\ \Big) \notin [0,h) \text{ x } [0,\epsilon)\ \right\}\ \right\}. \tag{2.33}$$

So our chosen stopping time τ is either the end of our time horizon T, or τ is the first time s after time t that either the state variables process X_s^α differs by ϵ from its value $X_t^\alpha = x$ at t, or after which the time interval $s - t$ is greater than h. Now

$$\lim_{h\to 0} \tau = t,\ \mathbb{P}\text{-a.s.} \tag{2.34}$$

As h shrinks towards 0, it becomes less likely that the process X will exit the ball $(x - \epsilon\ ,\ x + \epsilon)$ before time $(t + h)$. So $\tau = t + h$ when h is small enough.

We now use Itô's lemma (2.30) to write the value function for an arbitrary control $\alpha \in \mathcal{A}$ evaluated at the stopping time τ in terms of the value function at the starting point t:

$$\begin{aligned}&V\left(\tau, X_\tau^\alpha\right)\\&=V(t,x)+\int_t^\tau \left(\partial_s+\mathcal{L}_s^\alpha\right)V\left(s,X_s^\alpha\right)ds+\int_t^\tau \mathcal{D}_x V\left(s,X_s^\alpha\right)^{'}\sigma\left(s,X_s^\alpha,\alpha_s\right)dW_s.\end{aligned}\tag{2.35}$$

We choose an admissible control that is constant over the period $[t,\tau]$, i.e. for each $s\in[t,\tau]$, $\alpha_s=a$ for some $a\in\mathcal{A}$.

We apply the lower bound (2.23) to the value function when we operate under this control.

$$\begin{aligned}V(t,x)&\geq \sup_{\alpha\in\mathcal{A}_{(t,x)}}\mathbb{E}_{t,x}\left[V(\tau,X_\tau^\alpha)+\int_t^\tau f\left(s,X_s^\alpha,\alpha_s\right)ds\right]\\&\geq \mathbb{E}_{t,x}\left[V(\tau,X_\tau^a)+\int_t^\tau f\left(s,X_s^a,a\right)ds\right].\end{aligned}\tag{2.36}$$

Now we apply Itô's formula to $V(\tau,X_\tau^a)$ using (2.35), so our inequality becomes:

$$\begin{aligned}V(t,x)\geq \mathbb{E}_{t,x}\Bigg[&V(t,x)\\&+\int_t^\tau\left(\partial_s+\mathcal{L}_s^a\right)V\left(s,X_s^a\right)ds\\&+\int_t^\tau \mathcal{D}_xV\left(s,X_s^\alpha\right)^{'}\sigma\left(s,X_s^a,a\right)dW_s\\&+\int_t^\tau f\left(s,X_s^a,a\right)ds\Bigg].\end{aligned}\tag{2.37}$$

The integrand $\mathcal{D}_xV\left(s,X_s^\alpha\right)^{'}\sigma\left(s,X_s^a,a\right)$ in the stochastic integral is bounded because by the definition of the stopping time τ in (2.33), at each time $s\in[t,\tau]$, $|X_s^a-x|\leq\epsilon$. Thus, the stochastic integral is the stochastic integral of a martingale, and we can write its expectation as zero.

Also since $V(t,x)\in\mathcal{F}_t$, we can remove it from the conditional expectation on the RHS. Then

$$\begin{aligned}V(t,x)&\geq V(t,x)+\mathbb{E}_{t,x}\left[\int_t^\tau\left(\partial_s+\mathcal{L}_s^a\right)V\left(s,X_s^a\right)+f\left(s,X_s^a,a\right)ds\right]\\0&\geq \mathbb{E}_{t,x}\left[\int_t^\tau\left(\partial_s+\mathcal{L}_s^a\right)V\left(s,X_s^a\right)+f\left(s,X_s^a,a\right)ds\right].\end{aligned}\tag{2.38}$$

We recall that when h is sufficiently small, $\tau=(t+h)$, $\mathbb{P}-$a.s, because the process X will not exit the ball $(x-\epsilon\ ,\ x+\epsilon)$. Then

$$0\geq \mathbb{E}_{t,x}\left[\int_t^{t+h}\left(\partial_s+\mathcal{L}_s^a\right)V\left(s,X_s^a\right)+f\left(s,X_s^a,a\right)ds\right].\tag{2.39}$$

We now divide by h (keeping the factor $\frac{1}{h}$ within the expectation, and then take limits as $h \to 0$:

$$0 \geq \lim_{h\to 0} \mathbb{E}_{t,x}\left[\frac{1}{h}\int_t^{t+h} \left(\partial_s + \mathcal{L}_s^a\right) V\left(s, X_s^a\right) + f\left(s, X_s^a, a\right) ds\right]. \tag{2.40}$$

Now, we know that $X_t^a = x$. The condition $|X_s^a - x| \leq \epsilon$ implies that at the stopping time τ, $X_\tau^a \leq \epsilon + x$, and so is bounded.

Thus we can use the mean value theorem to write:

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} \omega_s = \omega_t. \tag{2.41}$$

Then our inequality becomes

$$0 \geq \left(\partial_t + \mathcal{L}_t^a\right) V(t,x) + f(t,x,a). \tag{2.42}$$

Since the above inequality holds for any control $a \in \mathcal{A}$, it holds true for the supremum over all admissible controls. So we have the inequality

$$0 \geq \partial_t V(t,x) + \sup_{\alpha\in\mathcal{A}} \left(\mathcal{L}_t^\alpha V(t,x) + f(t,x,\alpha)\right). \tag{2.43}$$

To show that we have an equality, we start with an optimal control $\alpha^* \in \mathcal{A}$, which gives us the following equality from the DPP:

$$V(t,x) = \mathbb{E}_{t,x}\left[V\left(\tau, X_\tau^{\alpha^*}\right) + \int_t^\tau f\left(s, X_s^{\alpha^*}, \alpha^*\right) ds\right]. \tag{2.44}$$

Following the same process as we have done for the control a, we obtain the equality

$$0 = \partial_t V(t,x) + \mathcal{L}_t^{\alpha^*} V(t,x) + f(t,x,\alpha^*). \tag{2.45}$$

Putting the inequality (2.43) and the equality (2.45) together gives us the HJB equation:

$$0 = \partial_t V(t,x) + \sup_{\alpha\in\mathcal{A}} \left(\mathcal{L}_t^\alpha V(t,x) + f(t,x,\alpha)\right). \tag{2.46}$$

The terminal condition

$$V(T,x) = g(X) \tag{2.47}$$

follows from the definition of the value function (2.9) applied at time $t = T$. The terminal penalty/reward function $g(T) \in \mathcal{F}_T$ and so comes out of the expectation, while the integral of the running penalty/reward function $f(t,x,a)$ vanishes to zero. □

2.1.6 Verification

We want to find a candidate value function and a candidate optimal control that solves the stochastic control problem (3.160). To verify that our candidate functions solve the problem, we will follow the method suggested in the proof in Pham's book ([Pha09]).

Theorem 2.1.4 (Verification Theorem)**.** *Let $\Phi(t,x) \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^n)$ and let $\Phi(t,x)$ satisfy a quadratic growth condition, that is for all for all $(t,x) \in [0,T] \times \mathbb{R}^n$, there exists a constant C such that:*

$$|\Phi(t,x)| \leq C\left(1+|x|^2\right). \tag{2.48}$$

Suppose also that $F : \mathbb{R}_+ \to \mathbb{R}^{n+p}$ and $G : \mathbb{R}^n \to \mathbb{R}$ are measurable uniformly bounded functions (where n is the number of controlled processes and p is the number of control processes in our model (2.1)), which satisfy:

$$\frac{\partial}{\partial t}\Phi(t,x) - \sup_{\alpha\in\mathcal{A}}\left(\mathcal{L}^{\alpha}\Phi(t,x) + F(t,x,\alpha)\right) \geq 0, \tag{2.49}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$, and

$$\Phi(T,x) \geq G(x) \text{ for all } x \in \mathbb{R}^n, \tag{2.50}$$

then

$$\Phi(t,x) \geq J(t,x,\alpha) \text{ on } [0,T] \times \mathbb{R}^n. \tag{2.51}$$

Now suppose that $\Phi(T,x) = G(x)$, and that there exists a measurable function $\alpha^*(t,x)$ for $(t,x) \in [0,T] \times \mathbb{R}^n$, with values in the admissible set $\mathcal{A}$, which satisfies:

$$\frac{\partial}{\partial t}\Phi(t,x) - \sup_{\alpha\in\mathcal{A}}\left(\mathcal{L}^{\alpha}\Phi(t,x) + F(t,x,\alpha)\right) = 0, \tag{2.52}$$

that is,

$$\frac{\partial}{\partial t}\Phi(t,x) - \left(\mathcal{L}^{\alpha^*(t,x)}\Phi(t,x) + F(t,x,\alpha^*(t,x))\right) = 0. \tag{2.53}$$

If the stochastic differential equation

$$dX_t^{\alpha^*} = \mu\left(t, X_t^{\alpha^*}, \alpha_t^*\right)dt + \sigma\left(t, X_t^{\alpha^*}, \alpha_t^*\right)dW_t, \quad X_0^{\alpha} = x, \tag{2.54}$$

admits a unique solution, denoted by $X_t^{\alpha^*}$, and if the process

$$\alpha^*\left(s, X_s^{\alpha^*}\right)_{t\leq s\leq T} \in \mathcal{A}, \tag{2.55}$$

then

$$\Phi(t,x) = V(t,x) \text{ on } [0,T] \times \mathbb{R}^n. \tag{2.56}$$

and $\alpha^*(t,x)$ is an optimal Markovian control.

Proof. We follow the proof laid out in Section 3.5 of Pham's book [Pha09]. Since $\Phi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^n)$, we can use Itô's lemma (2.29) to write for all $(t,x) \in [0,T] \times \mathbb{R}^n$ and all controls which are admissible over $[t,T]$ given that $X_t^{\alpha} = x$, ie. all $\alpha \in \mathcal{A}_{(t,x)}$, any time $s \in [t,T)$, and any stopping time $\tau \in [t,T]$,

$$\begin{aligned}\Phi\left(s\wedge\tau, X_{s\wedge\tau}^{\alpha}\right) = \Phi(t,x) &+ \int_t^{s\wedge\tau} \frac{\partial\Phi}{\partial u}\left(u, X_u^{\alpha}\right) + \mathcal{L}^{\alpha_u}\Phi\left(u, X_u^{\alpha}\right)du \\ &+ \int_t^{s\wedge\tau} \mathcal{D}_x\Phi\left(u, X_u^{\alpha}\right)'\sigma\left(u, X_u^{\alpha}, \alpha_u\right)dW_u. \end{aligned} \tag{2.57}$$

Given that the above is true for all stopping times $\tau \in [t, T]$, we choose the particular stopping time τ to be the first time after our initial time t that the integral of the square of the integrand of the stochastic integral component of Φ is larger than some integer n:

$$\tau = \tau_n := \inf \left\{ s \geq t : \int_t^s \left| \mathcal{D}_x \Phi \left(u, X_u^\alpha\right)' \sigma \left(u, X_u^\alpha, \alpha_u\right) \right|^2 du \geq n \right\}. \tag{2.58}$$

From this definition, as the integer $n \to \infty$, the stopping time $\tau_n \to \infty$. We can select the stopping time τ_n in this way because we have assumed that for all $(t, x) \in [0, T] \times \mathbb{R}^n$, the candidate value function $\Phi(t, x) \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$, so Φ has a continuous second derivative, and because by the definition of the admissible set $\mathcal{A}$, the volatility matrix σ of the state variables process satisfies quadratic growth (2.3), and hence $\sigma \left(u, X_u^\alpha, \alpha_u\right)$ is finite and integrable.

Then we can define each of our chosen stopping times τ_n as a member of a localizing sequence $\{\tau_n\}_{n \in \mathbb{N}}$ of stopping times such that the stopped stochastic integral process

$$\left\{ \int_t^{s \wedge \tau_n} \mathcal{D}_x \Phi \left(u, X_u^\alpha\right)' \sigma \left(u, X_u^\alpha, \alpha_u\right) dW_u \right\}_{t \leq s \leq T} \tag{2.59}$$

is a martingale. Thus its (conditional) expectation is zero:

$$\mathbb{E}_{t,x} \left[\int_t^{s \wedge \tau_n} \mathcal{D}_x \Phi \left(u, X_u^\alpha\right)' \sigma \left(u, X_u^\alpha, \alpha_u\right) dW_u \right] = 0. \tag{2.60}$$

So if we now take the expectation of the stopped candidate value function $\Phi \left(s \wedge \tau, X_{s \wedge \tau}\right)$ conditional on $X_t^\alpha = x$, we obtain:

$$\mathbb{E}_{t,x} \left[\Phi \left(s \wedge \tau_n, X_{s \wedge \tau_n}^\alpha \right) \right] = \Phi(t, x) + \mathbb{E}_{t,x} \left[\int_t^{s \wedge \tau_n} \frac{\partial \Phi}{\partial u} \left(u, X_u^\alpha\right) + \mathcal{L}^{\alpha_u} \Phi \left(u, X_u^\alpha\right) du \right]. \tag{2.61}$$

Now since we have assumed the function Φ satisfies the inequality (2.49) for the particular control that achieves the supremum, we have the same for all admissible controls $\alpha \in \mathcal{A}_{(t,x)}$:

$$\frac{\partial \Phi}{\partial t} \left(u, X_u^\alpha\right) + \mathcal{L}^{\alpha_u} \Phi \left(u, X_u^\alpha\right) + F \left(u, X_u^\alpha, \alpha_u\right) \leq 0. \tag{2.62}$$

Hence

$$\frac{\partial \Phi}{\partial t} \left(u, X_u^\alpha\right) + \mathcal{L}^{\alpha_u} \Phi \left(u, X_u^\alpha\right) \leq -F \left(u, X_u^\alpha, \alpha_u\right) \tag{2.63}$$

and so in our conditional expectation (2.61) above we get the inequality

$$\mathbb{E}_{t,x} \left[\Phi \left(s \wedge \tau_n, X_{s \wedge \tau_n}^\alpha \right) \right] \leq \Phi(t, x) - \mathbb{E}_{t,x} \left[\int_t^{s \wedge \tau_n} F \left(u, X_u^\alpha, \alpha_u\right) du \right]. \tag{2.64}$$

Since $s \wedge \tau_n \leq T$,

$$\left| \int_t^{s \wedge \tau_n} F \left(u, X_u^\alpha, \alpha_u\right) du \right| \leq \int_t^T \mid F \left(u, X_u^\alpha, \alpha_u\right) \mid du < \infty, \tag{2.65}$$

where the integrability of the right-hand side of the above inequality follows from the definition of the admissible set $A_{(t,x)}$ conditional on $\mathcal{F}_t$, (2.7).

Because $|\Phi(t,x)| \leq C\left(1+|x|^2\right)$, Φ satisfies quadratic growth and so we have a finite upper bound on the expected value of the stopped candidate value function $\Phi\left(s \wedge \tau_n, X^{\alpha}_{s\wedge\tau_n}\right)$:

$$\left|\Phi\left(s \wedge \tau_n, X^{\alpha}_{s\wedge\tau_n}\right)\right| \leq C\left(1+\sup_{s\in[t,T]}|X^{\alpha}_s|^2\right) < \infty, \tag{2.66}$$

where the finiteness of this upper bound follows from the properties of the strong solution (2.4).

Since $\lim_{n\to\infty}\tau_n = \infty$ and $s \wedge \tau_n \leq T$, $\lim_{n\to\infty} s\wedge\tau_n = s$ for all $s \in [t,T]$. We apply the Dominated Convergence Theorem and take limits as $n \to \infty$ on the inequality (2.64):

$$\begin{aligned} \lim_{n\to\infty}\mathbb{E}_{t,x}\left[\Phi\left(s\wedge\tau_n, X^{\alpha}_{s\wedge\tau_n}\right)\right] &\leq \Phi(t,x) - \lim_{n\to\infty}\mathbb{E}_{t,x}\left[\int_t^{s\wedge\tau_n} F\left(u, X^{\alpha}_u, \alpha_u\right)du\right]. \\ \implies \mathbb{E}_{t,x}\left[\Phi\left(s, X^{\alpha}_s\right)\right] &\leq \Phi(t,x) - \mathbb{E}_{t,x}\left[\int_t^{s} F\left(u, X^{\alpha}_u, \alpha_u\right)du\right]. \end{aligned} \tag{2.67}$$

We have assumed that Φ is continuous on $[0,T]\times\mathbb{R}_n$. Thus we can take limits as $s \to T$, and apply the Dominated Convergence Theorem again:

$$\begin{aligned} \lim_{s\to T}\mathbb{E}_{t,x}\left[\Phi\left(s, X^{\alpha}_s\right)\right] &\leq \Phi(t,x) - \lim_{s\to T}\mathbb{E}_{t,x}\left[\int_t^{s} F\left(u, X^{\alpha}_u, \alpha_u\right)du\right] \\ \implies \mathbb{E}_{t,x}\left[\Phi\left(T, X^{\alpha}_T\right)\right] &\leq \Phi(t,x) - \mathbb{E}_{t,x}\left[\int_t^{T} F\left(u, X^{\alpha}_u, \alpha_u\right)du\right], \end{aligned} \tag{2.68}$$

and since we assumed the terminal condition $\Phi(T,x) \geq G(x)$ for all $x \in \mathbb{R}^n$,

$$\mathbb{E}_{t,x}\left[G\left(X^{\alpha}_T\right)\right] \leq \Phi(t,x) - \mathbb{E}_{t,x}\left[\int_t^{T} F\left(u, X^{\alpha}_u, \alpha_u\right)du\right]. \tag{2.69}$$

Now since our choice of admissible control was arbitrary, we can say for all admissible controls $\alpha \in \mathcal{A}_{(t,x)}$ and all $(t,x) \in [0,T]\times\mathbb{R}^n$,

$$\begin{aligned} \Phi(t,x) &\geq \mathbb{E}_{t,x}\left[G\left(X^{\alpha}_T\right) + \int_t^{T} F\left(u, X^{\alpha}_u, \alpha_u\right)du\right] \\ \implies \Phi(t,x) &\geq J(t,x,\alpha). \end{aligned} \tag{2.70}$$

This means that given any admissible control, the candidate value function is always greater than or equal to the gain function obtained using that control.

To prove the equality, we select the measurable function $\alpha^*(t,x) \in \mathcal{A}$ as our candidate optimal control, and we perform the same steps to the process

$\Phi(u, X_u^{\alpha^*})$ for where u ranges between $t \in [0, T)$ and $s \in [t, T)$. We consider again the stopped process $\Phi\left(s \wedge \tau, X_{s\wedge\tau}^{\alpha^*}\right)$: Itô's Lemma yields us this:

$$\Phi\left(s \wedge \tau, X_{s\wedge\tau}^{\alpha^*}\right) = \Phi(t,x) + \int_t^{s\wedge\tau} \frac{\partial \Phi}{\partial u}\left(u, X_u^{\alpha^*}\right) + \mathcal{L}^{\alpha^*}\Phi\left(u, X_u^{\alpha^*}\right) du \\ + \int_t^{s\wedge\tau} \mathcal{D}_x\Phi\left(u, X_u^{\alpha^*}\right)' \sigma\left(u, X_u^{\alpha^*}, \alpha_u^*\right) dW_u. \qquad (2.71)$$

We again choose a stopping time $\tau \in [t, T]$ such that

$$\tau = \tau_n := \inf\left\{s \geq t : \int_t^s \left|\mathcal{D}_x\Phi\left(u, X_u^{\alpha^*}\right)' \sigma\left(u, X_u^{\alpha^*}, \alpha^*\right)\right|^2 du \geq n\right\}. \qquad (2.72)$$

Then we can define $\{\tau_n\}_{n\in\mathbb{N}}$ as a localizing sequence of stopping times such that the stopped stochastic integral is a martingale. So, taking the expectation conditional on $X_t^\alpha = x$ is gives us:

$$\mathbb{E}_{t,x}\left[\Phi\left(s \wedge \tau_n, X_{s\wedge\tau_n}^{\alpha^*}\right)\right] \\ = \Phi(t,x) + \mathbb{E}_{t,x}\left[\int_t^{s\wedge\tau_n} \frac{\partial \Phi}{\partial u}\left(u, X_u^{\alpha^*}\right) + \mathcal{L}^{\alpha^*}\Phi\left(u, X_u^{\alpha^*}\right) du\right]. \qquad (2.73)$$

Using the assumptions that $\Phi(t,x) \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^n)$ and that $|\Phi(t,x)| \leq C\left(1 + |x|^2\right)$, we take limits as $n \to \infty$ (recalling that $\lim_{n\to\infty} \tau_n = \infty$ and $s \wedge \tau_n \leq T$ implies that $\lim_{n\to\infty} s \wedge \tau_n = s$ for all $s \in [t, T)$) and apply the Dominated Convergence Theorem to obtain

$$\mathbb{E}_{t,x}\left[\Phi\left(s, X_s^{\alpha^*}\right)\right] = \Phi(t,x) + \mathbb{E}_{t,x}\left[\int_t^s \frac{\partial \Phi}{\partial u}\left(u, X_u^{\alpha^*}\right) + \mathcal{L}^{\alpha^*}\Phi\left(u, X_u^{\alpha^*}\right) du\right]. \qquad (2.74)$$

From the definition of our candidate optimal control α^* and (2.53),

$$\frac{\partial}{\partial t}\Phi(t,x) + \mathcal{L}^{\alpha^*(t,x)}\Phi(t,x) = -F(t,x,\alpha^*(t,x)). \qquad (2.75)$$

So we can express the conditional expectation of the candidate value function evaluated under the optimal control as

$$\mathbb{E}_{t,x}\left[\Phi\left(s, X_s^{\alpha^*}\right)\right] = \Phi(t,x) - \mathbb{E}_{t,x}\left[\int_t^s F(t,x,\alpha^*(t,x)) du\right].$$

We now take limits on the above equality as $s \to T$, and obtain

$$\lim_{s\to T} \mathbb{E}_{t,x}\left[\Phi\left(s, X_s^{\alpha^*}\right)\right] = \Phi(t,x) - \lim_{s\to T} \mathbb{E}_{t,x}\left[\int_t^s F(t,x,\alpha^*(t,x)) du\right]. \\ \implies \mathbb{E}_{t,x}\left[\Phi\left(T, X_T^{\alpha^*}\right)\right] = \Phi(t,x) - \mathbb{E}_{t,x}\left[\int_t^T F(t,x,\alpha^*(t,x)) du\right],$$

and since we assumed in this second optimal case the terminal condition $\Phi(T, x) = G(x)$ for all $x \in \mathbb{R}^n$ and all $\alpha \in \mathcal{A}$,

$$\mathbb{E}_{t,x}\left[G\left(X_T^{\alpha^*}\right)\right] = \Phi(t, x) - \mathbb{E}_{t,x}\left[\int_t^T F\left(u, X_u^{\alpha^*}, \alpha^*\right) du\right]. \tag{2.76}$$

Therefore we have our desired equality: for all $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$\begin{aligned}
\Phi(t, x) &= \mathbb{E}_{t,x}\left[G\left(X_T^{\alpha^*}\right) + \int_t^T F\left(u, X_u^{\alpha^*}, \alpha^*\right) du\right] \\
&= J(t, x, \alpha^*(t, x)) \\
&= V(t, x).
\end{aligned} \tag{2.77}$$

Therefore our candidate value function $\Phi(t, x)$ is the true value function $V(t, x)$ which solves the stochastic control problem, and our candidate optimal control $\alpha^*(t, x)$ is the true optimal control which achieves this value function.

□

Chapter 3

Optimal Execution for a Risk-Averse Trader using Market Orders

In this chapter we design optimal execution algorithms for a risk-averse trader using market orders. The trader aims to maximise his expected utility of terminal wealth. We incorporate the trader's risk aversion via the parameter of the utility functions used.

Consider a trader executing a large order to liquidate a long position in a particular asset by selling inventory of this asset within a time period $[0, T]$. Our trader wishes to liquidate the position in a way such as to minimise loss of revenue incurred due to price impact, when other market participants detect they are selling a large amount of shares and thus move their bids downwards, forcing the trader to sell at lower prices and hence gain less revenue. We assume our trader can control the rate at which they liquidate the position over time. Almgren and Chriss [AC00] and Cartea, Jaimungal, and Penalva [CJP15] in Section 6.5 of their book considered an agent who wishes to maximise terminal wealth in a risk-neutral way. Schied and Schoeneborn [SS09] found an explicit solution for the case of a risk-averse trader seeking to maximize exponential utility of terminal wealth. In Section 6.6 of their book [CJP15], Cartea et al found a link between their own risk-neutral trader trading with a quadratic running penalty function punishing non-zero inventory over the trading period, and the exponential utility case.

3.1 The Stochastic Model

Assume that we work on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(0\leq t\leq T)}, \mathbb{P})$ with the filtration $(\mathcal{F}_t)_{(0\leq t\leq T)}$ satisfying the "usual conditions" of being complete: each subfiltration $\mathcal{F}_t$ contains all of the "null" sets of probability zero of the full filtration $\mathcal{F}$, and being right-continuous, i.e. $\mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s = \mathcal{F}_t$.

Following Cartea et al [CJP15], we introduce the following stochastic model for the problem. We denote by v_t the trading rate at which the trader liquidates their position. This trading rate is the control process for the stochastic control problem. The rate v is a member of the admissible set $\mathcal{A}$. In our control problem, an admissible trading strategy must be square-integrable, and $\mathcal{F}$-predictable.

The trader's inventory or position is represented by Q_t^v, and satisfies the SDE:

$$dQ_t^v = -v_t dt, \quad Q_0^v = q. \tag{3.1}$$

The asset midprice is modelled as a generalised Brownian motion with constant volatility σ. We can assume the volatility is constant over the very short periods of time in which our high-frequency trader trades, as mentioned in Avellaneda and Stoikov's seminal paper on market-making [AS08]. [1]

The asset's growth rate is affected by the permanent impact $g(v_t)$ of the trading of market participants, which in turn is a function of our agent's speed of trading:

$$dS_t^v = -g(v_t)dt + \sigma dW_t, \quad S_0^v = S. \tag{3.2}$$

The trader is able to execute their selling trades at the execution price $\tilde{S}_t^v$, which is calculated by the midprice minus the temporary price impact $f(v_t)$ incurred by the trader's selling. [2]

$$\tilde{S}_t^v = S_t^v - f(v_t), \quad \tilde{S}_0^v = S. \tag{3.3}$$

The functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ and $g : \mathbb{R}_+ \to \mathbb{R}_+$ represent the temporary price impact and the permanent price impact respectively. Both functions are modelled to depend on the trader's rate of trading v_t.

Temporary price impact is incurred when our trader's sell orders lead to other market participants adjusting their offers to sell stock downwards. This action by other market participants impacts on the execution price at which our trader is able to sell stock. The more stock our trader sells in a given time period (i.e. the greater our rate of trading v_t), the greater the temporary impact is on their execution price. The impact is called "temporary" because it only affects our own trader's execution price at a given time, not the actual midprice of the asset.

In contrast, permanent impact is incurred by the general trading activity of all market participants and affects the market price of the asset. Because our trader's activity also contributes to permanent impact, we model permanent impact as a function of the trading rate v_t. In this case we incorporate the effect of other market participants in the volatility parameter σ of the midprice process (3.2). [3]

[1] It is possible to consider more sophisticated models for the volatility, including stochastic volatility, volatility driven by order flow Poisson processes.

[2] The execution price can be made much more sophisticated, and incorporate many aspects of market microstructure, including bid-ask spread and order imbalance.

[3] In more sophisticated models, for example in the order flow model by Cartea and Jaimungal [CJ16b], the action of other traders is encapsulated in a new variable ψ_t, which satisfies

From this, we can state that the trader's cash process X_t has the dynamics:

$$\begin{aligned} dX_t &= -\tilde{S}_t dQ_t && (3.5) \\ &= (S_t - f(v_t))v_t dt. && (3.6) \end{aligned}$$

We formulate the problem as a 3-dimensional system with three state variables: the total wealth X_t, the asset midprice S_t, and the inventory Q_t.

$$dZ_t = \begin{pmatrix} dX_t \\ dS_t \\ dQ_t \end{pmatrix} = \begin{pmatrix} S_t v_t - f(v_t)v_t \\ -g(v_t) \\ -v_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t. \tag{3.7}$$

In this way we have a vector of drifts $\mu(t, z, v)$ and a matrix of volatilites $\sigma(t, z, v)$ which are both Lipschitz-continuous and integrable, and which satisfy:

$$\mu(t, z, v) = \begin{pmatrix} S_t v_t - f(v_t)v_t \\ -g(v_t) \\ -v_t \end{pmatrix} \tag{3.8}$$

and

$$\sigma(t, z, v) = \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix}. \tag{3.9}$$

The admissible set $\mathcal{A}$ is a set of controls which satisfy certain properties making them realistic and regular. For example, admissible trading strategies must be square-integrable (so as not to be infinite and therefore impossible to implement in a world of finite resources and speed), and depend only on the information available up to a given time τ in the trading horizon $[t, T]$ (so a trader cannot look into the future in order to design a perfect strategy).

Using Huyên Pham's [Pha09] formulation for a finite-horizon problem, the admissible set $\mathcal{A}$ is the set of control processes $v \in A$ such that:

$$\mathbb{E}\left[\int_0^T |\mu(t, x, S, q, v_t)|^2 + |\sigma(t, x, S, q, v_t)|^2 \, dt\right] < \infty. \tag{3.10}$$

That is, the control processes v are such that the drift and the diffusion coefficients of the controlled processes X, S, Q in the underlying model are both square-integrable, and so is their sum, i.e. they satisfy a quadratic growth condition. This prevents all of the controlled processes from "blowing up" to infinity. These admissible controls ensure that the controlled processes have the properties which allow there to be a strong solution to the SDE **??** above, for

the equation

$$d\psi_t = -\kappa\psi_t dt + \eta dL_t, \tag{3.4}$$

where L_t are independent Poisson processes each with equal intensity λ, which represent buy and sell order flows, and κ the rate of decay of the order flow. Then the permanent impact function g_t becomes a function of both our own trader's rate of trading v_t and the order flow ψ_t coming into the limit order book.

all controls $v \in \mathcal{A}$ and any initial condition $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$, when the controlled processes begin at a particular value $X_t^v = x, S_t^v = S, Q_t^v = q$. We can follow Pham and denote this solution by $\{(X_u^v,\ S_u^v,\ Q_u^v)\ ;\ t \le u \le T\}$. The conditions of Lipschitz continuity (2.2) and quadratic growth (2.3) imply that the strong solution $X_s^{t,x}$ satisfies these properties:

$$\mathbb{E}_{t,x}\left[\sup_{t \le s \le T} ||(X_t^v, S_t^v, Q_t^v)||^2\right] < \infty \tag{3.11}$$

$$\lim_{h \downarrow 0^+} \mathbb{E}_{t,x}\left[\sup_{s \in [t,t+h]} ||(X_t^v - x,\ S_t^v - S,\ Q_t^v - q)||^2\right] = 0. \tag{3.12}$$

Now for $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$, we denote by $\mathcal{A}_{(t,x,S,q)}$ the subset of controls in $\mathcal{A}$ such that the function f has finite expected value on the interval [t,T], given that the state variables have attained particular values at time t:

$$\mathbb{E}_{t,x,S,q}\left[\int_t^T f(u, X_u^v, S_u^v, Q_u^v, v_u) ds\right] < \infty. \tag{3.13}$$

In this model, all of the randomness comes from the Brownian motion which drives the asset price. The three-dimensional stochastic model can then be expressed as a one-dimensional model. We can express each state variable as a function only of the control v and the Brownian motion W, and initial values for the asset price S_0, the trader's cash X_0, inventory Q_0.

Given a particular trading strategy v, we have for the asset price S_t^v:

$$S_t^v = S_0^v - \int_0^t g(v_u) du + \sigma \int_0^t dW_u, \tag{3.14}$$

the inventory process Q_t^v:

$$Q_t^v = Q_0^v - \int_0^t v_u du, \tag{3.15}$$

and the cash process X_t^v:

$$X_t^v = X_0^v + \int_{u=0}^{u=t} \left(S_0^v - \int_{r=0}^{r=u} g(v_r) dr + \sigma \int_{r=0}^{r=u} dW_r\right) - f(v_u) v_u du. \tag{3.16}$$

If the trader has not liquidated the entire position by the terminal time T, he or she is permitted to execute a final market order to immediately sell a quantity Q_T of the asset at price S_T, and incur a penalty given by $-\alpha Q_T$ for this transaction. This penalty takes into account any fee, slippage, or price impact on that final market order. This model follows that used by Cartea et al [CJP15] in their book. Hence our trader's terminal wealth is given by the sum of their cash at time T and this additional cash flow brought by the final sale at time T.

$$X_T^v + Q_T^v \left(S_T^v - \alpha Q_T^v\right). \tag{3.17}$$

3.2 The Stochastic Control problem

Our trader seeks to optimise the expected utility of his or her terminal wealth. He or she seeks to solve the stochastic control problem with performance criteria and value function:

$$H^v(t,x,S,q) = \mathbb{E}_{t,x,S,q}\left[U\left(X_T^v + Q_T^v\left(S_T^v - \alpha Q_T^v\right)\right)\right]; \tag{3.18}$$
$$H(t,x,S,q) = \sup_{v\in\mathcal{A}} H^v(t,x,S,q). \tag{3.19}$$

The Dynamic Programming Principle (2.1.1) implies that for all $(t,x,S,q) \in [0,T]\times\mathbb{R}^3$, and all stopping times $\tau \leq T$,

$$H(t,x,S,q) = \sup_{v\in\mathcal{A}_{(t,x)}} \mathbb{E}_{t,x,S,q}\left[H(\tau, X_\tau^v, S_\tau^v, Q_\tau^v)\right]. \tag{3.20}$$

Then if $v^* \in \mathcal{A}$ is the optimal control, which achieves the supremum over the interval $[0,T]$, the value function satisfies

$$H(t,x,S,q) = \mathbb{E}_{t,x,S,q}\left[H(T, X_T^{v^*}, S_T^{v^*}, Q_T^{v^*})\right]. \tag{3.21}$$

Using the Dynamic Programming Principle (2.1.1), we obtain the Hamilton-Jacobi-Bellman Equation for the value function (writing $\partial_\bullet H for \frac{\partial H}{\partial \bullet}$):

$$\partial_t H + \frac{1}{2}\sigma^2\partial_{SS}H + \sup_{v\in\mathcal{A}}\left(Sv\partial_x H - f(v)v\partial_x H - g(v)\partial_S H - v\partial_q H\right) = 0, \tag{3.22}$$

3.3 Verification

We want to find a candidate value function and a candidate optimal control that solves the stochastic control problem (3.18). To verify that our candidate functions solve the problem, we will follow the method suggested in the proof in Pham's book ([Pha09]).

Proposition 3.3.1 (Verification of our Candidate Value Function and Candidate Optimal Control)**.** *Let $\Phi(t,x,S,q) \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^3)$ and let $\Phi(t,x,S,q)$ satisfy a quadratic growth condition, that is for all for all $(t,x,S,q) \in [0,T]\times\mathbb{R}^3$, there exists a constant C such that:*

$$|\Phi(t,x,S,q)| \leq C\left(1 + ||(x,S,q)||^2\right). \tag{3.23}$$

We have 3 controlled processes and 1 control process in our model (3.7). We introduce the running penalty/reward function $F : [0,T]\times\mathbb{R}^3\times A \to \mathbb{R}$ and the terminal reward $G : \mathbb{R}^3 \to \mathbb{R}$, both measurable uniformly bounded functions. Suppose that F and G satisfy:

$$\frac{\partial}{\partial t}\Phi(t,x,S,q) - \sup_{v\in\mathcal{A}}\left(\mathcal{L}^v\Phi(t,x,S,q) + F(t,x,S,q,v)\right) \geq 0, \tag{3.24}$$

for all $(t,x,S,q) \in [0,T] \times \mathbb{R}^3$*, and*

$$\Phi(T,x,S,q) \geq G(x,s,q) \text{ for all } (x,S,q) \in \mathbb{R}^3, \tag{3.25}$$

then

$$\Phi(t,x,S,q) \geq H^v(t,x,S,q) \text{ on } [0,T] \times \mathbb{R}^3. \tag{3.26}$$

Now suppose that $\Phi(T,x,S,q) = G(x,S,q)$*, and that there exists a measurable function* $\nu^*(t,x,S,q)$ *for* $(t,x,S,q) \in [0,T] \times \mathbb{R}^3$*, with values in the admissible set* $\mathcal{A}$*, which satisfies:*

$$\frac{\partial}{\partial t}\Phi(t,x,S,q) - \sup_{\nu \in \mathcal{A}} \left(\mathcal{L}^\nu \Phi(t,x,S,q) + F(t,x,S,q,\nu)\right) = 0, \tag{3.27}$$

that is,

$$\frac{\partial}{\partial t}\Phi(t,x,S,q) - \left(\mathcal{L}^{\nu^*(t,x,S,q)}\Phi(t,x,S,q) + F(t,x,S,q,\nu^*(t,x,S,q))\right) = 0. \tag{3.28}$$

If the stochastic differential equation

$$dZ_t = \begin{pmatrix} dX_t \\ dS_t \\ dQ_t \end{pmatrix} = \begin{pmatrix} S_t\nu^*(t,x,S,q) - f(\nu^*(t,x,S,q))\nu^*(t,x,S,q) \\ -g(\nu^*(t,x,S,q)) \\ -\nu^*(t,x,S,q) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t \tag{3.29}$$

i.e.

$$dZ_t = \begin{pmatrix} dX_t \\ dS_t \\ dQ_t \end{pmatrix} = \begin{pmatrix} S_t\nu_t^* - f(\nu_t^*)\nu_t^* \\ -g(\nu_t^*) \\ -\nu_t^* \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t. \tag{3.30}$$

admits a unique solution, denoted by $Z_r^{\nu^*} = (X_r^{\nu^*}, S_r^{\nu^*}, Q_r^{\nu^*})$*, with* $Z_0 = z = (x,S,q)$*, then*

$$\Phi(t,x,S,q) = H(t,x,S,q) \text{ on } [0,T] \times \mathbb{R}^3. \tag{3.31}$$

and $\nu^*(t,x,S,q)$ *is an optimal control.*

We apply the verification theorem to our own stochastic control problem, given the model (3.7) above. Let $\Phi := \Phi(t,x,S,q)$ be our candidate value function. Applying Itô's Lemma to the candidate value function gives us:

$$d\Phi = \frac{\partial \Phi}{\partial t}dt + \frac{\partial \Phi}{\partial x}dX_t + \frac{\partial \Phi}{\partial S}dS_t + \frac{\partial \Phi}{\partial q}dQ_t + \frac{1}{2}\frac{\partial^2 \Phi}{\partial S^2}\sigma^2 dt. \tag{3.32}$$

We know from the model (3.7) the terms dX_t, dS_t, dQ_t, and that the mixed 2^{nd}-order partial derivatives vanish apart from for $(dS_t)^2$, which being driven by the Brownian motion W has quadratic variation T and hence $< dS, dS >_t = \sigma^2 dt$.

We have

$$\begin{aligned}
d\Phi = &\left(\frac{\partial \Phi}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi}{\partial S^2}\sigma^2\right) dt \\
&+ \frac{\partial \Phi}{\partial x}\big((S_t - f(v_t))\, v_t dt\big) \\
&+ \frac{\partial \Phi}{\partial S}\big(-g(v_t)dt + \sigma dW_t\big) \\
&+ \frac{\partial \Phi}{\partial q}\left(-v_t dt\right).
\end{aligned} \tag{3.33}$$

Collecting terms in dt and dW_t, we obtain:

$$\begin{aligned}
&d\Phi(t, X_t, S_t, Q_t) \\
&= \left(\frac{\partial \Phi}{\partial t} + \big(S_t v_t - f(v_t)v_t\big)\frac{\partial \Phi}{\partial x} - g(v_t)\frac{\partial \Phi}{\partial S} - v_t\frac{\partial \Phi}{\partial q} + \frac{1}{2}\sigma^2\frac{\partial^2 \Phi}{\partial S^2}\right) dt \\
&\quad + \sigma\frac{\partial \Phi}{\partial S} dW_t.
\end{aligned} \tag{3.34}$$

From this we can write the candidate value function in full integral form. At any time $t \in [0, T]$ and for an arbitrary control v in the admissible set $\mathcal{A}$:

$$\begin{aligned}
&\Phi(t, X_t, S_t, Q_t) = \Phi(0, X_0, S_0, Q_0) \\
&+ \int_0^t \left(\frac{\partial \Phi}{\partial u} + \big(S_u v_u - f(v_u)v_u\big)\frac{\partial \Phi}{\partial x} - g(v_u)\frac{\partial \Phi}{\partial S} - v_u\frac{\partial \Phi}{\partial q} + \frac{1}{2}\sigma^2\frac{\partial^2 \Phi}{\partial S^2}\right) du \\
&+ \int_0^t \sigma\frac{\partial \Phi}{\partial S} dW_u.
\end{aligned} \tag{3.35}$$

The infinitesimal generator of the performance criteria for this control problem:

$$\mathcal{L}_t^v \Phi(t, X_t, S_t, Q_t) = \big(Sv - f(v)v\big)\frac{\partial \Phi}{\partial X} - g(v)\frac{\partial \Phi}{\partial S} - v\frac{\partial \Phi}{\partial q} + \frac{1}{2}\sigma^2\frac{\partial^2 \Phi}{\partial S^2}. \tag{3.36}$$

Substituting this into the expression for the performance criteria we obtain the simpler representation

$$\Phi(t, X_t, S_t, Q_t) = \Phi(0, X_0, S_0, Q_0) + \int_0^t \left(\frac{\partial \Phi}{\partial u} + \mathcal{L}_u^v \Phi\right) du + \int_0^t \sigma\frac{\partial \Phi}{\partial S} dW_u. \tag{3.37}$$

Since $\Phi \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^3)$, we can use Itô's lemma to write for all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$ and all controls which are admissible over $[t, T]$ given that $X_t^v = x$, $S_t^v = S$, $Q_t^v = q$, ie. all $v \in \mathcal{A}_{(t,x,S,q)}$, any time $r \in [t, T)$, and any stopping time $\tau \in [t, T]$,

$$\begin{aligned}
&\mathbb{E}_{t,x,S,q}\left[\Phi(r\wedge\tau, X_{r\wedge\tau}, S_{r\wedge\tau}, Q_{r\wedge\tau})\right] = \Phi(t, X_t, S_t, Q_t) \\
&\quad + \mathbb{E}_{t,x,S,q}\left[\int_t^{r\wedge\tau}\left(\frac{\partial \Phi}{\partial u} + \mathcal{L}_u^v \Phi\right) du\right] + \mathbb{E}_{t,x,S,q}\left[\int_t^{r\wedge\tau}\sigma\frac{\partial \Phi}{\partial S} dW_u\right].
\end{aligned} \tag{3.38}$$

We choose the stopping time $\tau := \tau_n \in [t, T]$ to be such that

$$\tau_n := \inf \left\{ r \geq t \text{ such that } \int_t^r \left| \sigma \frac{\partial \Phi}{\partial S} \right|^2 du \geq n \right\}. \tag{3.39}$$

Then we can define $\{\tau_n\}_{n \in \mathbb{N}}$ as a localizing sequence of stopping times such that the stopped stochastic integral is a martingale. Then

$$\begin{aligned} &\mathbb{E}_{t,x,S,q} \left[\Phi(r \wedge \tau, X_{r \wedge \tau}, S_{r \wedge \tau}, Q_{r \wedge \tau}) \right] \\ &= \Phi(t, X_t, S_t, Q_t) + \mathbb{E}_{t,x,S,q} \left[\int_t^{r \wedge \tau} \left(\frac{\partial \Phi}{\partial u} + \mathcal{L}_u^v \Phi \right) du \right]. \end{aligned} \tag{3.40}$$

Now since we have assumed the function Φ satisfies the inequality (3.26) for the particular control that achieves the supremum, we have the same for all admissible controls $v \in \mathcal{A}_{(t,x,S,q)}$:

$$\frac{\partial \Phi}{\partial t} \left(u, X_u^v, S_u^v, Q_u^v \right) + \mathcal{L}^{v_u} \Phi \left(u, X_u^v, S_u^v, Q_u^v \right) \leq -F \left(u, X_u^v, S_u^v, Q_u^v, v_u \right) \tag{3.41}$$

In our current problem, $F \left(u, X_u^v, S_u^v, Q_u^v, v_u \right) = 0$, since there is no running penalty or reward function. The terminal function $G \left(T, X_T^v, S_T^v, Q_T^v, v_T \right) = U \left(X_T^v, S_T^v, Q_T^v \right)$. Thus for any time $r \in [t, T]$,

$$\mathbb{E}_{t,x,S,q} \left[\Phi(r, X_r, S_r, Q_r) \right] \leq \Phi(t, X_t, S_t, Q_t). \tag{3.42}$$

Using the Dominated Convergence Theorem as $r \to T$,

$$\mathbb{E}_{t,x,S,q} \left[\Phi(T, X_T, S_T, Q_T) \right] \leq \Phi(t, X_t, S_t, Q_t). \tag{3.43}$$

And since we assumed the terminal condition $\Phi(T, x, S, q) \geq G(x, S, q)$ for all $(x, S, q) \in \mathbb{R}^3$,

$$\mathbb{E}_{t,x,S,q} \left[G \left(T, X_T^v, S_T^v, Q_T^v, v_T \right) \right] \leq \Phi(t, X_t, S_t, Q_t). \tag{3.44}$$

Now since our choice of admissible control was arbitrary, we can say for all admissible controls $\alpha \in \mathcal{A}_{(t,x,S,q)}$ and all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$,

$$\begin{aligned} &\Phi(t, x, S, q) \geq \mathbb{E}_{t,x,S,q} \left[U \left(X_T^v, S_T^v, Q_T^v \right) \right] \\ \implies &\Phi(t, x, S, q) \geq H^v(t, x, S, q). \end{aligned} \tag{3.45}$$

Similar arguments give us the following equality for the optimal control v^*: for all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$,

$$\begin{aligned} &\Phi(t, x, S, q) \geq \mathbb{E}_{t,x,S,q} \left[U \left(X_T^{v^*}, S_T^{v^*}, Q_T^{v^*} \right) \right] \\ \implies &\Phi(t, x, S, q) = H(t, x, S, q). \end{aligned} \tag{3.46}$$

3.4 Solving the Stochastic Control Problem

3.4.1 Linear Temporary and Permanent Price Impact

We follow Cartea et al [CJP15] and assume that both temporary and permanent price impact are linear functions of the speed of trading: that is: $f(v_t) = \kappa v_t$ and $g(v_t) = \eta v_t$.

We then get the stochastic model

$$dZ_t = \begin{pmatrix} dX_t \\ dS_t \\ dQ_t \end{pmatrix} = \begin{pmatrix} S_t v_t - \kappa v_t^2 \\ -\eta v_t \\ -v_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t. \tag{3.47}$$

and the HJB equation:

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \sup_{v\in\mathcal{A}} \left(Sv\partial_x H - \kappa v^2 \partial_x H - \eta v \partial_S H - v\partial_q H \right) = 0. \tag{3.48}$$

3.4.2 Feedback form for the Optimal Control

We aim to solve the above HJB equation (3.48). First, we find the optimal control v^* in feedback form. Substituting that feedback form back into the HJB equation will yield us a (possibly non-linear) partial differential equation for the value function $H(t,x,S,q)$. When that PDE in H is solved, it will give us the optimized value function $H(t,x,S,q) = H^{v^*}(t,x,S,q)$ at each time $t \in [0,T]$. From this, we will attempt to find an explicit expression for the optimal control v^*.

To solve the HJB equation (3.48), first, we assume that v^* is the trading speed which achieves the supremum in the equation. Then

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \Big(Sv^*\partial_x H - \kappa {v^*}^2 \partial_x H - \eta v^* \partial_S H - v^*\partial_q H \Big) = 0. \tag{3.49}$$

The equation is a quadratic in v^*.

$$0 = \partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \left(S\partial_x H - \eta\partial_S H - \partial_q H\right) v^* - \left(\kappa \partial_x H\right) {v^*}^2. \tag{3.50}$$

Completing the square gives us:

$$\begin{aligned} 0 = &- \kappa\partial_x H \left(v^* - \frac{(S\partial_x H - \eta\partial_S H - \partial_q H)}{\partial_x H} \right)^2 \\ &+ \partial_t H + \frac{1}{2}\partial_{SS} H - \frac{(S\partial_x H - \eta\partial_S H - \partial_q H)^2}{-4\kappa\partial_x H}. \end{aligned} \tag{3.51}$$

Thus the value of v^* which satisfies the HJB equation is:

$$v^* = \frac{1}{2\kappa} \frac{(S\partial_x H - \eta\partial_S H - \partial_q H)}{\partial_x H}. \tag{3.52}$$

Substituting the optimal control in feedback form into the HJB equation gives us the non-linear PDE:

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \frac{1}{4\kappa}\frac{(S\partial_x H - \eta\partial_S H - \partial_q H)^2}{\partial_x H} = 0. \tag{3.53}$$

We now need to find the candidate value function $H(t, X_t, S_t, Q_t)$ which solves this PDE.

3.5 A general ansatz for the Value Function, based on the Utility Function

We introduce an ansatz for the value function:

$$H(t, x, S, q) = U\left(x + Sq + h(t, q)\right), \tag{3.54}$$

with terminal condition

$$H(T, x, S, q) = U\left(x + Sq - \alpha q^2\right). \tag{3.55}$$

This ansatz is based on the form of the ansatzes used in Cartea et al's book [CJP15], where at time t, x represents the cash obtained up to time t, Sq represents the marked-to-market book value of the units of the asset remaining to be liquidated, and the final term $h(t, q)$ represents the value added due to optimal execution of those shares.

Given this ansatz, we obtain the partial derivatives of the value function as follows. Let $z = z(t, x, S, q) := (x + Sq + h(t, q))$ and denote the first derivative of our utility function U with respect to z by U'_z and the second derivative of the utility function with respect to z by U''_{zz}. Also denote the derivative of z with respect to a variable u by z_u. Then we have the partial derivatives of the value function H as:

$$\begin{aligned}\partial_t H(t, x, S, q) &= U'_z \; z'_t \\ &= U'_z \; \partial_t h(t, q).\end{aligned} \tag{3.56}$$

$$\begin{aligned}\partial_x H(t, x, S, q) &= U'_z \; z'_x \\ &= U'_z \; 1.\end{aligned} \tag{3.57}$$

$$\begin{aligned}\partial_S H(t, x, S, q) &= U'_z \; z'_S \\ &= U'_z \; q.\end{aligned} \tag{3.58}$$

$$\begin{aligned}\partial_{SS} H(t,x,S,q) &= \partial_S U'_z\ q \\ &= q \partial_S U'_z \\ &= q U''_{zz}\ z'_S \\ &= q^2 U''_{zz}. \end{aligned} \tag{3.59}$$

$$\begin{aligned}\partial_q H(t,x,S,q) &= U'_z\ z'_q \\ &= U'_z\ \left(S + \partial_q h(t,q)\right). \end{aligned} \tag{3.60}$$

We substitute this ansatz (3.54) and its partial derivatives into the PDE (3.53) and obtain (where we suppress $H(t,x,S,q)$ to H):

$$\begin{aligned} 0 &= \partial_t h(t,q) U'_z \\ &+ \frac{1}{2}\sigma^2 q^2 U''_{zz} \\ &+ \frac{1}{4\kappa} \frac{\left(SU'_z - \eta\left(q - \left(S + \partial_q h(t,S,q)\right)\ U'_z\right)\right)^2}{U'_z}, \end{aligned} \tag{3.61}$$

Cancelling the terms in S and dividing through by the first derivative of the utility function U'_z gives us a new PDE in the simpler function $h(t,q)$:

$$0 = \partial_t h(t,q) + \frac{1}{2}\sigma^2 q^2 \frac{U''_{zz}}{U'_z} + \frac{1}{4\kappa}\left(\eta q + \partial_q h(t,q)\right)^2. \tag{3.62}$$

with terminal condition

$$h(T,q) = -\alpha q^2. \tag{3.63}$$

In this equation we see the absolute risk aversion coefficient $-\frac{U''_{zz}}{U'_z}$ together with the squared volatility in the second term.

The above equation (3.62) is a non-linear PDE whose solution depends on the form of the utility function U.

3.6 Utility Functions

We aim to solve the execution problem (3.18)under the stochastic model (3.7). To do so, we can solve the partial differential equation (3.62) using different utility functions $U(z)$ applied to the expression $z = x + Sq + h(t,q)$.

3.6.1 Linear Utility - The Risk-neutral approach and the Almgren-Chriss solution

When the utility function is simply identity, i.e.

$$U(z) = z, \tag{3.64}$$

we have a risk-neutral situation where the trader seeks to optimise his profits. The trader does not display risk aversion to price impact and so executes his trades evenly through time using a Time-Weighted Average Price (TWAP) algorithm, as shown in Section 6.3 of Cartea, Jaimungal, and Penalva [CJP15]. To incorporate risk-aversion to held inventory into the problem, Cartea et al incorporate a running inventory penalty of the form

$$\theta \int_t^T (Q_u^v)^2 \, du, \;\; \theta \geq 0. \tag{3.65}$$

This penalises inventories larger than 0 throughout the trading period. Cartea et al describe this function as representing the trader's urgency to execute the entire trade. The greater the value of θ, the quicker the optimal strategy liquidates the shares. They state that this can be seen as "equivalent to ambiguity aversion on the part of the agent ... over the midprice, which ... may have a non-zero stochastic drift".

3.7 Exponential Utility

We consider the situation as explored in Section 6.6 of [CJP15] where the execution trader is risk-averse with an exponential utility function, where they wants to maximise exponential utility of their terminal wealth:

$$U(z) = -\exp(-\gamma z). \tag{3.66}$$

The parameter γ here represents the risk-aversion of our trader. The greater γ is, the more risk-averse our trader is. Taking derivatives of the utility function with respect to z gives us

$$U'_z = \gamma \exp(-\gamma z) \tag{3.67}$$

$$U''_{zz} = -\gamma^2 \exp(-\gamma z). \tag{3.68}$$

Therefore we have the coefficient of absolute risk aversion

$$-\frac{U''_{zz}}{U'_z} = \gamma. \tag{3.69}$$

The exponential utility function leads the trader to have a constant absolute risk aversion (CARA). The trader's aversion to risk does not change with the level of the value of z. In our case, $z = x + Sq + h(t,q)$.

3.7.1 Exponential Utility: The Stochastic Control Problem

The trader's performance criteria becomes:

$$H^v(t,x,S,q) = \mathbb{E}_{(t,x,S,q)} \left[-\exp\left(-\gamma (X_T^v + Q_T^v \left(S_T^v - \alpha Q_T^v \right) \right) \right]. \tag{3.70}$$

The value function is:

$$H(t,x,S,q) = \sup_{v \in \mathcal{A}} H^v(t,x,S,q). \tag{3.71}$$

We use the ansatz

$$H(t,x,S,q) = -\exp(-\gamma(x + Sq + h(t,q)). \tag{3.72}$$

Using this ansatz and substituting the absolute risk aversion coefficient (3.69) associated with the exponential utility function into the PDE (3.62), we obtain the equation:

$$0 = -\partial_t h(t,q) + \frac{1}{2}\sigma^2\gamma q^2 - \frac{1}{4\kappa}\left(\eta q + \partial_q h(t,q)\right)^2. \tag{3.73}$$

At this point, Cartea, Jaimungal and Penalva assume that the ansatz solution $h(t,q)$ a quadratic function of the inventory q. Thus

$$h(t,q) = h_2(t)q^2, \tag{3.74}$$

where $h_2(t)$ is a deterministic function of time t, whose subscript implies that the function $h_2(t)$ is a coefficient of q^2. Their motivation for this assumption comes from the form of the terminal condition (3.63) for $h(t,q)$.

Substituting this new ansatz (3.74) into the PDE (3.73) gives us a Riccati equation in $h_2(t)$:

$$\begin{aligned} 0 &= -\partial_t(h_2(t)q^2) + \frac{1}{2}\sigma^2\gamma q^2 - \frac{1}{4\kappa}\left(\eta q + \partial_q(h_2(t)q^2)\right)^2 \\ 0 &= -\partial_t h_2(t) + \frac{1}{2}\sigma^2\gamma - \frac{1}{\kappa}\left(\frac{1}{2}\eta + h_2(t)\right)^2. \end{aligned} \tag{3.75}$$

3.7.2 Solving the Riccati equation in $h_2(t)$

This equation (3.75) is a Riccati equation in $h_2(t)$. To solve it, first we can write

$$\chi(t) = h_2(t) + \frac{1}{2}\eta. \tag{3.76}$$

Then we find the first derivative of $\chi(t)$ with respect to time t:

$$h_2(t) = \chi(t) - \frac{1}{2}\eta \tag{3.77}$$

$$\partial_t h_2(t) = \partial_t \chi(t). \tag{3.78}$$

Then we obtain a Riccati equation in the new variable $\chi(t)$:

$$0 = \partial_t \chi(t) - \frac{1}{2}\sigma^2\gamma + \frac{1}{\kappa}\left(\chi(t)\right)^2. \tag{3.79}$$

Rearranging gives us:

$$\begin{aligned}\partial_t \chi(t) + \frac{1}{\kappa}\left(\chi(t)\right)^2 &= \frac{1}{2}\sigma^2\gamma \\ \frac{\partial \chi(t)}{\partial t} &= \frac{1}{2}\sigma^2\gamma - \frac{1}{\kappa}\left(\chi(t)\right)^2 \\ \frac{\partial \chi(t)}{\partial t} &= \frac{1}{\kappa}\left(\frac{1}{2}\kappa\sigma^2\gamma - \left(\chi(t)\right)^2\right).\end{aligned} \tag{3.80}$$

We can then separate the variables $\chi(t)$ and t, and integrate:

$$\int \left(\frac{1}{\frac{1}{2}\kappa\sigma^2\gamma - \left(\chi(t)\right)^2}\right) d\chi(t) = \int \frac{1}{\kappa} dt. \tag{3.81}$$

The denominator $(\frac{1}{2}\kappa\sigma^2\gamma - \left(\chi(t)\right)^2)$ in the integral is equal to $-(\chi(t) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2})(\chi(t) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2})$, so we get:

$$\int \left(\frac{1}{\left(\chi(t) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right)\left(\chi(t) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right)}\right) d\chi(t) = -\int \frac{1}{\kappa} dt. \tag{3.82}$$

We evaluate the integral over the time interval [t,T]:

$$\int_{s=t}^{s=T} \left(\frac{1}{\left(\chi(t) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right)\left(\chi(t) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right)}\right) d\chi(t) = -\int_{s=t}^{s=T} \frac{1}{\kappa} dt. \tag{3.83}$$

Given that for $a \neq b$,

$$\int \left(\frac{1}{(x+a)(x+b)}\right) dx = \frac{1}{b-a}\left(\ln(x+a) - \ln(x+b)\right), \tag{3.84}$$

evaluating the integral gives us:

$$\begin{aligned}&\frac{1}{\sqrt{2\kappa\gamma\sigma^2}}\left(\ln\left(\chi(T) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right) - \ln\left(\chi(T) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right)\right) \\ &\qquad - \frac{1}{\sqrt{2\kappa\gamma\sigma^2}} - \left(\ln\left(\chi(t) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right) - \ln\left(\chi(t) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}\right)\right) \\ &\qquad = -\frac{1}{\kappa}(T-t).\end{aligned} \tag{3.85}$$

Multiplying through by $\sqrt{2\kappa\gamma\sigma^2}$ gives us on the right hand side (LHS):

$$\begin{aligned}\text{RHS} &= -\frac{\sqrt{2\kappa\gamma\sigma^2}}{\kappa}(T-t) \\ &= -2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t).\end{aligned} \tag{3.86}$$

This leaves us the left-hand side (LHS):

$$\begin{aligned}\text{LHS} &= \left(\ln\left[\frac{\chi(T) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}}{\chi(T) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}}\right] - \ln\left[\frac{\chi(t) - \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}}{\chi(t) + \frac{1}{\sqrt{2}}\sqrt{\kappa\gamma\sigma^2}}\right]\right) \\ &= \left(\ln\left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] - \ln\left[\frac{\chi(t) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(t) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right]\right) \qquad (3.87)\end{aligned}$$

Putting the two sides of the evaluated integral back together, we have:

$$\left(\ln\left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] - \ln\left[\frac{\chi(t) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(t) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right]\right) = -2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t). \qquad (3.88)$$

$$\implies \ln\left[\frac{\chi(t) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(t) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] = \ln\left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] + 2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t). \qquad (3.89)$$

Now we take exponentials:

$$\left[\frac{\chi(t) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(t) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] = \left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right). \qquad (3.90)$$

We continue to solve for $\chi(t)$.

$$\begin{aligned}&\chi(t) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2} \\ &\quad = \left(\chi(t) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\end{aligned} \qquad (3.91)$$

$$\begin{aligned}&\implies \chi(t)\left(1 - \left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right) \\ &\quad = \left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left(1 + \left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right).\end{aligned} \qquad (3.92)$$

Thus we have an expression for $\chi(t)$:

$$\chi(t) = \left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\frac{\left(1 + \left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right)}{\left(1 - \left[\frac{\chi(T) - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T) + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right)}. \qquad (3.93)$$

We use the terminal condition $h_2(T) = -\alpha$ to give us $\chi(T) = -\alpha + \frac{1}{2}b$, from our definition (3.76). Substituting the terminal condition for $\chi(T)$ into our expression for $\chi(t)$ yields us an explicit expression for the ansatz $h_2(t)$:

$$\begin{aligned} h_2(t) &= \chi(t) - \frac{1}{2}\eta \\ &= \left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right) \frac{\left(1 + \left[\frac{\chi(T)-\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T)+\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right)}{\left(1 - \left[\frac{\chi(T)-\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{\chi(T)+\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right)} - \frac{1}{2}\eta \qquad (3.94) \\ &= \left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right) \frac{\left(1 + \left[\frac{-\alpha+\frac{1}{2}\eta-\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{-\alpha+\frac{1}{2}\eta+\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right)}{\left(1 - \left[\frac{-\alpha+\frac{1}{2}\eta-\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{-\alpha+\frac{1}{2}\eta+\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}\right] \exp\left(2\sqrt{\frac{\gamma\sigma^2}{2\kappa}}(T-t)\right)\right)} - \frac{1}{2}\eta. \qquad (3.95) \end{aligned}$$

Now we define constants ξ and ζ as follows:

$$\xi = \sqrt{\frac{\gamma\sigma^2}{2\kappa}}, \qquad (3.96)$$

and

$$\zeta = \frac{-\alpha + \frac{1}{2}\eta - \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}{-\alpha + \frac{1}{2}\eta + \sqrt{\frac{1}{2}\kappa\gamma\sigma^2}}. \qquad (3.97)$$

Substituting ξ and ζ into the formula for $h_2(t)$, we obtain the second ansatz in the form obtained by Cartea et al's Section 6.6, page 151 [CJP15] (with slightly different expressions for the constants).

$$h_2(t) = \left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left(\frac{1 + \zeta\exp\left(2\xi(T-t)\right)}{1 - \zeta\exp\left(2\xi(T-t)\right)}\right) - \frac{1}{2}\eta. \qquad (3.98)$$

3.7.3 Explicit formulae for the Value Function and the Optimal Trading Speed

Now we use the fact that $h(t,q) = q^2 h_2(t)$ to get the original ansatz:

$$h(t,q) = q^2\left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left(\frac{1 + \zeta\exp\left(2\xi(T-t)\right)}{1 - \zeta\exp\left(2\xi(T-t)\right)}\right) - \frac{1}{2}\eta q^2. \qquad (3.99)$$

From this, we get the first derivative of $h(t,q)$ with respect to the quantity q from (3.74):

$$\partial_q h(t,q) = 2q\left(\left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left(\frac{1 + \zeta\exp\left(2\xi(T-t)\right)}{1 - \zeta\exp\left(2\xi(T-t)\right)}\right) - \eta\right). \qquad (3.100)$$

Using the relationship between the ansatz and the value function, we get a candidate value function

$$H(t,x,S,q) = -\exp\left(-\gamma\left(x+Sq+q^2\left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left(\frac{1+\zeta\exp\left(2\xi(T-t)\right)}{1-\zeta\exp\left(2\xi(T-t)\right)}\right)-\frac{1}{2}\eta q^2\right)\right). \tag{3.101}$$

Finally, we now find the optimal trading speed v^* from (3.52):

$$v_t^* = -\xi\left(\frac{1+\zeta\exp\left(2\xi(T-t)\right)}{1-\zeta\exp\left(2\xi(T-t)\right)}\right)Q_t^{v^*}. \tag{3.102}$$

This optimal trading speed is a non-linear function of time, and increases with the risk aversion parameter γ. A more risk-averse trader will prefer to execute more quickly. This speed is also proportional to the trader's held inventory Q_t. Since the trader is averse to the market risk of holding inventory, the greater the position the trader holds, the faster he trades to liquidate that position. The trader trades more quickly closer to the beginning of the trading horizon and then more slowly as time reaches the end of the horizon.

Interestingly, the asset price S_t does not appear in the optimal speed, so the trader does not take into account the price in order to obtain a higher sale price for the client. In our model the price was modelled by a Brownian motion with constant volatility σ and a drift $g(v_t)$ determined by a permanent price impact which is directly proportional to our trader's trading speed: $g(v_t) = gv$. That price impact is affected only by our own trader's activity. Cartea and Jaimungal [CJ16b] introduce models incorporating the order flow from other market participants, and this is a possible extension for this work. The volatility does appear in the formula for the optimal trading speed, and since our trader knows his or her trading speed, that volatility encapsulates all of the uncertainty over the price from the point of view of our trader.

3.7.4 A more explicit representation of the optimal inventory and trading strategy as deterministic functions of time and the initial inventory

We can write the optimal inventory $Q_t^{v^*}$ and the optimal speed v^* directly in terms of the initial inventory Q_0. This allows us to simulate the inventory process along its trajectory.

We find the partial derivatives of H using the ansatz $H(t,x,S,q) = x+Sq+h(t,q)$ to simplify the situation. The partial derivative of H with respect to time t is:

$$\begin{aligned}\partial_t H &= \partial_t\left(-\exp\{-\gamma(x+Sq+h(t,q))\}\right)\\ &= -\gamma\exp\{-\gamma(x+Sq+h(t,q)\}\ \ \partial_t h(t,q)\\ &= -\gamma H\partial_t h(t,q).\end{aligned} \tag{3.103}$$

The first partial derivative of H with respect to the stock price S is:

$$\begin{aligned}\partial_S H &= \partial_S \left(-\exp\{-\gamma(x+Sq+h(t,q))\}\right)\\ &= \partial_S \left(-\exp\{-\gamma x\}\exp\{-\gamma Sq\}\exp\{-\gamma h(t,q)\}\right)\\ &= -\gamma q \exp\{-\gamma(x+Sq-h(t,q))\}\\ &= -\gamma q H. \end{aligned} \tag{3.104}$$

The second partial derivative of H with respect to the stock price S is:

$$\begin{aligned}\partial_{SS} H &= \partial_{SS} \left(-\exp\{-\gamma(x+Sq+h(t,q))\}\right)\\ &= \partial_{SS} \left(-\exp\{-\gamma x\}\exp\{-\gamma Sq\}\exp\{-\gamma h(t,q)\}\right)\\ &= \gamma^2 q^2 \exp\{-\gamma(x+Sq-h(t,q))\}\\ &= \gamma^2 q^2 H. \end{aligned} \tag{3.105}$$

The partial derivative of H with respect to the trader's wealth x is:

$$\begin{aligned}\partial_x H &= \partial_x \left(-\exp\{-\gamma(x+Sq+h(t,q))\}\right)\\ &= \partial_x \left(-\exp\{-\gamma x\}\exp\{-\gamma Sq\}\exp\{-\gamma h(t,q)\}\right)\\ &= -\gamma \exp\{-\gamma(x+Sq-h(t,q))\}\\ &= -\gamma H. \end{aligned} \tag{3.106}$$

The partial derivative of H with respect to the trader's inventory q is:

$$\begin{aligned}\partial_q H &= \partial_q \left(-\exp\{-\gamma(x+Sq+h(t,q))\}\right)\\ &= -\exp\{-\gamma x\}\quad \partial_q \left(\exp\{-\gamma Sq\}\exp\{-\gamma h(t,q)\}\right)\\ &= -\exp\{-\gamma x\}\quad \left(-\gamma(S+\partial_q h(t,q))\exp\{-\gamma(Sq+h(t,q))\}\right)\\ &= -\gamma(S+\partial_q h(t,q))(\exp\{-\gamma(x+Sq+h(t,q))\})\\ &= -\gamma(S+\partial_q h(t,q))H \end{aligned} \tag{3.107}$$

Substituting these partial derivatives into the HJB equation (3.22) we obtain the simpler PDE

$$0 = -\partial_t h(t,q) + \frac{1}{2}\sigma^2\gamma q^2 + \sup_{v\in\mathcal{A}} \left(v(\eta q + \partial_q h(t,q) + \kappa v)\right). \tag{3.108}$$

This PDE in $h(t,q)$ has terminal condition:

$$h(T,q) = -\alpha q^2. \tag{3.109}$$

Now we can use the first-order condition to find an optimal trading speed v^* which satisfies the PDE (3.110) above.

First we assume that v^* is the maximiser of the supremum in the equation. Then we would have:

$$0 = -\partial_t h(t,q) + \frac{1}{2}\sigma^2\gamma q^2 + v^*(\eta q + \partial_q h(t,q) + \kappa v^*). \tag{3.110}$$

Differentiating with respect to v, then substituting v^* for v and solving, gives us the following equation for the optimal trading speed in feedback form:

$$v^* = -\frac{1}{2\kappa}\left(\eta q + \partial_q h(t,q)\right). \tag{3.111}$$

From the Riccati equation (3.75) and the definition of $\chi(t)$ in (3.76), we have the following relationship between the optimal trading speed v_t^* and the trajectory $Q_t^{v^*}$ of the inventory under this speed:

$$v_t^* = -\frac{\chi(t)Q_t^{v^*}}{\kappa}. \tag{3.112}$$

For each control $v \in \mathcal{A}$,

$$\begin{aligned} dQ_t^v &= -v_t dt \\ \implies v_t &= -\frac{dQ_t^v}{dt}. \end{aligned}$$

Choosing the optimal control v^*, we have:

$$\frac{dQ_t^{v^*}}{dt} = -\frac{\chi(t)Q_t^{v^*}}{\kappa}. \tag{3.113}$$

Separating variables in t and $Q_t^{v^*}$, and integrating over $[0,t]$ gives us:

$$\begin{aligned} \int_0^t \frac{1}{Q_u^{v^*}} dQ_u^{v^*} &= \int_0^t \frac{\chi(u)}{\kappa} du \\ \implies \ln\left(\frac{Q_t^{v^*}}{Q_0}\right) &= \int_0^t \frac{\chi(u)}{\kappa} du \\ \implies Q_t^{v^*} &= Q_0 \exp\left(\int_0^t \frac{\chi(u)}{\kappa} du\right). \end{aligned} \tag{3.114}$$

We now have an expression for the optimal inventory at any given time in terms of the trader's initial position multiplied by an accumulation factor. Now substituting the expression for $\chi(t)$ and using constants ξ and ζ (3.93, 3.96, 3.97), we obtain

$$\begin{aligned} \int_0^t \frac{\chi(u)}{\kappa} du &= \xi \int_0^t \left(\frac{1+\zeta\exp(2\xi(T-u))}{1-\zeta\exp(2\xi(T-u))}\right) du \\ &= \xi\left(\int_0^t \frac{1}{1-\zeta e^{2\xi(T-u)}} du + \int_0^t \frac{1}{\zeta e^{-2\xi(T-u)}-1} du\right) \\ &= \ln\left(\frac{\zeta e^{\xi(T-t)} - e^{-\xi(T-t)}}{\zeta e^{\xi T} - e^{-\xi T}}\right). \end{aligned} \tag{3.115}$$

From this we can find $Q_t^{v^*}$ explicitly as the initial inventory Q_0 multiplied by a deterministic function of time t elapsed from 0:

$$Q_t^{v^*} = \left(\frac{\zeta e^{\xi(T-t)} - e^{-\xi(T-t)}}{\zeta e^{\xi T} - e^{-\xi T}}\right) Q_0. \tag{3.116}$$

Finally, we have an explicit expression for the optimal trading speed as a function of the initial inventory Q_0:

$$v^* = \xi \left(\frac{\zeta e^{\xi(T-t)} + e^{-\xi(T-t)}}{\zeta e^{\xi T} - e^{-\xi T}} \right) Q_0. \tag{3.117}$$

The above expressions are identical in form to the ones found when using the quadratic running penalty function $\phi \int_0^t (Q_u^v)^2 \, du$ as used by Cartea et al [CJP15]. As Cartea et al state on page 152 of [CJP15], the function $h_2(t)$ are equal when $\phi = \frac{1}{2}\gamma\sigma^2$, and the two optimal trading strategies can be mapped isometrically to each other.

Taking limits as $\alpha \to \infty$, we see that $\zeta \to 1$. Then we have the following expressions for the optimal inventory

$$\lim_{\alpha \to \infty} Q_t^{v^*} = \left(\frac{\sinh(\xi(T-t))}{\sinh(\xi T)} \right) Q_0 \tag{3.118}$$

and trading speed

$$\lim_{\alpha \to \infty} v^* = \xi \left(\frac{\cosh(\xi(T-t))}{\sinh(\xi T)} \right) Q_0. \tag{3.119}$$

These optimal inventory and optimal trading speeds in the limit as $\alpha \to \infty$ are identical in form to the Almgren-Chriss model [AC00]. Using these explicit expressions, we can easily simulate the optimal inventory and investigate the way the trader behaves given a particular risk aversion parameter γ and terminal penalty parameter α.

3.8 Verification

We now verify that our candidate value function (3.101) and our candidate optimal control (3.102) actually solve the initial stochastic control problem (3.71). We follow the method suggested in the proof in Pham's book ([Pha09]).

Theorem 3.8.1. *Our candidate value function $\Phi(t,x,S,q)$, where $\Phi(t,x,S,q)$ replaces $H(t,x,S,q)$ in (3.101), and $\phi(t,x,S,q)$ replaces our ansatz $h(t,x,S,q)$ and our candidate optimal trading speed v_t^* in (3.102) satisfy the following conditions:*

Let $\Phi(t,x,S,q) \in \mathcal{C}^{1,2}([0,T] \text{ x } \mathbb{R}^n)$ and let $\Phi(t,x,S,q)$ satisfy a quadratic growth condition, that is for all for all $(t,x,S,q) \in [0,T] \text{ x } \mathbb{R}^n$, there exists a constant C such that:

$$|\Phi(t,x,S,q)| \leq C \left(1 + |(x,S,q)|^2 \right). \tag{3.120}$$

Suppose also that $F : \mathbb{R}_+ \to \mathbb{R}^{n+p}$ and $G : \mathbb{R}^n \to \mathbb{R}$ are measurable uniformly bounded functions (where n is the number of controlled processes and p is the number of control processes in our model (3.7)), which satisfy:

$$\frac{\partial}{\partial t}\Phi(t,x,S,q) - \sup_{v \in \mathcal{A}} \left(\mathcal{L}^v \Phi(t,x,S,q) + F(t,x,S,q,v) \right) \geq 0, \tag{3.121}$$

for all $(t, x, S, q) \in [0, T]$ x $\mathbb{R}^n$, *and*

$$\Phi(T, x, S, q) \geq G(x, s, q) \text{ for all } (x, S, q) \in \mathbb{R}^n, \tag{3.122}$$

then

$$\Phi(t, x, S, q) \geq H^v(t, x, S, q) \text{ on } [0, T] \text{ x } \mathbb{R}^n. \tag{3.123}$$

Now suppose that $\Phi(T, x, S, q) = G(x, S, q)$, *and that there exists a measurable function* $\nu^*(t, x, S, q)$ *for* $(t, x, S, q) \in [0, T]$ x $\mathbb{R}^n$, *with values in the admissible set* $\mathcal{A}$, *which satisfies:*

$$\frac{\partial}{\partial t}\Phi(t, x, S, q) - \sup_{\nu \in \mathcal{A}} \left(\mathcal{L}^\nu \Phi(t, x, S, q) + F(t, x, S, q, \nu)\right) = 0, \tag{3.124}$$

that is,

$$\frac{\partial}{\partial t}\Phi(t, x, S, q) - \left(\mathcal{L}^{\nu^*(t,x,S,q)}\Phi(t, x, S, q) + F(t, x, S, q, \nu^*(t, x, S, q))\right) = 0. \tag{3.125}$$

If the stochastic differential equation

$$dZ_t = \begin{pmatrix} dX_t \\ dS_t \\ dQ_t \end{pmatrix} = \begin{pmatrix} S_t\nu^*(t, x, S, q) - f(\nu^*(t, x, S, q))\nu^*(t, x, S, q) \\ -g(\nu^*(t, x, S, q)) \\ -\nu^*(t, x, S, q) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t \tag{3.126}$$

i.e.

$$dZ_t = \begin{pmatrix} dX_t \\ dS_t \\ dQ_t \end{pmatrix} = \begin{pmatrix} S_t\nu_t^* - \kappa\nu_t^{*2} \\ -\eta\nu_t^* \\ -\nu_t^* \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t. \tag{3.127}$$

admits a unique solution, denoted by $Z_r^{\nu^*} = (X_r^{\nu^*}, S_r^{\nu^*}, Q_r^{\nu^*})$, *with* $Z_0 = z = (x, S, q)$, *then*

$$\Phi(t, x, S, q) = H(t, x, S, q) \text{ on } [0, T] \text{ x } \mathbb{R}^n. \tag{3.128}$$

and $\nu^*(t, x, S, q)$ *is an optimal Markovian control.*

First, the candidate value function $\phi(t, x, S, q)$ in (3.101) is a continuous function of (t, x, S, q). We list the first-order partial derivatives of ϕ:

$$\partial_t\Phi = -\gamma\partial_t\phi(t, q)\Phi \tag{3.129}$$
$$\partial_x\Phi = -\gamma\Phi \tag{3.130}$$
$$\partial_S\Phi = -\gamma q\Phi \tag{3.131}$$
$$\partial_q\Phi = -\gamma\left(S + \partial_q\phi(t, q)\right)\Phi. \tag{3.132}$$

The second-order partial and mixed-partial derivatives of Φ:

$$\partial_{xx}\Phi = \gamma^2\Phi \tag{3.133}$$

$$\partial_{SS}\Phi = \gamma^2 q^2\Phi \tag{3.134}$$

$$\partial_{qq}\Phi = \left(\gamma^2(S + \partial_q\phi(t,q))^2 + \partial_{qq}\phi(t,q)\right)\Phi \tag{3.135}$$

$$\partial_{xS}\Phi = -\gamma^2 q\Phi \tag{3.136}$$

$$\partial_{xq}\Phi = -\gamma^2\left(S + \partial_q\phi(t,q)\right)\Phi \tag{3.137}$$

$$\partial_{Sq}\Phi = \gamma\left(\gamma q\left(S + \partial_q\phi(t,q)\right) - 1\right)\Phi. \tag{3.138}$$

Since the value function $\Phi(t,x,S,q)$ and the ansatz function $\phi(t,q)$ (which appear in the partial derivatives) are continuous functions, all of the required partial derivatives of Φ exist, and are continuous. Hence $\Phi(t,x,S,q)$ lies in the set $\mathcal{C}^{1,2}([0,T]\text{ x }\mathbb{R}^n)$.

In our model, we are optimising only the exponential utility of terminal wealth, and so

$$F(t,x,S,q,v) = 0 \tag{3.139}$$

$$G(x,s,q) = -\exp\left(-\gamma\left(x + q\left(S - \alpha q\right)\right)\right). \tag{3.140}$$

Both are measurable and uniformly bounded functions. Our situation is simplified relative to Pham's proof, because we do not have a running penalty/reward function. We can show immediately that the value function satisfies the required inequality (3.123) by verifying (3.121), and we do not require the local martingale argument Pham uses to deal with the running penalty $F(t,x,S,q,v)$.

Choosing a time $r \in [t,T]$, we apply Itô's lemma to our candidate value function Φ using the derivatives of Φ we calculated earlier:

$$\begin{aligned}\Phi(r,X_r,S_r,Q_r) = \Phi(t,x,S,q)&\\ + \int_t^r \frac{\partial\Phi}{\partial u}(u,X_u,S_u,Q_u) &+ \mathcal{L}^{v_u}\Phi(u,X_u,S_u,Q_u)du\\ + \int_t^r D_{x,S,q}\Phi(u,X_u,S_u,Q_u)'&\sigma(u,X_u,S_u,Q_u,v_u)dW_u,\end{aligned} \tag{3.141}$$

where $D_{x,S,q}\Phi(u,X_u,S_u,Q_u) = \left(\frac{\partial\Phi}{\partial X},\frac{\partial\Phi}{\partial S},\frac{\partial\Phi}{\partial Q}\right)$ is the Jacobian matrix of first derivatives of the candidate value function.

We take expectations of both sides, conditional on $(X_t = x,\quad S_t = S,\quad Q_t = q)$:

$$\begin{aligned}\mathbb{E}\left[\Phi(r,X_r^{t,x},S_r^{t,S},Q_r^{t,q})\right] &= \Phi(t,x,S,q)\\ &+ \mathbb{E}\left[\int_t^r \frac{\partial\Phi}{\partial u}(u,X_u^{t,x},S_u^{t,S},Q_u^{t,q}) + \mathcal{L}^{v_u}\Phi(u,X_u^{t,x},S_u^{t,S},Q_u^{t,q})du\right]\\ &+ \mathbb{E}\left[\int_t^r D_{x,S,q}\Phi(u,X_u^{t,x},S_u^{t,S},Q_u^{t,q})'\sigma(u,X_u^{t,x},S_u^{t,S},Q_u^{t,q},v_u)dW_u\right].\end{aligned} \tag{3.142}$$

We want to show that the stochastic integral part is a martingale. That will allow us to obtain the inequality (3.121) and then the equality (3.125) immediately.

We can show that the stochastic integral is a martingale by showing that its integrand is progressively measurable and square-integrable on the interval $[t,r]$. First we note that the at any time $u \in [t,r]$, the stochastic integrand equals

$$
\begin{aligned}
&D_{x,S,q}\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})' \sigma(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}, v_u) \\
&= \left(\frac{\partial \Phi}{\partial X}, \frac{\partial H}{\partial S}, \frac{\partial \Phi}{\partial Q}\right) \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} \\
&= \sigma \frac{\partial \Phi}{\partial S}(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) \\
&= -\sigma\gamma Q_u^{t,q} \Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) \\
&= \sigma\gamma Q_u^{t,q} \ldots \\
&\exp\left(-\gamma\left(X_u^{t,x} + S_u^{t,S} Q_u^{t,q} + \left(Q_u^{t,q}\right)^2 \left(\sqrt{\frac{1}{2}\kappa\gamma\sigma^2}\right)\left(\frac{1+\zeta\exp\left(2\xi(T-u)\right)}{1-\zeta\exp\left(2\xi(T-u)\right)}\right) - \frac{1}{2}\eta\left(Q_u^{t,q}\right)^2\right)\right).
\end{aligned} \tag{3.143}
$$

The integrand is progressively measurable, because it is a continuous function of the adapted state variable processes X, S, and Q.

Taking expectations of the Riemann integral of the square of the integrand, evaluated over the interval $[t,r]$ gives us:

$$
\begin{aligned}
&\mathbb{E}\left[\int_t^r \left|\sigma \frac{\partial \Phi}{\partial S}(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})\right|^2 du\right] \\
&\quad = \mathbb{E}\left[\int_t^r \left|\sigma\left(-\gamma Q_u^{t,q}\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})\right)\right|^2 du\right] \\
&\quad = \mathbb{E}\left[\int_t^r \left(\gamma^2\sigma^2\left(Q_u^{t,q}\right)^2\left(\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})\right)^2\right) du\right]
\end{aligned} \tag{3.144}
$$

The exponential function within $\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})$ is monotonically increasing over the real line $(-\infty, +\infty)$. Also the functions X_u, S_u, Q_u are driven only by the Wiener process W_u for all times $u \in [0,T]$ in our time horizon. A Wiener process is finite almost surely. Therefore the integrand in

$$
\mathbb{E}\left[\int_t^r \left(\gamma^2\sigma^2\left(Q_u^{t,q}\right)^2\left(\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})\right)^2\right) du\right] \tag{3.145}
$$

attains a finite maximum $\mathbb{P}$-almost surely at some point within the time interval

$[t, r]$. By the mean value theorem, we thus have:

$$\begin{aligned}
&\mathbb{E}\left[\int_t^r \left(\gamma^2\sigma^2 \left(Q_u^{t,q}\right)^2 \left(\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})\right)^2\right) du\right] \\
&\leq (r-t)\ \mathbb{E}\left[\sup_{u\in[t,r]} \left(\gamma^2\sigma^2 \left(Q_u^{t,q}\right)^2 \left(\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q})\right)^2\right)\right] \\
&< \infty. \qquad (3.146)
\end{aligned}$$

Since the integrand is square-integrable, we have:

$$\begin{aligned}
&\mathbb{E}\left[\Phi(r, X_r^{t,x}, S_r^{t,S}, Q_r^{t,q})\right] = \Phi(t, x, S, q) \\
&\quad + \mathbb{E}\left[\int_t^r \frac{\partial \Phi}{\partial u}(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) + \mathcal{L}^{v_u}\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) du\right].
\end{aligned}$$

We now apply the Dominated Convergence Theorem to take r to T.

$$\begin{aligned}
&\mathbb{E}\left[\Phi(T, X_T^{t,x}, S_T^{t,S}, Q_T^{t,q})\right] = \Phi(t, x, S, q) \\
&\quad + \mathbb{E}\left[\int_t^T \frac{\partial \Phi}{\partial u}(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) + \mathcal{L}^{v_u}\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) du\right], \qquad (3.147)
\end{aligned}$$

and since we assumed that $\Phi(T, x, S, q) \geq G(x, s, q)$ for all $(x, S, q) \in \mathbb{R}^n$, we have

$$\begin{aligned}
&\mathbb{E}\left[G(X_T^{t,x}, S_T^{t,S}, Q_T^{t,q})\right] \leq \Phi(t, x, S, q) \\
&\quad + \mathbb{E}\left[\int_t^T \frac{\partial \Phi}{\partial u}(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) + \mathcal{L}^{v_u}\Phi(u, X_u^{t,x}, S_u^{t,S}, Q_u^{t,q}) du\right]. \qquad (3.148)
\end{aligned}$$

Since we chose an arbitrary admissible control $v \in \mathcal{A}$, so that we can state that

$$\Phi(t, x, S, q) \geq H^v(t, x, S, q) \text{ on } [0, T] \text{ x } \mathbb{R}^n. \qquad (3.149)$$

To show that the candidate optimal control ν^* satisfies the equality (3.128), we apply Itô's formula to $\Phi(u, Z_u^{\nu^*}) = \Phi(u, X_u^{\nu^*}, S_u^{\nu^*}, Q_q^{\nu^*})$ between the times $t \in [0, T)$, and $r \in [t, T)$. Recall that the stochastic integral part of the expansion is a martingale, with mean 0. We revert to Cartea et al's notation for conditional expectation, i.e. $E_{t,Z}[.] = E[. \mid Z_t = z]$

$$\begin{aligned}
&\mathbb{E}_{t,x,S,q}\left[\Phi(r, X_r^{\nu^*}, S_r^{\nu^*}, Q_r^{\nu^*})\right] = \Phi(t, x, S, q) \\
&\quad + \mathbb{E}_{t,x,S,q}\left[\int_t^r \frac{\partial \Phi}{\partial u}(u, X_u^{\nu^*}, S_u^{\nu^*}, Q_u^{\nu^*}) + \mathcal{L}^{\nu_u}\Phi(u, X_u^{\nu^*}, S_u^{\nu^*}, Q_u^{\nu^*}) du\right]. \qquad (3.150)
\end{aligned}$$

Our assumption that the candidate optimal control satisfies the HJB equation 3.125 and the fact that we have no running penalty or reward function $F(t, x, S, q, \nu^*(t, x, S, q))$ gives us

$$\frac{\partial}{\partial t}\Phi(t,x,S,q) - \left(\mathcal{L}^{\nu^*(t,x,S,q)}\Phi(t,x,S,q) + F(t,x,S,q,\nu^*(t,x,S,q))\right) = 0$$
$$\frac{\partial}{\partial t}\Phi(t,x,S,q) - \mathcal{L}^{\nu^*(t,x,S,q)}\Phi(t,x,S,q) = 0. \tag{3.151}$$

From this we take get in the expectation above:

$$\mathbb{E}_{t,x,S,q}\left[\Phi(r, X_r^{\nu^*}, S_r^{\nu^*}, Q_r^{\nu^*})\right] = \Phi(t,x,S,q). \tag{3.152}$$

Now sending r to T in the above equation gives us

$$\mathbb{E}_{t,x,S,q}\left[\Phi(T, X_T^{\nu^*}, S_T^{\nu^*}, Q_T^{\nu^*})\right] = \Phi(t,x,S,q), \tag{3.153}$$

that is

$$G(X_T^{t,x}, S_T^{t,S}, Q_T^{t,q}) = \Phi(t,x,S,q), \tag{3.154}$$

and substituting

$$G(x,s,q) = -\exp\left(-\gamma\left(x + q\left(S - \alpha q\right)\right)\right) \text{ for } (t,x,S,q) \in [0,T] \text{ x } \mathbb{R}^3 \tag{3.155}$$

gives us the same terminal condition as for our desired value function $H(t,x,S,q)$

$$-\exp\left(-\gamma\left(X_T^{v^*} + Q_T^{v^*}\left(S_T^{v^*} - \alpha Q_T^{v^*}\right)\right)\right) = \Phi(t,x,S,q). \tag{3.156}$$

Finally, this implies that our candidate value function equals our desired value function and solves the stochastic control problem.

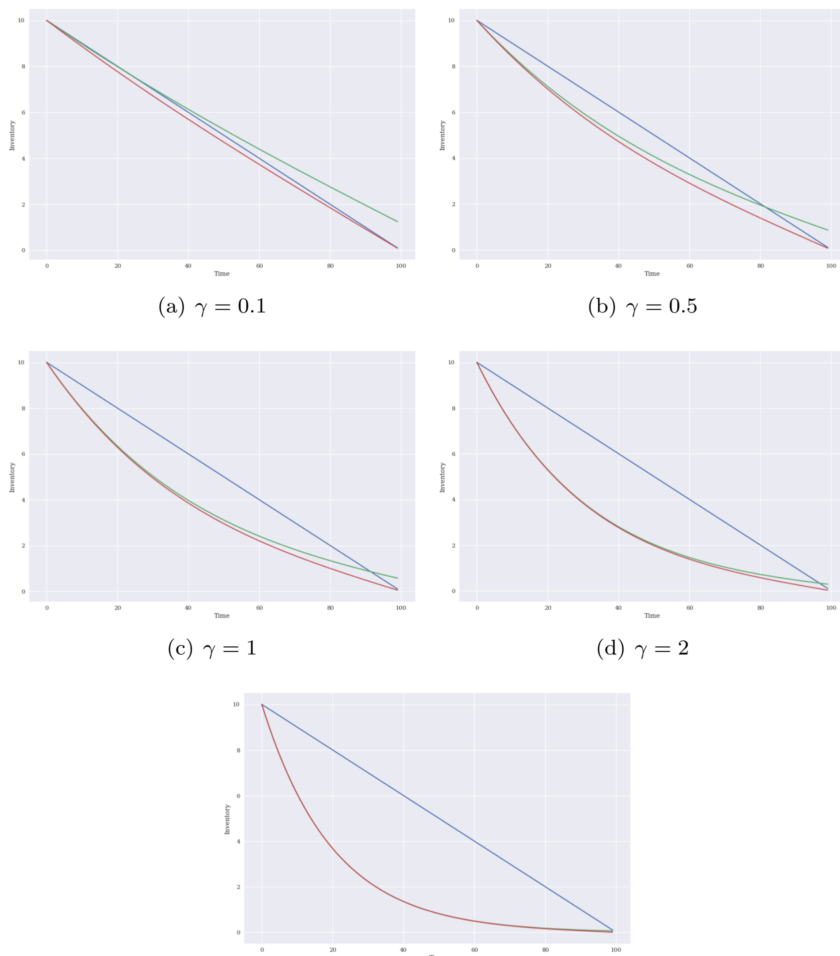
$$H(t,x,S,q,v^*) = \Phi(t,x,S,q). \tag{3.157}$$

3.8.1 Graphs of the Optimal Strategies

Here we plot graphs of the optimal strategies for a risk-averse trader liquidating a position using market orders, with the aim of maximising expected exponential utility of terminal wealth. From the graphs in figure (3.1) we can see that as the risk aversion parameter γ increases, the trader seeks to liquidate his position earlier. The trader trades more quickly earlier on to get rid of inventory and reduce market risk, but trades more slowly later on in order to reduce price impact.

When γ is small, the trader finds it optimal to hold some positive inventory Q_T (if we assume shares are infinitely divisible) at the terminal time T. This implies that the trader is prepared to incur the cost $-\alpha Q_T^2$ of executing the final market order at the end of the trading period, rather than to incur a potentially greater price impact by trading too quickly towards the end of the period.

Optimal Inventories given varying risk aversion parameter γ



(a) $\gamma = 0.1$

(b) $\gamma = 0.5$

(c) $\gamma = 1$

(d) $\gamma = 2$

(e) $\gamma - 5$

Figure 3.1: The optimal inventory Q_t^* which the trader holds at each time t, given he is trading to maximise expected exponential utility of terminal wealth, with the stated risk aversion parameter γ. The terminal penalty parameter $\alpha = 0.01$ and running penalty parameter $\phi = 0.01$. The time horizon is $T = 1$ minute. The trader wishes to liquidate an initial position of $Q_0 = 10$ shares. The temporary impact parameter $\kappa = 0.001$ and the permanent impact parameter $\eta = 0.001$. We plot three graphs on each figure: in blue, the inventory under TWAP, in green, the optimal inventory as required for the trader's risk aversion given the parameters, and in red, the limiting trajectory as $\alpha \to \infty$.

Interestingly, the "optimal" strategy sometimes trades more slowly than under TWAP as $\gamma \to 0$. In this situation the trader is relatively indifferent to the market risk of holding inventory and so is prepared to retain a positive inventory until the end of the trading horizon.

It is interesting to see that the limiting strategy as the terminal penalty $\alpha \to \infty$ is identical in shape to that under the Almgren-Chriss strategy. This limiting strategy always forces the trader to finish trading before the end of the trading period, because the penalty for having positive inventory is unbounded. This unbounded terminal penalty would be a good condition to impose in algorithms in practice in order to incentivize execution traders to complete transactions within the specified time horizon.

3.9 Power Utility

We consider the case where our trader's utility function for risk aversion is power utility, specifically what Merton [Mer69] calls the iso-elastic marginal utility function. Given the parameter γ, the utility of a variable z is given by:

$$u(z) = \begin{cases} \frac{z^{1-\gamma}-1}{1-\gamma}, & \gamma \geq 0, \quad \gamma \neq 1. \\ \ln(z), & \gamma = 1. \end{cases} \tag{3.158}$$

This utility function is a Constant Relative Risk Aversion function: that is:

$$-z\left(\frac{U''(z)}{U'(z)}\right) = \gamma. \tag{3.159}$$

This implies that the relative risk aversion does not scale with the size of the variable z. If z represents our trader's terminal wealth, then the use of this utility function implies that the trader seeks to optimise without taking into account the size of his or her wealth.

3.10 Power Utility: The Stochastic Control problem

Our trader seeks to optimise the expected iso-elastic utility function of his or her terminal wealth. He or she seeks to solve the stochastic control problem with performance criteria and value function:

$$H^v(t,x,S,q) = \mathbb{E}_{t,x,S,q}\left[U\left(X_T^v + Q_T^v\left(S_T^v - \alpha Q_T^v\right)\right)\right]; \tag{3.160}$$
$$H(t,x,S,q) = \sup_{v \in \mathcal{A}} H^v(t,x,S,q). \tag{3.161}$$

From the above stochastic model (3.7), and the utility function (3.158) above, we have the performance criteria:

$$H^v(t,x,S,q) = \begin{cases} \mathbb{E}_{(t,x,S,q)}\left[\frac{\left(X_T + S_T Q_T - \alpha Q_T^2\right)^{1-\gamma} - 1}{1-\gamma}\right], & \gamma \geq 0, \quad \gamma \neq 1. \\ \mathbb{E}_{(t,x,S,q)}\left[\ln\left(X_T + S_T Q_T - \alpha Q_T^2\right)\right], & \gamma = 1. \end{cases} \tag{3.162}$$

The value function is:

$$H(t,x,S,q) = \sup_{v \in \mathcal{A}} H^v(t,x,S,q) \tag{3.163}$$

Note that this iso-elastic marginal utility function prevents our trader from making a loss on the final trade: he cannot possibly execute the last trade if his terminal wealth does not satisfy the following condition after a potential market order:

$$\alpha < \frac{X_T + S_T Q_T}{Q_T^2}. \tag{3.164}$$

The Dynamic Programming Principle (2.1.1) implies that for all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$, and all stopping times $\tau \leq T$,

$$H(t, x, S, q) = \sup_{v \in \mathcal{A}_{(t,x)}} \mathbb{E}_{t,x,S,q} \left[H(\tau, X_\tau^v, S_\tau^v, Q_\tau^v) \right]. \tag{3.165}$$

Then if $v^* \in \mathcal{A}$ is the optimal control, which achieves the supremum over the interval $[0, T]$, the value function satisfies

$$H(t, x, S, q) = \mathbb{E}_{t,x,S,q} \left[H(T, X_T^{v^*}, S_T^{v^*}, Q_T^{v^*}) \right]. \tag{3.166}$$

Using the Dynamic Programming Principle (2.1.1), we obtain the Hamilton-Jacobi-Bellman Equation for the value function (writing $\partial_\bullet H for \frac{\partial H}{\partial \bullet}$):

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \sup_{v \in \mathcal{A}} \left(Sv\partial_x H - f(v)v\partial_x H - g(v)\partial_S H - v\partial_q H \right) = 0, \tag{3.167}$$

We assume that the temporary and permanent impact functions are linear functions of the speed of trading, i.e. $f(v_t) = \kappa v_t$ and $g(v_t) = \eta v_t$, respectively. We then get the HJB equation:

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \sup_{v \in \mathcal{A}} \big(Sv\partial_x H - \kappa v^2 \partial_x H - \eta v \partial_S H - v\partial_q H \big) = 0. \tag{3.168}$$

The terminal condition of the HJB PDE is:

$$H(T, x, S, q) = \begin{cases} \frac{\left(x + Sq - \alpha q^2\right)^{1-\gamma} - 1}{1-\gamma}, & \gamma \geq 0, \quad \gamma \neq 1. \\ \ln\left(x + Sq - \alpha q^2\right), & \gamma = 1. \end{cases} \tag{3.169}$$

3.11 Verification

We want to find a candidate value function and a candidate optimal control that solves the stochastic control problem (3.160).

We apply the verification theorem to our own stochastic control problem, given the model (3.47) above. Let $\Phi(t, x, S, q)$ be our candidate value function. Applying Itô's Lemma to the candidate value function gives us:

$$d\Phi = \frac{\partial \Phi}{\partial t} dt + \frac{\partial \Phi}{\partial x} dx + \frac{\partial \Phi}{\partial S} dS + \frac{\partial \Phi}{\partial q} dq + \frac{1}{2} \frac{\partial^2 \Phi}{\partial S^2} \sigma^2 dt. \tag{3.170}$$

We know from the model (3.7) the terms dx, dS, dq, and that the mixed 2^{nd}-order partial derivatives vanish apart from for $(dS)^2$, which being driven by the Brownian motion W has quadratic variation T and hence $< dS, dS >= \sigma^2 dt$.

We have

$$d\Phi = \left(\frac{\partial \Phi}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi}{\partial S^2}\sigma^2\right) dt + \frac{\partial \Phi}{\partial x}\big((S_t - f(v_t))\, v_t dt\big) + \frac{\partial \Phi}{\partial S}\big(-g(v_t)dt + \sigma dW_t\big) + \frac{\partial \Phi}{\partial q}\left(-v_t dt\right). \tag{3.171}$$

Collecting terms in dt and dW_t, we obtain:

$$d\Phi(t, X_t, S_t, Q_t) = \left(\frac{\partial \Phi}{\partial t} + \big(S_t v_t - f(v_t)v_t\big)\frac{\partial \Phi}{\partial x} - g(v_t)\frac{\partial \Phi}{\partial S} - v_t\frac{\partial \Phi}{\partial q} + \frac{1}{2}\sigma^2\frac{\partial^2 \Phi}{\partial S^2}\right) dt + \sigma\frac{\partial \Phi}{\partial S} dW_t. \tag{3.172}$$

From this we can write the candidate value function in full integral form. At any time $t \in [0, T]$ and for an arbitrary control v in the admissible set $\mathcal{A}$:

$$\Phi(t, X_t, S_t, Q_t) = \Phi(0, X_0, S_0, Q_0) + \int_0^t \left(\frac{\partial \Phi}{\partial u} + \big(S_u v_u - f(v_u)v_u\big)\frac{\partial \Phi}{\partial x} - g(v_u)\frac{\partial \Phi}{\partial S} - v_u\frac{\partial \Phi}{\partial q} + \frac{1}{2}\sigma^2\frac{\partial^2 \Phi}{\partial S^2}\right) du + \int_0^t \sigma\frac{\partial \Phi}{\partial S} dW_u. \tag{3.173}$$

The infinitesimal generator of the performance criteria for this control problem:

$$\mathcal{L}_t^v \Phi(t, X_t, S_t, Q_t) = \big(Sv - f(v)v\big)\frac{\partial \Phi}{\partial X} - g(v)\frac{\partial \Phi}{\partial S} - v\frac{\partial \Phi}{\partial q} + \frac{1}{2}\sigma^2\frac{\partial^2 \Phi}{\partial S^2}. \tag{3.174}$$

Substituting this into the expression for the performance criteria we obtain the simpler representation

$$\Phi(t, X_t, S_t, Q_t) = \Phi(0, X_0, S_0, Q_0) + \int_0^t \left(\frac{\partial \Phi}{\partial u} + \mathcal{L}_u \Phi\right) du + \int_0^t \sigma\frac{\partial \Phi}{\partial S} dW_u. \tag{3.175}$$

Since $\Phi \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^3)$, we can use Itô's lemma to write for all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$ and all controls which are admissible over $[t, T]$ given that $X_t^v = x$, $S_t^v = S$, $Q_t^v = q$, ie. all $v \in \mathcal{A}_{(t,x,S,q)}$, any time $r \in [t, T)$, and any stopping time $\tau \in [t, T]$,

$$\mathbb{E}_{t,x,S,q}\left[\Phi(r \wedge \tau, X_{r\wedge\tau}, S_{r\wedge\tau}, Q_{r\wedge\tau})\right] = \Phi(t, X_t, S_t, Q_t) + \mathbb{E}_{t,x,S,q}\left[\int_t^{r\wedge\tau} \left(\frac{\partial \Phi}{\partial u} + \mathcal{L}_u \Phi\right) du\right] + \mathbb{E}_{t,x,S,q}\left[\int_t^{r\wedge\tau} \sigma\frac{\partial \Phi}{\partial S} dW_u\right]. \tag{3.176}$$

We choose the stopping time $\tau := \tau_n \in [t, T]$ to be such that

$$\tau_n := \inf \left\{ r \geq t \text{ such that } \int_t^r \left| \sigma \frac{\partial \Phi}{\partial S} \right|^2 du \geq n \right\}. \tag{3.177}$$

Then we can define $\{\tau_n\}_{n \in \mathbb{N}}$ as a localizing sequence of stopping times such that the stopped stochastic integral is a martingale. Then

$$\begin{aligned} &\mathbb{E}_{t,x,S,q} \left[\Phi(r \wedge \tau, X_{r \wedge \tau}, S_{r \wedge \tau}, Q_{r \wedge \tau}) \right] \\ &= \Phi(t, X_t, S_t, Q_t) + \mathbb{E}_{t,x,S,q} \left[\int_t^{r \wedge \tau} \left(\frac{\partial \Phi}{\partial u} + \mathcal{L}_u \Phi \right) du \right]. \end{aligned} \tag{3.178}$$

Now since we have assumed the function Φ satisfies the inequality (**??**) for the particular control that achieves the supremum, we have the same for all admissible controls $v \in \mathcal{A}_{(t,x,S,q)}$:

$$\frac{\partial \Phi}{\partial t} \left(u, X_u^v, S_u^v, Q_u^v \right) + \mathcal{L}^{v_u} \Phi \left(u, X_u^v, S_u^v, Q_u^v \right) \leq -F \left(u, X_u^v, S_u^v, Q_u^v, v_u \right) \tag{3.179}$$

In our current problem, $F \left(u, X_u^v, S_u^v, Q_u^v, v_u \right) = 0$, since there is no running penalty or reward function. The terminal function $G \left(T, X_T^v, S_T^v, Q_T^v, v_T \right) = U \left(X_T^v, S_T^v, Q_T^v \right)$. Thus for any time $r \in [t, T]$,

$$\mathbb{E}_{t,x,S,q} \left[\Phi(r, X_r, S_r, Q_r) \right] \leq \Phi(t, X_t, S_t, Q_t). \tag{3.180}$$

Using the Dominated Convergence Theorem as $r \to T$,

$$\mathbb{E}_{t,x,S,q} \left[\Phi(T, X_T, S_T, Q_T) \right] \leq \Phi(t, X_t, S_t, Q_t). \tag{3.181}$$

And since we assumed the terminal condition $\Phi(T, x, S, q) \geq G(x, S, q)$ for all $(x, S, q) \in \mathbb{R}^3$,

$$\mathbb{E}_{t,x,S,q} \left[G \left(T, X_T^v, S_T^v, Q_T^v, v_T \right) \right] \leq \Phi(t, X_t, S_t, Q_t). \tag{3.182}$$

Now since our choice of admissible control was arbitrary, we can say for all admissible controls $\alpha \in \mathcal{A}_{(t,x,S,q)}$ and all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$,

$$\begin{aligned} &\Phi(t, x, S, q) \geq \mathbb{E}_{t,x,S,q} \left[U \left(X_T^v, S_T^v, Q_T^v \right) \right] \\ \Longrightarrow\ &\Phi(t, x, S, q) \geq H^v(t, x, S, q). \end{aligned} \tag{3.183}$$

Similar arguments give us the following equality for the optimal control v^*: for all $(t, x, S, q) \in [0, T] \times \mathbb{R}^3$,

$$\begin{aligned} &\Phi(t, x, S, q) \geq \mathbb{E}_{t,x,S,q} \left[U \left(X_T^{v^*}, S_T^{v^*}, Q_T^{v^*} \right) \right] \\ \Longrightarrow\ &\Phi(t, x, S, q) = H(t, x, S, q). \end{aligned} \tag{3.184}$$

3.12 Solving the Stochastic Control Problem

3.12.1 Feedback form for the Optimal Control

We aim to solve the above HJB equation (3.168). First, we find the optimal control v^* in feedback form. Substituting that feedback form back into the HJB equation will yield us a (possibly non-linear) partial differential equation for the value function $H(t,x,S,q)$. When that PDE in H is solved, it will give us the optimized value function $H(t,x,S,q) = H^{v^*}(t,x,S,q)$ at each time $t \in [0,T]$. From this, we will attempt to find an explicit expression for the optimal control v^*.

To solve the HJB equation (3.168), first, we assume that v^* is the trading speed which achieves the supremum in the equation. Then

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \Big(S v^* \partial_x H - \kappa {v^*}^2 \partial_x H - \eta v^* \partial_S H - v^* \partial_q H \Big) = 0. \quad (3.185)$$

The equation is a quadratic in v^*.

$$0 = \partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \left(S\partial_x H - \eta \partial_S H - \partial_q H \right) v^* - \left(\kappa \partial_x H \right) {v^*}^2. \quad (3.186)$$

Completing the square gives us:

$$\begin{aligned} 0 = & -\kappa \partial_x H \left(v^* - \frac{(S\partial_x H - \eta \partial_S H - \partial_q H)}{\partial_x H} \right)^2 \\ & + \partial_t H + \frac{1}{2}\partial_{SS} H - \frac{(S\partial_x H - \eta \partial_S H - \partial_q H)^2}{-4\kappa \partial_x H}. \end{aligned} \quad (3.187)$$

Thus the value of v^* which satisfies the HJB equation is:

$$v^* = \frac{1}{2\kappa} \frac{(S\partial_x H - \eta \partial_S H - \partial_q H)}{\partial_x H}. \quad (3.188)$$

Substituting the optimal control in feedback form into the HJB equation gives us the non-linear PDE:

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \frac{1}{4\kappa} \frac{(S\partial_x H - \eta \partial_S H - \partial_q H)^2}{\partial_x H} = 0. \quad (3.189)$$

The terminal condition of this PDE is:

$$H(T,x,S,q) = \begin{cases} \frac{\left(x + Sq - \alpha q^2\right)^{1-\gamma} - 1}{1-\gamma}, & \gamma \geq 0, \quad \gamma \neq 1. \\ \ln\left(x + Sq - \alpha q^2\right), & \gamma = 1. \end{cases} \quad (3.190)$$

We now need to find the candidate value function $H(t, X_t, S_t, Q_t)$ which solves this PDE.

Using the ansatz (3.54) and the constant relative risk aversion under power utility (3.159), we obtain from (3.62) the equation

$$0 = \partial_t h(t,q) + \frac{1}{2}\sigma^2 q^2 \left(\frac{\gamma}{x + Sq + h(t,q)} \right) + \frac{1}{4\kappa} \Big(\eta q + \partial_q h(t,q) \Big)^2. \quad (3.191)$$

This is no longer a Riccati equation as in the exponential case, but is a more complicated non-linear PDE, due to the presence of $h(t, q)$ in the denominator of the diffusion term.

3.12.2 Is a solution possible?

It might be the case that this equation is not solvable using the ansatz (3.54), and so a different ansatz or indeed an entirely different way of solving the problem might be required. In Section 9.5 of Cartea et al [CJP15], the authors try to solve a problem involving the maximisation of expected terminal exponential utility of terminal wealth while targeting a percentage of volume (PoV) strategy. In this case the authors find that one cannot add a linear penalty function to the problem because "the exponential utility and the linear penalty are in a sense incompatible, and ... the cash process does not factor out of the problem".

This is a similar problem to the one we have encountered while incorporating a quadratic terminal penalty function while trying to maximise expected power utility of terminal wealth. We have a PDE where the cash process and the asset midprice process remain in the problem. This casts doubt on the validity of the ansatz: there might not be a way to separate the value function $H(t, x, S, q)$ into the sum $x + Sq + h(t, q)$ because the function $h(t, q)$ might not exist.

It might be necessary to get rid of the penalty, or to try the technique that Cartea, Jaimungal and Penalva use for the aforementioned problem. They express the value function H_t as the continuous limit of a recursion, with H_T being the value function at time T. We do not attempt this here.

Chapter 4

Optimal Execution with Limit Orders

4.1 Introduction

In this chapter, we look at optimal execution of trades for a trader who uses limit orders rather than market orders. The trader posts orders to buy in an acquisition programme, or to sell in a liquidation programme, a specific quantity of an asset at a specified limit price. In doing this, the trader guarantees that his trade is executed at or better than the limit price, at the risk of the trade not being filled if the asset's market price does not reach the limit price.

4.2 The Stochastic Model

We use the model from Section 8.2 of Cartea et al [CJP15]. Our trader wants to liquidate a position of $\mathcal{Q}_0$ shares over the finite time horizon $[0, T]$.

4.2.1 Asset price

The share's midprice $S = \{S_t\}_{0 \leq t \leq T}$ obeys the following dynamics:

$$S_t = S_0 + \sigma W_t, \ \sigma > 0, \tag{4.1}$$

where $W = \{W_t\}_{0 \leq t \leq T}$ is a standard Brownian motion.

4.2.2 Order Depth

The trader aims to control the price level at which he places limit orders such that he maximises the expected value of his total revenue. As shown in Cartea et al [CJP15], Avellaneda and Stoikov [AS08], we can represent this price level via the "depth" in the limit order book, measured from the best bid or best ask price. Since we model the asset price using its midprice, we measure this depth from the midprice via the stochastic process $\delta = \{\delta\}_{0\leq t\leq T}$. At time t, our trader posts limit orders to sell at the price level $S_t + \delta_t$.

4.2.3 Market Orders and Fill Probability

Cartea et al [CJP15] model incoming market orders from other traders as a Poisson process $M = \{M_t\}_{0\leq t\leq T}$ with rate λ. Thus per unit time, we expect to see λ market orders sent to the exchange. Not every incoming market order will be filled by our trader's posted limit order. We assume that each buy market order of a price at or above our trader's sell limit order price level of $S_t + \delta_t$ matches with our trader's limit order and thus leads to a sale. We model the fill probability to follow an exponential distribution with rate κ:

$$P(\delta) = \exp(-\kappa\delta). \tag{4.2}$$

Those incoming market orders which match with our trader's posted limit orders are modelled by a counting process $N^\delta = \{N_t^\delta\}_{0\leq t\leq T}$.

Multiplying the fill probability $P(\delta)$ with the intensity λ for the incoming market orders M gives us a controlled stochastic process for the rate of filled market orders $\lambda^\delta = \{\lambda_t^\delta\}_{0\leq t\leq T}$, which satisfies:

$$\lambda_t^\delta = \lambda \exp\left(-\kappa\delta_t\right). \tag{4.3}$$

4.2.4 Cash process

We model the trader's cash process by the stochastic process $X^\delta = \{X_t^\delta\}_{0\leq t\leq T}$, which satisfies the SDE

$$dX_t^\delta = \left(S_t + \delta_t\right) dN_t^\delta. \tag{4.4}$$

4.2.5 Inventory

We model the trader's position in the stock by the stochastic process $Q^\delta = \{Q_t^\delta\}_{0\leq t\leq T}$, which satisfies $Q_t^\delta = \mathcal{Q}_0 - N_t^\delta$.

4.2.6 The full model

We formulate the problem as a 3-dimensional system with three state variables: the total wealth X_t^δ, the asset midprice S_t^δ, and the inventory Q_t^δ.

$$dZ_t^\delta = \begin{pmatrix} dX_t^\delta \\ dS_t^\delta \\ dQ_t^\delta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t + \begin{pmatrix} S_t + \delta_t \\ 0 \\ -1 \end{pmatrix} dN_t^\delta. \tag{4.5}$$

In this way we have a (3×1) vector of drifts $\mu(t, x, S, q, \delta)$, a (3×1) matrix of volatilites $\sigma(t, x, S, q, \delta)$ which are both Lipschitz-continuous and integrable, and a (3×1) matrix of counting processes $\gamma(t, x, S, q, \delta)$ which satisfy:

$$\mu(t, x, S, q, \delta) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{4.6}$$

and

$$\sigma(t, x, S, q, \delta) = \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} \tag{4.7}$$

and

$$\gamma(t, x, S, q, \delta) = \begin{pmatrix} S_t + \delta_t \\ 0 \\ -1 \end{pmatrix}. \tag{4.8}$$

4.3 The Optimal Control Problem

The trader wishes to maximise his profit from selling all $\mathcal{Q}_0$ shares during the interval $[0, T]$. We define the stopping time τ as the minimum of the final time T and the first time that the inventory hits zero:

$$\tau = T \wedge \min \left\{ t : Q_t^\delta = 0 \right\}. \tag{4.9}$$

If the trader does not liquidate his entire position by time T, then he incurs a penalty of αQ_T^δ per each unit of inventory to liquidate the remaining inventory via a market order (essentially he receives the price $\left(S_T - \alpha Q_T^\delta\right)$ for each unit in his inventory Q_T^δ). We also add a running penalty function which represents our trader's urgency to get rid of inventory to close his entire position. The function penalises positive (or negative) inventory held up until each stopping time τ: we subtract from the total revenue $-\phi \int_t^\tau \left(Q_t^\delta\right)^2 du$, for $\phi \geq 0$.

The trader's optimisation problem is to find the optimal depth δ_t above the best ask at which to post his sell limit orders at each time t so that his revenue is maximised when he stops trading at time τ:

$$H(t, x, S, q) = \sup_{\delta \in \mathcal{A}} \mathbb{E}_{t,x,S,q} \left[X_\tau^\delta + Q_\tau^\delta \left(S_\tau - \alpha Q_\tau^\delta\right) - \phi \int_t^\tau \left(Q_t^\delta\right)^2 du \right]. \tag{4.10}$$

From Cartea et al [CJP15], given a jump diffusion model

$$dX_t^\alpha = \mu\left(t, X_t^\alpha, \alpha_t\right) dt + \sigma\left(t, X_t^\alpha, \alpha_t\right) dW_t + \gamma\left(t, X_t^\alpha, \alpha_t\right) dN_t^\alpha, \quad X_0^\alpha = x, \tag{4.11}$$

where the stochastic process $W = \left(W_t^\alpha\right)_{0 \leq t \leq T}$ is a d-dimensional vector of independent Brownian motions on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(0 \leq t \leq T)}, \mathbb{P})$. The p-dimensional vector $\alpha = \left(\alpha_t\right)_{0 \leq t \leq T}$ represents the control processes, which are progressively measurable with respect to $\left(\mathcal{F}\right)_{0 \leq t \leq T}$, and are valued in a subset $A \subset R^p$. The process $\mu\left(t, X_t^\alpha, \alpha_t\right)$ is an $(d \times 1)$-dimensional vector of drifts, $\sigma\left(t, X_t^\alpha, \alpha_t\right)$ is a $(p \times p)$-dimensional matrix of volatilities, and $\gamma\left(t, X_t^\alpha, \alpha_t\right)$ is a

$(p \times d)$-dimensional matrix of jumps in the counting process N_t^α. Another process $\lambda(t, X_t^\alpha, \alpha)$ represents the controlled intensities of the counting processes.

We have the following HJB equation

$$\partial_t V(t,x) + \sup_{\alpha \in \mathcal{A}} \left(\mathcal{L}_t^\alpha V(t,x) + f(t,x,\alpha) \right) = 0, \tag{4.12}$$

$$V(T,x) = g(x). \tag{4.13}$$

where the infinitesimal generator acts as follows:

$$\begin{aligned} \mathcal{L}_t^\alpha V(t,x) &= \mu(t,x,\alpha)^{'}\ \mathcal{D}V(t,x) + \frac{1}{2}\mathbf{Tr}\left(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^{'}\ \mathcal{D}^2 V(t,x)\right) \\ &\quad + \sum_{j=1}^{p} \lambda_{\bullet,j}\left(t,x,\alpha\right)\left[V\left(t, x + \gamma_{\bullet,j}\left(t,x,\alpha\right)\right) - V(t,x)\right], \end{aligned} \tag{4.14}$$

where for any vector or matrix A, $A^{'}$ denotes its transpose, $\mathcal{D}\phi(x)$ denotes the n-dimensional column vector of first partial derivatives of $\phi(x)$, and $\mathcal{D}^2\phi(x)$ denotes the $(m \times n)$-dimensional matrix of second partial derivatives of $\phi(x)$, and $A_{\bullet,j}$ denotes the vector corresponding to the j^{th} column of the matrix A.

Given our model 4.5, this gives us the HJB equation

$$\begin{aligned} 0 &= \partial_t H(t,x,S,q) + \frac{1}{2}\sigma^2 \partial_{SS} H(t,x,S,q) - \phi q^2 \\ &\quad + \sup_{\delta \in \mathcal{A}} \left\{ \lambda \exp\left(-\kappa\delta\right)\left[H(t, x+S+\delta, S, q-1) - H(t,x,S,q)\right] \right\}, \end{aligned} \tag{4.15}$$

with boundary and terminal conditions

$$H(t,x,S,0) = 0, \tag{4.16}$$

$$H(T,x,S,q) = x + Sq - \alpha q^2. \tag{4.17}$$

4.4 Solving the HJB equation

4.4.1 An Ansatz for the Value Function

Cartea, Jaimungal, and Penalva introduce the ansatz

$$H(t,x,S,q) = x + Sq + h(t,q). \tag{4.18}$$

The value function at each time t is then a function of the cash x, proceeds Sq from selling a quantity q of the shares at price S, plus some function $h(t,q)$ of time and inventory which represents the value added due to optimal execution of the shares at time t.

Substituting this ansatz into the HJB equation (4.15) gives us:

$$\begin{aligned} 0 &= \partial_t h(t,q) + 0 - \phi q^2 \\ &\quad + \sup_{\delta \in \mathcal{A}} \left\{ \lambda e^{-\kappa\delta}\left((x + S + \delta + S(q-1) + h(t,q-1)) - (x + Sq + h(t,q)) \right) \right\} \end{aligned} \tag{4.19}$$

which becomes the simpler PDE in terms of time t and inventory q only:

$$\partial_t h(t,q) - \phi q^2 + \sup_{\delta \in \mathcal{A}} \left\{ \lambda e^{-\kappa\delta} \left(\delta + h(t,q-1) - h(t,q)\right) \right\} = 0, \tag{4.20}$$

$$h(t,0) = 0, \tag{4.21}$$

$$h(T,q) = -\alpha q^2. \tag{4.22}$$

4.4.2 The optimal depth δ^* in feedback form

Now we can find the optimal control in feedback form by assuming that δ^* is the value of δ which maximises the expression within the supremum, differentiating with respect to δ^*, equating the resulting expression to zero, and then solving for δ^*. As long as the second derivative of the function h with respect to δ^* is negative, we have a maximum.

$$0 = \partial_t h(t,q) - \phi q^2 + \lambda e^{-\kappa\delta^*} \left(\delta^* + h(t,q-1) - h(t,q)\right). \tag{4.23}$$

Differentiating with respect to δ^* gives us

$$-\kappa\lambda e^{-\kappa\delta^*} \left(\delta^* + h(t,q-1) - h(t,q)\right) + \lambda e^{-\kappa\delta^*} = 0. \tag{4.24}$$

Solving for δ^* yields the optimal control in feedback form:

$$\delta^*(t,q) = \frac{1}{\kappa} - \left(h(t,q-1) - h(t,q)\right). \tag{4.25}$$

4.4.3 Simplifying the PDE

We substitute the feedback form (4.25)into the simplified PDE (4.20) to get:

$$\partial_t h(t,q) - \phi q^2 + \frac{\lambda}{\kappa e} \exp(-\kappa \left(h(t,q) - h(t,q-1)\right) = 0. \tag{4.26}$$

$$h(t,0) = 0, \tag{4.27}$$

$$h(T,q) = -\alpha q^2. \tag{4.28}$$

We now follow Cartea et al [CJP15] and introduce the substitution

$$h(t,q) = \frac{1}{\kappa} \log \omega(t,q). \tag{4.29}$$

Then

$$\omega(t,q) = \exp\left(\kappa h(t,q)\right). \tag{4.30}$$

and

$$\partial_t h(t,q) = \frac{1}{\kappa} \frac{\partial_t \omega(t,q)}{\omega(t,q)}. \tag{4.31}$$

Our PDE then becomes

$$\partial_t \omega(t,q) + \frac{\lambda}{e} \omega(t,q-1) - \phi q^2 \kappa \omega(t,q) = 0. \tag{4.32}$$

$$\omega(t,0) = 1, \tag{4.33}$$

$$\omega(T,q) = \exp(-\kappa\alpha q^2). \tag{4.34}$$

We now aim to solve this equation.

4.4.4 A recursive equation for $\omega(t, q)$

We can find a recursive equation for $\omega(t, q)$ in terms of $\omega(t, q-1)$ by treating $\omega(t, q-1)$ as a separate variable and integrating. First, we use Cartea et al's [CJP15] notation of $\tilde{\lambda} = \frac{\lambda}{e}$. Our equation becomes

$$\partial_t \omega(t, q) = \kappa \phi q^2 \omega(t, q) - \tilde{\lambda} \omega(t, q-1).$$

We treat $\omega(t, q-1)$ as constant. Separating the variables and integrating over the trading period $[t, T]$:

$$\int_t^T \frac{\partial_u \omega(u, q)}{\kappa \phi q^2 \omega(u, q) - \tilde{\lambda} \omega(t, q-1)} = \int_t^T du$$

gives us

$$\frac{1}{\kappa \phi q^2} \log \left(\frac{\kappa \phi q^2 \omega(T, q) - \tilde{\lambda} \omega(t, q-1)}{\kappa \phi q^2 \omega(t, q) - \tilde{\lambda} \omega(t, q-1)} \right) = T - t.$$

Taking exponents gives us:

$$\frac{\kappa \phi q^2 \omega(T, q) - \tilde{\lambda} \omega(t, q-1)}{\kappa \phi q^2 \omega(t, q) - \tilde{\lambda} \omega(t, q-1)} = \exp\left(\kappa \phi q^2 (T - t) \right).$$

Finally we have a recursive equation for $\omega(t, q)$ in terms of $\omega(t, q-1)$,

$$\omega(t, q) = \frac{\kappa \phi q^2 \omega(T, q) - \tilde{\lambda} \omega(t, q-1) \left(1 - \exp\left(\kappa \phi q^2 (T - t)\right)\right)}{\kappa \phi q^2 \exp\left(\kappa \phi q^2 (T - t)\right)}.$$

Simplifying gives us:

$$\begin{aligned} \omega(t, q) = &\exp\left(-\kappa \phi q^2 (T - t)\right) \omega(T, q) \\ &+ \frac{\tilde{\lambda}}{\kappa \phi q^2} \left(1 - \exp\left(-\kappa \phi q^2 (T - t)\right)\right) \omega(t, q-1). \end{aligned} \tag{4.35}$$

Substituting the terminal conditions into this formula, we obtain

$$\begin{aligned} \omega(t, q) = &\exp\left(-\kappa q^2 \left(\alpha + \phi(T - t)\right)\right) \\ &+ \frac{\tilde{\lambda}}{\kappa \phi q^2} \left(1 - \exp\left(-\kappa \phi q^2 (T - t)\right)\right) \omega(t, q-1). \end{aligned} \tag{4.36}$$

Given that we have the boundary condition $\omega(t, 0) = 0$ when $q = 0$, at each time $t \in [0, T]$, given a position of size Q_t, we recursively calculate $\omega(t, q)$ for $q \in \{0, 1, ..., Q_t\}$.

4.5 Simulations and Graphs

In B.2, we present Python code which calculates and graphs the optimal depths using the recursive formula (4.36). We present examples and explain the consequences of them for the trader.

4.5.1 Optimal Depths given varying penalty parameters

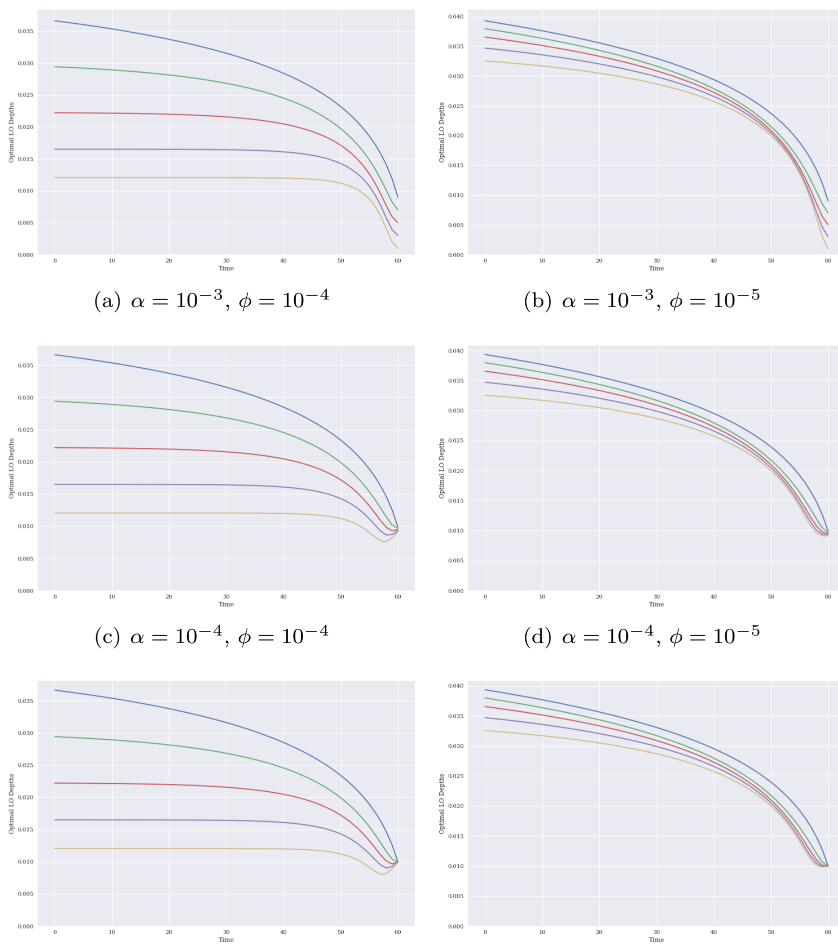
We use the same parameters as Cartea et al do in Section 8.2 of their book [CJP15], which looks at the same execution problem without a running penalty $\phi \int_t^T (Q_u)^2 du$ penalising inventories throughout the trading period.

The trader is looking to liquidate a position of $Q_0 = 5$ shares within a minute. The arrival rate of incoming market orders is $\lambda = 50$ per minute. Therefore, the trader is attempting to liquidate a position whose size is a tenth of the volume expected within the time horizon. The fill probability is exponentially distributed with parameter $\kappa = 100\$^{-1}$.

From the graphs in Figure 4.1, we see that the optimal strategy instructs the trader to post limit orders at depths which increase as the trader's position size decreases. The effect of the terminal penalty parameter α on the trajectory of the optimal depths $\delta^*(t, q)$ is subtle. As α increases, the graphs of the optimal depths flatten when further out from the end of the time horizon, but the overall depths themselves are marginally greater closer to the end of the time horizon T. The trader is trying to ensure that any sales made via sell limit orders towards the end of the time horizon are executed at a price $S_t + \delta_t$ that is sufficiently better than the midprice that those sales will cover the penalty $-\alpha Q_T^2$ incurred from being forced to liquidate any inventory Q_T held using a market order.

The effect of the running penalty parameter ϕ on the trader's optimal limit order execution strategy is more significant. The running penalty penalises all inventories greater than 0 throughout the trading period. Thus the trader is much more urgent to liquidate his position as quickly as possible. The trader posts limit orders which are much closer to the midprice in order to maximise the probability of being filled on his order. The optimal depths are asymptotically constant further away from the end of the time horizon, where the effect of the running penalty dominates that of the terminal penalty. As the end of the trading period looms, the trader narrows his spread in order to increase the probability of liquidating before the end of the period. For larger inventories, the trader's spread is actually humped lower than the limiting value and rises back up to maximise the probability that any sale covers the costs from the terminal market order.

Optimal Depths given varying penalty parameters



(a) $\alpha = 10^{-3}$, $\phi = 10^{-4}$ (b) $\alpha = 10^{-3}$, $\phi = 10^{-5}$

(c) $\alpha = 10^{-4}$, $\phi = 10^{-4}$ (d) $\alpha = 10^{-4}$, $\phi = 10^{-5}$

(e) $\alpha = 10^{-5}$, $\phi = 10^{-4}$ (f) $\alpha = 10^{-5}$, $\phi = 10^{-5}$

Figure 4.1: The optimal depths $\delta^*(t, q)$ at which the trader posts limit orders given an inventory q at time t, for the stated terminal penalty parameter α and running penalty parameter ϕ. The time horizon is $T = 1$ minute. The trader wishes to liquidate an initial position of $Q_0 = 5$ units of the asset. The fill probability parameter is $\kappa = 100\$^{-1}$. Incoming market orders arrive at rate $\lambda = 50$ per minute.

Conclusions and Model Improvements

We have used stochastic control to solve trade execution problems involving optimal trading speed v_t^* (and holdings Q_t^* through time) of trades using market orders, provided a general partial differential equation (3.62) and also found the optimal limit order posting depths $\delta^*(t,q)$ for a trader using limit orders to liquidate an inventory of an asset.

Risk aversion was incorporated into these problems using exponential and power utility for the market order (MO) problems and a running penalty function penalising non-zero inventories in the limit order (LO) problem. We have graphed the optimal solutions for the MO problem in Figure (3.1) and the LO problem in Figure (4.1).

In the MO problem, we have seen that an increase in risk aversion via the risk aversion coefficient greatly impacts the resulting optimal behaviour of the trader. Lower risk aversion leads to more risky behaviour, and the trader willing to take on market risk by holding a positive inventory closer to the end of the trading period, even to the point of incurring extra costs due to a forced market order at the end of the period.

In the LO problem, the running penalty function has a similarly impact on the optimal behaviour of the trader. Greater risk aversion leads to the trader posting much closer to the midprice for larger inventories, in order to liquidate the position as quickly as possible.

At this point we will discuss existing and potential improvements to the models. In the LO model, we did not account for the fact that the optimal depth might become negative. In fact, for higher levels of ϕ in the running penalty function, $\delta^*(t,q)$ becomes negative. A trader cannot post limit orders at negative depths in the limit order book. Cartea, Jaimungal, and Penalva ([CJP15]) make the same point in their own model in Section 8.2 (which does not involve a running penalty but also similarly does not account for negative values of δ^*). One could take negative depths to imply that the trader's urgency to liquidate at least one unit of asset is immediate; they want to get rid of at least one unit of the asset immediately. In Section 8.4 odf their book ([CJP15]) Cartea et al extend the model to incorporate market orders, so the trader can at particular stopping times τ, execute a market order to sell one or more units of the asset as well as posting limit orders at each time t, and solve the HJB equation accordingly. Gilbuad and Pham ([GP13]) also incorporate both limit orders and market orders in their approach to the execution problem. Note that Cartea et al's MO and LO problem does not include market impact for each market order. One could include that and make the problem more complete.

Appendix A

Probability

We list here some important mathematical concepts from probability theory which are used in the main text. A reference for this material is Jacod and Protter's "Probability Essentials" [JP04].

A.1 Basic concepts in probability theory

A random experiment is one whose outcomes cannot be predicted with certainty in advance.

- The **state space** Ω is the set of all possible outcomes ω of the experiment.
- An **event** A is a set of outcomes of the experiment, a subset of the state space: $A \subset \Omega$. It describes a property which can be observed either to be true or not true after the experiment has been done.

An **event** is a set of outcomes of the experiment, a subset of the state space. It describes a

A.2 Probability Spaces

Definition A.2.1 (Probability Space)**.** *A* ***probability space*** *is the triple* $(\Omega, \mathcal{A}, \mathbb{P})$.

A.3 Stochastic Processes

Definition A.3.1 (Wiener Process / Brownian motion)**.** *A stochastic process* $W = \{W_t\}_{0 \leq t \leq T}$*, defined on the filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(0 \leq t \leq T)}, \mathbb{P})$*, is a* ***Wiener Process*** *(or a standard Brownian motion) if it has the following properties:*

- $W_0 = 0$ $\mathbb{P}$*- almost surely;*

- *W has independent increments: for all times $0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_{n-1} \leq t_n = T$, the increments $W_{t_{k+1}} - W_{t_k}$ are independent random variables for $k = 0, \ldots, n-1$;*
- *W has stationary increments: for all $0 \leq t < t + h \leq T$, the increment $W_{t+h} - W_t$ has a probability distribution that is independent of the time t;*
- *the random variable W_t is Normally distributed with mean 0 and variance t, i.e. $W_t \sim \mathcal{N}(0, t)$;*
- *the function $t \to W_t$ is continuous almost surely.*

Appendix B

Python Code

B.1 Optimal Execution using Market Orders: Exponential Utility

Listing B.1: Optimal Execution using Market Orders: Exponential Utility

```
import numpy as np
import pandas as pd
import math
import random

import gzip
import csv
import json

import matplotlib as mpl
mpl.__version__

import matplotlib.pyplot as plt

plt.style.use('seaborn')

mpl.rcParams['font.family'] = 'serif'


""" variables:
    Stochastic Process:
        Brownian motion: W_t
    Brownian motion Parameters:
        volatility: sigma
        drift: mu
```

```
        Brownian motion initial value: W_0

    State Variables:
        asset midprice S_t
        inventory Q_t
        cash X_t
    Control Process:
        speed v_t

    Execution Parameters:
        risk aversion: gamma
        temporary impact: kappa
        permanent impact: eta
        terminal penalty: alpha
        running penalty: phi

    Simulation Parameters:
        terminal time: T
        Brownian motion initial value: W_0
        initial inventory Q_0
        initial stock price S_0
        initial cash process X_0
        number of time steps: nSteps
        number of simulations: nSims
    """


# initialise python random seed to the first one
# def initialiseRandomSeeds:
random.seed(0)
# initialise numpy random seed to the first one.
# note the python random seed and np random seed are completely independent.
np.random.seed(0)

# def simulateBrownianMotion:

# Execution Parameters:
T = 1 # terminal time
riskAversion = 1 # risk aversion parameter gamma
# Note that in our formulation of exp utility as
# U(x) = -exp(-gamma x), gamma can be any real number.
terminalPenalty = 0.01 # terminal penalty parameter alpha
runningPenalty = 0.01 # running penalty parameter phi

temporaryImpact = 0.001 # temporary impact parameter kappa
permanentImpact = 0.001 # permanent impact parameter eta
```

```
# Brownian motion parameters
volatility = 0.1
drift = 0
W_0 = 0 # Brownian motion starts at 0.

# constants from parameters: zeta and xi
xi = math.sqrt( (riskAversion*math.pow(volatility ,2)) / (2*temporaryImpact) )
zetaNumerator = ( -terminalPenalty
                  + (1/2)*permanentImpact
                  - (math.sqrt(temporaryImpact
                               *riskAversion*math.pow(volatility ,2))) )
zetaDenominator = ( -terminalPenalty
                    + (1/2)*permanentImpact
                    + (math.sqrt(temporaryImpact
                                 *riskAversion*math.pow(volatility ,2))) )

zeta  = zetaNumerator / zetaDenominator

# Simulation parameters
# set number of simulations
nSims = 1000
# set number of time steps
nSteps = 100
# generate standard normal random variates for each step of each simulation.
epsilon = np.random.normal(0, 1, (nSteps,nSims))

dt = T / nSteps # time step size



# State Variable Starting Points
inventory_0 = 10 # initial inventory
midprice_0 = 30 # initial asset midprice
cash_0 = 0 # initial cash



# initialise State variables

W = np.zeros((nSteps, nSims)) # brownian motion W
inventory = np.zeros((nSteps, nSims)) # inventory process Q
midprice = np.zeros((nSteps, nSims)) # asset midprice process S
cash = np.zeros((nSteps, nSims)) # cash process X

# simulate Brownian motion
```

```
for j in range(0, nSims):
    W[0][j] = W_0
    for t in range(1, nSteps):
        W[t][j] = W[t-1][j] + epsilon[t][j]*math.sqrt(dt)

# TWAP

# initialise control under TWAP
speedTWAP = np.zeros(nSteps)
# initialise state variables under TWAP
inventoryTWAP = np.zeros(nSteps)
midpriceTWAP = np.zeros((nSteps, nSims))
cashTWAP = np.zeros((nSteps, nSims))

# simulate trading speed under TWAP
for t in range(0, nSteps):
    speedTWAP[t] = inventory_0 / T

# simulate inventory under TWAP
for t in range(0, nSteps):
    inventoryTWAP[t] = ( (1-(t*dt)/T)*inventory_0 )

# simulate midprice under TWAP

for j in range(0, nSims):
    midpriceTWAP[0][j] = midprice_0
    for t in range(1, nSteps):
        midpriceTWAP[t][j] = ( midpriceTWAP[t-1][j]
                               - permanentImpact*speedTWAP[t-1]*dt
                               + volatility*(W[t][j] - W[t-1][j]) )

# simulate cash process under TWAP
for j in range(0, nSims):
    cashTWAP[0][j] = cash_0
    for t in range(1, nSteps):
        cashTWAP[t][j] = ( cashTWAP[t-1][j] +
                          (midpriceTWAP[t-1][j]*speedTWAP[t-1]
                          - temporaryImpact*math.pow(speedTWAP[t-1], 2))*dt )

# initialise Optimal State Variables
inventoryOpt = np.zeros(nSteps) # optimal inventory process Q*

speedOpt = np.zeros(nSteps) # Optimal Trading Speed v*

midpriceOpt = np.zeros((nSteps, nSims))
# asset midprice process S* given optimal trading trajectory
```

```
cashOpt = np.zeros((nSteps, nSims))
# cash process X* given optimal trading trajectory

# simulate optimal inventory
inventoryOpt[0] = inventory_0

for t in range(1,nSteps):
    inventoryOpt[t] = ( inventory_0
                        *( (zeta*math.exp(xi*(T-t*dt))
                            - math.exp(-xi*(T-t*dt)))
                          / (zeta*math.exp(xi*T) - math.exp(-xi*T)) ))

# simulate optimal Trading Speed
for t in range(0,nSteps):
    speedOpt[t] = ( inventory_0*xi
                    *((zeta*math.exp(xi*(T-t*dt)) + math.exp(-xi*(T-t*dt)))
                      / (zeta*math.exp(xi*T) - math.exp(-xi*T)) ) )



# simulate optimal Stock price
for j in range(0, nSims):
    midpriceOpt[0][j] = midprice_0
    for t in range(1, nSteps):
        midpriceOpt[t][j] = ( midpriceOpt[t-1][j]
                              - permanentImpact*speedOpt[t-1]*dt
                              + volatility*(W[t][j] - W[t-1][j]) )


# simulate optimal cash flow
for j in range(0, nSims):
    cashOpt[0][j] = cash_0
    for t in range(1, nSteps):
        cashOpt[t][j] = ( cashOpt[t-1][j] +
                          (midpriceOpt[t-1][j]*speedOpt[t-1]
                           - temporaryImpact*math.pow(speedOpt[t-1], 2))*dt )




# simulate optimal inventory with limiting penalty
inventoryOptLimit = np.zeros(nSteps)
speedOptLimit = np.zeros(nSteps)

midpriceOptLimit = np.zeros((nSteps, nSims))
# asset midprice process S* given optimal trading trajectory
```

```
cashOptLimit = np.zeros((nSteps, nSims))
# cash process X* given optimal trading trajectory

inventoryOptLimit[0] = inventory_0

for t in range(1,nSteps):
    inventoryOptLimit[t] = ( inventory_0
                            *( (math.sinh(xi*(T-t*dt)))
                              / (math.sinh(xi*(T))) ) )

for t in range(0,nSteps):
    speedOptLimit[t] = ( xi*inventory_0
                        *( (math.cosh(xi*(T-t*dt))) / (math.sinh(xi*(T))) ) )


#simulating limiting optimal stock prices and cash flows
midpriceOptLimit[0] = midprice_0
for j in range(0, nSims):
    midpriceOptLimit[0][j] = midprice_0
    for t in range(1, nSteps):
        midpriceOptLimit[t][j] = ( midpriceOptLimit[t-1][j]
                                  - permanentImpact*speedOptLimit[t-1]*dt
                                  + volatility*(W[t][j] - W[t-1][j]) )

cashOptLimit[0] = cash_0
for j in range(0, nSims):
    cashOptLimit[0][j] = cash_0
    for t in range(1, nSteps):
        cashOptLimit[t][j] = ( cashOptLimit[t-1][j]
                              + (midpriceOptLimit[t-1][j]*speedOptLimit[t-1]
                                 - (temporaryImpact
                                    *math.pow(speedOptLimit[t-1], 2))*dt ))


plt.figure(figsize=(12,8))
plt.plot(inventoryTWAP, label="inventory under TWAP")
plt.plot(inventoryOpt, label="Optimal Inventory")
plt.plot(inventoryOptLimit, label="Optimal Inventory as alpha -> infty")

plt.legend(bbox_to_anchor=(1.3, 0.5))
plt.xlabel("Time")
plt.ylabel("Inventory")
plt.title("Optimal Inventory given a terminal penalty of "
          + str(terminalPenalty) )
plt.show()
```

B.2 Optimal Execution using Limit Orders

Listing B.2: Optimal Execution using Limit Orders

```
# -*- coding: utf-8 -*-
"""
Created on Thu Jun 29 20:29:04 2023

@author: AdamNiiArmahHesse
"""

""" Cartea, Jaimungal, Penalva Ex8.1

This is a solution to the Optimal Execution
problem with Limit Orders,
with a running penalty function for the inventory held by an
execution trader.

The trader is choosing the optimal depths
at which to execute his child orders.

We solve by using a recursive formula for Omega(t,q)
and then finding h(t,q) and the optimal depth delta*(t,q).

Thurs 29 June 2023, 20:33
"""
import numpy as np
import pandas as pd
import math

import gzip
import csv
import json

import matplotlib as mpl
mpl.__version__

import matplotlib.pyplot as plt

plt.style.use('seaborn')

mpl.rcParams['font.family'] = 'serif'

"""
mpl.use("pgf")
```

```
mpl.rcParams.update({
    "pgf.texsystem": "pdflatex",
    'font.family': 'serif',
    'text.usetex': True,
    'pgf.rcfonts': False,
})
"""
inventoryToLiquidate = 5
# terminal time is 1 minute = 60 seconds.
terminalTime = 60
numberOfSeconds = 60

midpriceZero = 30
#
volatility = 0.01

terminalPenalty = 0.001
runningPenalty =  0.00001

# 50 incoming MOs per minute on average.
intensityMO = (50 / 60)*(terminalTime/numberOfSeconds)
fillProbParameter = 100
# fillProb = math.exp(-fillProbParameter*depth[t])
# incomingMO =
# filledMO =

nSteps = terminalTime
nSimulations=10000
dt = terminalTime/nSteps


grid = np.zeros((inventoryToLiquidate + 1, terminalTime + 1))
gridDimensions = (inventoryToLiquidate +1, terminalTime +1)

omegaGrid = np.zeros(gridDimensions)
hGrid = np.zeros(gridDimensions)
depthGrid = np.zeros(gridDimensions)


# One simulation for now.
epsilon = np.zeros(nSteps)
inventory = np.zeros(nSteps)
brownianMotion = np.zeros(nSteps)
midprice = np.zeros(nSteps)
midprice[0] = 30
#the control process is the depth at which the trader posts in the LOB
```

```
depth = np.zeros(nSteps)
incomingMO = np.zeros(nSteps) # Poisson process for incoming MOs,
# intensity given by intensityMO
filledMO = np.zeros(nSteps) # Counting process for filled MO.
cash = np.zeros(nSteps)
# cash[t+1] = cash[t] + (midprice[t] + depth[t])*filledMO[t]

np.random.seed(0) # set random number generator to the first generator

epsilon = np.random.normal(loc=0, scale=1, size=nSteps)

# Simulating midprice (note there are no feedback effects
# from the control (depth) on the midprice here
# so we can pre-simulate the entire price path.)
for t in range(nSteps-1):
    brownianMotion[t+1] = brownianMotion[t] + epsilon[t]*math.sqrt(dt)
    midprice[t+1] = midprice[0] + volatility*(brownianMotion[t+1])

# Generating the Poisson process for incoming market orders
for t in range(nSteps -1):
    uniformVariate = np.random.uniform(low=0.0, high=1.0, size=None)
    inverseExpCDF = (-math.log(1-uniformVariate, math.e))/(intensityMO)
    incomingMO[t] = math.floor(inverseExpCDF)

# Setting boundary conditions
for t in range(0, terminalTime +1):
    omegaGrid[0][t] = 1


for q in range(1, inventoryToLiquidate +1):
    omegaGrid[q][terminalTime] = math.exp(-fillProbParameter
                                    *terminalPenalty*pow(q,2))


    for t in range(0, terminalTime+1):

        omegaGrid[q][t] = (
            math.exp(-fillProbParameter
                    * runningPenalty
                    * pow(q,2)
                    * (terminalTime - t))
                * ( math.exp(-fillProbParameter
                        * terminalPenalty * pow(q,2)))
                  + ((intensityMO/math.e)/(fillProbParameter
                                    * runningPenalty
                                    * pow(q,2)))
```

```
                *(1
                  - math.exp(-fillProbParameter
                             * runningPenalty
                             * pow(q, 2)
                             * (terminalTime - t)))
                * omegaGrid[q-1][t]
            )

for q in range(0, inventoryToLiquidate+1):
    for t in range(0, terminalTime +1):
        hGrid[q][t] = ((1/fillProbParameter)
            *math.log(omegaGrid[q][t], math.e)
        )
# delta grid
for q in range(1, inventoryToLiquidate+1):
    for t in range(0, terminalTime+1):
        depthGrid[q][t] = ((1 / fillProbParameter)
            + (hGrid[q][t] - hGrid[q-1][t])
        )

plt.figure(figsize=(12,8))
for q in range(1, inventoryToLiquidate+1):
    plt.plot(depthGrid[q])
#plt.legend(bbox_to_anchor=(1.3, 0.5))
plt.xlabel("Time")
plt.ylabel("Optimal LO Depths")
#plt.title("Optimal Depths for LOs")
plt.ylim(ymin=0)  # force y-axis to start at 0.
plt.show()
```

Bibliography

[AC00] Robert Almgren and Neill Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 51(3):5–39, 2000.

[Ald13] Irene Aldridge. *High-frequency trading: a practical guide to algorithmic trading strategies and trading systems.* John Wiley & Sons, 2013.

[AS08] Marco Avellaneda and Sasha Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.

[BL98] Dimitris Bertsimas and Andrew W. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.

[CJ16a] Álvaro Cartea and Sebastian Jaimungal. A closed-form execution strategy to target volume weighted average price. 7:760–785, 2016. `http://www.siam.org/journals/sifin/7/M105840.html`.

[CJ16b] Álvaro Cartea and Sebastian Jaimungal. Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, 2016.

[CJP15] Álvaro Cartea, Sebastian Jaimungal, and José Penalva. *Algorithmic and High Frequency Trading.* Cambridge University Press, 2015.

[CW12] René Carmona and Kevin Webster. High frequency market making. 2012.

[Don19] Brent Donnelly. *The Art of Currency Trading.* John Wiley & Sons, 2019.

[GLFT13] Olivier Guéant, Charles-Albert Lehalle, and Joaquin Fernandez-Tapia. Dealing with the inventory risk: a solution to the market making problem. *Mathematics and Financial Economics*, 7(4):477–507, 2013.

[GLSW17] Xin Guo, Tze Leung Lai, Howard Shek, and Samuel Po-Shing Wong. *Quantitative Trading: algorithms, analytics, data, models, optimization.* CRC Press, 2017.

[GP13] Fabien Guilbaud and Huyên Pham. Optimal high-frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79–94, 2013.

[Hul12] John C. Hull. *Options, Futures, and Other Derivatives, Eighth Edition, Global Edition*. Pearson, 2012.

[JP04] Jean Jacod and Philip Protter. *Probability Essentials*. Springer-Verlag, 2004.

[Lef23] Edwin Lefevre. *Reminiscences of a Stock Operator*. George H. Doran Company, 1923.

[Lew15] Martin Lewis. *Flash Boys*. Penguin Books, 2015.

[Men13] Albert J. Menkveld. High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712–740, 2013.

[Mer69] Robert Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–257, 1969.

[OW13] Anna A. Obizhaeva and Jiang Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32, 2013.

[Pha09] Huyên Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer, 2009.

[Shr08] Steve Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance, 2008.

[SS09] Alexander Schied and Torsten Schoeneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastics*, 13(2):181–204, 2009.