A Partial Differential Equation Approach to Three Problems in Finance: Barrier Option Pricing, Optimal Asset Liquidation and Insider Trading

Weston Barger

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Matthew Lorig, Chair

Tim Leung

Hong Qian

Program Authorized to Offer Degree:
Applied Mathematics

University of Washington

Abstract

A Partial Differential Equation Approach to Three Problems in Finance: Barrier Option Pricing, Optimal Asset Liquidation and Insider Trading

Weston Barger

Chair of the Supervisory Committee:
Associate Professor Matthew Lorig
Applied Mathematics

We examine three problems in mathematical finance. These problems broadly fall under the sub-disciplines of contract pricing and optimal execution of orders on an exchange under price impact.

The first problem deals with the pricing of contingent claims, a classical problem in mathematical finance. After a brief introduction to pricing, we derive asymptotic expansions for the prices of a variety of European and barrier-style claims in a general local-stochastic volatility setting. Our method combines Taylor series expansions of the diffusion coefficients with an expansion in the correlation parameter between the underlying asset and volatility process. We provide rigorous accuracy results for European-style claims which depend explicitly on the correlation parameter. For barrier-style claims, we include several numerical examples to illustrate the accuracy and versatility of our approximations.

We then turn our attention to the problem of optimal execution. We assume a continuous-time price impact model similar to that of Almgren-Chriss but with the added assumption that the price impact parameters are stochastic processes modeled as correlated scalar Markov diffusions. For a fixed trading horizon, we perform coefficient expansion on the Hamilton-Jacobi-Bellman equation associated with the trader's value function. The coefficient expansion yields a sequence of partial differential equations that we solve to give

closed-form approximations to the value function and optimal liquidation strategy, which is given in feedback form. We examine some special cases of the optimal liquidation problem and give financial interpretations of the approximate liquidation strategies in these cases. We then provide numerical examples to demonstrate the efficacy of our approximations.

The third problem we consider deals with insider trading under price impact. We consider an exponentially risk-averse (or risk-neutral) trader who wishes to capitalize on inside information about the true value of an asset that is traded on an exchange. The insider faces a market maker who desires to set a fair price for the asset and attempts to do so by leveraging information contained in the aggregate order flow. We also assume that the insider faces a trading cost that is a linear function of the insider's trading speed. We give equilibrium strategies for the insider and the market maker in both the single auction and continuous-time settings. While the discrete-time equilibrium was given previously, we expand upon it by performing an expansion in small trading cost to analyze the behavior as the cost tends towards zero. A continuous-time equilibrium is given in terms of the solution to a forward-backward ordinary differential equation, which we arrive at by applying nonlinear filtering and dynamic programming in continuous time. We then give some comparative statics about the insider and market maker's strategies and numerically explore the effects of varying model parameters.

TABLE OF CONTENTS

Page

LIST OF FIGURES

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Matthew Lorig. Matt has taught me a significant amount of mathematics, but, more generally, he has imparted on me his organized and structured approach to attacking challenging, technical problems. These skills have carried over into many aspects of my life and have already paid dividends, figuratively and literally. I am without a doubt smarter as a direct result of his of mentorship. Matt encouraged me throughout my phd to pursue projects and interests about which I was passionate, even when those projects and interests did not align with his, an advising approach that I never took for granted. Matt and I also share a plethora of tastes and personal interests, which routinely turned planned short research meetings into hours long conversations about life, weight lifting, sports, music and so on. The ways in which Matt has impacted my professional life are numerous and significant but perhaps not as significant than the ways in which he has impacted my personal life. Matt helped give me courage to unapologetically be my true self and inspired me to shed worry about that which I cannot control. I could not have been more fortunate to have an academic advisor who taught me so much and advocated for me so strongly, but who I can also count as a close, personal friend.

I am extremely grateful for my friend and co-author Ryan Donnelly. Ryan provided me a great deal of support on our project together, a work I am particularly proud of, as well as much needed reprieve from the stresses of graduate school and job hunting by accompanying me to the College Inn Pub.

I received help from every member of the Applied Mathematics staff. In particular, I would like to thank Lauren Lederer for her administrative support, particularly during the hectic graduation process, and Karen Beaudry, for her invaluable help in securing my first graduate school internship, an achievement which positively altered my career trajectory.

I would like to thank my dear, personal friend and fellow metalhead Bobby "Bob" Baraldi. On Bobby's visiting weekend, he commented on my Skeletonwitch t-shirt, and we immediately became fast friends. Having massive guitar nerd in the next office was more than I could have ever hoped for in graduate school. Bobby accompanied me to virtually every concert I attended, and together we experienced the most technical and beautiful music the world has to offer. His observational prowess and relentless humor never failed to brighten even my most gloomy days, and his work ethic and sheer grit in the face of seemingly insurmountable adversity was, and continues to be, deeply moving and inspiring.

I am also grateful for my friendship with Diya Sashidhar. We share passions for coffee, spicy chicken and college football, and I cherished our Orange King/Parnassus lunch routine, where those passions converged. Our shared experience growing up in the South was deeply comforting to me during times of homesickness. I have never met someone with such a combination of compassion; social, emotional and intellectual intelligence; and humor. Diya uplifts everyone around her, and I am exceptionally lucky to have been one of those people.

I am immensely thankful for my friendship with Matt Lukac. Matt and my decade long friendship began in Arkansas, but with great fortune has continued on in the Pacific Northwest. Our semi-annual meet ups in Eugene and Seattle are among the happiest memories of my life. Matt is one of the funniest people I have ever met, and knowing that I can pick up the phone and dial up a laugh has provided me a great deal of solace during turbulent times. I have been fortunate to have someone

near me during graduate school with a shared background and common interests in mathematics, Razorbacks football, music and cycling. Matt's determination and work ethic is contagious, and as such our friendship has motivated me to continue to challenge and improve myself in a variety of ways.

And where would I be without my friends, roommates and incoming 2014 class companions Brian de Silva and Jeremy Upsal? I cannot imagine better people with which to experience our program's first year than Brian and Jeremy. They fostered a spirit of friendly competitiveness, which inspired my best performance. Brian and Jeremy both have extraordinary moral and ethical compasses, and they are steadfast in their principals, even when it is not convenient. I will forever hold dear the mornings we spent standing in our kitchen, starting the day with good coffee and laughter.

Finally, I would like to express my love and gratitude for my mother Michele, her husband Jim, and my father Steve. I cannot put to words the amount of love and support that I receive from them on a daily basis. I am thankful for them every day.

DEDICATION

I dedicate this thesis to my mother Michele and my father Steven.

Chapter 1

INTRODUCTION

This thesis is collection of problems in the field of mathematical finance. The work herein falls broadly under the mathematical finance sub-disciplines of option pricing and optimal trade execution, and each of the chapters that follow correspond to a problem in their respective sub-fields. As such, each chapter is largely self-contained, and we save the necessary background information, literature review, and modeling considerations for the beginning of each chapter. As this thesis spans multiple sub-disciplines, the author performed a balancing act with regards to notation. The author attempted to conform each chapter to the notational conventions of the respective literature while maintaining consistency throughout the document. The reader should assume that the notional scope is at the chapter level, including the appendix.

Chapter 2 falls under the classical mathematical finance sub-discipline of option pricing, and this work is largely contained in the paper [8]. In Chapter 2, we present a contract pricing framework—which includes European and barrier-style claims—where the underlying asset is prescribed local-stochastic volatility dynamics. The coefficients of the asset model are specified only up to a class of function i.e. specific dynamics are not stipulated. We assume that the coefficients of the underlying price process are such that the associated pricing partial differential equation (PDE) is parabolic and known to have well-behaved solutions. However, due to the generality of the coefficients closed form solutions of the pricing PDE are unavailable, and we turn to asymptotic expansions to obtain closed form approximations. We develop closed form approximations to the price of European claims as well as single and double-barrier claims by using a Taylor series expansion on the coefficients of the pricing PDE.

We shift our focus away from option pricing in Chapter 3 and move to the problem of optimal trade execution under price impact models. We consider a trader who seeks to liquidate a large, long position in a risky security (the optimal acquisition problem is identical). The trader would like to maximize the expected terminal wealth resulting from asset liquidation at the end of a fixed trading horizon. In our market model, the trader's order submissions to sell put downward pressure on the midprice of the asset. Furthermore, a fee that is a function of the rate order submissions is also paid by the trader. In their seminal work, the authors Almgren and Chriss of [4] model price impact as linear functions of the order submission rate, formulate the optimal control problem associated with wealth maximization and obtain the optimal control via dynamic programming. But, as distinct from [4], we take the coefficients of the linear price-impact functions to be diffusion processes with general coefficients in order to capture the dynamic nature of order books on security exchanges. This prevents us from explicitly solving the associated dynamic programming (partial differential) equation in closed form. So, in the same spirit as the problem of Chapter 2, we turn to Taylor series expansions to develop a closed form approximation of the optimal trading strategy. The difference here, though, is that the PDE considered in Chapter 3 is fully nonlinear. Furthermore, we use the approximation formulas for the value function to give approximations to the optimal control, which is given in feedback form. This work is largely contained in the paper [9].

Like Chapter 3, we also consider price impact models in Chapter 4. However, the setting there is fundamentally different. In Chapter 4, we consider a trader who obtains information unavailable to the public about the value of a risky asset to be realized at a fixed date in the future. We assume that the trader would like to optimally capitalize on their informational advantage over the rest of the market. However, we assume that there is a market maker who views the aggregate order flow for the asset and seeks to set a fair price. While the insider's trades are obscured by those of noise traders, information about the true value of the underlying leaks to the market maker as the insider trades. The market maker incorporates this information into prices to the disadvantage of the insider. That is, as the insider trades they affect price impact. We present equilibrium strategies for the insider and the market maker. Chapter 4 expands on the existing literature by assuming that the insider also faces a transaction cost, felt only by the insider, associated with submitting orders to the exchange. The introduction of the transaction fee qualitatively changes equilibrium solutions from those developed in previous works. Similarly to Chapter 3, we employ dynamic programming to develop optimal order submission strategies for the insider. We obtain the optimal control in feedback form and then reduce the solution of the dynamic programing PDE to the solution of an ordinary differential equation (ODE) system, for which we show the existence and uniqueness of a solution.

Without further ado, let us present the first problem.

Chapter 2

APPROXIMATE PRICING OF EUROPEAN AND BARRIER CLAIMS

2.1 Background

In finance, an *option* is a contract between two parties which gives the owner of the contract the right to buy or sell an asset from the other party at an agreed upon price on or before an agreed upon date, given that the conditions of the contract are met. The asset on which the contract is written is known as the *underlying*. It is important to note that the owner of the contract does not have the obligation to exercise their right. However, no reasonable trader would decline to exercise their right if the payoff at expiry associated with exercising is positive.

An investor may use an option contract to speculate on future price movements. The investor can purchase options, which are often cheaper than the underlying, to bet on price movements. Furthermore, an investor who would like hedge their position in an asset can buy options as insurance against adverse price movements. After their instantiation, option contracts may be bought or sold. The use of options in modern investing is so ubiquitous that there are exchanges dedicated entirely to options trading (e.g. The Chicago Board Options Exchange).

Because the owner of the option contract has the potential to make a profit in the future, the option contract itself has a value. Due to the ubiquity of option trading in modern finance, there has been a great deal of academic (and otherwise) research on computing the fair price of an option. To compute the price of an option, it is necessary to prescribe a model for the underlying asset.

The *Brownian motion* is an indispensable tool in asset modeling. In this thesis, Brownian motion refers to a continuous-time stochastic process W with the properties

1. $W_0 = 0$,

2. the path of W_t is continuous almost-surely, and

3. for each $0 \leq t \leq s < \infty$, $W_s - W_t \sim \mathcal{N}(0, s-t)$.

Brownian motion—named for the botanist Robert Brown, who gave a qualitative description of the phenomenon by observing particles floating on top of water behaving erratically—was first mathematically formalized by Thiele (see [46]). Bachelier independently developed the Brownian motion in his thesis [6] and was also the first to employ the use of Brownian motion to model asset prices. Einstein brought Brownian motion to the attention of physicists in his paper [30].

However, it was in 1973 that the authors Black and Scholes of [15] first utilized Brownian motion in the computation of option prices. The pricing formula is foundational in mathematical finance and is widely known as the Black-Scholes or the Black-Scholes-Merton formula, to honor the contributions of Robert Merton [54]. They used a geometric Brownian motion (GBM), first used to model asset prices in [63], to model the price of the risky underlying asset and subsequently constructed a dynamically traded portfolio consisting of the underlying and a bond that replicated the terminal value of the option, known as the option payoff. They argued that the under the assumption of a lack of arbitrage in the market the price of the option at a given time must be equal to the value of the replicating portfolio, as any other option price would lead to an arbitrage opportunity.

Constructing an option-replicating portfolio and applying a no-arbitrage argument is one method by which option prices can be computed. The authors of [38] showed that the absence of arbitrage on a market is equivalent to the existence of a probability measure such that each non-dividend paying asset on the market denominated in units of a common non-dividend paying asset forms a martingale. This probability measure is known as a *risk-neutral measure*. The present time value of an option can be found by computing the expected value of the (discounted) option payoff with respect to the risk-neutral measure conditioned on the path of the underlying. Generally, this procedure is referred to as *risk-neutral pricing*, and it is this approach that we take in the upcoming chapter.

It is often the case that authors in the option-pricing literature model the underlying asset as a stochastic differential equation (SDE) and specify the dynamics of an asset under a risk-neutral measure directly, and align ourselves with those authors in this chapter. For example, consider a market is composed of a single risky asset S, a bond B, which for the sake of simplicity we assume to pay an interest rate of $r = 0$, and a European option contract V expiring at time $t = T$ written on S with payoff $\varphi(S_T)$. Suppose further that under a risk-neutral measure $\mathbb{P}$ the asset evolves according to the dynamics

$$\mathrm{d}S_t = \mu(t, S_t)\,\mathrm{d}t + \sigma(t, S_t)\,\mathrm{d}W_t,$$

where $W = (W_t)_{0 \le t}$ is a standard Brownian motion under $\mathbb{P}$. Well, as V (denominated by the bond $B = 1$) is a martingale under the risk-neutral measure $\mathbb{P}$ and $V_T = \varphi(S_T)$, the time $t \le T$ value of the of the option is given by

$$V_t = \mathbb{E}^{\mathbb{P}}\left[V_T \mid \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\varphi(S_T) \mid \mathcal{F}_t\right],$$

where $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ is the filtration generated by S. As S is the solution to an SDE, and is hence a Markov process, there exists a function $u \equiv u(t, s)$ such that

$$u(t, s) = \mathbb{E}^{\mathbb{P}}\left[\varphi(S_T) \mid S_t = s\right] = \mathbb{E}^{\mathbb{P}}\left[\varphi(S_T) \mid \mathcal{F}_t\right] = V_t.$$

The function u is known to satisfy the parabolic PDE

$$\partial_t u(t, s) + \frac{1}{2}\sigma^2(t, s)\partial_s^2 u(t, s) + \mu(t, s)\partial_s u(t, s) = 0, \qquad u(T, s) = \varphi(s), \quad (2.1.1)$$

known as the *Kolmogorov backward equation*. In this setting, computing the time-t value of the option V_t is therefore equivalent to solving the PDE problem (2.1.1). In this chapter, we will cast the problem of pricing options as PDE problems and apply PDE approximation methods to compute approximate prices. We refer the reader to the textbooks [65] and [14] for detailed introductions to risk-neutral option pricing.

The most simple option contracts allow the holder to buy (sell) the asset from (to) the contract writer within a given time frame. It is for the lack of attached bells and whistles that these contracts earned the name *vanilla options*. The European flavored option gives the holder their right to exercise *only* at a specified time, and its American counterpart allows the holder to exercise their right at any moment up to a specified time. More complicated contracts are referred to as *exotic* options. The payoff of exotic options are often dependent on the realized path of the underlying asset as opposed to simply the value of an asset at a specified time.

A common example of an exotic option is a barrier-style option. For barrier options, an event is triggered by the price of the underlying asset reaching a predetermined level. For example, a single-barrier knock-out option with lower barrier L written on an underlying S with a payoff $\varphi(S_T)$ is a contract such that the holder of the contract receives $\varphi(S_T)$ at time $t = T$ only if the asset price never touches the lower barrier L in the time interval $[0, T]$. Otherwise, the contract owner receives no payout.

Barrier-style options are some of the most liquid path-dependent claims, and barrier-style claims are often cheaper than their European counterparts. As such, the barrier-style claims are popular among speculators who wish to bet on market movements while taking advantage of the lower prices barrier-style claims entail. Despite the widespread use of barrier-style claims they remain challenging to price.

In his seminal work, [54] was the first to value a down-and-out call in closed form for a stock that follows a GBM. Static hedging results are available for barrier claims in GBM and similar settings. [16] show that in the GBM setting a down-and-out call can be replicated by going long a European call with the same underlying, maturity and strike while going short a specified number of puts, where the number of puts and the put strike price are functions of the barrier and strike of the down-and-out call. [18] extend this result by showing that the static hedge of [16] works for any asset price model with local volatility when the volatility function is symmetric in the log of the futures price relative to the barrier. [19] make clear that the aforementioned local volatility symmetry condition is sufficient but not necessary.

The replication strategy described above for a down-and-out call is valid when there are no jumps over the barrier and the call and put have the same implied volatility at the first passage time to the barrier. This condition is known as as Put Call Symmetry (PCS) and was introduced to finance by [11]. [20] give static replication strategies for barrier-style claims for a general class of local volatility models, and [17] present semi-static hedges for barrier options on price and volatility.

The CEV ([27]), Heston ([40]), and SABR ([37]) models are among the most frequently used models for pricing European claims. These models do not posses the symmetry conditions discussed in the previous paragraph. However, a number of pricing formulas are known and the associated hedging strategies are dynamic. For example, the authors of [29] price barrier-style claims under the CEV model by using eigenfunctions expansions. In the zero correlation setting, [31] prices barrier options can be priced via Fourier Sine series for double barrier options or via Fourier Sine transforms for single barrier options under the Heston model.

When underlyings are modeled as zero-correlation stochastic volatility processes, the induce induce symmetric implied volatility smiles [62]. This is inconsistent with empirical evidence from equity markets where smiles exhibit at-the-money skews [61]. For accurate pricing, it is therefore advantageous to permit the underlying to be correlated with the volatility-driving process. When the correlation of the price and volatility processes is non-zero, closed-form pricing formulas for barrier option prices are not available and perturbation methods are often employed. [47] presents an approximate pricing formula for barrier options by expanding the pricing partial differential equation in the vol-of-vol parameter. The authors of [32] price barrier options under a fast mean-reverting volatility model. And [49] values barrier options and other claims for a class of multiscale stochastic volatility models (see [33] for a review of these models). However, the methods of [47] and [32] cannot be applied in the CEV or SABR settings, and the results in [49] require a separation of time scales between the price process and the corresponding fast and slow factors of volatility, which may not be realistic in certain markets.

In this chapter, we consider a very general class of local-stochastic volatility models, which includes the CEV, Heston and SABR models. We approximately price barrier-style claims by expanding the coefficients of the infinitesimal generator of the underlying as a Taylor series about a fixed point. The Taylor series expansion method was initially developed for European-style claims in scalar diffusion setting in [59] and later extended to d-dimensional diffusions in [52] and [51].

Adapting the methods developed in [51], which deal with diffusions on $\mathbb{R}^d$, to diffusions on strict subsets of $\mathbb{R}^d$ is not straightforward. In particular, the zeroth order approximation to the transition density of a diffusion in $\mathbb{R}^d$ is a Gaussian kernel that is a function of the difference of the forward and backward variables. This symmetry between forward and backward variables is exploited to simplify the computations required to obtain higher order corrections to the transition density. But for diffusions in a strict subset of $\mathbb{R}^d$, the zeroth order transition density approximation will no longer be a function of the difference of the forward and backward variables and depends on the geometry of the subset. As a result, the computations needed to obtain higher order corrections to the transition density are significantly more involved.

The rest of this chapter proceeds as follows. In Section 2.2, we introduce a general local-stochastic volatility model and describe the option-pricing problems we wish to solve. In Section 2.3, we develop an asymptotic expansion for options prices. This expansion leads to a sequence of nested PDE problems, which we solve explicitly in Section 2.4. In Section 2.5, we establish the asymptotic accuracy of our approximation for European options. Finally, in Section 2.6, we provide several numerical illustrations of our pricing approximation for barrier-style claims and compare our results to prices obtained via Monte Carlo simulation. Some concluding remarks are offered in Section 2.7.

2.2 Market Model

In this chapter, we consider a market defined on a complete, filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t}, \mathbb{P})$. Here, the measure $\mathbb{P}$ represents the market's chosen pricing measure. Let $S = (S_t)_{t \geq 0}$ be the value of a risky asset. We suppose the dynamics of S are given by

$$\begin{aligned} S_t &= f(X_t), \\ \mathrm{d}X_t &= \mu(X_t, Y_t)\mathrm{d}t + \sigma(X_t, Y_t)\mathrm{d}W_t, \\ \mathrm{d}Y_t &= c(X_t, Y_t)\mathrm{d}t + g(X_t, Y_t)\mathrm{d}B_t, \\ \mathrm{d}\left\langle W, B\right\rangle_t &= \rho\,\mathrm{d}t. \end{aligned}$$

where the function f must be positive, strictly increasing and $f \in C^2(\mathbb{R})$. The processes $W = (W_t)_{t \geq 0}$ and $B = (B_t)_{t \geq 0}$ are driftless $(\mathbb{P}, \mathbb{F})$-Brownian motions with constant correlation $\rho \in (-1, 1)$. We assume the dynamics of $(X, Y) = (X_t, Y_t)_{t \geq 0}$ are such that (X, Y) has a unique strong solution until the first exit time of X of some interval $I \subseteq \mathbb{R}$.

For simplicity, we take the risk-free rate of interest to be zero. Thus, in order to preclude the possibility of arbitrage, the risky asset S must be a martingale under $\mathbb{P}$. As a result, the function μ must satisfy

$$\mu = \frac{-f''\sigma^2}{2f'}.$$

The condition on μ can be easily derived by computing $\mathrm{d}f(X_t)$ and setting the $\mathrm{d}t$-term to zero. Typical choices for f are $f(x) = \mathrm{e}^x$ and $f(x) = x$, which give rise to the conditions $\mu = -\frac{1}{2}\sigma^2$ and $\mu = 0$, respectively.

We are interested in computing the price of a barrier-style contingent claim, whose payoff at the maturity date T is given by

$$\text{Payoff :} \qquad \mathbb{I}_{\{\tau > T\}}\varphi(X_T), \qquad \tau = \inf\{t \geq 0 : X_t \notin I\}, \tag{2.2.1}$$

where I is an interval in $\mathbb{R}$. For a single-barrier claim with a lower barrier $L < X_0$ we have $I = (L, \infty)$, and for a single-barrier claim with an upper barrier $X_0 < U$ we have $I = (-\infty, U)$. For a double-barrier claim, we have $I = (L, U)$ where $L < X_0 < U$. We also allow for the possibility that $I = (-\infty, \infty)$, which corresponds to a European claim on X.

Remark 2.2.1. When $I \neq \mathbb{R}$, payoffs of the form (2.2.1) are *knock-out* style payoffs. A *knock-in* style payoff is a payoff of the form

$$\text{Payoff}: \qquad \mathbb{I}_{\{\tau \leq T\}}\varphi(X_T).$$

Let V^{KO} and V^{KI} be a knock-in and knock-out claim, respectively, each with strike K and barrier L and payoff $\varphi(X_T)$. Then

$$V_T^{KO} + V_T^{KI} = \mathbb{I}_{\{\tau > T\}}\varphi(X_T) + \mathbb{I}_{\{\tau \leq T\}}\varphi(X_T) = \varphi(X_T) = V_T^E,$$

where V^E is a European claim on X with payoff $\varphi(X_T)$. Thus, via a no arbitrage argument, by pricing both knock-out and European style claims we have also priced knock-in style claims.

The value V_t of the claim with payoff (2.2.1) at time $t \leq T$ is given by

$$V_t = \mathbb{I}_{\{\tau > t\}} u(t, X_t, Y_t), \qquad u(t,x,y) := \mathbb{E}\left(\mathbb{I}_{\{\tau > T\}}\varphi(X_T) | X_t = x, Y_t = y, \tau > t\right),$$

Under mild conditions, the function u is the unique classical solution of the Kolmogorov Backward equation

$$0 = (\partial_t + \mathcal{A})u, \qquad u(T, \cdot) = \varphi, \tag{2.2.2}$$

where $\mathcal{A}$ is the generator of (X, Y), is given explicitly by

$$\mathcal{A} = \mu(x,y)\partial_x + \tfrac{1}{2}\sigma^2(x,y)\partial_x^2 + c(x,y)\partial_y + \tfrac{1}{2}g^2(x,y)\partial_y^2 + \rho\sigma(x,y)g(x,y)\partial_x\partial_y,$$

and is defined to act on functions that are twice differentiable and satisfy the boundary conditions

$$\text{dom}(\mathcal{A}) := \{g \in C^2(I) : \lim_{x \to \partial I} g(x,y) = 0\}.$$

Here we use the notation ∂I to indicate a *finite* endpoint of I. So, for example, if $I = (L, \infty)$, then $\mathcal{A}$ acts on functions that satisfy $\lim_{x \searrow L} g(x,y) = 0$. Throughout this chapter, we assume a unique classical solution to (2.2.2) exists. Solving the PDE (2.2.2) explicitly would give an explicit expression for V. As no explicit solution of (2.2.2) exists for general coefficients (μ, σ, c, g), we shall seek instead an explicit approximation for u.

2.3 Formal asymptotic expansion

In this section, we present a formal asymptotic expansion for u. To begin, let us introduce some notation. For any coefficient function of $\mathcal{A}$ we define

$$\chi^\varepsilon(x,y) := \chi(\bar{x}+\varepsilon(x-\bar{x}), \bar{y}+\varepsilon(y-\bar{y})), \qquad \chi \in \{\mu, \tfrac{1}{2}\sigma^2, c, \tfrac{1}{2}g^2, \sigma g\},$$

where $(\bar{x},\bar{y})$ is a fixed point and $\varepsilon \in [0,1]$. Next, we introduce an operator $\mathcal{A}^{\varepsilon,\rho}$, which is given explicitly by

$$\mathcal{A}^{\varepsilon,\rho} = \mu^\varepsilon(x,y)\partial_x + (\tfrac{1}{2}\sigma^2)^\varepsilon(x,y)\partial_x^2 + c^\varepsilon(x,y)\partial_y + (\tfrac{1}{2}g^2)^\varepsilon(x,y)\partial_y^2 + \rho(\sigma g)^\varepsilon(x,y)\partial_x\partial_y,$$

where $\mathrm{dom}(\mathcal{A}^{\varepsilon,\rho}) := \mathrm{dom}(\mathcal{A})$. Consider, now, a family of PDE problems, indexed by (ε,ρ)

$$0 = (\partial_t + \mathcal{A}^{\varepsilon,\rho})u^{\varepsilon,\rho}, \qquad u^{\varepsilon,\rho}(T,\cdot,\cdot) = \varphi. \tag{2.3.1}$$

Noting that $\mathcal{A}^{\varepsilon,\rho}|_{\varepsilon=1} = \mathcal{A}$, it follows from (2.2.2) and (2.3.1) that $u^{\varepsilon,\rho}|_{\varepsilon=1} = u$. Rather than seeking an approximate solution to PDE problem (2.2.2) directly, we shall instead seek an approximate solution to PDE problem (2.3.1) by expanding $u^{\varepsilon,\rho}$ in powers of ε and ρ as follows

$$u^{\varepsilon,\rho} = \sum_{i=0}^\infty \sum_{j=0}^\infty \varepsilon^i \rho^j u_{i,j}, \tag{2.3.2}$$

where the functions $u_{i,j}$ are (at present) unknown. Once we obtain an approximation for $u^{\varepsilon,\rho}$, we will obtain an approximation for u by setting $\varepsilon = 1$.

We shall see later that the approximation we obtain for u does not require the coefficients of $\mathcal{A}$ to be analytic. However, making the assumption of analytic coefficients simplifies the presentation considerably so we will temporarily proceed with it. We may then write

$$\mathcal{A}^{\varepsilon,\rho} = \sum_{n=0}^\infty \varepsilon^n \left(\mathcal{A}_{n,0} + \rho\mathcal{A}_{n,1}\right), \tag{2.3.3}$$

$$\mathcal{A}_{n,0} = \mu_n(x,y)\partial_x + (\tfrac{1}{2}\sigma^2)_n(x,y)\partial_x^2 + c_n(x,y)\partial_y + (\tfrac{1}{2}g^2)_n(x,y)\partial_y^2, \tag{2.3.4}$$

$$\mathcal{A}_{n,1} = (\sigma g)_n(x,y)\partial_x\partial_y, \tag{2.3.5}$$

where $\text{dom}(\mathcal{A}_{0,0}) := \text{dom}(\mathcal{A})$ and have introduced the notation

$$\chi_n(x,y) := \frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}\varepsilon^n}\chi^\varepsilon(x,y)\Big|_{\varepsilon=0} = \sum_{i=0}^{n} \frac{\partial_{\bar{x}}^i \partial_{\bar{y}}^{n-i}\chi(\bar{x},\bar{y})}{i!(n-i)!}(x-\bar{x})^i(y-\bar{y})^{n-i}, \tag{2.3.6}$$

for $\chi \in \{\mu, \frac{1}{2}\sigma^2, c, \frac{1}{2}g^2, \sigma g\}$. Observe that χ_n is the nth order term in the Taylor series expansion of χ about the point $(\bar{x},\bar{y})$.

Inserting expansions (2.3.2) and (2.3.3) into PDE problem (2.3.1) and collecting terms with like powers of ε and ρ, we obtain

$$\mathcal{O}(1): \qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{0,0}, \qquad u_{0,0}(T,\cdot,\cdot) = \varphi, \tag{2.3.7}$$

$$\mathcal{O}(\varepsilon^n\rho^k): \qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{n,k} + \sum_{i=0}^{n}\sum_{j=0}^{k}(1-\delta_{i+j,0})\mathcal{A}_{i,j}u_{n-i,k-j}, \qquad u_{n,k}(T,\cdot,\cdot) = 0. \tag{2.3.8}$$

For clarity, we present the lowest order terms explicitly here

$$\begin{aligned}
\mathcal{O}(1): &\qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{0,0}, && u_{0,0}(T,\cdot,\cdot) = \varphi,\\
\mathcal{O}(\varepsilon): &\qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{1,0} + \mathcal{A}_{1,0}u_{0,0}, && u_{1,0}(T,\cdot,\cdot) = 0, \quad (2.3.9)\\
\mathcal{O}(\rho): &\qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{0,1} + \mathcal{A}_{0,1}u_{0,0}, && u_{0,1}(T,\cdot,\cdot) = 0, \quad (2.3.10)\\
\mathcal{O}(\varepsilon^2): &\qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{2,0} + \mathcal{A}_{2,0}u_{0,0} + \mathcal{A}_{1,0}u_{1,0}, && u_{2,0}(T,\cdot,\cdot) = 0,\\
\mathcal{O}(\varepsilon\rho): &\qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{1,1}, && u_{1,1}(T,\cdot,\cdot) = 0,\\
&\qquad\quad + \mathcal{A}_{1,1}u_{0,0} + \mathcal{A}_{1,0}u_{0,1} + \mathcal{A}_{0,1}u_{1,0} &&\\
\mathcal{O}(\rho^2): &\qquad 0 = (\partial_t + \mathcal{A}_{0,0})u_{0,2} + \mathcal{A}_{0,1}u_{0,1} + \mathcal{A}_{0,2}u_{0,0}, && u_{0,2}(T,\cdot,\cdot) = 0.
\end{aligned}$$

We note that for any value of ε, the PDE (2.3.1) stipulates that $u^{\varepsilon,\rho}(T,\cdot,\cdot) = \varphi$. This is why we ascribe the terminal condition $u_{0,0}(T,\cdot,\cdot) = \varphi$ and a terminal condition of zero for all higher order terms.

The above computation motivates the following definition.

Definition 2.3.1. Let u be the unique classical solution of PDE problem (2.2.2). We define $\bar{u}_N^\rho$, the *Nth order approximation of* u, as

$$\bar{u}_N^\rho(t,x,y) := \sum_{i=0}^{N}\sum_{j=0}^{i} \varepsilon^j \rho^{i-j} u_{j,i-j}(t,x,y)\Big|_{(\bar{x},\bar{y},\varepsilon)=(x,y,1)}, \tag{2.3.11}$$

where $u_{0,0}$ satisfies (2.3.7) and $u_{n,k}$ satisfies (2.3.8) for $(n,k) \neq (0,0)$.

From classical PDE theory, we get that the zeroth order problem (2.3.7) has a unique classical solution (see [34]).

Remark 2.3.2. Often in asymptotic expansions, one assumes that the solution to a problem is a power series in a small parameter that is inherent in the given problem. Approximations to the the solution of a problem should be depend only on the parameters of the original problem. As ε does not appear in the original PDE problem (2.2.2) it should not appear in the approximation for u. Observe that we have set $\varepsilon = 1$ in (2.3.11), and as a result, the parameter ε plays no role in the definition for $\bar{u}_N^\rho$. We think of this as $\varepsilon = 1$ being the value of ε inherent in our problem. Indeed, ε was introduced merely as an accounting tool in the formal asymptotic expansion above.

Remark 2.3.3. Note that we have set $(\bar{x},\bar{y}) = (x,y)$ in (2.3.11). The small-time behavior of a diffusion is predominantly determined by the local behavior of the diffusion coefficients near the starting point of the diffusion (x,y). In turn, the most accurate Taylor series expansion of any function in a neighborhood of the point (x,y) is the Taylor series expansion centered at $(\bar{x},\bar{y}) = (x,y)$. So first, one should solve the sequence of nested PDE problems (2.3.7)–(2.3.8) with $(\bar{x},\bar{y})$ fixed. For clarity, let us make explicit the dependence of our approximations on the point $(\bar{x},\bar{y})$ by denoting the solution of the $\mathcal{O}(\varepsilon^n\rho^k)$ PDE by $u_{n,k}^{\bar{x},\bar{y}}$. If one is then interested in the approximate value of u at the point (x,y), one should then compute $u_{n,k}^{\bar{x},\bar{y}}(x,y)|_{(\bar{x},\bar{y})=(x,y)}$ in the sum (2.3.11).

As previously mentioned, analyticity of the coefficients of $\mathcal{A}$ is not required. The construction of the Nth order approximation $\bar{u}_N^\rho$ relies only on the operators $\mathcal{A}_{i,j}$ for $i \leq N$ and

$j = 0, 1$. Thus, the Nth order approximation of u requires only that the coefficients of $\mathcal{A}$ belong to $C^N(\mathbb{R}^2)$. This leads us to the proposition.

Proposition 2.3.4. *Fix $N \geq 0$. If the coefficients of $\mathcal{A}$ belong to $C^N(\mathbb{R}^2)$, then the N-th order approximation $\bar{u}_N^\rho$ (2.3.11) exists.*

Proof. Classical results give that the zeroth order problem (2.3.7) has a unique solution $u_{0,0}$ and that $u_{0,0} \in C^{1,2}(\mathbb{R}^2)$. Now, for any $0 \leq i, j$ we have that $\mathcal{A}_{i,j} u_{0,0}$ is continuous. So, classical results give that $u_{1,0}$ and $u_{0,1}$ exist uniquely and that $u_{1,0}, u_{0,1} \in C^{1,2}(\mathbb{R}^2)$. Induction shows that for any $0 \leq n, k$, $u_{n,k}$ exists uniquely and $u_{n,k} \in C^{1,2}(\mathbb{R}^2)$. Therefore, $\bar{u}_N^\rho$ exists. □

2.4 Explicit expressions

In this section, we provide explicit expressions for the functions $u_{n,k}$ required to compute $\bar{u}_N^\rho$, the Nth order approximation of u. We begin this section with a proposition where we give an expressions for the solutions $u_{n,k}$ for $0 \leq n, k$ as integrals in time in terms of the operators $\mathcal{A}_{i,j}$ and the semigroup $\mathcal{P}_{0,0}$ generated by the operator $\mathcal{A}_{0,0}$ on I. The expressions given in the aforementioned proposition are for a general domain I. Afterwards, we examine the European, single-barrier and double-barrier cases, respectively. In each of the respective cases, we give spectral representations for $\mathbb{P}_{0,0}$ and discuss how to use these representations to compute $u_{n,k}$.

Before proceeding, it is helpful to review Duhamel's principle. Let $\Gamma_{0,0}$ be the fundamental solution of parabolic operator $(\partial_t + \mathcal{A}_{0,0})$. That is,

$$0 = (\partial_t + \mathcal{A}_{0,0})\Gamma_{0,0}(\cdot, \cdot, \cdot; T, \xi, \eta), \qquad \Gamma_{0,0}(T, \cdot, \cdot; T, \xi, \eta) = \delta_{\xi,\eta}.$$

Duhamel's principle states that the the unique classical solution to

$$0 = (\partial_t + \mathcal{A}_{0,0})u + f, \qquad u(T, \cdot, \cdot) = h,$$

is given by

$$u(t, x, y) = \mathcal{P}_{0,0}(t, T)h(x, y) + \int_t^T \mathrm{d}s \, \mathcal{P}_{0,0}(t, s) f(s, x, y),$$

where we have introduced $\mathcal{P}_{0,0}$ the *semigroup* generated by $\mathcal{A}_{0,0}$, which is defined as follows

$$\mathcal{P}_{0,0}(t,s)\varphi(x,y) = \int_I \mathrm{d}\xi \int_{\mathbb{R}} \mathrm{d}\eta \, \Gamma_{0,0}(t,x,y;s,\xi,\eta)\varphi(\xi,\eta), \tag{2.4.1}$$

where $0 \leq t \leq s \leq T$.

We are now ready to give an expressions for the functions $u_{n,k}$ for $0 \leq n, k$

Proposition 2.4.1. *Fix $N \geq 0$ and suppose that the coefficients of $\mathcal{A}$ belong to $C^N(\mathbb{R}^2)$, and let the functions $\{u_{n,k}\}_{0\leq n,k\leq N}$ be the the unique classical solution of the nested sequence of PDE problems (2.3.7)–(2.3.8). Then, omitting the spacial arguments (x,y) to ease notation, the function $u_{0,0}$ is given by*

$$u_{0,0}(t) = \mathcal{P}_{0,0}(t,T)\varphi, \tag{2.4.2}$$

and for $(n,k) \in \{(i,j) \mid 0 \leq i,j \leq N\} \setminus \{(0,0)\}$, we have

$$u_{n,k}(t) = \sum_{j=1}^{n+k} \sum_{I_{n,k,j}} \int_t^T \mathrm{d}s_1 \int_{s_1}^T \mathrm{d}s_2 \cdots \int_{s_{j-1}}^T \mathrm{d}s_j \tag{2.4.3}$$
$$\mathcal{P}_{0,0}(t,s_1)\mathcal{A}_{n_1,k_1}\mathcal{P}_{0,0}(s_1,s_2)\mathcal{A}_{n_2,k_2} \cdots \mathcal{P}_{0,0}(s_{j-1},s_j)\mathcal{A}_{n_j,k_j}\mathcal{P}_{0,0}(s_j,T)\varphi,$$

with $I_{n,k,j}$ given by

$$I_{n,k,j} = \left\{ \begin{pmatrix} n_1, \ldots, n_j \\ k_1, \ldots, k_j \end{pmatrix} \in \mathbb{N}_0^{2\times j} \;\middle|\; \begin{array}{l} n_1 + \cdots + n_j = n, \\ k_1 + \cdots + k_j = k, \\ 1 \leq n_i + k_i, \text{ for all } 1 \leq i \leq j \end{array} \right\}.$$

Proof of Proposition 2.4.1. We proceed by induction. Suppose $(n,k) = (0,0)$. The operator $\mathcal{A}_{0,0}$ has constant coefficients so the existence of the semi-group operator $\mathcal{P}_{0,0}$ generated by $\mathcal{A}_{0,0}$ on I is a standard result in PDE theory. The formula (2.4.2) is a straightforward application of Duhamel's principle.

We now note that formula (2.4.3) holds for $(n,k) = (1,0)$ and $(n,k) = (0,1)$ by applying Duhamel's principle to (2.3.9) and (2.3.10). Next, assume as an inductive hypothesis that for non-negative integers n and k such that $n+k \geq 1$ formula (2.4.3) holds for pairs of non-negative integers (m,j) such that $m+j \leq n+k$. For an integer b, define the set

$$A^b_{n,k} := \{(i,j) \mid 0 \leq i \leq n,\, 0 \leq j \leq k,\, 1 \leq i+j \leq b\}.$$

Applying Duhamel's principle to (2.3.8), we see that

$$\begin{aligned}
u_{n+1,k} &= \int_t^T \mathrm{d}s \mathcal{P}_{0,0}(t,s) \left(\sum_{i=0}^{n+1} \sum_{j=0}^{k} (1-\delta_{i+j,0}) \mathcal{A}_{i,j} u_{n-i+1,k-j} \right) \\
&= \int_t^T \mathrm{d}s \mathcal{P}_{0,0}(t,s) \left(\sum_{\substack{(i,j)\in \\ A^{n+k+1}_{n+1,k}}} \mathcal{A}_{i,j} u_{n-i+1,k-j} \right) \\
&= \int_t^T \mathrm{d}s \mathcal{P}_{0,0}(t,s) \mathcal{A}_{n+1,k} \mathcal{P}_{0,0}(s,T)\varphi \qquad (2.4.4) \\
&\quad + \sum_{\substack{(i,j)\in \\ A^{n+k+1}_{n+1,k}}} \sum_{l=1}^{n+k-i-j+1} \sum_{I_{n-i+1,k-j,l}} \int_t^T \mathrm{d}s \int_s^T \mathrm{d}s_1 \cdots \int_{s_{l-1}}^T \mathrm{d}s_l \\
&\quad\quad \mathcal{P}_{0,0}(t,s)\mathcal{A}_{i,j}\mathcal{P}_{0,0}(s,s_1)\mathcal{A}_{(n-i+1)_1,(k-j)_1} \cdots \mathcal{P}_{0,0}(s_{l-1},s_l)\mathcal{A}_{(n-i+1)_l,(k-j)_l}\mathcal{P}_{0,0}(s_l,T)\varphi,
\end{aligned}$$

where (2.4.4) follows from our inductive hypothesis. Reordering the sums in (2.4.4) we obtain

$$\begin{aligned}
u_{n+1,k} &= \int_t^T \mathrm{d}s \mathcal{P}_{0,0}(t,s) \mathcal{A}_{n+1,k} \mathcal{P}_{0,0}(s,T)\varphi \qquad (2.4.5) \\
&\quad + \sum_{l=1}^{n+k} \sum_{\substack{(i,j)\in \\ A^{n+k-\ell+1}_{n+1,k}}} \sum_{I_{n-i+1,k-j,l}} \int_t^T \mathrm{d}s \int_s^T \mathrm{d}s_1 \cdots \int_{s_{l-1}}^T \mathrm{d}s_l \\
&\quad\quad \mathcal{P}_{0,0}(t,s)\mathcal{A}_{i,j}\mathcal{P}_{0,0}(s,s_1)\mathcal{A}_{(n-i+1)_1,(k-j)_1} \cdots \mathcal{P}_{0,0}(s_{l-1},s_l)\mathcal{A}_{(n-i+1)_l,(k-j)_l}\mathcal{P}_{0,0}(s_l,T)\varphi.
\end{aligned}$$

Next, note that

$$I_{n+1,k,\ell} = \bigcup_{(i,j)\in A^{n+k-\ell+1}_{n+1,k}} \left\{ \begin{pmatrix} i & n_1 & n_2 & \cdots & n_{\ell-1} \\ j & k_1 & k_2 & \cdots & k_{\ell-1} \end{pmatrix} \,\middle|\, \begin{pmatrix} n_1 & n_2 & \cdots & n_{\ell-1} \\ k_1 & k_2 & \cdots & k_{\ell-1} \end{pmatrix} \in I_{n-i+1,k-j,\ell-1} \right\}. \qquad (2.4.6)$$

Therefore, combining (2.4.6) with (2.4.5) we obtain

$$u_{n+1,k} = \int_t^T \mathrm{d}s \mathcal{P}_{0,0}(t,s)\mathcal{A}_{n+1,k}\mathcal{P}_{0,0}(s,T)\varphi + \sum_{l=1}^{n+k} \sum_{I_{n+1,k,l+1}} \int_t^T \mathrm{d}s \int_s^T \mathrm{d}s_1 \cdots \int_{s_{l-1}}^T \mathrm{d}s_l$$

$$\mathcal{P}_{0,0}(t,s)\mathcal{A}_{(n+1)_1,k_1}\mathcal{P}_{0,0}(s,s_1)\mathcal{A}_{(n+1)_2,k_2} \cdots \mathcal{P}_{0,0}(s_{l-1},s_l)\mathcal{A}_{(n+1)_{l+1},k_{l+1}}\mathcal{P}_{0,0}(s_l,T)\varphi.$$

Relabeling $(s, s_1, s_2, \cdots, s_\ell) \to (s_1, s_2, \cdots, s_{\ell+1})$ and reindexing and gives

$$u_{n+1,k} = \int_t^T \mathrm{d}s_1 \mathcal{P}_{0,0}(t,s_1)\mathcal{A}_{n+1,k}\mathcal{P}_{0,0}(s_1,T)\varphi + \sum_{l=2}^{n+k+1} \sum_{I_{n+1,k,l}} \int_t^T \mathrm{d}s_1 \int_{s_1}^T \mathrm{d}s_2 \cdots \int_{s_{l-1}}^T \mathrm{d}s_l$$

$$\mathcal{P}_{0,0}(t,s_1)\mathcal{A}_{(n+1)_1,k_1}\mathcal{P}_{0,0}(s_1,s_2)\mathcal{A}_{(n+1)_2,k_2} \cdots \mathcal{P}_{0,0}(s_{l-1},s_l)\mathcal{A}_{(n+1)_l,k_l}\mathcal{P}_{0,0}(s_l,T)\varphi$$

$$= \sum_{l=1}^{n+k+1} \sum_{I_{n+1,k,l}} \int_t^T \mathrm{d}s_1 \int_{s_1}^T \mathrm{d}s_2 \cdots \int_{s_\ell}^T \mathrm{d}s_l$$

$$\mathcal{P}_{0,0}(t,s_1)\mathcal{A}_{(n+1)_1,k_1}\mathcal{P}_{0,0}(s_1,s_2)\mathcal{A}_{(n+1)_2,k_2} \cdots \mathcal{P}_{0,0}(s_{l-1},s_l)\mathcal{A}_{(n+1)_l,k_l}\mathcal{P}_{0,0}(s_l,T)\varphi,$$

which is (2.4.3) for the case $(n+1,k)$. The proof for the case $(n,k+1)$ is analogous. □

2.4.1 European claims

In this section, we consider the case $I = \mathbb{R}$. As $\tau = \infty$ when $I = \mathbb{R}$, we see from (2.2.1) that this case corresponds to a European claim written on X. We begin with the following lemma.

Lemma 2.4.2. *Let $\mathcal{H}$ be the linear operator*

$$\mathcal{H} = b\partial_x + a\partial_x^2, \qquad \mathrm{dom}(\mathcal{H}) = C^2(\mathbb{R}).$$

Then

$$\mathcal{H}\psi_\omega = \lambda_\omega\psi_\omega, \qquad \omega \in \mathbb{R},$$

$$\psi_\omega(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathtt{i}\omega x}, \qquad \lambda_\omega(b,a) = b\mathtt{i}\omega - a\omega^2. \tag{2.4.7}$$

Moreover, we have

$$\langle\psi_\omega, \psi_\gamma\rangle = \delta(\omega - \gamma), \qquad \langle f, g\rangle := \int_{\mathbb{R}} \mathrm{d}x\, \overline{f}(x)g(x), \tag{2.4.8}$$

where $\delta(\omega - \gamma)$ is a Dirac delta function and $\overline{f}$ denotes the complex conjugate of f.

Proof. The lemma is checked by direct computation. □

Proposition 2.4.3. *Let $\mathcal{P}_{0,0}$ be the semigroup generated by $\mathcal{A}_{0,0}$ with* $\mathrm{dom}(\mathcal{A}_{0,0}) = C^2(\mathbb{R}^2)$*. Then we have*

$$\mathcal{P}_{0,0}(t,T)f = \int_{\mathbb{R}} \mathrm{d}\omega \int_{\mathbb{R}} \mathrm{d}\gamma \, \mathrm{e}^{\Lambda_{\omega,\gamma}(T-t)} \left\langle \Psi_{\omega,\gamma}, f \right\rangle \Psi_{\omega,\gamma}, \tag{2.4.9}$$

$$\langle f, g \rangle := \int_{\mathbb{R}} \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y \, \overline{f}(x,y) g(x,y),$$

where $\Psi_{\omega,\gamma}$ and $\Lambda_{\omega,\gamma}$ are given by

$$\Psi_{\omega,\gamma}(x,y) = \psi_\omega(x)\psi_\gamma(y), \qquad \Lambda_{\omega,\gamma} = \lambda_\omega(b_1,a_1) + \lambda_\gamma(b_2,a_2),$$

$$(b_1,a_1) = \left(\mu_0, (\tfrac{1}{2}\sigma^2)_0\right), \qquad (b_2,a_2) = \left(c_0, (\tfrac{1}{2}g^2)_0\right),$$

with ψ_ω and $\lambda_\omega(b,a)$ as defined in (2.4.7)*.*

Proof. Using Lemma 2.4.2 and the fact that

$$\int_{\mathbb{R}} \mathrm{d}\omega \, \overline{\psi}_\omega(x)\psi_\omega(y) = \delta(x-y),$$

one can check by direct computation that

$$\Gamma_{0,0}(t,x,y;T;\xi,\eta) := \int_{\mathbb{R}} \mathrm{d}\omega \int_{\mathbb{R}} \mathrm{d}\gamma \, \mathrm{e}^{\Lambda_{\omega,\gamma}(T-t)} \Psi_{\omega,\gamma}(x,y)\Psi_{\omega,\gamma}(\xi,\eta), \tag{2.4.10}$$

is the fundamental solution of $\partial_t + \mathcal{A}_{0,0}$. Expression (2.4.9) follows directly by inserting (2.4.10) into (2.4.1). □

Now, from Proposition 2.4.1, we see that the $(u_{n,k})$ are represented by a sum of terms of the form

$$A = \left(\prod_{i=1}^{j} \int_{s_{i-1}}^{T} \mathrm{d}s_i \, \mathcal{P}_{0,0}(s_{i-1},s_i)\mathcal{A}_{n_i,k_i} \right) \mathcal{P}_{0,0}(s_j,T)\varphi \qquad s_0 = t. \tag{2.4.11}$$

Using (2.4.9), we can write these terms as

$$A = \int_{\mathbb{R}} \mathrm{d}\omega_{j+1} \int_{\mathbb{R}} \mathrm{d}\gamma_{j+1} \left(\prod_{i=1}^{j} \int_{s_{i-1}}^{T} \mathrm{d}s_i \int_{\mathbb{R}} \mathrm{d}\omega_i \int_{\mathbb{R}} \mathrm{d}\gamma_i \, \mathrm{e}^{\Lambda_{\omega_i,\gamma_i}(s_i - s_{i-1})} \left\langle \Psi_{\omega_i,\gamma_i}, \mathcal{A}_{n_i,k_i} \Psi_{\omega_{i+1},\gamma_{i+1}} \right\rangle \right)$$

$$\left\langle \Psi_{\omega_{j+1},\gamma_{j+1}}, \varphi \right\rangle \mathrm{e}^{\Lambda_{\omega_{j+1},\gamma_{j+1}}(T-s_j)} \Psi_{\omega_1,\gamma_1}. \tag{2.4.12}$$

Although the multiple integral may seem unwieldy, we shall see that all but a single integral collapses when we compute the elements

$$\left\langle \Psi_{\omega_i,\gamma_i}, \mathcal{A}_{n_i,k_i} \Psi_{\omega_{i+1},\gamma_{i+1}} \right\rangle,$$

which appear in (2.4.12)

Lemma 2.4.4. *Let $\Psi_{\omega,\gamma}$ and $\langle \cdot, \cdot \rangle$ be as defined in Proposition 2.4.3. Define the operator*

$$\mathcal{B} := x^i y^j \partial_x^k \partial_y^l, \qquad i, j, k, l \in \mathbb{N}_0. \tag{2.4.13}$$

Then we have

$$\langle \Psi_{\omega',\gamma'}, \mathcal{B}\Psi_{\omega,\gamma} \rangle = (\mathtt{i}\omega)^k (\mathtt{i}\gamma)^l (-\mathtt{i}\partial_\omega)^i (-\mathtt{i}\partial_\gamma)^j \delta(\omega - \omega')\delta(\gamma - \gamma').$$

Proof. The proof is a straightforward computation. Recalling that $\Psi_{\omega,\gamma} = \frac{1}{2\pi} \mathrm{e}^{\mathtt{i}\omega x + \mathtt{i}\gamma y}$, we have

$$\begin{aligned}
\langle \Psi_{\omega',\gamma'}, \mathcal{B}\Psi_{\omega,\gamma} \rangle &= \int_{\mathbb{R}} \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y \, \overline{\Psi}_{\omega',\gamma'}(x, y) x^i y^j \partial_x^k \partial_y^l \Psi_{\omega,\gamma}(x, y) \\
&= (\mathtt{i}\omega)^k (\mathtt{i}\gamma)^l \int_{\mathbb{R}} \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y \, \overline{\Psi}_{\omega',\gamma'}(x, y) x^i y^j \Psi_{\omega,\gamma}(x, y) \\
&= (\mathtt{i}\omega)^k (\mathtt{i}\gamma)^l (-\mathtt{i}\partial_\omega)^i (-\mathtt{i}\partial_\gamma)^j \int_{\mathbb{R}} \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y \, \overline{\Psi}_{\omega',\gamma'}(x, y) \Psi_{\omega,\gamma}(x, y) \\
&= (\mathtt{i}\omega)^k (\mathtt{i}\gamma)^l (-\mathtt{i}\partial_\omega)^i (-\mathtt{i}\partial_\gamma)^j \delta(\omega - \omega')\delta(\gamma - \gamma').
\end{aligned}$$

where, in the last step, we have used (2.4.8). □

Remark 2.4.5. The derivative of a Dirac delta function is defined as follows:

$$\int_{\mathbb{R}} \mathrm{d}\omega \, f(\omega) \partial_\omega^n \delta(\omega - \omega') = \int_{\mathbb{R}} \mathrm{d}\omega \, \delta(\omega - \omega')(-\partial_\omega)^n f(\omega) = (-\partial_{\omega'})^n f(\omega'),$$

where we have integrated by parts.

Note from (2.3.4), (2.3.5) and (2.3.6) that the operators $(\mathcal{A}_{n,k})$ are finite linear combinations of operators of the form (2.4.13). Thus, in light of Lemma 2.4.4 and Remark 2.4.5, we see that the integrals in (2.4.12) with respect to ω_i and γ_i, for $i = 1, 2, \ldots j$ collapse due to the Dirac delta functions. The integral with respect to γ_{j+1} also collapses, due to the fact that the payoff function φ does not depend on y. And the iterated integrals with respect to s_i for $i = 1, 2, \ldots j$ involve only exponentials and are explicitly computable. Thus, what remains is the integral with respect to ω_{j+1}, which, in general, must be computed numerically. We note that in the special case $\varphi(x) = x^n \mathrm{e}^{px}$ for some $n \in \mathbb{N}_0$ and $p \in \mathbb{R}$, the integral with respect to ω_{j+1} can be evaluated analytically.

2.4.2 Single-barrier claims

In this section, we consider the case $I = (L, \infty)$, which corresponds to a single-barrier knock-out claim written on X with a lower barrier $L < X_0$. The case $I = (-\infty, U)$ with $U > X_0$ can be handled analogously. We begin with the following lemma.

Lemma 2.4.6. *Let $\mathcal{H}$ be the following linear operator*

$$\mathcal{H} = b\partial_x + a\partial_x^2, \qquad \mathrm{dom}(\mathcal{H}) = \{f \in C^2(I) \mid \lim_{x \downarrow L} f(x) = 0\}, \qquad I = (L, \infty).$$

The following holds

$$\begin{aligned} \mathcal{H}\eta_\omega &= \mu_\omega \eta_\omega, & \omega &\in \mathbb{R}_+, \\ \eta_\omega(x; b, a) &= \sqrt{\frac{2}{\pi}} \mathrm{e}^{-bx/(2a)} \sin\Big(\omega(x - L)\Big), & \mu_\omega(b, a) &= -\frac{b^2}{4a} - a\omega^2. \end{aligned} \tag{2.4.14}$$

Moreover, we have

$$\langle \eta_\omega, \eta_\gamma \rangle = \delta(\omega - \gamma), \qquad \langle f, g \rangle := \int_I \mathrm{d}x\, f(x) g(x) \mathfrak{m}(x), \qquad \mathfrak{m}(x) = \mathrm{e}^{bx/a}. \tag{2.4.15}$$

Here, $\delta(\omega - \gamma)$ is a Dirac delta function.

Proof. The lemma is checked by direct computation. □

Proposition 2.4.7. *Let $\mathcal{P}_{0,0}$ the semigroup generated by $\mathcal{A}_{0,0}$ with*

$$\mathrm{dom}(\mathcal{A}_{0,0}) = \{f \in C^2(E) \mid \lim_{x \downarrow L} f(x,y) = 0\}, \qquad E = (L,\infty) \times \mathbb{R}.$$

Then

$$\mathcal{P}_{0,0}(t,T)f = \int_{\mathbb{R}_+} \mathrm{d}\omega \int_{\mathbb{R}} \mathrm{d}\gamma \, \mathrm{e}^{\Lambda_{\omega,\gamma}(T-t)} \langle \Psi_{\omega,\gamma}, f \rangle \, \Psi_{\omega,\gamma}, \tag{2.4.16}$$

$$\langle f, g \rangle := \int_I \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y \, \overline{f}(x,y) g(x,y) \mathfrak{m}(x), \qquad I = (L,\infty),$$

where $\Psi_{\omega,\gamma}$ and $\Lambda_{\omega,\gamma}$ are given by

$$\Psi_{\omega,\gamma}(x,y) = \eta_\omega(x; b_1, a_1)\psi_\gamma(y), \qquad \Lambda_{\omega,\gamma} = \mu_\omega(b_1,a_1) + \lambda_\gamma(b_2,a_2),$$

$$(b_1,a_1) = \left(\mu_0, (\tfrac{1}{2}\sigma^2)_0\right), \qquad (b_2,a_2) = \left(c_0, (\tfrac{1}{2}g^2)_0\right),$$

with ψ_γ and λ_γ as defined in (2.4.7), η_ω and γ_ω as defined in (2.4.14) and $\mathfrak{m}$ as defined in (2.4.15).

Proof. Using Lemma 2.4.6, we check by direct computation that

$$\Gamma_{0,0}(t,x,y;T;\xi,\eta) := \int_{\mathbb{R}_+} \mathrm{d}\omega \int_{\mathbb{R}} \mathrm{d}\gamma \, \mathrm{e}^{\Lambda_{\omega,\gamma}(T-t)} \Psi_{\omega,\gamma}(x,y)\Psi_{\omega,\gamma}(\xi,\eta)\mathfrak{m}(\xi), \tag{2.4.17}$$

is the fundamental solution of $\partial_t + \mathcal{A}_{0,0}$. Expression (2.4.16) follows directly by inserting (2.4.17) into (2.4.1). □

Once again, to compute the $(u_{n,k})$, we must examine terms of the form (2.4.11). Using (2.4.16), we write these terms as

$$A = \int_{\mathbb{R}} \mathrm{d}\omega_{j+1} \int_{\mathbb{R}} \mathrm{d}\gamma_{j+1} \left(\prod_{i=1}^{j} \int_{s_{i-1}}^{T} \mathrm{d}s_i \int_{\mathbb{R}} \mathrm{d}\omega_i \int_{\mathbb{R}} \mathrm{d}\gamma_i \, \mathrm{e}^{\Lambda_{\omega_i,\gamma_i}(s_i - s_{i-1})} \left\langle \Psi_{\omega_i,\gamma_i}, \mathcal{A}_{n_i,k_i} \Psi_{\omega_{i+1},\gamma_{i+1}} \right\rangle \right)$$
$$\left\langle \Psi_{\omega_{j+1},\gamma_{j+1}}, \varphi \right\rangle \mathrm{e}^{\Lambda_{\omega_{j+1},\gamma_{j+1}}(T-s_j)} \Psi_{\omega_1,\gamma_1}, \tag{2.4.18}$$

where Ψ_{ω_i,γ_i} and $\Lambda_{\omega_i,\gamma_i}$ are as in Proposition 2.4.7. Noting that each $\mathcal{A}_{n,k}$ can be expressed as a finite linear combination of operators with the form of $\mathcal{B}$, which was defined in (2.4.13), we must compute inner products of the form

$$\left\langle \Psi_{\omega',\gamma'}, \mathcal{B}\Psi_{\omega,\gamma} \right\rangle .$$

This motivates the following lemma.

Lemma 2.4.8. *Let* $\Psi_{\omega,\gamma}$ *and* $\langle\cdot,\cdot\rangle$ *be as defined in Proposition 2.4.7 and* $\mathcal{B}$ *be defined as in* (2.4.13)*. Then*

$$\langle \Psi_{\omega',\gamma'}, \mathcal{B}\Psi_{\omega,\gamma}\rangle = (\mathtt{i}\gamma)^l(-\mathtt{i}\partial_\gamma)^j\delta(\gamma-\gamma')C_{\omega',\omega,i,k}, \tag{2.4.19}$$

where

$$\begin{aligned} C_{\omega',\omega,i,k} &= \sum_{m=0}^{i}\binom{i}{m}L^{i-m}C^{(1)}_{\omega',\omega,m,k}, \\ C^{(1)}_{\omega',\omega,m,k} &= \frac{m!}{\pi}c^{(e)}_{\omega,k}\left((\omega+\omega')^{-m-1}-|\omega-\omega'|^{-m-1}\right)\sin\left(\frac{m\pi}{2}\right) \\ &\quad + \frac{m!}{\pi}c^{(o)}_{\omega,k}\left((\omega+\omega')^{-m-1}-\operatorname{sign}(\omega-\omega')|\omega-\omega'|^{-m-1}\right)\cos\left(\frac{m\pi}{2}\right), \quad (2.4.20) \\ c^{(o)}_{\omega,k} &= \sum_{m=1}^{\lfloor\frac{k+1}{2}\rfloor}\binom{k}{2m-1}(-1)^{k-m}\left(\frac{b}{2a}\right)^{k-2m+1}\omega^{2m-1}, \quad (2.4.21) \\ c^{(e)}_{\omega,k} &= \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\binom{k}{2m}(-1)^{k-m}\left(\frac{b}{2a}\right)^{k-2m}\omega^{2m}. \quad (2.4.22) \end{aligned}$$

Proof. Recalling the definition of Ψ from Proposition 2.4.3, we compute

$$\begin{aligned} \langle \Psi_{\omega',\gamma'}, \mathcal{B}\Psi_{\omega,\gamma}\rangle &= \int_I \mathrm{d}x\int_{\mathbb{R}}\mathrm{d}y\ \mathfrak{m}(x)\overline{\Psi}_{\omega',\gamma'}(x,y)x^iy^j\partial_x^k\partial_y^l\Psi_{\omega,\gamma}(x,y) \\ &= (\mathtt{i}\gamma)^l(-\mathtt{i}\partial_\gamma)^j\delta(\gamma-\gamma')\int_I\mathrm{d}x\ \mathfrak{m}(x)x^i\eta_{\omega'}(x)\partial_x^k\eta_\omega(x). \end{aligned} \tag{2.4.23}$$

By direct computation, we find that

$$\partial_x^k\eta_\omega(x) = \sqrt{\frac{2}{\pi}}\mathrm{e}^{-\frac{b}{2a}x}\left(c^{(o)}_{\omega,k}\cos\left(\omega(x-L)\right)+c^{(e)}_{\omega,k}\sin\left(\omega(x-L)\right)\right),$$

where $c^{(o)}_{\omega,k}$ and $c^{(e)}_{\omega,k}$ are given by (2.4.21) and (2.4.22), respectively. Thus,

$$\begin{aligned} \int_I\mathrm{d}x\ \mathfrak{m}(x)x^i\eta_{\omega'}(x)\partial_x^k\eta_\omega(x) &= \frac{2}{\pi}\int_I\mathrm{d}x\ x^i\sin\left(\omega'(x-L)\right)\left(c^{(o)}_{\omega,k}\cos\left(\omega(x-L)\right)+c^{(e)}_{\omega,k}\sin\left(\omega(x-L)\right)\right) \\ &= \sum_{m=0}^{i}\binom{i}{m}L^{i-m}\frac{2}{\pi}\int_0^\infty\mathrm{d}x\ x^m\sin(\omega'x)\left(c^{(o)}_{\omega,k}\cos(\omega x)+c^{(e)}_{\omega,k}\sin(\omega x)\right) \end{aligned}$$

$$= \sum_{m=0}^{i} \binom{i}{m} L^{i-m} C^{(1)}_{\omega',\omega,m,k}, \tag{2.4.24}$$

where $C^{(1)}_{\omega',\omega,m,k}$ is given by (2.4.20). Inserting (2.4.24) into (2.4.23) gives (2.4.19). □

Note from (2.3.4), (2.3.5) and (2.3.6), the operators $\mathcal{A}_{n,k}$ are finite linear combinations of operators of the form (2.4.13). We see from (2.4.19) that the integrals with respect to γ_i, $i = 1, 2, 3, \cdots, j$ in (2.4.18) collapse due to the Dirac delta functions. As φ is independent of y, the integral with respect to γ_{j+1} also collapses. Furthermore, the iterated integrals with respect to s_i, $i = 1, 2, 3, \cdots j$ in (2.4.18) involve only exponentials in s_i and can therefore be evaluated explicitly. We are left only with integrals with respect to ω_i, $i = 1, 2, 3, \cdots, j+1$, which can be evaluated numerically.

2.4.3 Double-barrier claims

In this section, we consider the case $I = (L, U)$, which corresponds to a double-barrier knock-out claim written on X with a barriers L and U satisfying $L < X_0 < U$. We begin with the following lemma.

Lemma 2.4.9. *Let $\mathcal{H}$ be the following linear operator*

$$\mathcal{H} = b\partial_x + a\partial_x^2, \qquad \text{dom}(\mathcal{H}) = \{f \in C^2(I) \mid \lim_{x\downarrow L} f(x) = 0, \lim_{x\uparrow U} f(x) = 0\}, \qquad I = (L, U).$$

The following holds

$$\mathcal{H}\phi_\ell = \nu_\ell \phi_\ell, \qquad \ell \in \mathbb{N}$$
$$\phi_\ell(x; b, a) = \sqrt{\frac{2}{U-L}} \mathrm{e}^{-\frac{bx}{2a}} \sin\left(\frac{\pi\ell(x-L)}{U-L}\right), \qquad \nu_\ell(b, a) = -\frac{b^2}{4a} - \frac{a\pi^2\ell^2}{(U-L)^2} \tag{2.4.25}$$

Moreover, we have

$$\langle \phi_\ell, \phi_k \rangle = \delta_{\ell,k}, \qquad \langle f, g \rangle := \int_I \mathrm{d}x\, \mathfrak{m}(x) f(x) g(x).$$

Here, $\delta_{\ell,k}$ is a Kronecker delta function and $\mathfrak{m}$ is given by (2.4.15).

Proof. The lemma can be checked by direct computation. □

Proposition 2.4.10. *Let $\mathcal{P}_{0,0}$ the semigroup generated by $\mathcal{A}_{0,0}$ with*

$$\text{dom}(\mathcal{A}_{0,0}) = \{f \in C^2(E) \mid \lim_{x \downarrow L} f(x,y) = 0, \lim_{x \uparrow U} f(x,y) = 0\}, \qquad E = (L,R) \times \mathbb{R}.$$

Then we have

$$\mathcal{P}_{0,0}(t,T)f = \sum_{\ell=1}^{\infty} \int_{\mathbb{R}} \mathrm{d}\gamma \, \mathrm{e}^{\Lambda_{\ell,\gamma}(T-t)} \langle \Psi_{\ell,\gamma}, f \rangle \, \Psi_{\ell,\gamma}, \tag{2.4.26}$$

$$\langle f,g \rangle := \int_I \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y \, \overline{f}(x,y) g(x,y) \mathfrak{m}(x), \qquad I = (L,U),$$

where $\Psi_{\ell,\gamma}$ and $\Lambda_{\ell,\gamma}$ are given by

$$\Psi_{\ell,\gamma}(x,y) = \phi_\ell(x; b_1, a_1)\psi_\gamma(y), \qquad \Lambda_{\ell,\gamma} = \nu_\ell(b_1,a_1) + \lambda_\gamma(b_2,a_2),$$

$$(b_1,a_1) = \left(\mu_0, (\tfrac{1}{2}\sigma^2)_0\right), \qquad (b_2,a_2) = \left(c_0, (\tfrac{1}{2}g^2)_0\right),$$

and ψ_γ and $\lambda_\gamma(b,a)$ are given in (2.4.7), ϕ_ℓ and ν_ℓ are given in (2.4.25) and $\mathfrak{m}$ is given in (2.4.15).

Proof. Using Lemma 2.4.9, we check by direct computation that

$$\Gamma_{0,0}(t,x,y;T;\xi,\eta) := \sum_{\ell=1}^{\infty} \int_{\mathbb{R}} \mathrm{d}\gamma \, \mathrm{e}^{\Lambda_{\ell,\gamma}(T-t)} \Psi_{\ell,\gamma}(x,y) \Psi_{\ell,\gamma}(\xi,\eta) \mathfrak{m}(\xi), \tag{2.4.27}$$

is the fundamental solution of $\partial_t + \mathcal{A}_{0,0}$. Expression (2.4.26) follows directly by inserting (2.4.27) into (2.4.1). □

As with the European and single-barrier cases, to compute the functions $(u_{n,k})$ we must evaluate terms of the form (2.4.11). Using (2.4.26) we write these terms as

$$A = \sum_{\ell_{j+1}=1}^{\infty} \int_{\mathbb{R}} \mathrm{d}\gamma_{j+1} \left(\prod_{i=1}^{j} \int_{s_{i-1}}^{T} \mathrm{d}s_i \sum_{\ell_i=1}^{\infty} \int_{\mathbb{R}} \mathrm{d}\gamma_i \, \mathrm{e}^{\Lambda_{\ell_i,\gamma_i}(s_i - s_{i-1})} \left\langle \Psi_{\ell_i,\gamma_i}, \mathcal{A}_{n_i,k_i} \Psi_{\ell_{i+1},\gamma_{i+1}} \right\rangle \right)$$
$$\left\langle \Psi_{\ell_{j+1},\gamma_{j+1}}, \varphi \right\rangle \mathrm{e}^{\Lambda_{\ell_{j+1},\gamma_{j+1}}(T-s_j)} \Psi_{\omega_1,\gamma_1}. \tag{2.4.28}$$

As each $\mathcal{A}_{n,k}$ is a finite linear combination of operators form of $\mathcal{B}$, which is defined in (2.4.13), we must compute terms of the form $\left\langle \Psi_{\ell_i,\gamma_i}, \mathcal{B}\Psi_{\ell_{i+1},\gamma_{i+1}} \right\rangle$.

Lemma 2.4.11. *Let* $\Psi_{\ell,\gamma}$ *and* $\langle\cdot,\cdot\rangle$ *be as defined in Proposition 2.4.10 and* $\mathcal{B}$ *be as defined in* (2.4.13)*. Then*

$$\langle\Psi_{\ell',\gamma'},\mathcal{B}\Psi_{\ell,\gamma}\rangle=(\mathtt{i}\gamma)^{l}(-\mathtt{i}\partial_\gamma)^{j}\delta(\gamma-\gamma')C_{\ell',\ell,i,k}, \tag{2.4.29}$$

where

$$C_{\ell',\ell,i,k}=\sum_{m=0}^{i}\binom{i}{m}L^{i-m}\left(\frac{U-L}{\pi}\right)^{m+1}\left(C^{(1)}_{\ell',\ell,m,k}\mathbb{I}_{\ell'\neq\ell}+C^{(2)}_{\ell,m,k}\delta_{\ell',\ell}\right),$$

and

$$\begin{aligned}C^{(1)}_{\ell',\ell,m,k}&=\frac{1}{4}\pi^{m+\frac{3}{2}}\Gamma_E\left(\frac{m+1}{2}\right)c^{(e)}_{\ell,k}\,{}_1\widetilde{F}_2\left(\frac{m+1}{2};\frac{1}{2},\frac{m+3}{2};-\frac{1}{4}(\ell-\ell')^2\pi^2\right) &(2.4.30)\\&-\frac{1}{4}\pi^{m+\frac{3}{2}}\Gamma_E\left(\frac{m+1}{2}\right)c^{(e)}_{\ell,k}\,{}_1\widetilde{F}_2\left(\frac{m+1}{2};\frac{1}{2},\frac{m+3}{2};-\frac{1}{4}(\ell+\ell')^2\pi^2\right)\\&+\frac{\pi^{m+2}}{2(m+2)}\left(\ell'c^{(o)}_{\ell,k}-\ell c^{(o)}_{\ell,k}\right)\,{}_1F_2\left(\frac{m}{2}+1;\frac{3}{2},\frac{m}{2}+2;-\frac{1}{4}(\ell-\ell')^2\pi^2\right)\\&+\frac{\pi^{m+2}}{2(m+2)}\left(\ell'c^{(o)}_{\ell,k}+\ell c^{(o)}_{\ell,k}\right)\,{}_1F_2\left(\frac{m}{2}+1;\frac{3}{2},\frac{m}{2}+2;-\frac{1}{4}(\ell+\ell')^2\pi^2\right),\\C^{(2)}_{\ell,m,k}&=\frac{\pi^{m+2}\ell c^{(o)}_{l,k}}{m+2}\,{}_1F_2\left(\frac{m}{2}+1;\frac{3}{2},\frac{m}{2}+2;-\pi^2\ell^2\right) &(2.4.31)\\&-\frac{2\pi^{m+3}\ell^2c^{(e)}_{l,k}}{m^2+4m+3}\,{}_1F_2\left(\frac{m}{2}+\frac{3}{2};\frac{3}{2},\frac{m}{2}+\frac{5}{2};-\pi^2\ell^2\right)\\c^{(o)}_{\ell,k}&=\sqrt{\frac{2}{U-L}}\sum_{j=1}^{\lfloor\frac{k+1}{2}\rfloor}\binom{k}{2j-1}(-1)^{k-j}\left(\frac{b}{2a}\right)^{k-2j+1}\left(\frac{\pi\ell}{U-L}\right)^{2j-1},\\c^{(e)}_{\ell,k}&=\sqrt{\frac{2}{U-L}}\sum_{j=0}^{\lfloor\frac{k}{2}\rfloor}\binom{k}{2j}(-1)^{k-j}\left(\frac{b}{2a}\right)^{k-2j}\left(\frac{\pi\ell}{U-L}\right)^{2j}.\end{aligned}$$

Here, Γ_E *is the Euler gamma function, and* ${}_1\mathrm{F}_2$ *and* ${}_1\widetilde{\mathrm{F}}_2$ *are hypergeometric and regularized hypergeometric functions, respectively.*

Proof. Recalling the definition of Ψ from Proposition 2.4.10, we compute

$$\langle\Psi_{\ell',\gamma'},\mathcal{B}\Psi_{\ell,\gamma}\rangle=\int_I\mathrm{d}x\int_{\mathbb{R}}\mathrm{d}y\;\mathfrak{m}(x)\overline{\Psi}_{\ell',\gamma'}(x,y)x^iy^j\partial_x^k\partial_y^j\Psi_{\ell,\gamma}(x,y)$$

$$= (\mathtt{i}\gamma)^l(-\mathtt{i}\partial_\gamma)^j\delta(\gamma-\gamma')\int_I \mathrm{d}x\ \mathfrak{m}(x)x^i\phi_{\ell'}(x)\partial_x^k\phi_\ell(x). \tag{2.4.32}$$

We see by direct computation that

$$\partial_x^k\phi_\ell(x;b,a) = \mathrm{e}^{-\frac{b}{2a}x}\left(c_{\ell,k}^{(o)}\cos\left(\frac{\ell\pi(x-L)}{U-L}\right) + c_{\ell,k}^{(e)}\sin\left(\frac{\ell\pi(x-L)}{U-L}\right)\right).$$

Thus, we have

$$\begin{aligned}
&\int_I \mathrm{d}x\ \mathfrak{m}(x)x^i\phi_{\ell'}(x)\partial_x^k\phi_\ell(x)\\
&= \int_I \mathrm{d}x\ x^i\sin\left(\frac{\ell'\pi(x-L)}{U-L}\right)\left(c_{\ell,k}^{(o)}\cos\left(\frac{\ell\pi(x-L)}{U-L}\right) + c_{\ell,k}^{(e)}\sin\left(\frac{\ell\pi(x-L)}{U-L}\right)\right)\\
&= \sum_{m=0}^{i}\binom{i}{m}L^{i-m}\left(\frac{U-L}{\pi}\right)^{m+1}\int_0^\pi \mathrm{d}x\ x^m\sin(\ell'x)\left(c_{\ell,k}^{(o)}\cos(\ell x) + c_{\ell,k}^{(e)}\sin(\ell x)\right)\\
&= \sum_{m=0}^{i}\binom{i}{m}L^{i-m}\left(\frac{U-L}{\pi}\right)^{m+1}\left(C_{\ell',\ell,m,k}^{(1)}\mathbb{I}_{\ell'\neq\ell} + C_{\ell,m,k}^{(2)}\delta_{\ell',\ell}\right),
\end{aligned} \tag{2.4.33}$$

where the formulas for $C_{\ell',\ell,m,k}^{(1)}$ and $C_{\ell',\ell,m,k}^{(2)}$ are given in (2.4.30) and (2.4.31), respectively. Inserting (2.4.33) into (2.4.32) yields (2.4.29). □

Remark 2.4.12. The functions ${}_1F_2$ and ${}_1\widetilde{F}_2$, which appear in the expression for $C_{\ell',\ell,i,k}$, arise from computing integrals of the form $\int_0^\pi \mathrm{d}x\, x^m\sin(\ell'x)\cos(\ell x)$ and $\int_0^\pi \mathrm{d}x\, x^m\sin(\ell'x)\sin(\ell x)$. For any $m,\ell,\ell'\in\mathbb{N}_0$, these integrals are equal to finite sums of terms containing powers of x, sines and cosines (as can be seen by integrating by parts). Thus, the functions ${}_1F_2$ and ${}_1\widetilde{F}_2$ can be evaluated with minimal computational effort.

Note from (2.3.4), (2.3.5) and (2.3.6), the operators $\mathcal{A}_{n,k}$ are finite linear combinations operators of the form (2.4.13). We see from (2.4.29) that the integrals with respect to γ_i, $i=1,2,3,\cdots,j$ in (2.4.28) collapse due to the Dirac delta functions. Since φ is independent of y, the integral with respect to γ_{j+1} also collapses. Furthermore, the iterated integrals with respect to s_i, $i=1,2,3,\cdots j$ in (2.4.18) involve only exponentials in s_i and can therefore be evaluated explicitly. Thus, (2.4.28) is an explicit sum and does not require any numerical integration.

Remark 2.4.13. The fundamental solution corresponding to the parabolic operator $(\partial_t + \mathcal{A}_{0,0} + \rho\mathcal{A}_{0,1})$ can be obtained explicitly in the European, single barrier and double barrier cases. As such, one might wonder why we expand the operator $\mathcal{A}$ in powers of $(x - \bar{x})$ and $(y - \bar{y})$ as well as in powers ρ (as opposed to expanding in powers of $(x - \bar{x})$ and $(y - \bar{y})$ only). The reason we expand in powers of ρ is that without this expansion the integrals in (2.4.3) with respect to $s_1, s_2, \ldots, s_j$ cannot be computed explicitly in the single or double-barrier cases. Thus, by expanding in ρ avoid having to evaluate multidimensional numerical integrals in time. We trade additional bookkeeping for analytic tractability.

2.5 Accuracy results

In this section, we establish the accuracy of our formal pricing approximation for European options. Before stating our accuracy result, let us introduce some additional notation. For a set $D \subset \mathbb{R}^d$, denote by $C_b^{n,1}(D)$ the class of bounded functions on D with globally Lipschitz continuous derivatives of order less than or equal to n. Let $\|f\|_{C_b^{n,1}}$ denote the sum of the L^∞-norms of the derivatives of f up to order n. We also denote by $C_b^{-1,1}(D) = L^\infty(D)$ and we set $\|\cdot\|_{C_b^{-1,1}} = \|\cdot\|_{L^\infty}$. The following theorem describes the accuracy of the Nth order approximation of the price of a European option written on an asset with local-stochastic volatility dynamics.

Theorem 2.5.1. *Let $I = \mathbb{R}$, and suppose for some non-negative integer N that $\sigma, \mu, c, g \in C_b^{N,1}(\mathbb{R}^2)$ and that there exists a positive constant M such that*

$$\frac{1}{M} \leq \|\sigma\|_{C_b^{N,1}},\ \|\mu\|_{C_b^{N,1}},\ \|c\|_{C_b^{N,1}},\ \|g\|_{C_b^{N,1}} \leq M.$$

Furthermore, assume that $\varphi \in C_b^{h-1,1}(\mathbb{R}^2)$ for some $0 \leq h \leq 2$. Then we have

$$|(u - \bar{u}_0^\rho)(t, x, y)| \leq C\,(T - t)^{\frac{h+1}{2}}, \qquad 0 \leq t < T, \qquad x \in I, y \in \mathbb{R}. \quad (2.5.1)$$

For $N \geq 1$, we have

$$|(u - \bar{u}_N^\rho)(t, x, y)| \leq C\left((T - t)^{\frac{1}{2}} + |\rho|\right) \sum_{i=0}^{N} |\rho|^i (T - t)^{\frac{N-i+h}{2}}, \quad (2.5.2)$$

where $0 \leq t < T$, $x \in I$, *and* $y \in \mathbb{R}$. *The positive constants* C *in* (2.5.1) *and* (2.5.2) *depend only on* M, N *and* $\|\varphi\|_{C_b^{h-1,1}}$.

Our strategy for proving Theorem 2.5.1 is to adapt the proof of asymptotic accuracy in [51] to our present situation. As such, many of the propositions and lemmas needed for the proof of Theorem 2.5.1 follow from analogous propositions and lemmas contained in [51].

For the remainder of this section, we let $z = (x, y)$, $\bar{z} = (\bar{x}, \bar{y})$ and $\zeta = (\xi, \eta)$ be elements of $\mathbb{R}^2$. It will also be convenient to introduce multi-index notation for the operators $\mathcal{A}$ and $\mathcal{A}_{n,k}$. We write

$$\mathcal{A} = \sum_{|\alpha| \leq 2} a_\alpha(z) \mathcal{D}_z^\alpha, \qquad \mathcal{A}_{n,k} = \sum_{\alpha \in A_k} a_{\alpha,n}(z) \mathcal{D}_z^\alpha, \qquad \mathcal{D}_z^\alpha = \partial_{z_1}^{\alpha_1} \partial_{z_2}^{\alpha_2},$$

where

$$A_0 = \{(1,0), (0,1), (2,0), (0,2)\}, \qquad A_1 = \{(1,1)\},$$

$$\alpha = (\alpha_1, \alpha_2) \in A_0 \cup A_1, \qquad |\alpha| = \alpha_1 + \alpha_2.$$

Before proving Theorem 2.5.1, we require some preliminary results. In what follows, we denote by Γ the fundamental solution corresponding to the parabolic operator $\partial_t + \mathcal{A}$.

Lemma 2.5.2. *For any* $\delta > 0$, *and* $\alpha, \beta \in \mathbb{N}_0^2$ *with* $\beta \leq N + 2$, *we have*

$$\left| (z - \zeta)^\alpha \mathcal{D}_z^\beta \Gamma(t, z; T, \zeta) \right| \leq C \cdot (T - t)^{\frac{|\alpha| - |\beta|}{2}} \widehat{\Gamma}(t, z; T, \zeta), \tag{2.5.3}$$

and

$$\left| (z - \zeta)^\alpha \mathcal{D}_\zeta^\beta \Gamma(t, z; T, \zeta) \right| \leq C \cdot (T - t)^{\frac{|\alpha| - |\beta|}{2}} \widehat{\Gamma}(t, z; T, \zeta), \tag{2.5.4}$$

for $0 \leq t < T$, *and* $z, \zeta \in I \times \mathbb{R}$, *where* $\widehat{\Gamma}$ *is the fundamental solution of the operator* $(\partial_t + (M + \delta)(\partial_{z_1}^2 + \partial_{z_2}^2))$, *and* C *is a positive constant dependent only on* M, N, δ*and* $|\beta|$.

Proof. The inequality (2.5.3) is given in [51, Lemma 6.21]. The inequality (2.5.4) can be seen by examining the Kolmogorov forward equation. □

The following fact will also be helpful. Let a and b be constants such that $a, b \geq 1/2$. Then, for $0 \leq t < T$

$$\int_t^T \mathrm{d}s\, (T-s)^a (s-t)^b = \frac{\Gamma_E(a+1)\Gamma_E(b+1)}{\Gamma_E(a+b+2)}(T-t)^{a+b+1}, \tag{2.5.5}$$

where Γ_E is the Euler gamma function.

Proposition 2.5.3. *Under the assumptions of Theorem 2.5.1, for any multi-index $\beta \in \mathbb{N}_0^2$ we have*

$$\left|\mathcal{D}_z^\beta u_{0,0}(t,z)\right| \leq C \cdot (T-t)^{\frac{\min\{h-|\beta|,0\}}{2}}, \qquad 0 \leq t < T, \qquad z, \bar{z} \in \mathbb{R}^2. \tag{2.5.6}$$

If $N \geq 1$, then for any $n, k \in \mathbb{N}$, $1 \leq n+k \leq N$, we have

$$\left|\mathcal{D}_z^\beta u_{n,k}(t,z)\right| \leq C \cdot (T-t)^{\frac{n+h-|\beta|}{2}} \left(1 + |z-\bar{z}|^n (T-t)^{-\frac{n}{2}}\right), \tag{2.5.7}$$

for $0 \leq t < T$, $z, \bar{z} \in \mathbb{R}^2$. The constants in (2.5.6) and (2.5.7) depend only on $M, N, |\beta|$ and $\|\varphi\|_{C_b^{h-1,1}}$.

Proof. The proof is analogous to the proof of [51, Lemma 6.24]. □

Proposition 2.5.4. *Define for $i \geq 0$*

$$\mathcal{A}_i^\rho := \mathcal{A}_{i,0} + \rho \mathcal{A}_{i-1,1}, \qquad \bar{\mathcal{A}}_n^\rho := \sum_{i=0}^n \mathcal{A}_i^\rho, \tag{2.5.8}$$

$$u_n^\rho := \sum_{i=0}^n \varepsilon^i \rho^{n-i} u_{i,n-i}\Big|_{\varepsilon=1}, \tag{2.5.9}$$

with the convention that $\mathcal{A}_{-1,1} = 0$. Then for $N \geq 0$, we have

$$(u - \bar{u}_N^\rho)(t,z) = \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta\, \Gamma(t,z;s,\zeta) \sum_{i=0}^N (\mathcal{A} - \bar{\mathcal{A}}_i^\rho) u_{N-i}^\rho(s,\zeta). \tag{2.5.10}$$

Proof. We will show that

$$(\partial_t + \mathcal{A})(u - \bar{u}_N^\rho) + \sum_{i=0}^N \left(\mathcal{A} - \bar{\mathcal{A}}_i^\rho\right) u_{N-i}^\rho = 0, \qquad (u - \bar{u}_N^\rho)(T, \cdot) = 0, \tag{2.5.11}$$

from which (2.5.10) follows by an application of Duhamel's principal. Note that (2.5.11) follows if we show

$$(\partial_t + \mathcal{A})\bar{u}_N^\rho = \sum_{i=0}^{N} \left(\mathcal{A} - \bar{\mathcal{A}}_i^\rho\right) u_{N-i}^\rho, \tag{2.5.12}$$

because $(\partial_t + \mathcal{A})u = 0$ and $u(T,\cdot) = \bar{u}_N^\rho(T,\cdot) = \varphi$. From equations (2.3.7), (2.3.8), (2.5.8) and (2.5.9), we deduce

$$(\partial_t + \mathcal{A}_{0,0})u_n^\rho + \sum_{i=1}^{n} \mathcal{A}_i^\rho u_{n-i}^\rho = 0. \tag{2.5.13}$$

We now proceed to show (2.5.12) by induction. When $N = 0$, since $u_0^\rho = \bar{u}_0^\rho$, we have

$$(\partial_t + \mathcal{A})\,\bar{u}_0^\rho = \left(\mathcal{A} - \bar{\mathcal{A}}_0^\rho\right) u_0^\rho.$$

Assume now that (2.5.12) holds for $N \geq 1$. Then we have by (2.5.13) that

$$\begin{aligned}
(\partial_t + \mathcal{A})\,\bar{u}_{N+1}^\rho &= (\partial_t + \mathcal{A})\,\bar{u}_N^\rho + (\partial_t + \mathcal{A})\,u_{N+1}^\rho \\
&= \sum_{i=0}^{N} \left(\mathcal{A} - \bar{\mathcal{A}}_i^\rho\right) u_{N-i} + \left(\mathcal{A} - \bar{\mathcal{A}}_0^\rho\right) u_{N+1}^\rho - \sum_{i=1}^{N+1} \mathcal{A}_i^\rho u_{N-i+1}^\rho \\
&= \sum_{i=1}^{N+1} \left(\mathcal{A} - \bar{\mathcal{A}}_{i-1}^\rho\right) u_{N-i+1}^\rho + \left(\mathcal{A} - \bar{\mathcal{A}}_0^\rho\right) u_{N+1}^\rho - \sum_{i=1}^{N+1} \mathcal{A}_i^\rho u_{N-i+1}^\rho \\
&= \sum_{i=0}^{N+1} \left(\mathcal{A} - \bar{\mathcal{A}}_i^\rho\right) u_{N-i+1}^\rho.
\end{aligned}$$

Therefore, (2.5.12) holds for all N. □

We now prove Theorem 2.5.1.

Proof of Theorem 2.5.1. From (2.5.8) and (2.5.10) we have

$$(u - \bar{u}_N^\rho)(t,z) = \sum_{k=0}^{N} \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta\, \Gamma(t,z;s,\zeta) \left(\mathcal{A} - \sum_{j=0}^{k} (\mathcal{A}_{j,0} + \rho\mathcal{A}_{j-1,1})\right) \sum_{i=0}^{N-k} \rho^i u_{N-k-i,i}(s,\zeta). \tag{2.5.14}$$

Let $T_k^{a_\alpha}(z)$ be the k-th Taylor polynomial approximation of $a_\alpha(z)$ with the convention that $T_{-1}^{a_\alpha}(z) = 0$. We rewrite (2.5.14) as

$$(u - \bar{u}_N^\rho)(t,z) = \sum_{k=0}^{N}\sum_{i=0}^{N-k} \rho^{i+1} \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \left(a_{(1,1)} - T_{k-1}^{a_{(1,1)}}\right)(\zeta)\Gamma(t,z;s,\zeta)\mathcal{D}_\zeta^{(1,1)} u_{N-k-i,i}(s,\zeta)$$

$$+ \sum_{k=0}^{N}\sum_{i=0}^{N-k}\sum_{\alpha\in A_0} \rho^{i} \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \left(a_\alpha - T_k^{a_\alpha}\right)(\zeta)\Gamma(t,z;s,\zeta)\mathcal{D}_\zeta^{\alpha} u_{N-k-i,i}(s,\zeta)$$

$$= \sum_{k=0}^{N}\sum_{i=0}^{N-k} \rho^{i+1} J_{i,k}^{(1)} + \sum_{k=0}^{N}\sum_{i=0}^{N-k} \rho^{i} \left(J_{i,k,1}^{(2)} + J_{i,k,2}^{(2)}\right), \tag{2.5.15}$$

where

$$J_{i,k}^{(1)} := \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \left(a_{(1,1)} - T_{k-1}^{a_{(1,1)}}\right)(\zeta)\Gamma(t,z;s,\zeta)\mathcal{D}_\zeta^{(1,1)} u_{N-k-i,i}(s,\zeta),$$

$$J_{i,k,1}^{(2)} := \sum_{|\alpha|\le 1} \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \left(a_\alpha - T_k^{a_\alpha}\right)(\zeta)\Gamma(t,z;s,\zeta)\mathcal{D}_\zeta^{\alpha} u_{N-k-i,i}(s,\zeta),$$

$$J_{i,k,2}^{(2)} := \sum_{\substack{|\alpha|=2\\ \alpha\neq(1,1)}} \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \left(a_\alpha - T_k^{a_\alpha}\right)(\zeta)\Gamma(t,z;s,\zeta)\mathcal{D}_\zeta^{\alpha} u_{N-k-i,i}(s,\zeta).$$

We first consider $J_{i,k}^{(1)}$. We note that $J_{i,0} = 0$ since $\mathcal{A}_{-1,1} = 0$ by convention. For $k \ge 1$, we perform integration by parts to obtain for $|\alpha_1| = |\alpha_2| = 1$,

$$J_{i,k}^{(1)} = -\int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \left[\mathcal{D}_\zeta^{\alpha_1}\left(a_{(1,1)} - T_{k-1}^{a_{(1,1)}}\right)(\zeta)\Gamma(t,z;s,\zeta)\right]\left[\mathcal{D}_\zeta^{\alpha_2} u_{N-k-i,i}(s,\zeta)\right].$$

By the product rule and (2.5.4), evaluating at $z = \bar{z}$ gives

$$|J_{i,k}^{(1)}| \le C_1 \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, |z-\zeta|^{k-1}\Gamma(t,z;s,\zeta)\left|\mathcal{D}_\zeta^{\alpha_2} u_{N-k-i,i}(s,\zeta)\right|$$

$$+ C_2 \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, |z-\zeta|^{k}\left|\mathcal{D}_z^{\alpha_1}\Gamma(t,z;s,\zeta)\right|\left|\mathcal{D}_\zeta^{\alpha_2} u_{N-k-i,i}(s,\zeta)\right|.$$

Applying (2.5.3) and (2.5.7) gives

$$|J_{i,k}^{(1)}| \le C_3 \int_t^T \mathrm{d}s \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \widehat{\Gamma}(t,z;s,\zeta)(s-t)^{\frac{k-1}{2}}(T-s)^{\frac{N+h-k-i-1}{2}}\left(1 + |z-\zeta|^{N-k-i}(T-s)^{\frac{N-k-i}{2}}\right)$$

$$\leq C_4 \int_t^T \mathrm{d}s \left[(s-t)^{\frac{k-1}{2}} (T-s)^{\frac{N+h-k-i-1}{2}} + (s-t)^{\frac{N-i-1}{2}} (T-s)^{\frac{h-1}{2}} \right] \int_{\mathbb{R}^2} \mathrm{d}\zeta \, \widehat{\Gamma}_2(t,z;s,\zeta)$$

$$\leq C_5 \, (T-t)^{\frac{N+h-i}{2}},$$

where the last line comes from (2.5.5). Similar arguments show

$$|J^{(2)}_{i,k,1}| \leq C_6 \, (T-t)^{\frac{N+h-i+2}{2}}, \qquad |J^{(2)}_{i,k,2}| \leq C_7 \, (T-t)^{\frac{N+h-i+1}{2}},$$

for $0 \leq k \leq N$. When $N = 0$, $J^{(1)}_{i,k} = 0$, so by (2.5.15) we have

$$|(u - \bar{u}^{\rho}_0)(t,z)| \leq C_8 \, (T-t)^{\frac{h+1}{2}}.$$

When $N \geq 1$, by (2.5.15) we have

$$\begin{aligned}
|(u - \bar{u}^{\rho}_N)(t,z)| &\leq C_9 \sum_{k=0}^{N} \sum_{i=0}^{N-k} \left(|\rho|^{i+1} (T-t)^{\frac{N-i+h}{2}} + |\rho|^{i} (T-t)^{\frac{N+h-i+1}{2}} \right) \\
&\leq C_{10} \left((T-t)^{\frac{1}{2}} + |\rho| \right) \sum_{k=0}^{N} \sum_{i=0}^{N-k} |\rho|^{i} (T-t)^{\frac{N-i+h}{2}} \\
&= C_{11} \left((T-t)^{\frac{1}{2}} + |\rho| \right) \sum_{k=0}^{N} (N-k+1) |\rho|^{k} (T-t)^{\frac{N-k+h}{2}} \\
&\leq C_{12} \left((T-t)^{\frac{1}{2}} + |\rho| \right) \sum_{k=0}^{N} |\rho|^{k} (T-t)^{\frac{N-k+h}{2}},
\end{aligned}$$

which proves Theorem 2.5.1. □

2.6 Numerical examples

2.6.1 Heston model

In this section, we implement our pricing approximation for an underlying $X = \log S$ that has Heston dynamics [40]. Specifically, we suppose that (X, Y) satisfies

$$\mathrm{d}X_t = -\frac{1}{2} Y_t \, \mathrm{d}t + \sqrt{Y_t} \, \mathrm{d}W_t, \qquad \mathrm{d}Y_t = \kappa(\theta - Y_t) \, \mathrm{d}t + \delta \sqrt{Y_t} \, \mathrm{d}B_t, \qquad \mathrm{d} \left\langle W, B \right\rangle_t = \rho \, \mathrm{d}t.$$

We require that Y satisfies the Feller condition i.e. $2\kappa\theta \geq \delta^2$, which implies that Y remains strictly positive. In our numerical experiments, we consider double-barrier knock-out calls and puts with the following parameters fixed

X_0	Y_0	K	T	ρ	κ	θ	δ
0.62	0.04	.62	0.083	-0.4	1.15	0.04	0.2

where e^K represents the strike and T represents the maturity date. We first consider call payoffs $\varphi(x) = (\mathrm{e}^x - \mathrm{e}^K)^+$ with the lower barrier $L = 0$ fixed and the upper barrier $U > K$ varying. We compute both our zeroth and second order price approximation $\bar{u}_0^\rho$ and $\bar{u}_2^\rho$ as well as the "exact" price u, which we obtain via Monte Carlo simulation.

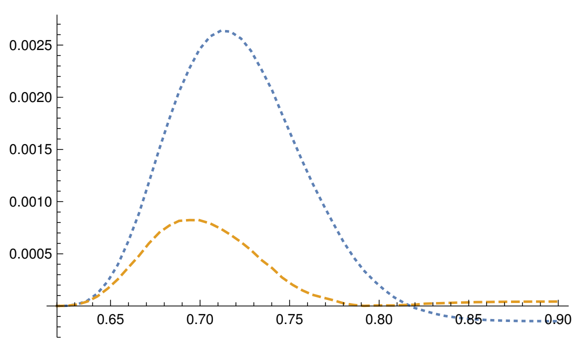


Figure 2.1: For the Heston model considered in Section 2.6.1, we plot $u - \bar{u}_0^\rho$ (blue dotted) and $u - \bar{u}_2^\rho$ (orange dotted-dashed) as a function of the upper barrier U for a call option.

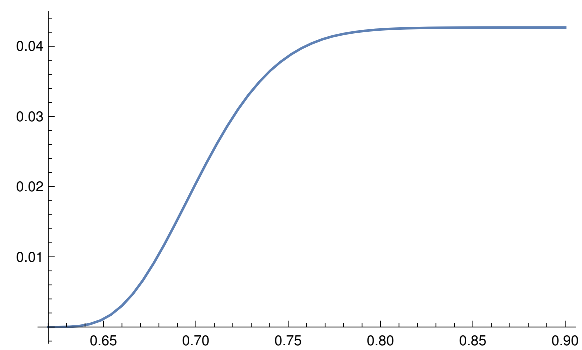


Figure 2.2: For the Heston model considered in Section 2.6.1, we plot u as a function of the upper barrier U for a call option.

In Figure 2.1, we plot the error $u - \bar{u}_0^\rho$ and $u - \bar{u}_2^\rho$ of our zeroth and second order approximations, respectively, as a function of the upper barrier U. To get a sense of the scale of the error, we also plot in Figure 2.2 the exact price u as a function of U. In Figures 2.3 and 2.4, we provide analogous plots for put payoffs $\varphi(x) = (\mathrm{e}^K - \mathrm{e}^x)^+$ with the upper barrier $U = 1$ fixed while varying the lower barrier $L < K$. We see from Figures 2.1 and

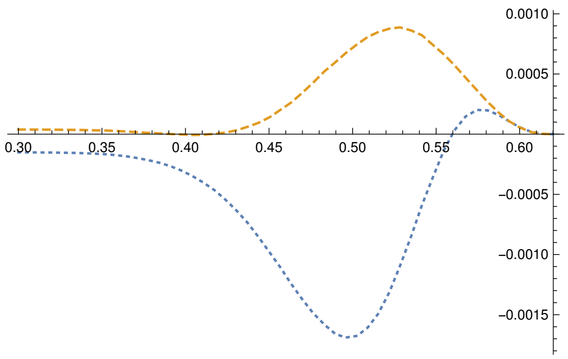


Figure 2.3: For the Heston model considered in Section 2.6.1, we plot $u - \bar{u}_0^\rho$ (blue dotted) and $u - \bar{u}_2^\rho$ (orange dotted-dashed) as a function of the lower barrier L for a put option.

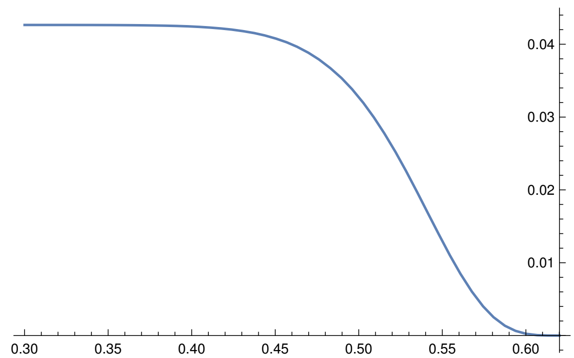


Figure 2.4: For the Heston model considered in Section 2.6.1, we plot u as a function of the lower barrier L for a put option.

2.3 that $\bar{u}_2$ provides a more accurate approximation of u than $\bar{u}_0^\rho$ for both puts and calls at nearly all levels of L and U.

Remark 2.6.1. We omit the first order approximation $\bar{u}_1^\rho$ in Figures 2.1 and 2.3 for the following reason. The difference $|\bar{u}_0^\rho - \bar{u}_1^\rho|$ is small compared to $|\bar{u}_0^\rho - \bar{u}_2^\rho|$ because when the payoff function φ depends only on x (as is the case for call and put payoffs), we have $u_{0,1} = 0$. Therefore, the first correlation correction term in our approximation appears at the second order in $u_{1,1}$. The effect of including the first correlation correction is large compared to the first correction due to y-dependence in the coefficients of $\mathcal{A}$.

2.6.2 CEV Model

In this section, we implement our pricing approximation for an underlying $S = \mathrm{e}^X$ that has Constant Elasticity of Variance (CEV) dynamics [27]. Specifically, we suppose that X satisfies

$$\mathrm{d}X_t = -\frac{1}{2}\sigma^2 \mathrm{e}^{2X_t(\gamma-1)}\,\mathrm{d}t + \sigma \mathrm{e}^{X_t(\gamma-1)}\,\mathrm{d}W_t.$$

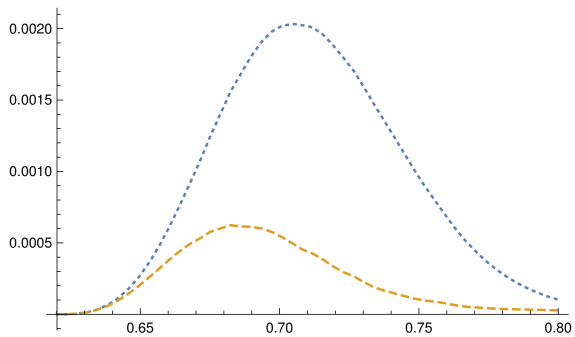


Figure 2.5: For the CEV model considered in Section 2.6.2, we plot $u - \bar{u}_0$ (blue dotted) and $u - \bar{u}_2$ (orange dashed) as a function of the upper barrier U for a call option.

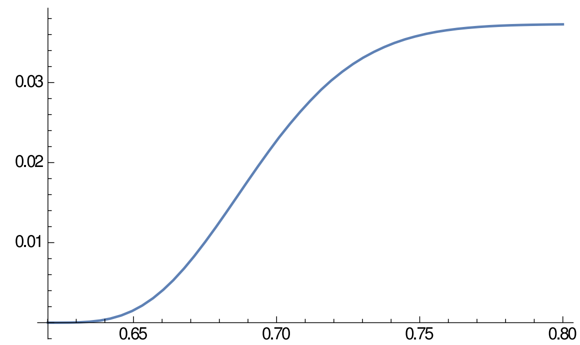


Figure 2.6: For the CEV model considered in Section 2.6.2, we plot u as a function of the upper barrier U for a call option.

where $\sigma > 0$ and $\gamma > 0$. We consider double-barrier knock-out calls and puts with the following parameters fixed

X_0	K	T	σ	γ
0.62	0.62	0.083	0.32	0.019

where e^K represents the strike and T represents the maturity date. We first consider call payoffs $\varphi(x) = (e^x - e^K)^+$ with the lower barrier $L = 0$ fixed and the upper barrier $U > K$ varying. We compute the zeroth and second order price approximation $\bar{u}_0$ and $\bar{u}_2$, respectively, as well as "exact" price u, which we obtain via Monte Carlo simulation. Note that we omit the superscript ρ from u and $\bar{u}$ as correlation plays no role in a local volatility setting. In Figure 2.5, we plot the error $u - \bar{u}_0$ and $u - \bar{u}_2$ of our zeroth and second order approximations as functions of the upper barrier U. To get a sense of the scale of the error, we also plot in Figure 2.6 the exact price u as a function of U. In Figures 2.7 and 2.8, we provide analogous plots for put payoffs $\varphi(x) = (e^K - e^x)^+$ with the upper barrier $U = 1$ fixed while varying the lower barrier $L < K$. We see from Figures 2.5 and 2.7 that in both the call

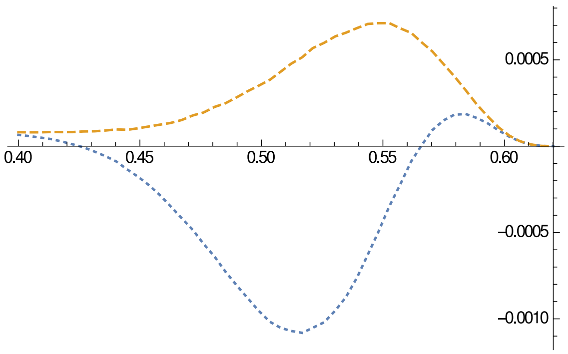


Figure 2.7: For the CEV model considered in Section 2.6.2, we plot $u - \bar{u}_0$ (blue dotted) and $u - \bar{u}_2$ (orange dashed) as a function of the lower barrier L for a put option.

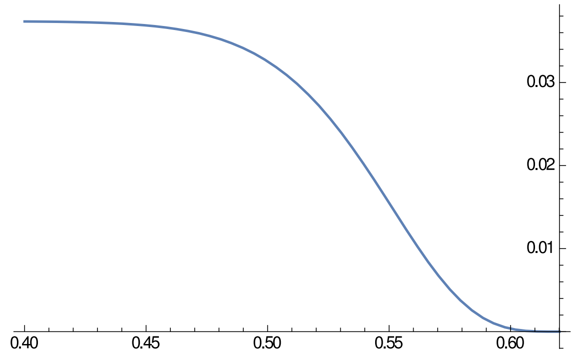


Figure 2.8: For the CEV model considered in Section 2.6.2, we plot u as a function of the lower barrier L for a put option.

and put cases the second order approximation outperforms the zeroth order approximation.

2.7 Conclusion

In this chapter, we have presented a formal pricing approximation for European and barrier-style claims in a local-stochastic volatility setting. We have provided rigorous accuracy results for European-style claims. And we have provided several numerical examples illustrating the accuracy and versatility of our approximation for barrier-style claims.

We set up a rather general pricing framework where the coefficients of the SDEs modeling asset prices were specified only up to a class of function, and we developed formulas for the approximate solution to a PDE in terms of the coefficients. If we now prescribe specific dynamics for the price of the underlying, our formulas immediately yield a closed form approximation to the price of a European or barrier-style claim. This expansion approach can be applied to obtain approximation formulas for other PDE problems. In the next chapter, we apply the Taylor expansion method to the dynamic programming equation corresponding to an optimal control problem to obtain approximately optimal controls in terms of the SDE coefficients of the controlled process.

Chapter 3

OPTIMAL LIQUIDATION UNDER STOCHASTIC PRICE IMPACT

3.0.1 Introduction

When institutional traders execute large market orders, they are faced with transactional frictions. Direct frictions, such as exchange and brokerage fees, are known in advance and can be incorporated into a trading strategy. Traders also incur indirect costs, and such costs are, in general, unknown in advance and may be difficult to quantify even after the trading is complete. The opportunity cost that arises from waiting to execute trades and the price impact that results from trading are both examples of indirect costs. Price impact typically affects traders adversely. Selling an asset puts downward pressure on the price thereby lowering revenues while purchasing an asset pushes its price upward, resulting in higher costs. We focus on price impact costs in this chapter. Specifically, we examine how a trader should optimally liquidate a large position in a market in which price impact is stochastic.

The optimal liquidation problem under price impact has been studied extensively in the literature. [13] use a linear price impact model and solve a discrete optimal control problem to minimize expected trading costs. [3, 4, 42] also use a linear price impact model but consider the variance in trading costs. [5] employs nonlinear impact functions, and the continuous-time limit of the aforementioned models are discussed in [3, 4] in more detail. [2] considers optimal liquidation in a market with stochastic liquidity and stochastic volatility. [55] include price impact by modeling the limit order book directly (see also the published version, [56]). The authors of [1] extend the work of [55] to allow for general limit order book shapes. [23] use a continuous time linear impact model and incorporate stochastic order flow.

For an overview of continuous-time price impact models, see [24] and the references therein.

In this chapter, we assume a continuous-time price impact model where the price impact parameters are stochastic. Specifically, the temporary and permanent price impact parameters are modeled as scalar Markov diffusions. We allow the temporary and permanent price impact parameters to be correlated, as empirical evidence suggests they are (see [23]). In this setting, we define a trader's value function and formulate the associated Hamilton-Jacobi-Bellman (HJB) PDE (see Appendix A for a brief review of the HJB PDE and dynamic programming in continuous-time). We find an approximate solution of the HJB equation by applying coefficient expansion techniques that were first developed for one-dimensional, linear parabolic PDEs in [59] and extended to d-dimensional PDEs in [51], to the ultra-parabolic case [58], and to nonlinear problems in [53, 50] and [36]. This, in turn, yields approximations to the associated optimal trading strategy. The resulting optimal strategy approximations are explicit and do not require numerical integration.

The zeroth order approximation to the optimal strategy can be interpreted as an Almgren-Chriss strategy for which the price impact parameters are recalibrated in continuous time. Successive terms in higher-order approximations can therefore be viewed as corrections to the strategy of Almgren and Chriss. Higher-order strategy approximations are influenced by the local behavior of the coefficients of the SDEs modeling the price impact parameter's dynamics, allowing traders to take advantage of periods of relatively high or low price impact.

The rest of the chapter is organized as follows. In Section 3.1, we state our modeling assumptions, define the trader's value function, and provide the associated HJB equation. In Section 3.2, we develop an asymptotic expansion for solutions of the HJB equation. This expansion leads to a sequence of PDEs, which we solve recursively in Section 3.2.2. The solution of these PDEs allows us to construct approximations to the optimal liquidation strategy. We discuss limiting cases of the optimal strategy approximations in Section 3.2.3. In Section 3.3, we demonstrate the effectiveness of the approximate optimal strategies by performing a Monte Carlo study. Some concluding remarks are offered in Section 3.4.

3.1 Market model and trader's value function

To begin, we fix a trading horizon $T > 0$ and filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$. We suppose an institutional trader holds $Q_0 > 0$ shares of a stock S that they wish to liquidate. The trader does not post limit orders but trades exclusively via market orders. They must choose the speed at which they send market orders with the aim of liquidating all Q_0 shares by the end of a trading horizon T. We assume that the trader trades in continuous time, and we denote by $\nu = (\nu_t)_{0\leq t\leq T}$ the rate at which the trader sends market orders (i.e., the liquidation speed). The inventory $Q^\nu = (Q_t^\nu)_{0\leq t\leq T}$ depends on the trading strategy ν and is given by

$$\mathrm{d}Q_t^\nu = -\nu_t \mathrm{d}t.$$

A positive trading rate $\nu_t > 0$ at a time t corresponds to selling shares of S, and a negative rate $\nu_t < 0$ corresponds to buying shares of S. Although we shall restrict ourselves to the liquidation problem in this chapter, we mention that the set up for the acquisition problem is similar. The trader wishes to choose ν such that they minimize the indirect costs they incur as a result of trading. We incorporate price impact in the model by explicitly including temporary and permanent price impact parameters.

3.1.1 Permanent price impact

We assume that when the trader sends market orders there is a permanent impact on the midprice of the stock. Sell orders put downward pressure on the midprice of the stock, and, conversely, buy orders put upward pressure on the midprice. For example, suppose that a trader submits a large sell order for S, and suppose further that other traders on the market have similar signals and also post market orders to sell. Liquidity providers fill the gap in the book by posting limit orders to sell at lower prices, thus moving the midprice down.

We model the midprice of the stock as a stochastic process $S^\nu = (S_t^\nu)_{0\leq t\leq T}$ with the dynamics

$$\mathrm{d}S_t^\nu = -g(b_t)\nu_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \tag{3.1.1}$$

where the constant $\sigma > 0$ is positive, the function g is continuous and real-valued, the Markov diffusion $b = (b_t)_{0\le t\le T}$ has the dynamics

$$\mathrm{d}b_t = \eta(b_t)\,\mathrm{d}t + \psi(b_t)\,\mathrm{d}B_t^{(1)}, \tag{3.1.2}$$

and the standard Brownian motions $W = (W_t)_{0\le t\le T}$ and $B^{(1)} = (B_t^{(1)})_{0\le t\le T}$ are uncorrelated.

We think of the Brownian motion W as market noise due to the reshuffling of limit orders. Permanent price impact is modeled by the process $g(b) = (g(b_t))_{0\le t\le T}$, the magnitude of which corresponds the severity of the impact. We require that $g(b_t) \ge 0$ for all $0 \le t \le T$, because a negative permanent price impact would imply that selling shares of an asset would push the midprice upwards, which is unrealistic. We perform asymptotic expansions in Section 3.2 for general g, but we give the linear model $g(z) = z$ special attention in Section 3.3.

3.1.2 Temporary impact

In addition to permanent price impact, we assume that the trader also faces a temporary price impact. Temporary price impact is the cost directly associated with each trade, and, unlike the permanent impact, temporary impact does not carry over into subsequent trades. Temporary impact can be understood as follows: the number of shares available at the best bid is limited, and if the trader's market order is large enough then the trader walks the book (i.e. they deplete the outstanding limit orders nearest the midprice). We include temporary price impact in the model by defining the execution price $\widehat{S}^\nu = (\widehat{S}_t^\nu)_{0\le t\le T}$ of the asset to be

$$\widehat{S}_t^\nu = S_t^\nu - f(a_t)\nu_t, \tag{3.1.3}$$

where the function f is continuous and real-valued and the Markov diffusion $a = (a_t)_{0\le t\le T}$ has the dynamics

$$\mathrm{d}a_t = \mu(a_t)\,\mathrm{d}t + \omega(a_t)\,\mathrm{d}B_t^{(2)}. \tag{3.1.4}$$

Here, the standard Brownian motion $B^{(2)} = (B_t^{(2)})_{0\le t\le T}$ is uncorrelated with the Brownian motion W that drives the midprice $\mathrm{d}\left\langle W, B^{(2)}\right\rangle = 0$ but is correlated with the Brownian motion $B^{(1)}$ that drives the permanent price impact $\mathrm{d}\left\langle B^{(1)}, B^{(2)}\right\rangle = \rho\,\mathrm{d}t$ where $\rho \in (-1, 1)$. Taking the temporary price impact to be a stochastic process allows us to incorporate stochastic liquidity into our model.

We require that $f(a_t) > 0$ for all $0 \le t \le T$ to reflect the fact that market orders are not frictionless. Assuming the trader wishes to minimize the cost associated with temporary price impact, if the temporary price impact ever reached zero, the trader would liquidate their entire inventory immediately resulting in blowup in the optimal strategy. The magnitude of the process $f(a_t)$ describes the severity of the temporary price impact. It is suggested in [23] that temporary and permanent price impact are correlated, so we allow temporary and permanent price impact to be correlated with parameter ρ. Generally, temporary and permanent price impact are positively correlated, although our model does not require that.

Remark 3.1.1. As the coefficients of the SDEs governing (a, b) are left in a general form, it may appear that the functions f and g serve merely as changes of variables and that both f and g can be taken to be the identity function without a loss of generality. However, there are a few reasons to include f and g. First, a practitioner may have in mind a particular model for permanent and temporary price impact in the form of (3.1.1), (3.1.2), (3.1.3), (3.1.4). Thus, leaving the model as specified above facilitates practical implementation. Second, while it is only the dynamics of $f(a) = (f(a_t))_{0\le t\le T}$ and $g(b) = (g(b_t))_{0\le t\le T}$ that affect the dynamics of the midprice S^ν and the execution price $\widehat{S}^\nu$, not the individual specification of the processes (a, b) and the functions (f, g), the approximations we develop in subsequent sections for the optimal liquidation strategies *do* depend on the particular choice of (a, b) and (f, g). Among equivalent model specifications, there may be a particular choice of (a, b) and (f, g) whose approximate optimal liquidation strategy is closest to exact optimal liquidation strategy.

The above framework stipulates that the temporary impact is only felt by the trader who initiates the market order. Furthermore, our model assumes that the limit order book rebalances infinitely fast to the state before the arrival of the market order. This assumption is known as order book resilience. See [1], [5], [35], [44], and [64] for further study and relaxations of the resilience assumption. Asymptotic expansions are performed in Section 3.2 for general f, but we give special attention to the linear case $f(z) = z$ in Section 3.3.

In the framework described above, one easily derives that the trader's cash position $X^\nu = (X_t^\nu)_{0 \le t \le T}$ is given by

$$\mathrm{d}X_t^\nu = \nu_t \widehat{S}_t^\nu \mathrm{d}t = \nu_t \left(S_t^\nu - f(a_t)\nu_t \right) \mathrm{d}t.$$

3.1.3 Trader's value function

We consider a trader who wishes to liquidate Q_0 shares of S^ν under the model described in Section 3.1. We assume that the trader wishes to maximize their expected cash at the terminal time T subject to penalties for holding inventory. For a given trading strategy ν, we define the trader's performance criteria H^ν to be

$$H^\nu(t,x,s,q,a,b) := \mathbb{E}_{t,x,s,q,a,b} \left[X_T^\nu + Q_T^\nu \left(S_T^\nu - \kappa Q_T^\nu \right) - \varphi \int_t^T \left(Q_s^\nu \right)^2 \mathrm{d}s \right], \qquad (3.1.5)$$

where the constants $\kappa > 0$ and $\varphi > 0$ are positive and $\mathbb{E}_{t,x,s,q,a,b}$ is shorthand for expectation conditioned on $(X_t^\nu, S_t^\nu, Q_t^\nu, a_t, b_t) = (x, s, q, a, b)$. From left to right, the following three terms are present in the trader's performance criteria (3.1.5): terminal cash, the proceeds from liquidating the remaining shares at the terminal time T, and an integral term penalizing the holding of inventory. The proceeds from liquidation at time T are subject to temporary price impact, which is incorporated through the parameter κ. The third term $\varphi \int_t^T \mathrm{d}s (Q_s^\nu)^2$ imposes a running penalty for holding inventory. When φ is large, optimal strategies will trade quickly at the beginning of the trading horizon rather than face holding large inventories. In [21], the authors show that including the inventory penalty term is equivalent to the trader considering alternate models with stochastic drifts but penalizing models that are far from the reference model in the sense of relative entropy. In that context, larger values of φ correspond to a trader who is less confident about the drift of the S^ν. The authors of [22] introduce the inventory penalty term heuristically and justify it by showing that it is proportional to the variance of the book value of the inventory over the trading horizon.

The trader's value function is given by

$$H(t,x,s,q,a,b) = \sup_{\nu \in \mathcal{A}_t} H^\nu(t,x,s,q,a,b),$$

where $\mathcal{A}_t = \{\nu \mid \nu \text{ is } \mathbb{F}\text{-adapted and } \int_t^T |\nu_s| \, \mathrm{d}s < \infty, \mathbb{P} - \text{a.s.}\}$ is the set of admissible strategies.

3.1.4 The Hamilton-Jacobi-Bellman equation

In this section, we give the HJB equation associated with the value function H. Let $\mathcal{H}^\nu$ denote the infinitesimal generator for the process $(X^\nu, S^\nu, Q^\nu, a, b)$ with ν fixed. Explicitly, $\mathcal{H}^\nu$ is given by

$$\begin{aligned} \mathcal{H}^\nu = \; & \frac{1}{2}\sigma^2 \partial_s^2 - g(b)\nu\partial_s - \nu\partial_q + \nu\left(s - f(a)\nu\right)\partial_x \\ & + \frac{1}{2}\omega^2(a)\partial_a^2 + \rho\,\omega(a)\psi(b)\partial_a\partial_b + \frac{1}{2}\psi^2(b)\partial_b^2 + \mu(a)\partial_a + \eta(b)\partial_b. \end{aligned}$$

As we shall see in later in this section, it is convenient to write $\mathcal{H}^\nu$ as the sum of two operators

$$\mathcal{H}^\nu = \mathcal{A}^\nu + \mathcal{L},$$

$$\mathcal{A}^\nu := \frac{1}{2}\sigma^2 \partial_s^2 - g(b)\nu\partial_s - \nu\partial_q + \nu\left(s - f(a)\nu\right)\partial_x, \tag{3.1.6}$$

$$\mathcal{L} := \frac{1}{2}\omega^2(a)\partial_a^2 + \rho\,\omega(a)\psi(b)\partial_a\partial_b + \frac{1}{2}\psi^2(b)\partial_b^2 + \mu(a)\partial_a + \eta(b)\partial_b. \tag{3.1.7}$$

The operator $\mathcal{L}$ is the infinitesimal generator of the process (a,b) and the operator $\mathcal{A}^\nu$ is the infinitesimal generator of (X^ν, S^ν, Q^ν) with (ν, a, b) fixed. When the process (a,b) is constant (i.e., $(\mu, \eta, \omega, \psi) = (0,0,0,0)$), we have $\mathcal{H}^\nu = \mathcal{A}^\nu$, and the model reduces to the continuous-time Almgren-Chriss model (see [4]).

The HJB equation associated with the trader's value function H is

$$\left(\partial_t + \mathcal{L}\right) H + \sup_\nu \left(\mathcal{A}^\nu H - \varphi q^2\right) = 0, \qquad H(T,x,s,q,\cdot,\cdot) = x + q\left(s - \kappa q\right), \tag{3.1.8}$$

where $\mathcal{A}^\nu$ and $\mathcal{L}$ are given by (3.1.6) and (3.1.7), respectively. We assume that (3.1.8) admits a unique classical solution which coincides with the trader's value function (see [60]).

Following [23], we make the following ansatz

$$H(t,x,s,q,a,b) = x + qs + q^2 h(t,a,b), \tag{3.1.9}$$

for some function h to be determined. Henceforth, we refer to h as the *transformed value function*. Inserting (3.1.9) into (3.1.8) yields the following PDE problem for h:

$$q^2(\partial_t + \mathcal{L})h + \sup_\nu \left(-q(2\nu h + \nu g + q\varphi) - \nu^2 f\right) = 0, \qquad h(T,\cdot,\cdot) = -\kappa. \tag{3.1.10}$$

The optimal strategy ν^*, obtained by maximizing the supremum in (3.1.10), is given in feedback form as

$$\nu^*(t,q,a,b) = -\left(\frac{g(b) + 2h(t,a,b)}{2f(a)}\right) q. \tag{3.1.11}$$

Inserting (3.1.11) into (3.1.10) we obtain

$$0 = (\partial_t + \mathcal{L})h + \mathcal{N}(h) - \varphi, \qquad h(T,\cdot,\cdot,) = -\kappa, \tag{3.1.12}$$

$$\mathcal{N}(h) = \frac{1}{f}h^2 + \frac{g}{f}h + \frac{g^2}{4f}. \tag{3.1.13}$$

Note that we have reduced the HJB equation (3.1.8) to a PDE that involves only three variables: (t,a,b).

3.2 Asymptotics

For general $(f,g,\omega,\mu,\psi,\eta,\rho)$, there is no closed-form solution to (3.1.12). In this section, we develop a formal asymptotic expansion for the transformed value function h and the corresponding optimal execution strategy ν^* by performing polynomial expansions on the coefficients of (3.1.12). As mentioned in Chapter 2, the authors of [51] use this approach for the European option pricing problem in a general local-stochastic volatility setting. This is the approach we take in the previous chapter. One key difference here is that, unlike classical option pricing PDEs, which are linear, the PDE (3.1.12) is nonlinear. Our approach is similar to that of [53] and [50], who apply the polynomial coefficient expansion method to the Merton problem and indifference pricing problem, both of which are nonlinear.

3.2.1 Coefficient Taylor series expansions

This section proceeds similarly to Section 2.3. In Section 2.3, we performed a double expansion by expanding the solution of a parabolic PDE in the correlation parameter ρ and a polynomial expansion on the PDE coefficients.

In this section, we perform a polynomial expansion of the PDE (3.1.10) in a single parameter. However, we perform a polynomial expansion of the coefficients of a nonlinear PDE, which requires a different treatment. Furthermore, in Section 2.3 we were interested in obtaining exact formulas for the approximate value of the solution u of the PDE (2.2.2). In this setting, we are not only interested in approximately evaluating h, but in obtaining formulas for approximations to the optimal trading strategy ν^* in (3.1.11). The setting here differs enough from that of Section 2.3 that some concepts bear repeating.

As in Section 2.3, we assume for the sake of simplicity in the following formal computations that the coefficients of (3.1.12) are analytic on $\mathbb{R}^2$. We shall see later that the Nth-order approximations we obtain for h and ν^* require only that the coefficients of (3.1.12) belong to $C^N(D)$ where D is some open set in $\mathbb{R}^2$ and $C^N(D)$ is the set of functions $f : D \to \mathbb{R}$ whose derivatives up to order N exist and are continuous.

Let χ be a placeholder for any of the coefficients appearing in PDE (3.1.12)

$$\chi \in \{\tfrac{1}{2}\omega^2, \rho\omega\psi, \tfrac{1}{2}\psi^2, \mu, \eta, f^{-1}, f^{-1}g, 4^{-1}f^{-1}g^2\}, \tag{3.2.1}$$

and fix a point $(\overline{a}, \overline{b}) \in \mathbb{R}^2$. For any $\varepsilon \in [0, 1]$, we define

$$\chi^\varepsilon(a, b) := \chi(\overline{a} + \varepsilon(a - \overline{a}), \overline{b} + \varepsilon(b - \overline{b})). \tag{3.2.2}$$

Formally, Taylor expanding χ^ε in ε about the point $\varepsilon = 0$ gives

$$\chi^\varepsilon(a, b) := \sum_{n=0}^{\infty} \varepsilon^n \chi_n(a, b), \qquad \varepsilon \in [0, 1], \tag{3.2.3}$$

$$\chi_n(a, b) := \sum_{k=0}^{n} \chi_{n-k,k} \cdot (a - \overline{a})^{n-k}(b - \overline{b})^k, \qquad \chi_{n-k,k} := \frac{1}{(n-k)!k!}\partial_a^{n-k}\partial_b^k \chi(\overline{a}, \overline{b}). \tag{3.2.4}$$

In particular, evaluating (3.2.3) at $\varepsilon = 1$ yields the Taylor series expansion of χ about the point $(\overline{a}, \overline{b})$. Consider now the family of PDEs indexed by ε:

$$(\partial_t + \mathcal{L}^\varepsilon)h^\varepsilon + \mathcal{N}^\varepsilon(h^\varepsilon) - \varphi = 0, \qquad h^\varepsilon(T, \cdot, \cdot, \cdot) = -\kappa, \qquad \varepsilon \in [0, 1], \quad (3.2.5)$$

where $\mathcal{L}^\varepsilon$ and $\mathcal{N}^\varepsilon(\cdot)$ are the operators obtained by replacing the coefficients of $\mathcal{L}$ and $\mathcal{N}(\cdot)$ in (3.1.7) and (3.1.13), respectively, with their ε-counterparts. Explicitly, we make the replacements

$$\begin{aligned} &\{\tfrac{1}{2}\omega^2, \rho\omega\psi, \tfrac{1}{2}\psi^2, \mu, \eta, f^{-1}, f^{-1}g, 4^{-1}f^{-1}g^2\} \\ &\qquad \mapsto \{(\tfrac{1}{2}\omega^2)^\varepsilon, (\rho\omega\psi)^\varepsilon, (\tfrac{1}{2}\psi^2)^\varepsilon, \mu^\varepsilon, \eta^\varepsilon, (f^{-1})^\varepsilon, (f^{-1}g)^\varepsilon, (4^{-1}f^{-1}g^2)^\varepsilon\} \end{aligned}$$

in (3.1.12) to obtain (3.2.5). Using (3.2.3), the linear operator $\mathcal{L}^\varepsilon$ in the PDE (3.2.5) can be written as

$$\mathcal{L}^\varepsilon = \sum_{n=0}^{\infty} \varepsilon^n \mathcal{L}_n, \qquad (3.2.6)$$

where we have defined

$$\mathcal{L}_n := (\tfrac{1}{2}\omega^2)_n \partial_a^2 + (\rho\omega\psi)_n \partial_{ab}^2 + (\tfrac{1}{2}\psi^2)_n \partial_b^2 + \mu_n \partial_a + \eta_n \partial_b,$$

and the subscript notation χ_n is as described in (3.2.4). The expansion of the nonlinear operator $\mathcal{N}^\varepsilon$ is more involved, and we handle it below.

We construct an expansion for the function h^ε, the solution to the PDE (3.2.5), as the power series in ε

$$h^\varepsilon(t, a, b) = \sum_{n=0}^{\infty} \varepsilon^n h_n(t, a, b), \qquad \varepsilon \in [0, 1]. \qquad (3.2.7)$$

Here, the sequence of functions $(h_n)_{n=0}^\infty$ are not polynomials in (a, b) but rather functions to be determined which, in particular, are independent of ε. We shall eventually construct the asymptotic approximation to the transformed value function h by truncating (3.2.7) for some $n = N$ and setting $\varepsilon = 1$.

We insert (3.2.6) and (3.2.7) into (3.2.5), expand the terms in $\mathcal{N}^\varepsilon(h^\varepsilon)$ in powers of ε, and collect terms of like order in ε. As the equality in (3.2.5) holds for all for all $\varepsilon \in [0,1]$, we obtain the following sequence of PDEs:

$$O(\varepsilon^0): \qquad 0 = (\partial_t + \mathcal{L}_0)\, h_0 + (f^{-1})_0 h_0^2 + (f^{-1}g)_0 h_0 + (4^{-1} f^{-1} g^2)_0 - \varphi, \qquad h_0(T,\cdot,\cdot) = -\kappa, \tag{3.2.8}$$

$$O(\varepsilon^n): \qquad 0 = \left(\partial_t + \widehat{\mathcal{L}}_0\right) h_n + F_n, \qquad h_n(T,\cdot,\cdot) = 0, \tag{3.2.9}$$

for $n \geq 1$, where we have defined the differential operator

$$\widehat{\mathcal{L}}_0 := \mathcal{L}_0 + 2(f^{-1})_0 h_0 + (f^{-1}g)_0, \tag{3.2.10}$$

and the functions

$$F_n := \sum_{i=0}^{n-1} \mathcal{L}_{n-i} h_i + \sum_{i=0}^{n-1} \sum_{j=0}^{i} (f^{-1})_{n-i} h_{i-j} h_j + f_0 \sum_{i=1}^{n-1} h_{n-i} h_i + \sum_{i=0}^{n-1} (f^{-1}g)_{n-i} h_i + (4^{-1} f^{-1} g^2)_n. \tag{3.2.11}$$

We are now in position to define the Nth order approximation for h.

Definition 3.2.1. Let N be a non-negative integer, and assume that the coefficients of $\mathcal{L}$ and $\mathcal{N}$ are $C^N(D)$ where D is an open set in $\mathbb{R}^2$. For any $(a,b) \in D$, we define the Nth-order approximation of the transformed value function h by

$$\overline{h}_N(t,a,b) := \sum_{n=0}^{N} \varepsilon^n h_n(t,a,b)\Big|_{\left(\varepsilon,\overline{a},\overline{b}\right)=(1,a,b)},$$

where h_0 is the solution to (3.2.8) and h_n, for $n \geq 1$, is the solution to (3.2.9).

We now focus on developing an Nth order approximation for the optimal liquidation strategy ν^*. To this end, recalling the expression (3.1.11) for the optimal execution strategy ν^*, we define

$$(\nu^*)^\varepsilon(\cdot, q, \cdot, \cdot) := -\left(\frac{g^\varepsilon + 2h^\varepsilon}{2 f^\varepsilon}\right) q, \qquad \varepsilon \in [0,1],$$

where f^ε and g^ε are given by (3.2.2) and h^ε is the solution to (3.2.5).

Definition 3.2.2. Let N be a non-negative integer, and assume that the coefficients of (3.1.11) are $C^N(D)$ where D is an open set in $\mathbb{R}^2$. For any $(a,b) \in D$, we define the define the Nth-order approximation of the optimal control ν^* as

$$\overline{\nu_N^*}(\cdot,\cdot,a,b) := \sum_{n=0}^{N} \varepsilon^n \nu_n^*(\cdot,\cdot,a,b)\Big|_{(\varepsilon,\overline{a},\overline{b})=(1,a,b)}, \tag{3.2.12}$$

where, for every n, the function ν_n^* is the nth-order coefficient in the Taylor series expansion of $(\nu^*)^\varepsilon$ about $\varepsilon = 0$.

Remark 3.2.3. As we noted at the beginning of Section 3.2.1, for a given N, the analyticity of the coefficients of $\mathcal{L}$ and $\mathcal{N}$ is not required to construct the approximation $\overline{h}_N$. Indeed, to construct the N-th order approximations $\overline{h}_N$ and $\overline{\nu_N^*}$ one only needs that the coefficients are $C^N(D)$ for some $D \subseteq \mathbb{R}^2$.

Remark 3.2.4. The construct the approximations $\overline{h}_N$ and $\overline{\nu_N^*}$ by summing up the solutions to the $O(\varepsilon^i)$ problems for $i = 0, 1, ..., N$ and evaluation the sum at $(\varepsilon, \overline{a}, \overline{b}) = (1, a, b)$. This may be a point of confusion. However, we performed the analogous evaluation of the asymptotic approximations of Chapter 2 and discussed this evaluation in Remarks 2.3.2 and 2.3.3.

We now give a representation of the approximate strategy $\overline{\nu_N^*}$ (3.2.12) in terms of the functions (h_n).

Proposition 3.2.5. *Fix $N \geq 0$, and suppose the coefficients functions appearing in (3.2.1) belong to $C^N(D)$ for some open set $D \subset \mathbb{R}^2$. Then for any $(a,b) \in D$, the approximate strategy $\overline{\nu_N^*}$ in (3.2.12) is given by*

$$\overline{\nu_N^*}(t,q,a,b) = -\frac{1}{f(a)}\left(\frac{1}{2}g(b) + \sum_{n=0}^{N} h_n(t,a,b)\Big|_{(\overline{a},\overline{b})=(a,b)}\right) q, \tag{3.2.13}$$

where the h_0 is the solution to the PDE (3.2.8), and h_n is the solution to the PDE (3.2.9) for $n \geq 1$.

Proof. Suppose that $1 \leq k \leq N$. By (3.2.4), we have that

$$\chi_k\Big|_{(\overline{a},\overline{b})=(a,b)} = 0.$$

Thus, for $0 \leq n \leq N$,

$$\begin{aligned} \nu_n^*(t,q,a,b)\Big|_{(\overline{a},\overline{b})=(a,b)} &= -\left(\frac{1}{2}(f^{-1}g)_n + \sum_{i=0}^{n}(f^{-1})_{n-i}h_i(t,a,b)\right)q\Bigg|_{(\overline{a},\overline{b})=(a,b)} \\ &= -\frac{1}{f(a)}\left(\frac{1}{2}g(b)\mathbb{1}_{\{n=0\}} + h_n(t,a,b)\Big|_{(\overline{a},\overline{b})=(a,b)}\right)q, \end{aligned} \tag{3.2.14}$$

where $\mathbb{1}$ is the indicator function. By inserting (3.2.14) into (3.2.12) and evaluating at $\varepsilon = 1$, we arrive at (3.2.13). □

3.2.2 Expressions for h_n

We begin this section by solving (3.2.8) explicitly for h_0, which yields an explicit representation of the operator $\widehat{\mathcal{L}}_0$. We then give a recursive, integral expression for h_n and evaluate the integral explicitly for h_1. We use the expressions for h_0 and h_1 to construct $\overline{\nu_0^*}$ and $\overline{\nu_1^*}$. For readability, we opt not to give h_2 or higher order approximations to the transformed value function. However, while tedious to obtain, their explicit computation is straightforward.

Proposition 3.2.6. *The classical solution h_0 to (3.2.8) is*

$$h_0(t) = -\frac{1}{2}g_0 + \sqrt{\varphi f_0}\,\theta_0(t), \qquad \theta_0(t) = \frac{1+\zeta e^{2\gamma(T-t)}}{1-\zeta e^{2\gamma(T-t)}}, \tag{3.2.15}$$

where we have defined the constants

$$\gamma := \sqrt{\frac{\varphi}{f_0}}, \qquad \zeta := \frac{\kappa - \frac{1}{2}g_0 + \sqrt{\varphi f_0}}{\kappa - \frac{1}{2}g_0 - \sqrt{\varphi f_0}}. \tag{3.2.16}$$

Proof. As both the forcing term and the terminal condition in (3.2.8) are independent of (a,b), we conclude that h_0 is a function of t only. Therefore, $\mathcal{L}_0 h_0 = 0$. The PDE (3.2.8) thus reduces to the constant coefficient ODE

$$h_0' + \frac{1}{f_0}h_0^2 + \frac{g_0}{f_0}h_0 + \frac{g_0^2}{4f_0} - \varphi = 0, \qquad h_0(T) = -\kappa, \tag{3.2.17}$$

which the reader will recognize (3.2.17) as a Ricatti equation. We check by direct substitution that (3.2.15) satisfies (3.2.17). □

Corollary 3.2.7. *The operator $\widehat{\mathcal{L}}_0$, defined in (3.2.10), is an elliptic operator and has the explicit representation*

$$\widehat{\mathcal{L}}_0 := \mathcal{L}_0 + 2\gamma\theta_0,$$

where $\mathcal{L}_0$ is defined in (3.1.7), θ_0 is defined in (3.2.15), and γ is given in (3.2.16).

With an explicit expression for h_0 in hand, we are able to write the zeroth order approximation $\overline{\nu_0^*}$ to the optimal liquidation strategy ν^*. By (3.2.12) and (3.2.15), we have

$$\overline{\nu_0^*}(t,q,a,b) = -\gamma \frac{1+\zeta \mathrm{e}^{2\gamma(T-t)}}{1-\zeta \mathrm{e}^{2\gamma(T-t)}} q \Bigg|_{(\overline{a},\overline{b})=(a,b)}, \tag{3.2.18}$$

where γ and ζ are given in (3.2.16). We note that $\overline{\nu_0^*}(t,q,\overline{a},\overline{b})$ is the continuous-time Almgren-Chriss strategy, and we define

$$\nu_{AC}(t,q) := -\gamma \frac{1+\zeta \mathrm{e}^{2\gamma(T-t)}}{1-\zeta \mathrm{e}^{2\gamma(T-t)}} q. \tag{3.2.19}$$

Thus, the strategy $\overline{\nu_0^*}$ can be viewed as an implementation of ν_{AC} in which the price impact parameters are recalibrated in continuous time.

We now give a recursive expression for h_n. The reader may with to refer to the beginning of Section 2.4 for a review of Duhamel's principle.

Proposition 3.2.8. *The classical solution h_n to (3.2.9), which is unique within the class of non-rapidly increasing functions, is*

$$\begin{aligned} h_n(t,a,b) &= \int_t^T \mathcal{P}_0(t,s) F_n(s,a,b)\,\mathrm{d}s \\ &= \int_t^T \int_{\mathbb{R}^2} \widehat{\Gamma}_0(t,a,b;s,\alpha,\beta) F_n(s,\alpha,\beta)\,\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}s. \end{aligned} \tag{3.2.20}$$

Here, $\mathcal{P}_0$ is the semigroup operator generated by $\widehat{\mathcal{L}}_0$, F_n is given in (3.2.11), and $\widehat{\Gamma}_0$ is the fundamental solution associated with the operator $\partial_t + \widehat{\mathcal{L}}_0$. The function $\widehat{\Gamma}_0$ is given explicitly *by*

$$\widehat{\Gamma}_0(t, \mathbf{a}; s, \mathbf{y}) = \frac{\Psi_0(t,s)}{2\pi\sqrt{\det \mathbf{C}(t,s)}} \times \exp\left(-\frac{1}{2}\left\langle \mathbf{C}^{-1}(t,s)(\mathbf{y}-\mathbf{a}-\mathbf{m}(t,s)), \mathbf{y}-\mathbf{a}-\mathbf{m}(t,s)\right\rangle\right), \tag{3.2.21}$$

where

$$C := \begin{pmatrix} \omega_0^2 & \rho(\omega\psi)_0 \\ \rho(\omega\psi)_0 & \psi_0^2 \end{pmatrix}, \qquad \mathbf{C}(t,s) := \int_t^s C \, \mathrm{d}s = C(s-t), \tag{3.2.22a}$$

$$\mathbf{a} := \begin{pmatrix} a & b \end{pmatrix}^\top, \qquad \mathbf{y} := \begin{pmatrix} \alpha & \beta \end{pmatrix}^\top, \tag{3.2.22b}$$

$$m := \begin{pmatrix} \mu_0 & \eta_0 \end{pmatrix}^\top, \qquad \mathbf{m}(t,s) := \int_t^s m \, \mathrm{d}s = (s-t)m, \tag{3.2.22c}$$

and

$$\Psi_0(t,s) := e^{-2\gamma(s-t)} \left(\frac{\zeta e^{2\gamma T} - e^{2\gamma s}}{\zeta e^{2\gamma T} - e^{2\gamma t}}\right)^2. \tag{3.2.23}$$

Proof. Let $\widehat{\Gamma}_0$ be the fundamental solution associated with $\partial_t + \widehat{\mathcal{L}}_0$. As the coefficients of the operator $\widehat{\mathcal{L}}_0$ are constant in (a,b), classical results give

$$\widehat{\Gamma}_0(t, \mathbf{a}; s, \mathbf{y}) = \frac{1}{2\pi\sqrt{\det \mathbf{C}(t,s)}} \times \exp\left(2\gamma \int_t^s \theta_0(r)\, \mathrm{d}r - \frac{1}{2}\left\langle \mathbf{C}^{-1}(t,s)(\mathbf{y}-\mathbf{a}-\mathbf{m}(t,s)), \mathbf{y}-\mathbf{a}-\mathbf{m}(t,s)\right\rangle\right),$$

where $\mathbf{C}, \mathbf{a}, \mathbf{y}$ and $\mathbf{m}$ are given in (3.2.22). Using (3.2.15), we compute explicitly

$$\exp\left(2\gamma \int_t^s \theta_0(r)\, \mathrm{d}r\right) = \Psi_0(t,s),$$

which yields the expression (3.2.21). Applying Duhamel's principle to (3.2.9) gives (3.2.20).

□

Corollary 3.2.9. *Define*

$$c^{(1)}(t,a,b) := -\gamma^2 f'(\overline{a})\mu_0, \qquad c^{(2)}(t,a,b) := -\gamma^2 f'(\overline{a})(a-\overline{a}-t\mu_0),$$
$$c^{(3)}(t,a,b) := \gamma g'(\overline{b})\eta_0, \qquad c^{(4)}(t,a,b) := \gamma g'(\overline{b})(b-\overline{b}-t\eta_0),$$

and

$$\begin{aligned}
I^{(1)}(t) &:= \frac{\mathrm{e}^{2\gamma(t+T)}\left(4\gamma^2\zeta\left(T^2-t^2\right)-\zeta^2(2\gamma T+1)+2\gamma T-1\right)}{4\gamma^2\left(\mathrm{e}^{2\gamma t}-\zeta\mathrm{e}^{2\gamma T}\right)^2} \\
&\qquad + \frac{\mathrm{e}^{4\gamma t}(1-2\gamma t)+\zeta^2(2\gamma t+1)\mathrm{e}^{4\gamma T}}{4\gamma^2\left(\mathrm{e}^{2\gamma t}-\zeta\mathrm{e}^{2\gamma T}\right)^2} \\
I^{(2)}(t) &:= -\frac{\mathrm{e}^{4\gamma t}+\mathrm{e}^{2\gamma(t+T)}\left(\zeta^2+4\gamma\zeta(t-T)-1\right)+\zeta^2\left(-\mathrm{e}^{4\gamma T}\right)}{2\gamma\left(\mathrm{e}^{2\gamma t}-\zeta\mathrm{e}^{2\gamma T}\right)^2}, \\
I^{(3)}(t) &:= \frac{\mathrm{e}^{4\gamma t}(1-2\gamma t)+\zeta^2(2\gamma t+1)\left(-\mathrm{e}^{4\gamma T}\right)}{4\gamma^2\left(\mathrm{e}^{2\gamma t}-\zeta\mathrm{e}^{2\gamma T}\right)^2} \\
&\qquad + \frac{\mathrm{e}^{2\gamma(t+T)}\left(\zeta^2+2\gamma\left(\zeta^2+1\right)T-1\right)}{4\gamma^2\left(\mathrm{e}^{2\gamma t}-\zeta\mathrm{e}^{2\gamma T}\right)^2}, \\
I^{(4)}(t) &:= -\frac{\left(\mathrm{e}^{2\gamma t}-\mathrm{e}^{2\gamma T}\right)\left(\mathrm{e}^{2\gamma t}-\zeta^2\mathrm{e}^{2\gamma T}\right)}{2\gamma\left(\mathrm{e}^{2\gamma t}-\zeta\mathrm{e}^{2\gamma T}\right)^2},
\end{aligned}$$

where θ_0 and Ψ_0 are given in (3.2.15) and (3.2.23), respectively. Then the solution h_1 to (3.2.9) is given by

$$h_1(t,a,b) = \sum_{i=1}^{4} c^{(i)}(t,a,b)\, I^{(i)}(t). \tag{3.2.26}$$

Proof. By (3.2.20), we have

$$\begin{aligned}
h_1(t,a,b) &= \int_t^T \mathcal{P}_0(t,s)F_1(s,a,b)\,\mathrm{d}s \\
&= \int_t^T \int_{\mathbb{R}^2} \widehat{\Gamma}_0(t,a,b;s,\alpha,\beta)\left((f^{-1})_1(\alpha)h_0^2(s)\right. \\
&\qquad\qquad \left.+(f^{-1}g)_1(\alpha,\beta)h_0(s)+(4^{-1}f^{-1}g^2)_1(\alpha,\beta)\right)\,\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}s.
\end{aligned} \tag{3.2.27}$$

Using (3.2.4), we see that

$$\left(f^{-1}\right)_1(\alpha) = -\frac{f'(\overline{a})}{f^2(\overline{a})}(\alpha-\overline{a}), \tag{3.2.28a}$$

$$(f^{-1}g)_1(\alpha,\beta) = \frac{f'(\overline{a})g(\overline{b})}{f^2(\overline{a})}(\alpha-\overline{a}) + \frac{g'(\overline{b})}{f(\overline{a})}(\beta-\overline{b}), \tag{3.2.28b}$$

$$\left(4^{-1}f^{-1}g^2\right)_1(\alpha,\beta) = -\frac{f'(\overline{a})g^2(\overline{b})}{4f^2(\overline{a})}(\alpha-\overline{a}) + \frac{g(\overline{b})g'(\overline{b})}{2f(\overline{a})}(\beta-\overline{b}). \tag{3.2.28c}$$

Additionally, by (3.2.21),

$$\int_{\mathbb{R}^2} \alpha\widehat{\Gamma}_0(t,a,b;s,\alpha,\beta)\,\mathrm{d}\alpha\mathrm{d}\beta = \Psi_0(t,s)\left(a+\mu_0(s-t)\right), \tag{3.2.29}$$

$$\int_{\mathbb{R}^2} \beta\widehat{\Gamma}_0(t,a,b;s,\alpha,\beta)\,\mathrm{d}\alpha\mathrm{d}\beta = \Psi_0(t,s)\left(b+\eta_0(s-t)\right), \tag{3.2.30}$$

where Ψ_0 is given in (3.2.23). Recalling the expression (3.2.15) for h_0 and applying (3.2.28), (3.2.29) and (3.2.30) to (3.2.27), we evaluate the integrals with respect α and β present in (3.2.27) to obtain

$$\begin{aligned} h_1(t,a,b) = \int_t^T \Psi_0(t,s)\big(&-\gamma^2 f'(\overline{a})\left(a-\overline{a}+\mu_0(s-t)\right)\theta_0^2(s) \\ &+\gamma g'(\overline{b})\left(b-\overline{b}+\eta_0(s-t)\right)\theta_0(s)\big)\,\mathrm{d}s. \end{aligned} \tag{3.2.31}$$

We can rewrite the integrand of (3.2.31) as

$$\begin{aligned} &c^{(1)}(t,a,b)\,s\theta_0^2(s)\Psi_0(t,s) + c^{(2)}(t,a,b)\,\theta_0^2(s)\Psi_0(t,s) \\ &\quad + c^{(3)}(t,a,b)\,s\theta_0(s)\Psi_0(t,s) + c^{(4)}(t,a,b)\,\theta_0(s)\Psi_0(t,s) \end{aligned} \tag{3.2.32}$$

where the $c^{(i)}$ are defined in (3.2.24), and θ_0 and Ψ_0 are defined in (3.2.15) and (3.2.23), respectively. From (3.2.31) and (3.2.32), we see–with the arguments of the functions $c^{(i)}$ suppressed for notational convenience–that

$$\begin{aligned} h_1(t,a,b) &= \int_t^T \big(c^{(1)}s\theta_0^2(s)\Psi_0(t,s) + c^{(2)}\theta_0^2(s)\Psi_0(t,s) \\ &\qquad +c^{(3)}s\theta_0(s)\Psi_0(t,s) + c^{(4)}\theta_0(s)\Psi_0(t,s)\big)\,\mathrm{d}s \\ &= c^{(1)}\int_t^T s\theta_0^2(s)\Psi_0(t,s)\,\mathrm{d}s + c^{(2)}\int_t^T \theta_0^2(s)\Psi_0(t,s)\,\mathrm{d}s \\ &\quad + c^{(3)}\int_t^T s\theta_0(s)\Psi_0(t,s)\,\mathrm{d}s + c^{(4)}\int_t^T \theta_0(s)\Psi_0(t,s)\,\mathrm{d}s. \end{aligned} \tag{3.2.33}$$

Evaluating the integrals in (3.2.33), we see that

$$I^{(1)}(t) = \int_t^T s\theta_0^2(s)\Psi_0(t,s)\,\mathrm{d}s, \qquad I^{(2)}(t) = \int_t^T \theta_0^2(s)\Psi_0(t,s)\,\mathrm{d}s,$$
$$I^{(3)}(t) = \int_t^T s\theta_0(s)\Psi_0(t,s)\,\mathrm{d}s, \qquad I^{(4)}(t) = \int_t^T \theta_0(s)\Psi_0(t,s)\,\mathrm{d}s,$$

where the functions $I^{(i)}$ are defined in equation (3.2.25), which establishes (3.2.26). □

With an explicit expression for h_1 in hand, we are able to construct the first order approximation $\overline{\nu_1^*}$ to the optimal liquidation strategy ν^*. By (3.2.13) and (3.2.18), we have

$$\overline{\nu_1^*}(t,q,a,b) = -\left(\gamma\frac{1+\zeta\mathrm{e}^{2\gamma(T-t)}}{1-\zeta\mathrm{e}^{2\gamma(T-t)}} + \frac{1}{f(a)}\sum_{i=1}^{4} c^{(i)}\left(t,a,b\right) I^{(i)}\left(t\right)\right) q\Bigg|_{(\overline{a},\overline{b})=(a,b)}. \quad (3.2.34)$$

Higher order approximations to the transformed value function h can be computed explicitly. For the sake of readability, we do not carry out these calculations and instead focus on zero and first order approximations, which are sufficient to capture the lowest order effects of stochastic price impact.

3.2.3 Analysis of some limiting cases

The asymptotic approximations of the transformed value function h (and hence the approximations to ν^*) developed in Section 3.2 depend on the parameters κ, which controls the penalty for liquidation occurring at the trading horizon T, and φ, which controls the penalty for holding shares of S^ν throughout trading. In this section, we develop strategies that are independent of one or both of these parameters by taking the limit of the optimal strategy approximations $\overline{\nu_N^*}$ as the parameters κ and φ tend to ∞ and 0, respectively. One motivation for considering such strategies is financial. Taking the limit of the optimal strategy ν^* as $\kappa \to \infty$ corresponds to a setting in which the trader is adamant about liquidating their entire inventory before $t = T$. Subsequently taking $\varphi \to 0$ corresponds to a setting in which the trader both demands complete liquidation by the trading horizon and is indifferent about holding inventory.

Another motivation for developing limiting strategies is analytic tractability. One attractive feature of asymptotic expansions is that they often yield correction terms that have physical (in our case financial) interpretations, leading to qualitative insights about the behavior of solutions to the unperturbed problem. Ideally, the financial interpretation of the terms appearing in the Nth order correction would be given without considering limiting cases. However, due to the nonlinear nature of the optimal liquidation problem, the number and complexity of terms in the approximate liquidation strategy grows rapidly with the approximation order N. As a result, even the first order approximate strategy in the general case (3.2.34) has sufficiently many terms that giving an easy financial interpretation is prohibitive. However, we shall see later in this section that the approximate liquidation strategies in limiting cases $\kappa \to \infty$ and $(\kappa, \varphi) \to (\infty, 0)$ have far fewer terms than the corresponding nonlimiting case, thereby facilitating a financial interpretation of the resulting expressions.

For the remainder of this section, we make the dependence on parameters κ and φ of the PDE solutions h_n and strategies ν explicit with the superscipt notation

$$h_n \equiv h_n^{(\kappa,\varphi)}, \qquad\qquad \nu \equiv \nu^{(\kappa,\varphi)}.$$

Definition 3.2.10. Fix $N \geq 0$. We define the limiting strategies as

$$\begin{aligned}
\nu_{AC}^{(\infty,\varphi)} &:= \lim_{\kappa\to\infty} \nu_{AC}^{(\kappa,\varphi)}, & \nu_{AC}^{(\infty,0)} &:= \lim_{\varphi\to 0}\lim_{\kappa\to\infty} \nu_{AC}^{(\kappa,\varphi)}, \\
(\nu^*)^{(\infty,\varphi)} &:= \lim_{\kappa\to\infty} (\nu^*)^{(\kappa,\varphi)}, & (\nu^*)^{(\infty,0)} &:= \lim_{\varphi\to 0}\lim_{\kappa\to\infty} (\nu^*)^{(\kappa,\varphi)}, \\
\overline{\nu_N^*}^{(\infty,\varphi)} &:= \lim_{\kappa\to\infty} \overline{\nu_N^*}^{(\kappa,\varphi)}, & \overline{\nu_N^*}^{(\infty,0)} &:= \lim_{\varphi\to 0}\lim_{\kappa\to\infty} \overline{\nu_N^*}^{(\kappa,\varphi)}.
\end{aligned}$$

For n, $0 \leq n \leq N$, we also define the limiting PDE solutions

$$h_n^{(\infty,\varphi)}(t) := \lim_{\kappa\to\infty} h_n^{(\kappa,\varphi)}(t), \qquad\qquad h_n^{(\infty,0)}(t) := \lim_{\varphi\to 0}\lim_{\kappa\to\infty} h_n^{(\kappa,\varphi)}(t). \tag{3.2.35}$$

We refer to the PDE solutions $h_n^{(\kappa,\varphi)}$ and strategies $\nu^{(\kappa,\varphi)}$ as nonlimiting PDE solutions and nonlimiting strategies, respectively.

Below, we provide explicit expressions for $\overline{\nu_N^*}^{(\infty,\varphi)}$ and $\overline{\nu_N^*}^{(\infty,0)}$ for $N \in \{0,1\}$. As the strategies $\overline{\nu_N^*}^{(\kappa,\varphi)}$ depend on (κ,φ) only through the functions $h_n^{(\kappa,\varphi)}$, and $\overline{\nu_N^*}^{(\kappa,\varphi)}$ depend on the functions $h_n^{(\kappa,\varphi)}$ continuously, we can obtain the strategies $\overline{\nu_N^*}^{(\infty,\varphi)}$ and $\overline{\nu_N^*}^{(\infty,0)}$ by replacing $h_n^{(\kappa,\varphi)}$ in (3.2.13) with $h_n^{(\infty,\varphi)}$ and $h_n^{(\infty,0)}$, respectively. This leads us to make the following proposition.

Proposition 3.2.11. *Define*

$$I^{(1)}_{(\infty,\varphi)}(t) := \frac{e^{2\gamma(t+T)}\left(4\gamma^2\left(T^2-t^2\right)-2\right)}{4\gamma^2\left(e^{2\gamma t}-e^{2\gamma T}\right)^2} + \frac{e^{4\gamma t}(1-2\gamma t)+(2\gamma t+1)e^{4\gamma T}}{4\gamma^2\left(e^{2\gamma t}-e^{2\gamma T}\right)^2}, \tag{3.2.36a}$$

$$I^{(2)}_{(\infty,\varphi)}(t) := \frac{-e^{4\gamma t}-4\gamma(t-T)e^{2\gamma(t+T)}+e^{4\gamma T}}{2\gamma\left(e^{2\gamma t}-e^{2\gamma T}\right)^2}, \tag{3.2.36b}$$

$$I^{(3)}_{(\infty,\varphi)}(t) := \frac{e^{4\gamma t}(1-2\gamma t)+4\gamma T e^{2\gamma(t+T)}-(2\gamma t+1)e^{4\gamma T}}{4\gamma^2\left(e^{2\gamma t}-e^{2\gamma T}\right)^2}, \tag{3.2.36c}$$

$$I^{(4)}_{(\infty,\varphi)}(t) := -\frac{1}{2\gamma}. \tag{3.2.36d}$$

We have

$$h_0^{(\infty,\varphi)}(t) = -\frac{g_0}{2} - f_0\gamma\coth\left(\gamma(T-t)\right), \qquad h_0^{(\infty,0)}(t) = -\frac{g_0}{2} - \frac{f_0}{T-t}, \tag{3.2.37}$$

and

$$h_1^{(\infty,\varphi)}(t,a,b) = \sum_{i=1}^{4} c^{(i)}(t,a,b) I^{(i)}_{(\infty,\varphi)}(t), \tag{3.2.38}$$

$$h_1^{(\infty,0)}(t,a,b) = -\frac{f'(\overline{a})}{2(T-t)}\left(2(a-\overline{a})+\mu_0(T-t)\right) - \frac{g'(\overline{b})}{6}\left(3(b-\overline{b})+\eta_0(T-t)\right), \tag{3.2.39}$$

where $c^{(i)}$ and $I^{(i)}_{(\infty,\varphi)}$ are given in (3.2.24) and (3.2.36), respectively.

Proof. Equations (3.2.37) follow by direct computation of the limits in (3.2.35). Let us now compute $h_1^{(\infty,\varphi)}$. We see that the $c^{(i)}$ in (3.2.24) do not depend on the parameter κ. Therefore, to compute $h_1^{(\infty,\varphi)}$ it suffices to compute $\lim_{\kappa\to\infty} I^{(i)}$, for $i = 1,2,3,4$. By direct computation, we see that

$$I^{(i)}_{(\infty,\varphi)}(t) = \lim_{\kappa\to\infty} I^{(i)}(t), \qquad \forall i \in \{1,2,3,4\},$$

which gives us equation (3.2.38).

We now compute $h_1^{(\infty,0)}$ i.e. $\lim_{\varphi\to 0} h_1^{(\infty,\varphi)}$. By direct computation, we obtain

$$\begin{aligned}
\lim_{\varphi\to 0} c^{(1)}(t,a,b) I^{(1)}_{(\infty,\varphi)}(t) &= -\frac{f'(\overline{a})\mu_0(T+t)}{2(T-t)}, \\
\lim_{\varphi\to 0} c^{(2)}(t,a,b) I^{(2)}_{(\infty,\varphi)}(t) &= -\frac{f'(\overline{a})(a-\overline{a}-t\mu_0)}{T-t}, \\
\lim_{\varphi\to 0} c^{(3)}(t,a,b) I^{(3)}_{(\infty,\varphi)}(t) &= -\frac{1}{6} g'(\overline{b})\eta_0(T+2t), \\
\lim_{\varphi\to 0} c^{(4)}(t,a,b) I^{(4)}_{(\infty,\varphi)}(t) &= -\frac{1}{2} g'(\overline{b})(b-\overline{b}-t\eta_0),
\end{aligned}$$

where the $c^{(i)}$ are given in (3.2.24). Summing the terms in (3.2.40) yields the expression for $h_1^{(\infty,0)}$ given in (3.2.39), which concludes the proof. □

For $n \in \{0,1\}$, replacing $h_n^{(\kappa,\varphi)}$ with $h_n^{(\infty,\varphi)}$ in (3.2.13) yields the strategies

$$\begin{aligned}
\overline{\nu_0^*}^{(\infty,\varphi)}(t,q,a,b) &= \gamma \coth\left(\gamma(T-t)\right) q \Big|_{(\overline{a},\overline{b})=(a,b)}, \\
\overline{\nu_1^*}^{(\infty,\varphi)}(t,q,a,b) &= \left(\gamma \coth\left(\gamma(T-t)\right) - \frac{1}{f(a)} \sum_{i=1}^{4} c^{(i)}(t,a,b) I^{(i)}_{(\infty,\varphi)}(t)\right) q \Bigg|_{(\overline{a},\overline{b})=(a,b)},
\end{aligned}$$

and replacing $h_n^{(\kappa,\varphi)}$ in (3.2.13) with $h_n^{(\infty,0)}$ in (3.2.13) yields the strategies

$$\overline{\nu_0^*}^{(\infty,0)}(t,q,a,b) = \frac{1}{T-t} q, \tag{3.2.41}$$

$$\overline{\nu_1^*}^{(\infty,0)}(t,q,a,b) = \left(\frac{1}{T-t} + \frac{1}{2}\frac{\mu(a) f'(a)}{f(a)} + \frac{1}{6}(T-t)\frac{\eta(b) g'(b)}{f(a)}\right) q. \tag{3.2.42}$$

The reader will recognize (3.2.41) as the *time-weighted average* strategy. We can thus view equation (3.2.42) as a first order correction to the time-weighted average strategy. The second term in (3.2.42) instructs the trader to adjust their trading speed in proportion to the product of the slope f' of the temporary impact function and the drift μ of the process a. For instance, suppose that at time t, $\mu(a_t) < 0$ and $f'(a_t) > 0$. In that case, the temporary price impact process a is drifting downwards, and the price impact $f(a)$ will decrease with a. The second term in (3.2.42) instructs the trader to slow down liquidation because they expect a lower temporary impact in the near future. If $\mu(a_t) > 0$ and $f'(a_t) > 0$, then a trader following $\overline{\nu_1^*}^{(\infty,0)}$ will speed up trading as they expect higher price impact soon. Of course, if $\mu(a_t)f'(a_t)$ is small relative to $f(a)$, then the contributions from this term are small. The third term instructs the trader to adjust their trading speed proportional to the product $\eta(b)g'(b)$. Like the second term, the adjustments of the third term are weighted relative to the temporary price impact. But unlike the second term, the third term's influence on the trading speed diminishes as time approaches the trading horizon. This is intuitive as permanent impact matters not to a trader who is soon to exit the market. We note that the third term in strategy (3.2.42) causes more dramatic deviations from the time-weighted strategy when the permanent price impact is large relative to the temporary price impact early in the trading period.

3.3 Numerical examples

In this section, we provide examples of simulated trading using the strategies we developed in Sections 3.2.2 and 3.2.3. In Example 3.3.2, we simulate a single trading day to illustrate the effects of the correction terms present in the first order limiting strategy $\overline{\nu_1^*}^{(\infty,0)}$. In Example 3.3.3, we simulate a large number of trading days to demonstrate the improvement over the Almgren-Chriss strategy (3.2.19) our approximations give in the nonlimiting and limiting cases. Furthermore, we demonstrate an improvement of the first order strategies over the zeroth order strategies in the nonlimiting and limiting cases.

Throughout this section, we assume the price impact processes a and b are Cox-Ingersoll-Ross (see [28]) processes with the dynamics

$$\mu(z) = \lambda_a\left(\theta_a - z\right), \qquad \omega(z) = \sigma_a\sqrt{z},$$
$$\eta(z) = \lambda_b\left(\theta_b - z\right), \qquad \psi(z) = \sigma_b\sqrt{z}.$$

where the constants $\lambda_a, \theta_b, \theta_a, \theta_b, \sigma_a$ and σ_b are all positive, and the Brownian motions $B_t^{(1)}$ and $B_t^{(2)}$ are correlated with parameter ρ. Furthermore, we require that the coefficients

$\lambda_a, \theta_b, \theta_a, \theta_b, \sigma_a$ and σ_b satisfy the Feller condition

$$2\lambda_a\theta_a > \sigma_a^2, \qquad\qquad 2\lambda_b\theta_b > \sigma_b^2, \tag{3.3.1}$$

so that a and b are strictly positive processes. Explicitly, we have

$$\mathrm{d}a_t = \lambda_a\left(\theta_a - a_t\right)\mathrm{d}t + \sigma_a\sqrt{a_t}\mathrm{d}B_t^{(2)}, \qquad \mathrm{d}b_t = \lambda_b\left(\theta_b - b_t\right)\mathrm{d}t + \sigma_b\sqrt{b_t}\mathrm{d}B_t^{(1)}. \tag{3.3.2}$$

We also take

$$f(a) = a, \qquad\qquad g(b) = b.$$

Note that because we require the Feller condition (3.3.1) to be satisfied we have $f(a_t) > 0$ and $g(b_t) > 0$ for all $0 \leq t \leq T$. Thus, both temporary and permanent price impact processes $f(a)$ and $g(b)$ remain strictly positive.

Remark 3.3.1. In the following examples, we select numerical values for the parameters in (3.3.2) and perform simulated trading. We do not discuss how price impact is measured from market data as that is discussed elsewhere in the literature. For example, [23] and [24] discuss how to approximate price impact parameters from data, and the latter reports temporary and permanent price impact parameters for a selection of Nasdaq stocks. For the dynamics specified in (3.3.2), the processes a and b are mean reverting, and the long-run means θ_a and θ_b are similar to the values given in [23]. The remaining parameters in (3.3.2) are chosen so that (a, b) hover near their long-run means (θ_a, θ_b), but stray enough to demonstrate the effects (a, b) have on our approximations.

Example 3.3.2. Let us suppose the trader demands complete liquidation by T and is indifferent to holding inventory. In this case, the zeroth order strategy $\overline{\nu_0^*}^{(\infty,0)}$ is equal to time-weighted average strategy (3.2.41) and thus does not depend on (a, b). Under the dynamics (3.3.2), the first order strategy $\overline{\nu_1^*}^{(\infty,0)}$ is given by

$$\overline{\nu_1^*}^{(\infty,0)}(t, q, a, b) = \left(\frac{1}{T-t} + \frac{\lambda_a\left(\theta_a - a\right)}{2a} + (T-t)\frac{\lambda_b\left(\theta_b - b\right)}{6a}\right)q. \tag{3.3.3}$$

The second and third terms in (3.3.3) are a time-dependent linear combination of the distance of the price impact processes from their respective long-run means relative to a. The mean reversion parameters λ_a and λ_b control the aggressiveness of the adjustment. When the mean reversion parameters are large, we expect that deviations of a and b from their respective means θ_a and θ_b to be short lived. In this case, the strategy $\overline{\nu_1^*}^{(\infty,0)}$ adjusts quickly to take advantage of these deviations.

When a trader is following the time-weighted average strategy, at each instant they sell a fraction of their inventory that is inversely proportional to the remaining time $T-t$. As such $\overline{\nu_0^*}^{(\infty,0)} > 0$ for all t in the trading period. In some instances, however, the strategy $\overline{\nu_1^*}^{(\infty,0)}$ instructs the trader to purchase shares of S^ν. This occurs in times of relatively large price impact. Although it may seem counter-intuitive for a trader who wishes to liquidate a position to buy shares, this strategy can increase the objective function H^ν if, for example, the trader buys shares of S^ν during a period of relatively high price impact, putting upward pressure on the midprice S^ν, then subsequently sells shares rapidly during a period of low price impact.

In Figure 3.1, we provide a simulated path of (a, b) and the paths (X^ν, S^ν, Q^ν) that result from following strategies $\nu = \overline{\nu_0^*}^{(\infty,0)}$ and $\nu = \overline{\nu_1^*}^{(\infty,0)}$ using parameters

$$\left.\begin{aligned} &\lambda_a = 10, \quad \theta_a = 2\times 10^{-6}, \quad \sigma_a = 1.5\times 10^{-3},\\ &\lambda_b = 10, \quad \theta_b = 5\times 10^{-5}, \quad \sigma_b = 3\times 10^{-3},\\ &T = 1, \qquad \rho = 0.7, \qquad\quad \sigma = 0.01. \end{aligned}\right\} \tag{3.3.4}$$

This simulation demonstrates how the first order strategy $\overline{\nu_1^*}^{(\infty,0)}$ responds to the high and low price impact values encountered early in the trading period. The strategy $\overline{\nu_1^*}^{(\infty,0)}$ instructs the trader to purchase shares of S^ν when price impact is relatively high early in the trading period. The price impact processes subsequently decrease below their long run means, and the trader following $\overline{\nu_1^*}^{(\infty,0)}$ liquidates shares at rate in excess of the trading speed dictated by $\overline{\nu_0^*}^{(\infty,0)}$. In this example, $\overline{\nu_1^*}^{(\infty,0)}$ is more profitable than $\overline{\nu_0^*}^{(\infty,0)}$.

While the optimal strategy approximations we have developed in this chapter suggest that, under certain market conditions, the trader should buy shares of the stock when the price impact is high, the authors of [12] note that in practice, if a trader wants to sell a block of securities then it is usually antithetical to their stance as a seller to purchase shares of the security during the trading period. In some cases, it is a violation of a manager's fiduciary responsibility to their client and is hence illegal. As such, a liquidation strategy ν that can sometimes instruct a trader to buy could be modified to be $\max(0, \nu)$. A truncated strategy may not be optimal with respect to the objective functional H^ν that we have defined, but in our numerical simulations the periods in which our approximate strategies instruct the trader to buy are short-lived.

Example 3.3.3. In this example, we carry out a number of Monte Carlo simulations to evaluate the performance of the liquidation strategies $\overline{\nu_N^*}$ as well as the limiting strategies $\overline{\nu_N^*}^{(\infty,\varphi)}$ and $\overline{\nu_N^*}^{(\infty,0)}$ for $N \in \{0, 1\}$. We demonstrate the relative improvement a trader gains by following $\overline{\nu_0^*}$ over ν_{AC} and the relative improvement a trader gains by following $\overline{\nu_1^*}$ over $\overline{\nu_0^*}$. We repeat this experiment in the limiting case $\kappa \to \infty$. We see from (3.2.41) that in the limiting case $(\kappa, \varphi) \to (\infty, 0)$, both $\nu_{AC}^{(\infty,0)}$ and $\overline{\nu_0^*}^{(\infty,0)}$ are equal to the time-weighted average strategy (i.e. $\nu_{AC}^{(\infty,0)} = \overline{\nu_0^*}^{(\infty,0)} = q(T-t)^{-1}$). So, we demonstrate the relative performance increase a trader gains by following $\overline{\nu_1^*}^{(\infty,0)}$ over $\overline{\nu_0^*}^{(\infty,0)}$.

To this end, let us define the random variables

$$\begin{aligned} \Phi(\nu) &:= X_T^\nu + Q_T^\nu(S_T^\nu - \kappa Q_T^\nu) - \varphi \int_0^T (Q_s^\nu)^2 \, \mathrm{d}s, \\ \Phi^{(\infty,\varphi)}(\nu) &:= X_T^\nu - \varphi \int_0^T (Q_s^\nu)^2 \, \mathrm{d}s, \\ \Phi^{(\infty,0)}(\nu) &:= X_T^\nu. \end{aligned}$$

For a fixed strategy ν, the random variables $\Phi(\nu)$, $\Phi^{(\infty,\varphi)}(\nu)$ and $\Phi^{(\infty,0)}(\nu)$ give the value a trader following ν achieves on a single path of $(X_T^\nu, S_T^\nu, Q_T^\nu, a, b)$. In both of the limiting cases $\kappa \to \infty$ and $(\kappa, \varphi) \to (\infty, 0)$, the optimal strategies ensure liquidation by the terminal time T. Therefore, the term $Q_T^\nu(S_T^\nu - \kappa Q_T^\nu)$ that accounts for liquidation of the remaining shares at time T does not appear in either $\Phi^{(\infty,\varphi)}(\nu)$ or $\Phi^{(\infty,0)}(\nu)$. Let us also define the sample mean of our Monte carlo simulations as follows

$$\widehat{\Phi}(\nu) := \frac{1}{M}\sum_{i=1}^{M} \Phi_i(\nu),$$

where $\Phi_i(\nu)$ is the value of $\Phi(\nu)$ obtained by the i-th independent path of $(X_T^\nu, S_T^\nu, Q_T^\nu, a, b)$. Observe that $\widehat{\Phi}(\nu)$ is a statistical estimate of the performance criteria H^ν. The definitions for $\widehat{\Phi}^{(\infty,\varphi)}(\nu)$ and $\widehat{\Phi}^{(\infty,0)}(\nu)$, our statistical estimators for H^ν in the limiting cases $\kappa \to \infty$ and $(\kappa, \varphi) \to (\infty, 0)$, are analogous.

In this example, we take the following parameters

$$\left.\begin{array}{lll} \lambda_a = 1, & \theta_a = 1 \times 10^{-4}, & \sigma_a = 8 \times 10^{-3}, \\ \lambda_b = 1, & \theta_b = 5 \times 10^{-4}, & \sigma_b = 8 \times 10^{-3}, \\ T = 1, & \rho = 0.7, & \varphi = 0.01, \\ \kappa = 10, & \sigma = 0.2. & \end{array}\right\} \tag{3.3.5}$$

We note that the processes a and b under the parameter choice (3.3.5) both satisfy the Feller condition (3.3.1). Furthermore, we choose the initial conditions

$$t = 0, \qquad X_0 = 0, \qquad S_0 = 40, \qquad Q_0 = 5000, \qquad a_0 = \theta_a, \qquad b_0 = \theta_b. \tag{3.3.6}$$

The price impact parameters in the Almgren-Chriss strategy (3.2.19) are constant, and we take them to be $(a_0, b_0) = (\theta_a, \theta_b)$. In total, we run $M = 10,000$ sample paths.

In our Monte Carlo simulations, we obtain

$$\frac{\widehat{\Phi}\left(\overline{\nu_0^*}\right) - \widehat{\Phi}\left(\nu_{AC}\right)}{\widehat{\Phi}\left(\nu_{AC}\right)} \cdot 10^4 = 6.0385, \qquad \frac{\widehat{\Phi}\left(\overline{\nu_1^*}\right) - \widehat{\Phi}\left(\overline{\nu_0^*}\right)}{\widehat{\Phi}\left(\overline{\nu_0^*}\right)} \cdot 10^4 = 0.0224, \tag{3.3.7}$$

in the nonlimiting case,

$$\frac{\widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right) - \widehat{\Phi}^{(\infty,\varphi)}\left(\nu_{AC}^{(\infty,\varphi)}\right)}{\widehat{\Phi}^{(\infty,\varphi)}\left(\nu_{AC}^{(\infty,\varphi)}\right)} \cdot 10^4 = 6.0367, \tag{3.3.8}$$

$$\frac{\widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_1^*}^{(\infty,\varphi)}\right) - \widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right)}{\widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right)} \cdot 10^4 = 0.0224, \tag{3.3.9}$$

in the limiting case $\kappa \to \infty$, and

$$\frac{\widehat{\Phi}^{(\infty,0)}\left(\overline{\nu_1^*}^{(\infty,0)}\right) - \widehat{\Phi}^{(\infty,0)}\left(\overline{\nu_0^*}^{(\infty,0)}\right)}{\widehat{\Phi}^{(\infty,0)}\left(\overline{\nu_0^*}^{(\infty,0)}\right)} \cdot 10^4 = 0.8131. \tag{3.3.10}$$

in the limiting case $(\kappa, \varphi) \to (\infty, 0)$.

Equations (3.3.7), (3.3.8), (3.3.9) and (3.3.10) demonstrate that in the nonlimiting case and both limiting cases, the trader gains a relative value increase from following the zeroth order strategy approximation over the Almgren-Chriss strategy and from following the first order strategy approximation over the zeroth order strategy approximation.

In Figures 3.2a and 3.2b, we plot histograms of the relative performance

$$\frac{\Phi(\overline{\nu_0^*}) - \Phi(\nu_{AC})}{\Phi(\nu_{AC})} \cdot 10^4, \qquad \frac{\Phi(\overline{\nu_1^*}) - \Phi(\overline{\nu_0^*})}{\Phi(\overline{\nu_0^*})} \cdot 10^4, \tag{3.3.11}$$

respectively, in Figures 3.3a and 3.3b we plot histograms of the relative performance

$$\frac{\Phi^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right) - \Phi^{(\infty,\varphi)}\left(\nu_{AC}^{(\infty,\varphi)}\right)}{\Phi^{(\infty,\varphi)}\left(\nu_{AC}^{(\infty,\varphi)}\right)} \cdot 10^4, \tag{3.3.12a}$$

$$\frac{\Phi^{(\infty,\varphi)}\left(\overline{\nu_1^*}^{(\infty,\varphi)}\right) - \Phi^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right)}{\Phi^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right)} \cdot 10^4, \tag{3.3.12b}$$

respectively, and in Figure 3.4, we plot a histogram of the relative performance

$$\frac{\Phi^{(\infty,0)}\left(\overline{\nu_1^*}^{(\infty,0)}\right) - \Phi^{(\infty,0)}\left(\overline{\nu_0^*}^{(\infty,0)}\right)}{\Phi^{(\infty,0)}\left(\overline{\nu_0^*}^{(\infty,0)}\right)} \cdot 10^4. \tag{3.3.13}$$

Figures 3.2a and 3.2b show that in addition to the expected value increases seen in (3.3.7), $\Phi_i(\overline{\nu_0^*}) > \Phi_i(\nu_{AC})$ and $\Phi_i(\overline{\nu_1^*}) > \Phi_i(\overline{\nu_0^*})$ more often than not. We see the same result in both the limiting cases $\kappa \to \infty$ and $(\kappa, \varphi) \to (\infty, 0)$.

For the chosen parameters (3.3.5) and initial conditions (3.3.6), the relative improvement gained by following a first order strategy approximation is muted compared to the relative improvement of the zeroth order strategy over Almgren-Chriss. When $(a_0, b_0) = (\theta_a, \theta_b)$, the price impact parameters typically hover around their respective long-run means, keeping the correction terms in $\overline{\nu_1^*}$, $\overline{\nu_1^*}^{(\infty,\varphi)}$ and $\overline{\nu_1^*}^{(\infty,0)}$ small. Let us keep the parameter values (3.3.5) but modify the initial conditions as follows

$$t = 0, \qquad X_0 = 0, \qquad S_0 = 40, \qquad Q_0 = 5000, \qquad a_0 = 1.5\theta_a, \qquad b_0 = 1.5\theta_b. \tag{3.3.14}$$

We note that the difference between the initial conditions (3.3.6) and (3.3.14) are the values of a_0 and b_0. With the initial conditions (3.3.14), the price impact processes a and b start above their long-run means θ_a and θ_b and will typically float downwards towards their respective means throughout the trading period. When the price impact parameters begin away from their long-run means, the correction terms present in the first order strategy approximations have a more pronounced influence on the trading strategy, and we see a larger relative improvement of the first order strategies over the zeroth order strategies.

We repeat the above experiments with the initial conditions (3.3.14) and obtain

$$\frac{\widehat{\Phi}\left(\overline{\nu_1^*}\right) - \widehat{\Phi}\left(\overline{\nu_0^*}\right)}{\widehat{\Phi}\left(\overline{\nu_0^*}\right)} \cdot 10^4 = 0.2682, \qquad \frac{\widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_1^*}^{(\infty,\varphi)}\right) - \widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right)}{\widehat{\Phi}^{(\infty,\varphi)}\left(\overline{\nu_0^*}^{(\infty,\varphi)}\right)} \cdot 10^4 = 0.2683,$$

$$\frac{\widehat{\Phi}^{(\infty,0)}\left(\overline{\nu_1^*}^{(\infty,0)}\right) - \widehat{\Phi}^{(\infty,0)}\left(\overline{\nu_0^*}^{(\infty,0)}\right)}{\widehat{\Phi}^{(\infty,0)}\left(\overline{\nu_0^*}^{(\infty,0)}\right)} \cdot 10^4 = 3.541.$$

The relative performance of the first order strategies increases by an order of magnitude when going from the initial conditions (3.3.6) to (3.3.14). In Figures 3.5, 3.6 and 3.7, we plot histograms of the relative performance (3.3.11) (right), (3.3.12) and (3.3.13), respectively, with the initial conditions (3.3.14). We see that in all three cases, the first order strategy out-performs the zeroth order strategy more often than not.

3.4 Conclusion

In this chapter, we have analyzed an optimal liquidation problem in the spirit of [3]. In contrast to [3] and other papers in this line of research, we have allowed for temporary and permanent price impact to vary stochastically in time, as is empirically observed. Although an explicit optimal liquidation strategy can not be found in the general case, we have been able obtain explicit approximations to the optimal liquidation strategy using a novel coefficient polynomial expansion technique.

One advantage of the analysis we have performed is that we have not needed to specify a particular price impact model. As long as the drift and diffusion coefficients that appear in our general framework are sufficiently differentiable, the approximations we have developed can be implemented.

Our theoretical analysis and numerical tests have led to some surprising insights. For instance, we have seen in Example 3.3.2 that in times of high price impact, it can be optimal for a trader who wishes to exit a position to buy shares of a stock in times of unusually high price impact in anticipation of selling these shares when the price impact returns to normal levels. Such a situation can never arise in a constant price impact setting.

While the numerical results we have presented support the use of our approximations, we have not yet established rigorous accuracy bounds. Such accuracy bounds, if obtained, would provide a theoretical justification for using the approximations we have developed. We believe that the techniques we have developed in this chapter can be applied to other stochastic control problems, particularly control problems that are generalizations of problems with explicit solutions. For example, our expansion techniques could potentially be applied to models for which the trading speed affects the asset price's volatility.

Our approach in this chapter was to model price impact effects via a decomposition into temporary and permanent price impact and modeling each as a linear function of the trader's order submission speed. However, there is more than one way to Wal-Mart, so to speak. In the next chapter, we incorporate permanent price impact into our model by directly modeling a market maker who seeks to set fair asset prices based on the aggregate order flow for an asset on an exchange.

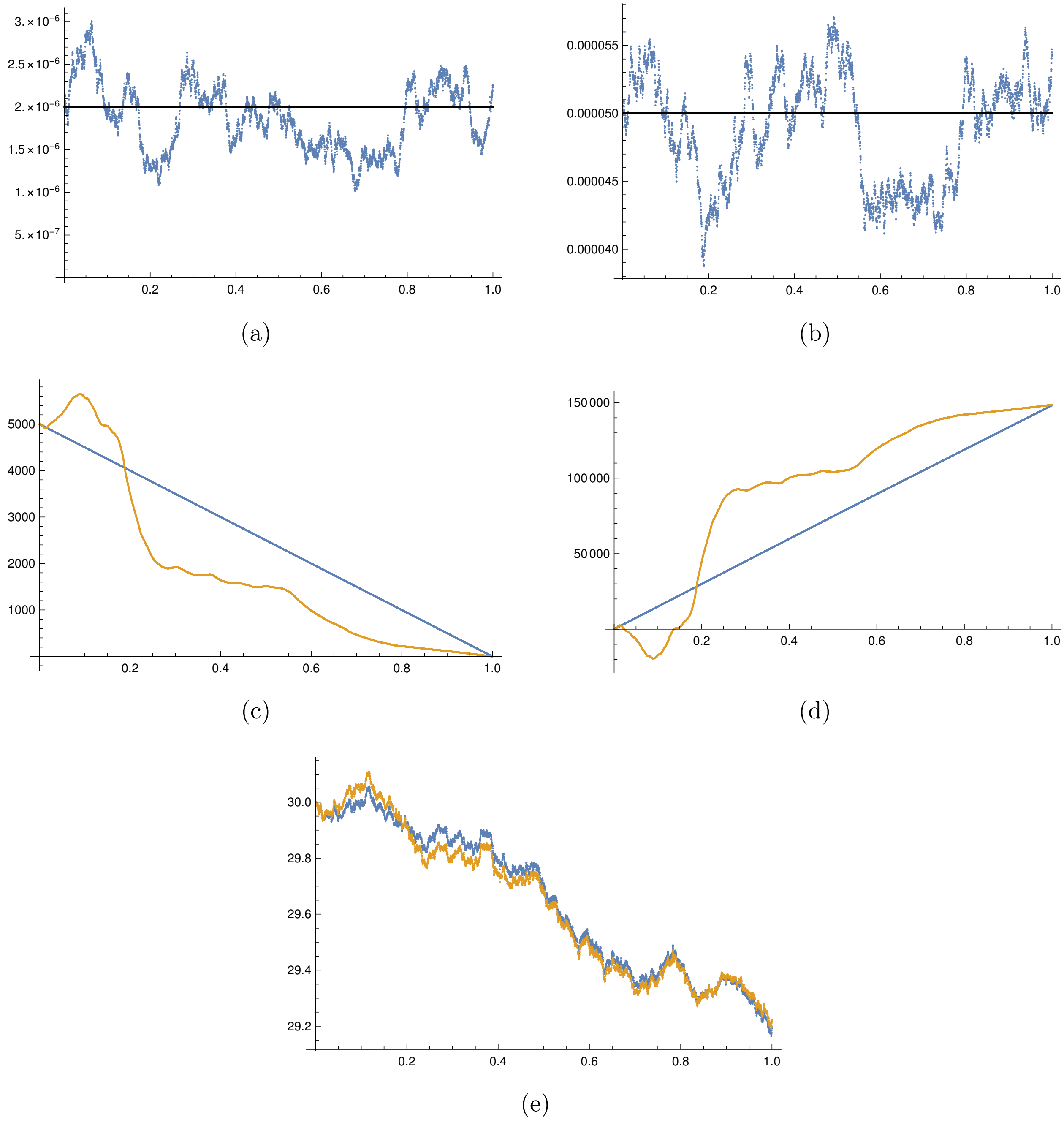
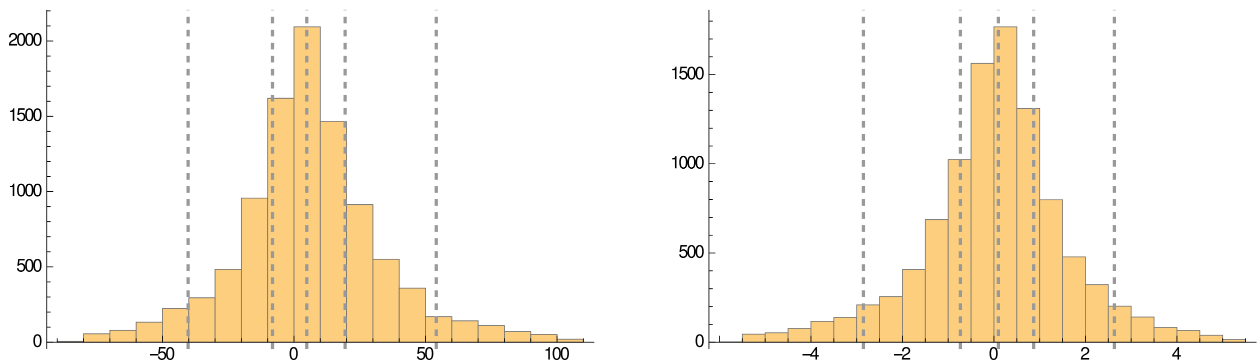
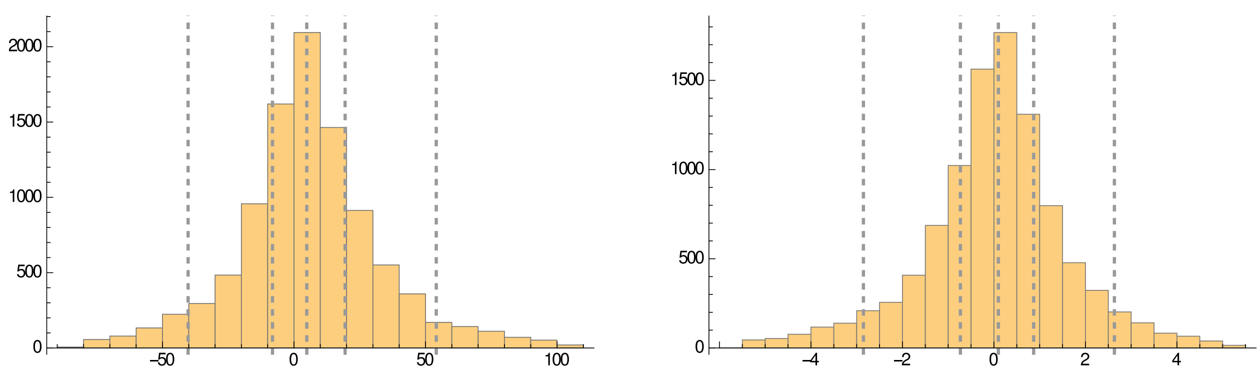


Figure 3.1: Here we plot a single sample path of (a, b) and the paths of (X^ν, S^ν, Q^ν) that result from following $\nu = \overline{\nu_0^*}^{(\infty,0)}$ (blue) and $\nu = \overline{\nu_1^*}^{(\infty,0)}$ (**orange**) with dynamics (3.3.2) and parameters (3.3.4). In Figure 3.1a, we plot the temporary price impact a, and in Figure 3.1b we plot the permanent price impact b. We plot the trader's inventory Q^ν in Figure 3.1c and the trader's cash X^ν in Figure 3.1d. In Figure 3.1e, we plot the stock prices S^ν.



(a) (b)

Figure 3.2: Here we plot histograms of the relative performance criteria given in (3.3.11) with the initial conditions (3.3.6). In Figure 3.2a we plot the performance of $\overline{\nu_0^*}$ relative to ν_{AC}. In Figure 3.2b, we plot performance of $\overline{\nu_1^*}$ relative to $\overline{\nu_0^*}$. The vertical, dashed lines represent the 5%, 25%, 50%, 75%, and 95% quantiles, respectively.



(a) (b)

Figure 3.3: Here we plot histograms of the relative performance criteria given in (3.3.12) with the initial conditions (3.3.6). In Figure 3.3a, we plot the performance (3.3.12a) of $\overline{\nu_0^*}^{(\infty,\varphi)}$ relative to $\nu_{AC}^{(\infty,\varphi)}$. In Figure 3.3b, we plot performance (3.3.12b) of $\overline{\nu_1^*}^{(\infty,\varphi)}$ relative to $\overline{\nu_0^*}^{(\infty,\varphi)}$. The vertical, dashed lines represent the 5%, 25%, 50%, 75%, and 95% quantiles, respectively.

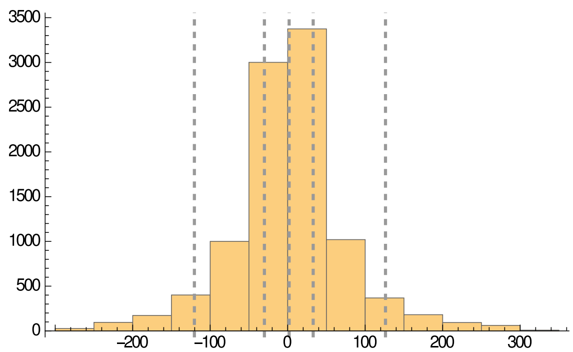


Figure 3.4: Here we plot the performance of $\overline{\nu_1^*}^{(\infty,0)}$ relative to $\overline{\nu_0^*}^{(\infty,0)}$ with respect to the performance criteria (3.3.13) with initial conditions (3.3.6). The vertical, dashed lines represent the 5%, 25%, 50%, 75%, and 95% quantiles, respectively.

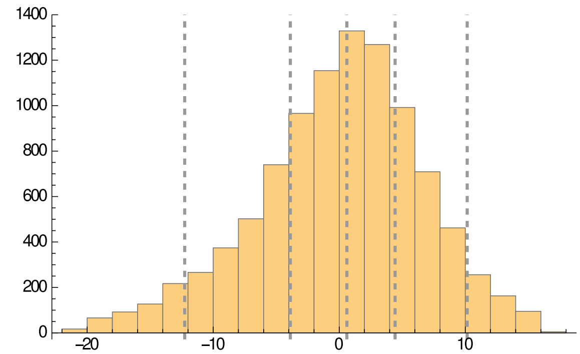


Figure 3.5: Here we plot the performance of $\overline{\nu_1^*}$ relative to $\overline{\nu_0^*}$ with respect to the performance criteria (3.3.11) (right) with initial conditions (3.3.14). The vertical, dashed lines represent the 5%, 25%, 50%, 75%, and 95% quantiles, respectively.

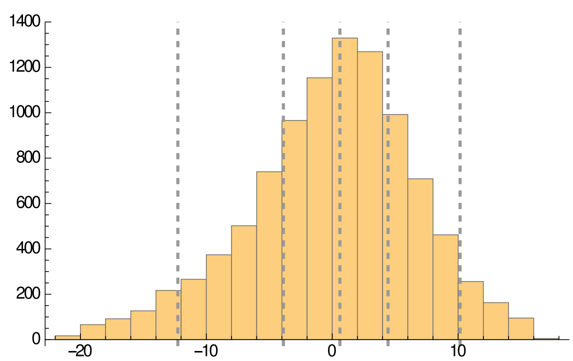


Figure 3.6: Here we plot the performance of $\overline{\nu_1^*}^{(\infty,\varphi)}$ relative to $\overline{\nu_0^*}^{(\infty,\varphi)}$ with respect to the performance criteria (3.3.12b) with initial conditions (3.3.14). The vertical, dashed lines represent the 5%, 25%, 50%, 75%, and 95% quantiles, respectively.

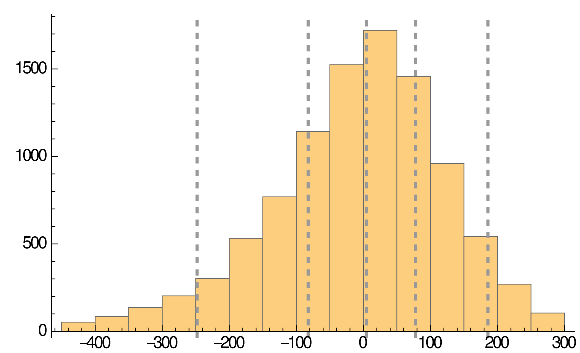


Figure 3.7: Here we plot the performance of $\overline{\nu_1^*}^{(\infty,0)}$ relative to $\overline{\nu_0^*}^{(\infty,0)}$ with respect to the performance criteria (3.3.13) with initial conditions (3.3.14). The vertical, dashed lines represent the 5%, 25%, 50%, 75%, and 95% quantiles, respectively.

Chapter 4

INSIDER TRADING WITH TEMPORARY PRICE IMPACT

4.1 Background

As discussed in Chapter 3, when traders place market orders on a securities exchange, they face transactional frictions. Direct frictions include exchange and brokerage fees, which are easily incorporated into a trading strategy. But, consumers of liquidity also experience indirect costs. Market makers adjust their limit orders to reflect the information contained in incoming market orders by moving the midprice of the asset in the direction of the market order. This adversely affects traders who submit sequences of market orders as the next market order will be transacted at a less favorable price. This effect is known as price impact. Additionally, if a market order is large enough, then it consumes all of the liquidity at the best available price and the remainder of the order is executed at sequentially worse prices. This can be thought of as a transaction cost which is dependent on the size of the market order and the state of the order book. In this chapter, we consider an insider trading model in the spirit of [45] for which a trader with inside information is risk-averse, faces risk-neutral market makers, and pays a transaction cost proportional to the size of their market order.

Risk-aversion and transaction costs have been previously studied in the insider trading literature. The authors of [41] extend the discrete-time model of [45] to include an exponentially risk-averse insider. Furthermore, their model allows for multiple insiders with the same level of risk-aversion who all receive identical information. [66] further extends the model of [41] to include a quadratic transaction cost for a risk-averse insider in discrete-time. [10] extends the continuous-time model which was first given by [45] and generalized by [7]. [26] considers an activist insider who is able to change the value of the asset after trading is complete by expending effort. [25] gives market equilibrium interactions between a trader with inside information and an uniformed trader performing portfolio rebalancing.

In this chapter, we model an exponentially risk-averse (or risk-neutral) insider who trades via market orders and faces a transaction cost that is proportional to the size of the order. We allow for the risk aversion parameter to be zero, which corresponds to a risk-averse insider. We first present a single-auction model and give the unique linear equilibrium. We show that the market maker's equilibrium pricing rule is the unique positive root of a polynomial equation whose coefficients depend on the model parameters. An asymptotic expansion of the roots of the aforementioned polynomial is performed for small transaction cost, which allows us to examine the effects of the transaction cost relative to frictionless models.

We then formulate an analogous model in continuous-time and present a linear equilibrium classified by the solution to a forward-backward ordinary differential equation (FBODE). We show that the resulting FBODE has a unique solution and is explicitly solvable when the insider is risk-neutral. When the insider is risk-averse and we cannot explicitly solve the FBODE. We then numerically analyze the effects of varying the model parameters on equilibrium using numerical solutions of the associated FBODE.

The models formulated in this chapter are also related to those of optimal execution literature, which we discussed in Chapter 3. Often in the optimal execution literature, the pressure on the asset price exerted by order flow is referred to as permanent price impact, while the immediate cost associated with market microstructure is referred to as temporary price impact. In Chapter 3, we modeled permanent and temporary price impact by defining two distinct price processes: the midprice and the transaction price. The midprice process represents the midpoint between the best quoted bid and ask prices set by liquidity providers, and the transaction price represents the average price per unit of asset at which the trader collects proceeds from trades. We modeled permanent impact by letting the drift of the midprice be an exogenous function of the trader's order flow, and temporary impact was modeled by defining the transaction price of trades to be equal to the midprice plus an exogenous function of the trader's order volume.

Our model also includes midprice and transaction price processes. As distinct from the model of Chapter 3, we directly model price setting market makers which allows for the permanent price impact to be endogenous. However, we define a transaction price process that is analogous the model of [4] by explicitly introducing an exogenous transaction cost. To be consistent with the insider trading literature, we refer the permanent price impact effect simply as price impact and temporary price impact effect as transaction cost.

The continuous-time version of our model is a direct generalization of the continuous-time models of the papers [45], [7] (when the insider's signal is Gaussian), and [10] (with constant volatility of noise trading). As such, some qualitative features of equilibrium in these works also arise in the present chapter, but there are also some notable differences. In particular, a key feature of many other models with asymmetric information is that the asset price is always fully revealing of the insider's signal at the end of the trading horizon (the insider always has incentive to exploit their informational advantage). This is not the case in our model when the transaction cost is positive. As a consequence, revelation of the insider's signal contains information not already incorporated in the publicly available price.

The rest of the chapter is organized as follows. In Section 4.2, we develop a single-auction model, present the unique linear equilibrium, and analyze the effects of the transaction cost by performing an asymptotic expansion on the equilibrium for small transaction cost. We subsequently shift our focus to continuous-time and begin by presenting a continuous-time model in Section 4.3. We begin Section 4.3.1 by developing the mathematical machinery necessary for the development of the linear equilibrium, and we finish the section by presenting a linear continuous-time equilibrium. In Section 4.4, we demonstrate the effects on the equilibrium of the previous section of varying the model parameters. Some concluding remarks are offered in Section 4.5.

4.2 Single-Auction

In this section, we consider a single-auction market where the realized price of the insider's trades is the midprice minus penalty which is linear with respect to the trade volume. After presenting the model that describes the dynamics of trade and the objective of the insider and market maker, we prove the existence of a unique equilibrium in this setting. We then investigate the effects of the transaction cost on the associated equilibrium by performing an asymptotic expansion of the equilibrium strategies.

The single-auction model in this section is similar to [66] for the case of a single agent. However, [66] never considers a model in which exponentially risk-averse insiders interact with noise traders. When the insider is risk-averse, the model of [66] assumes that the volume traded by the noise traders is directly observed by the insider before submitting their own trade, and the uncertainty in the model comes from an endowment to the insider. Some of the results included in this section are analogous to [66], but there are some distinctions. For example, our model guarantees equilibrium whereas the lack of noise traders in [66] can give rise to situations with no equilibrium (see Lemma 2 of [66]). We include the single-auction results so that we may perform a more in-depth analysis of the equilibrium through an asymptotic expansion, and also for the sake of completeness before investigating a continuous-time version of the model. The proof structure of the continuous-time equilibrium is directly analogous to that of the single auction case. However, the development of the continuous-time equilibrium requires more sophisticated mathematical machinery, and hopefully the reader will find the presentation of the continuous-time equilibrium more digestible after the single-auction appetizer.

4.2.1 Model

In the spirit of [45], we consider a single-auction on a market with one risky asset that is traded on an exchange with three types of traders: market makers who set the asset's midprice, an insider who has information about the future value of the asset, and noise traders. We let v denote the *ex-post* liquidation value of the asset, and we assume that $v \sim \mathcal{N}(v_0, \Sigma_0^v)$.

The insider receives the realization of v before the auction takes place, and this information is unavailable to the public. The insider wishes to utilize their informational advantage by trading the asset on auction. We denote the insider's order by Δx. We assume that the number of noise traders on the exchange is large, and we denote the aggregate order of the noise traders by Δz, which we assume to be distributed as $\Delta z \sim \mathcal{N}(0, \sigma^2)$. The market maker observes only the total quantity of orders Δy submitted to the exchange, where

$$\Delta y = \Delta x + \Delta z.$$

The auction takes place in two phases. First, the insider and the noise traders submit orders to the exchange. Second, the market maker observes the aggregate order Δy and sets the midprice p.

We assume that the insider pays a transaction cost proportional to the size of their order. The effective price at which the insider's trades are executed can be written as

$$\widehat{p} = p + c\Delta x,$$

where $c > 0$ is constant. We refer to $\widehat{p}$ as the insider's transaction price, and the constant c is referred to as the transaction cost parameter. The transaction cost is felt only by the insider. We also note that as the insider sells shares of the asset the transaction price is lower than the midprice, and, conversely, as the insider purchases shares the transaction price is higher than the midprice.

The discrepancy between the quoted price p and the transaction price $\widehat{p}$ can arise from a variety of sources. [66] views the transaction cost as a tax imposed by a regulator. In previous works, it is assumed that a large number of risk-neutral market makers are in perfect competition, and together they all quote the same price. On real exchanges, market makers may have difference price preferences, creating a demand structure that depend on the totality of their quotes. We capture this behavior by including a linear penalty on trade volume in lieu of modeling this phenomenon directly.

We assume without a lose of generality that the insider holds no shares of the asset prior to the auction. The wealth of the insider after the trades are executed is thus

$$w = (v - \widehat{p})\Delta x.$$

The insider would like to maximize the utility of their expected wealth w. That is, the insider chooses Δx to achieve

$$\max_{\Delta x} \mathbb{E}\left[U(w)|v\right], \tag{4.2.1}$$

where the insider's utility function U is increasing and concave. The market maker is tasked with setting prices efficiently. That is, the market maker must choose the price p such that

$$p = \mathbb{E}\left[v|\Delta y\right]. \tag{4.2.2}$$

4.2.2 Single-Auction Equilibrium

We begin this section by defining what it means for a pricing rule and a trading strategy to form an equilibrium. A linear equilibrium is then given for an exponentially risk-averse (or risk-neutral) insider.

We assume that the market maker and insider choose pricing rules and trading strategies, respectively, as functions of information available to them during the auction. That is, the market maker chooses the price p as a function of Δy, and the insider chooses their order size Δx as a function of v. Let P and X be functions such that $p = P(\Delta y)$ and $\Delta x = X(v)$.

Definition 4.2.1. A *single-auction equilibrium* (P, X) consists of a pricing rule P and trading strategy X such that

- given a pricing rule P, the order Δx given by the trading strategy $X(v)$ achieves the maximum in (4.2.1), and

- given a trading strategy X, the price p given by the pricing rule $P(\Delta y)$ satisfies the efficiency condition (4.2.2).

Suppose that $\mathcal{P}$ is the set of pricing rules $P \equiv P(\Delta y)$ and let $\mathcal{P}_0 \subset \mathcal{P}$ be the set of pricing rules for which there exists a corresponding trading strategy X that satisfies (4.2.1). Similarly, let $\mathcal{X}$ be the set of trading strategies $X \equiv X(v)$ and let $\mathcal{X}_0 \subset \mathcal{X}$ be the set of trading strategies for which there exists a pricing rule P that satisfies (4.2.2). The sets $\mathcal{P}_0$ and $\mathcal{X}_0$ induce mappings $\rho : \mathcal{P}_0 \to \mathcal{X}$ and $\xi : \mathcal{X}_0 \to \mathcal{P}$. A pricing rule P and trading strategy X form a single-auction equilibrium if $\rho(P) \in \mathcal{X}_0$, $\xi(X) \in \mathcal{P}_0$ and $(P, X) = (\xi(X), \rho(P))$.

We will restrict our focus to the risk-neutral and exponentially risk-averse insider. For any constant $A \geq 0$, we define the utility function

$$U(w) = \begin{cases} w, & A = 0 \\ -\exp\left(-Aw\right), & A > 0 \end{cases} . \tag{4.2.3}$$

We refer to A as the risk aversion parameter. The insider's optimal strategy depends on their chosen utility of wealth function U. When $A = 0$ and $U(w) = w$, the insider wishes only to maximize their expected post-auction wealth and is completely indifferent to the risk posed by the noise trader's influence on the market maker's price $P(\Delta y)$. That is, they are risk-neutral. Alternatively, when constant $A > 0$ and $U(w) = -\exp(-Aw)$, then, intuitively, the insider penalizes outcomes that lead to lower wealth more than they reward outcomes that lead to higher wealth. The severity of this mismatched outcome weighting is controlled by A, where higher values of A correspond to an insider who is more averse to risk.

It is helpful to introduce the constants

$$\lambda_K = \frac{1}{2}\sqrt{\frac{\Sigma_0^v}{\sigma^2}}, \qquad\qquad \beta_K = \frac{1}{2\lambda_K}.$$

before presenting the single-auction, linear equilibrium. The constants λ_K and β_K correspond to the single-auction pricing rule and trading strategy, respectively, of [45]. In that work, the market maker's single-auction equilibrium pricing rule for a risk-neutral insider is $P_K(\Delta y) = v_0 + \lambda_K \, \Delta y$, and the corresponding optimal trading strategy is $X_K(v) = \beta_K(v - v_0)$. We will write the equilibrium pricing rule and trading strategy of our model in terms of λ_K which will help to illuminate the effect of the added transaction cost.

Theorem 4.2.2. *Choose $A \geq 0$ and let U, defined in (4.2.3), be the insider's utility function. Then the unique linear single-auction equilibrium is given by*

$$P(\Delta y) = v_0 + \lambda \, \Delta y \, ,$$

$$X(v) = \beta\,(v - v_0)$$

where

$$\beta = \frac{1}{2\,(\lambda + c) + A\,\sigma^2\,\lambda^2}\,, \tag{4.2.4}$$

and λ is the unique positive root of the polynomial

$$r(x) = A^2\,\sigma^4\,x^5 + 4\,A\,\sigma^2\,x^4 + 4\,(1 + A\,c\,\sigma^2)\,x^3 + 4\,(2\,c - A\,\sigma^2\,\lambda_K^2)\,x^2 + 4\,(c^2 - \lambda_K^2)\,x - 8\,c\,\lambda_K^2\,. \tag{4.2.5}$$

Proof. First we show that r has a unique positive root. Define the function

$$s(\lambda) = 2\,(\lambda + c) + A\,\sigma^2\,\lambda^2\,.$$

We will see later that $s(\lambda) > 0$ is the second order condition required for optimality of the trading strategy. The polynomial r can be written as

$$r(\lambda) = \lambda\,(s^2(\lambda) + 4\,\lambda_K^2) - 4\,\lambda_K^2\,s(\lambda)\,,$$

and we note that $r(\lambda) = 0$ if and only if

$$\lambda\,(s^2(\lambda) + 4\,\lambda_K^2) = 4\,\lambda_K^2\,s(\lambda)\,. \tag{4.2.6}$$

The left and right sides of (4.2.6) are both polynomials in λ with positive coefficients, and therefore are both strictly increasing functions of $\lambda \geq 0$. Furthermore, for any $A \geq 0$ the degree of the left hand side polynomial is strictly greater than the degree of the right hand side polynomial, and when $\lambda = 0$ the left hand side of (4.2.6) is zero and the right hand side is positive. Therefore, (4.2.6) has exactly one positive solution, and hence r has exactly one positive root.

The rest of the proof is divided into the two cases $A = 0$ and $A > 0$. Some related expressions are different between the two cases, but the structure of the proofs are identical. First, we suppose that the insider's trading strategy is a linear function of v and show the efficient pricing rule is linear. Then, we suppose the market maker's pricing rule is linear and show that the insider's optimal trading strategy is linear. Finally, matching coefficients gives the result.

Case $A = 0$: First, we suppose that the market maker chooses p as a linear function of Δy. Namely, we let $P(\Delta y) = \mu + \lambda\,\Delta y$, where μ and λ are constants. When the market maker follows the pricing rule P, the insider's transaction price is given by

$$\widehat{p} = \mu + \lambda\,\Delta y + c\,\Delta x = \mu + (\lambda + c)\,\Delta x + \lambda\,\Delta z\,. \tag{4.2.7}$$

Thus,

$$\begin{aligned}\mathbb{E}[U(w)\,|\,v] = \mathbb{E}[(v - \widehat{p})\,\Delta x|\,v] &= \mathbb{E}[(v-\mu)\,\Delta x - (\lambda + c)\,\Delta x^2 - \lambda\,\Delta x\,\Delta z\,|\,v] \\ &= (v-\mu)\,\Delta x - (\lambda + c)\,\Delta x^2\,. \end{aligned} \tag{4.2.8}$$

The second order condition for optimality is $\lambda + c > 0$, which is equivalent to $s(\lambda) > 0$. Assuming this is satisfied, the choice of

$$\Delta x \equiv X(v) = \frac{v - \mu}{2\,(\lambda + c)}\,, \tag{4.2.9}$$

maximizes (4.2.8) for any v. We note that for linear pricing rules P linear trading strategies X are optimal even if we allow X to be a nonlinear function.

Now we assume that the insider chooses Δx to be a linear function X of the *ex-post* price v. Specifically, the insider chooses $X(v) = \alpha + \beta\,v$, where α and β are constants. Then, by the projection theorem for normal random variables we have

$$\begin{aligned}\mathbb{E}[v\,|\,\Delta y] = \mathbb{E}[v\,|\,\alpha + \beta\,v + \Delta z] &= \mathbb{E}[v] + \frac{\mathbb{E}[(v - \mathbb{E}[v])(\Delta y - \mathbb{E}[\Delta y])]}{\mathbb{E}[(\Delta y - \mathbb{E}[\Delta y])^2]}(\Delta y - \mathbb{E}[\Delta y]) \\ &= v_0 + \frac{4\,\lambda_K^2\,\beta}{1 + 4\,\lambda_K^2\,\beta^2}(\Delta y - \alpha - \beta\,v_0) \\ &= \frac{v_0 - 4\,\lambda_K^2\,\alpha\,\beta}{1 + 4\,\lambda_K^2\,\beta^2} + \frac{4\,\lambda_K^2\,\beta}{1 + 4\,\lambda_K^2\,\beta^2}\,\Delta y\,. \end{aligned} \tag{4.2.10}$$

By examining (4.2.9) and (4.2.10), we see that for $(P, X) \equiv (P(\Delta y), X(v)) = (\mu + \lambda\,\Delta y, \alpha + \beta\,v)$ to be an equilibrium, it must be that

$$\alpha = -\beta\,\mu, \qquad \mu = \frac{v_0 - 4\,\lambda_K^2\,\alpha\,\beta}{1 + 4\,\lambda_K^2\,\beta^2}\,, \tag{4.2.11a}$$

$$\beta = \frac{1}{2\,(\lambda + c)}, \qquad \lambda = \frac{4\,\lambda_K^2\,\beta}{1 + 4\,\lambda_K^2\,\beta^2}\,, \tag{4.2.11b}$$

subject to the constraint $s(\lambda) > 0$.

Immediately, we get from (4.2.11a) that in equilibrium $\mu = v_0$ and $\alpha = -\beta\,v_0$, which gives the insider's trading strategy $X(v) = \beta\,(v - v_0)$. Recall that the insider's second order condition is satisfied if $s(\lambda) > 0$. Inserting β into the expression for λ in (4.2.11b), we see that λ satisfies (4.2.6). If $\lambda \leq 0$ then from (4.2.6) we see that $s(\lambda) \leq 0$, contradicting optimality of Δx. Therefore, λ is the unique positive root of the polynomial r.

Case $A > 0$: First, suppose that the market maker chooses p as a linear function of Δy so that $P(\Delta y) = \mu + \lambda\Delta y$, where μ and λ are constants. When the market maker follows the pricing strategy P, then the insider's transaction price $\widehat{p}$ is given by (4.2.7). Thus,

$$\begin{aligned} \mathbb{E}[U(w)|v] &= \mathbb{E}[-\exp\{-A(v - \widehat{p})\Delta x\}|v] \\ &= -\exp\{-A(v-\mu)\Delta x + A(\lambda + c)\Delta x^2 + \tfrac{1}{2}A^2\lambda^2\sigma^2\Delta x^2\}\,. \end{aligned} \tag{4.2.12}$$

The second order condition for optimality is $2(\lambda + c) + A\sigma^2\lambda^2 > 0$ which is equivalent to $s(\lambda) > 0$. Assuming this is satisfied, the choice of

$$\Delta x \equiv X(v) = \frac{v - \mu}{2\,(\lambda + c) + A\,\sigma^2\,\lambda^2} \tag{4.2.13}$$

maximizes (4.2.12) for any v. We note that for linear pricing rules P linear trading strategies X are optimal even if we allow X to be a nonlinear function.

Now, suppose that the insider chooses the linear trading strategy $X(v) = \alpha + \beta\,v$. Then $\mathbb{E}[v\,|\,\Delta y]$ is given by (4.2.10).

By examining (4.2.10) and (4.2.13), we see that for $(P, X) \equiv (P(\Delta y), X(v)) = (\mu + \lambda\, \Delta y, \alpha + \beta\, v)$ to be an equilibrium we must have

$$\alpha = -\beta\, \mu, \qquad \mu = \frac{v_0 - 4\,\lambda_K^2\, \alpha\, \beta}{1 + 4\,\lambda_K^2\, \beta^2}\,,$$
$$\beta = \frac{1}{2\,(\lambda + c) + A\, \sigma^2\, \lambda^2}\,, \qquad \lambda = \frac{4\,\lambda_K^2\, \beta}{1 + 4\,\lambda_K^2\, \beta^2}\,, \tag{4.2.14a}$$

where $s(\lambda) > 0$. The rest of the proof is identical to the case $A = 0$. □

Intuitively, a risk-averse trader prefers to submit a smaller than the risk-neutral trader in order to reduce the risk posed by noise traders. So one might expect that the insider's order size decreases in A. Correspondingly, less information about the true value of the asset would be contained in the order signal Δy received by the market maker, and they thus reduce the severity of the price adjustment. One might also expect that as the transaction cost parameter c increases it becomes less worth it for a the insider to submit larger orders, regardless of their risk preference, thus leading to a smaller β and λ. Both of these statements are indeed true and are summarized by the following proposition.

Proposition 4.2.3. *Let $P(\Delta y) = v_0 + \lambda \Delta y$ and $X(v) = \beta(v - v_0)$ be the pricing rule and trading strategy forming the unique single-auction equilibrium given by Theorem 4.2.2. If $A > 0$ then*

$$\frac{\partial \beta}{\partial c} < 0\,, \qquad \frac{\partial \beta}{\partial A} < 0\,,$$
$$\frac{\partial \lambda}{\partial c} < 0\,, \qquad \frac{\partial \lambda}{\partial A} < 0\,.$$

Proof. In the proof of Theorem 4.2.2, we saw that when $A > 0$ the equilibrium β and λ satisfy (4.2.14a). Inserting λ into β in (4.2.14a) we get that the equilibrium β satisfies

$$q(\beta) := 32c\lambda_K^4 \beta^5 + 16\lambda_K^4 \beta^4 + 16\lambda_K^2 \left(A\lambda_K^2 \sigma^2 + c\right) \beta^3 + 2c\beta - 1 = 0. \tag{4.2.15}$$

We see that for any fixed $x > 0$, $\partial q(x)/\partial c > 0$, implying that β decreases as the transaction cost c increases. Furthermore, for any fixed $x > 0$, $\partial q(x)/\partial A > 0$, which implies that β

decreases as the risk-aversion parameter increases. Taking the derivative of λ in (4.2.14a) with respect to A we see that

$$\frac{\partial \lambda}{\partial A} < 0 \qquad \Longleftrightarrow \qquad \beta < \frac{1}{2\lambda_K} = \beta_K,$$

where β_K is the insider's strategy in the case that $(A, c) = (0, 0)$ i.e. the model of [45]. As β strictly decreases in both A and c, we have that $\partial\lambda/\partial A < 0$. Applying the same argument for the transaction cost shows that $\partial\lambda/\partial c < 0$. We note that when $A = 0$ then β and λ are both still decreasing in c. □

It is natural to ask how the addition of a transaction cost affects the pricing rule relative to the Kyle pricing rule λ_K. The classification of equilibrium given by Theorem 4.2.2 is in terms of a root of a fifth degree polynomial. In general the quantities λ and β involved in the equilibrium will not have closed form expressions in terms of model parameters. However, we are able to find approximations to these quantities which hold when certain model parameters are small.

Let λ be the pricing rule of Theorem 4.2.2. Suppose that $c \ll 1$ and that λ can be expanded as a power series in c as

$$\lambda = \sum_{i=0}^{\infty} c^i \lambda_i,$$

where the λ_i are constants to be determined. We know that $r(\lambda) = 0$, where r is given in (4.2.5). We formally expand $r(\lambda)$ in c and collect like powers of c. As $r(\lambda) = 0$ for all c, we get a sequence of algebraic problems corresponding to each power of c. Explicitly, at the first three orders we have

$$O(c^0): \quad \frac{1}{4}A^4\sigma^4\lambda_0^5 + A^2\sigma^2\lambda_0^4 + \lambda_0^3 - A^2\lambda_K^2\sigma^2\lambda_0^2 - \lambda_K^2\lambda_0 = 0 \tag{4.2.16}$$

$$O(c^1): \quad \lambda_1\left(\frac{1}{4}\lambda_0^2\left(A^2\sigma^2\lambda_0 + 2\right)\left(5A^2\sigma^2\lambda_0 + 6\right) - \lambda_K^2\left(2A^2\sigma^2\lambda_0 + 1\right)\right) + \lambda_0^2\left(A^2\sigma^2\lambda_0 + 2\right) - 2\lambda_K^2 = 0, \tag{4.2.17}$$

$$O(c^2): \quad q_2(\lambda_2; \lambda_0, \lambda_1) = 0 \tag{4.2.18}$$

where q_2 is a polynomial in λ_2 whose coefficients depend on λ_0 and λ_1. We omit the expression for q_2 as it is cumbersome in the case $A > 0$. However, when $A = 0$ we have

$$q_2|_{A=0}(\lambda_2; \lambda_0, \lambda_1) = 4\left(3\lambda_0^2 - \lambda_K^2\right)\lambda_2 + 4\lambda_0\left(3\lambda_1^2 + 4\lambda_1 + 1\right)$$

We note that the algebraic problem (4.2.16) is $r|_{c=0}(\lambda_0) = 0$. It is still the case that $r|_{c=0}$ has a unique positive root, and that root corresponds to the market maker's equilibrium strategy when $c = 0$. When $A = 0$, $\lambda_0 = \lambda_K$, and when $A > 0$ then λ_0 is the market maker's strategy in [41]. Taking λ_0 to be the single positive root of $r|_{c=0}$, we solve (4.2.17) and (4.2.18) to get that

$$\lambda = \begin{cases} \lambda_K - \frac{c^2}{2\lambda_K} + O(c^3), & A = 0 \\ \lambda_0 + c\left(\frac{4\lambda_0^2\left(A^2\sigma^2\lambda_0+2\right)-8\lambda_K^2}{\lambda_K^2(8A^2\sigma^2\lambda_0+4)-\lambda_0^2(A^2\sigma^2\lambda_0+2)(5A^2\sigma^2\lambda_0+6)}\right) + O(c^2), & A > 0 \end{cases}.$$

Inserting λ into (4.2.4) and expanding in c gives

$$\beta = \begin{cases} \frac{1}{2\lambda_K} - \frac{c}{2\lambda_K^2} + \frac{3c^2}{4\lambda_K^3} + O(c^3) & A = 0 \\ \frac{1}{2\lambda_0 + A\sigma^2\lambda_0^2} - \frac{2c\left(A\lambda_0\lambda_1\sigma^2+\lambda_1+1\right)}{\lambda_0^2(A\lambda_0\sigma^2+2)^2} + O(c^2) & A > 0 \end{cases}.$$

One of the most striking features of the expansion above is that when the insider is risk-neutral $\lambda - \lambda_K = O(c^2)$. That is, the introduction of a transaction has a small impact on the pricing strategy, even though the insider's adjustment $\beta - \beta_K = O(c)$. Denote the N-th order approximation of β and λ, respectively, as

$$\overline{\beta}_N = \sum_{i=0}^{N} c^i \beta_i, \qquad \overline{\lambda}_N = \sum_{i=0}^{N} c^i \lambda_i.$$

In Figures 4.1 and 4.2, we plot the relative third order error $(\beta - \overline{\beta}_2)/\beta$ and $(\lambda - \overline{\lambda}_2)/\lambda$, respectively, as functions of c for $A = 0$ to demonstrate the accuracy of our approximations. Performing the above expansion in the polynomial (4.2.15) for equilibrium β yields the same result.

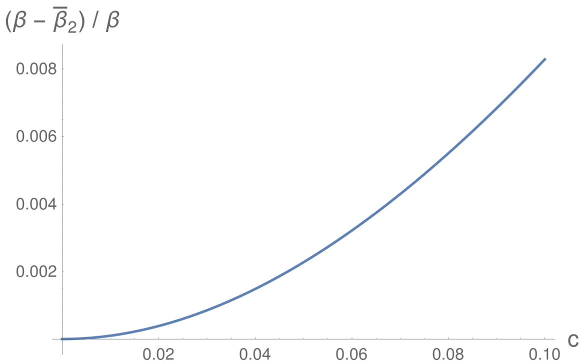


Figure 4.1: Here we take $(A, \sigma, \lambda_K) = (0, 0.5, 1)$ and plot $(\beta - \overline{\beta}_2)/\beta$ has a function of c.

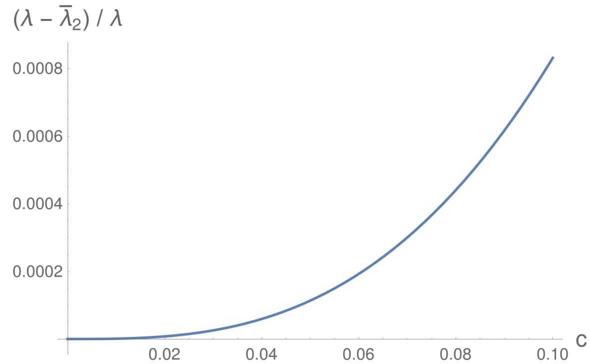


Figure 4.2: Here we take $(A, \sigma, \lambda_K) = (0, 0.5, 1)$ and plot $(\lambda - \overline{\lambda}_2)/\lambda$ has a function of c.

4.3 Continuous-Time Auction

In this section, we present a continuous-time model which incorporates a transaction cost which is analogous to the transaction cost introduced in the previous section. After introducing the model in Section 4.3, we give the notion of market equilibrium that we consider, a notion which is analogous to that of the single-auction. We then establish some mathematical machinery that is necessary for presenting equilibrium in continuous-time and then solve for linear equilibrium in closed form for a risk-neutral insider. When the insider is risk-averse, the equilibrium is given as a solution to an ODE, which we prove has a unique solution. Finally, we numerically analyze the dependence on the model parameters of the equilibrium solution.

Consider an exchange in which orders for a single risky asset are submitted in continuous-time, and assume that there are three types of traders: market makers who set the midprice of the asset, an insider who has information about the future value of the asset, and noise traders. We let the time T value of the asset v be normally distributed with mean v_0 and variance Σ_0^v, and we suppose that the insider is privy to the realization of v at the initial time $t = 0$.

Let $P = (P_t)_{0 \le t \le T}$ be the midprice of the asset, let $X = (X_t)_{0 \le t \le T}$ be the insider's inventory, and let $Z = (Z_t)_{0 \le t \le T}$ be a standard Brownian motion. We model the cumulative orders submitted by the noise traders up to time t as σZ_t, where $\sigma > 0$ is constant. We denote by $Y = (Y_t)_{0 \le t \le T}$ the total number of shares submitted to the market up to time t. That is,

$$Y_t = X_t - X_0 + \sigma Z_t. \tag{4.3.1}$$

Similarly to the single-auction case, we assume that the market maker observes the aggregate order flow Y but neither of the components X or Z.

It is helpful to explicitly denote the sets of information available to the various market participants at a given time. We define the filtrations $\mathcal{F}_t^Y$ and $\mathcal{F}_t^Z$ to be the filtrations generated by the processes Y and Z, respectively. As the market maker observes only Y, we define $\mathcal{F}_t^M = \mathcal{F}_t^Y$ to be the market maker's information. The insider is aware of their own trades, so they can back out the liquidity trades by observing historical prices (see [7]), and the insider has received the realization of v before trading begins. So we set the insider's information to be $\mathcal{F}_t^I = \sigma(\mathcal{F}_t^Z \cup \sigma(v))$.

We assume that insider only considers trading strategies that yield an almost surely absolutely continuous inventory process. To this end, we write

$$\mathrm{d}X_t = \theta_t \, \mathrm{d}t, \tag{4.3.2}$$

where θ is $(\mathcal{F}_t^I)_{0 \le t \le T}$ adapted, and, analogously to the single-auction model, we define the insider's transaction price $\widehat{P} = (\widehat{P}_t)_{0 \le t \le T}$ to be

$$\widehat{P}_t = P_t + c\theta_t, \tag{4.3.3}$$

where c is a positive constant. Taking X to have the form (4.3.2) allows us to formulate the continuous-time transaction price analogously to the transaction price in single-auction model. However, the assumption that there exists a θ such that X is given by (4.3.2) is not particularly restrictive. It is often the case that even when assumptions on X are relaxed the equilibrium inventory processes have the form (4.3.2). For example, in [7] the author assumes a general distribution of v and assumes only that the insider's inventory process X is a semi-martingale and gives an equilibrium X, which is unique in a certain class, that has the form (4.3.2).

At the end of the trading horizon, the insider's wealth is $X_T v$ minus the cost of trading throughout the period. Thus, the terminal wealth W_T is given by

$$W_T = X_T v - \int_0^T \widehat{P}_s \theta_s \, \mathrm{d}s. \tag{4.3.4}$$

The insider chooses a strategy to maximize their expected utility of terminal wealth. Expressly, θ is chosen to achieve

$$\sup_{\theta \in \mathcal{A}} \mathbb{E}\left[U(W_T) \middle| \mathcal{F}_0^I\right], \tag{4.3.5}$$

where U is an increasing, concave function, and the set of admissible strategies $\mathcal{A}$ is given by

$$\mathcal{A} = \{\theta \mid \theta \text{ is } (\mathcal{F}_t^I)_{0 \le t \le T} \text{ adapted and } \mathbb{E}\left[\int_0^T \theta_t^2 \mathrm{d}t\right] < \infty \text{ a.s.}\}$$

is the set of admissible trading strategies.

The market maker is tasked with setting the price of the asset efficiently at all times $0 \le t \le T$ and so should choose the price according to

$$P_t = \mathbb{E}[v \mid \mathcal{F}_t^M]. \tag{4.3.6}$$

We now define the concept of equilibrium in continuous-time. This concept is analogous to the single-auction equilibrium concept of Section 4.2.2.

Definition 4.3.1. A *continuous-time equilibrium* (P, X) consists of a price process P and an inventory process X such that

- given a price process P, the inventory process X achieves the supremum in (4.3.5), and
- given an inventory process X, the price process P is efficient, i.e. P satisfies (4.3.6).

4.3.1 Continuous-Time Equilibrium

The goal of this section is to present a linear equilibrium in continuous-time for an exponentially risk-averse or risk-neutral insider. We will restrict our consideration to pricing rules and trading strategies that have a form which are analogous to those of the single-auction equilibrium presented in Theorem 4.2.2. Namely, increments of the midprice set by the market maker will be linear with respect to increments of trade volume, and the insider's trading strategy will be linear with respect to the price mismatch $v - P_t$. Before stating the classification of equilibrium in continuous-time we need to develop some additional mathematical machinery. These are contained in Lemmas 4.3.2, 4.3.3, and 4.3.5.

Using expressions (4.3.2), (4.3.3), and (4.3.4), we rewrite the expression for terminal wealth W_T as

$$W_T = \int_0^T (v - P_s - c\theta_s)\,\theta_s\,\mathrm{d}s,$$

where we have taken $X_0 = 0$. The computations that follow can be carried out with $X_0 \neq 0$. This choice will not affect the insider's optimal trading strategy, but, rather, for risk-neutral or an exponentially risk-averse insider taking $X_0 \neq 0$ will scale or shift the insider's value function.

We will develop optimal insider trading strategies by first fixing the market maker's pricing rule and then solving the HJB equation associated with the optimization problem (4.3.5). We will subsequently verify that the solution is optimal. This motivates us to introduce the dynamic version of 4.3.5 as

$$H(t,P) = \sup_{\theta\in\mathcal{A}} \mathbb{E}\left[U\left(\int_t^T (v - P_s - c\,\theta_s)\,\theta_s\,ds \right) \middle| \mathcal{F}_t^I \right]. \tag{4.3.7}$$

For a given time $t \in [0,T]$ and midprice P_t, the quantity $H(t,P_t)$ gives the optimal expected utility that the insider can achieve by trading during the time interval $[t,T]$. We refer to H as the insider's value function.

Lemma 4.3.2. *Let $A \geq 0$ and let the utility function U be as defined in (4.2.3). Let λ be a positive, bounded, deterministic function such that the Riccati differential equation*

$$\frac{dh(t)}{dt} = -\frac{(1 - 2\,A\,c\,\sigma^2)\lambda^2(t)}{c}\,h^2(t) + \frac{\lambda(t)}{c}\,h(t) - \frac{1}{4\,c}\,, \qquad h(T) = 0\,, \tag{4.3.8}$$

has a global solution $h : [0,T] \to \mathbb{R}$. Suppose the midprice process P is given by

$$P_t = v_0 + \int_0^t \lambda(s)\,dY_s,$$

where Y is given in (4.3.1). Then the insider's value function (4.3.7) is given by

$$H(t,P) = \begin{cases} -\exp\left\{-A\Big((v-P)^2\,h(t) + \sigma^2\,\int_t^T \lambda^2(s)\,h(s)\,ds\Big)\right\}\,, & A > 0\,, \\ (v-P)^2\,h(t) + \sigma^2\,\int_t^T \lambda^2(s)\,h(s)\,ds\,, & A = 0\,, \end{cases} \tag{4.3.9}$$

and the optimal strategy in feedback form is given by

$$\theta_t^* = \beta(t)\,(v - P_t), \qquad \beta(t) = \frac{1 - 2\,\lambda(t)\,h(t)}{2\,c}\,.$$

Proof. We consider the cases $A = 0$ and $A > 0$ separately, however the proof strategy is the same in both cases. We write down the dynamic programming equation, the HJB PDE, associated with the insider's control problem. We then obtain a candidate optimal control in feedback control and reduce the HJB PDE to a Riccati ODE. Finally, we verify that the candidate control is indeed optimal via an application of the verification theorem. The reader may wish to refer to Appendix A for a brief review of dynamic programming in continuous time, the HJB equation, and the verification theorem.

Case $A = 0$: When $A = 0$, $U(w) = w$, and the value function (4.3.7) becomes

$$H(t,P) = \sup_{\theta \in \mathcal{A}} \mathbb{E}\left[\int_t^T (v - P_s - c\,\theta_s)\,\theta_s\,ds \,\middle|\, \mathcal{F}_t^I\right].$$

Associated with this stochastic control problem is the HJB partial differential equation

$$\partial_t H + \sup_{\theta}\left\{\frac{1}{2}\,\sigma^2\,\lambda^2(t)\,\partial_{PP}H + \theta\,\lambda(t)\,\partial_P H + (v - P - c\,\theta)\,\theta\right\} = 0\,, \tag{4.3.10}$$

$$H(T, \cdot) = 0\,.$$

We check by direct substitution that the solution of this equation is given by (4.3.9) when h satisfies the ODE (4.3.8). The supremum in (4.3.10) is achieved at

$$\theta^*(t, P) = \beta(t)\,(v - P), \qquad\qquad \beta(t) = \frac{1 - 2\,\lambda(t)\,h(t)}{2\,c}\,.$$

All that remains to be shown is that this feedback form of θ^* yields an admissible trading strategy. Optimality of θ^* then follows from a standard verification argument (see [60]). To this end, define an auxiliary process $Q = (Q_t)_{0 \le t \le T}$ by

$$Q_t = P_t - v\,.$$

Then under the control θ^*, the dynamics of Q are given by

$$dQ_t = -\lambda(t)\,\beta(t)\,Q_t\,dt + \lambda(t)\,\sigma\,dZ_t\,, \qquad\qquad Q_0 = v_0 - v\,.$$

This stochastic differential equation is linear and therefore has a unique strong solution. Further, since $v_0 - v$ is Gaussian, the resulting solution is a Gaussian process (see [43] Section 5.6). This gives

$$\mathbb{E}\left[\int_0^T (\theta_s^*)^2\,ds\right] = \mathbb{E}\left[\int_0^T \beta^2(s) Q_s^2\,ds\right] < \infty\,.$$

In addition, Q has continuous paths and thus the trading strategy is predictable, and therefore admissible.

<u>Case $A > 0$</u>: When $A > 0$, $U(w) = -\exp(-Aw)$, and the value function (4.3.7) becomes

$$H(t, P) = \sup_{\theta \in \mathcal{A}} \mathbb{E}\left[-\exp\left\{-A\int_t^T (v - P_s - c\,\theta_s)\,\theta_s\,ds\right\}\,\middle|\,\mathcal{F}_t^I\right].$$

This stochastic control problem has the associated HJB partial differential equation

$$\partial_t H + \sup_\theta\left\{\frac{1}{2}\,\sigma^2\,\lambda^2(t)\,\partial_{PP} H + \theta\,\lambda(t)\,\partial_P H - A\,(v - P - c\,\theta)\,\theta\,H\right\} = 0,$$

$$H(T,\cdot) = -1\,.$$

Once again, we check by direct substitution that this equation has solution given by (4.3.9). The resulting feedback form of the control is again linear with respect to $v - P$, and thus the remainder of the proof is identical to the risk-neutral case. □

In the next Lemma, we provide the details of the market maker's setting of efficient prices. In the proof of Theorem 4.2.2, we showed that the pricing rule was efficient by assuming that the insider's trading strategy was linear in v and applying the projection theorem for normal random variables directly to the expectation $\mathbb{E}[v \mid \Delta y]$. Our approach in continuous-time is analogous, but we need a generalization of the projection theorem to continuous-time. In following Lemma, we assume the insider follows a linear trading strategy, and we apply optimal filtering theory to compute $\mathbb{E}[v \mid \mathcal{F}_t^M]$.

Lemma 4.3.3. *Suppose that insider's inventory process is specified by*

$$dX_t = \beta(t)\,(v - P_t)\,dt\,,$$

for a deterministic function β. Then the process P specified by the dynamics

$$dP_t = \lambda(t)\,dY_t, \qquad\qquad P_0 = v_0\,,$$

where Y is given by (4.3.1) and

$$\lambda(t) = \frac{\beta(t)\Sigma(t)}{\sigma^2}\,, \qquad \frac{d\Sigma(t)}{dt} = -\sigma^2\lambda^2(t)\,, \qquad \Sigma(0) = \Sigma_0^v\,,$$

satisfies the efficiency condition (4.3.6). Furthermore, the function Σ is equal to the posterior variance of v given the market maker's information:

$$\Sigma(t) = \mathbb{E}[(v - P_t)^2 | \mathcal{F}_t^M]\,.$$

Proof. Lemma 4.3.3 is a direct application of [48, Theorem 12.1]. □

The previous two lemmas are analogous the the steps in the single-auction case. The reader will recall that the proof of Theorem 4.2.2 was done in two steps. The first step was the computation of the insider's optimal strategy for a fixed, linear pricing rule. In discrete time, this came down to solving an algebraic equation, but Lemma 4.3.2 is the continuous-time analog. The second step was the computation of an efficient pricing rule for a fixed, linear trading strategy, and Lemma 4.3.3 is the corresponding continuous-time analog.

We reduced the computation of the linear single-auction equilibrium to coupled algebraic equations, which were subsequently reduced to a single algebraic equation. In continuous-time, we reduce the computation of an equilibrium to two coupled, nonlinear ODEs for which one ODE is prescribed an initial condition and the other is prescribed a terminal condition. This motivates the following definition.

Definition 4.3.4. Let $T, \xi_0, \xi_T \in \mathbb{R}$ be constants with $T > 0$. Let $x : \mathbb{R} \to \mathbb{R}^2$ and $F : \mathbb{R}^2 \to \mathbb{R}^2$. A *forward-backward ordinary differential equation (FBODE)* is a system of the form

$$\frac{dx(t)}{dt} = F(x(t)), \qquad \begin{cases} x_1(0) = \xi_0 \\ x_2(T) = \xi_T \end{cases}.$$

In the following lemma, we state the exact FBODE that will appear in our continuous-time equilibrium and prove the existence and uniqueness of a solution.

Lemma 4.3.5. *Choose constants $A \geq 0$ and $\xi_0 > 0$. For $x \in \mathbb{R}^2$, let us define the function $F : \mathbb{R}^2 \mapsto \mathbb{R}^2$ as*

$$F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \end{pmatrix}, \qquad F_1(x) = -\frac{\sigma^2 x_1^2}{4\left(c\sigma^2 + x_1 x_2\right)^2}, \qquad F_2(x) = \frac{2A\sigma^2 x_1^2 x_2^2 - c\sigma^4}{4\left(c\sigma^2 + x_1 x_2\right)^2}.$$

Then the FBODE

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = F(x(t)), \qquad \begin{cases} x_1(0) = \xi_0 \\ x_2(T) = 0 \end{cases}, \tag{4.3.11}$$

has a unique solution. Furthermore, if x satisfies (4.3.11), then $x_1(t) > 0$ for all $t \in [0,T]$, and $x_2(t) > 0$ for all $t \in [0,T)$. When $A = 0$,

$$x_1(t) = 2c\lambda\sigma^2 + \lambda^2\sigma^2(T-t), \qquad x_2(t) = \frac{T-t}{2\lambda(T-t)+4c}, \tag{4.3.12}$$

where

$$\lambda = \sqrt{\frac{\xi_0}{T\sigma^2} + \frac{c^2}{T^2}} - \frac{c}{T}.$$

Before presenting the proof of Lemma 4.3.5, we first present the following lemmas to aid in the proof.

Lemma 4.3.6. *If x satisfies the FBODE (4.3.11), then $x_1(t) > 0$ for all $t \in [0,T]$ and $x_2(t) > 0$ for all $t \in [0,T)$. Furthermore, there exists a function $\rho : \mathbb{R}^+ \to \mathbb{R}$ such that $x_2(t) = \rho(x_1(t))$ for all $t \in [0,T]$.*

Proof of Lemma 4.3.6. Suppose that x satisfies (4.3.11). Inspection of F_1 reveals that x_1 is strictly decreasing, and because $x_1(T) = \xi_0 > 0$, we have $x_1(t) > 0$ for all $t \in [0,T]$. Define $I_1 = \{x_1(t) | t \in [0,T]\}$, which is a compact interval because x_1 is continuous. By examining F_2 we see that for all sufficiently small $\delta > 0$ we have $F_2(y) < 0$ for all $y \in I_1 \times (-\delta, \delta)$. If there exists a $t \in [0,T)$ for which $x_2(t) < -\delta$, then $x_2(s) < -\delta$ for all $s \in [t,T]$, which violates the boundary condition $x_1(T) = 0$. Therefore, $x_2(t) > 0$ for all $t \in [0,T)$.

Because x_1 is strictly decreasing it is invertible on I_1, thus we have that $\rho = x_2 \circ x_1^{-1}$. □

Lemma 4.3.7. *Choose $\xi_0 > 0$, $k > 0$ and $\gamma > 1$, and define $f_1 : \mathbb{R}^+ \to \mathbb{R}$ by*

$$f_1(x_1) = -\frac{16A^2\sigma^2 x_1^2\left(x_1^\gamma + k\right)^2}{\left((\gamma+1)^2 x_1^\gamma + (\gamma-1)^2 k\right)^2}. \tag{4.3.13}$$

Then,

$$\frac{dx_1(t)}{dt} = f_1(x_1(t)), \qquad x_1(0) = \xi_0, \tag{4.3.14}$$

has a unique solution on $[0,T]$ satisfying $x_1(t) > 0$ for all $t \in [0,T]$.

Proof of Lemma 4.3.7. Inspection of f_1 shows that it is Lipschitz continuous and satisfies $f_1(x_1) \leq 0$ with equality if and only if $x_1 = 0$. The desired results (existence, uniqueness, and positive) are then immediate. □

Lemma 4.3.8. *Let $E = (1, \infty) \times (0, \infty)$, and suppose that x_1 satisfies (4.3.14). Then $x_1(t; \gamma, k)$ is continuously differentiable in γ and k on $(t, \gamma, k) \in (0, T] \times E$. Furthermore, $x_1(t; \gamma, k)$ is strictly decreasing in k for $(t, \gamma, k) \in (0, T] \times E$.*

Proof of Lemma 4.3.8. We note that f_1 is continuous in variables (x_1, γ, k) on $(0, \infty) \times E$ and has continuous partial derivatives with respect to x_1, γ and k for $(x_1, \gamma, k) \in (0, \infty) \times E$. By Lemma 4.3.7, for any choice of $(\xi_0, \gamma, k) \in (0, \infty) \times E$ the IVP (4.3.14) has a unique solution which exists up to (and past) time $t = T$. Therefore, $x_1(T; c, k)$ is continuously differentiable with respect to γ and k (see [39]).

Additionally, we have that $z = \partial x_1(t; \gamma, k)/\partial k$ satisfies

$$\begin{aligned} \frac{dz(t;\gamma,k)}{dt} &= \frac{\partial f_1(x_1(t;\gamma,k);\gamma,k)}{\partial x_1} z(t;\gamma,k) + \frac{\partial f_1(x_1(t;\gamma,k);\gamma,k)}{\partial k}, \\ z(0;\gamma,k) &= 0, \end{aligned} \tag{4.3.15}$$

where

$$\begin{aligned} \frac{\partial f_1(x_1;\gamma,k)}{\partial x_1} &= -\frac{32A^2\sigma^2 x_1 \left(k + x_1^\gamma\right)\left((\gamma+1)x_1^\gamma - (\gamma-1)k\right)^2}{\left((\gamma+1)^2 x_1^\gamma + (\gamma-1)^2 k\right)^3}, \\ \frac{\partial f_1(x_1;\gamma,k)}{\partial k} &= -\frac{128A^2\gamma\sigma^2 x_1^{\gamma+2}\left(x_1^\gamma + k\right)}{\left((\gamma+1)^2 x_1^\gamma + (\gamma-1)^2 k\right)^3} \end{aligned}$$

By Lemma 4.3.7, we know that $x_1(t; \gamma, k) > 0$ for $t \in [0, T]$. Therefore,

$$\begin{aligned} \frac{\partial f_1(x_1(t;\gamma,k);\gamma,k)}{\partial x_1} &\leq 0, \qquad && (t,\gamma,k) \in [0,T] \times E, \\ \frac{\partial f_1(x_1(t;\gamma,k);\gamma,k)}{\partial k} &< 0, \qquad && (t,\gamma,k) \in [0,T] \times E. \end{aligned}$$

Thus, by examining (4.3.15), we see that $z(t; \gamma, k) < 0$ for all $(t, \gamma, k) \in (0, T] \times E$. □

With Lemmas 4.3.6, 4.3.7 and 4.3.8 in hand, we now present the proof of Lemma 4.3.5.

Proof of Lemma 4.3.5. If x solves (4.3.11), then by Lemma 4.3.6 there exists a function ρ such that $x_2(t) = \rho(x_1(t))$ for all $t \in [0, T]$. By the chain rule,

$$\frac{d\rho(x_1)}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{F_2(x_1, \rho(x_1))}{F_1(x_1, \rho(x_1))} = -2\,A\,\rho^2(x_1) + \frac{c\,\sigma^2}{x_1^2}\,. \tag{4.3.16}$$

We will break the remainder of the proof into two cases: $A = 0$ and $A > 0$.

Case $A = 0$: Let $A = 0$. Then (4.3.16) is satisfied by

$$\rho(x_1) = k - \frac{c\,\sigma^2}{x_1}\,, \tag{4.3.17}$$

for any $k \in \mathbb{R}$. Thus

$$\frac{dx_1(t)}{dt} = F_1(x_1(t), \rho(x_1(t))) = -\frac{\sigma^2}{4\,k^2}, \qquad\qquad x_1(0) = \xi_0\,.$$

This ODE has the unique solution

$$x_1(t) = \xi_0 - \frac{\sigma^2\,t}{4\,k^2}\,. \tag{4.3.18}$$

We must now choose k such that $x_2(T) = \rho(x_1(T)) = 0$. Substituting (4.3.18) into (4.3.17) yields

$$x_2(T) = \frac{4\,\xi_0\,k^3 - 4\,c\,\sigma^2\,k^2 - \sigma^2\,T\,k}{4\,\xi_0\,k^2 - \sigma^2\,T}\,.$$

There are three values of k which satisfy $x_2(T) = 0$. We can rule out $k = 0$ immediately due to the form of x_1 in (4.3.18). The other two roots are

$$k = \frac{c\,\sigma^2 \pm \sqrt{c^2\,\sigma^4 + \xi_0\,\sigma^2\,T}}{2\,\xi_0}\,.$$

If we take either of these values of k and evaluate $x_1(T)$ we have

$$x_1(T) = \frac{2\,c^2\,\sigma^4 \pm 2\,c\,\sigma^2\sqrt{c^2\,\sigma^4 + \xi_0\,\sigma^2\,T}}{4\,k^2\,\xi_0}\,.$$

The negative sign on the square root gives $x_1(T) < 0$, so we must take the value of k corresponding to the positive sign on the square root. Substituting this value of k into the expression for $x_1(t)$ in (4.3.18) gives

$$x_1(t) = \frac{\xi_0 \left(2\,c^2\,\sigma^4 + 2\,c\,\sigma^2 \sqrt{c^2\,\sigma^4 + \xi_0\,\sigma^2\,T} + \xi_0\,\sigma^2\,(T-t) \right)}{(c\,\sigma^2 + \sqrt{c^2\,\sigma^4 + \xi_0\,\sigma^2\,T})^2}\,.$$

This can be simplified by rationalizing the denominator which yields

$$x_1(t) = 2\,c\,\sigma^2\lambda + \lambda^2\,\sigma^2\,(T-t)\,,$$

where we have defined

$$\lambda = \sqrt{\frac{c^2}{T^2} + \frac{\xi_0}{\sigma^2\,T}} - \frac{c}{T}\,,$$

as in (4.3.12). We then evaluate $x_2(t) = \rho(x_1(t))$ to find

$$x_2(t) = \frac{T-t}{4\,c + 2\,\lambda\,(T-t)}\,,$$

as desired.

<u>Case $A > 0$</u>: Let $A > 0$. Then (4.3.16) has the solution

$$\rho(x_1) = \frac{(\gamma+1)x_1^\gamma - (\gamma-1)k}{4Ax_1(x_1^\gamma + k)}, \qquad \gamma = \sqrt{1 + 8\,A\,c\,\sigma^2}, \tag{4.3.19}$$

for any $k \in \mathbb{R}$. With this expression for ρ we have

$$\frac{dx_1(t)}{dt} = F_1(x_1(t), \rho(x_1(t))) = f_1(x_1(t))\,, \qquad x_1(0) = \xi_0\,, \tag{4.3.20}$$

where f_1 is given in (4.3.13) and γ is given in (4.3.19).

We note that in order for the terminal condition $x_2(T) = \rho(x_1(T)) = 0$ to be satisfied, then k must be chosen such that

$$g(k) := x_1^\gamma(T;k) - \frac{\gamma-1}{\gamma+1}k = 0\,,$$

where we have made the dependence of $x_1(T)$ on k explicit. If $k = 0$, the ODE (4.3.20) has an explicit solution, but the terminal condition $x_2(T) = 0$ cannot be satisfied for $\xi_0 > 0$. If $k < 0$, then we must have $x_1(T;k) < 0$ to satisfy the terminal condition $x_2(T) = 0$. Thus k must be chosen to be positive if we are to solve the constrained two-point boundary value problem (4.3.11) with $x_1(T;k) > 0$. For each $k > 0$, Lemma 4.3.7 gives that the ODE (4.3.20) has a unique positive solution on $[0,T]$. Therefore, for each $k > 0$ the function $x(t) = (x_1(t), \rho(x_1(t)))$ satisfies the ODE $dx(t)/dt = F(x(t))$. What remains to be shown is that there is a unique k that solves $g(k) = 0$, implying $x_2(T) = 0$.

By Lemma 4.3.8, g is continuous and strictly decreasing in k, and we have the bounds $0 < x_1(T;k) < \xi_0$ for all $k > 0$. As such, a constant k_r can be chosen such that

$$0 < \frac{\gamma+1}{\gamma-1}\xi_0^\gamma < k_r,$$

and $g(k_r) < 0$. Now choose

$$k_l = \frac{\gamma+1}{\gamma-1}x_1^\gamma(T;k_r).$$

We note that

$$0 < k_l < \frac{\gamma+1}{\gamma-1}\xi_0^\gamma < k_r.$$

Therefore,

$$0 < g(k_l) = x_1^\gamma(T;k_l) - x_1^\gamma(T;k_r).$$

As g is continuous and monotone in k, there exists a unique $k^* \in (k_l, k_r)$ such that $g(k^*) = 0$.

All that remains to be shown is that $x_2 > 0$ on $[0,T)$. By examining ρ in (4.3.19), we see that as $x_1(t;k^*) > 0$ for all $t \in [0,T]$, $x_2(t;k^*) \leq 0$ only if

$$x_1(t;k^*) - \frac{\gamma+1}{\gamma-1}k^* \leq 0. \tag{4.3.21}$$

As $x_1(t;k^*)$ is strictly decreasing in t, if there exist a $t \in (0,T)$ for which (4.3.21) holds, then $x_2(s;k^*) < 0$ for all $s \in (t,T]$. However, we have chosen k^* such that $x_2(T;k^*) = 0$. Therefore, $x_2(t) > 0$ for $t \in [0,T)$. □

Before giving a continuous-time equilibrium it is useful to define the constant

$$\Lambda_K = \sqrt{\frac{\Sigma_0^v}{\sigma^2\, T}}\,.$$

The constant Λ_K is the continuous-time pricing rule of [45]. That is, when a risk-neutral insider faces a risk-neutral market maker on a frictionless exchange, the processes specified by the dynamics

$$dP_t = \Lambda_K\, dY_t\,, \qquad\qquad P_0 = v_0\,,$$

$$dX_t = \frac{1}{\Lambda_K\,(T-t)}\,(v - P_t)\,dt\,, \qquad\qquad X_0 = 0\,,$$

form an equilibrium. Writing the risk-neutral continuous-auction equilibrium in terms of Λ_K will help illuminate the effect that transaction costs have on the market participant's respective strategies.

We now present a continuous-time equilibrium.

Theorem 4.3.9. *Choose $A \geq 0$, let U be as given in (4.2.3), and let $x = (\Sigma, h)$ be the unique solution to the FBODE (4.3.11) with $\xi_0 = \Sigma_0^v$. Then the midprice process P and inventory process X specified by*

$$dP_t = \lambda(t)\, dY_t\,, \qquad\qquad P_0 = v_0\,,$$

$$dX_t = \beta(t)\,(v - P_t)\,dt\,, \qquad\qquad X_0 = 0\,,$$

where Y is given in (4.3.1) and where

$$\beta(t) = \frac{\sigma^2}{2\,(c\,\sigma^2 + \Sigma(t)\,h(t))}\,, \tag{4.3.23a}$$

$$\lambda(t) = \frac{\Sigma(t)}{2\,(c\,\sigma^2 + \Sigma(t)\,h(t))}\,, \tag{4.3.23b}$$

form a continuous-time equilibrium. Furthermore, $\Sigma(t) = \mathbb{E}[(v - P_t)^2 \,|\, \mathcal{F}_t^M]$ and the value function H is given by (4.3.9). When $A = 0$, $\lambda(t) \equiv \lambda$ is a constant and

$$\beta(t) = \frac{1}{\lambda\,(T-t) + 2\,c}\,,$$

$$\lambda = \sqrt{\Lambda_K^2 + \frac{c^2}{T^2}} - \frac{c}{T}\,. \tag{4.3.24a}$$

Proof. First we fix (Σ, h) to be the unique solution to the FBODE (4.3.11) with $\xi_0 = \Sigma_0^v$. Suppose the price dynamics are given by

$$dP_t = \lambda(t)\, dY_t\,, \qquad\qquad P_t = v_0\,,$$

with $\lambda(t)$ given by

$$\lambda(t) = \frac{\Sigma(t)}{2\,(c\,\sigma^2 + \Sigma(t)\,h(t))}\,,$$

as in the statement of the Theorem. Because (Σ, h) is a solution to the FBODE (4.3.11) with $\xi_0 = \Sigma_0$ we have

$$\frac{dh(t)}{dt} = \frac{2\,A\,\sigma^2\,\Sigma^2(t)\,h^2(t) - c\,\sigma^4}{4\,(c\,\sigma^2 + \Sigma(t)\,h(t))^2}\,.$$

In addition, a straightforward computation shows

$$-\frac{(1 - 2\,A\,c\,\sigma^2)\,\lambda^2(t)}{c}\,h^2(t) + \frac{\lambda(t)}{c}\,h(t) - \frac{1}{4\,c} = \frac{2\,A\,\sigma^2\,\Sigma^2(t)\,h^2(t) - c\,\sigma^4}{4\,(c\,\sigma^2 + \Sigma(t)\,h(t))^2}\,.$$

Thus we have that h satisfies the ODE (4.3.8). Therefore, by Lemma 4.3.2, the optimal trading strategy is given by

$$\theta_t^* = \beta(t)\,(v - P_t)\,dt\,,$$

with

$$\beta(t) = \frac{1 - 2\,\lambda(t)\,h(t)}{2\,c}\,,$$

and the insider's value function is given by (4.3.9).

Now, again with (Σ, h) as the unique solution to the FBODE (4.3.11) with $\xi_0 = \Sigma_0$, suppose that the insider's trading strategy is given by

$$\theta_t = \beta(t)\,(v - P_t)\,dt\,, \tag{4.3.25}$$

with

$$\beta(t) = \frac{\sigma^2}{2\,(c\,\sigma^2 + \Sigma(t)\,h(t))}\,.$$

From (4.3.11) we see that Σ satisfies

$$\frac{d\Sigma(t)}{dt} = -\frac{\sigma^2\,\Sigma^2(t)}{4\,(c\,\sigma^2 + \Sigma(t)\,h(t))^2} = -\sigma^2\,\lambda^2(t)\,. \tag{4.3.26}$$

In addition we have by their definitions in the statement of the Theorem that

$$\lambda(t) = \frac{\beta(t)\,\Sigma(t)}{\sigma^2}\,. \tag{4.3.27}$$

By Lemma 4.3.3, the relations in (4.3.25), (4.3.26), and (4.3.27) imply that the efficient pricing rule is given by

$$dY_t = \lambda(t)\,dY_t\,, \qquad\qquad P_0 = v_0\,,$$

and that the function Σ also yields the market maker's conditional variance:

$$\Sigma(t) = \mathbb{E}[(v - P_t)^2|\mathcal{F}_t^M]\,.$$

Finally, when $A = 0$, applying Lemma 4.3.5 gives the closed form expressions

$$\beta(t) = \frac{1}{\lambda\,(T-t) + 2\,c}\,,$$
$$\lambda = \sqrt{\Lambda_K^2 + \frac{c^2}{T^2}} - \frac{c}{T}\,.$$

as desired. □

Remark 4.3.10. Recall that in Lemma 4.3.2 we removed from consideration many pricing rules in which an associated ODE did not have a solution, as failure to do so would result in the ODE solutions blowing up within the interval $[0,T]$. Our method of classifying equilibrium in Theorem 4.3.9 avoids this issue by taking the function h as given (a unique solution to an FBODE which does not blow up), and then specifying the pricing rule λ which yields h as the solution to the ODE (4.3.8).

Definition 4.3.1 defines what it means for processes (P, X) to form a continuous-time equilibrium. In Theorem 4.3.9, we gave deterministic functions β and λ such that the processes defined by $dP_t = \lambda(t)\, dY_t$ and $dX_t = \beta(t)\,(v - P_t)dt$ with $(P_0, X_0) = (v_0, 0)$ form a continuous-time equilibrium. While the functions (β, λ) do not themselves form an equilibrium, the functions (β, λ) correspond directly to a linear equilibrium (P, X). In the sequel, we will refer to linear equilibria (P, X) using (β, λ), and we refer to β as an *equilibrium trading rule* and λ as an *equilibrium pricing rule.*

We will now discuss and interpret some characteristics of the continuous-time equilibrium of Theorem 4.3.9.

Proposition 4.3.11. *Let β and λ be the trading and pricing rules, respectively, corresponding to the continuous-time equilibrium of Theorem 4.3.9. Then,*

1. *β is a strictly increasing function of t,*

2. *if $A > 0$ then λ is a strictly decreasing function of t.*

Proof. First, let Σ and h be the functions in (4.3.23a) and (4.3.23b) corresponding to the equilibrium rules β and λ. It will be convenient to define the function $m = \Sigma h$. We see that if $x = (\Sigma, h)$ satisfies (4.3.11) with $\xi_0 = \Sigma_0^v$, then $y = (\Sigma, m)$ satisfies the FBODE

$$\frac{dy(t)}{dt} = G(y(t))\,, \qquad \begin{cases} y_1(0) = \Sigma_0^v \\ y_2(T) = 0 \end{cases}, \tag{4.3.28}$$

where

$$G(y) = \begin{pmatrix} G_1(y) \\ G_2(y) \end{pmatrix}, \quad G_1(y) = -\frac{\sigma^2\, y_1^2}{4\,(c\,\sigma^2 + y_2)^2}, \quad G_2(y) = -\frac{\sigma^2\, y_1 (c\,\sigma^2 + y_2 - 2\,A\,y_2^2)}{4\,(c\,\sigma^2 + y_2)^2}.$$

In the new variables, we write

$$\beta(t) = \frac{\sigma^2}{2\,(c\,\sigma^2 + m(t))}, \qquad \lambda(t) = \frac{\Sigma(t)}{2\,(c\,\sigma^2 + m(t))}. \tag{4.3.29}$$

By (4.3.29), we see that β is increasing if and only if m is decreasing. As we know that $m(t) > 0$ for all $t \in [0, T)$, examining the numerator of G_2 gives that if

$$m(t) < K\,, \qquad K = \frac{1 + \sqrt{1 + 8\,A\,c\,\sigma^2}}{4\,A},$$

for all $t \in [0,T]$ then m is decreasing on $[0,T]$. We observe by inspecting (4.3.28) that if there exists a $t^* \in [0,T)$ for which $m(t^*) \geq K$, then $m(s)$ will be non-decreasing for all $s \in [t^*,T]$. In that case, m cannot satisfy the terminal condition $m(T)=0$. Therefore, β is increasing.

By examining (4.3.28) and (4.3.29), we compute

$$\frac{d\lambda(t)}{dt} = -\frac{A\,\sigma^2\,\Sigma^2(t)\,m^2(t)}{4\,(c\,\sigma^2 + m(t))^4}\,,$$

and we have that λ is decreasing. □

In equilibrium, β is positive, increasing in time and $\beta(T) = 1/2\,c$, so

$$0 < \beta(t) < \frac{1}{2\,c}, \qquad\qquad t \in [0,T)\,. \tag{4.3.30}$$

And (4.3.24a) and Proposition 4.3.11 tell us that the equilibrium pricing rule λ never increases in time; in the risk-neutral case, the cost to move the price of the asset remains constant throughout the trading window, and in the risk-averse setting, it becomes strictly more expensive to move the price of the asset. So, the insider never finds it advantageous to buy (sell) the overvalued (undervalued) asset with the intention of pushing the midprice P away from v to profit from the misprice at a later time. Hence, $0 < \beta(t)$ for $t \in [0,T]$.

We also see from (4.3.30) that the equilibrium trading rule β is bounded above by $1/2\,c$. To gain some intuition, suppose that the market maker takes either the risk-neutral or risk-averse equilibrium pricing rule λ. Suppose the insider chooses a strategy of the form $\theta_t = b(t)\,(v - P_t)$, where b is a continuous, deterministic function defined on $[0,T]$. Then, recalling that the insider's terminal wealth W_T is given by (4.3.4), we have

$$W_T = \int_0^T b(s)\,(1 - c\,b(s))\,(v - P_s)^2\,ds\,.$$

If at a time t the insider submits orders at a rate $b(t) > 1/c$ then they accumulate wealth at a negative instantaneous rate. Furthermore, as $b(t) > 0$, the insider affects price impact and moves the midprice P towards v. In effect, the insider has paid to relinquish part of their informational advantage, something a trader acting optimally would never do.

At time t, the insider accumulates wealth at an instantaneous rate of $b(t)\,(1 - c\,b(t))\,(v - P_t)^2$. The function $b\,(1 - c\,b)$ is symmetric in b about $b = 1/2\,c$. If at time t the insider wishes to accumulate wealth at a specified instantaneous rate, then they have two choices of $b(t)$ to achieve their goal. An insider trading in equilibrium chooses $b(t) < 1/2\,c$ in order to affect lower price impact than the corresponding higher trading speed, allowing the insider to retain more of their informational advantage at the same wealth accumulation rate.

In the absence of transaction costs, the authors of [45] and [10] show that in equilibrium the risk-neutral and exponentially risk-averse insider, respectively, take trading rules β that blow up as t approaches T. This forces the midprice to v and gives that $\Sigma(T) = 0$. For a non-zero transaction cost c, we saw in the discussion above that an insider of any risk tolerance never wishes to choose $\beta > 1/2c$, as the cost of trading outweighs the benefits. In this setting, the insider does not reveals the true value of the asset to the market maker by the end of the trading period i.e. $\Sigma(T) > 0$.

4.4 Parameter Dependence

We now wish to investigate the dependence of the continuous-time equilibrium rules β and λ on the model parameters. The amounts to analyzing the dependence of solutions $x = (\Sigma, h)$ to the FBODE (4.3.11) on its parameters. Let $\Theta = (A, c, \sigma, \Sigma_0^v)$ be the vector of model parameters. One approach is to treat solutions $x(t) \equiv x(t; \Theta)$ as functions of the parameters and differentiate both sides of (4.3.11) with respect to Θ_i to derive an ODE for $\partial x/\partial\Theta_i$ in time. However, while linear in $\partial x/\partial\Theta_i$, the new ODE system is a forward-backward system with time dependent coefficients and forcing terms. Due to the difficulty of this approach, we numerically investigate the dependence of the continuous-time equilibrium rules β and λ on the model parameters Θ.

We begin by showing the effect varying the risk-aversion parameter A has on β and λ. Let us fix the parameters $(T, c, \sigma, \Sigma_0^v) = (1, 0.2, 1, 0.5)$. In Figures 4.3 and 4.4, we plot β and λ, respectively, as functions of t while varying A. We see that as the insider becomes more risk averse, they take advantage of the price mismatch $v - P$ more aggressively. The more

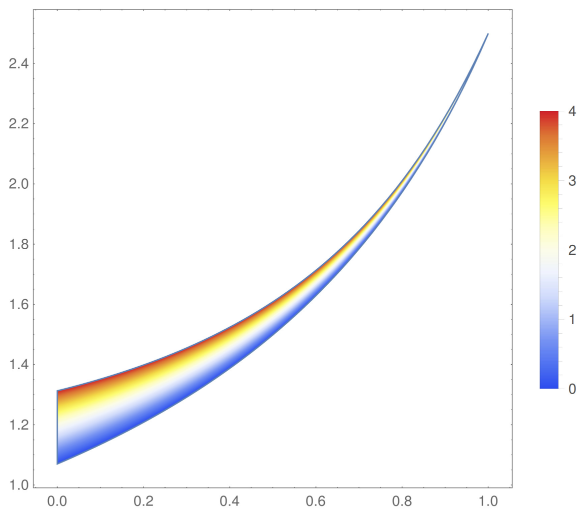


Figure 4.3: Here we plot β as a function of t for the risk-averse insider and vary the risk-aversion parameter $A > 0$. Cooler colors correspond to lower values of A and warmer colors correspond to higher values of A.

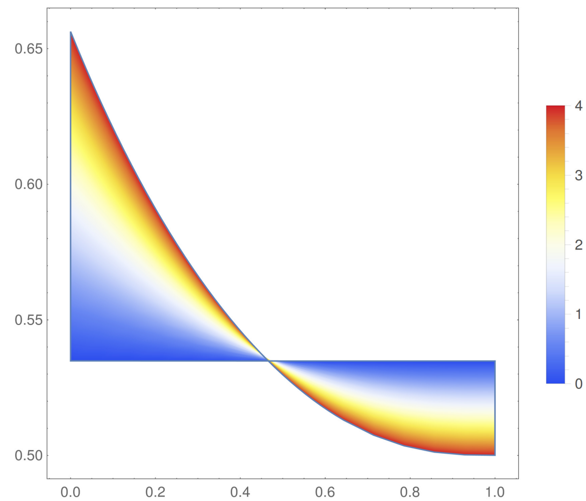


Figure 4.4: Here we plot λ as a function of t for the market maker facing a risk-averse insider and vary the risk-aversion parameter $A > 0$. Cooler colors correspond to lower values of A and warmer colors correspond to higher values of A.

risk-averse insider prefers to lock in profits near the beginning of the trading horizon rather than wait in an attempt to maximize profit. Correspondingly, as the insider trades more aggressively, the market maker has more information about the true price v. Thus, as the insider becomes more risk averse, the market maker increases λ early in the trading period and lowers λ towards the end of the period. It appears that β is monotonically increasing in A, while λ is increasing in A for the first part of the trading period and decreasing the second part.

Now let us fix the parameters $(T, A, \sigma, \Sigma_0^v) = (1, 1, 1, 0.5)$ and investigate the behavior of equilibrium rules β and λ as we vary the transaction cost parameter c. In Figures 4.5 and 4.6, we plot β and λ, respectively, as functions of t while varying c. We see that as the transaction cost increases, the insider monotonically slows their trading. Furthermore, Figure

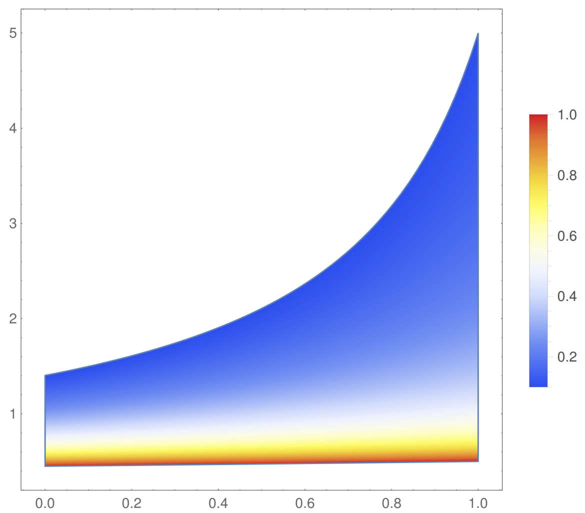


Figure 4.5: Here we plot β as a function of t for the risk-averse insider and vary the temporary price impact parameter $c > 0$. Cooler colors correspond to lower values of c and warmer colors correspond to higher values of c.

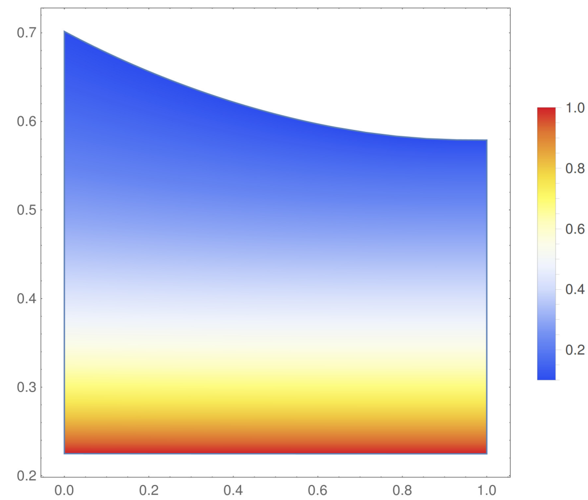


Figure 4.6: Here we plot λ as a function of t for the market maker facing a risk-averse insider and vary the temporary price impact parameter $c > 0$. Cooler colors correspond to lower values of c and warmer colors correspond to higher values of c.

4.5 suggests that in addition to decreasing β, the insider also decreases $\partial\beta/\partial t$ monotonically in c. As the transaction cost c increases and the insider slows their trading, less information flows onto the market. It is thus no surprise that Figure 4.6 shows λ decreasing monotonically in c.

Now let us fix the parameters $(T, A, c, \Sigma_0^v) = (1, 1, 0.2, 0.5)$ and investigate the behavior of equilibrium rules β and λ as we vary the noise parameter σ. In Figures 4.7 and 4.8, we plot β and λ, respectively, as functions of t while varying σ. Figure 4.7 shows that as σ increases, so does the insider's trading speed. As the volume of noise trades increases, each share that the insider trades contributes less to the signal Y visible to the market maker, which corresponds to less information leaking onto the market. Thus, an insider with a fixed transaction costs and risk tolerance can afford to trade more rapidly. As σ increases, the

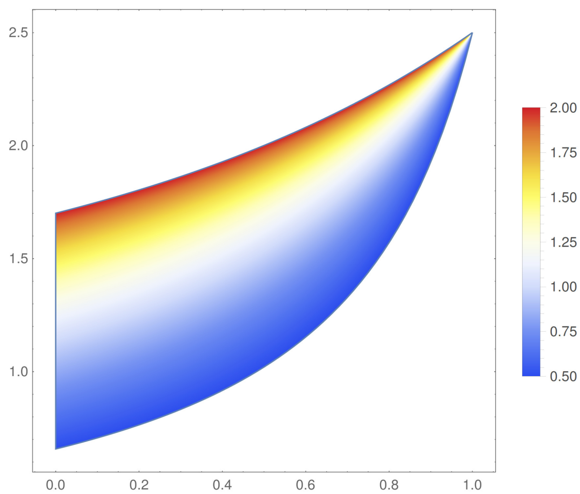


Figure 4.7: Here we plot β as a function of t for the risk-averse insider and vary the noise parameter $\sigma > 0$. Cooler colors correspond to lower values of σ and warmer colors correspond to higher values of σ.

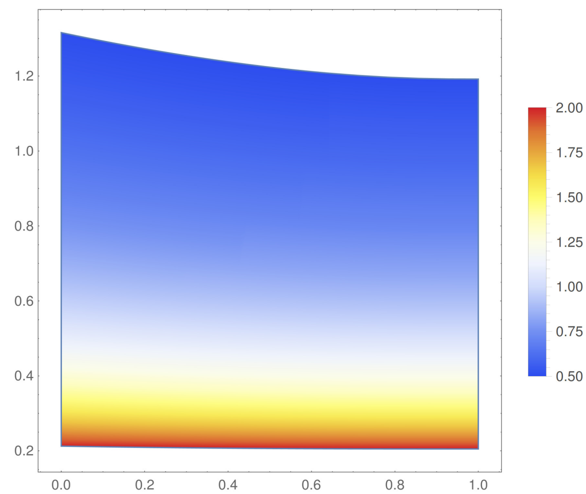


Figure 4.8: Here we plot λ as a function of t for the market maker facing a risk-averse insider and vary the noise parameter $\sigma > 0$. Cooler colors correspond to lower values of σ and warmer colors correspond to higher values of σ.

market maker gains less information from the signal Y about v. As such, the market maker decreases price pressure as noise trading increases, and this is demonstrated in Figure 4.8.

Now let us fix the parameters $(T, A, c, \sigma) = (1, 1, 0.2, 1)$ and investigate the behavior of equilibrium rules β and λ as we vary the initial variance Σ_0^v. In Figures 4.9 and 4.10, we plot β and λ, respectively, as functions of t while varying Σ_0^v. We see in Figure 4.10 that the market maker increases price pressure as the initial midprice variance Σ_0^v increases. When the market maker is relatively sure about v, they choose λ to be small as they have likely set the midprice near the true value of the asset. This allows the insider to trade more aggressively on their information as the negative effects of price impact are lessened. We see in Figure 4.9 that β increases with Σ_0^v.

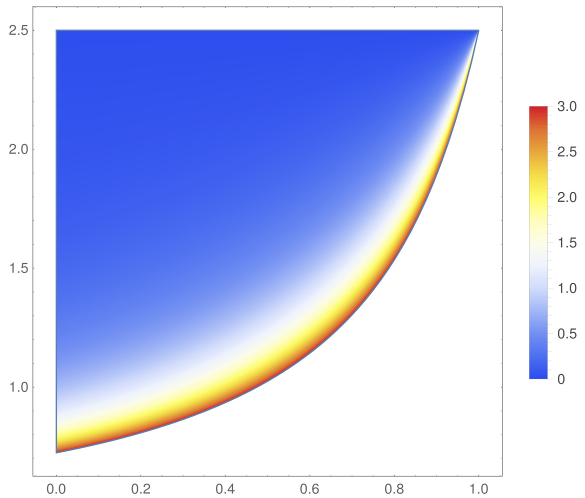


Figure 4.9: Here we plot β as a function of t for the risk-averse insider and vary the initial price variance parameter $\Sigma_0^v > 0$. Cooler colors correspond to lower values of Σ_0^v and warmer colors correspond to higher values of Σ_0^v.

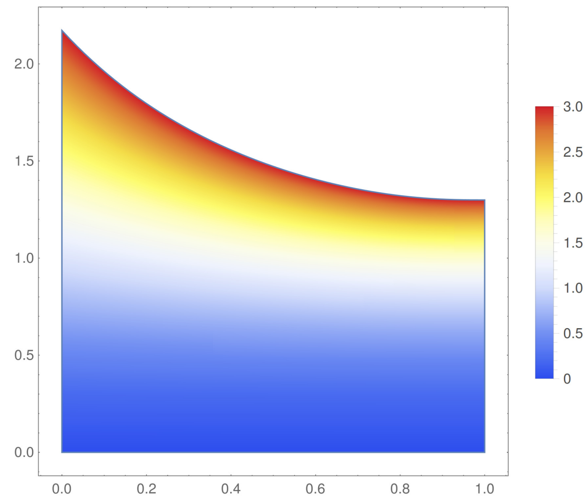


Figure 4.10: Here we plot λ as a function of t for the market maker facing a risk-averse insider and vary the initial price variance parameter $\Sigma_0^v > 0$. Cooler colors correspond to lower values of Σ_0^v and warmer colors correspond to higher values of Σ_0^v.

4.5 Conclusion

In this chapter, we extended the results of [45], [41] and [10] by adding friction to the market in the form of a transaction cost which is linear in the insider's order size. We considered a exponentially risk-averse and risk-neutral insiders. We began by modeling a single-auction exchange and providing the corresponding unique market equilibrium in the form of the solution to an algebraic equation. We then demonstrated the effect of the transaction cost on the equilibrium by asymptotically expanding the algebraic equation about the frictionless equilibrium for small transaction cost. This procedure uncovered the equilibrium relationship between the insider's risk preferences and the transaction cost. Namely, we see that when the insider is risk-neutral the first order correction term in the market maker's strategy disappears.

We then formulated an analogous market model in continuous time. We developed an optimal trading strategy for an exponentially risk-averse or risk-neutral insider by explicitly solving the HJB equation associated with optimization problem (4.3.5). We then gave the filtering equations that provide an efficient price process for a fixed insider strategy. The main result is Theorem 4.3.9, which gives a linear, continuous-time market equilibrium in terms of the solution to the FBODE (4.3.11), for which we showed there exists a unique solution. We explored the effects of varying other model parameters by numerically solving the ODEs associated with the equilibrium strategies.

There are several qualitative similarities between the equilibria of the frictionless case and ours. First, linear equilibria exist in both cases and are given in terms of the solutions to differential equations. Additionally, the equilibrium insider trading rule β is increasing in time, and the equilibrium market maker's pricing rule λ is constant in the risk-neutral setting and decreasing in time in the risk-averse setting. The insider in equilibrium never shorts an undervalued asset or goes long an overvalued asset in order to get favorable price impact i.e. $\beta \geq 0$. Furthermore, β and λ exhibit similar dependence on the overlapping model parameters. For example, in the continuous-time setting, as the risk-aversion parameter A increases, the $\lambda(0)$ increases and $\lambda(T)$ decreases as is the case in [10].

But not all of the properties of equilibria carry over from the frictionless case. The market maker's conditional variance Σ has the property $\Sigma(T) > 0$, as distinct from the frictionless case in which $\Sigma(T) = 0$. This implies that the true value of the asset v is not reveled to the market maker by time t. This is because the addition of a transaction cost puts an upper bound on the insider's equilibrium trading rule β. The insider never finds it worth it to take the trading rule $\beta > 1/2c$ because the cost of trading becomes prohibitively high. Additionally, the insider's trading rule remains bounded as $A \to \infty$, converging to the constant $\beta = 1/2c$.

BIBLIOGRAPHY

[1] Aurélien Alfonsi, Antje Fruth, and Alexander Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2):143–157, 2010.

[2] Robert Almgren. Optimal trading with stochastic liquidity and volatility. *SIAM Journal on Financial Mathematics*, 3(1):163–181, 2012.

[3] Robert Almgren and Neil Chriss. Value under liquidation. *Risk*, 12(12):61–63, 1999.

[4] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[5] Robert F Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied mathematical finance*, 10(1):1–18, 2003.

[6] L. Bachelier. Théorie de la spéculation. In *Annales Scientifiques de L'Ecole Normale Supérieure*, volume 17, pages 21–88, 1900.

[7] Kerry Back. Insider trading in continuous time. *The Review of Financial Studies*, 5(3):387–409, 1992.

[8] Weston Barger and Matthew Lorig. Approximate pricing of european and barrier claims in a local-stochastic volatility setting. *International Journal of Financial Engineering*, 4(02n03):1750018, 2017.

[9] Weston Barger and Matthew Lorig. Optimal liquidation under stochastic price impact. *International Journal of Theoretical and Applied Finance*, 22(2), 2019.

[10] Shmuel Baruch. Insider trading and risk aversion. *Journal of Financial Markets*, 5(4):451–464, 2002.

[11] David Bates. The crash premium: Option pricing under asymmetric processes, with applications to options on deutschemark futures. *Working Paper, University of Pennsylvania*, 1988.

[12] Dimitris Bertsimas, Paul Hummel, and Andrew W Lo. Optimal control of execution costs for portfolios. *Computing in Science & Engineering*, 1(6):40–53, 1999.

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

[14] T. Björk. *Arbitrage theory in continuous time*. Oxford Univ Pr, 2009.

[15] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, 81(3):637–654, 1973.

[16] J. Bowie and P. Carr. Static simplicity. *Risk*, 8 1994.

[17] P. Carr and M. Lorig. Robust replication of barrier-style claims on price and volatility. *ArXiv e-prints*, August 2015.

[18] Peter Carr, Katrina Ellis, and Vishal Gupta. Static hedging of exotic options. *The Journal of Finance*, 53(3):1165–1190, 1998.

[19] Peter Carr and Roger Lee. Put-call symmetry: Extensions and applications. *Mathematical Finance*, 19(4):523–560, 2009.

[20] Peter Carr and Sergey Nadtochiy. Static hedging under time-homogeneous diffusions. *SIAM Journal on Financial Mathematics*, 2(1):794–838, 2011.

[21] Alvaro Cartea, Ryan Francis Donnelly, and Sebastian Jaimungal. Algorithmic trading with model uncertainty. 2014.

[22] Álvaro Cartea and Sebastian Jaimungal. Risk metrics and fine tuning of high-frequency trading strategies. *Mathematical Finance*, 25(3):576–611, 2015.

[23] Álvaro Cartea and Sebastian Jaimungal. Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, 10(3):339–364, 2016.

[24] Álvaro Cartea, Sebastian Jaimungal, and José Penalva. *Algorithmic and high-frequency trading*. Cambridge University Press, 2015.

[25] Jin Hyuk Choi, Kasper Larsen, and Duane J Seppi. Information and trading targets in a dynamic market equilibrium. *Journal of Financial Economics*, 2018.

[26] Pierre Collin-Dufresne and Vyacheslav Fos. Shareholder activism, informed trading, and stock prices. *Swiss Finance Institute Research Paper*, (13-70), 2015.

[27] J. Cox. Notes on option pricing I: Constant elasticity of diffusions. *Unpublished draft, Stanford University*, 1975. A revised version of the paper was published by the Journal of Portfolio Management in 1996.

[28] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407, 1985.

[29] D. Davydov and V. Linetsky. Pricing and hedging path-dependent options under the cev process. *Management science*, 47(7):949–965, 2001.

[30] A. Einstein. On the theory of the Brownian movement. *Ann. Phys*, 19:371, 1906.

[31] Oliver Faulhaber. *Analytic methods for pricing double barrier options in the presence of stochastic volatility*. PhD thesis, Technische Universität Kaiserslautern, 2002.

[32] Jean-Pierre Fouque, Sebastian Jaimungal, and Matthew Lorig. Spectral decomposition of option prices in fast mean-reverting stochastic volatility models. *SIAM Journal on Financial Mathematics*, 2(1), 2011.

[33] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Solna. *Multiscale stochastic volatility for equity, interest rate, and credit derivatives.* Cambridge University Press, Cambridge, 2011.

[34] Avner Friedman. *Partial differential equations of parabolic type.* Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[35] Jim Gatheral, Alexander Schied, and Alla Slynko. Transient linear price impact and fredholm integral equations. *Mathematical Finance*, 22(3):445–474, 2012.

[36] Emmanuel Gobet and Stefano Pagliarani. Analytical approximations of non-linear sdes of mckean–vlasov type. *Journal of Mathematical Analysis and Applications*, 2018.

[37] P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woodward. Managing smile risk. *Wilmott Magazine*, 1000:84–108, 2002.

[38] J.M. Harrison and S.R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, 11(3):215–260, 1981.

[39] P. Hartman. *Ordinary Differential Equations.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2002.

[40] Steven Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 6(2):327–343, 1993.

[41] Craig W. Holden and Avanidhar Subrahmanyam. Risk aversion, imperfect competition, and long-lived information. *Economics Letters*, 1994.

[42] Gur Huberman and Werner Stanzl. Optimal liquidity trading. *Review of Finance*, 9(2):165–200, 2005.

[43] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[44] Idris Kharroubi and Huyên Pham. Optimal portfolio liquidation with execution cost and risk. *SIAM Journal on Financial Mathematics*, 1(1):897–931, 2010.

[45] Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

[46] Steffen L Lauritzen. Time series analysis in 1880: A discussion of contributions made by tn thiele. *International Statistical Review/Revue Internationale de Statistique*, pages 319–331, 1981.

[47] Alexander Lipton, Andrey Gal, and Andris Lasis. Pricing of vanilla and first-generation exotic options in the local stochastic volatility framework: survey and new results. *Quantitative Finance*, 14(11):1899–1922, 2014.

[48] Robert Liptser and Albert Shiryaev. *Statistics of Random Processes : II. Applications.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[49] Matthew Lorig. Pricing derivatives on multiscale diffusions: An eigenfunction expansion approach. *Mathematical Finance*, 24(2):331–363, 2014.

[50] Matthew Lorig. Indifference prices and implied volatilities. *Mathematical Finance*, pages n/a–n/a, 2016.

[51] Matthew Lorig, Stefano Pagliarani, and Andrea Pascucci. Analytical expansions for parabolic equations. *SIAM Journal on Applied Mathematics*, 75:468–491, 2015.

[52] Matthew Lorig, Stefano Pagliarani, and Andrea Pascucci. Explicit implied volatilities for multifactor local-stochastic volatility models. *To appear: Mathematical Finance*, 2015.

[53] Matthew Lorig and Ronnie Sircar. Portfolio optimization under local-stochastic volatility: Coefficient taylor series approximations and implied sharpe ratio. *SIAM Journal on Financial Mathematics*, 7(1):418–447, 2016.

[54] R.C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[55] Anna Obizhaeva and Jiang Wang. Optimal trading strategy and supply/demand dynamics. Working Paper 11444, National Bureau of Economic Research, June 2005.

[56] Anna A Obizhaeva and Jiang Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32, 2013.

[57] B.K. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*. Springer Verlag, 2005.

[58] S Pagliarani, Andrea Pascucci, and M Pignotti. Intrinsic expansions for averaged diffusion processes. *Stochastic Processes and their Applications*, 127(8):2560–2585, 2017.

[59] Stefano Pagliarani and Andrea Pascucci. Analytical approximation of the transition density in a local volatility model. *Cent. Eur. J. Math.*, 10(1):250–270, 2012.

[60] Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

[61] Riccardo Rebonato. *Volatility and correlation: In the pricing of equity, FX and interest-rate options*. John Wiley & Sons, 1999.

[62] Eric Renault and Nizar Touzi. Option hedging and implied volatilities in a stochastic volatility model 1. *Mathematical Finance*, 6(3):279–302, 1996.

[63] P. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Management Review*, 6(2):41–49, 1965.

[64] Alexander Schied. Robust strategies for optimal order execution in the almgren–chriss framework. *Applied Mathematical Finance*, 20(3):264–286, 2013.

[65] S.E. Shreve. *Stochastic calculus for finance II: Continuous-time models.* Springer Verlag, 2004.

[66] Avanidhar Subrahmanyam. Transaction taxes and financial market equilibrium. *The Journal of Business*, 71(1):81–118, 1998.

Appendix A

DYNAMIC PROGRAMMING IN CONTINUOUS TIME

In this section, we review some of the fundamental concepts of dynamic programming in continuous time. Our presentation is standard material and is adapted primarily from the textbooks [60] and [57]. We give this background information here as a supplement for a reader who is unfamiliar with these concepts a brief overview. We opt to omit several technical details and proofs, giving largely heuristic arguments referring to the aforementioned texts for details. The notation in this section was chosen to be similar to the reference texts [60] and [57] while also trying to be as similar to the problem of this chapter.

In this section, we let $|\cdot|$ denote the Frobenius matrix norm. Explicitly, for $M \in \mathbb{R}^{k_1 \times k_2}$ we have

$$|M| = \sqrt{\text{Trace}(MM^T)}.$$

The dimension corresponding to $|\cdot|$ is inferred from the object on which the norm is computed. We allow for the possibility of $k_i = 1$ for either or both of $i = 1, 2$. When $k_2 = 1$ then M is a vector, and when $k_1 = k_2 = 1$ then $M \in \mathbb{R}$ and $|\cdot|$ is the absolute value.

Let us fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t}, \mathbb{P})$. Assume that $A \subset \mathbb{R}^m$ is a bounded set, and let

$$\mu : \mathbb{R}^n \times A \mapsto \mathbb{R}^n, \qquad\qquad \sigma : \mathbb{R}^n \times A \mapsto \mathbb{R}^{n \times d},$$

be measurable functions such that there exists a $K \geq 0$ where

$$|\mu(x, a) - \mu(y, a)| + |\sigma(x, a) - \sigma(y, a)| \leq K|x - y|, \qquad \forall (x, y, a) \in \mathbb{R}^n \times \mathbb{R}^n \times A.$$

Remark A.0.1. Although we have assumed that A is bounded, this assumption is not necessary. We make this assumption to simplify the derivation of the dynamic programming equation. See [60] for a the derivation with the boundedness assumption on A relaxed.

Let $\nu = (\nu_t)_{0\leq t}$ be a progressively measurable process with respect to $\mathbb{F}$, and let $\mathcal{T}_{t,T}$ be the set of stopping times in $[t, T]$, where $0 \leq t \leq T$. We define the set

$$\mathcal{A}_T = \left\{ \nu \,\middle|\, \mathbb{E}\left[\int_0^T \left(|\mu(0,\nu_t)|^2 + |\sigma(0,\nu_t)|^2 \right) \mathrm{d}t \right] < \infty \right\}.$$

The above assumptions guarantee that for all $\nu \in \mathcal{A}_T$, the SDE

$$\mathrm{d}X_t^\nu = \mu(X_t^\nu, \nu_t)\,\mathrm{d}t + \sigma(X_t^\nu, \nu_t)\,\mathrm{d}W_t,$$

where $W = (W_t)_{0\leq t}$ is a standard d-dimensional Brownian motion adapted to $\mathbb{F}$, will have a unique, strong solution up to time $t = T$ for any initial condition $(t, x) \in [0, T] \times \mathbb{R}^n$.

We refer to ν as a *control process* or simply a *control*. We tend to think of ν as being an input to the system that an agent chooses, and, often, an agent will choose ν in an attempt to steer X towards achieving some objective. The superscript ν on X^ν is to make explicit the dependence of the state variable X^ν on the choice of the control process ν.

Let us now define an objective function for which an agent will choose ν to optimize. Suppose that

$$f : [0, T] \times \mathbb{R}^n \times A \mapsto \mathbb{R}, \qquad\qquad g : \mathbb{R}^n \mapsto \mathbb{R},$$

are functions such that there exists a $(C_1, C_2) \in \mathbb{R} \times (0, \infty)$ such that

$$\begin{aligned} C_1 &\leq g(x), && \forall x \in \mathbb{R}^n, \\ |g(x)| &\leq C_2(1 + |x|^2), && \forall x \in \mathbb{R}^n. \end{aligned}$$

For $(t, x) \in [0, T] \times \mathbb{R}^n$, define

$$\mathcal{A}_T(t,x) = \left\{ \nu \,\middle|\, \mathbb{E}\left[\int_t^T |f(s, X_s^\nu, \nu_s)|\, \mathrm{d}s \,\middle|\, X_t = x \right] < \infty \right\}$$

Finally, let us define the functional $\mathbb{H} : [0, T] \times \mathbb{R}^n \times \mathcal{A}_T(t, x) \mapsto \mathbb{R}$ as

$$\mathbb{H}(t, x, \nu) = \mathbb{E}\left[\int_t^T f(s, X_s^\nu, \nu_s)\, \mathrm{d}s + g(X_T) \,\middle|\, X_t = x \right].$$

The functional $\mathbb{H}$ is sometimes referred to as the *objective functional* or the *gain functional*. For a given initial state $X_t = x$, the objective functional gives the expected reward an agent receives at time $t = T$ by taking a control ν. We would like to choose ν to maximize the objective functional, which leads us to define the *value function*

$$H(t, x) = \sup_{\nu \in \mathcal{A}_T(t,x)} \mathbb{H}(t, x, \nu). \tag{A.0.1}$$

The function H can be thought of as the expected reward or value that an agent receives at time $t = T$ by starting at a state $X_t = x$ and acting optimally over the entire horizon $[t, T]$.

Definition A.0.2. If $\widehat{\nu} \in \mathcal{A}_T(t, x)$ is such that $H(t, x) = \mathbb{H}(t, x, \widehat{\nu})$, then $\widehat{\nu}$ is said to be an *optimal control.*

Definition A.0.3. Let $(t, x) \in [0, T] \times \mathbb{R}^n$ and suppose that $X_t = x$. If for a control process ν there exists a measurable function $a : [0, T] \times \mathbb{R}^n \mapsto A$ such that $\nu_t = a(t, X_t^\nu)$ then ν is said to be a *Markovian control.*

We now state the well-known dynamic programming principle (see [60] for proof).

Theorem A.0.4 (Dynamic Programming Principle)**.** *For any stopping time $\tau \in \mathcal{T}_{t,T}$,*

$$H(t, x) = \sup_{\nu \in \mathcal{A}_\tau(t,x)} \mathbb{E}\left[\int_t^\tau f(s, X_s^\nu, \nu_s)\, \mathrm{d}s + H(\tau, X_\tau^\nu) \,\middle|\, X_t = x \right]. \tag{A.0.2}$$

Theorem A.0.4 has a straightforward interpretation. Let $\tau \in \mathcal{T}_{t,T}$ and suppose that the agent knows how act optimally on $[\tau, T]$ with respect to (A.0.1). Then Theorem A.0.4 says that if the agent acts optimally on the interval $[t, \tau]$ and then acts optimally on the interval $[\tau, T]$ then they have acted optimally on $[t, T]$. Essentially, Theorem A.0.4 allows us to break the optimization problem (A.0.1) in to chunks which correspond to sub-optimization problems

We now use the dynamic programming principle to derive infinitesimal conditions that necessarily must be satisfied by smooth value functions. We do this with the hope that finding functions that satisfy the aforementioned infinitesimal conditions will allow us to significantly narrow our search for the value function (and corresponding optimal control). To this end, choose $h > 0$, and let $\tau = t + h$. For some arbitrary $a \in A$, we have by (A.0.2) that

$$H(t,x) \geq \mathbb{E}\left[\int_t^{t+h} f\left(s, X_s^a, a\right)\, \mathrm{d}s + H\left(t+h, X_{t+h}^a\right) \,\middle|\, X_t = x\right]. \tag{A.0.3}$$

The left hand side of (A.0.3) is the expected reward an agent receives by starting at a $X_t = x$ and acting optimally until $t = T$, and the right hand side of (A.0.3) is the expected reward an agent receives by starting at a state $X_t = x$, taking the constant control a for h amount of time, and then acting optimally on $[t + h, T]$.

Assuming that H is smooth enough, we can apply Itô's formula to write

$$H(t+h, X_{t+h}^a) = H(t,x) + \int_t^{t+h} \left(\partial_t + \mathcal{A}^a\right) H(s, X_s^a)\, \mathrm{d}s + (\text{martingale}), \tag{A.0.4}$$

where

$$\mathcal{A}^a = \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^T)_{ij}(x,a)\partial^2_{x_i x_j} + \sum_{i=1}^n \mu_i(x,a)\partial_{x_i}$$

is the infinitesimal generator of the diffusion X^a. By inserting (A.0.4) into (A.0.3), we get the inequality

$$\mathbb{E}\left[\int_t^{t+h} \left[\left(\partial_t + \mathcal{A}^a\right) H(s, X_s^a) + f\left(s, X_s^a, a\right)\right]\, \mathrm{d}s \,\middle|\, X_t = x\right] \leq 0.$$

Dividing by h and taking $h \to 0$ we get

$$\left(\partial_t + \mathcal{A}^a\right) H(t,x) + f(t,x,a) \leq 0.$$

As $a \in A$ was arbitrary, we have

$$\partial_t H(t,x) + \sup_{a \in A}\left[\mathcal{A}^a H(t,x) + f(t,x,a)\right] \leq 0.$$

Now, suppose that $\widehat{\nu}$ is an optimal control. Then taking $\widehat{\nu}$ as the control in (A.0.3) instead of the constant control $\nu = a$ gives us an equality (as opposed to an inequality). Then applying the same line of reasoning as above, we get that

$$\partial_t H(t,x) + \mathcal{A}^{\widehat{\nu}_t} H(t,x) + f(t,x,\widehat{\nu}_t) = 0.$$

As $\widehat{\nu}$ takes values in a bounded set A and $H(T,x) = g(x)$ we have that H should satisfy

$$\partial_t H(t,x) + \sup_{a\in A}\left[\mathcal{A}^a H(t,x) + f(t,x,a)\right] = 0, \qquad H(T,x) = g(x). \tag{A.0.5}$$

The equation (A.0.5) is known as the *Hamilton-Jacobi-Bellman (HJB)* equation.

The above discussion shows that if H is sufficiently smooth then H satisfies the HJB equation (A.0.5). However, if we already know the value function H then the fact that, it satisfies a specific PDE is likely of little use as we already have the value function in hand. What would be useful is the ability to cast the optimal control problem (A.0.1) as a PDE problem (namely (A.0.5)). When the control space is a singleton $\mathcal{A} = \{a_0\}$, the supremum in (A.0.5) is trivial and the PDE is reduced to the linear parabolic Cauchy problem for whom existence and uniqueness results are standard. Unfortunately, for interesting control problems the HJB equation is nonlinear, and, as such, solutions to the HJB equation corresponding to a control problem do not necessarily correspond to the value function. But, not all hope is lost! Fortunately, there are known sufficiency conditions which allow us to assert that a given solution to the HJB PDE corresponds to the value function of the associated control problem. This set of conditions is called the verification theorem. In addition to the verification of a candidate value function, the verification theorem allows us to obtain an optimal Markovian control. We state the verification theorem here (see [60] for proof).

Theorem A.0.5 (Verification Theorem)**.** *Let* $G \in C^{1,2}([0,T) \times \mathbb{R}^n) \cap C^0([0,T] \times \mathbb{R}^n)$. *If*

- $G(T,x) = g(x)$,
- *there exists a constant* $C > 0$ *such that*

$$|G(t,x)| \le C(1 + |x|^2), \qquad \forall (t,x) \in [0,T] \times \mathbb{R}^n,$$

- *there exists a measurable function* $\widehat{\nu} : [0,T) \times \mathbb{R}^n \mapsto A$ *such that*

$$\partial_t G(t,x) + \sup_{a\in A}\left[\mathcal{A}^a G(t,x) + f(t,x,a)\right]$$

$$= \partial_t G(t,x) + \mathcal{A}^{\widehat{\nu}(t,x)} G(t,x) + f(t,x,\widehat{\nu}(t,x))$$
$$= 0,$$

- *the SDE*

$$\mathrm{d}X_s^{\widehat{\nu}} = \mu\left(X_s^{\widehat{\nu}}, \widehat{\nu}(t, X_s^{\widehat{\nu}})\right) \mathrm{d}s + \sigma\left(X_s^{\widehat{\nu}}, \widehat{\nu}(t, X_s^{\widehat{\nu}})\right) \mathrm{d}W_s, \qquad X_t = x$$

  admits a unique solution, and

- *the process* $\{\widehat{\nu}(s, X_s^{\widehat{\nu}})\}_{t \le s \le T} \in \mathcal{A}_T(t,x)$,

then $G = H$ *on* $[0,T] \times \mathbb{R}^n$ *and* $\widehat{\nu}$ *is an optimal Markovian control.*

To recap, to the control problem (A.0.1) we have an associated dynamic programming equation, the HJB PDE (A.0.5), and a set of conditions which allows us to verify that a solution to the dynamic programming equation corresponds to the value function. For a given control problem (A.0.1), a common strategy is to write down the corresponding dynamic programming equation (A.0.5) and maximize the expression inside the supremum of (A.0.5). Doing so yields a deterministic function $\widehat{\nu}$ in terms of the value function H, which we refer to as a *feedback control*. We insert this expression back in to the HJB equation (i.e. we take the supremum in (A.0.5)) resulting in a PDE for H no longer dependent on ν. Solving the resulting PDE gives us a candidate value function and an associated Markovian control to which we can apply the verification theorem.

We provide a toy example to make the above procedure more clear.

Example A.0.6. Suppose that the process X^ν is given by the dynamics

$$\mathrm{d}X_t^\nu = (\nu_t - \alpha)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

where $\alpha > 0$. When the agent takes no action, the process X^ν drifts down at a rate α. Otherwise, the process has a drift of $\nu_t - \alpha$. Let

$$f(t,x,\nu) = -\varphi\nu^2, \qquad g(x) = x, \qquad \varphi > 0.$$

The value functional and corresponding value function are

$$\mathbb{H}(t,x,\nu) = \mathbb{E}\left[X_T^\nu - \varphi \int_t^T \nu_s^2 \, \mathrm{d}s \,\middle|\, X_t^\nu = x\right],$$

$$H(t,x) = \sup_{\nu \in \mathcal{A}_T(t,x)} \mathbb{H}(t,x,\nu). \tag{A.0.6}$$

An agent whose value function is (A.0.6) wishes to achieve a maximal expected value of X_T^ν from a starting state $X_t^\nu = x$ but is subject to a running quadratic penalty on the control.

The associated dynamic programming equation is

$$\partial_t H(t,x) + \sup_{\nu \in \mathbb{R}} \left[\frac{1}{2}\sigma^2 \partial_x^2 H(t,x) + (\nu - \alpha)\partial_x H(t,x) - \varphi \nu^2\right] = 0, \quad H(T,x) = x. \tag{A.0.7}$$

We see that the expression

$$\frac{1}{2}\sigma^2 \partial_x^2 H(t,x) + (\nu - \alpha)\partial_x H(t,x) - \varphi \nu^2$$

inside the supremum is quadratic in ν with a negative coefficient on the leading term and for any (t,x) has the unique maximizer

$$\widehat{\nu}(t,x) = \frac{\partial_x H(t,x)}{2\varphi}. \tag{A.0.8}$$

Inserting (A.0.8) into (A.0.7) (i.e. taking the supremum in (A.0.7)) yields

$$\partial_t H(t,x) + \frac{1}{2}\sigma^2 H(t,x) - \alpha \partial_x H(t,x) + \frac{1}{4\varphi}\left(\partial_x H(t,x)\right)^2 = 0, \qquad H(T,x) = x. \tag{A.0.9}$$

If we assume that $H(t,x) = x + h(t)$, then (A.0.9) is reduced to the ODE

$$\frac{dh(t)}{dt} - \alpha + \frac{1}{4\varphi} = 0, \qquad h(T) = 0,$$

whose solution is given by

$$h(t) = \left(\frac{1}{4\varphi} - \alpha\right)(T - t).$$

We have a candidate value function and optimal control

$$H(t,x) = x + \left(\frac{1}{4\varphi} - \alpha\right)(T - t), \qquad \widehat{\nu} = \frac{1}{2\varphi},$$

respectively. Applying Theorem A.0.5 gives that H and $\widehat{\nu}$ do indeed correspond to the value function and optimal Markovian control corresponding to the problem (A.0.6).