Optimal Execution of Backstopped Block Trades

by

Xuran Lin

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

Master of Science

in

Mathematical Finance

Department of Mathematical and Statistical Sciences
University of Alberta

© Xuran Lin, 2016

Abstract

In this thesis, we introduce and study a model for a broker who executes a client order and takes over its execution risk at some transition time. Such agreements between clients and brokers are often called backstopped trades. To minimize risk, it may be beneficial for the broker to trade on his/her own book, even before taking over the execution risk of the order. The broker is not allowed to trade in the same stock while executing on behalf of the client, but the broker may trade in a different, correlated stock. We formulate this question as a mean-variance optimization problem with two correlated stocks, also incorporating permanent and temporary market impacts.

We consider this problems in three different cases. In the first case, we assume that the transition time is deterministic. We then manage to find an explicit formula for the optimal trading strategy and analyze it in a numerical example. In the second case, we consider a stochastic transition time, but restrict the analysis to deterministic strategies. Under this assumption, we can characterize the optimal trading strategy through an ordinary differential equation. Finally, in the most general case of a stochastic transition time with stochastic strategies, we derive the Hamilton-Jacobi-Bellman equation corresponding to the optimization problem and describe a numerical implementation to find an approximate solution.

Acknowledgements

At the moment of my Master's thesis being close to completion, I would like to take this valuable opportunity to express my gratitude to people who helped me during my Master's study. Without their support, I could not make these accomplishments.

First of all, I reserve the most gratitude to my supervisor Dr. Christoph Frei. I feel very lucky and grateful of being accepted as his Master's student. I also would like to thank for his great patience, kindness and intelligence for the last two years. I benefited so much from him. He taught me on how to give an effective presentation, how to write a better thesis and how to work more efficiently. In addition, he guided me on my future career, provided invaluable help to me in job applications and offered insightful suggestions to aid my preparation for the interviews. His excellent teaching in the course Computational Finance inspired me not only to obtain knowledge itself, but also to possess programming skills, which proved to be an asset in my research and job hunting.

Besides my supervisor, I also feel grateful to all other committee members of my thesis defense, Dr. Tahir Choulli, Dr. Alexander Melnikov and Dr. Gary Smith, for their precious time and helpful comments.

Also, I would like to express my gratitude again for Dr. Abel Cadenillas, Dr. Alexander Melnikov and Dr. Gary Smith. Their intriguing courses Mathematical Finance, Stochastic Analysis and Managerial Finance gave me a deeper understanding of the fields of mathematical finance and finance, providing me with a solid foundation for my research.

Lastly, I want to thank all support staff of the Department of Mathematical and Statistical Sciences at the University of Alberta. It is my great pleasure to study at such a wonderful institute. I am thankful for my friends and my family. Their priceless support gave me strength and meant a lot to me. My experience in Edmonton is unforgettable.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Large client orders, known as *block trades*, can take different forms. In a *bought deal*, the broker acts as manager of the buy order (sell order is analogous), sells the shares to the client at the beginning and buys the shares later on the market. The entire risk is with the broker. The other extreme is *accelerated book building*, where the broker merely acts as an agent for the client and does not take any risk him-/herself. Between the two forms, there are *backstopped deals*, where the broker takes the shares on his/her own book at some transition time. From that moment onward, the risk is with the broker.[1] This leaves the broker and the client share the risk; see the Financial Times article [7]. The quantitative modeling and analysis of such backstopped deals is the topic of this thesis.

At the transition time, the broker commits to give the remaining shares to

[1] A sample agreement for backstopped deals is provided by the Association for Financial Markets in Europe (AMFE); see www.afme.eu/Documents/Standard-forms-and-documents.aspx.

the client for a fixed price. Therefore, the broker has the risk of price changes in buying these shares after the transition time. An idea might be that the broker could trade on his/her own book already before the transition time, in order to minimize the risk of the trades after the transition time. However, due to the regulatory restrictions, the broker is prohibited from trading the same stock when he/she is also trading on behalf of a client. So, it may be beneficial that the broker trades a second stock, which is correlated to the first stock.

For example, the client wants to buy a large volume of stock A. Before the transition time, the broker buys a certain quantity of a stock B which has high positive correlation with stock A. When the transition time becomes effective, the broker takes over the remaining order. After the transition time, the broker buys the shares in stock A at his/her own risk, but at the same time sells his/her position in stock B, thereby minimizing the risk of price fluctuations because of the high positive correlation. Therefore, the broker reduces the volatility of the cost due to price changes in the underlying stock. We note that the higher trade volume from also trading in stock B inevitably leads on average to more cost of the whole trade due to the price impact in trading in stock B. So, if the quantity of trade in the second stock is too big, the cost of the whole trade will more than offset the benefit from the reduction in risk. Therefore, there is a tradeoff between risk minimization and cost of trading due to the price impact. We decide to use mean-variance optimization to find optimal strategies and to take the tradeoff between expected costs and risk into account. Thus, we consider only the first two moments (mean and variance) of the trading costs while neglecting higher-order moments. This is standard in the literature (see Section 1.2 below) and allows for a tractable framework, which captures the main criteria used for measuring expected costs and execution risk.

The remainder of this thesis is structured as follows. In the following sections in this chapter, we will review some related literature and formulate the problem in a mathematical model. We then consider the problem based on three different assumptions on transition time and trading strategies. In Chapter 2, we assume both transition time and trading strategies are deterministic. In Chapter 3, we solve the problem under the assumption of a stochastic transition time with deterministic trading strategies. In Chapter 4, we consider the situation when both transition time and trading strategies are stochastic. Finally, we conclude and summarize the optimal solutions in Chapter 5. The Appendix contains MATLAB code used in the numerical examples of Chapters 3 and 4.

1.2 Some related literature

As equity markets are growing fast and new technology for algorithmic trading is becoming available, optimal order execution of large stock volume has become an important topic in both trading practice and academic research. We give next a very brief summary of some research developments in this area. Further relevant literature and a good overview can be found in the recent book by Cartea et al. [8].

In 1998, Bertsimas and Lo [6] used stochastic dynamic programming to find optimal trading strategies, which minimize the expected cost of trading a large block of equity over a fixed time horizon, given a price impact function. Almgren and Chriss [4] analyzed optimal execution of portfolio transactions with the aim of minimizing a combination of volatility risk and transaction costs arising from permanent and temporary price impact. They used a mean-variance formulation with the unimpacted stock price process modelled by arithmetic Brownian motion, and the arrival price chosen as a benchmark. Our situation differs from Almgren and Chriss [4] by the fact that the broker bears the execution risk of the order only from some transition time onward and can trade in another correlated stock, even before the transition time. The seminal paper by Almgren and Chriss [4] was generalized in various other works. For example, Almgren [2] incorporated a nonlinear price impact function. Bayraktar and Ludkovski [5] studied the optimal trade execution strategies in financial markets with discrete order flow, and Gatheral and Schied [11] extended the Almgren-Chriss setting from arithmetic Brownian motion to geometric Brownian motion as a model for the unaffected price process.

Further recent studies in this area include Schied [19], who investigates the robustness of the strategy which was derived in Gatheral and Schied [11]. Schied [19] proved that the strategy remains optimal whenever the unaffected price process is a square-integrable martingale and also found an explicit solution to the problem of minimizing the expected liquidation costs when the unaffected price process is a square-integrable semimartingale. In the same year 2013, Gatheral and Schied [12] studied the regularity of dynamical market impact models and their associated optimal order execution strategies. In addition, Guo et al. [13] proposed and studied an optimal placement problem in a limit order book with two simple models, one with price impact and one without price impact. Around the same time, a best execution problem in the fixed income markets was brought up in Almgren [3]. This paper mainly focused on interest rate markets and in particular on interest rates futures markets.

Curato et al. [9] explored the problem of the optimal execution of a large trade in the propagator model with nonlinear transient impact in 2014. Very recently, Frei and Westray [10] analyzed the optimal liquidation of a position of stock (long or short) where trading has a temporary market impact on the price. Their aim was to minimize a combination of the mean and variance of the order slippage with respect to a benchmark given by the market volume-weighted average price (VWAP).

1.3 The model and problem formulation

We consider a buy order of x shares of some stock. The order needs to be executed over a finite time horizon $[0, T]$. We assume that the stock prices are given by

$$S^u(t) = S(0) + \sigma B_1(t) + \Gamma_1 \int_0^t u(s)\,\mathrm{d}s + \gamma_1 u(t), \quad 0 \le t \le T. \tag{1.1}$$

where B_1 is a standard Brownian motion and $u(s)$ is the instantaneous rate of buying at time s.[2] The constants Γ_1 and γ_1 are the coefficients of permanent and temporary market impact. Note that the temporary market impact $\gamma_1 u(t)$ vanishes instantaneously if $u(t)$ becomes zero while the permanent market impact $\Gamma_1 \int_0^t u(s)\,\mathrm{d}s$ depends on the entire volume traded up to time t. The number of shares remaining to purchase at time t is denoted by $X^u(t)$ and is

[2]Formally, we are working on a filtered probability space $(\Omega, \mathscr{F}_T, (\mathscr{F}_s)_{0\le s\le T}, P)$, satisfying the usual conditions of completeness and right-continuity and containing two Brownian motions B_1 and B_2 with instantaneous correlation ρ (compare Section I.1 and I.3 of Protter [17]). The control process u needs to be progressively measurable with respect to this filtration.

given by

$$X^u(0) = x, \quad \mathrm{d}X^u(t) = -u(t)\,\mathrm{d}t, \quad X^u(T) = 0. \tag{1.2}$$

In a backstopped block order, the broker first acts as an agent for the client, but at some transition time τ, the broker takes the risk of the remaining order. In addition, we assume that when the broker acts merely as an agent, the buy orders are done at a constant rate (TWAP order) so that $u(s) = x/T$, which is in line with the model agreement mentioned in footnote 1.

For legal reasons, the broker is not allowed to do proprietary trading in the stock which the client is buying. However, the broker may trade a second, correlated stock. We assume that such a second, correlated stock has dynamics

$$P^v(t) = P_0 + \nu B_2(t) + \Gamma_2 \int_0^t v(s)\,\mathrm{d}s + \gamma_2 v(t), \quad 0 \le t \le T,$$

where B_2 is a Brownian motion with instantaneous correlation ρ to B_1. We assume that trading in the first stock (and the second stock) has a price impact only on the same stock. Otherwise, the broker could affect the transition time τ by trading in the second stock. If τ is a stopping time depending on the price of the second stock, we want to exclude this possibility. The broker starts and ends with zero exposure in the second stock, hence, the investment will result in a net loss on average due to the market impact. However, an investment in the second stock can still be beneficial for the broker since it may allow for a risk reduction because of the correlation between the two stocks. The (negative) holdings of shares in the second stock are

$$Z^v(0) = 0, \quad \mathrm{d}Z^v(s) = -v(s)\,\mathrm{d}s, \quad Z^v(T) = 0,$$

using the same sign convention as in (1.2).

As is classical, we assume that the broker uses a mean-variance criterion to minimize expected cost and risk. Hence, the broker's minimization problem is

$$E[Y^{u,v}] + \lambda \mathrm{Var}(Y^{u,v})$$

for a mean-variance tradeoff parameter $\lambda > 0$, where the costs of trading are

$$Y^{u,v} = \int_0^T P^v(s)v(s)\,\mathrm{d}s + \mathbf{1}_{\tau \le T} \int_\tau^T S^u(s)u(s)\,\mathrm{d}s.$$

The parameter λ models how much the broker cares about the execution risk (square root of $\mathrm{Var}(Y^{u,v})$) relative to the expected costs $E[Y^{u,v}]$.

1.4 Target problem

In this section, we introduce a vector notation for the stock prices, which allows us to write the minimization problem in a compact way.
First, we split $Y^{u,v}$ into two parts $Y^{u,v} = Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v}$, where we define

$$Y_1^v = \int_0^{\tau \wedge T} P^v(s)v(s)\,\mathrm{d}s$$

and

$$Y_2^{u,v} = \int_\tau^T P^v(s)v(s)\,\mathrm{d}s + \int_\tau^T S^u(s)u(s)\,\mathrm{d}s.$$

Our problem is two-dimensional with the positions of the stock given by

$$X(t) = \begin{bmatrix} X^u(t) \\ Z^v(t) \end{bmatrix}, \quad X(0) = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad X(\tau) = \begin{bmatrix} x\Big(1 - \dfrac{\tau}{T}\Big) \\ Z^v(\tau) \end{bmatrix}.$$

The prices of the stocks are

$$S(t)^\top = \big[S^u(t), P^v(t)\big].$$

The permanent market impact coefficients are

$$\Gamma = \begin{bmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{bmatrix}.$$

The temporary market impact coefficients are

$$\gamma = \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix}.$$

The Brownian motions are

$$B(t)^\top = \big[B_1(t), B_2(t)\big].$$

The instantaneous rates of buying stocks at time t are

$$w(t)^\top = \big[u(t), v(t)\big],$$

and the volatility matrix is

$$\Sigma = \begin{bmatrix} \sigma\sqrt{(1-\rho^2)} & \rho\sigma \\ 0 & \nu \end{bmatrix}.$$

Then we are able to express the prices of these two stocks in a vector notation:

$$S(t) = S(0) + \Sigma B(t) + \Gamma \int_0^t w(s)\,\mathrm{d}s + \gamma w(t).$$

Our minimization problem becomes $E[Y^{u,v}] + \lambda \mathrm{Var}(Y^{u,v})$, where

$$\begin{aligned} Y^{u,v} =& Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v} \\ =& \int_0^{\tau \wedge T} P^v(s) v(s)\,\mathrm{d}s + \mathbf{1}_{\tau \le T} \left(\int_\tau^T P^v(s) v(s)\,\mathrm{d}s + \int_\tau^T S^u(s) u(s)\,\mathrm{d}s \right) \\ =& \int_0^{\tau \wedge T} P^v(s) v(s)\,\mathrm{d}s + \mathbf{1}_{\tau \le T} \int_\tau^T S(s)^\top w(s)\,\mathrm{d}s. \end{aligned}$$

In conclusion, the broker wants to minimize

$$E[Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v}] + \lambda \mathrm{Var}[Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v}]$$

over $u(t)$ and $v(t)$, subject to $u(t) = x/T$ for $0 \le t < \tau$ because the broker is not allowed to trade the first stock as long as he/she is executing on behalf of the client. Moreover, the strategies u and v need to be progressively measurable and satisfy $E\Big[\int_0^T |u(t)|^2\,\mathrm{d}t \Big] < \infty$ and $E\Big[\int_0^T |v(t)|^2\,\mathrm{d}t \Big] < \infty$ to be admissible.

Chapter 2

Deterministic transition time

Recall that the transition time τ is the time when the broker takes over the risk of the remaining order of the first stock. As a starting point, we assume that τ is deterministic. We will see, under this assumption, we can find an explicit solution to the optimization problem, similarly to the solution in Section 1.2 in Almgren [2].

2.1 First case: $\tau \geq T$

When $\tau \geq T$, the transition time does not become effective, which means the deal becomes an "accelerated book building" and the client bears all the cost and risk. Thus, the broker does not need to do anything to minimize the objective function. Indeed, he/she is not allowed to trade in the first stock and chooses $v(t) = 0$ for all $t \in [0, T]$ (no trading in the second stock) to minimize the expected costs while risk cannot be reduced by trading in the second stock as there is no risk for the broker from the first stock.

2.2 Second case: $\tau < T$

2.2.1 Objective function

In this section, we assume τ is deterministic and $\tau \le T$. Using these assumptions, we can rearrange the terms of the costs of the trading:

$$\begin{aligned} Y^{u,v} =& Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v} \\ =& \int_0^{\tau \wedge T} P^v(s) v(s)\, \mathrm{d}s + \mathbf{1}_{\tau \le T} \int_\tau^T S(s)^\top w(s)\, \mathrm{d}s \\ =& \int_0^{\tau} P^v(s) v(s)\, \mathrm{d}s + \int_\tau^T S(s)^\top w(s)\, \mathrm{d}s. \end{aligned}$$

Since we have already set up the price model for the two stocks in the previous chapter, we now use it to simplify the function, which helps us study the minimization problem. We have

$$\begin{aligned} \int_0^\tau P^v(s) v(s)\, \mathrm{d}s =& - P_0 Z^v(\tau) - \nu B_2(\tau) Z^v(\tau) \\ &+ \int_0^\tau \nu Z^v(t)\, \mathrm{d}B_2(t) - \int_0^\tau \Gamma_2 Z^v(t) v(t)\, \mathrm{d}t + \int_0^\tau \gamma_2 v(t)^2\, \mathrm{d}t. \end{aligned}$$

Similarly for $\int_\tau^T S(s)^\top w(s)\, \mathrm{d}s$ but using matrix operations here, we can rewrite

$$\begin{aligned} \int_\tau^T S(s)^\top w(s)\, \mathrm{d}s =& \Big(S(0)^\top + B(\tau)^\top \Sigma\Big) X(\tau) + \int_\tau^T X(t)^\top \Sigma\, \mathrm{d}B(t) \\ &+ \int_\tau^T (X(0) - X(t))^\top \Gamma w(t)\, \mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t)\, \mathrm{d}t. \end{aligned}$$

Thus, after combining the two parts and organizing them, we can write $Y^{u,v}$ as a sum of some integrals plus two terms which are $\mathscr{F}_\tau$-measurable:

$$\begin{aligned} Y^{u,v} = &\int_0^\tau \nu Z^v(t)\,\mathrm{d}B_2(t) - \int_0^\tau \Gamma_2 Z^v(t)v(t)\,\mathrm{d}t + \int_0^\tau \gamma_2 v(t)^2\,\mathrm{d}t \\ &\int_\tau^T X(t)^\top \Sigma\,\mathrm{d}B(t) + \int_\tau^T (X(0)-X(t))^\top \Gamma w(t)\,\mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t)\,\mathrm{d}t \\ &+ S^u(0)X^u(\tau) + \big(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau)\big)B_1(\tau). \end{aligned}$$

Next, we use $Y^{u,v}$ to get its expectation and variance, which together compose the optimization problem. First, we know the main drivers of the variance are the stochastic integrals with Brownian motion as integrator by the approximation of Section 1.1 in Almgren [2]. Using this and the conditional variance formula, we can approximate the variance by

$$\begin{aligned} \mathrm{Var}(Y^{u,v}) &\approx E[a_1^v + a_2^{u,v}], \\ a_1^v &:= \int_0^\tau \nu^2 Z^v(s)^2\,\mathrm{d}s, \\ a_2^{u,v} &:= \int_\tau^T \big(X(t)^\top \Sigma\Sigma^\top X(t)\big)\,\mathrm{d}s + \big(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau)\big)^2\tau. \end{aligned}$$

Because $E\Big[\int_0^T \alpha(t)\,\mathrm{d}B(t)\Big] = 0$ for any progressively measurable process α with $E\Big[\int_0^T |\alpha(t)|^2\,\mathrm{d}t\Big] < \infty$, we have

$$\begin{aligned} E[Y^{u,v}] &= E[b_1^v + b_2^{u,v}], \\ b_1^v &:= -\int_0^\tau \Gamma_2 Z^v(t)v(t)\,\mathrm{d}t + \int_0^\tau \gamma_2 v(t)^2\,\mathrm{d}t, \\ b_2^{u,v} &:= \int_\tau^T (X(0)-X(t))^\top \Gamma w(t)\,\mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t)\,\mathrm{d}t + S^u(0)X^u(\tau). \end{aligned}$$

Note that since all of a_1^v, $a_2^{u,v}$, b_1^v and $b_2^{u,v}$ do not involve any stochastic expression, the optimal strategy will be deterministic and we can restrict ourselves to optimizing over deterministic strategies. Indeed, for any stochastic u and v, we have

$$a_1^v(\omega) + a_2^{u,v}(\omega) + \lambda\big(b_1^v(\omega) + b_2^{u,v}(\omega)\big) \geq \min_{\tilde{u},\tilde{v} \text{ deterministic}} a_1^{\tilde{v}} + a_2^{\tilde{u},\tilde{v}} + \lambda(b_1^{\tilde{v}} + b_2^{\tilde{u},\tilde{v}})$$

for all possible outcomes $\omega \in \Omega$ and hence

$$\begin{aligned} &\min_{u,v} E\big[a_1^v(\omega) + a_2^{u,v}(\omega) + \lambda\big(b_1^v(\omega) + b_2^{u,v}(\omega)\big)\big] \\ &\geq \min_{\tilde{u},\tilde{v} \text{ deterministic}} a_1^{\tilde{v}} + a_2^{\tilde{u},\tilde{v}} + \lambda(b_1^{\tilde{v}} + b_2^{\tilde{u},\tilde{v}}). \end{aligned}$$

On the other hand, deterministic strategies are special cases of stochastic strategies so that

$$\begin{aligned} &\min_{u,v} E\big[a_1^v(\omega) + a_2^{u,v}(\omega) + \lambda\big(b_1^v(\omega) + b_2^{u,v}(\omega)\big)\big] \\ &\leq \min_{\tilde{u},\tilde{v} \text{ deterministic}} a_1^{\tilde{v}} + a_2^{\tilde{u},\tilde{v}} + \lambda(b_1^{\tilde{v}} + b_2^{\tilde{u},\tilde{v}}), \end{aligned}$$

and thus

$$\begin{aligned} &\min_{u,v} E\big[a_1^v(\omega) + a_2^{u,v}(\omega) + \lambda\big(b_1^v(\omega) + b_2^{u,v}(\omega)\big)\big] \\ &= \min_{\tilde{u},\tilde{v} \text{ deterministic}} a_1^{\tilde{v}} + a_2^{\tilde{u},\tilde{v}} + \lambda(b_1^{\tilde{v}} + b_2^{\tilde{u},\tilde{v}}). \end{aligned}$$

Therefore, we can restrict ourselves to minimizing over deterministic strategies, and we can split the objective function

$$OF = E[Y^{u,v}] + \lambda E[a_1^v + a_2^{u,v}],$$

into two parts by two time periods, $[0, \tau]$ and $[\tau, T]$:

$$OF = OF_1 + OF_2,$$

where

$$\begin{aligned}
OF_1 =& E[b_1^v + \lambda a_1^v] \\
=& -\int_0^\tau \Gamma_2 Z^v(t) v(t) \,\mathrm{d}t + \int_0^\tau \gamma_2 v(t)^2 \,\mathrm{d}t + \int_0^\tau \lambda \nu^2 Z^v(s)^2 \,\mathrm{d}s, \\
OF_2 =& E[b_2^{u,v} + \lambda a_2^{u,v}] \\
=& \int_\tau^T (X(0) - X(t))^\top \Gamma w(t) \,\mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t) \,\mathrm{d}t \\
&+ \int_\tau^T \lambda \big(X(t)^\top \Sigma \Sigma^\top X(t) \big) \,\mathrm{d}s \\
&+ \lambda \big(\sigma \sqrt{1-\rho^2} X^u(\tau) + \rho \sigma Z^v(\tau) \big)^2 \tau + S^u(0) X^u(\tau).
\end{aligned}$$

In the next step, we analyze the trading strategies on these two periods and solve them backwards. If we just solved them separately, we would obtain suboptimal solutions for the minimization problem, since these two periods are related and not independent. Therefore, we solve the second period first, and use that solution to find the optimal strategy for the first period and then combine them.

2.2.2 Optimal strategy from τ to T

For the second period $[\tau, T]$, using this and the conditional expectation formula, we make some modifications:

$$
\begin{aligned}
OF_2 &= E[b_2^{u,v} + \lambda a_2^{u,v}] \\
&= E\big[E[b_2^{u,v} + \lambda a_2^{u,v} | \mathscr{F}_\tau]\big].
\end{aligned}
$$

To solve the problem above, we just need

$$
\min E[Y_2^{u,v} + \lambda a_2^{u,v} | \mathscr{F}_\tau]
$$

over u and v on $[\tau, T]$ for given $Z^v(\tau)$. As we know,

$$
\lambda\big(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau)\big)^2\tau + S^u(0)X^u(\tau)
$$

is $\mathscr{F}_\tau$-measurable. Then our problem becomes to minimize the integral:

$$
I_1 = \int_\tau^T \big(\lambda X(t)^\top \Sigma\Sigma^\top X(t) + (X(0) - X(t))^\top \Gamma w(t) + w(t)^\top \gamma w(t)\big)\, \mathrm{d}t.
$$

Although our problem is a two-dimensional optimization, we can generalize it to a multidimensional situation and find the general solution. Once we get the solution, we apply it to our problem and get our optimal strategy.

Lemma 2.1. *At time t, where $\tau \leq t \leq T$, consider a multidimensional setting with the positions of the stock $X(t)$, the prices of the stocks $S(t)$, and the instantaneous rates of buying stocks $w(t)$ are n-dimension column vectors; the permanent coefficients Γ, and the temporary coefficients γ are $(n \times n)$-*

dimensional diagonal matrices; the Brownian motions $B(t)$ is a d-dimensional column vector. Still, Σ is an $(n \times n)$-dimensional matrix. The multidimensional problem is to minimize

$$I = \int_{\tau}^{T} \Big(\lambda\big(X(t)^{\top}\Sigma\Sigma^{\top}X(t)\big) + (X(0) - X(t))^{\top}\Gamma w(t) + w(t)^{\top}\gamma w(t)\Big)\,\mathrm{d}t.$$

Then the optimal solution is

$$X(t) = \sqrt{\gamma}^{-1} Q H(t) Q^{\top} \sqrt{\gamma} X(\tau) \tag{2.1}$$

where $H(t)$ is a diagonal matrix with entries

$$\begin{aligned}&\Big[H_1(t), H_2(t), \ldots, H_n(t)\Big]\\ =&\left[\frac{\sinh(\kappa_1(T-t))}{\sinh(\kappa_1(T-\tau))}, \frac{\sinh(\kappa_2(T-t))}{\sinh(\kappa_2(T-\tau))}, \ldots, \frac{\sinh(\kappa_n(T-t))}{\sinh(\kappa_n(T-\tau))}\right],\end{aligned}$$

$\kappa_i^2 = \lambda D_{ii}$ *, for* $1 \le i \le n$ *, and* $(\sqrt{\gamma})^{-1}\Sigma\Sigma^{\top}(\sqrt{\gamma})^{-1} = QDQ^{\top}$*, where Q is an orthogonal matrix and D is a diagonal matrix.*

Proof. This is a minimization problem and the target function is

$$\int_{\tau}^{T} \Big(\lambda\big(X(t)^{\top}\Sigma\Sigma^{\top}X(t)\big) + (X(0) - X(t))^{\top}\Gamma w(t) + w(t)^{\top}\gamma w(t)\Big)\,\mathrm{d}t. \tag{2.2}$$

Since $X'(t) = -w(t)$, using the Euler Lagrange equation, we can get

$$X''(t) = \lambda(\gamma)^{-1}\Sigma\Sigma^{\top}X(t).$$

Recall that the Euler Lagrange equation gives a differential equation for the optimizer of an integral of the form (2.2), where a function $X(t)$ and its derivative

$X'(t) = -w(t)$ appear. Further information on the use of the Euler Lagrange equation can be found in Roubicek [18].
By the eigendecomposition, we can write $(\sqrt{\gamma})^{-1}\Sigma\Sigma^\top(\sqrt{\gamma})^{-1} = QDQ^\top$, where Q is an orthogonal matrix and D is a diagonal matrix since $(\sqrt{\gamma})^{-1}\Sigma\Sigma^\top(\sqrt{\gamma})^{-1}$ is a symmetric matrix. Set $U(t) = Q^\top\sqrt{\gamma}X(t)$, the problem becomes

$$U''(t) = \lambda U(t),$$

which has the solution $U(t) = H(t)U(\tau)$, and shows (2.1). $\square$

We now apply Lemma 2.1 with $n = 2$ and $d = 2$. In addition, we choose an orthogonal matrix Q, and diagonal matrices D and H such that $(\sqrt{\gamma})^{-1}\Sigma\Sigma^\top(\sqrt{\gamma})^{-1} = QDQ^\top$ with

$$Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix}, D = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix},$$

$$H(t) = \begin{bmatrix} H_1(t) & 0 \\ 0 & H_2(t) \end{bmatrix} = \begin{bmatrix} \frac{\sinh(\kappa_1(T-t))}{\sinh(\kappa_1(T-\tau))} & 0 \\ 0 & \frac{\sinh(\kappa_2(T-t))}{\sinh(\kappa_2(T-\tau))} \end{bmatrix},$$

where $\kappa_1^2 = \lambda D_{11}$ and $\kappa_2^2 = \lambda D_{22}$.

Finally, using Lemma 2.1, we get that the optimal solution for the stock positions $X(t)$ is

$$X(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} X(\tau), \tag{2.3}$$

where

$$
\begin{aligned}
A_{11}(t) &= H_1(t)Q_1^2 + H_2(t)Q_2^2, \\
A_{12}(t) &= (H_1(t)Q_1Q_3 + H_2(t)Q_2Q_4)\sqrt{\frac{\gamma_2}{\gamma}}, \\
A_{21}(t) &= (H_1(t)Q_1Q_3 + H_2(t)Q_2Q_4)\sqrt{\frac{\gamma}{\gamma_2}}, \\
A_{22}(t) &= H_1(t)Q_3^2 + H_2(t)Q_4^2.
\end{aligned}
$$

Besides, the optimal solution for the instantaneous rates $V(t)$ of trading stocks is

$$
V(t) = \begin{bmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & B_{22}(t) \end{bmatrix} X(\tau) \tag{2.4}
$$

where

$$
\begin{aligned}
B_{11}(t) &= \kappa_1 H_1(t)Q_1^2 + \kappa_2 H_2(t)Q_2^2, \\
B_{12}(t) &= (\kappa_1 H_1(t)Q_1Q_3 + \kappa_2 H_2(t)Q_2Q_4)\sqrt{\frac{\gamma_2}{\gamma}}, \\
B_{21}(t) &= \kappa_1 H_1(t)Q_1Q_3 + \kappa_2 H_2(t)Q_2Q_4, \\
B_{22}(t) &= \kappa_1 H_1(t)Q_3^2 + \kappa_2 H_2(t)Q_4^2\sqrt{\frac{\gamma}{\gamma_2}}.
\end{aligned}
$$

2.2.3 Value of the objective function from τ to T

After we have determined the solution over $[\tau, T]$, we use it to get the minimal value of $OF_2 = E[Y_2^{u,v} + \lambda a_2^{u,v} | \mathscr{F}_\tau]$ over that period for given $Z^v(\tau)$. Because our ultimate goal is to find the optimal strategy for the whole period, the objective function OF_2 will be needed later. We have to combine it with the objective function OF_1 over the first period $[0,\tau]$ to determine the optimal holding position of the second stock.

We plug the solution back into the function

$$\begin{aligned}
OF_2 =& E[b_2^{u,v} + \lambda a_2^{u,v}] \\
=& \int_\tau^T \Big(\lambda\big(X(t)^\top \Sigma\Sigma^\top X(t)\big) + (X(0) - X(t))^\top \Gamma w(t) + w(t)^\top \gamma w(t)\Big)\,\mathrm{d}t \\
&+ \lambda(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau))^2\tau + S^u(0)X^u(\tau).
\end{aligned}$$

Assume

$$A(t)^\top \Sigma\Sigma^\top A(t) = \begin{bmatrix} M_{11}(t) & M_{12}(t) \\ M_{21}(t) & M_{22}(t) \end{bmatrix},$$

then accordingly,

$$\begin{aligned}
M_{11}(t) = \quad & A_{11}(t)^2\sigma^2 + 2A_{12}(t)A_{21}(t)\rho\sigma\nu + \\
& A_{21}(t)^2\nu^2, \\
M_{12}(t) = \quad & A_{11}(t)A_{12}(t)\sigma^2 + A_{11}(t)A_{22}(t)\rho\sigma\nu + \\
& A_{12}(t)A_{21}(t)\rho\sigma\nu + A_{21}(t)A_{22}(t)\nu^2, \\
M_{21}(t) = \quad & A_{11}(t)A_{12}(t)\sigma^2 + A_{11}(t)A_{22}(t)\rho\sigma\nu + \\
& A_{12}(t)A_{21}(t)\rho\sigma\nu + A_{21}(t)A_{12}(t)\nu^2, \\
M_{22}(t) = \quad & A_{12}(t)^2\sigma^2 + 2A_{12}(t)A_{22}(t)\rho\sigma\nu + \\
& A_{22}(t)^2\nu^2.
\end{aligned}$$

Since $X^u(\tau) = x\Big(1 - \dfrac{\tau}{T}\Big)$, so $OF_2 = N_1 Z^v(\tau)^2 + N_2 Z^v(\tau) + N_3$, where

$$
\begin{aligned}
N_1 = &\int_\tau^T \Big[\lambda M_{22}(t) - \Gamma_1 A_{12}(t)B_{12}(t) - \Gamma_2 A_{22}(t)B_{22}(t) \\
&+ (B_{12}(t)^2\gamma_1 + B_{22}(t)^2\gamma_2)\Big]\,\mathrm{d}t + \lambda\sigma_1^2\rho^2\tau, \\
N_2 = &\int_\tau^T \Big[\lambda(M_{12}(t) + M_{21}(t))x\Big(1 - \frac{\tau}{T}\Big) + B_{12}(t)\Gamma_1 x \\
&- \Gamma_1 A_{11}(t)B_{12}(t)x\Big(1 - \frac{\tau}{T}\Big) - \Gamma_1 B_{11}(t)A_{12}(t)x\Big(1 - \frac{\tau}{T}\Big) \\
&- \Gamma_2 B_{22}(t)A_{21}(t)x\Big(1 - \frac{\tau}{T}\Big) - \Gamma_2 B_{21}(t)A_{22}(t)x\Big(1 - \frac{\tau}{T}\Big) \\
&+ 2\gamma_1 B_{12}(t)B_{11}(t)x\Big(1 - \frac{\tau}{T}\Big) + 2\gamma_1 B_{22}(t)B_{21}(t)x\Big(1 - \frac{\tau}{T}\Big)\Big]\,\mathrm{d}t \\
&+ 2\lambda\rho\sigma^2 x\Big(1 - \frac{\tau}{T}\Big)\tau\sqrt{1-\rho^2}, \\
N_3 = &\int_\tau^T \Big[\lambda M_{11}(t)x^2\Big(1 - \frac{\tau}{T}\Big)^2 + \Gamma_1 B_{11}(t)x^2\Big(1 - \frac{\tau}{T}\Big) \\
&- \Gamma_1 A_{11}(t)B_{11}(t)x^2\Big(1 - \frac{\tau}{T}\Big)^2 - \Gamma_2 A_{21}(t)B_{21}(t)x^2\Big(1 - \frac{\tau}{T}\Big)^2 \\
&+ (B_{11}(t)^2\gamma_1 + B_{21}(t)^2\gamma_2)x^2\Big(1 - \frac{\tau}{T}\Big)^2\Big]\,\mathrm{d}t \\
&+ \lambda\sigma^2(1-\rho^2)x^2\Big(1 - \frac{\tau}{T}\Big)^2\tau + S(0)x\Big(1 - \frac{\tau}{T}\Big).
\end{aligned}
$$

2.2.4 Optimal strategy from 0 to τ

In this subsection, we consider the objective function over $[0, \tau]$ and assume $\tau > 0$. We start by simplifying our objective function:

$$
\begin{aligned}
OF_1 =& E[b_1^v + \lambda a_1^v] \\
=& \int_0^\tau \left[\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2\right]\mathrm{d}t.
\end{aligned}
$$

As mentioned earlier, we have to combine the objective functions for both periods and then find the optimal solution on $[0,\tau]$, namely,

$$
\begin{aligned}
OF =& OF_1 + OF_2 \\
=& \int_0^\tau \left[\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2\right] \mathrm{d}t \\
& + N_1 Z^v(\tau)^2 + N_2 Z^v(\tau) + N_3.
\end{aligned}
$$

We notice that this objective function is mixed of Lagrange form and Mayer form. Because it involves a running cost (Lagrange form) and a terminal cost (Mayer form). We transform the Mayer form to the Lagrange form so that we just aim to minimize an integral, which is much easier for us to calculate.

First, set $K(Z^v(\tau)) = N_1 Z^v(\tau)^2 + N_2 Z^v(\tau) + N_3$, then $K(Z^v(0)) = N_3$. Secondly, we make the transformation of $K(Z^v(\tau))$:

$$
\begin{aligned}
K(Z^v(\tau)) &= K(Z^v(0)) + \int_0^\tau \frac{\mathrm{d}}{\mathrm{d}t}\big(K(Z^v(t)\big)\mathrm{d}t \\
&= \int_0^\tau \big(2N_1 Z^v(t) Z^v(t)' + N_2 Z^v(t)'\big)\mathrm{d}t + N_3.
\end{aligned}
$$

Since $Z^v(t)' = -v(t)$,

$$
\begin{aligned}
OF = \int_0^\tau & \left[\lambda\nu^2 Z^v(t)^2 + (2N_1 + \Gamma_2) Z^v(t) Z^v(t)' + \gamma_2 Z^v(t)'^2\right. \\
& \left. + N_2 Z^v(t)'\right]\mathrm{d}t + N_3.
\end{aligned}
$$

It becomes solvable by the Euler Lagrange equation:

$$
Z^v(t)'' = \frac{\lambda\nu^2}{\gamma_2} Z^v(t).
$$

Then set $\theta = \sqrt{\dfrac{\lambda\nu^2}{\gamma_2}}$, the solution is $Z^v(t) = c_1 \mathrm{e}^{\theta t} + c_2 \mathrm{e}^{-\theta t}$. Using the initial condition $Z^v(0) = 0$, our solution becomes

$$\begin{aligned} Z^v(t) &= 2c_1 \frac{\mathrm{e}^{\theta t} - \mathrm{e}^{-\theta t}}{2} = 2c_1 \sinh(\theta t), \\ Z^v(t)' &= 2c_1\theta \frac{\mathrm{e}^{\theta t} + \mathrm{e}^{-\theta t}}{2} = 2c_1\theta \cosh(\theta t). \end{aligned}$$

Because we only have one initial condition, so that there still is one unknown variable c_1 remaining. Therefore, we have to find another way to determine that constant. Plugging the solutions back into OF, we get a quadratic function of c_1,

$$\begin{aligned} OF =& \Big[\int_0^\tau 4\lambda\nu^2 \sinh(\theta t)^2 + 4(\Gamma_2 + 2N_1)\theta \sinh(\theta t)\cosh(\theta t) \\ &+ 4\gamma_2\theta^2 \cosh(\theta t)^2 \,\mathrm{d}t \Big] c_1^2 + \Big[\int_0^\tau 2N_2\theta \cosh(\theta t)\,\mathrm{d}t \Big] c_1 + N_3 \\ =& N_4 c_1^2 + N_5 c_1 + N_3. \end{aligned}$$

We know that the minimal value of this quadratic function can be achieved with $c_1 = -\dfrac{N_5}{2N_4}$. Finally, the solution becomes

$$\begin{aligned} Z^v(t) &= -\frac{N_5(\tau)}{N_4(\tau)} \sinh(\theta t), \\ Z^v(t)' &= -\frac{N_5(\tau)}{N_4(\tau)} \theta \cosh(\theta t). \end{aligned}$$

2.2.5 Conclusions

In conclusion, to minimize the objective function OF, the solution is

$$Z^v(t) = -\frac{N_5(\tau)}{N_4(\tau)}\sinh(\theta t),\ Z^v(t)' = -\frac{N_5(\tau)}{N_4(\tau)}\theta\cosh(\theta t)$$

for $t \in [0, \tau]$, and

$$X(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} X(\tau),\ V(t) = \begin{bmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & B_{22}(t) \end{bmatrix} X(\tau)$$

for $t \in [\tau, T]$, where

$$X(\tau) = \begin{bmatrix} x\left(1 - \frac{\tau}{T}\right) \\ Z^v(\tau) \end{bmatrix},$$

and the matrices A and B can be seen in (2.3) and (2.4).

2.2.6 Numerical examples and explanations

This subsection consists of two main parts. One part is to treat the optimal strategy of two stocks as a function of λ, so we can see how it changes with varied values of λ. The other part is to compare the optimal strategy we get and the strategy with only one stock with respect to three aspects: expected costs, approximate variances and objective functions. For this example, we choose some parameters based on Chapter 4 in Almgren and Chriss [4].

$S(0)$	$P(0)$	ρ	T	τ	x
50	30	0.78	5	3	10^6
σ	ν	Γ_1	γ_1	Γ_2	γ_2
0.95	0.83	2.5×10^{-7}	2.3×10^{-6}	2.8×10^{-7}	2×10^{-6}

Table 2.1: Chosen values of the model parameters

First, we get three different trading strategies with three different λ, the stock positions are as follows:

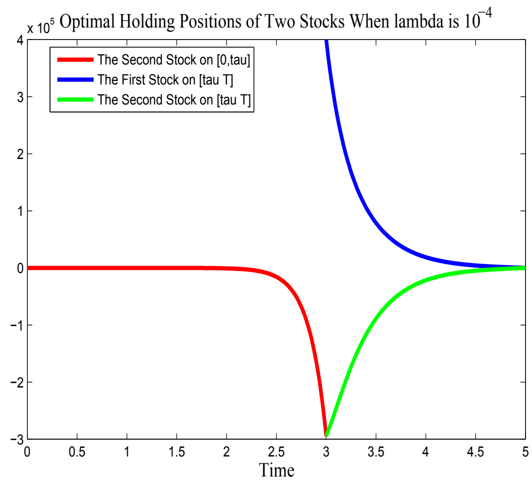

Figure 2.1: Optimal positions for mean-variance tradeoff parameter $\lambda = 10^{-4}$

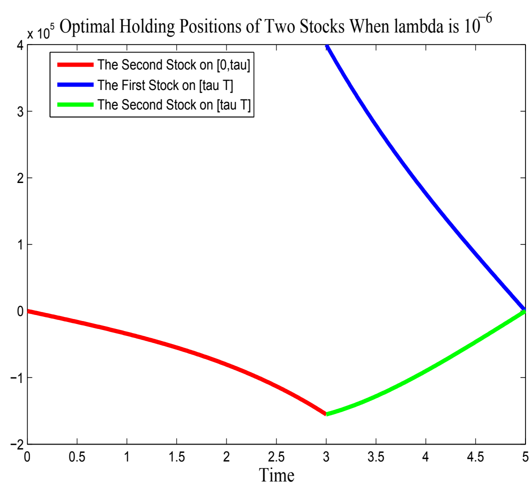

Figure 2.2: Optimal positions for mean-variance tradeoff parameter $\lambda = 10^{-6}$

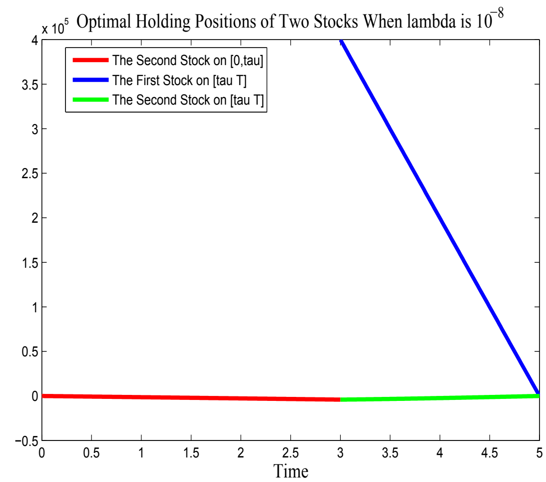

Figure 2.3: Optimal positions for mean-variance tradeoff parameter $\lambda = 10^{-8}$

The figures show when λ decreases, the holding position of the second stock at time τ declines; the rate of trading the first stock is high at the beginning and decreases as time progresses.

Explanations:

(1) λ is a risk-aversion parameter, so when it is bigger, the broker pays more attention to risk. The function of the second stock is to help minimize the risk. When the broker cares more about the risk, we should trade more volume of the second stock to reduce the variance (the square of risk).

(2) As for the first stock, if we split the order and trade most in the early period, we can reduce the risk thanks to the lower exposure to the stock price fluctuations.

Secondly, I analyze the solution as a function of λ from 0 to 10^{-4}, and get the Figures 2.4–2.6, which further confirm the explanations above.

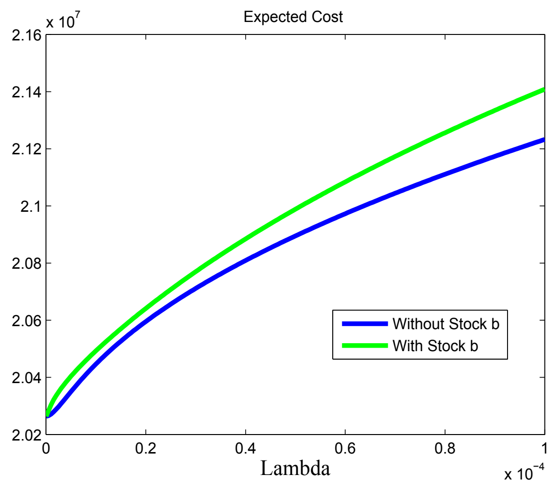

Figure 2.4: Expected cost of optimal strategy

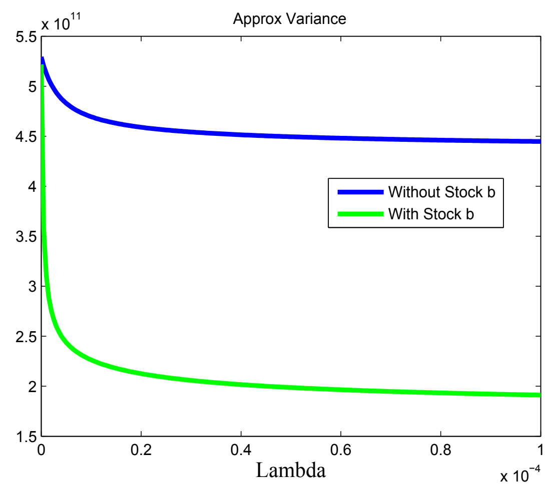

Figure 2.5: Variance of optimal strategy

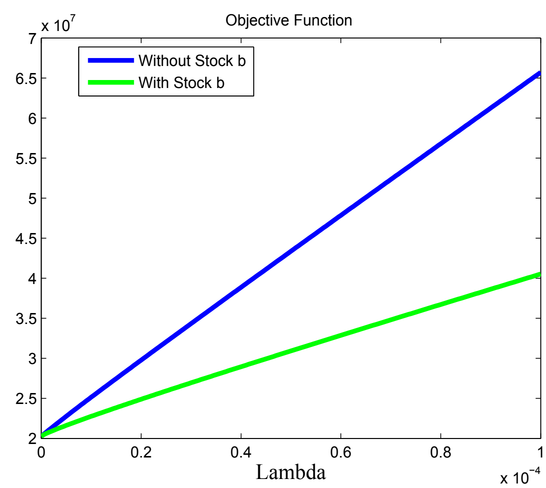

Figure 2.6: Objective function

Although the expected cost of the whole trade increases when investments in the second stocks are made, the objective function is still taking much lower values than when trading only one stock because the position of the second stock allows the broker to reduce the risk from the opposite position in the first stock due to the correlation between the two stocks.

Chapter 3

Stochastic transition time with deterministic strategies

As we know, τ is often not fixed in reality, so it makes sense to also analyze the situation where the transition time τ is stochastic, which we do next. To model this, we assume the client and broker agree that when the stock price hits the level $S_0 + L$, the broker takes the risk of the remaining order and the client receives the remaining order at price $S_0 + L$ from the broker. For the problem to make sense, we need to have $L \geq \gamma x/T$, as otherwise, the broker would take over the order immediately at the beginning. The time when the broker takes over the order is

$$\tau = \inf\{t \geq 0 : S^u(t) \geq S_0 + L\}. \tag{3.1}$$

Note that in this definition of τ, we only use $S^u(t)$ before time τ, which is given by

$$\begin{aligned} S^u(t) &= S(0) + \sigma B_1(t) + \Gamma_1 \int_0^t u(s)\,\mathrm{d}s + \gamma_1 u(t) \\ &= S(0) + \sigma B_1(t) + \frac{\Gamma_1 t x}{T} + \frac{\gamma_1 x}{T} \end{aligned}$$

because $u(t) = x/T$ before time τ. This shows that $S^u(t)$ used in the definition (3.1) of τ does not depend on the strategy that the broker is using.

Lemma 3.1 (Density of hitting time)**.** *The density function of τ is*

$$f_\tau(t) = \frac{L - \gamma_1 x/T}{\sigma\sqrt{2\pi t^3}} \exp\left(- \frac{(L - \gamma_1 x/T + \Gamma_1 x t/T)^2}{2\sigma^2 t} \right). \tag{3.2}$$

Proof. First, we recall the density function of the stopping time without drift, which is $\nu = \inf\{t \geq 0 : B_1(t) \geq b\}$, where b is some constant. We use the reflection principle, similarly to Section 2.6.A in Karatzas and Shreve [15]. We know that

$$P[\nu < t] = P[\nu < t, B_1(\nu) > b] + P[\nu < t, B_1(\nu) < b].$$

On the one hand, $P[\nu < t, B_1(\nu) > b] = P[B_1(\nu) > b]$; on the other hand, if $\nu < t$ and $B_1(\nu) < b$, then sometime before time ν the Brownian path reached the level b and in the remaining time, it moved from b to some point, which is less than b. Since the path of $B_1(t)$ with respect to b is symmetric starting at b, the probability that $B_1(\nu)$ lies above b is the same as that below b. Thus,

$$P[\nu < t, B_1(\nu) < b] = P[\nu < t, B_1(\nu) > b] = P[B_1(\nu) > b],$$

which gives us

$$P[\nu < t] = 2P[B_1(\nu) > b] = \sqrt{\frac{2}{\pi}} \int_{bt^{-\frac{1}{2}}}^{\infty} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x.$$

The density function becomes

$$f_\nu(t) = \frac{|b|}{\sqrt{2\pi t^3}} \exp\Big(-\frac{b^2}{2t}\Big). \tag{3.3}$$

Secondly, in our case, $\tau = \inf\{t \geq 0 : S^u(t) \geq S_0 + L\}$, which leads to $\tau = \inf\big\{t \geq 0 : B_1(t) + \Gamma_1 \frac{xt}{\sigma T} \geq \frac{L - \gamma_1 \frac{x}{T}}{\sigma}\big\}$. Then set $\tilde{W}_t = B_1(t) - ut$, $u = -\Gamma_1 \frac{x}{\sigma T}$ and $b = \frac{L - \gamma_1 \frac{x}{T}}{\sigma}$ so that $\tau = \inf\{t \geq 0 : \tilde{W}_t \geq b\}$ has been written as a hitting time of a Brownian motion with drift.

Based on Corollary 3.5.2 in Karatzas and Shreve [15], the process $(\tilde{W}_t)_{t\geq 0}$ is a Brownian motion under the probability measure $P^{(u)}(A) = E[\mathbf{1}_A Z_T]$, $A \in \mathscr{F}_T^{B_1}$, where $Z_T = \exp\big(uB_1(T) - \frac{1}{2}u^2 T\big)$. In addition, due to $\{\tau < t\} \in \mathscr{F}_t^{\tilde{W}} \cap \mathscr{F}_t^{B_1}$, we get $Z_{t\wedge\tau} = Z_\tau$. Besides, note that

$$\frac{1}{Z_t} = \exp\Big(-uB_1(t) + \frac{1}{2}u^2 t\Big) = \exp\Big(-u\tilde{W}_t - \frac{1}{2}u^2 t\Big)$$

is a martingale under $P^{(u)}$. Therefore, using the optional sampling theorem, we have for all $t \le T$ that

$$
\begin{aligned}
P[\tau \le t] &= E[\mathbf{1}_{\tau \le t}] = E^{(u)}\left[\frac{\mathbf{1}_{\tau \le t}}{Z_T}\right] \\
&= E^{(u)}\left[\mathbf{1}_{\tau \le t} E^{(u)}\left[\frac{1}{Z_T}\Big|\mathscr{F}^{\tilde{W}}_{t\wedge\tau}\right]\right] = E^{(u)}\left[\mathbf{1}_{\tau \le t}\frac{1}{Z_\tau}\right] \\
&= E^{(u)}\left[\mathbf{1}_{\tau \le t}\exp\Big(-ub - \frac{1}{2}u^2\tau\Big)\right] \\
&= \int_0^t \exp\Big(-ub - \frac{1}{2}u^2 s\Big) f(s)\,\mathrm{d}s, \qquad (3.4)
\end{aligned}
$$

where $E^{(u)}$ denotes the expectation under $P^{(u)}$. Then we combine (3.3) and (3.4) and we can conclude that for $\tilde{W}_t = B_1(t) - ut$, the density function of $\tau = \inf\{t \ge 0 : \tilde{W}_t \ge b\}$ is

$$
f_\tau(t) = \frac{|b|}{\sqrt{2\pi t^3}} \exp\left(-\frac{(b+ut)^2}{2t}\right),
$$

which leads to (3.2), given $u = -\Gamma_1 \frac{x}{\sigma T}$ and $b = \frac{L - \gamma_1 \frac{x}{T}}{\sigma}$. $\square$

Actually, there is something else we should note. Because in real life, the trading strategy of the second stock depends on the course of the price of the first stock. If the price of the first stock rises a lot and becomes really close to the level $S(0)+L$, we should have a large short position in the second stock to hedge the risk, because there is a higher chance that the price of the first stock will hit the level $S(0)+L$ soon. Conversely, if the price of the first stock decreases a lot and it seems unlikely that it hits the level $S(0)+L$, we then should have only a small short position or even none of the second stock.

However, in this chapter, we will assume the trading strategy of the second stock is deterministic, which means it will not be affected by the movement of the price of the first stock.

3.1 Objective function

We still face the same minimization problem:

$$\min E[Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v}] + \lambda \mathrm{Var}[Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v}]$$

over $u(t)$ and $v(t)$, subject to $u(t) = x/T$ for $0 \le t < \tau$, where

$$\begin{aligned} Y^{u,v} =& Y_1^v + \mathbf{1}_{\tau \le T} Y_2^{u,v} \\ =& \int_0^{\tau \wedge T} P^v(s) v(s) \, \mathrm{d}s + \mathbf{1}_{\tau \le T} \int_\tau^T S^u(s) u(s) \, \mathrm{d}s. \end{aligned}$$

Similarly as in Chapter 2, we need to find the objective function. Although τ here is stochastic, τ actually has no influence during the trading because our trading strategy is deterministic at the beginning of our the trade. Firstly, we separate our trade into two parts, over $[\tau, T]$ and $[0, \tau]$. For $[\tau, T]$, we still get

$$\begin{aligned} Y_2^{u,v} &= \int_\tau^T S(s) w(s) \, \mathrm{d}s \\ &= \big(S(0)^\top + B(\tau)^\top \Sigma\big) X(\tau) + \int_\tau^T X(t)^\top \Sigma \, \mathrm{d}B(t) \\ &\quad + \int_\tau^T (X(0) - X(t))^\top \Gamma w(t) \, \mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t) \, \mathrm{d}t. \end{aligned}$$

For $[0,\tau]$, we should notice that the upper limit of the integral is $\tau \wedge T$ instead, thus

$$
\begin{aligned}
Y_1^v = &\int_0^{\tau\wedge T} P^v(s)v(s)\,\mathrm{d}s \\
= &- P_0 Z^v(\tau\wedge T) - \nu B_2(\tau\wedge T) Z^v(\tau\wedge T) \\
&+ \int_0^{\tau\wedge T} \nu Z^v(t)\,\mathrm{d}B_2(t) - \int_0^{\tau\wedge T} \Gamma_2 Z^v(t)v(t)\,\mathrm{d}t + \int_0^{\tau\wedge T} \gamma_2 v(t)^2\,\mathrm{d}t.
\end{aligned}
$$

Secondly, we combine them, rearrange them by

$$
Z^v(\tau\wedge T) = \mathbf{1}_{\tau\le T} Z^v(\tau) + \mathbf{1}_{\tau>T} Z^v(T) = \mathbf{1}_{\tau\le T} Z^v(\tau),
$$

and get a similar result to that in Chapter 2, namely,

$$
\begin{aligned}
Y^{u,v} = &\int_0^{\tau\wedge T} \nu Z^v(\tau)\,\mathrm{d}B_2(t) - \int_0^{\tau\wedge T} \Gamma_2 Z^v(t)v(t)\,\mathrm{d}t + \int_0^{\tau\wedge T} \gamma_2 v(t)^2\,\mathrm{d}t \\
&\mathbf{1}_{\tau\le T}\Bigg(\int_\tau^T X(t)^\top \Sigma\,\mathrm{d}B(t) + \int_\tau^T (X(0)-X(t))^\top \Gamma w(t)\,\mathrm{d}t \\
&\quad + \int_\tau^T w(t)^\top \gamma w(t)\,\mathrm{d}t + X^u(\tau)S^u(0) \\
&\quad + \big(\sigma\sqrt{1-\rho^2} X^u(\tau) + \rho\sigma Z^v(\tau)\big) B_1(\tau)\Bigg).
\end{aligned}
$$

In the next step, we use $Y^{u,v}$ to get its expectation and variance, which together compose the objective function. Assuming

$$
E\big[\mathbf{1}_{\tau\le T}\big(\sigma\sqrt{1-\rho^2} X^u(\tau) + \rho\sigma Z^v(\tau)\big) B_1(\tau)\big] = 0,
$$

we approximate the expectation by

$$
\begin{aligned}
E[Y^{u,v}] &\approx E[b_3^v + \mathbf{1}_{\tau\le T} b_4^{u,v}],\\
b_3^v &:= -\int_0^{\tau\wedge T} \Gamma_2 Z^v(t)v(t)\,\mathrm{d}t + \int_0^{\tau\wedge T} \gamma_2 v(t)^2\,\mathrm{d}t,\\
b_4^{u,v} &:= \int_\tau^T \big(X(0)-X(t)\big)^\top \Gamma w(t)\,\mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t)\,\mathrm{d}t + S^u(0)X^u(\tau).
\end{aligned}
$$

Because the main driver of the variance is the Brownian motion and u, v are deterministic by assumption, we approximate the variance by

$$
\begin{aligned}
\mathrm{Var}(Y^{u,v}) &\approx E[a_3^v + \mathbf{1}_{\tau\le T} a_4^{u,v}],\\
a_3^v &:= \int_0^{\tau\wedge T} \nu^2 Z^v(s)^2\,\mathrm{d}s,\\
a_4^{u,v} &:= \int_\tau^T \big(X(t)^\top \Sigma\Sigma^\top X(t)\big)\,\mathrm{d}s + \big(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau)\big)^2\tau.
\end{aligned}
$$

Finally, our objective function becomes $OF = OF_1 + OF_2$ with

$$
\begin{aligned}
OF_1 =& E[b_3^v + \lambda a_3^v]\\
=& E\Big[-\int_0^{\tau\wedge T} \Gamma_2 Z^v(t)v(t)\,\mathrm{d}t + \int_0^{\tau\wedge T} \gamma_2 v(t)^2\,\mathrm{d}t + \int_0^{\tau\wedge T} \lambda\nu^2 Z^v(s)^2\,\mathrm{d}s\Big],
\end{aligned}
$$

$$
\begin{aligned}
OF_2 =& E[\mathbf{1}_{\tau\le T}(b_4^{u,v} + \lambda a_4^{u,v})]\\
=& E\Big[\mathbf{1}_{\tau\le T}\Big(\int_\tau^T (X(0)-X(t))^\top \Gamma w(t)\,\mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t)\,\mathrm{d}t\\
&+ \int_\tau^T \lambda\big(X(t)^\top \Sigma\Sigma^\top X(t)\big)\,\mathrm{d}s\\
&+ \lambda(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau))^2\tau + S^u(0)X^u(\tau)\Big)\Big].
\end{aligned}
$$

We still analyze the problem $\min OF$ separately on two periods $[0,\tau]$ and $[\tau, T]$.

3.2 Trading from τ to T

Over the period $[\tau, T]$, the result from the last chapter still applies here because we adapted conditional expectation as in the last chapter. In this section, we can use the same approach by

$$\begin{aligned} OF_2 =& E\big[E[\mathbf{1}_{\tau\leq T}(b_4^{u,v} + \lambda a_4^{u,v})|\mathscr{F}_\tau]\big] \\ =& E\big[\mathbf{1}_{\tau\leq T} E[(b_4^{u,v} + \lambda a_4^{u,v})|\mathscr{F}_\tau]\big]. \end{aligned}$$

We just need to $\min E[b_4^{u,v} + \lambda a_4^{u,v}|\mathscr{F}_\tau]$ first over $u(t)$ and $v(t)$ on $[\tau, T]$ for given τ and $Z^v(\tau)$. Therefore, we adapt the same trading strategy:

$$X(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} X(\tau)$$

for $t \in [\tau, T]$, where matrix A can be seen in (2.3). The value of the objective function from τ to T is

$$OF_2 = E\Big[\mathbf{1}_{\tau\leq T}\big(N_1(\tau)Z^v(\tau)^2 + N_2(\tau)Z^v(\tau) + N_3(\tau)\big)\Big],$$

the coefficient, instead of constants N_1, N_2 and N_3, are functions of τ now. Thus, some changes and differences accordingly should be made in the following sections.

3.3 Trading from 0 to τ

On the time interval $[0,\tau]$, for the second stock, the objective function is:

$$
\begin{aligned}
OF_1 =& E[b_3^v + \lambda a_3^v] \\
=& E\Big[\int_0^{\tau\wedge T} \big(\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2\big)\,\mathrm{d}t\Big].
\end{aligned}
$$

For the whole period, we add both objective functions OF_1 and OF_2 together:

$$
\begin{aligned}
OF &= OF_1 + OF_2 \\
&= E\Big[\int_0^{\tau\wedge T} (\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2)\,\mathrm{d}t \\
&\qquad + \mathbf{1}_{\tau\leq T}\Big(N_1(\tau)Z^v(\tau)^2 + N_2(\tau)Z^v(\tau) + N_3(\tau)\Big)\Big].
\end{aligned}
$$

We still need to transform the Mayer form to Lagrange form. However, this time we set $K(\tau, Z^v(\tau)) = N_1(\tau)Z^v(\tau)^2 + N_2(\tau)Z^v(\tau) + N_3(\tau)$, then we can get

$$
\begin{aligned}
K(\tau, Z^v(\tau)) &= K(0, Z^v(0)) + \int_0^{\tau} \frac{\mathrm{d}}{\mathrm{d}t}\big(K(t, Z^v(t)\big)\,\mathrm{d}t \\
&= \int_0^{\tau} \big(N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) + N_2(t)Z^v(t)' \\
&\qquad + 2N_1(t)Z^v(t)Z^v(t)'\big)\,\mathrm{d}t + N_3(0).
\end{aligned}
$$

Our goal is to find a deterministic strategy in the second stock, however, the OF we get here is still stochastic due to the upper limit of the integral. Therefore, we need to make some adjustments. Since $Z^v(t)' = -v(t)$, we can rearrange our objective function:

$$
\begin{aligned}
OF =& OF_1 + OF_2 \\
=& E\Big[\int_0^{\tau\wedge T} (\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2)\,\mathrm{d}t \\
&\quad + \mathbf{1}_{\tau\le T}\int_0^{\tau} N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) \\
&\quad\quad + N_2(t)Z^v(t)' + 2N_1(t)Z^v(t)Z^v(t)'\,\mathrm{d}t + N_3(0)\Big] \\
=& E\Big[\int_0^{T} \mathbf{1}_{\tau\ge t}(\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2) \\
&\quad + \int_0^{T} \mathbf{1}_{\tau\le T}\,\mathbf{1}_{\tau\ge t}\Big(N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) + N_2(t)Z^v(t)' \\
&\quad\quad + 2N_1(t)Z^v(t)Z^v(t)'\Big)\,\mathrm{d}t + \mathbf{1}_{\tau\le T}N_3(0)\Big] \\
=& \int_0^T \Bigg(E[\mathbf{1}_{\{\tau\ge t\}}]E\big[\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2\big] \\
&\quad\quad + E[\mathbf{1}_{\{t\le\tau\le T\}}]E\big[N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) + N_2(t)Z^v(t)' \\
&\quad\quad + 2N_1(t)Z^v(t)Z^v(t)'\big]\Bigg)\,\mathrm{d}t + P[\tau\le T]N_3(0) \\
=& \int_0^T \Bigg(P[\tau\ge t]\big(\lambda\nu^2 Z^v(t)^2 + \Gamma_2 Z^v(t)Z^v(t)' + \gamma_2 Z^v(t)'^2\big) \\
&\quad\quad + P[t\le\tau\le T]\big(N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) + N_2(t)Z^v(t)' \\
&\quad\quad + 2N_1(t)Z^v(t)Z^v(t)'\big)\Bigg)\,\mathrm{d}t + P[\tau\le T]N_3(0).
\end{aligned}
$$

After these preparations, as $P[\tau \le T]$ and $N_3(0)$ are constant, we just need to minimize the integral in the middle, which is a total Lagrange form with deterministic integral limits. Using the Euler Lagrange equation and

$$P(t\le\tau) = 1 - P(\tau\le t) = 1 - \int_0^t f_\tau(s)\,\mathrm{d}s$$

as well as

$$P[t \leq \tau \leq T] = \int_t^T f_\tau(s)\,\mathrm{d}s,$$

we get our target ODE:

$$R_1(t)Z^v(t)'' + R_2(t)Z^v(t)' + R_3(t)Z^v(t) = R_4(t), \tag{3.5}$$

where

$$\begin{aligned}
R_1(t) &= 2\gamma_2(1 - P[\tau \leq t]),\\
R_2(t) &= -2\gamma_2 f_\tau(t),\\
R_3(t) &= -2\lambda\nu^2(1 - P[\tau \leq t]) - f_\tau(t)\big(\Gamma_2 + 2N_1(t)\big),\\
R_4(t) &= f_\tau(t)N_2(t).
\end{aligned}$$

Thus, our optimal trading strategy of the second stock is the solution of the target ODE (3.5). Note that (3.5) is a linear, second-order ODE with continuous coefficient, and thus has a unique solution; compare Chapter 13 of Nagle et al [16]. After time τ, the trading strategy of the first stock is $X^u(t) = A_{11}(t)\Big(1 - \dfrac{\tau}{T}\Big) + A_{12}(t)Z^v(t)$, where $A_{11}(t)$ and $A_{12}(t)$ can be seen in (2.3).

3.4 Numerical solution and explanation

Although there is no explicit result from the ODE, we still can get numerical answers by using MATLAB. The relevant parameters are the same as in Table 2.1 and λ is 10^{-6}. Since as mentioned before, $L \geq \gamma x/T = 0.46$, we select L

over $[0.5, 10]$.

Firstly, we pick some different values of L, which are 1, 3, 6, 10 in our case, and use Matlab with the ode45 function to get numerical solutions of the holding positions of the second stock and draw graphs to make it more clear, which are shown in Figures 3.1 and 3.2:

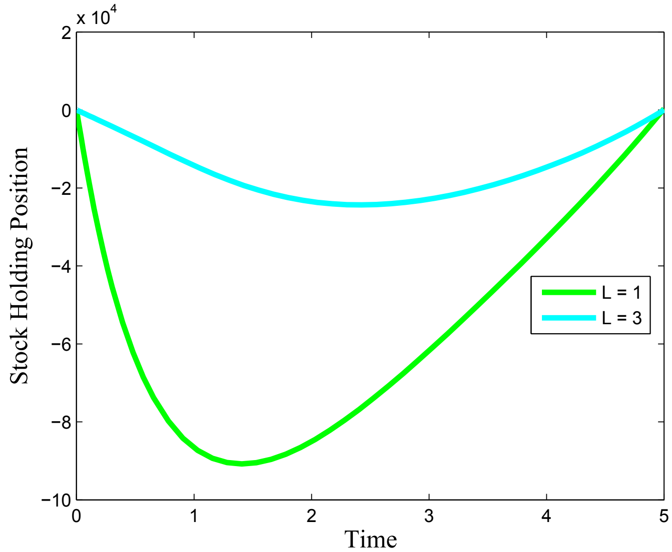

Figure 3.1: Optimal positions for threshold values $L = 1$ and $L = 3$.

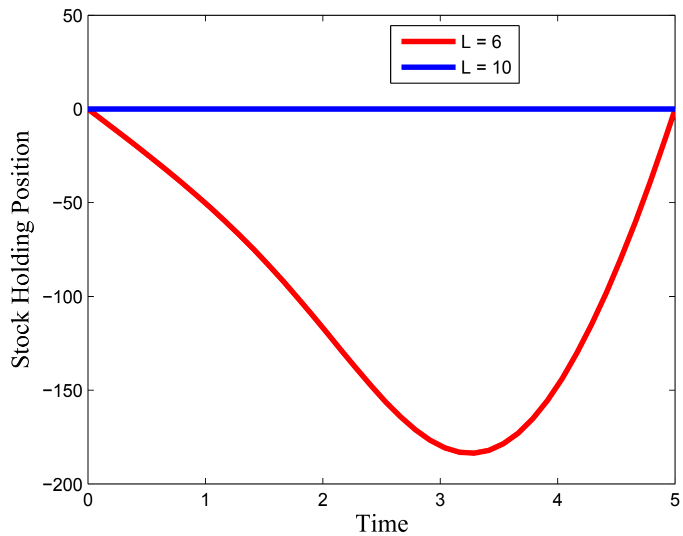

Figure 3.2: Optimal positions for threshold values $L = 6$ and $L = 10$.

There are some comments on the interpretation of Figure 3.2: Firstly, we can get that as L increases, the trading volume of the second stock gets down quickly; especially, when L equals to 10, the volume is almost zero. This is because, when L increases, the initial price becomes further away from the threshold level, which means that the probability $P[\tau < T]$ declines. We should have a smaller short position in the second stock with the decreasing probability. Because if there is little chance that the stock price of the first stock can hit the given level, we do not have to short sell any shares of the second stock to hedge the risk. For example, based on the model I set up and use, the probability $P[\tau < T]$ is low enough with L equal to 10 or T equal to 5, to make the trade volume quite small, almost zero.

Secondly, we notice that as L increases, the largest position is taken at a later moment. It is still because when L becomes bigger, the probability $P[\tau < T]$ declines and $E[\tau]$ increases, which means that we do not need to short sell the second stock early. Late trades still meet our goal to use the second stock and reduce the variance part of the objective function, while they are less costly than building a large short position early.

In addition, in Figures 3.3 and 3.4, I analyze the time and volume of the maximum short position, as functions of L from 0.5 to 10. The graph of time further confirms our second conclusion above, that with bigger values of L, the maximum short position appears later. However, our first conclusion just partially explains the Figure 3.4, where the largest short position of the second stock declines. There is a short period when the largest short position of the second stock actually increases. That is because when the value of L is relatively small, in our case $L < 1$, the time to build the maximum short position is also small, although we should take larger position due to the higher probability $P[\tau < T]$. This means we have to trade the second stock during a shorter time. If we do so, we would have a large temporary market impact and then the expected cost would get much bigger so as to make our objective function increase. The increase in expected cost dominates the decrease in approximate variance. Thus, to avoid this circumstance, we have to reduce the volume of the maximum short position to minimize the objective function. Therefore, the volume of the maximum short position at first goes up and gets down afterwards.

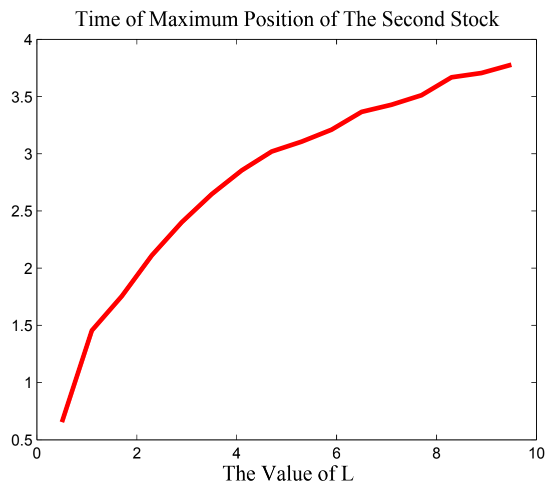

Figure 3.3: The time of the largest position

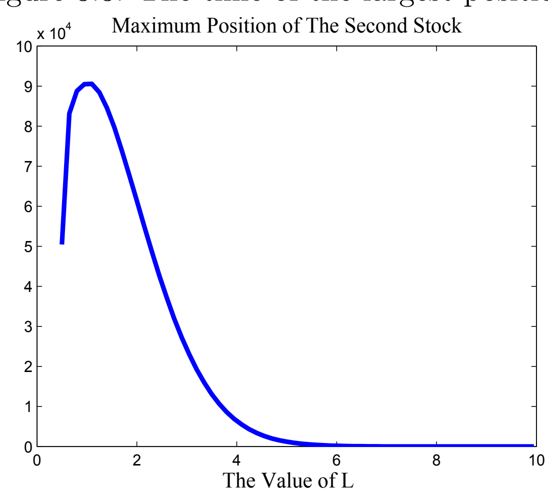

Figure 3.4: The volume of the largest position

Chapter 4

Stochastic transition time with stochastic strategies

In the previous chapter, we assumed that the trading strategy of the second stock is deterministic, which does not need to be the case in reality. As a matter of fact, the movement of the first stock does affect the trading strategies on both stocks. Thus, in this chapter, we would like to find the optimal strategies under the conditions that both transition time τ and trading strategies $u(t)$ and $v(t)$ are stochastic.

4.1 The HJB equation for the Almgren-Chriss problem

The Hamilton-Jacobi-Bellman (HJB) equation is a well-known method to solve optimal control problems. The solution of the HJB equation is the value function which gives the minimum value for a given dynamical system with an associated function. We first illustrate this method on the Almgren-Chriss problem. It is an alternative way to the solution via discrete approximations as in Almgren and Chriss [4] and the calculus of variation approach applied in Almgren [2]. We use the same notation as in Section 1.1 in Almgren [2].

First, we know the objective function is

$$J\big(t, x(t), v\big) = \int_t^T \big(\eta v^2(s) + \lambda\sigma^2 x^2(s)\big)\,\mathrm{d}s$$

from Section 1.2 in Almgren [2]. Then we define the value function as

$$V\big(t, x(t)\big) = \inf_v \int_t^T \big(\eta v^2(s) + \lambda\sigma^2 x^2(s)\big)\,\mathrm{d}s. \tag{4.1}$$

Since $x(t) = -\int_0^t v(s)\,\mathrm{d}s$, with the application of the partial derivative with respect to $V\big(t, x(t)\big)$, the HJB equation is

$$V_t + \min_v\{V_x(-v) + \eta v^2 + \lambda\sigma^2 x^2\} = 0,$$

where we write $V_t = V_t(t, x)$ and $V_x = V_x(t, x)$ for notational simplicity. Writing

$$V_x(-v) + \eta v^2 + \lambda\sigma^2 x^2 = \eta\Big(v - \frac{V_x}{2\eta}\Big)^2 - \frac{V_x^2}{4\eta} + \lambda\sigma^2 x^2,$$

we see that the minimizer is

$$v = \frac{V_x}{2\eta}.$$

Plugging $v = \dfrac{V_x}{2\eta}$ back into the HJB equation, we get a PDE

$$V_t - \frac{V_x^2}{4\eta} + \lambda\sigma^2 x^2 = 0$$

with boundary condition $V(t,0) = 0$.

To find a solution to this PDE, we make the ansatz

$$V\big(t,x\big) = x^2 g(t)$$

because the problem (4.1) is quadratic in the initial value x. We obtain

$$x^2 g'(t) - \frac{g^2(t)x^2}{\eta} + \lambda\sigma^2 x^2 = 0$$

so that

$$g'(t) - \frac{g^2(t)}{\eta} + \lambda\sigma^2 = 0.$$

A solution to this Riccati equation is given by

$$g(t) = \eta\kappa \coth\big(\kappa(T-t)\big),$$

where $\kappa = \dfrac{\lambda\sigma^2}{\eta}$, and therefore a solution of this PDE is

$$V(t,x) = \eta\kappa x^2 \coth\big(\kappa(T-t)\big).$$

Thus, due to $v = \dfrac{V_x}{2\eta}$, the optimal trading strategy becomes

$$v(t) = \kappa x(t) \coth\big(\kappa(T-t)\big).$$

4.2 The HJB equation for our problem

In this section, we consider the HJB equation for our minimization problem. As we know, the cost of trading is

$$\begin{aligned} Y^{u,v} =& Y_1^v + \mathbf{1}_{\tau\le T} Y_2^{u,v} \\ =& \int_0^{\tau\wedge T} P^v(t)v(t)\,\mathrm{d}s + \mathbf{1}_{\tau\le T}\int_\tau^T S^u(s)u(s)\,\mathrm{d}s. \end{aligned}$$

Our problem is to

$$\min E[Y^{u,v}] + \lambda \mathrm{Var}[Y^{u,v}]$$

over $u(t)$ and $v(t)$, subject to $u(t) = x/T$ for $0 \le t < \tau$.

Then we use that the main driver of the variance is the Brownian motion as in the approximation of Section 1.1 in Almgren [2] and obtain the objective function in the form of

$$\begin{aligned} & J\big(t, B_1(t), B_2(t), X^u(t), Z^v(t); u(t), v(t)\big) \\ & = E\bigg[\int_t^T L\big(s, B_1(s), B_2(s), X^u(s), Z^v(s); u(s), v(s)\big)\,\mathrm{d}s\bigg|\mathscr{F}_t\bigg], \end{aligned}$$

where

$$\begin{aligned} & L\big(s, B_1(s), B_2(s), X^u(s), Z^v(s); u(s), v(s)\big) \\ & = \mathbf{1}_{\tau\ge s}\big(\lambda\nu^2 Z^v(s)^2 - \Gamma_2 Z^v(s)v(s) + \gamma_2 v(s)^2\big) + \mathbf{1}_{\tau\ge s} E[\mathbf{1}_{\tau\le T}|\mathscr{F}_s] \\ & \quad \times \big(N_1'(s)Z^v(s)^2 + N_2'(s)Z^v(s) + N_3'(s) - N_2(s)v(s) - 2N_1(s)Z^v(s)v(s)\big). \end{aligned}$$

We define the value function as

$$V\big(t, B_1(t), B_2(t), X^u(t), Z^v(t)\big) = \inf_{(u,v)\in\mathscr{A}} J\big(t, B_1(t), B_2(t), X^u(t), Z^v(t); u(t), v(t)\big)$$

with $\mathscr{A} \triangleq$ the set of all such pairs of functions $\big(u(.), v(.)\big)$. Then our HJB equation is

$$\begin{aligned} & V_t + \frac{1}{2}V_{x_1^2} + \frac{1}{2}V_{x_2^2} + \rho V_{x_1,x_2} \\ & + \inf_{(u,v)}\{V_{x_3}(-u) + V_{x_4}(-v) + L\big(t, B_1(t), B_2(t), X^u(t), Z^v(t); u(t), v(t)\big)\} = 0. \end{aligned}$$

Next, we take derivatives with respect to u and v to obtain

$$\begin{aligned} L_u\big(t, B_1(t), B_2(t), X^u(t), Z^v(t); u(t), v(t)\big)\} &= V_{x_3} \\ L_v\big(t, B_1(t), B_2(t), X^u(t), Z^v(t); u(t), v(t)\big)\} &= V_{x_4} \end{aligned}$$

with the boundary condition that $V\big(\tau, B_1(\tau), B_2(\tau), X^u(\tau), Z^v(\tau)\big)$ takes the same value as the solution in the Almgren-Chriss problem with two stocks as in Almgren [2].

Theoretically, we may determine the optimal strategy (u, v) in terms of V_{x_3} and V_{x_4} and get a PDE for V. After solving this PDE, we could get the value function and the optimal strategies. However, this problem appears to be very complicated. Hence, we study in the next section an approximation via a binomial model.

4.3 Approximation using a binomial model

Since it is too difficult to solve our problem directly by the HJB equation, we choose a binomial model to approximate the prices of the first stock during the whole time.

4.3.1 Trading from τ to T

Actually, the objective functions we use here are the same as those in Chapter 3, because we have already assumed there that the transition time τ is stochastic. In addition, because of a similar reason as we mentioned in Chapter 3, the trading strategy from τ to T will not be affected, either. Thus, we will use the conclusions from Section 3.2.

Therefore, we still have our objective functions $OF = OF_1 + OF_2$ with

$$\begin{aligned}
OF_1 &= E[b_3^v + \lambda a_3^v] \\
&= E\Big[- \int_0^{\tau\wedge T} \Gamma_2 Z^v(t) v(t)\,\mathrm{d}t + \int_0^{\tau\wedge T} \gamma_2 v(t)^2\,\mathrm{d}t + \int_0^{\tau\wedge T} \lambda\nu^2 Z^v(t)^2\,\mathrm{d}t \Big], \\
OF_2 &= E[\mathbf{1}_{\tau\le T}(b_4^{u,v} + \lambda a_4^{u,v})] \\
&= E\Big[\mathbf{1}_{\tau\le T}\Big(\int_\tau^T (X(0) - X(t))^\top \Gamma w(t)\,\mathrm{d}t + \int_\tau^T w(t)^\top \gamma w(t)\,\mathrm{d}t \\
&\qquad + \int_\tau^T \lambda\big(X(t)^\top \Sigma\Sigma^\top X(t)\big)\,\mathrm{d}t \\
&\qquad + \lambda(\sigma\sqrt{1-\rho^2}X^u(\tau) + \rho\sigma Z^v(\tau))^2\tau + S^u(0)X^u(\tau)\Big)\Big].
\end{aligned}$$

Moreover, the formula for the density function of τ still holds too,

$$f_\tau(t) = \frac{L - \gamma_1 x/T}{\sigma\sqrt{2\pi t^3}} \exp\Big(- \frac{(L - \gamma_1 x/T + \Gamma_1 xt/T)^2}{2\sigma^2 t} \Big).$$

Therefore, we adapt the same trading strategy

$$X(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} X(\tau),$$

for $t \in [\tau, T]$, where the matrix A can be seen in (2.3).

The value of the objective function from τ to T is

$$OF_2 = E\Big[\mathbf{1}_{\tau \le T}\big(N_1(\tau)Z^v(\tau)^2 + N_2(\tau)Z^v(\tau) + N_3(\tau)\big)\Big].$$

Then we still can get

$$\begin{aligned} OF =& OF_1 + OF_2 \\ =& E\Big[\int_0^T \mathbf{1}_{\tau \ge t}\Big((\lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2) \\ &+ \mathbf{1}_{\tau \le T}\big(N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) + N_2(t)Z^v(t)' \\ &+ 2N_1(t)Z^v(t)Z^v(t)'\big)\Big)\,\mathrm{d}t\Big] + N_3(0). \end{aligned}$$

4.3.2 Trading from 0 to τ

Things go differently from here. Because the trading strategy is stochastic now, the objective function OF has to be rearranged in another way.

To make it look simpler, since $Z^v(t)' = -v(t)$, we set

$$L_1(v(t), Z^v(t)) = \lambda\nu^2 Z^v(t)^2 - \Gamma_2 Z^v(t)v(t) + \gamma_2 v(t)^2$$

and

$$L_2(t, v(t), Z^v(t)) = N_1'(t)Z^v(t)^2 + N_2'(t)Z^v(t) + N_3'(t) - N_2(t)v(t) \\ - 2N_1(t)Z^v(t)v(t).$$

The objective function becomes

$$OF = E\bigg[\int_0^T \mathbf{1}_{\tau\geq t}\big(L_1(v(t), Z^v(t)) + \mathbf{1}_{\tau\leq T}L_2(t, v(t), Z^v(t))\big)\,\mathrm{d}t\bigg] + P[\tau \leq T]N_3(0).$$

Since $P[\tau \leq T]$ and $N_3(0)$ are constants, our problem is

$$\min I_2 = E\bigg[\int_0^T \mathbf{1}_{\tau\geq t}\Big(L_1(v(t), Z^v(t)) + \mathbf{1}_{\tau\leq T}L_2(t, v(t), Z^v(t))\Big)\,\mathrm{d}t\bigg]. \tag{4.2}$$

We adopt a binomial model this time, so we are going to determine the trading strategy of the second stock afterwards based on each simulated situation on every node. Thus, we have to take conditional expectations with respect to $\mathscr{F}_t$ and obtain

$$\begin{aligned} I_2 &= E\bigg[\int_0^T \mathbf{1}_{\tau\geq t}\Big(L_1(v(t), Z^v(t)) + \mathbf{1}_{\tau\leq T}L_2(t, v(t), Z^v(t))\Big)\,\mathrm{d}t\bigg] \\ &= E\bigg[\int_0^T E\Big[\mathbf{1}_{\tau\geq t}\Big(L_1(v(t), Z^v(t)) + \mathbf{1}_{\tau\leq T}L_2(t, v(t), Z^v(t))\Big)\Big|\mathscr{F}_t\Big]\,\mathrm{d}t\bigg] \\ &= E\bigg[\int_0^T \mathbf{1}_{\tau\geq t}\Big(L_1(v(t), Z^v(t)) + P\big[\tau \leq T|\mathscr{F}_t\big]L_2(t, v(t), Z^v(t))\Big)\,\mathrm{d}t\bigg]. \end{aligned} \tag{4.3}$$

4.3.3 The general idea for the simulation

First, we use a binomial tree to simulate the movement of the Brownian motion B_1, then incorporate the permanent market impact to the first stock. Therefore, we get a tree for the price of the first stock during the whole trading time, from 0 to T, so that the transition time τ can be observed. The jump between each time point is $\sigma\sqrt{T/N}$, where σ is the coefficient of B_1 in (1.1) and N is the number of steps in the binomial tree. When the trade is over, the temporary market impact disappears almost instantaneously and does not change the stock price. We have to consider a permanent market impact because it changes the stock prices permanently. As we assume that before time τ, the buy orders are done at a constant rate (TWAP order) so that $u(t) = x/T$ for $0 \leq t \leq \tau$. Thus, we add $\Gamma_1 \dfrac{xt}{T}$ to each node of the price tree.

Second, we assume for each node in the tree, we have three choices for the trading rate. Based on the deterministic strategy we got from the previous chapter, the three options are $v(t)$, $v(t) + std$ and $v(t) - std$, where $v(t)$ is the average of the trading speed at the same time point and that at the next time point in the deterministic strategy for $t = 0, T/N, 2T/N, \ldots, T(N-1)/N$, and std is the standard deviation of the averages $v(t)$, for $t = 0, T/N, 2T/N, \ldots, T(N-1)/N$.

Third, we calculate the conditional expectations I_2 at each time point over different combinations of $v(t)$, for $t = 0, T/N, 2T/N, \ldots, T(N-1)/N$.

Finally, based the conditional expectations, we choose the optimal trading strategies that make the objective function minimal.

4.3.4 Example

Let us take a four-step binomial model for instance, which means $N = 4$. The relevant parameters are the same as in Table 2.1 and λ is 10^{-6}.

We simulate a binomial tree to represent the price movement without market impacts. Every node in the tree has the same probability for up or down and the jump is $\sigma\sqrt{T/N} = 1.0621$. Next, we incorporate the permanent market impacts by adding $\Gamma_1 xj/N = 0.0625t$ to the according nodes, for $j = 0, 1, 2, 3, 4$. Besides, we also use the price tree to calculate the indicator function and conditional probability in (4.3) preparing for the integral calculation later. Secondly, I pick up the according $v(t)$ from the deterministic strategy we got in the previous chapter, for $t = 0, 1.25, 2.5, 3.75, 5$, which are -3.8716, -3.5261, 0.1528, 2.6937, 5.3711. After taking the average of the adjacent two $v(t)$, a sample standard deviation is obtained as 3.4064. Thus, the choice matrix for $v(t)$ is

$$\begin{bmatrix} -0.2924 & -3.6989 & -7.1053 \\ 1.7198 & -1.6867 & -5.0931 \\ 4.8297 & 1.4232 & -1.9832 \\ 7.4388 & 4.0324 & 0.6260 \end{bmatrix}.$$

Finally, we just need to calculate the conditional expectation for each possibility of $v(t)$ combination and select the optimal strategy backward. The related codes can be seen in Appendix A.2.3.

Chapter 5

Conclusion

In this thesis, we discussed backstopped deals from the broker's perspective under different assumptions. Differently from the Almgren-Chriss problem, we introduced a second stock to hedge the risk of the trading in order to minimize the objective function.

In the first chapter, we introduced the definition of backstopped deals, elaborated on the motivation and then formulated the mathematical problem. The target problem has been presented, which is used to find the objective function in the following chapters.

The second chapter described the situation when the assumption is that the transition time τ and the optimal strategy are both deterministic. We divided into two cases, when $\tau \geq T$ and when $\tau < T$. For $\tau \geq T$, the broker need not do anything because the client bears all the risk. For $\tau < T$, we obtained an explicit solution by using the Euler Lagrange equation and tested the solution with different parameters. In addition, we compared our result with the situation when there was no second stock and explained the resulting differences.

The third chapter presented the further study under the assumption that, while optimal strategy was still deterministic, the transition time τ was stochastic, which is a situation that also occurs in practice. We used the conclusion from Chapter 2 for the first stock during the time interval $[\tau, T]$. For the second stock, we got an ODE with no explicit solution. Thus, we exploited the numerical answers with various hitting price levels, also made comparisons among them and explained the outcomes.

Finally, we assumed a stochastic transition time with stochastic strategies in the fourth chapter. First, we set up the HJB equation for the Almgren-Chriss problem to show the viability. However, after getting the HJB equation for our problem, we found it too complicated to deal with. Therefore, we chose an alternative, an approximation using a binomial model, to solve our problem. Next, we stated the fact that the trading strategy from τ to T will not be affected and then the general idea about how to implement the model to find an approximate solution. Lastly, a simple example was shown to further explain the binomial model with four steps.

Appendix A

MATLAB code

A.1 To Chapter 3: stochastic transition time with deterministic strategies

A.1.1 $N_1(t)$, $N_2(t)$ and $N_3(t)$

```
function a = N1(t,LN)
%The N1 function returns the coefficient of Z^v(\tau)^2 in ...
    the objective function 2 in Section 3.2. t is the time ...
    from the beginning and LN is the excess price by which ...
    the hit level is above the orinigal price S(0). The ...
    other parameters used already are included in the ...
    m-file, which are from Table 2.1.
S0=50;
P0=30;
L =LN;
sigma = 0.95;
```

```
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
per= [ Gamma 0 ; 0 Xi];
tem= [ gamma 0 ; 0 xi];
Sigma= [ sigma*sqrt(1 - rho^2) rho*sigma ; 0 nu];
duichen = sqrt(tem)^-1 * Sigma * Sigma.' * sqrt(tem)^-1 ;
[Q,D] = eig(duichen);
kappa =sqrt(lambda * D);
theta = sqrt(lambda*nu^2/xi);
N = 10^3;
s = t : (T-t)/N:T;
Δ = (T-t)/N;
%%
l1= length(t);
l2 =length(s);
%%
function b = n1(x1,y1)
H1  = sinh(kappa(1,1)*(T-y1))/sinh(kappa(1,1)*(T-x1));
H2  = sinh(kappa(2,2)*(T-y1))/sinh(kappa(2,2)*(T-x1));
H3  = cosh(kappa(1,1)*(T-y1))/sinh(kappa(1,1)*(T-x1));
H4  = cosh(kappa(2,2)*(T-y1))/sinh(kappa(2,2)*(T-x1));
H = [H1 0;0 H2];
Hu = [H3 0;0 H4];
A = sqrt(tem)^-1*Q*H*Q.'*sqrt(tem);
```

```
B = sqrt(tem)^-1*Q*kappa*Hu*Q.'*sqrt(tem);
M = A.'*Sigma*Sigma.'*A;
b = lambda*M(2,2)-Gamma*A(1,2)*B(1,2) - Xi*A(2,2)*B(2,2)...
    +B(1,2)^2*gamma+B(2,2)^2*xi;
end
a = zeros(l1,1);
for i = 1 : l1
    for j = 2 : l2
    a(i) = a(i) + n1(t(i),s(j))*Δ;
    end
    a(i) = a(i)+lambda*sigma^2*rho^2*t;
end
end
```

```
function a = N2(t,LN)
%The N2 function returns the coefficient of Z^v(\tau) in ...
    the objective function 2 in Section 3.2. t is the time ...
    from the beginning and LN is the excess price by which ...
    the hit level is above the orinigal price S(0). The ...
    other parameters used already are included in the ...
    m-file, which are from Table 2.1.
S0=50;
P0=30;
L = LN;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
```

```
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
per= [ Gamma 0 ; 0 Xi];
tem= [ gamma 0 ; 0 xi];
Sigma= [ sigma*sqrt(1 - rho^2) rho*sigma ; 0 nu];
duichen = sqrt(tem)^-1 * Sigma * Sigma.' * sqrt(tem)^-1 ;
[Q,D] = eig(duichen);
kappa =sqrt(lambda * D);
theta = sqrt(lambda*nu^2/xi);
N = 10^3;
s = t : (T-t)/N:T;
Δ = (T-t)/N;
%%
l1= length(t);
l2 =length(s);
%%
function b = n2(x1,y1)
H1  = sinh(kappa(1,1)*(T-y1))/sinh(kappa(1,1)*(T-x1));
H2  = sinh(kappa(2,2)*(T-y1))/sinh(kappa(2,2)*(T-x1));
H3  = cosh(kappa(1,1)*(T-y1))/sinh(kappa(1,1)*(T-x1));
H4  = cosh(kappa(2,2)*(T-y1))/sinh(kappa(2,2)*(T-x1));
H = [H1 0;0 H2];
Hu = [H3 0;0 H4];
A = sqrt(tem)^-1*Q*H*Q.'*sqrt(tem);
B = sqrt(tem)^-1*Q*kappa*Hu*Q.'*sqrt(tem);
M = A.'*Sigma*Sigma.'*A;
b = lambda*(M(1,2)+M(2,1))*x*(1-x1/T)...
    - Gamma*A(1,1)*B(1,2)*x*(1-x1/T) - ...
```

```
        Gamma*B(1,1)*A(1,2)*x*(1-x1/T)...
    - Xi*A(2,1)*B(2,2)*x*(1-x1/T) - ...
        Xi*B(2,1)*A(2,2)*x*(1-x1/T)...
    +2*B(1,1)*B(1,2)*gamma*x*(1-x1/T)...
    +2*B(2,2)*B(2,1)*xi*x*(1-x1/T)+B(1,2)*x*Gamma;
end
a = zeros(l1,1);
for i = 1 : l1
    for j = 2 : l2
    a(i) = a(i) + n2(t(i),s(j))*∆;
    end
    a(i) = a(i)+2*lambda*x*(1-t/T)*sigma^2*...
        rho*t*sqrt(1-rho^2);
end
end
```

```
function a = N3(t,LN)
%The N3 function returns the coeficient of the constant ...
    term (without  Z^v(\tau)) in the objective function 2 in ...
    Section 3.2. t is the time from the beginning and LN is ...
    the excess price by which the hit level is above the ...
    orinigal price S(0). The other parameters used already ...
    are included in the m-file, which are from Table 2.1.
S0=50;
P0=30;
L = LN;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
```

```
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
per= [ Gamma 0 ; 0 Xi];
tem= [ gamma 0 ; 0 xi];
Sigma= [ sigma*sqrt(1 - rho^2) rho*sigma ; 0 nu];
duichen = sqrt(tem)^-1 * Sigma * Sigma.' * sqrt(tem)^-1 ;
[Q,D] = eig(duichen);
kappa =sqrt(lambda * D);
theta = sqrt(lambda*nu^2/xi);
N = 10^3;
s = t : (T-t)/N:T;
Δ = (T-t)/N;
%%
l1= length(t);
l2 =length(s);
%%
function b = n3(x1,y1)
H1  = sinh(kappa(1,1)*(T-y1))/sinh(kappa(1,1)*(T-x1));
H2  = sinh(kappa(2,2)*(T-y1))/sinh(kappa(2,2)*(T-x1));
H3  = cosh(kappa(1,1)*(T-y1))/sinh(kappa(1,1)*(T-x1));
H4  = cosh(kappa(2,2)*(T-y1))/sinh(kappa(2,2)*(T-x1));
H = [H1 0;0 H2];
Hu = [H3 0;0 H4];
A = sqrt(tem)^-1*Q*H*Q.'*sqrt(tem);
B = sqrt(tem)^-1*Q*kappa*Hu*Q.'*sqrt(tem);
M = A.'*Sigma*Sigma.'*A;
```

```
b = lambda*M(1,1)*x^2*(1-x1/T)^2 ...
    - Gamma*A(1,1)*B(1,1)*x^2*(1-x1/T)^2 - ...
       Xi*A(2,1)*B(2,1)*x^2*(1-x1/T)^2 ...
    +(B(1,1)^2*gamma+B(2,1)^2*xi)*x^2*(1-x1/T)^2 ...
    +Gamma*B(1,1)*x^2*(1-x1/T);
end
a = zeros(l1,1);
for i = 1 : l1
    for j = 2 : l2
    a(i) = a(i) + n3(t(i),s(j))*Δ;
    end
    a(i) = a(i)+lambda*x^2*(1-t/T)^2*Etau*sigma^2*(1-rho^2) ...
       + S0*x*(1-t/T);
end
end
```

A.1.2 Density function and cumulative distribution function of τ

```
function a = ftau(t,LN)
%The ftau returns the value of the density of \tau, based ...
   on the density function proved by Lemma 3.1 in Chapter ...
   3. t is the time from the beginning and LN is the excess ...
   price by which the hit level is above the orinigal price ...
   S(0). The other parameters used already are included in ...
   the m-file, which are from Table 2.1. In the case t=0, I ...
   specify that the density is zero
S0=50;
```

```
P0=30;
L = LN ;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
l = length(t);
a= zeros(l,1);
for i = 1:l
    if t(i) == 0
        a(i) = 0;
    else
        a(i) = (L - gamma * x/T)/(sigma*sqrt(2*pi*t(i)^3))* ...
            exp(-(L-gamma*x/T - Gamma * x * ...
                t(i)/T)^2/(2*sigma^2*t(i)));
    end
end
```

```
function a = Ptau(t,LN)
%The Ptau returns the value of the probability of \tau ...
    smaller than t. t is the time from the beginning and LN ...
    is the excess price by which the hit level is above the ...
    orinigal price S(0). Riemann sum ia used by splitting ...
```

```
    the interval into 1000 parts.
N = 10^3;
b = 0:t/N:t;
a = sum(ftau(b,LN))*t/N;
```

A.1.3 Target ODE and the function needed to find $Z^v(0)'$

```
function xprime = ODE2(t,Z,LN)
%The function ODE2 returns a first-order system, which ...
    consist of two equations. Taking Z1(t) = Z(t) and Z2(t) ...
    = Z'(t), the first one is  Z'1(t) =Z2(t). The second one ...
    is the target second-order ODE. t is the time from the ...
    beginning, Z is the objective function and LN is the ...
    excess price by which the hit level is above the ...
    orinigal price S(0). The other parameters used already ...
    are included in the m-file, which are from Table 2.1.
S0=50;
P0=30;
L = LN;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
```

```
R1 = 2*xi*(1 - Ptau(t,LN));
R2=  - 2*xi*ftau(t,LN);
R3  = - (2*lambda*nu^2*(1 - Ptau(t,LN))+ ftau(t,LN)*(Xi + ...
    2*N1(t,LN)));
R4 =  ftau(t,LN)*N2(t,LN);
xprime = [Z(2);(-R2*Z(2)- R3*Z(1)+R4)/R1];
```

```
function Z0 = Zprime0(h,l,LN)
%The Zprime0 returns the value of Z^v(0)' that makes the ...
    Z^v(T) close to zero (within some pre-specified error ...
    range, which is 10^-4 in this case). h is the value of ...
    Z^v(0)' that makes the Z^v(T) positive outside the ...
    range, l is the value of Z^v that makes the Z^v(T) ...
    negative outside the range, LN is the excess price by ...
    which the hit level is above the orinigal price S(0).
tspan = [0,5];
Z01 = [0;h];
Z02 = [0;l];
Z03 = (Z01+Z02)/2;
[t,Z] = ode45(@(t,Z)ODE2(t,Z,LN),tspan,Z03);
if abs(Z(end,1)) < 10^-4
     Z0 = Z03;
elseif Z(end,1) >  10^-4
     Z0 = Zprime0(Z03(2),l,LN);
else
     Z0 = Zprime0(h,Z03(2),LN);
end
```

A.2 To Chapter 4: stochastic transition time with stochastic strategies

A.2.1 $N_1(t)'$, $N_2(t)'$ and $N_3(t)'$

```
function a = N1prime(t,LN)
%The N1prime function returns the value of partial ...
    derivative of function N1 with respect to t. t is the ...
    time from the beginning and LN is the excess price by ...
    which the hit level is above the orinigal price S(0). ...
    The other parameters used already are included in the ...
    m-file, which are from Table 2.1.
S0=50;
P0=30;
L =LN;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
per= [ Gamma 0 ; 0 Xi];
tem= [ gamma 0 ; 0 xi];
Sigma= [ sigma*sqrt(1 - rho^2) rho*sigma ; 0 nu];
duichen = sqrt(tem)^-1 * Sigma * Sigma.' * sqrt(tem)^-1 ;
```

```
[Q,D] = eig(duichen);
kappa =sqrt(lambda * D);
theta = sqrt(lambda*nu^2/xi);
N = 10^3;
s = t : (T-t)/N:T;
Δ = (T-t)/N;
%%
l1= length(t);
l2 =length(s);
%%
syms x2 y2
H1  = sinh(kappa(1,1)*(T-y2))/sinh(kappa(1,1)*(T-x2));
H2  = sinh(kappa(2,2)*(T-y2))/sinh(kappa(2,2)*(T-x2));
H3  = cosh(kappa(1,1)*(T-y2))/sinh(kappa(1,1)*(T-x2));
H4  = cosh(kappa(2,2)*(T-y2))/sinh(kappa(2,2)*(T-x2));
H = [H1 0;0 H2];
Hu = [H3 0;0 H4];
A = sqrt(tem)^-1*Q*H*Q.'*sqrt(tem);
B = sqrt(tem)^-1*Q*kappa*Hu*Q.'*sqrt(tem);
M = A.'*Sigma*Sigma.'*A;
b = lambda*M(2,2)-Gamma*A(1,2)*B(1,2) - Xi*A(2,2)*B(2,2)...
    +B(1,2)^2*gamma+B(2,2)^2*xi;
b2 =diff(b,x2);
%%
a = zeros(l1,1);
for i = 1 : l1
    b3 = subs(b2,x2,t(i));
    for j = 2 : l2
    a(i) = a(i) + subs(b3,y2,s(j))*Δ;
    end
```

```
    a(i) = ...
        a(i)−subs(b,[x2,y2],[t(i),t(i)])+lambda*sigma^2*rho^2;
end
end
```

```
function a = N2prime(t,LN)
%The N2prime function returns the value of partial ...
    derivative of function N2 with respect to t. t is the ...
    time from the beginning and LN is the excess price by ...
    which the hit level is above the orinigal price S(0). ...
    The other parameters used already are included in the ...
    m−file, which are from Table 2.1.
S0=50;
P0=30;
L = LN;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^−7;
gamma = 2.3*10^−6;
Xi = 2.8 * 10^−7;
xi = 2*10^−6;
T = 5;
lambda= 10^−6;
x= 10^6;
per= [ Gamma 0 ; 0 Xi];
tem= [ gamma 0 ; 0 xi];
Sigma= [ sigma*sqrt(1 − rho^2) rho*sigma ; 0 nu];
duichen = sqrt(tem)^−1 * Sigma * Sigma.' * sqrt(tem)^−1 ;
```

```
[Q,D] = eig(duichen);
kappa =sqrt(lambda * D);
theta = sqrt(lambda*nu^2/xi);
N = 10^3;
s = t : (T-t)/N:T;
Δ = (T-t)/N;
%%
l1= length(t);
l2 =length(s);
%%
syms x2 y2
H1  = sinh(kappa(1,1)*(T-y2))/sinh(kappa(1,1)*(T-x2));
H2  = sinh(kappa(2,2)*(T-y2))/sinh(kappa(2,2)*(T-x2));
H3  = cosh(kappa(1,1)*(T-y2))/sinh(kappa(1,1)*(T-x2));
H4  = cosh(kappa(2,2)*(T-y2))/sinh(kappa(2,2)*(T-x2));
H = [H1 0;0 H2];
Hu = [H3 0;0 H4];
A = sqrt(tem)^-1*Q*H*Q.'*sqrt(tem);
B = sqrt(tem)^-1*Q*kappa*Hu*Q.'*sqrt(tem);
M = A.'*Sigma*Sigma.'*A;
b = lambda*(M(1,2)+M(2,1))*x*(1-x2/T)...
    - Gamma*A(1,1)*B(1,2)*x*(1-x2/T) - ...
       Gamma*B(1,1)*A(1,2)*x*(1-x2/T)...
    - Xi*A(2,1)*B(2,2)*x*(1-x2/T) - ...
       Xi*B(2,1)*A(2,2)*x*(1-x2/T)...
    +2*B(1,1)*B(1,2)*gamma*x*(1-x1/T)...
    +2*B(2,2)*B(2,1)*xi*x*(1-x2/T)+B(1,2)*x*Gamma;
b2 = diff(b,x2);
a = zeros(l1,1);
for i = 1 : l1
    b3 = subs(b2,x2,t(i));
```

```
    for j = 2 : l2
    a(i) = a(i) + subs(b3,y2,s(j))*Δ;
    end
    a(i) = a(i)-subs(b,[x2,y2],[t(i),t(i)])...
        +2*lambda*x*(1-t/T)*sigma^2*rho*sqrt(1-rho^2);
end
end
```

```
function a = N3prime(t,LN)
%The N3prime function returns the value of partial ...
    derivative of function N3 with respect to t. t is the ...
    time from the beginning and LN is the excess price by ...
    which the hit level is above the orinigal price S(0). ...
    The other parameters used already are included in the ...
    m-file, which are from Table 2.1.
S0=50;
P0=30;
L = LN;
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma = 2.5 * 10^-7;
gamma = 2.3*10^-6;
Xi = 2.8 * 10^-7;
xi = 2*10^-6;
T = 5;
lambda= 10^-6;
x= 10^6;
per= [ Gamma 0 ; 0 Xi];
```

```
tem= [ gamma 0 ; 0 xi];
Sigma= [ sigma*sqrt(1 - rho^2) rho*sigma ; 0 nu];
duichen = sqrt(tem)^-1 * Sigma * Sigma.' * sqrt(tem)^-1 ;
[Q,D] = eig(duichen);
kappa =sqrt(lambda * D);
theta = sqrt(lambda*nu^2/xi);
N = 10^3;
s = t : (T-t)/N:T;
Δ = (T-t)/N;
%%
l1= length(t);
l2 =length(s);
%%
syms x2 y2
H1  = sinh(kappa(1,1)*(T-y2))/sinh(kappa(1,1)*(T-x2));
H2  = sinh(kappa(2,2)*(T-y2))/sinh(kappa(2,2)*(T-x2));
H3  = cosh(kappa(1,1)*(T-y2))/sinh(kappa(1,1)*(T-x2));
H4  = cosh(kappa(2,2)*(T-y2))/sinh(kappa(2,2)*(T-x2));
H = [H1 0;0 H2];
Hu = [H3 0;0 H4];
A = sqrt(tem)^-1*Q*H*Q.'*sqrt(tem);
B = sqrt(tem)^-1*Q*kappa*Hu*Q.'*sqrt(tem);
M = A.'*Sigma*Sigma.'*A;
b = lambda*M(1,1)*x^2*(1-x2/T)^2 ...
    - Gamma*A(1,1)*B(1,1)*x^2*(1-x2/T)^2 - ...
       Xi*A(2,1)*B(2,1)*x^2*(1-x2/T)^2 ...
    +(B(1,1)^2*gamma+B(2,1)^2*xi)*x^2*(1-x2/T)^2 ...
    +Gamma*B(1,1)*x^2*(1-x2/T);
b2 = diff(b,x2);
a = zeros(l1,1);
for i = 1 : l1
```

```
    b3 = subs(b2,x2,t(i));
    for j = 2 : l2
    a(i) = a(i) + subs(b3,y2,s(j))*Δ;
    end
    a(i) = a(i)-subs(b,[x2,y2],[t(i),t(i)])...
        +lambda*x^2*(1-t/T)^2*sigma^2*(1-rho^2)...
        + lambda*x^2*(2*t/T^2 - 2/T)*t*sigma^2*(1-rho^2)- ...
            (S0*x)/T...

end
end
```

A.2.2 $L_1(v(t), Z^v(t))$ and $L_2(t, v(t), Z^v(t))$

```
function a = L1(vt,zt)
%The L1 fuction returns the value of L1 function in ...
    equation (4.2) with respect to vt and zt. vt is the ...
    value of the trading rate at time t and zt is the ...
    position of the second stock at time t.
sigma = 0.95;
nu = 0.83;
rho = 0.78;
Gamma1 = 2.5 * 10^-7;
gamma1 = 2.3*10^-6;
Gamma2 = 2.8 * 10^-7;
gamma2 = 2*10^-6;
lambda= 10^-6;
a = lambda * nu^2 *zt.^2 - Gamma2*zt.*vt + gamma2*vt.^2;
```

```
function a = L2(N,vt,zt,LN,T)
%The L2 fuction returns the value of L2 function in ...
    equation (4.2) with respect to N,vt, zt, LN and T. N is ...
    the number of steps of the binomail model. vt is the ...
    value of the trading rate at time t. zt is the position ...
    of the second stock at time t.LN is the excess price by ...
    which the hit level is above the orinigal price S(0) and ...
    T is the ternial time.
L=LN;
N1p = zeros(N,1);
N2p = zeros(N,1);
N3p = zeros(N,1);
N1m = zeros(N,1);
N2m = zeros(N,1);
for c = 0:N-1
    N1p(c+1) = N1prime(c/N*T,L);
    N2p(c+1) = N2prime(c/N*T,L);
    N3p(c+1) = N3prime(c/N*T,L);
    N1m(c+1) = N1(c/N*T,L);
    N2m(c+1) = N2(c/N*T,L);
end
N1p = repmat(N1p,1,3^(N-1));
N2p = repmat(N2p,1,3^(N-1));
N3p = repmat(N3p,1,3^(N-1));
N1m = repmat(N1m,1,3^(N-1));
N2m = repmat(N2m,1,3^(N-1));
a = N1p.* zt.^2 + N2p.*zt + N3p - N2m.*vt - 2*N1m.*zt.*vt;
```

A.2.3 Generating the binomial tree and selecting the optimal strategy

```
N = 4; % The binomial model has N steps
T = 5;  % The terminal time
L = 3; % The excess prive above S(0)
x= 10^6; % Total trading volume
sigma = 0.95; % The coefficient of B_1(t)
nu = 0.83; % The coefficient of B_2(t)
rho = 0.78; % The correlation of B_1(t) and B_2(t)
Gamma1 = 2.5 * 10^-7;  % Permenant impact coefficient of ...
    the first stock
gamma1 = 2.3*10^-6; % Temporary impact coefficient of the ...
    first stock
Gamma2 = 2.8 * 10^-7; % Permenant  impact coefficient of ...
    the second stock
gamma2 = 2*10^-6; % Temporary impact coefficient of the ...
    second stock
lambda= 10^-6;
Δ = sigma * sqrt(T/N);
tree= zeros(N+1,2^N); % 2^N paths
for i = 1:N
     a = 2^N / 2^i;
     b = 0:a:2^N;
     l= length(b);
    for j =2:2:l
        tree(i+1,b(j-1)+1:b(j)) = Δ;
        tree(i+1,b(j)+1:b(j+1))= -Δ;
    end
```

```
end
tree2 = cumsum(tree); %Cumulative sum of tree
tree3= zeros(N+1,2^N); % Tree with permenant market impact
for j = 1:N
    tree3(j+1,:)= tree2(j+1,:)+ Gamma1 * x *j/N;
end
treev = tree(2:end,:);
tree4= tree(1:end-1,:);
tree5= tree2(1:end-1,:);
tree6= tree3(1:end-1,:);
%% indicator function
I1= zeros(N,2^N);
I1(tree6≥3)=0;
I1(tree6<3)=1;
I2 = cumprod(I1);
%% conditional probability
cp =zeros(N,2^N);
for m= 1:N
    for n = 1: 2^N
        if tree6(m,n)≥3
        cp(m,n) = 1;
        else
        cp(m,n) = Ptau(T-T*(m-1)/N, L-tree6(m,n));
        end
    end
end
%% The different choice for vt
v0 = [-3.8716, 0.25*(-3.3242)+0.75*(-3.5934), ...
    0.1528,0.25*2.5254+0.75*2.7498,5.3711];
v = zeros(N,1);
for i = 1:N %Take the average of the adjacent two v(t)
```

```
    v(i,1) = (v0(i)+v0(i+1))/2;
end
standv = std(v); %
v2 = zeros(N,3);
v2(:,1) = v+standv;
v2(:,2) = v;
v2(:,3) = v-standv;
v3 = v2(1:end-1,:);
```

```
%% conditional expectation
k=3;
vt1 = zeros(N,k^(N-1));
for i = 1:N-1
    vt1(i,:) = ...
        reshape(repmat(v3(i,:),[k^(N-1-i),k^(i-1)]),1,k^(N-1));
end
vt1(N,:) = -(sum(vt1(1:N-1,:)));
zt1=[zeros(1,3^(N-1)); cumsum(vt1)*T/N];
zt1=zt1(1:end-1,:) ;
%L1
L1v = L1(vt1,zt1);
%L2
L2v = L2(N,vt1,zt1,L,T);
%% for upward at first step
%for upward at second step
I21 =repmat(I2(:,1),1,3^(N-1));
I22 =repmat(I2(:,3),1,3^(N-1));
cp1 =repmat(cp(:,1),1,3^(N-1));
cp2 =repmat(cp(:,3),1,3^(N-1));
```

```
it1 = I21.*(L1v+cp1.*L2v)*T/N;
it2=  I22.*(L1v+cp2.*L2v)*T/N;
ita = sum((it1+it2)/2);
%for downward at second step
I23 =repmat(I2(:,5),1,3^(N-1));
I24 =repmat(I2(:,7),1,3^(N-1));
cp3 =repmat(cp(:,5),1,3^(N-1));
cp4 =repmat(cp(:,7),1,3^(N-1));
it3 = I23.*(L1v+cp3.*L2v)*T/N;
it4=  I24.*(L1v+cp4.*L2v)*T/N;
itb = sum((it3+it4)/2);
%% for downward at first step
%for upward at second step
I25 =repmat(I2(:,9),1,3^(N-1));
I26 =repmat(I2(:,11),1,3^(N-1));
cp5 =repmat(cp(:,9),1,3^(N-1));
cp6 =repmat(cp(:,11),1,3^(N-1));
it5 = I25.*(L1v+cp3.*L2v)*T/N;
it6=  I26.*(L1v+cp4.*L2v)*T/N;
itc = sum((it5+it6)/2);
%for downward at second step
I27 =repmat(I2(:,13),1,3^(N-1));
I28 =repmat(I2(:,15),1,3^(N-1));
cp7 =repmat(cp(:,13),1,3^(N-1));
cp8 =repmat(cp(:,15),1,3^(N-1));
it7 = I27.*(L1v+cp7.*L2v)*T/N;
it8=  I28.*(L1v+cp8.*L2v)*T/N;
itd = sum((it7+it8)/2);
%% select the optimal strategy
%upward at fisrt step
num1= zeros(9,2);
```

```
v1= zeros(9,2);
m1=0;
for m1 = 1:9 %find v(2.5) minimize the I2 among same ...
    v(0)and v(1.25).
     [v1(m1,1),up1] = min(ita(:,3*m1-2:3*m1));
     [v1(m1,2),up2] = min(itb(:,3*m1-2:3*m1));
     num1(m1,1)= up1*m1;
     num1(m1,2)= up2*m1;
end
v12 = (v1(:,1)+v1(:,2))/2;
num2= zeros(3,2);
v2= zeros(3,1);
for m2 = 1:3  %find v(1.25) minimize the I2 among same v(0).
     [v2(m2,1),up3] = min(v12(3*m2-2:3*m2,:));
     numa = up3*m2;
     num2(m2,:) =num1(numa,:);
end
%downward at first step
num3= zeros(9,2);
v3= zeros(9,2);
for m3 = 1:9 %find v(2.5) minimize the I2 among same ...
    v(0)and v(1.25).
     [v3(m3,1),dn1] = min(itc(:,3*m3-2:3*m3));
     [v3(m3,2),dn2] = min(itd(:,3*m3-2:3*m3));
     num3(m3,1)=dn1*m3;
     num3(m3,2)=dn2*m3;
end
v32 = (v3(:,1)+v3(:,2))/2;
num4= zeros(3,2);
v4= zeros(3,1);
for m4 = 1:3 %find v(1.25) minimize the I2 among same v(0).
```

```
    [v4(m4,1),dn3] = min(v32(3*m4-2:3*m4,:));
    numb = dn3*m4;
    num4(m4,:) =num3(numb,:);
end
%% at the beginning
v5 = (v2+v4)/2;
num5 = zeros(4,1);
[vf,numc] = min(v5); %find v(0) minimize the I2
numf =[num2(numc,:);num4(numc,:)];
```

Bibliography

[1] R. Almgren: Optimal execution with nonlinear impact functions and trading-enhanced risk, *Applied Mathematical Finance*, **10**, 1–18, 2003.

[2] R. Almgren: Optimal trading with stochastic liquidity and volatility, *SIAM Journal on Financial Mathematics*, **3**, 163–181, 2012.

[3] R. Almgren: Execution strategies in fixed income markets, *in: High Frequency Trading; New Realities for Trades, Markets and Regulators.* Eds.: D. Easley, M. Prado and M. O'Hara, Risk Books, 2013.

[4] R. Almgren and N. Chriss: Optimal execution of portfolio transactions, *Journal of Risk*, **3**, 5–39, 2000/2001.

[5] E. Bayraktar and M. Ludkovski: Optimal trade execution in illiquid markets, *Mathematical Finance*, **21**, 681–701, 2011.

[6] D. Bertsimas, and A. Lo: Optimal control of execution cost, *Journal of Financial Markets*, **1**, 1–50, 1998.

[7] C. Brown-Humes, L. Saigol and T. Tassell: Investment banks put their balance sheets on the block, *Financial Times*, Available at www.ft.com/intl/cms/s/0/258be1fc-5ebe-11d9-9c66-00000e2511c8.html, January 5, 2005.

[8] Á. Cartea, S. Jaimungal and J. Penalva: Algorithmic and High-Frequency Trading, Cambridge University Press, 2015.

[9] G. Curato, J. Gatheral and F. Lillo: Optimal execution with nonlinear transient market impact, *Available at SSRN 2539240*, 2014.

[10] C. Frei, and N. Westray: Optimal execution of a VWAP order: a stochastic control approach, *Mathematical Finance*, **25**, 612–639, 2015.

[11] J. Gatheral and A. Schied: Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework, *International Journal of Theoretical and Applied Finance*, **14**, 353–368, 2011.

[12] J. Gatheral and A. Schied: Dynamical models for market impact and algorithms for optimal order execution, *Handbook on Systemic Risk.* Eds.: J.-P. Fouque and J. Langsam, Cambridge University Press, 579–602, 2013.

[13] X. Guo, A. Larrard and Z. Ruan: Optimal placement in a limit order book, Preprint, 2013.

[14] C. Holmes and P. Castellon: Block trades: an introduction, *PLC magazine. Thomson Reuters.*, May 2006.

[15] I. Karatzas and S. Shreve: Brownian Motion and Stochastic Calculus, Second Edition, *Graduate Texts in Mathematics*, Springer, 1991.

[16] R. Nagle, E. Saff and A. Snider: Fundamentals of Differential Equations and Boundary Value Problems, Sixth Edition, Pearson, 2011.

[17] P. Potter: Stochastic Integration and Differential Equations. Second Edition, Springer, 2004.

[18] T. Roubicek: Calculus of variations, *Mathematical Tools for Physicists*, Ed.: M. Grinfeld, Wiley, 551–588, 2014.

[19] A. Schied: Robust strategies for optimal order execution in the Almgren-Chriss framework, *Applied Mathematical Finance*, **20**, 264–286, 2013.