Stochastic Modeling and Methods for Portfolio Management in Cointegrated Markets

Bahman Angoshtari

St Anne's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2013

Dedicated to my parents, Toghyan and Shahin

Acknowledgements

I would like to thank my advisor, Thaleia Zariphopoulou, for all her help, guidance and support. Her dedication and enthusiasm for academic life is both inspiring and contagious. I aspire to be as good an advisor to my students as she has been to me.

I would also like to thank Xunyu Zhou, Ronnie Sircar, Jan Obłój and Lajos Gergely Gyurkó for accepting to act as my examiners and providing invaluable feedback which has improved my thesis. Furthermore, I am greatly indebted to Anders Rahbek, Kumar Muthuraman, Gordan Žitković and Michael Monoyios who has offered me many good suggestions and comments. I have also greatly benefited from discussions with Remy Cottet, Anthony Ledford, Torsten Schöneborn, Giuliana Bordigoni and Pablo Verdes who were kind enough to share with me their expertise and insights on the practice of quantitative investment.

I would like to thank all the members of the Oxford-Man Institute of Quantitative Finance, especially its founding director, Neil Shephard, and current director, Terry Lyons, for creating such an inspiring place which have made my stay in Oxford more enjoyable and more productive. Financial support from the OMI Scholarship is also gratefully acknowledged.

I would also like to thank my brother and his wife, Arzhang and Marmar, and my sister, Roshan, for being such good friends whom I can always rely on. Finally, I would like to thank my parents for supporting and encouraging me in my studies for these many years. I can never repay them for all the sacrifices they have made to bring me up and to make me who I am today.

I remain responsible for all possible errors and omissions.

Abstract

In this thesis we study the utility maximization problem for assets whose prices are cointegrated, which arises from the investment practice of convergence trading and its special forms, pairs trading and spread trading.

The major theme in the first two chapters of the thesis, is to investigate the assumption of market-neutrality of the optimal convergence trading strategies, which is a ubiquitous assumption taken by practitioners and academics alike. This assumption lacks a theoretical justification and, to the best of our knowledge, the only relevant study is Liu and Timmermann (2013) which implies that the optimal convergence strategies are, in general, not market-neutral.

We start by considering a minimalistic pairs-trading scenario with two cointegrated stocks and solve the Merton investment problem with power and logarithmic utilities. We pay special attention to when/if the stochastic control problem is well-posed, which is overlooked in the study done by Liu and Timmermann (2013). In particular, we show that the problem is ill-posed if and only if the agent's risk-aversion is less than a constant which is an explicit function of the market parameters. This condition, in turn, yields the necessary and sufficient condition for well-posedness of the Merton problem for all possible values of agent's risk-aversion. The resulting well-posedness condition is surprisingly strict and, in particular, is equivalent to assuming the optimal investment strategy in the stocks to be market-neutral. Furthermore, it is shown that the well-posedness condition is equivalent to applying Novikov's condition to the market-price of risk, which is a ubiquitous sufficient condition for imposing absence of arbitrage. To the best of

our knowledge, these are the only theoretical results for supporting the assumption of market-neutrality of convergence trading strategies. We then generalise the results to the more realistic setting of multiple cointegrated assets, assuming risk factors that effects the asset returns, and general utility functions for investor's preference. In the process of generalising the bivariate results, we also obtained some well-posedness conditions for matrix Riccati differential equations which are, to the best of our knowledge, new.

In the last chapter, we set up and justify a Merton problem that is related to spread-trading with two futures assets and assuming proportional transaction costs. The model possesses three characteristics whose combination makes it different from the existing literature on proportional transaction costs: 1) finite time horizon, 2) Multiple risky assets 3) stochastic opportunity set. We introduce the HJB equation and provide rigorous arguments showing that the corresponding value function is the viscosity solution of the HJB equation. We end the chapter by devising a numerical scheme, based on the penalty method of Forsyth and Vetzal (2002), to approximate the viscosity solution of the HJB equation.

Contents

Chapter 0

Introduction

This thesis is a contribution to portfolio management using assets whose price processes are *cointegrated*. Such processes have the property that linear combinations of them are stationary. Intuitively speaking, two cointegrated processes are tied together, will never go too far from each other and have a long-run equilibrium with respect to each other. Many economic and financial data series are known to exhibit these properties. Examples include interest rates (Engle and Granger (1987) and Hall et al. (1992)), foreign exchange rates (Baillie and Bollerslev (1989)), equities (Cerchi and Havenner (1988)), equity indices (Taylor and Tonks (1989)), future and spot prices (Brenner and Kroner (1995)), and commodities (MacDonald and Taylor (1988)).

In portfolio management, there are specific strategies for trading assets which have co-movement in their prices. When there are only two cointegrated assets, these strategies are called **pairs trading** or **spread trading**. For the general case of several cointegrated assets, the strategies are commonly referred to as **convergence trading**. Convergence trading is, in many aspects, the ancestor of all the proprietary statistical arbitrage tools used by active portfolio managers such as hedge funds today. The strategy involves identifying two or more assets whose prices are driven by common economic forces, and then trading on any temporary deviation of the prices from their long-run equilibrium. Convergence trading strategies have been around in one form or another since the beginning of listed markets, but the hedge fund boom has given a new face to these strategies as well as the specific vehicle needed to demonstrate their successes and failures.

We refer the reader to Ehrman (2006), Vidyamurthy (2004) and Whistler (2004) for a detailed exposition on pairs-trading from a practitioner's point of view, as well as on historical insights.

There are two major themes in the convergence trading literature:

- Empirical studies on profitability of convergence trading.
- Theoretical studies on optimal convergence trading.

The first extensive empirical study on convergence trading was provided by Gatev et al. (1999) where they documented economically significant profits from implementing a very simple pairs-trading rule in the US equity market over an extended period of time. Subsequently, in Gatev et al. (2006), the researchers extended their analysis to a more recent data sample and confirmed steady profitability. Papadakis and Wysocki (2007) and Engelberg et al. (2009) used the same pairs-trading rule (a modified version for the latter study) to examine the impact of idiosyncratic news and liquidity on pairs-trading. Khandani and Lo (2007) investigated the large losses of quantitative strategies during August 2007 and suggested that the losses resulted from the sudden unwinding of large long-short equity portfolios held by multi-strategy funds. More recently, Do and Faff (2010) focused on the declining trend in the profitability of pairs-trading in recent years and attempted to identify the underlying forces. Despite confirming the downward trend in profitability, they found that the strategy performs strongly during periods of prolonged turbulence, including the recent global financial crisis. Avellaneda and Lee (2010) used more sophisticated statistical techniques to generate trading signals, and also highlighted the relationship between the performance of pairs-trading and the stock market cycle which confirms the findings of Khandani and Lo (2007). Finally, Galenko et al. (2012) used cointegration analysis to obtain mean-reverting signals, and evaluated a pre-assumed convergence trading strategy on a handful of exchange-traded funds tracking equity indices.

The trading strategies in the empirical studies above were all pre-assumed rather than being the outcome of some sort of portfolio optimization. Theoretical studies on convergence-trading in continuous-time optimal portfolio choice settings include Xiong (2001), Liu and Longstaff (2004), Jurek and Yang (2007), Mudchanatongsuk et al. (2008), Chiu and Wong (2011), and Liu and Timmermann (2013). Assuming that the spread is an Ornstein-Uhlenbeck (O-U) process and that the investors have logarithmic utility, Xiong (2001) formulated a general equilibrium model and solved it numerically. The results showed that pairs-trading can have destabilizing effects on the market. Liu and Longstaff (2004) modeled the spread by a Brownian bridge process and provided analytical solution for the associated Merton problem with logarithmic utility. Jurek and Yang (2007) and Mudchanatongsuk et al. (2008) considered an O-U spread and solved the optimal expected terminal utility problem for power utilities in closed form. The former study, provided analytical evidence for the potential destabilizing behavior of the convergence traders, consistent with the numerical solution of Xiong (2001) general equilibrium model. Finally, Chiu and Wong (2011) and Liu and Timmermann (2013) modeled the original cointegrated prices by a **continuous-time error correction model (CTECM)**. The former study solved the associated mean-variance portfolio selection problem while the latter considered the Merton problem with power utility. Both studies derived the optimal pairs-trading strategies in closed form.

All of the empirical and theoretical studies above, apart from Chiu and Wong (2011) and Liu and Timmermann (2013), implicitly or explicitly assumed the investor's strategy to be **market-neutral**. The exact definition of market-neutrality depends on the form of the mean-reverting signal used for convergence trading, and can be classified as follows:

- When the mean-reverting signal is the logarithm of the price-spread (i.e. price differences), market-neutrality is interpreted as **dollar-neutrality** which requires the (monetary) investments or portfolio weights in the assets to offset each other. This assumption is most common when trading equities.

- When the mean-reverting signal is the price-spread, market-neutrality is interpreted as **share-neutrality** in which case the number of shares or contracts in different assets offset each other. This assumption is relevant to futures markets, or when the assets are almost identical (e.g. trading Siamese twins stocks).

The main advantage of market-neutral strategies, and the reason for their popularity among practitioners and academics alike, is that the profits and losses of such strategies only depend on the change in the mean-reverting signal used. Therefore, by assuming the strategy to be market-neutral, it is enough to only consider the dynamics of the mean-reverting combination(s) of the prices, and ignore the individual asset prices altogether. This reduces the dimensionality of the problem and, more importantly, facilitates the process of model estimation and calibration, as the mean-reverting signal is a stationary process while the original price processes are not. Detailed discussion on market-neutrality will be given in Chapters 1 and 2.

Despite its convenience and practical relevance, the market-neutrality of the optimal convergence trading strategy is yet to be justified. As far as we know, the only relevant study is Liu and Timmermann (2013), which is a "negative" result. Assuming the CTECM market setting and an agent maximizing her power-utility of terminal wealth, they have shown that the optimal strategy is, in general, not market-neutral.

In the first two chapters of the thesis, we investigate the assumption of market-neutrality of the optimal convergence trading strategies. Chapter 1, takes a minimalistic pairs-trading market setting with two cointegrated stocks and solves the Merton investment problem with power and logarithmic utilities. We pay special attention to when/if the stochastic control problem is **well-posed**, which is overlooked in the study done by Liu and Timmermann (2013). Indeed, it is well known that, when the market price of risk is an Ornstein-Uhlenbeck process, the maximal expected terminal utility might explode in finite time which leads to the existence of the so-called **nirvana** strategies, c.f. Kim and Omberg (1996). We derive the necessary and sufficient condition for ill-posedness of the Merton problem (i.e. existence of nirvana strategies). In particular, it is shown that the problem is ill-posed if and only if the agent's risk-aversion is less than a constant which is an explicit function of the market parameters. This condition, in turn, yields the necessary and sufficient condition for well-posedness of the Merton problem for all possible values of agent's risk-aversion (i.e. the condition for absence of nirvana strategies). The resulting well-posedness condition is surprisingly strict and, in particular, is equivalent to assuming the optimal convergence strategies to be market-neutral. Furthermore, it is shown that the well-posedness condition is equivalent to applying Novikov's condition to the market-price of risk, which is a ubiquitous sufficient condition for imposing absence of arbitrage. To summarize, our study shows a strong connection between the market-neutrality of the optimal convergence trading strategies, well-posedness of the Merton problem, and absence of arbitrage in CTECM market setting. To the best of our knowledge, these are the only theoretical results for justifying the market-neutrality assumption in convergence trading.

Chapter 2, generalizes the results of Chapter 1 to the more realistic setting of multiple cointegrated assets, assuming risk factors that effects the asset returns, and general utility functions for investor's preference. In particular, we consider a convergence trading scenario (i.e. multiple cointegrated assets) and, similar to Liu and Timmermann (2013), introduce tradable risk factors and assume that stock prices follow a factor model on the long run (i.e. in the equilibrium state), but, short-term deviations from the factor model may occur which are captured by cointegrating relations between the stocks. We start by considering the Merton problem with power utility. Unlike the bivariate case of Chapter 1, the solution of this problem can not be expressed in closed form. Instead, we show that well-posedness of the Merton problem is equivalent to a particular **matrix Riccati differential equation (MRDE)** to have a stabilizing solution. We then prove the necessary and sufficient condition for the existence of a stabilizing solution for the MRDE, which seems to be a new result for matrix Riccati equations and can be of independent interest. This condition, in turn, yields the sufficient condition for well-posedness of the Merton problem for all power utilities which, similar to the bivariate case, is a rather strict condition on the market parameters. Next, using the general results of Kramkov and Schachermayer (1999), we prove that the well-posedness condition for power utilities is also the sufficient condition for well-posedness of the Merton problem with general utility. Finally, it is shown that the well-posedness condition is equivalent to Novikov's condition and is also equivalent to the optimal convergence trading strategies to be market-neutral as well as beta-neutral.

Chapter 3 deals with transaction costs, which are the main obstacle for implementing active trading strategies such as convergence trading that require frequent rebalancing of the portfolio. Although there is an extensive literature on portfolio choice models with transaction costs, starting with seminal works of Magill and Constantinides (1976) and Davis and Norman (1990), the majority of the results assume **static opportunity sets**, i.e. that the market price of risk is deterministic. Such results are not appropriate for convergence trading scenarios when the market price of risk is explicitly stochastic. The literature on portfolio choice with transaction costs and **stochastic opportunity set** is very thin and, to the best of our knowledge, the only study directly applicable to convergence trading is Martin and Schoneborn (2011) who considered a single Ornstein-Uhlenbeck asset and proportional transaction costs. Although, to make such connection, one must assume that the optimal strategy is market-neutral, which is yet to be established in the presence of transaction costs.

In Chapter 3, we consider spread-trading with two futures assets and assuming proportional transaction costs. We set up and justifying a Merton problem that is appropriate for portfolio managers and traders rather than individual investors. The model possesses three characteristics whose combination makes it different from the existing literature: 1) finite time horizon, 2) Multiple risky assets 3) stochastic opportunity set. After formulating the problem as a singular stochastic control problem, we introduce the associated Hamilton-Jacobi-Bellman (HJB) equation and, in particular, provide a rigorous proof of the viscosity property of the value function. By general comparison results, it then follows that the value function is the unique continuous viscosity solution of the HJB equation. We then devise a numerical scheme, based on the penalty method of Forsyth and Vetzal (2002), to approximate the viscosity solution of the HJB equation.

Chapter 1

Portfolio Choice with One Pair of Cointegrated Assets

In portfolio management, there are specific strategies for trading two assets which have co-movement in their prices. These strategies are commonly referred to as **pairs trading** or **spread trading**, depending on the context. Generally speaking, these strategies try to exploit the relative mispricing of the two assets by taking a long position in the over-priced asset and a short position in the under-priced one, while maintaining **market neutrality** by taking offsetting long/short positions. The goal of the strategy is to make profit from temporary deviations of prices from their equilibrium state, while hedging against other market movements.

As mentioned in the introduction, almost all of the quantitative analyses on pairs trading restrict the portfolio strategies to market-neutral long/short strategies. Although this approach is intuitively appealing and there are various good reasons that support it, there is, from a theoretical point of view, an unanswered fundamental question. *How can one justify this investment practice in a theoretical portfolio choice framework? In other words, can one identify a market model and a preference criterion for the investor which support pairs-trading?* The answer to this question will be the main focus of this chapter. In other words, the main motivation is to provide *a theoretical ground* for pairs-trading, *without a priori restricting the portfolio strategies.*

To provide such a framework, the first step is to identify an appropriate market model for assets which have co-movement in their prices. To this end, the concepts of cointegration and error-correction models of the seminal studies of Granger

(1981) and Engle and Granger (1987) and, in particular, their continuous-time extensions by Phillips (1991), Comte (1999), Kessler and Rahbek (2001), and Kessler and Rahbek (2004), are quite relevant. In the context of mathematical finance, these **Continuous Time Error Correction Models (CTECM)** has been used by Duan and Pliska (2004) for valuation of spread options. They can also be seen as multivariate generalisations of the Schwartz exponential Ornstein-Uhlenbeck process; see Schwartz (1997) and Benth and Karlsen (2005). More recently, Chiu and Wong (2011) and Liu and Timmermann (2013) have used these models to study pairs trading, as discussed in the introduction.

With CTECM market model at hand, one can readily apply the classical portfolio choice approach. That is, one assumes an investment horizon and choose a utility function at the end of the trading horizon, say a logarithmic or power utility. In turn, one aims at maximising the expected utility of terminal wealth and finds the optimal admissible strategy. From the mathematical point of view, such results are not new, and can be seen as a special case of the general results obtained, for example, by Goll and Kallsen (2003), Karatzas and Kardaras (2007), and Nutz (2010).

However, as pointed out by Liu and Timmermann (2013), the results obtained in this way do not support the practice of pairs-trading and, in particular, the assumption of market neutrality. Furthermore, they exhibit some unpleasant characteristics. For example, for some investors who are less risk averse than an investor with logarithmic preference, the optimal expected terminal utility increases rapidly with the investment horizon and approaches infinity at a finite critical horizon. This phenomenon is well-known for the case where the market price of risk is an Ornstein-Uhlenbeck process, and was first pointed out by Kim and Omberg (1996) who coined the terminology **nirvana strategies** for strategies which achieve infinite expected utility in finite time. From a stochastic control point of view, the problem is said to be **ill posed**, in that a minute change in the investment horizon may change the value function by a large amount.

Our main contribution to the literature of pairs trading is twofold. Firstly, we provide the necessary and sufficient condition for the market-neutrality of the optimal strategy in the framework mentioned above. Since this condition only involves market parameters, it paves the way for empirically testing the assumption of market-neutrality of the optimal convergence trading strategy. Secondly, we provide economic viability for the assumption of market-neutrality of optimal pairs-trading strategies, by showing that it is the necessary and sufficient condition for the absence of nirvana strategies (i.e. for well-posedness of Merton's problem for all power utilities), as well as a sufficient condition for the absence of risk-free arbitrage opportunities.

The existing finance literature related to our market setting, including Kim and Omberg (1996), Liu (2007), and Liu and Timmermann (2013), do not provide the so-called **verification step** in solving continuous-time portfolio optimisation problems. Neglecting the verification step does not necessarily mean that the results are wrong but, as illustrated in Korn and Kraft (2004) for a different problem, it remains an open question under which conditions the obtained results are valid. Indeed, investigating such conditions for the power utility case brought us to our main contribution and new insights on the characteristic of the portfolio choice model, which would otherwise be neglected. In Theorems 1.7 and 1.8, we provide verification results for the corresponding stochastic control problems while trying to keep the arguments as simple and accessible as possible, which we believe is a contribution on its own.

It must be mentioned that our results are complementary to, rather than conflicting with, the ones obtained by Liu and Timmermann (2013). We provide minimal economic interpretations for the results and refer the reader to Liu and Timmermann (2013), as they provide a thorough discussion on the resulting formulas and various cases that might happen. On the other hand, we focus more on the technical aspect of the market model and the associated stochastic control problem for portfolio choice. More importantly, we emphasise the role of the market-neutrality condition, as well as the possibility of the expected utility to explode in finite time horizon (i.e. the nirvana strategies), which are missing in the analysis provided by Liu and Timmermann (2013).

Finally, it is worth mentioning the link between the current study and the literature of mean-reverting returns including Kim and Omberg (1996), Wachter (2002), Liu (2007), and Jurek and Viceira (2011). Although, the CTECM market setting can be thought of as a special case of the market setting assumed in those studies, the extra structure (i.e cointegration) assumed here requires different justification and modeling, and provides more room for interpretation and mathematical scrutiny. As an example, the conditions for the existence of nirvana strategies in Kim and Omberg (1996) are in the form of various inequalities, whereas in the CTECM setting they reduce to a single equality which is much easier to interpret and test.

The rest of this chapter is organised as follows. In Section 1.1, we introduce the market model. In Section 1.2, we explain the main ideas behind pairs-trading and the approach taken by practitioners. In Section 1.3, we pose the portfolio choice problem, and point out the deficiencies in the associated optimal strategies, while in Section 1.4 we introduce an extra condition which amends the deficiencies and provides justification for pairs-trading. Finally, in Section 1.5, we present a numerical example using both simulated and real data.

1.1 The market setting

The market consists of a riskless asset that pays no interest,[1] and two stocks whose price processes, $(S_t^\top) = (S_t^1, S_t^2)_{t\geq 0}$ satisfy

$$\frac{dS_t^1}{S_t^1} = \alpha_1 Z_t dt + \sigma_1 dW_t^1, \tag{1.1}$$

and

$$\frac{dS_t^2}{S_t^2} = \alpha_2 Z_t dt + \sigma_2 \rho dW_t^1 + \sigma_2 \sqrt{1-\rho^2}\, dW_t^2, \tag{1.2}$$

where the **log-spread** $(Z_t)_{t\geq 0}$ is defined as

$$Z_t := \log S_t^1 - c \log S_t^2 + \frac{1}{2}\left(\sigma_1^2 - c\sigma_2^2\right) t. \tag{1.3}$$

On various occasions, the matrix form of (1.1) and (1.2) will be handy:

$$dS_t = \operatorname{diag}(S_t)\left(\alpha Z_t dt + \Sigma dW_t\right), \tag{1.4}$$

where

$$\Sigma := \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2 \rho & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}, \quad \text{and} \quad \alpha := \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}. \tag{1.5}$$

[1] Assuming zero short rate is not restrictive. If the short-rate is non-zero or even time-varying, one can impose the zero short-rate assumption by using the discounted prices.

Here, $(W_t)_{t\geq 0} = (W_t^1, W_t^2)_{t\geq 0}^\top$ is a two dimensional standard Brownian motion in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by W that is augmented with $\mathbb{P}$-null sets. All the coefficients are constant and the following assumption is standing throughout.

Assumption 1.1. *The following conditions hold:*

(i) $\sigma_1, \sigma_2 > 0$ *and* $|\rho| < 1$

(ii) $\alpha_1 < c\alpha_2$

(iii) $Z_0 :=$ *is a Gaussian random variable with mean zero and variance*

$$\frac{\sigma_1^2 + c^2\sigma_2^2 - 2c\rho\sigma_1\sigma_2}{2(c\alpha_2 - \alpha_1)},$$

and it is independent of $(W_t)_{t\geq 0}$.

The main implication of Assumption 1.1.(i) is the existence of the **market price of risk** given by the vector process $(Z_t\lambda)_{t\geq 0}$, where

$$\lambda := \Sigma^{-1}\alpha = \begin{pmatrix} \dfrac{\alpha_1}{\sigma_1} \\ \\ \dfrac{\alpha_2\sigma_1 - \rho\alpha_1\sigma_2}{\sigma_1\sigma_2\sqrt{1-\rho^2}} \end{pmatrix}. \tag{1.6}$$

Moreover, the associated **state price density** $(Y_t)_{t\geq 0}$ is given by the stochastic exponential:

$$Y := \mathcal{E}\left(-\int_0^{\cdot} Z_t\lambda \cdot dW_t\right). \tag{1.7}$$

The role of Assumption 1.1.(ii) and (iii), and the central feature of the price dynamics (1.1) and (1.2), is that $(Z_t)_{t\geq 0}$ is enforced to be a stationary Ornstein-Uhlenbeck process. We state this simple yet important result in the following proposition. This result is a special case of (Kessler and Rahbek, 2001, Theorem 1), but, for the sake of completeness, we provide the proof.

Proposition 1.2. *The log-spread* $(Z_t)_{t\geq 0}$ *satisfies*

$$dZ_t = -\kappa Z_t dt + \sigma_Z dW_t^Z, \tag{1.8}$$

where

$$\kappa := c\alpha_2 - \alpha_1, \quad \sigma_Z^2 := \sigma_1^2 + c^2\sigma_2^2 - 2c\rho\sigma_1\sigma_2, \tag{1.9}$$

and

$$W_t^Z := \frac{1}{\sigma_Z}\left\{(\sigma_1 - c\sigma_2\rho)\, W_t^1 - c\sigma_2\sqrt{1-\rho^2}W_t^2\right\} \tag{1.10}$$

is a standard Brownian motion. In particular, $(Z_t)_{t\geq 0}$ *is a stationary Ornstein-Uhlenbeck process given by*

$$Z_t = e^{-\kappa t}\left(Z_0 + \sigma_Z\int_0^t e^{\kappa s}dW_s^Z\right), \tag{1.11}$$

which is a Gaussian process with

$$\mathbb{E}\left(Z_t\right) = 0 \quad \text{and} \quad \mathbb{E}\left(Z_t Z_s\right) = \frac{\sigma_Z^2}{2\kappa}e^{-\kappa|t-s|}, \quad t,s \geq 0. \tag{1.12}$$

Proof. Applying Itô's lemma to (1.1) and (1.2) yields

$$d\log S_t^1 = \frac{dS_t^1}{S_t^1} - \frac{1}{2}\sigma_1^2 dt = \left(\alpha_1 Z_t - \frac{1}{2}\sigma_1^2\right)dt + \sigma_1 dW_t^1$$

and

$$d\log S_t^2 = \frac{dS_t^2}{S_t^2} - \frac{1}{2}\sigma_2^2 dt = \left(\alpha_2 Z_t - \frac{1}{2}\sigma_2^2\right)dt + \sigma_2\rho dW_t^1 + \sigma_2\sqrt{1-\rho^2}\, dW_t^2.$$

Then, (1.8)-(1.10) follow from combining these two equations. It is well-known that the strong solution of (1.8) is given by (1.11), which is a Gaussian process and, thanks to Assumption 1.1.(ii), is non-explosive.
A simple calculation shows that

$$\mathbb{E}\left(Z_t\right) = \mathbb{E}\left(Z_0\right)e^{-\kappa t} \quad \text{and} \quad Var\left(Z_t\right) = \frac{\sigma_Z^2}{2\kappa} + \left(Var\left(Z_0\right) - \frac{\sigma_Z^2}{2\kappa}\right)e^{-2\kappa t},$$

and for $t, s \geq 0$,

$$Cov\left(Z_s, Z_t\right) = \frac{\sigma_Z^2}{2\kappa}e^{-\kappa|t-s|} + \left(Var\left(Z_0\right) - \frac{\sigma_Z^2}{2\kappa}\right)e^{-\kappa(s+t)}.$$

Finally, substituting for the moments of Z_0 from Assumption 1.1.(iii) yields (1.12) which, in turn, implies the stationarity of z. □

We recall that two stochastic processes are said to be **cointegrated** if a linear combination of them is a stationary process. Hence, Proposition 1.2 implies that the stock log-prices are cointegrated. As it was mentioned in the introduction, more can be said about the connection of the market model considered herein and the theory of cointegration. Indeed, as shown in Kessler and Rahbek (2001), Kessler and Rahbek (2004) and Duan and Pliska (2004), the price dynamics given by (1.1) and (1.2) is the diffusion limit of a so-called *error correction model.* These models are discrete-time representations of systems of cointegrated processes. We, refer the interested reader to Hamilton (1994), Johansen (1995), and Juselius (2006) for a more detailed exposition on cointegration. This connection with econometrics will be particularly useful when we estimate the parameters in Section 1.5.

Remark 1.3. *The linear term $(\sigma_1^2 - c\sigma_2^2)t/2$ in (1.3) is specifically chosen such that the long-term risk premia of the stocks (i.e. the expectation of the market-price of risk) are zero. In this **pure pairs-trading** scenario, the only reason for investing in the stock is for capturing **short-term risk premium** which arises when the stock prices deviate from their equilibrium state $z = 0$. Needless to say, this is a simplistic case chosen to facilitate the analysis of pairs-trading by isolating the effect of cointegration.*
To clarify and motivate our model, consider a more realistic scenario:

$$\frac{dS_t^1}{S_t^1} = (\mu_1 + \alpha_1 Z_t)\, dt + \sigma_1 dW_t^1, \tag{1.13}$$

$$\frac{dS_t^2}{S_t^2} = (\mu_2 + \alpha_2 Z_t)\, dt + \sigma_2 \rho dW_t^1 + \sigma_2 \sqrt{1-\rho^2}\, dW_t^2, \tag{1.14}$$

and

$$Z_t := bt + \log S_t^1 - c \log S_t^2. \tag{1.15}$$

Note that the long-term risk premia of the stocks are not μ_1 and μ_2, since the stationary mean of $(Z_t)_{t\geq 0}$ is not zero, and can be shown to be:

$$m := \frac{b + \mu_1 - c\mu_2 - \frac{1}{2}(\sigma_1^2 - c\sigma_2^2)}{c\alpha_2 - \alpha_1}.$$

Since $(Z_t - m)_{t\geq 0}$ has stationary mean of zero, one may rewrite (1.13) and (1.14) as

$$\frac{dS_t^1}{S_t^1} = \left(\mu_1 + \alpha_1 m + \alpha_1 \left(Z_t - m\right)\right) dt + \sigma_1 dW_t^1$$

and

$$\frac{dS_t^2}{S_t^2} = \left(\mu_2 + \alpha_2 m + \alpha_2 \left(Z_t - m\right)\right) dt + \sigma_2 \rho dW_t^1 + \sigma_2 \sqrt{1-\rho^2}\, dW_t^2,$$

to deduce that the long-term risk premia of the stocks are $\mu_1 + \alpha_1 m$ and $\mu_2 + \alpha_2 m$. To obtain the pure pairs-trading market setting (1.1) and (1.2), we require the long-term risk premium to be zero, i.e.

$$\begin{cases} \mu_1 + \alpha_1 \dfrac{b + \mu_1 - c\mu_2 - (\sigma_1^2 - c\sigma_2^2)/2}{c\alpha_2 - \alpha_1} = 0, \\ \\ \mu_2 + \alpha_2 \dfrac{b + \mu_1 - c\mu_2 - (\sigma_1^2 - c\sigma_2^2)/2}{c\alpha_2 - \alpha_1} = 0. \end{cases}$$

Assuming $c\alpha_2 \neq \alpha_1$ and a direct computation yield

$$\begin{cases} c(\alpha_2\mu_1 - \alpha_1\mu_2) = -\alpha_1 \left(b - \frac{1}{2}\left(\sigma_1^2 - c\sigma_2^2\right)\right), \\ \\ (\alpha_2\mu_1 - \alpha_1\mu_2) = -\alpha_2 \left(b - \frac{1}{2}\left(\sigma_1^2 - c\sigma_2^2\right)\right) \end{cases}$$

Therefore, there are only two possibilities if the long-term risk premia are zero:

(i) $\alpha_2\mu_1 \neq \alpha_1\mu_2$ *and* $b \neq (\sigma_1^2 - c\sigma_2^2)/2$, *therefore* $c\alpha_2 = \alpha_1$ *which is, however, a contradiction,*

and

(ii) $\alpha_2\mu_1 = \alpha_1\mu_2$ *and* $b = (\sigma_1^2 - c\sigma_2^2)/2$.

Our market setting assumes $\mu_1 = \mu_2 = 0$, and $b = (\sigma_1^2 - c\sigma_2^2)/2$, which satisfies (ii). Of course, another possibility is to assume $\mu_1, \mu_2 \neq 0$, and add $\alpha_2\mu_1 = \alpha_1\mu_2$ as an assumption, but we decided against adding the extra parameters μ_1 and μ_2 in order to keep the notation light. It must be mentioned that in Chapter 2, we will extend our analysis to a setting which is even more general than (1.13)-(1.15).

1.2 Pairs-trading in investment practice

As mentioned earlier, the main motivation of this paper is to provide a theoretical ground for pairs-trading and spread-trading rules (the difference between the two will be clarified shortly). Before we provide such a justification, we explain the main ideas behind pairs-trading and, specifically, the approach taken by practitioners. To make the arguments more precise, we assume that, in accordance to the model introduced earlier, we have only one pair of assets, say S^1 and S^2. Moreover, to keep the concepts intuitive, the arguments will be presented in an informal way. We refer the reader to Ehrman (2006), Vidyamurthy (2004) and Whistler (2004) for a detailed exposition on pairs-trading.

The key step in pairs-(or spread-)trading strategies is to find a way to quantify the **relative price** of the pair. Note that co-movement in prices implies that there should be a way to combine the two prices to obtain a mean-reverting process in a way highlighted in Proposition 1.2. *It is this mean-reverting process that is used to quantify the relative price of the pair.*

There are two common assumptions regarding this relative price indicator:

i) A linear combination of the asset prices S^1 and S^2 is mean-reverting. In other words, there is a constant c such that the process $(s_t)_{t\geq 0}$, given by

$$s_t = S_t^1 - cS_t^2,$$

is mean-reverting,

or,

ii) A linear combination of the logarithm of the prices is mean-reverting. In other words, there is a constant c such that the process $(Z_t)_{t\geq 0}$, defined as

$$Z_t = \log S_t^1 - c \log S_t^2, \tag{1.16}$$

is mean-reverting.

To differentiate between the two cases, we will be referring to s_t as the *spread* and to Z_t as the *log-spread*. In analogy, we will be referring to market settings with assumption (i) as **spread-trading**, and those with assumption (ii) as **pairs-trading**. In practice, pairs-trading is more suitable for trading equities, while spread-trading is more suitable for trading futures or when the assets are almost identical (e.g. trading Siamese twins stocks). In this chapter and the next, we work under assumption (ii), that is pairs-trading. In Chapter 3, we will adapt a spread-trading setting.

Next, we explain the practitioners' approach. A pairs-trader starts by identifying the residual Z_t of (1.16), and then ignores the individual prices S_t^1 and S_t^2. Note that modeling Z_t is essentially equivalent to determining the price of one of the assets in terms of the other. For this reason, any model for Z_t is often called a **partial pricing model**.

A benchmark partial pricing model is to assume that Z_t, $t \geq 0$, is a stationary Ornstein–Uhlenbeck process, given by

$$dZ_t = \kappa \left(\bar{z} - Z_t\right) dt + \sigma_z dW_t^z, \tag{1.17}$$

with $Z_0 = \log S_0^1 - c \log S_0^2$. Here (with a slight abuse of notation) κ, $\bar{z}$, and σ_z are constants, $\kappa > 0$, and W_t^z is a standard Brownian motion.

It is crucial to differentiate between the partial pricing model (1.17) and the seemingly identical equation (1.8). The former is an assumption about the market per se, while the latter is a direct consequence of the price equations (1.1) and (1.2). As mentioned earlier, a pairs-trader does **not** model S_t^1 and S_t^2 separately. Instead, he takes (1.17) as the market model. On the contrary, we model the original prices via (1.1) and (1.2).

After assuming such a model, the pairs-trader restricts the candidate market strategies to the so-called **pairs-trading strategies**.

Definition 1.4. *Let the mean-reverting signal be given by (1.16) with $c > 0$ and consider a trading strategy represented by (π_t^1, π_t^2), where π_t^i is the portfolio weight of the i-th asset at t. Then, the strategy is a* ***pairs-trading strategy*** *if the following two properties hold:*

1. *The strategy maintains a short position in the over-priced stock and a long position in the under-priced one, as indicated by the sign of the residual Z_t. More specifically, $\pi_t^1 Z_t \leq 0$ and $\pi_t^2 Z_t \geq 0$, for all t.*

*2. The strategy is **market-neutral**, namely:*

$$\pi_t^2 = -c\,\pi_t^1. \tag{1.18}$$

The idea behind the first property, i.e. keeping long (resp. short) positions in the under-priced (resp. over-priced) asset, is simple. By maintaining such positions, one will make profit as the log-spread converges to its equilibrium. Therefore, the relative mispricing of the pair will be constantly exploited.

The main advantage of the second property, i.e. market-neutrality, is that the profits and losses of such strategies only depend on the change in the log-spread. We use a simple discrete-time argument to illustrate the idea. Let r^p, r^1 and r^2 be the excess returns over time period (t_1, t_2) of the portfolio, asset S^1, and asset S^2, respectively. Further, let π^i be the portfolio weight of the i-th asset over the same time interval, and Z be the log-spread given by (1.16). Then, one has

$$r^p = \pi^1 r^1 + \pi^2 r^2 \approx \pi^1 \Delta \ln S^1 + \pi^2 \Delta \ln S^2 = \pi^1 \Delta Z + \left(\pi^2 + c\,\pi^1\right) \Delta \ln S^2.$$

For a market-neutral strategy satisfying (1.18), the last term vanishes, i.e.

$$r^p = \pi^1 \Delta Z.$$

Therefore, by assuming the strategy to be market-neutral, it is enough to only consider the dynamics of the log-spread, and ignore the individual asset prices altogether. This reduces the dimensionality of the problem and, more importantly, facilitates the process of model estimation and calibration, as the spread is a stationary process while the original price processes are not.

Although this approach is intuitively appealing and there are various good reasons that support it, there is, from a theoretical point of view, an unanswered fundamental question. *How can one justify this investment practice in a theoretical portfolio choice framework? In other words, can one identify a market model and a preference criterion for the investor which support pairs-trading?* This idea will be the theme for the rest of the chapter. Specifically, we try to justify pairs-trading **without** a priori restricting the portfolio strategies. For this, we first need to consider a full pricing model which implies the partial pricing model (1.17). Note that according to Proposition 1.2, the market setting of section 1.1 fulfills this requirement.

1.3 Optimal strategies for CRRA investors

To provide a theoretical framework for the investment practice of pairs-trading, we consider a risk preference criterion for the investor, analyze the associated maximal expected utility problem, and explore the connection between pairs-trading and the optimal investment strategies.

Assume there is an agent who invests in the market of Section 1.1 over a fixed trading horizon $[0,T]$ and with an initial endowment $x>0$. An **admissible strategy** is defined as an integrable process $(\pi_t^\top) = (\pi_t^1, \pi_t^2)_{t\in[0,T]}$, i.e. an $(\mathcal{F}_t)$-adapted process satisfying

$$\int_0^T \left(|\pi_t^\top \alpha Z_t| + \pi_t^\top \Sigma\Sigma^\top \pi_t\right) dt < \infty, \quad \mathbb{P}\text{-almost surely.} \tag{1.19}$$

π_t^i is interpreted as the **proportion** of agent's wealth invested in the i-th stock at t. Then, $(1-\pi_t^1-\pi_t^2)_{t\in[0,T]}$ is the proportion of wealth invested in the bank account. The set of admissible strategies is denoted by $\mathcal{A}$.

For any admissible strategy $\pi = (\pi^1,\pi^2)\in\mathcal{A}$, the agent's wealth $(X_t^\pi)_{t\in[0,T]}$ is given by the stochastic exponential

$$X^\pi = x\,\mathcal{E}\left(\int_0^\cdot \pi_t^\top \alpha Z_t dt + \int_0^\cdot \pi_t^\top \Sigma dW_t\right), \tag{1.20}$$

which is positive, $\mathbb{P}$-almost surely.

We consider the classical Merton investment problem:

Definition 1.5. *The optimal strategy is defined as*

$$(\pi^\star_{\gamma,t})_{t\in[0,T]} := \arg\max_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^\pi\right)\right), \tag{1.21}$$

assuming that

$$\mathbb{E}\Big(\max_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^\pi\right)|Z_0\right)\Big) < \infty. \tag{1.22}$$

Otherwise, we say that the optimal strategy does not exists. Here, the agent's terminal utility $U_\gamma(.)$ is either power or logarithmic utility:

$$U_\gamma(x) := \begin{cases} \dfrac{x^{1-\gamma}-1}{1-\gamma}, & \gamma\in(0,\infty)\backslash\{1\}, \\ \log(x), & \gamma = 1, \end{cases} \tag{1.23}$$

where $\gamma = -\frac{xU_\gamma''(x)}{U_\gamma'(x)} > 0$ is the investor's relative risk aversion.

Remark 1.6. *In the original formulation of the Merton problem, the initial values are assumed to be known while, by Assumption 1.1.(iii), Z_0 is random. Note that, in general, one has*

$$\sup_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^\pi\right)\right) \leq \mathbb{E}\left(\operatorname*{ess\,sup}_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^\pi\right)|Z_0\right)\right).$$

Therefore, if

$$\arg\max_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^\pi\right)|Z_0 = z\right) \tag{1.24}$$

is well-defined for all $z \in \mathbb{R}$ in feedback form such that (1.22) holds, then (1.21) is also well defined and the two strategies coincide. In the sequel, we solve the unconditional problem (1.21) by considering the problem (1.24) and then showing the existence of an optimal feedback control satisfying (1.22).

Next, we introduce the *value function*. For $(\pi, x, z, s) \in \mathcal{A} \times \mathbb{R}^+ \times \mathbb{R} \times [0, T]$, define $(X_t^{\pi,x,z,s})_{t\in[s,T]}$ as the wealth process of the agent from s to T if, at s, her wealth is x, the log-spread is z, and she follows an admissible strategy π. Similarly, let $(Z_t^{z,s})$ be the log-spread from s to T if the log-spread is z at s. In other words, $(X_t^{\pi,x,z,s}, Z_t^{z,s})_{t\in[s,T]}$ is the unique strong solution of the stochastic differential equations

$$\frac{dX_t^{\pi,x,z,s}}{X_t^{\pi,x,z,s}} = \pi_t^\top \alpha\, Z_t^{z,s}\, dt + \pi_t^\top \Sigma\, dW_t, \tag{1.25}$$

and

$$dZ_t^{z,s} = (1, -c)\alpha\, Z_t^{z,s}\, dt + (1, -c)\Sigma\, dW_t, \tag{1.26}$$

with the initial conditions $X_s^{\pi,x,z,s} = x$ and $Z_s^{z,s} = z$. We then define the value function

$$u(t, x, z; T, \gamma) := \sup_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^{\pi,x,z,t}\right)\right). \tag{1.27}$$

We consider the logarithmic and power case separately in the following subsections.

1.3.1 Logarithmic utility

The following theorem gives the optimal portfolio strategy for logarithmic utility, i.e. $\gamma = 1$.

Theorem 1.7. *The optimal strategy for the logarithmic case is given by:*

$$\pi^{\star}_{1,t} = \Sigma^{-1\top}\lambda Z_t = \frac{1}{\sigma_1\sigma_2\left(1-\rho^2\right)}\begin{pmatrix}\alpha_1\frac{\sigma_2}{\sigma_1}-\rho\alpha_2\\ \alpha_2\frac{\sigma_1}{\sigma_2}-\rho\alpha_1\end{pmatrix}Z_t, \tag{1.28}$$

and the value function is given by

$$u(t,x,z;T,1) = \log x + \frac{1}{2}\|\lambda\|^2\mathbb{E}\Big(\int_t^T (Z_s^{z,t})^2 ds\Big). \tag{1.29}$$

In particular, the value function is bounded for any $(t,x,z) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}$.

Merton's problem with logarithmic utility is a well-studied problem, and for existing results on general market settings see Goll and Kallsen (2003) and Karatzas and Kardaras (2007). Although we could derive Theorem 1.7 from these general results, we opted not to do so, because of the simplicity of our market setting and the aim to make our arguments accessible to the more applied finance community. On the other hand, the existing finance literature related to our market setting, including Kim and Omberg (1996), Liu (2007), and more recently Liu and Timmermann (2013), do not provide the so-called *verification lemmas* for the corresponding stochastic control problems. Neglecting the verification lemmas does not necessarily mean that the results are wrong but, as illustrated in Korn and Kraft (2004) for a different problem, it remains an open question that *under which conditions the obtained results are valid.* Indeed, investigating such conditions for the power utility case brought us to our main contribution and new insights on the characteristic of the portfolio choice model, which was otherwise neglected. In summary, our main reason for providing the proofs in detail, is to highlight the importance of the verification results while keeping the arguments simple and accessible, which we believe is a contribution on its own.

Proof of Theorem 1.7. The proof is divided in two steps. In the first step we provide an upper bound for the value function, while in the second we show that this upper bound is attained by an admissible strategy. The second step is based on Angoshtari (2009) which, in turn, has been motivated by Jamshidian (private communication, 2009).

Step 1: Define the process $(Y_t^{z,s})_{t\in[s,T]}$ by

$$\begin{aligned} Y_t^{z,s} &:= \mathcal{E}\left(-\int_s^\cdot Z_u^{z,s}\lambda \cdot dW_u\right)_t \\ &= \exp\Big(-\tfrac{1}{2}\int_s^t \|Z_u^{z,s}\lambda\|^2 du - \int_s^t Z_u^{z,s}\lambda \cdot dW_u\Big). \end{aligned} \tag{1.30}$$

By the product rule, for $(\pi, x, z, s) \in \mathcal{A} \times \mathbb{R}^+ \times \mathbb{R} \times [0, T]$, one has

$$\begin{aligned} d(Y_t^{z,s} X_t^{\pi,x,z,s}) &= Y_t^{z,s} dX_t^{\pi,x,z,s} + X_t^{\pi,x,z,s} dY_t^{z,s} + d\langle Y^{z,s}, X^{\pi,x,z,s}\rangle_t \\ &= Y_t^{z,s} X_t^{\pi,x,z,s}\left(\pi_t^\top \Sigma - Z_t \lambda^\top\right) dW_t. \end{aligned}$$

Therefore, the discounted wealth process $(Y_t^{z,s} X_t^{\pi,x,z,s})_{t\in[s,T]}$ is a non-negative local martingale, and, by Fatou's lemma, a supermartingale. In particular,

$$\mathbb{E}\left(Y_T^{z,s} X_T^{\pi,x,z,s}\right) \le x. \tag{1.31}$$

Let $V_1 : \mathbb{R}^+ \to \mathbb{R}$ be the convex conjugate of the logarithmic utility function $U_1(x) = \log(x)$, i.e.

$$V_1(y) := \sup_{x\in\mathbb{R}^+} \{U_1(x) - xy\} = -1 - \log y, \quad y > 0. \tag{1.32}$$

The following duality argument is quite useful. For $(\pi, x, y, z, t) \in \mathcal{A} \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \times [0, T]$, one has

$$\begin{aligned} \mathbb{E}\Big(U_1\left(X_T^{\pi,x,z,t}\right)\Big) &\le \mathbb{E}\Big(U_1(X_T^{\pi,x,z,t})\Big) + y\Big(x - \mathbb{E}(Y_T^{z,t} X_T^{\pi,x,z,t})\Big) \\ &= \mathbb{E}\Big(U_1(X_T^{\pi,x,z,t}) - y Y_T^{z,t} X_T^{\pi,x,z,t}\Big) + xy \\ &\le \mathbb{E}\Big(V_1(y Y_T^{z,t})\Big) + xy = -\mathbb{E}(\log Y_T^{z,t}) + xy - \log y - 1, \end{aligned}$$

where we used the supermartingale property (1.31) and the definition of the convex conjugate function (1.32) to obtain the first and the second inequalities, respectively. Maximising the left side of this inequality among admissible controls $\pi \in \mathcal{A}$ and minimising the right side with respect to $y \in \mathbb{R}^+$ yield an upper bound for the value function:

$$\begin{aligned}
u(t,x,z;T,1) &:= \sup_{\pi\in\mathcal{A}} \mathbb{E}\left(U_1\left(X_T^{\pi,x,z,t}\right)\right) \\
&\le \inf_{y\in\mathbb{R}^+} \left\{-\mathbb{E}(\log Y_T^{z,t}) + xy - \log y - 1\right\} \\
&= \log x - \mathbb{E}(\log Y_T^{z,t}) \\
&= \log x + \frac{1}{2}\|\lambda\|^2 \mathbb{E}\left(\int_t^T (Z_s^{z,t})^2 ds\right) + \mathbb{E}\int_t^T Z_s^{z,t}\lambda^\top dW_s \\
&= \log x + \frac{1}{2}\|\lambda\|^2 \mathbb{E}\left(\int_t^T (Z_s^{z,t})^2 ds\right) < \infty, \qquad (1.33)
\end{aligned}$$

where in the last step, we used the square integrability of the (non-explosive) Ornstein-Uhlenbeck process, i.e.

$$\mathbb{E}\int_0^T Z_s^2 ds < \infty. \qquad (1.34)$$

Step 2: By (1.20) we have

$$\mathbb{E}\log X_T^{\pi,x,z,t} = \log x + \mathbb{E}\int_t^T \left(\pi_s^\top \alpha Z_s^{z,t} - \frac{1}{2}\pi_s^\top \Sigma\Sigma^\top \pi_s\right) ds + \mathbb{E}\int_t^T \pi_s^\top \Sigma dW_s. \quad (1.35)$$

For the moment, we assume that the square integrability condition

$$\mathbb{E}\Big[\int_0^T \pi_s^\top \Sigma\Sigma^\top \pi_s ds\Big] < \infty \qquad (1.36)$$

holds. We will shortly verify that the optimal strategy indeed satisfies this condition. Then, $\mathbb{E}\int_t^T \pi_s^\top \Sigma dW_s = 0$, and (1.35) becomes

$$\begin{aligned}
\mathbb{E}\log X_T^{\pi,x,z,t} &= \log x + \mathbb{E}\int_t^T \left(\pi_s^\top \alpha Z_s^{z,t} - \frac{1}{2}\pi_s^\top \Sigma\Sigma^\top \pi_s\right) ds \\
&= \log x + \frac{1}{2}\mathbb{E}\int_t^T \left(\|\lambda\|^2 (Z_s^{z,t})^2 - \|\Sigma^\top \pi_s - \lambda Z_s^{z,t}\|^2\right) ds. \qquad (1.37)
\end{aligned}$$

This expectation is maximised by the strategy:

$$\pi_s^{z,t} := \Sigma^{-1\top}\lambda Z_s^{z,t}, \quad s\in[t,T].$$

In particular, $(\pi_s^{z,t})_{s\in[t,T]}$ satisfies the square integrability condition (1.36) because of (1.34). Substituting $\pi^{z,t}$ in (1.37) yields

$$\mathbb{E}\left(U_1\left(X_T^{\pi^{z,t},x,z,t}\right)\right) = \log x + \frac{1}{2}\|\lambda\|^2 \mathbb{E}\left(\int_t^T (Z_s^{z,t})^2 ds\right).$$

In other words, $(\pi_s^{z,t})_{s\in[t,T]}$ attains the upper bound in (1.33), which implies that it must be the optimal strategy for $Z_t = z$. Hence, the value function is the upper bound found in the first step of the proof. Finally, the optimal strategy is $(\pi_t^\star) := (\pi_t^{Z_0,0})_{t\in[0,T]})$, which coincides with (1.28). □

The well-known property of the log-optimal allocation is that it is **myopic**, in that a long term strategy can be thought of as a sequence of short-term strategies executed one after another, cf. Mossin (1968). Indeed, assume that there are two investors, both with logarithmic utility but with two different time horizons, T' and T, where $T' < T$ (so the first investor is short-sighted or myopic compare to the second investor). Note that the investment horizon T does not appear in equation (1.28). Therefore, both investors will follow precisely the same investment strategy on the interval $t \in [0, T')$. It follows that one can think of a long-term log-optimal strategy as a sequence of short-term log-optimal strategies executed one after another.

It is important to observe that the optimal strategy of Theorem 1.7 does *not* justify pairs-trading as defined in Section 1.2. Indeed, the optimal portfolio (1.28) firstly does not satisfy the market-neutral condition (1.18) and, secondly, it depends on α and Σ. Hence, it cannot be identified if one only knows the partial pricing model (1.17). We provide the remedy for these deficiencies in Section 1.4.

1.3.2 Power utility

Next, we consider the Merton investment problem (1.21) with power utility, i.e. $\gamma \in (0,1) \cup (1,+\infty)$. Different values for γ are interpreted as follows: The logarithmic utility can be considered as the limiting case of the power utility when $\gamma \to 1$. If $\gamma > 1$ (resp. $0 < \gamma < 1$), the investor is more risk averse (resp. more risk seeking) than a log-utility investor. If $\gamma \downarrow 0$, the investor becomes risk neutral.

Our main insight of this section is identifying the exact **well-posedness** conditions for the Merton problem with power utility. In particular, we introduce the **critical relative risk aversion**:

$$\gamma_0 := 1 - \left(\frac{\kappa}{\sigma_Z\|\lambda\|}\right)^2 = 1 - \left(\frac{(1,-c)\alpha}{\|(1,-c)\Sigma\|\,\|\Sigma^{-1}\alpha\|}\right)^2 \in [0,1). \tag{1.38}$$

The fact that $\gamma_0 \in [0,1)$ follows from the Cauchy-Schwarz inequality:

$$\kappa := -(1,-c)\alpha \leq \|(1,-c)\Sigma\| \, \|\Sigma^{-1}\alpha\| =: \sigma_Z \|\lambda\|.$$

We show that the optimal strategy (1.21) and the associated value function (1.27) are classified based on the relative risk aversion γ as follows:

(i) If $\gamma \geq \gamma_0$, then the Merton problem is **well-posed**, in the sense that the value function is finite, for *any choice* of time horizon $T > 0$. See Theorem 1.8.

(ii) If $\gamma < \gamma_0$, then, as shown in Theorem 1.9, the Merton problem is **ill-posed** and, in the terminology of Kim and Omberg (1996), the optimal strategies are **nirvana strategies**. This means that the agent's expected terminal utility of wealth increases rapidly with the investment horizon and approaches infinity at a finite critical horizon, denoted by $T_{\text{nirvana}}(\gamma)$, which is explicitly calculated in (1.52).

Note that since $\gamma_0 < 1$, the logarithmic case as well as power utilities with $\gamma > 1$ are always well-posed; the ill-posed case can only happen for power utilities with $0 < \gamma < 1$, i.e. when the agent is more **risk seeking** than a log-utility investor.

Similar to the case of logarithmic utility, the case of power utility is also a well-studied problem. We refer, among others, to the PhD thesis Nutz (2010) for results in general semi-martingale markets and the relevant literature. But, these general results are based on the assumption that the **value function is finite**. As already mentioned, the value function can be infinite in our market setting . Therefore, the existing results are not directly applicable, and it is essential that we provide the proofs in detail, with special attention to when/if the problem is **well-posed**.

Theorem 1.8. *(The well-posed case)*
For $(T,\gamma) \in (0,\infty) \times [\gamma_0, 1) \cup (1,\infty)$, *the value function* $u(.,.,.;T,\gamma) : [0,T] \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ *is bounded and given by:*

$$u(t,x,z;T,\gamma) = \frac{x^{1-\gamma}\left(e^{g_{w.p.}(T-t,\gamma)+\frac{1}{2}h_{w.p.}(T-t,\gamma)z^2}\right)^{\gamma} - 1}{1-\gamma}, \tag{1.39}$$

and the optimal strategy is

$$\pi^\star_{\gamma,t} = \frac{1}{\gamma}\pi^\star_{1,t} + h_{w.p.}(T-t,\gamma) Z_t \begin{pmatrix} 1 \\ -c \end{pmatrix}, \quad t \in [0,T]. \tag{1.40}$$

Here, $g_{w.p.}(\cdot,\gamma)$ *and* $h_{w.p.}(\cdot,\gamma)$ *are given by (1.88) and (1.89), respectively, and* $(\pi^\star_{1,t})_{t\in[0,T]}$ *is the optimal strategy for the logarithmic utility given by (1.28).*

Proof. The proof is divided into three steps. In the first step, an upper bound for the value function is obtained, in terms of an expectation of the state price density $(Y_t)_{t\in[0,T]}$. In the second step, an explicit formula for the expectation in the upper bound is obtained and, in particular, it is shown that the upper bound is **bounded** itself. In the third step, an admissible strategy is constructed which attains the upper bound found in the first step, hence proving that the upper bound is the value function.

Step 1: Let $V_\gamma : \mathbb{R}^+ \to \mathbb{R}$ be the convex conjugate of the power utility function $U_\gamma(x) = (x^{1-\gamma}-1)/(1-\gamma)$, i.e.

$$V_\gamma(y) := \sup_{x\in\mathbb{R}^+} \{U_\gamma(x) - xy\} = \frac{\gamma\, y^{\frac{\gamma-1}{\gamma}} - 1}{1-\gamma}, \quad y > 0,$$

and consider the process $(Y^{z,s}_t)_{t\in[s,T]}$ given by (1.30). Then, a similar argument as in Step 1 of the proof of Theorem 1.7 yields that for any $(\pi, x, y, z, t) \in \mathcal{A} \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \times [0,T]$:

$$\begin{aligned} \mathbb{E}\Big(U_\gamma\left(X^{\pi,x,z,t}_T\right)\Big) &\leq \mathbb{E}\Big(V_\gamma(yY^{z,t}_T)\Big) + xy \\ &= \frac{\gamma\,\mathbb{E}\Big((Y^{z,t}_T)^{\frac{\gamma-1}{\gamma}}\Big)\, y^{\frac{\gamma-1}{\gamma}} + (1-\gamma)xy - 1}{1-\gamma}. \end{aligned}$$

Maximising the left side for all $\pi \in \mathcal{A}$ and minimising the right side among all

$y \in \mathbb{R}^+$ yield an upper bound for the value function:

$$
\begin{aligned}
u(t,x,z;T,\gamma) &:= \sup_{\pi\in\mathcal{A}} \mathbb{E}\left(U_\gamma\left(X_T^{\pi,x,z,t}\right)\right) \\
&\leq \inf_{y\in\mathbb{R}^+}\left\{\frac{\gamma\,\mathbb{E}\Big((Y_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big)\,y^{\frac{\gamma-1}{\gamma}}+(1-\gamma)xy-1}{1-\gamma}\right\} \\
&= \frac{\Big(\mathbb{E}\Big[(Y_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big]\Big)^{\gamma}x^{1-\gamma}-1}{1-\gamma}. \qquad (1.41)
\end{aligned}
$$

Step 2: The goal of this step is to show the boundedness of the expectation $\mathbb{E}\Big[(Y_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big]$ for $\gamma \geq \gamma_0$. To this end, define the function $\psi : [0,T]\times\mathbb{R}^+\times\mathbb{R}\to\mathbb{R}$

$$
\psi(t,y,z) := \mathbb{E}\Big[(yY_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big]. \qquad (1.42)
$$

We then study the related Cauchy problem:

$$
\psi_t + \frac{1}{2}y^2z^2\|\lambda\|^2\psi_{yy} - \kappa z\,\psi_z + \frac{1}{2}\sigma_Z^2\psi_{zz} + \kappa yz\psi_{yz} = 0; \qquad (1.43)
$$

$(t,y,z)\in[0,T)\times\mathbb{R}^+\times\mathbb{R}$, with terminal condition $\psi(T,y,z) = y^{\frac{\gamma-1}{\gamma}}$. If (1.43) has a (classical) solution, then the Feynman-Kac formula yields the stochastic representation (1.42).

Substituting the ansatz

$$
\psi(t,y,z) = y^{\frac{\gamma-1}{\gamma}}\varphi(t,z),
$$

yields the following PDE for the unknown function $\varphi : [0,T]\times\mathbb{R}\to\mathbb{R}$:

$$
\varphi_t - \frac{1}{\gamma}\kappa z\varphi_z + \frac{1}{2}\sigma_Z^2\varphi_{zz} + \frac{1-\gamma}{2\gamma^2}z^2\|\lambda\|^2\varphi = 0; \quad (t,z)\in[0,T)\times\mathbb{R}, \qquad (1.44)
$$

with the terminal conditions $\varphi(T,z) = 1$. The solution of this PDE is given in Appendix 1.A. In particular, by taking $\mathbf{a} = (1/\gamma)\lambda$, $\mathbf{b}^\top = (1,-c)\Sigma$, and $\xi = 1-\gamma$, one may re-write (1.44) as (1.71). The corresponding escape criterion discriminant defined by (1.73) is:

$$
\mathfrak{D} = \frac{\kappa^2}{\gamma^2} - \frac{1-\gamma}{\gamma^2}\sigma_Z^2\|\lambda\|^2 = \frac{\sigma_Z^2\|\lambda\|^2}{\gamma^2}(\gamma-\gamma_0). \qquad (1.45)
$$

Because of the assumption $\gamma \geq \gamma_0$, we have that $\mathfrak{D} \geq 0$. Therefore, Proposition 1.18 applies and PDE (1.44) is well-posed with the unique solution

$$\varphi_{\text{w.p.}}(t, z) = e^{g_{\text{w.p.}}(T-t,\gamma)+\frac{1}{2}z^2 h_{\text{w.p.}}(T-t,\gamma)}; \quad t \in [0, T], \tag{1.46}$$

where $g_{\text{w.p.}}(\cdot, \cdot)$ and $h_{\text{w.p.}}(\cdot, \cdot)$ are given by (1.88) and (1.89), respectively. It follows that $\psi(t, y, z)$ of (1.42) is bounded and, in particular,

$$\mathbb{E}\Big[(Y_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big] = e^{g_{\text{w.p.}}(T-t,\gamma)+\frac{1}{2}z^2 h_{\text{w.p.}}(T-t,\gamma)}; \quad t \in [0, T]. \tag{1.47}$$

Step 3: To construct the optimal strategy, we follow the classical stochastic control approach through the Hamilton-Jacobi-Bellman (HJB) equation:

$$u_t + H(x, z, u_x, u_{xx}, u_z, u_{zz}, u_{xz}) = 0, \quad (t, x, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}, \tag{1.48}$$

with the terminal condition $u(T, x, z; T, \gamma) = \frac{x^{1-\gamma}-1}{1-\gamma}$. Here, the Hamiltonian is

$$\begin{aligned} H(x, z, u_x, u_{xx}, u_z, u_{zz}, u_{xz}) &= -\kappa z u_z + \tfrac{1}{2}\sigma_Z^2 u_{zz} \\ &+ \sup_{\pi\in\mathbb{R}^2}\Big\{\big(z u_x \alpha^\top + u_{xz}(1, -c)\Sigma\Sigma^\top\big)x\pi + \tfrac{1}{2}x^2\pi^\top\Sigma\Sigma^\top\pi u_{xx}\Big\}. \end{aligned} \tag{1.49}$$

Optimising the right side of this equation yields the candidate optimal strategy:

$$\pi^\star(x, z, u_x, u_{xx}, u_z, u_{zz}, u_{xz}) := \frac{-u_x}{x u_{xx}} z(\Sigma\Sigma^\top)^{-1}\alpha - \frac{u_{xz}}{x u_{xx}}(1, -c)^\top. \tag{1.50}$$

Substituting $\pi^\star$ into (1.49) and then into (1.48) yields

$$u_t - \kappa z u_z + \frac{1}{2}\sigma_Z^2 u_{zz} - \frac{1}{2}\|\lambda\|^2 z^2 \frac{u_x^2}{u_{xx}} - \frac{1}{2}\sigma_Z^2\frac{u_{xz}^2}{u_{xx}} + \kappa z\frac{u_x u_{xz}}{u_{xx}} = 0; \tag{1.51}$$

$(t, x, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}$, with the terminal condition $u(T, x, z; T, \gamma) = \frac{x^{1-\gamma}-1}{1-\gamma}$ as before.

With (1.41) in mind, substituting the ansatz

$$u(t, x, z; T, \gamma) = \frac{\big(\varphi(t, z)\big)^\gamma x^{1-\gamma} - 1}{1-\gamma}$$

into (1.51) yields that the unknown function $\varphi(\cdot, \cdot)$ satisfies (1.44) and, therefore, is given by (1.46). In turn, it follows that the solution of the HJB equation (1.51) is (1.39). Substituting into (1.50) yields the optimal solution as in (1.40), which is square integrable thanks to (1.34) and the boundedness of $h_{\text{w.p.}}$ given by (1.89). Finally, the solution of (1.51) coincides with the upper bound (1.41) found in Step 1, which verifies that (1.39) is indeed the value function. □

Theorem 1.9. *(The ill-posed case and nirvana strategies)*
For $\gamma \in (0, \gamma_0)$ and $T \in (0, T_{nirvana}(\gamma))$, where

$$T_{nirvana}(\gamma) := \frac{\gamma}{\sigma_Z \|\lambda\| \sqrt{\gamma_0 - \gamma}} \left(\frac{\pi}{2} + \arctan \left(\frac{\kappa \gamma}{\sigma_Z \|\lambda\| \sqrt{\gamma_0 - \gamma}} \right) \right), \tag{1.52}$$

the value function $u(.,.,.;T,\gamma) : [0,T] \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is bounded and given by:

$$u(t,x,z;T,\gamma) = \frac{x^{1-\gamma} \left(e^{g_{i.p.}(T-t,\gamma) + \frac{1}{2} h_{i.p.}(T-t,\gamma) z^2} \right)^{\gamma} - 1}{1 - \gamma}, \tag{1.53}$$

and the optimal strategy is

$$\pi^{\star}_{\gamma,t} = \frac{1}{\gamma} \pi^{\star}_{1,t} + h_{i.p.}(T-t,\gamma) Z_t \begin{pmatrix} 1 \\ -c \end{pmatrix}, \quad t \in [0,T]. \tag{1.54}$$

Here, $g_{i.p.}(\cdot,\gamma)$ and $h_{i.p.}(\cdot,\gamma)$ are given by (1.91) and (1.92), respectively, and $(\pi^{\star}_{1,t})_{t \in [0,T]}$ is the optimal strategy for the logarithmic utility given by (1.28). Furthermore,

$$\lim_{T \uparrow T_{nirvana}(\gamma)} u(0,x,z;T,\gamma) = +\infty, \quad \forall (x,z) \in \mathbb{R}^+ \times \mathbb{R}. \tag{1.55}$$

Proof. The proof is similar to the proof of Theorem 1.8. The only difference is that for $\gamma < \gamma_0$, the escape time discriminant $\mathfrak{D}$ in (1.45) is negative and, therefore, (1.44) is ill-posed and its solution is given by Proposition 1.19. The rest of the arguments are similar to the well-posed case and are, thus, omitted. □

We end this section by a discussion on the optimal trading strategies. As seen by (1.40) and (1.54), the form of the optimal allocation for the well-posed and ill-posed cases are similar. It shows that replacing the logarithmic preference with power utility will change the optimal portfolio in two ways:

i) The optimal portfolio is scaled by the factor $1/\gamma$.

ii) The investor also invests in a market-neutral strategy

$$h(T-t,\gamma) Z_t \left(\begin{smallmatrix} 1 \\ -c \end{smallmatrix} \right), \tag{1.56}$$

where depending on the value of γ, h should be replaced by $h_{\text{w.p.}}$ or $h_{\text{i.p.}}$ given by (1.89) and (1.92), respectively.

The following Lemma provides further properties of $h_{\text{w.p.}}$ and $h_{\text{i.p.}}$. Its proof is immediate and, thus, omitted.

Lemma 1.10. *If $\gamma \in [\gamma_0, 1)$ (resp. $\gamma > 1$), then $h_{w.p.}(\cdot;\gamma)$ is positive and increasing (resp. negative and decreasing). In both cases:*

$$\lim_{t\to\infty} h_{w.p.}(t;\gamma) = (1-\gamma)\frac{\|\lambda\|^2}{\gamma^2(\sqrt{\mathfrak{D}}+\kappa)}. \tag{1.57}$$

If $\gamma \in (0, \gamma_0)$, then $h_{i.p.}(\cdot,\gamma) : [0, T_{nirvana}(\gamma)) \to \mathbb{R}$ is positive and increasing.

This lemma explains the characteristics of the market-neutral component (1.56). In particular:

a) If $\gamma > 1$, then $h(T-t,\gamma) < 0$, $t \geq 0$. This implies that the market-neutral term (1.56) is a pairs-trade (i.e. it buys the under-priced stock and sells the over-priced one). Hence, the market-neutral component mitigates the decrease in the pairs-trade which comes from dividing the positions in the stocks by γ.

b) If $\gamma < 1$, then $h(T-t,\gamma) > 0$, $t \geq 0$. This implies that the market-neutral term (1.56) is the opposite of a pairs-trade (i.e. it sells the under-priced stock and buys the over-priced one). Therefore, the market-neutral component mitigates the increase in the pairs-trade resulting from dividing the positions in the stocks by γ.

Furthermore, note that the optimal allocation for power utility is not myopic, as the market-neutral term (1.56) is time-varying.

From a practical point of view, these results are not satisfactory. By the same argument as in the logarithmic case, we deduce that they are not consistent with the practice of pairs-trading. Furthermore, the optimal policies might have unpleasant properties, i.e. blow-ups for nirvana strategies. In the next section, we provide a way to amend these deficiencies and find theoretical ground for pairs-trading.

1.4 Well posedness condition and pairs-trading

Aiming at remedying the deficiencies of the strategies obtained in the previous section, we introduce the following condition.

Condition 1.11. *The following* ***equivalent*** *relationships hold between the market parameters:*

(i) $\alpha_1/\alpha_2 = (\sigma_1^2 - c\sigma_1\sigma_2\rho)/(\sigma_1\sigma_2\rho - c\sigma_2^2)$.

(ii) There exists $\xi \in \mathbb{R}$ *such that* $\alpha = \Sigma\Sigma^\top(1,-c)^\top\xi$.

(iii) $\alpha = \Sigma\Sigma^\top(1,-c)^\top(-\kappa/\sigma_Z^2)$.

In Condition 1.11, the relationships $(i) \Leftrightarrow (ii)$ and $(iii) \Rightarrow (ii)$ are trivial. To see $(ii) \Rightarrow (iii)$, left-multiply (ii) by $(1,-c)$ to obtain

$$-\kappa = (1,-c)\alpha = (1,-c)\Sigma\Sigma^\top(1,-c)^\top\xi = \sigma_Z^2\xi,$$

which yields $\xi = (-\kappa/\sigma_Z^2)$.

The following theorem is the main result of this chapter. It characterises the central role of Condition 1.11.

Theorem 1.12. *The following statements are equivalent:*

(i) Condition 1.11 holds.

(ii) For all $T \in (0,\infty)$*, the Novikov condition holds, i.e.*

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \|\lambda\|^2 Z_s^2 ds\Big)\Big] < \infty, \quad \forall T \in (0,\infty)\,. \tag{1.58}$$

(iii) For all $\gamma \in (0,\infty)$*, the Merton problem (1.21) is well-posed, i.e. there is no nirvana strategy.*

(iv) The optimal strategy (1.21) (which is for power and log utilities) is market-neutral, i.e. it satisfies (1.18).

Proof. **(i)** $\Leftrightarrow$ **(ii)**: Define the function $\varphi : [0,T] \times \mathbb{R} \to \mathbb{R}$ by

$$\varphi(t,z) := \mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_t^T \|\lambda\|^2 (Z_s^{t,z})^2 ds\Big)\Big]. \tag{1.59}$$

Similar to the second step in the proof of Theorem 1.8, we study the related Cauchy problem:

$$\varphi_t - \kappa z \varphi_z + \frac{1}{2}\sigma_Z^2 \varphi_{zz} + \frac{1}{2} z^2 \|\lambda\|^2 \varphi = 0; \quad (t,z) \in [0,T) \times \mathbb{R}, \tag{1.60}$$

with the terminal conditions $\varphi(T,z) = 1$. By Feynman-Kac formula, the (classical) solution of (1.60) satisfies the stochastic representation (1.59) and, conversely, if φ is well defined by (1.59), then it is the (classical) solution of (1.60).
Equation (1.60) is a special case of the PDE solved in Appendix 1.A, with $\mathbf{a} = \lambda$, $\mathbf{b}^\top = (1,-c)\Sigma$, and $\xi = 1$. By Propositions 1.18 and 1.19, it immediately follows that (1.60) has a solution for all T (i.e. it is well-posed) if and only if the corresponding escape criterion discriminant is non-negative, i.e.,

$$\mathfrak{D} = \kappa^2 - \sigma_Z^2 \|\lambda\|^2 \geq 0. \tag{1.61}$$

On the other hand, the Cauchy-Schwarz inequality yields

$$\kappa = -(1,-c)\alpha \leq \|(1,-c)\Sigma\| \, \|\Sigma^{-1}\alpha\| = \sigma_Z \|\lambda\|.$$

Therefore, (1.61) holds if and only if

$$(1,-c)\Sigma\lambda = \|(1,-c)\Sigma\| \|\lambda\|. \tag{1.62}$$

This equation is equivalent to the linear dependence of $\Sigma^\top (1,-c)^\top$ and λ which is, in turn, equivalent to Condition 1.11.(ii).
(i) $\Leftrightarrow$ **(iii)** By Theorems 1.8 and 1.9, the Merton problem is well posed for $\gamma \in (0,\infty)$ if and only if $\gamma_0 = 0$. A simple calculation then shows that $\gamma_0 = 0$ if and only if (1.62) holds, which, as already shown, is equivalent to Condition 1.11.
(i) $\Leftrightarrow$ **(iv)** By (1.40) and (1.54), the optimal strategy for power utility is market-neutral if and only if the optimal strategy $(\pi_{1,t}^\star)_{t\in[0,T]}$ for the logarithmic utility is market-neutral which is, in turn, equivalent to Condition 1.11.(ii). $\square$

Theorem 1.12 provides economic viability for the assumption that the optimal pairs-trading strategy is market-neutral. Indeed, real investors neither attain nirvana strategies nor take infinite positions. The implications of the nirvana solutions are therefore that the parameter combinations producing nirvana solutions do not occur in the real world. This means that either there is no investor with $\gamma < \gamma_0$, or Condition 1.11 holds, which, in turn, implies the market-neutrality assumption.

Another line of argument is through the absence of (risk-free) arbitrage opportunities. A common sufficient condition to impose no-arbitrage assumption is the Novikov condition (1.58), which is shown to be equivalent to the market-neutrality assumption. From a theoretical point of view, it would be interesting to investigate whether Condition 1.11 is also necessary for the market to be arbitrage-free. In general, the Novikov condition is a rather strong condition, and usually not necessary. Nonetheless, the appearance of nirvana solutions when the condition fails, might lead to arbitrage. It is possible to find the necessary and sufficient condition for the absence of arbitrage through the Feller test, see Karatzas and Ruf (2013). We do not pursuit more results in this direction and include it as a possible future research topic.

Remark 1.13. *Condition 1.11 is fundamentally different from conditions (15)-(17) in Theorem 4.4 of Benth and Karlsen (2005) (henceforth, B-K). Indeed, the former is equivalent to the* ***well-posedness*** *of the Merton problem while the latter is a technical* sufficient *condition for the uniform integrability assumption needed for the verification result (Benth and Karlsen, 2005, Theorem 4.2, p. 696).*

In particular, well-posedness, i.e. the boundedness of the value function, which is the main focus of this chapter, is not an issue in B-K. Indeed, as shown in Appendix 1.C below,[2] *the Merton problem in B-K is well-posed even without imposing conditions (15)-(17) therein. This fact has been acknowledged in (Benth and Karlsen, 2005, paragraph 1, p. 689):*

[2]Note that the well-posedness for Schwartz model does not directly follow from the previous results of this chapter, although the arguments are quite similar. The reason is that we assumed that $\sigma_1, \sigma_2 \neq 0$ while to get the Schwartz model, one needs to assume that one of the stocks has zero volatility.

It is worth emphasizing that our candidate solution exists as a classical solution for ***general choices*** *of parameters. This provides us with an upper bound for the value function, since the verification theorem tells us that any classical solution dominates the value function.*

Next, we turn our attention to the optimal strategies if Condition 1.11 holds. Indeed, by Theorem 1.12, the Merton problem is always well-posed under Condition 1.11, and imposing Condition 1.11.(iii) on Theorems 1.7 and 1.8 yields the following result.

Proposition 1.14. *Assume that Condition 1.11 holds. Then, the optimal strategies for both power and logarithmic utilities are given by:*

$$\pi^{\star}_{\gamma,t} = \left(\frac{-\kappa}{\sigma_Z^2}\right) \frac{1 + 1/\sqrt{\gamma}\coth\left(\frac{\kappa}{\sqrt{\gamma}}(T-t)\right)}{1 + \sqrt{\gamma}\coth\left(\frac{\kappa}{\sqrt{\gamma}}(T-t)\right)} Z_t \begin{pmatrix} 1 \\ -c \end{pmatrix}, \tag{1.63}$$

for $\gamma \in (0,\infty)$ *and* $t \in [0,T]$.

Proposition 1.14 provides a solid ground for the practitioners' approach to pairs-trading as explained in Section 1.2. From one hand, the optimal strategy (1.63) is a genuine pairs-trading strategy as defined by Definition 1.4. Indeed, it is a market-neutral strategy, cf. (1.18), and since

$$\left(\frac{-\kappa}{\sigma_Z^2}\right) \frac{1 + 1/\sqrt{\gamma}\coth\left(\frac{\kappa}{\sqrt{\gamma}}(T-t)\right)}{1 + \sqrt{\gamma}\coth\left(\frac{\kappa}{\sqrt{\gamma}}(T-t)\right)} < 0,$$

it always shorts the over-priced stock and longs the under-priced one. On the other hand, the pairs-trading strategy depends solely on the parameters c, κ, and σ_z^2. This, in turn, yields that in order to identify the log-optimal pairs-trade, we only need to specify the partial pricing model (1.17), as is done in practice.

The form of the optimal strategy (1.63) is quite intuitive and deserves attention of its own. The factor $-\kappa/\sigma_z^2$ tells us that the long-short positions should be bigger if the mean-reversion rate κ is bigger, and they should be smaller if the variance rate of the residual, σ_z^2, is larger. Furthermore, the time varying coefficient:

$$f(t,\gamma) := \frac{1 + 1/\sqrt{\gamma}\coth\left(\frac{\kappa}{\sqrt{\gamma}}(T-t)\right)}{1 + \sqrt{\gamma}\coth\left(\frac{\kappa}{\sqrt{\gamma}}(T-t)\right)}$$

has the following properties:

i) $\lim_{\gamma \to 1} f(t,\gamma) = f(t,1) = 1$. Hence, the log-optimal pairs-trade is myopic, as it is expected.

ii) For $\gamma > 1$, the function $f(t,\gamma)$ is decreasing in t and satisfies

$$f(T,\gamma) = \frac{1}{\gamma} < f(t,\gamma) < \frac{1}{\sqrt{\gamma}} = f(-\infty,\gamma) < 1, \quad \text{for } t \in (-\infty, T)\,.$$

Therefore, more risk-averse investors take smaller long-short positions if compared to a log-utility investor. Furthermore, they tend to reduce the size of their pairs-trade as time increases.

iii) For $\gamma < 1$, the function $f(t,\gamma)$ is increasing in t and satisfies

$$1 < f(-\infty,\gamma) = \frac{1}{\sqrt{\gamma}} < f(t,\gamma) < \frac{1}{\gamma} = f(T,\gamma), \quad \text{for } t \in (-\infty, T)\,.$$

This means that more risk-seeking investors take larger long-short positions if compared to a log-utility investor. Furthermore, they tend to increase the size of their pairs-trade as time increases.

We end this section by pointing out that the main result, i.e. the equivalence of Condition 1.11, well-posedness of the Merton problem, and market-neutrality of the optimal strategies is *not* restricted to CRRA utilities. Indeed, as the following theorem shows, these results holds for general utility functions.

Theorem 1.15. *Consider the Merton problem:*

$$(\pi_t^\star)_{t\in[0,T]} := \arg\max_{\pi \in \mathcal{A}} \mathbb{E}\left(U\left(X_T^\pi\right)\right). \tag{1.64}$$

Here, the utility function $U : \mathbb{R}^+ \to \mathbb{R}$ is strictly increasing, strictly concave, continuously differentiable, and satisfies the Inada conditions $\lim_{x\downarrow 0} U'(x) = \infty$ and $\lim_{x\to\infty} U'(x) = 0$. Without loss of generality, we also assume $U(1) = 0$ and $U'(1) = 1$. Finally, we assume the asymptotic elasticity of U to be less than 1, i.e.

$$\limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1. \tag{1.65}$$

Then, the following statements are equivalent:

(i) Condition 1.11 holds.

(ii) For any $T > 0$ and any utility function $U(.)$ (satisfying the assumptions), the Merton problem (1.64) is well-posed.

(iii) For any $T > 0$ and any utility function $U(.)$ (satisfying the assumptions), the optimal strategy is market-neutral.

Proof. **(ii)** $\Rightarrow$ **(i)** and **(iii)** $\Rightarrow$ **(i)**: If Condition 1.11 fails, then, by Theorem 1.12, there exists power utilities for which the Merton problem is ill-posed and the optimal strategies is not market-neutral.

(i) $\Rightarrow$ **(ii)** and **(i)** $\Rightarrow$ **(iii)**: Since the market is complete, one may apply Theorem 2.0 in Kramkov and Schachermayer (1999) to show the regularity of the value function

$$u(t, x, z; T) := \sup_{\pi \in \mathcal{A}} \mathbb{E}\left(U\left(X_T^{\pi,x,z,t}\right)\right).$$

But first, one must check the validity of conditions (2.2), (2.4) and (2.5) therein. Condition (2.4) is the Inada condition (2.21) and Condition (2.2), i.e. existence of a risk-neutral measure, follows from the Novikov condition (2.61). To show condition (2.5), namely, that $u(0, x, z; T) < \infty$ for $(x, z) \in \mathbb{R}^+ \times \mathbb{R}$, note that, by Lemma 2.7 in the next chapter, there exist $\gamma > 0$ such that, for any $T > 0$ and $(x, z) \in \mathbb{R}^+ \times \mathbb{R}$:

$$u(0, x, z; T) \leq u(t, x, z; T, \gamma). \tag{1.66}$$

Recall that $u(., ., .; T, \gamma)$, given by (1.27), is the value function for CRRA utilities. Since Condition 1.11 holds, it follows from Theorem 1.12 that the right side is bounded. Hence, condition (2.5) in Kramkov and Schachermayer (1999) is also valid. It then follows from Theorem 2.0 therein that the value function (2.39) and its dual are bounded and smooth, and that the dual value function is given by

$$v(t, y, z; T) = \mathbb{E}\left(V\left(yY_T^{z,t}\right)\right), \tag{1.67}$$

where V is the convex conjugate of U, and $(Y_s^{t,z})_{s \in [t,T]}$ is given by (1.30).

It only remains to find the optimal strategy. The HJB equation associated with the value function (1.66) is the same as (1.48) with the terminal condition

$u(T, x, z; T) = U(x)$. The same argument as in the Step 3 of the proof of Theorem 1.8, yields the candidate optimal strategy

$$\pi^\star(x, z, u_x, u_{xx}, u_z, u_{zz}, u_{xz}) := \frac{-u_x}{xu_{xx}} z(\Sigma\Sigma^\top)^{-1}\alpha - \frac{u_{xz}}{xu_{xx}}(1, -c)^\top. \tag{1.68}$$

which, in turn, simplifies the HJB equation to

$$u_t - \kappa z u_z + \frac{1}{2}\sigma_Z^2 u_{zz} - \frac{1}{2}\|\lambda\|^2 z^2 \frac{u_x^2}{u_{xx}} - \frac{1}{2}\sigma_Z^2 \frac{u_{xz}^2}{u_{xx}} + \kappa z \frac{u_x u_{xz}}{u_{xx}} = 0; \tag{1.69}$$

$(t, x, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}$, with the terminal condition $u(T, x, z; T) = U(x)$. Applying the Legendre transform, i.e.

$$v(t, y, z) = \sup_x \Big\{ u(t, x, z; T) - xy \Big\},$$

to the simplified HJB equation, yields the dual HJB equation

$$v_t - \kappa z v_z + \frac{1}{2}\sigma_Z^2 v_{zz} + \frac{1}{2}\|\lambda\|^2 z^2 y^2 v_{yy} + \kappa y z v_{zy} = 0; \tag{1.70}$$

$(t, y, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}$, with the terminal condition $v(T, y, z) = V(y)$. By the Feynman-Kac theorem, the solution of (1.70) is the dual value function (1.67). Therefore, the solution of (1.69) is indeed the value function, and hence, (1.68) is the optimal portfolio strategy in feedback form. In particular, by applying Condition 1.11, the optimal strategy is

$$\left(\frac{u_x}{xu_{xx}} \frac{\kappa}{\sigma^2} z - \frac{u_{xz}}{xu_{xx}} \right) \begin{pmatrix} 1 \\ -c \end{pmatrix}$$

which is market-neutral. $\square$

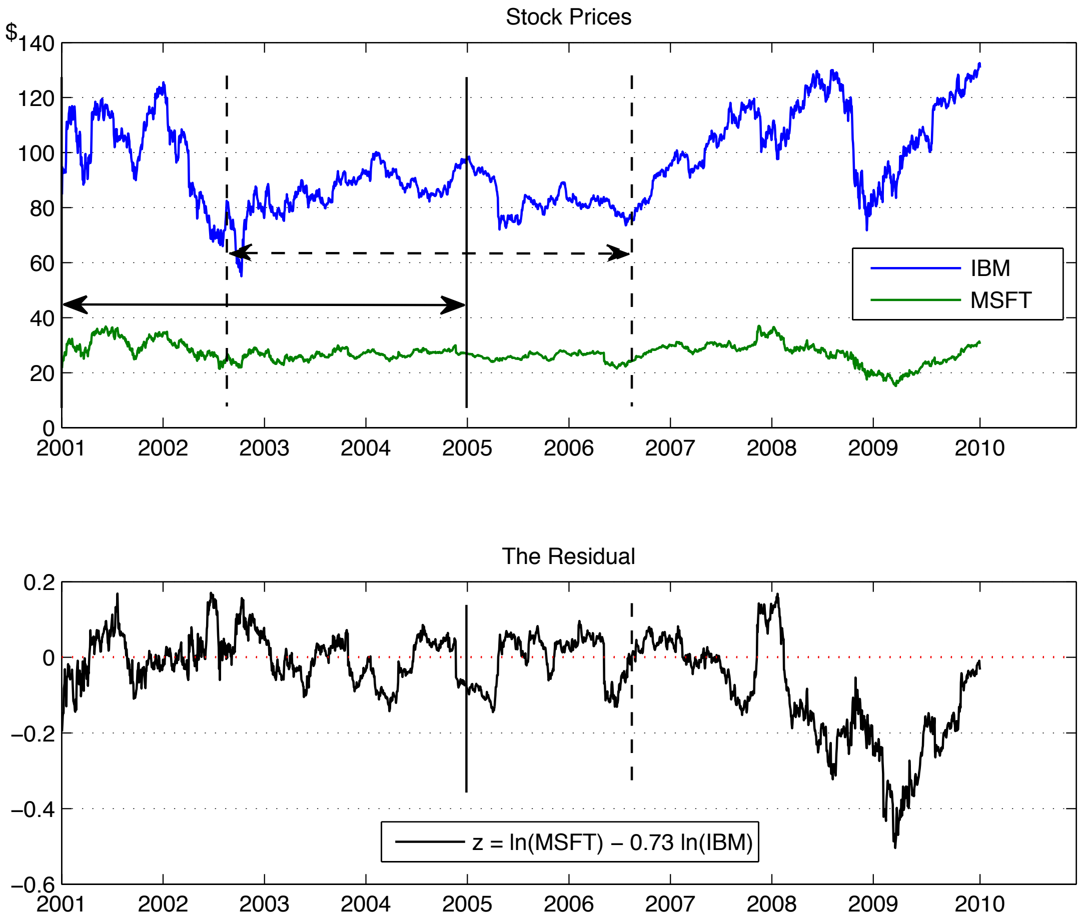


Figure 1.1: Stock prices of Microsoft and IBM (top), and the associated residual (bottom) from Jan 2001 to Dec 2009.

1.5 Numerical example

This section provides an illustration using real market data in order to give some insights on the ideas presented in the previous sections. We use the daily stock prices of Microsoft and IBM for a period of nine years (from Jan. 2, 2001 to Dec. 31, 2009). The data series are extracted from the CRSP [3] database and are adjusted for splits and cash dividends. The top part of Figure 1.1 shows these

[3]Source: CRSP, Center for Research in Security Prices. Booth School of Business, The University of Chicago. Used with permission. All rights reserved. www.crsp.chicagobooth.edu

time series.

To estimate the parameters in (1.1)-(1.3), we follow the Engle-Granger two-step procedure, cf. Engle and Granger (1987), while imposing the well-posedness condition (1.11). More specifically, after establishing that the processes $(\log S_t^1)$ and $(\log S_t^2)$ are $I(1)$ using the augmented Dickey-Fuller test, we use the Phillips-Ouliaris variance ratio (or P_u) and trace statistic (or P_z) tests for testing for cointegration, cf. Phillips and Ouliaris (1990). If the tests imply cointegration, then c and the residual process (Z_t) can be obtained by regressing $\log S_t^1$ over $\log S_t^2$. To find κ and σ_z, we fit a first order autoregressive model to the time series (Z_t) obtained by regression in the previous step. To obtain σ_1, σ_2 and ρ, we regress $\Delta \log S_t^1$ and $\Delta \log S_t^2$ over Z_{t-1}. Then, α_1 and α_2 are calculated according to Condition 1.11, i.e.

$$\begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix} = \frac{-\hat{\kappa}}{\hat{\sigma}_Z^2} \begin{pmatrix} \sigma_1^2 & \hat{\rho}\hat{\sigma}_1\hat{\sigma}_2 \\ \hat{\rho}\hat{\sigma}_1\hat{\sigma}_2 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{c} \end{pmatrix}.$$

To check the performance of the estimation method, we try it on simulated data, generated by an Euler scheme from the SDE given by the price equations (1.1)-(1.3), with the market parameters given in Table 1.1. Figure 1.2 shows ten simulated sample paths for each stock along with the estimated log-spread process.

Table 1.1: Values of market parameters used for simulation.

c	σ_1	σ_2	ρ	α_1	α_2	σ_Z^2	κ	S_0^1	S_0^2
-0.73	0.25	0.24	0.5	-4.49	1.33	0.0494	5.46	21.7	84.8

We calculate out-of-sample test statistics and estimates as follows: On each day from Dec. 12, 2004 to Dec. 12, 2009, we take the last four years of data and run the estimation process discussed above. The results are shown in Figures 1.3 to 1.6. The results suggest that the estimates for c, σ_z, σ_1, σ_2, and ρ are acceptable while the estimates for κ, α_1 and α_2 are occasionally far off.

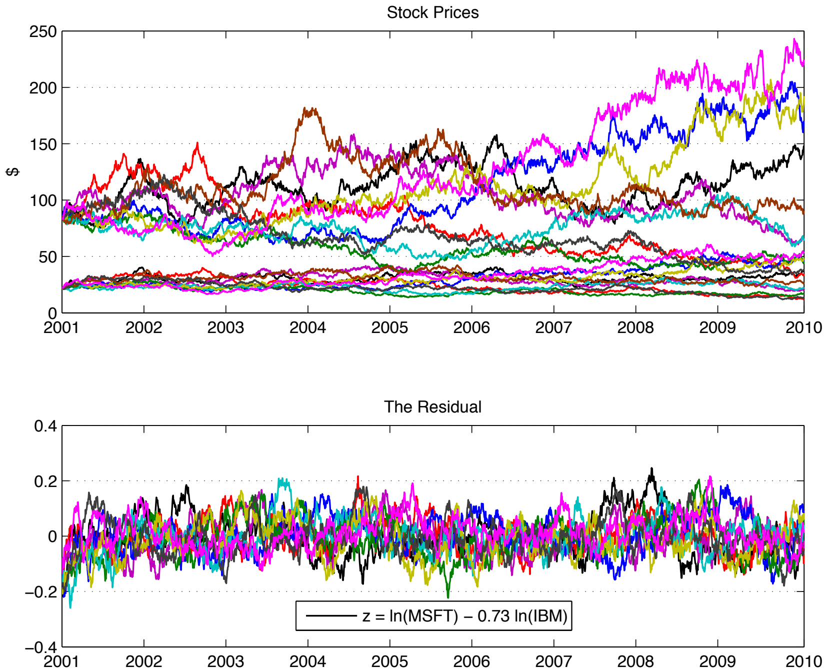


Figure 1.2: Ten sample paths of simulated stock prices of Microsoft/IBM pair (top), and the associated residual (bottom).

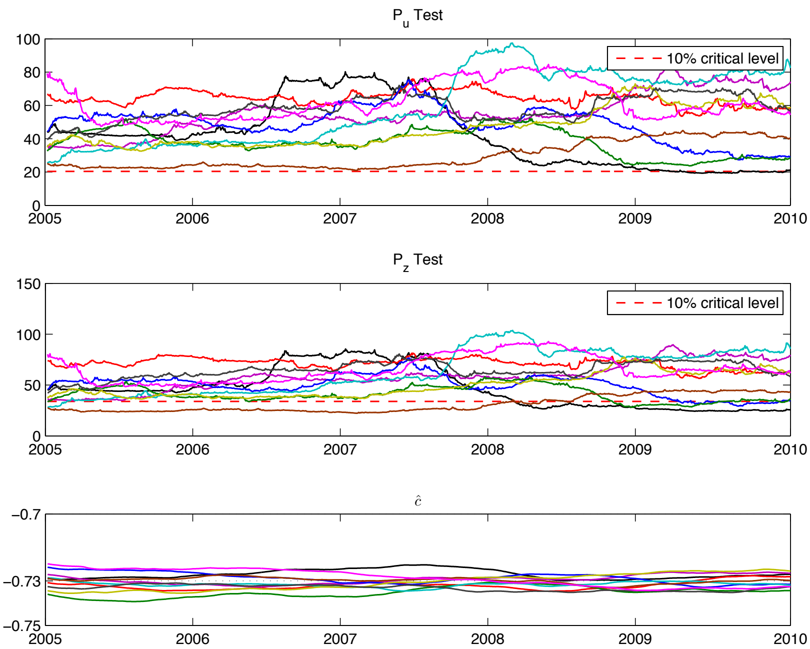


Figure 1.3: Phillips-Ouliaris P_u and P_z cointegration tests, run for ten simulated sample paths, and the estimated cointegration coefficient $\hat{c}$.

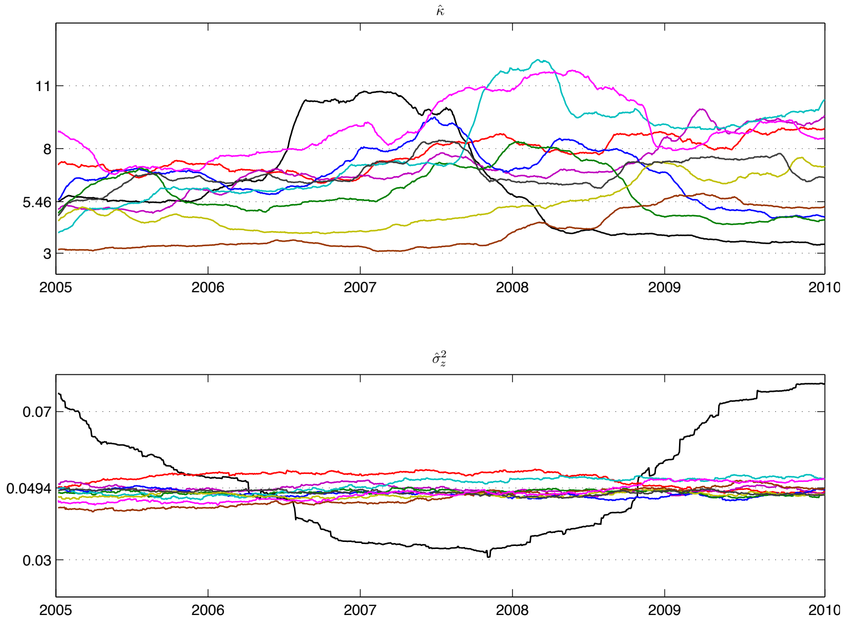


Figure 1.4: Estimation of mean reversion rate κ and variance rate σ_z^2 of the residual, for ten sample paths.

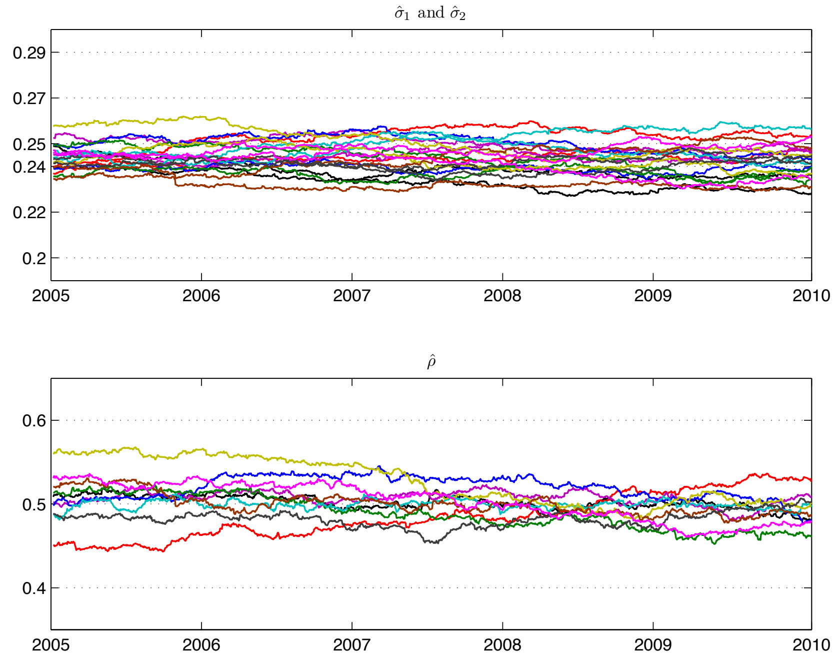


Figure 1.5: Estimation of volatilities σ_1 and σ_2 and correlation ρ, for ten sample paths.

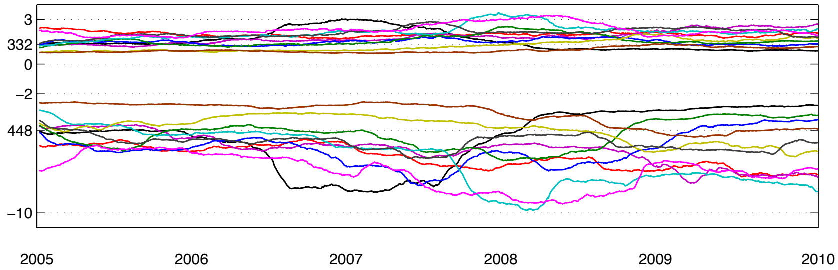


Figure 1.6: Estimation of α for 10 sample paths.

Next, we evaluate the portfolio value for the optimal strategy (1.63) with $\gamma = 1$ (i.e. logarithmic utility) and using simulated data. We assume an initial wealth of \$100. Figure 1.7 shows the log-optimal portfolio value for each simulation using the real parameters (i.e. from Table 1). In all scenarios except one, the portfolio does not lose more than half of its initial value, while all scenarios end up with the terminal wealth of at least \$1000. Figure 1.8 shows the log-optimal portfolio value, for each simulation, by using the out-of-sample estimates.

The results are quite different from the case of using real values of parameters. In five out of ten scenarios, it is observed that the portfolio ends up losing more than 90% of its initial value at some point during the trading horizon. This observation highlights the importance of having good estimates.

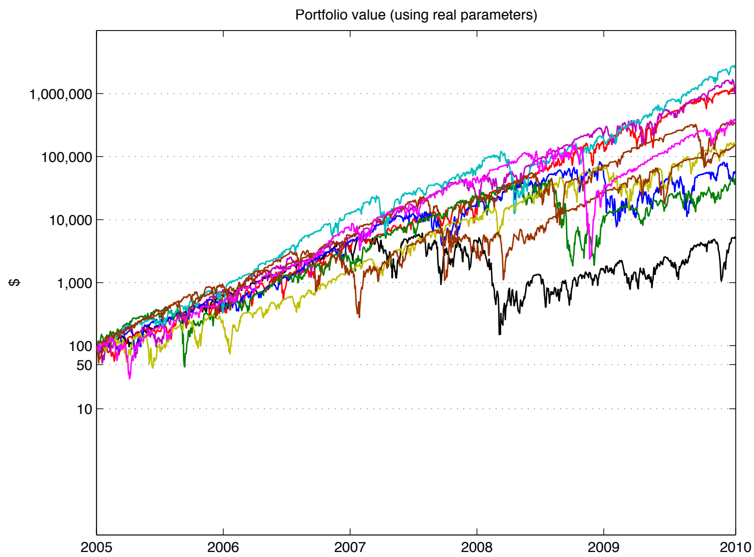


Figure 1.7: Portfolio value of the log-optimal strategy for ten sample paths, using the real values of the parameters.

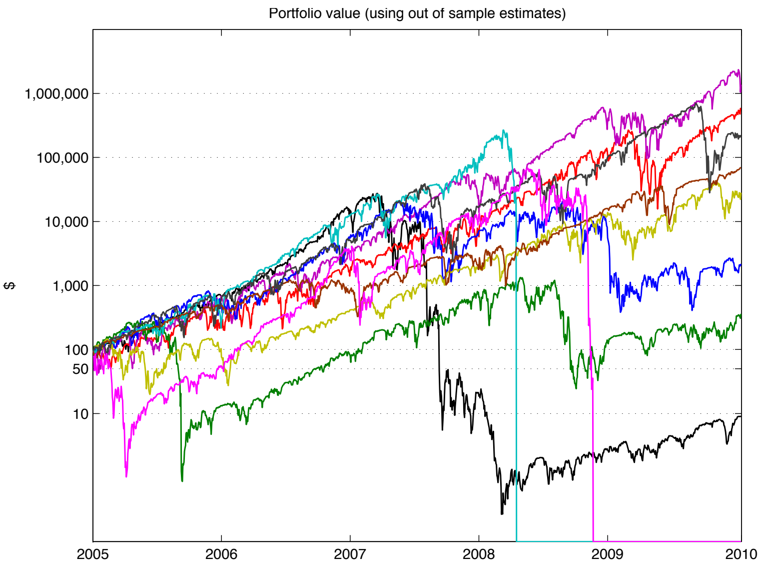


Figure 1.8: Portfolio value of the log-optimal strategy for ten sample paths, using out of sample estimates of the parameters.

We have conducted the same procedure for the real data series of Figure 1.1. The results are shown in Figures 1.9 to 1.12. Note that, the estimator for the constant c is quite robust, while the estimates for σ_z, σ_1, σ_2, and ρ suggest that these parameters vary significantly during the estimation period. Moreover, the test statistics imply that the cointegration relation ceased to exist somewhere during the estimation period.

Figure 1.13 shows the performance of the log-optimal pairs-trading strategy, i.e. (1.63) with $\gamma = 1$, using real data of Figure 1.1. The portfolio weights are calculated by using the out-of-sample estimates discussed above. We consider three scenarios with different assumptions on transaction costs and frequency of trade. In the first scenario, associated with the top solid line, it is assumed that there are no transaction costs and that the investor is adjusting his/her portfolio daily. In the second scenario, showed in the bottom line, it is assumed that the investor is buying with the daily high price and selling with the daily low price. As it can be seen, this strategy is not profitable due to the high transaction cost. Finally, the third scenario (the middle line) refers to the case that there are transaction costs, but the investor adjusts his/her portfolio every two weeks.

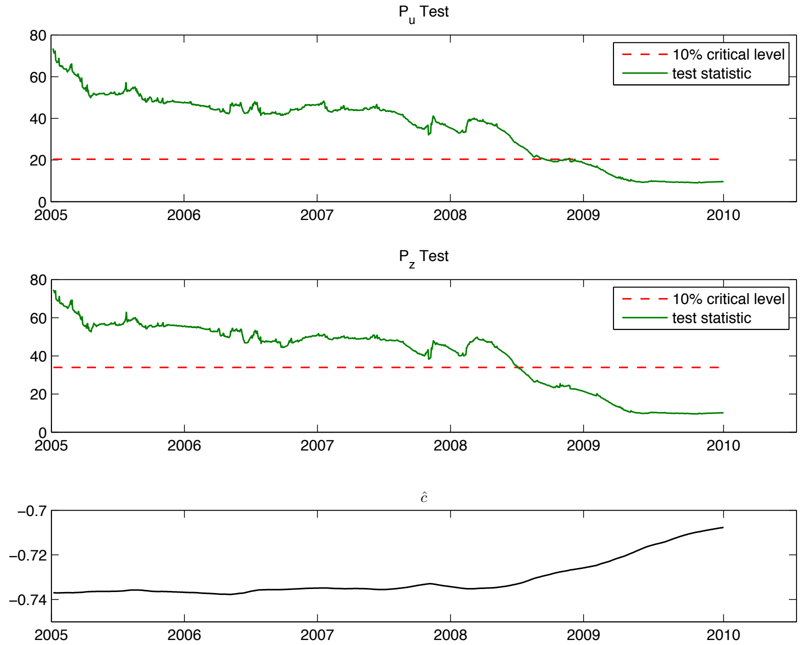


Figure 1.9: Phillips-Ouliaris P_u and P_z cointegration tests, run on a period from Jan 2005 to Dec 2009 (at each day the last four years of data is considered), and the estimated cointegration coefficient $\hat{c}$.

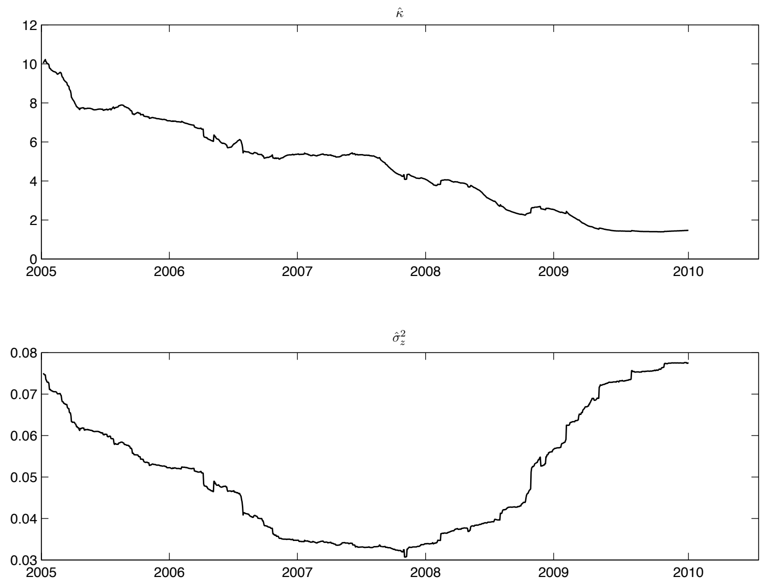


Figure 1.10: Estimation of mean reversion rate κ and variance rate σ_z^2 of the residual.

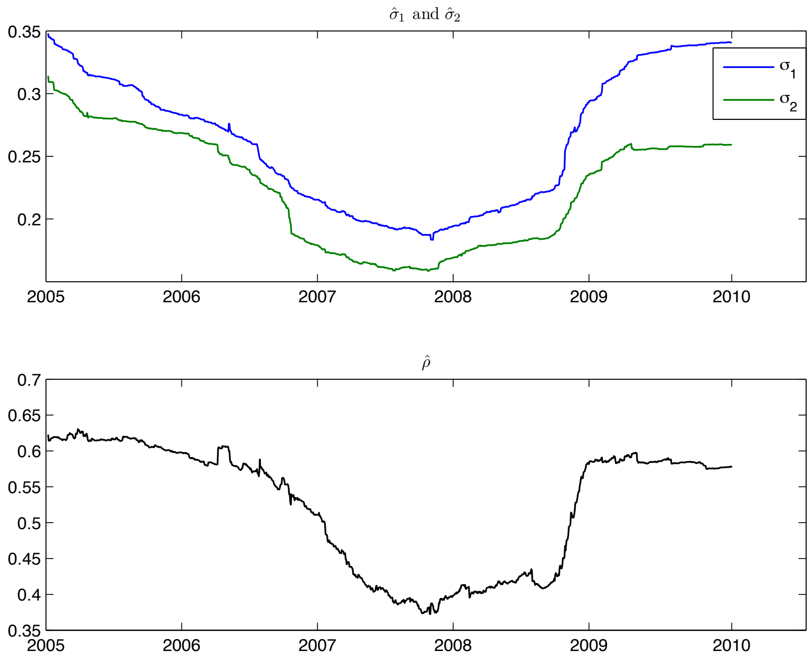


Figure 1.11: Estimation of volatilities σ_1 (for MSFT) and σ_2 (for IBM) and correlation ρ.

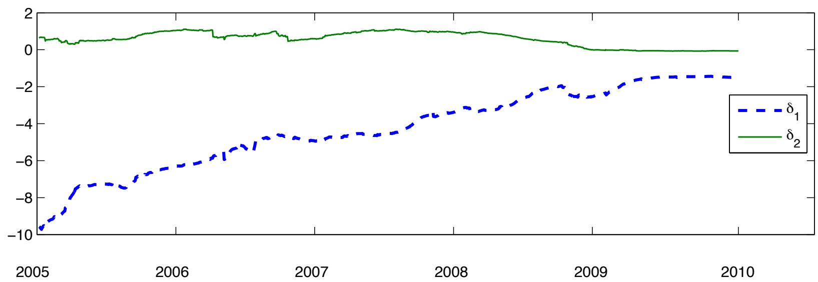


Figure 1.12: Estimation of α.

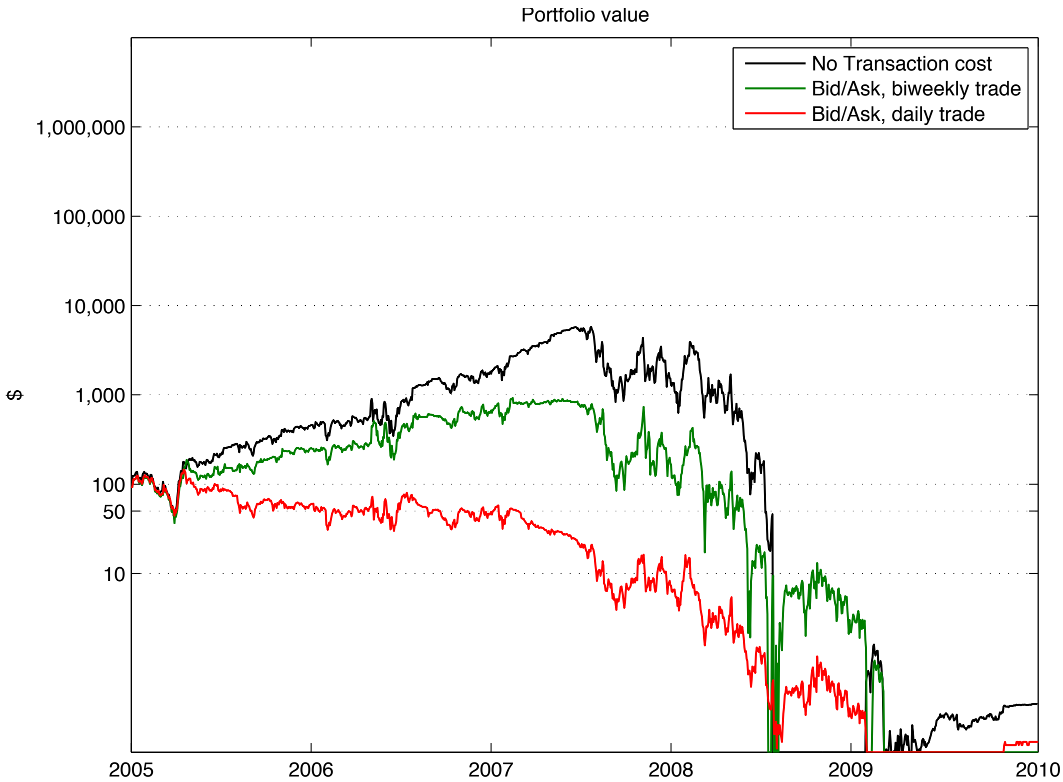


Figure 1.13: Portfolio value of the log-optimal strategy for the MSFT/IBM pair, assuming that: a) there is no transaction cost (black line). b) buying with the daily high price and selling with the daily low price (red line). c) same as (b), but trading biweekly.

1.6 Conclusion

We considered the problem of optimal investment in a market with two cointegrated risky assets, with the motivation of finding a theoretical ground for the so-called pairs-trading strategies. For this, we formulated the classical Merton problem of expected utility of terminal wealth and investigated whether this model supports, in terms of optimal choice, pairs-trading strategies. We focused on the class of homothetic utilities and found that such models do not support, in general, pairs trading policies. Moreover, the optimal policies might have abnormal properties (blow ups for "nirvana solutions").

Aiming at remedying these deficiencies, we introduced an extra condition, i.e. Condition 1.11, on the market coefficients. This condition, which is one of the main contributions of this section, can be obtained and interpreted, in three seemingly unconnected ways. Firstly, it is equivalent to the so-called Novikov condition which guarantees that the market is arbitrage-free. Secondly, it is the necessary and sufficient condition under which the optimal portfolios in the underlying Merton problem indeed justify pairs-trading policies. Thirdly, this condition is, also, necessary and sufficient in order to exclude nirvana solutions and ensure that the Merton problem is well-posed.

We showed that, the optimal pairs-trading strategies obtained by imposing this condition, have intuitive properties and transparent structure, and can be interpreted easily. We concluded with numerical examples including both simulated and real data.

In terms of future research directions, several interesting questions arise. Specifically, a theoretical question is whether Condition 1.11 is also necessary for the market to be arbitrage-free. In more practical directions, one might generalise the market model, with possible extensions including, among others, allowing for several risky assets, stochastic volatility, jump-diffusion stock prices, and regime-switching. As a more challenging task, but very relevant in practice, one might incorporate transaction costs. Other possible research directions include the development of robust estimators for the market parameters and optimal strategies, and statistical tests for the validity of the no-arbitrage condition. Developing these tools will make it possible to conduct empirical studies in order to check the relevance of the results obtained herein.

In the next two chapters, we consider two of these research direction. In particular, Chapter 2 extends the analysis of this chapter to multiple assets, more realistic market setting, and general utility functions. In Chapter 3, we consider spread-trading with futures assets under proportional transaction costs.

1.A Auxiliary PDE

This section provides the explicit solution for the auxiliary PDE

$$\varphi_t + z(\mathbf{a}\cdot\mathbf{b})\varphi_z + \frac{1}{2}\|\mathbf{b}\|^2\varphi_{zz} + \frac{\xi}{2}z^2\|\mathbf{a}\|^2\varphi = 0; \quad (t,z)\in[0,T)\times\mathbb{R}, \tag{1.71}$$

with the terminal condition $\varphi(T, z) = 1$. It is assumed that $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$, $\mathbf{a} \cdot \mathbf{b} < 0$, $\xi \in \mathbb{R}\backslash\{0\}$, and $T \in \mathbb{R}^+$.

The Riccati differential equation

$$h'(t) = 2(\mathbf{a} \cdot \mathbf{b})h(t) + \|\mathbf{b}\|^2 h^2(t) + \xi\|\mathbf{a}\|^2; \quad t \in [0, T), \tag{1.72}$$

with the initial condition $h(0) = 0$, plays a pivotal rule in characterizing the solution of (1.71). Following the terminology of Sasagawa (1982), we define the **escape criterion discriminant** of (1.72) by:

$$\mathfrak{D} := (\mathbf{a} \cdot \mathbf{b})^2 - \xi\|\mathbf{a}\|^2\|\mathbf{b}\|^2. \tag{1.73}$$

The solution of the Riccati equation (1.72) in case $\mathfrak{D} \geq 0$ (resp. $\mathfrak{D} < 0$), is provided by Lemma 1.16 (resp. Lemma 1.17). The proofs of the lemmas are by direct substitution and are left for the reader.

Lemma 1.16. *Assume that $\mathfrak{D} \geq 0$. Then:*

(i) The solution of (1.72) exists for $T = \infty$ and is given by:

$$h_{w.p.}(t) = \begin{cases} \dfrac{\xi\|\mathbf{a}\|^2}{-\mathbf{a}\cdot\mathbf{b} + \sqrt{\mathfrak{D}}\coth\left(t\sqrt{\mathfrak{D}}\right)}; & \text{if } \mathfrak{D} > 0, \\ \dfrac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{b}\|^2}\left(\dfrac{1}{1 - (\mathbf{a}\cdot\mathbf{b})t} - 1\right); & \text{if } \mathfrak{D} = 0. \end{cases} \tag{1.74}$$

The subscript w.p. stands for well-posed.

(ii) $h_{w.p.}(\cdot)$ is uniformly bounded, in particular:

$$|h_{w.p.}(t)| \leq \frac{\left|\mathbf{a}\cdot\mathbf{b} + \sqrt{\mathfrak{D}}\right|}{\|\mathbf{b}\|^2}, \quad t \in [0, \infty). \tag{1.75}$$

(iii) If $\xi > 0$ (resp. $\xi < 0$), then $h_{w.p.}(\cdot)$ is positive and strictly increasing (resp. negative and strictly decreasing). Furthermore,

$$\lim_{t\to\infty} h_{w.p.}(t) = \frac{\xi\|\mathbf{a}\|^2}{\sqrt{\mathfrak{D}} - \mathbf{a}\cdot\mathbf{b}}. \tag{1.76}$$

Lemma 1.17. *Assume that $\mathfrak{D} < 0$. Then:*

(i) (1.72) has a solution iff $T < T_{escape}$, where the escape time T_{escape} is given by:

$$T_{escape} := \frac{1}{\sqrt{-\mathfrak{D}}}\left(\frac{\pi}{2} + \arctan\left(\frac{-\mathbf{a}\cdot\mathbf{b}}{\sqrt{-\mathfrak{D}}}\right)\right). \tag{1.77}$$

The solution is given by

$$h_{i.p.}(t) = -\frac{\sqrt{-\mathfrak{D}}}{\|\mathbf{b}\|^2}\tan\left(\arctan\left(\frac{-\mathbf{a}\cdot\mathbf{b}}{\sqrt{-\mathfrak{D}}}\right) - \sqrt{-\mathfrak{D}}t\right) - \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{b}\|^2}, \tag{1.78}$$

for $0 \leq t \leq T < T_{escape}$.

(ii) $h_{i.p.}(\cdot)$ given by (1.78) escapes to infinity at T_{escape}, in particular:

$$\lim_{T\uparrow T_{escape}} h_{i.p.}(T) = +\infty. \tag{1.79}$$

We are now ready to fully identify the solution of (1.71). Parallel to lemmas (1.16) and (1.17), the following **dichotomy** holds for PDE (1.71):

Well-posed case: If $\mathfrak{D} \geq 0$, then (1.71) admits a solution for any $T \in (0, \infty)$. Proposition 1.18 characterizes this well-posed solution.

Ill-posed case: If $\mathfrak{D} < 0$, then (1.71) has a solution only if $T \in (0, T_{\text{escape}})$, with the escape time T_{escape} given by (1.77). Proposition 1.19 characterizes the ill-posed solution, which escapes to infinity as $T \uparrow T_{\text{escape}}$.

Proposition 1.18. *Assume that $\mathfrak{D} \geq 0$. Then, for any $T \in (0, \infty)$, the solution of PDE (1.71) is:*

$$\varphi_{w.p.}(t, z) = e^{g_{w.p.}(T-t)+\frac{1}{2}z^2 h_{w.p.}(T-t)}; \quad t \in [0, T], \tag{1.80}$$

where

$$g_{w.p.}(t) = -\frac{1}{2}(\mathbf{a}\cdot\mathbf{b})t - \frac{1}{2}\begin{cases} \log\Big(\cosh\left(t\sqrt{\mathfrak{D}}\right) - \\ \qquad \frac{\mathbf{a}\cdot\mathbf{b}}{\sqrt{\mathfrak{D}}}\sinh\left(t\sqrt{\mathfrak{D}}\right)\Big); & \text{if } \mathfrak{D} > 0, \\ \log\Big(1 - (\mathbf{a}\cdot\mathbf{b})t\Big); & \text{if } \mathfrak{D} = 0. \end{cases} \tag{1.81}$$

and $h_{w.p.}(\cdot)$ is given by (1.74).

Proof. Substituting the ansatz

$$\varphi(t,z) = e^{g(T-t)+\frac{1}{2}z^2h(T-t)}, \tag{1.82}$$

where $g, h \in \mathcal{C}^1[0,T]$, into (1.71) yields:

$$\begin{aligned}\varphi(t,z)\Big\{ &- g'(T-t) - \frac{1}{2}z^2 h'(T-t) + z^2(\mathbf{a}\cdot\mathbf{b})h(T-t)\\ &+ \frac{1}{2}\|\mathbf{b}\|^2\Big(z^2h^2(T-t) + h(T-t)\Big) + \frac{c}{2}z^2\|\mathbf{a}\|^2\Big\} =\end{aligned}$$

$$\begin{aligned}\varphi(t,z)\Big\{\frac{1}{2}z^2\Big[-h'(T-t) + 2(\mathbf{a}\cdot\mathbf{b})h(T-t) + \|\mathbf{b}\|^2h^2(T-t) + c\|\mathbf{a}\|^2\Big]&\\ - g'(T-t) + \frac{1}{2}\|\mathbf{b}\|^2h(T-t)\Big\} = 0,&\end{aligned}$$

for all $(t,z) \in [0,T)\times\mathbb{R}$. Therefore, h must satisfy the Riccati equation (1.72) and g must be given by

$$g(t) = \frac{1}{2}\|\mathbf{b}\|^2\int_0^t h(s)ds. \tag{1.83}$$

Since $\mathfrak{D} \geq 0$, Lemma 1.16 applies and $h \equiv h_{\text{w.p.}}$. Finally, $g \equiv g_{\text{w.p.}}$ follows by substituting $h_{\text{w.p.}}$ in (1.83). For the case $\mathfrak{D} > 0$, the following integral is useful:

$$\int_0^t \frac{ds}{1 + a\coth(s)} = \frac{1}{1-a^2}\Big\{t - a\log\Big(\cosh(t) + \frac{1}{a}\sinh(t)\Big)\Big\}, \quad a \in (0,\infty)\backslash\{1\}.$$

□

Proposition 1.19. *Assume that $\mathfrak{D} < 0$ and let T_{escape} be as in (1.77). Then, for $T \in (0, T_{escape})$, the solution of PDE (1.71) is:*

$$\varphi_{i.p.}(t,z;T) = e^{g_{i.p.}(T-t)+\frac{1}{2}z^2h_{i.p.}(T-t)}; \quad 0 \leq t \leq T < T_{escape}, \tag{1.84}$$

where

$$g_{i.p.}(t) = -\frac{1}{2}\Big((\mathbf{a}\cdot\mathbf{b})t + \log\Big|\cos\big(t\sqrt{-\mathfrak{D}}\big) - \frac{\mathbf{a}\cdot\mathbf{b}}{\sqrt{-\mathfrak{D}}}\sin\big(t\sqrt{-\mathfrak{D}}\big)\Big|\Big), \tag{1.85}$$

and $h_{i.p.}(\cdot)$ is given by (1.78). Furthermore, $\varphi_{i.p.}(0,z;T)$ escapes to infinity as $T \uparrow T_{escape}$, i.e.

$$\lim_{T\uparrow T_{escape}} \varphi_{i.p.}(0,z;T) = +\infty, \quad \forall z \in \mathbb{R}. \tag{1.86}$$

Proof. As shown in the proof of Proposition 1.18, assuming the ansatz (1.82) yields that the functions h and g must satisfy (1.72) and (1.83), respectively. Since $\mathfrak{D} < 0$, Lemma 1.17 applies and $h \equiv h_{\text{i.p.}}$. Substituting $h_{\text{i.p.}}(\cdot)$ in (1.83) and using the formula

$$\int_0^t \tan(a - bs)ds = \frac{1}{b}\log\Big|\cos(bt) + \tan(a)\sin(bt)\Big|,$$

yields $g \equiv g_{\text{i.p.}}$. Finally, as $T \uparrow T_{\text{escape}}$, both $h_{\text{i.p.}}(T) \to +\infty$ and $g_{\text{i.p.}}(T) \to +\infty$, therefore $\varphi_{\text{i.p.}}(0, z; T) \to +\infty$ for all $z \in \mathbb{R}$. $\square$

1.B Auxiliary functions

Let

$$\mathfrak{D} = \frac{\sigma_Z^2 \|\lambda\|^2}{\gamma^2}(\gamma - \gamma_0). \tag{1.87}$$

Then, the functions $g_{\text{w.p.}}(\cdot,\cdot)$ and $h_{\text{w.p.}}(\cdot,\cdot)$ of Theorem 1.8 are given by:

$$g_{\text{w.p.}}(t,\gamma) = \frac{1}{2}\frac{\kappa}{\gamma}t - \frac{1}{2}\begin{cases} \log\Big(\cosh\big(t\sqrt{\mathfrak{D}}\big) + \\ \qquad \frac{\kappa}{\gamma\sqrt{\mathfrak{D}}}\sinh\big(t\sqrt{\mathfrak{D}}\big)\Big); & \text{if } \gamma > \gamma_0, \\ \log\Big(1 + \frac{\kappa}{\gamma}t\Big); & \text{if } \gamma = \gamma_0, \end{cases} \tag{1.88}$$

and

$$h_{\text{w.p.}}(t,\gamma) = \begin{cases} \dfrac{(1-\gamma)\|\lambda\|^2}{\kappa\gamma + \gamma^2\sqrt{\mathfrak{D}}\coth\Big(t\sqrt{\mathfrak{D}}\Big)}; & \text{if } \gamma > \gamma_0, \\ \dfrac{\kappa}{\gamma\sigma_Z^2}\left(1 - \dfrac{\gamma}{\gamma + \kappa t}\right); & \text{if } \gamma = \gamma_0. \end{cases} \tag{1.89}$$

In particular, note that

$$h_{\text{w.p.}}(t, 1) = g_{\text{w.p.}}(t, 1) = 0, \quad t \geq 0. \tag{1.90}$$

Furthermore, the functions $g_{\text{i.p.}}(\cdot,\cdot)$ and $h_{\text{i.p.}}(\cdot,\cdot)$ of Theorem 1.9 are given by

$$g_{\text{i.p.}}(t,\gamma) = \frac{\kappa}{2\gamma}t - \frac{1}{2}\log\Big|\cos\big(t\sqrt{-\mathfrak{D}}\big) + \frac{\kappa}{\gamma\sqrt{-\mathfrak{D}}}\sin\big(t\sqrt{-\mathfrak{D}}\big)\Big|, \tag{1.91}$$

and

$$h_{\text{i.p.}}(t,\gamma) = -\frac{\sqrt{-\mathfrak{D}}}{\sigma_Z^2}\tan\left(\arctan\left(\frac{\kappa}{\gamma\sqrt{-\mathfrak{D}}}\right) - \sqrt{-\mathfrak{D}}t\right) + \frac{\kappa}{\gamma\sigma_Z^2}. \tag{1.92}$$

1.C Well-posedness of the Merton problem in Benth and Karlsen (2005)

The goal of this short note is to show that the Merton problem considered in Benth and Karlsen (2005) is well-posed *under general conditions*. In particular, the Merton problem is well-posed even if conditions (15)-(17) in (Benth and Karlsen, 2005, Theorem 4.4, p. 698) *do not hold*. Note that we can re-write the Schwartz model (c.f. (Benth and Karlsen, 2005, Eq. (1), p. 689)) as follows:

$$\frac{dS_t}{S_t} = \left(\frac{\sigma^2}{2} - \alpha Z_t\right)dt + \sigma dW_t, \tag{1.93}$$

where

$$Z_t = \frac{\sigma^2}{2\alpha} - \mu + \ln S_t. \tag{1.94}$$

Then (Z_t) is an O-U process:

$$dZ_t = -\alpha Z_t dt + \sigma dW_t. \tag{1.95}$$

Define, also,

$$\lambda(z) := \frac{\sigma}{2} - \frac{\alpha}{\sigma}z, \tag{1.96}$$

such that the market-price of risk is $(\lambda(Z_t))$ and the state-price density is:

$$\frac{dY_t}{Y_t} = -\lambda(Z_t)dW_t. \tag{1.97}$$

Next, let a trading strategy be represented by (π_t) which is the portfolio weight of the stock. The admissible strategies are defined as usual, and the discounted wealth process is given by:

$$\frac{dX_t^\pi}{X_t^\pi} = \pi_t\left(\frac{\sigma^2}{2} - \alpha Z_t\right)dt + \pi_t\sigma dW_t. \tag{1.98}$$

After introducing the value function $u(t,x,z)$, one can use the duality argument in Step 1 of the proof of Theorem 1.8 to obtain:

$$u(t,x,z) := \sup_{\pi\in\mathcal{A}} \mathbb{E}\left(\frac{\big(X_T^{\pi,x,z,t}\big)^{1-\gamma}-1}{1-\gamma}\right) \leq \frac{\left(\mathbb{E}\Big[(Y_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big]\right)^{\gamma} x^{1-\gamma}-1}{1-\gamma}, \tag{1.99}$$

for $(t,x,z)\in[0,T]\times\mathbb{R}^+\times\mathbb{R}$ and $\gamma\in(0,\infty)$. Therefore, the well-posedness of the Merton problem relies on the boundedness of the expectation on the right side. Let us define,

$$\psi(t,y,z) := \mathbb{E}\Big[(yY_T^{z,t})^{\frac{\gamma-1}{\gamma}}\Big]. \tag{1.100}$$

By the Feynman-Kac theorem, this expectation is related to the Cauchy problem

$$\psi_t + \frac{1}{2}y^2\lambda^2(z)\psi_{yy} - \alpha z\,\psi_z + \frac{1}{2}\sigma^2\psi_{zz} - y\sigma\lambda(z)\psi_{yz} = 0; \tag{1.101}$$

$(t,y,z)\in[0,T)\times\mathbb{R}^+\times\mathbb{R}$, with terminal condition $\psi(T,y,z) = y^{\frac{\gamma-1}{\gamma}}$. We investigate the solution to this PDE. Substituting the ansatz

$$\psi(t,y,z) = y^{\frac{\gamma-1}{\gamma}}\varphi(t,z), \tag{1.102}$$

yields the following PDE for the unknown function $\varphi : [0,T]\times\mathbb{R}\to\mathbb{R}$:

$$\varphi_t + \Big(\frac{(1-\gamma)\sigma^2}{2\gamma} - \frac{1}{\gamma}\alpha z\Big)\varphi_z + \frac{1}{2}\sigma^2\varphi_{zz} + \frac{1-\gamma}{2\gamma^2}\lambda^2(z)\varphi = 0; \quad (t,z)\in[0,T)\times\mathbb{R}, \tag{1.103}$$

with the terminal conditions $\varphi(T,z) = 1$. Using the second ansatz:

$$\varphi(t,z) := \exp\left(f(T-t) + g(T-t)z + \frac{1}{2}h(T-t)z^2\right),$$

yields three ordinary differential equations for the unknown functions $f(.)$, $g(.)$ and $h(.)$. In particular, $h(.)$ satisfies the scalar Riccati equation:

$$h'(t) = -2\frac{\alpha}{\gamma}h(t) + \sigma^2h^2(t) + \frac{(1-\gamma)\alpha^2}{2\gamma^2\sigma^2}, \quad t\in(0,T], \tag{1.104}$$

with the initial condition $h(0) = 0$. This Riccati equation is a special case of equation (1.72), with $\mathbf{a} = (\frac{-\alpha}{\gamma\sigma})$, $\mathbf{b} = \sigma$, and $\xi = 1-\gamma$. The associated *escape criterion discriminant* is:

$$\mathfrak{D} := (\mathbf{a}\cdot\mathbf{b})^2 - \xi\|\mathbf{a}\|^2\|\mathbf{b}\|^2 = \frac{\alpha^2}{\gamma} > 0. \tag{1.105}$$

Since $\mathfrak{D} > 0$, there is no finite escape time for $h(.)$. By using their associated differential equations, it can be shown that $f(.)$ and $g(.)$ also don't have finite escape time. Hence, $\psi(t, y, z)$ of (1.100) is bounded for all T (and bounded (y, z)). Finally, by (1.99), the Merton problem is well-posed.

Chapter 2

On the Market-Neutrality of Optimal Convergence Trading Strategies

In this chapter, we generalise the results of Chapter 1 to the more realistic setting of multiple cointegrated assets, assuming risk factors that effects the asset returns, and general utility functions for investor's preference. These extensions will come at a costs: the closed forms that were the cornerstones of the arguments presented in Chapter 1 are no longer available. Therefore, instead of trying to obtain the generalised counterparts of every result found in the previous chapter, we focus on the main result, i.e to establish the market-neutrality of optimal convergence trading strategies and its connection to the well-posedness of the Merton problem in the multi-asset market setting.

The notion of a market-neutral strategy for two cointegrated assets was defined by (1.18). We need to extend this notion to multiple cointegrated assets. To this end, consider a convergence trading scenario where the mean-reverting signals are given by the $r \times 1$ "log-spread" process

$$Z_t = \eta^\top \log S_t. \tag{2.1}$$

Here, $(S_t^\top) = (S_t^1, \ldots, S_t^n)$ are the asset prices and η is an $n \times r$ matrix such that $\mathrm{rank}(\eta) = r$. Without loss of generality, one may decompose η as follows:

$$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \tag{2.2}$$

where η_1 is an $r \times r$ non-singular matrix and η_2 is an $(n-r) \times r$ matrix.

Definition 2.1. *Let a trading strategy be represented by the $n \times 1$ vector process $(\pi_t^\top) = (\pi_t^1, \dots, \pi_t^n)$, where π_t^i is the portfolio weight of the i-th asset at t, and consider the decomposition:*

$$\pi_t = \begin{pmatrix} \pi_{1,t} \\ \pi_{2,t} \end{pmatrix},$$

for an $r \times 1$ vector process $(\pi_{1,t})$. Then, the trading strategy is market-neutral if:

$$\pi_{2,t} = \eta_2 \eta_1^{-1} \pi_{1,t}, \tag{2.3}$$

or, equivalently:

$$\pi_t = \eta \xi_t, \tag{2.4}$$

for some $r \times 1$ vector process (ξ_t). Here, η, η_1, and η_2 are as in (2.1) and (2.2).

The main idea behind this definition is the same as the one for the bivariate case, i.e. *the profits and losses of market-neutral strategies only depend on the change in the log-spread.*

The following simple discrete-time argument illustrates this point. Let R^p and R^i, $i \in \{1, \dots, n\}$, be the excess returns over time period $[t_1, t_2]$ of the portfolio and the i-th asset, respectively. Let also π^i, $i \in \{1, \dots, n\}$, be the portfolio weight of the i-th asset over the same time interval and Z be the log-spread given by (2.1). Furthermore, consider the decomposition $S^\top = (S_1^\top, S_2^\top)$, $R^\top = (R_1^\top, R_2^\top)$, and $\pi^\top = (\pi_1^\top, \pi_2^\top)$. Here, $R_1^\top = (R^1, \dots, R^r)$ and $R_2^\top = (R^{r+1}, \dots, R^n)$. The components S_1, S_2, π_1, and π_2 are defined similarly. Then, one has that

$$\begin{aligned} R^p = \pi_1^\top R_1 + \pi_2^\top R_2 &\approx \pi_1^\top \Delta \log S_1 + \pi_2^\top \Delta \log S_2 \\ &= \left(\eta_1^{-1} \pi_1\right)^\top \Delta Z + \left(\pi_2 - \eta_2 \eta_1^{-1} \pi_1\right)^\top \Delta \log S_2. \end{aligned}$$

For a market-neutral strategy satisfying (2.3), the last term vanishes, and one has $R^p = (\eta_1^{-1}\pi_1)^\top \Delta Z$. This shows our claim, i.e. that the portfolio return of a market-neutral strategy only depends on the change in the mean-reverting signal.

In Section 2.2 we introduce a rather general continuous time error-correction model (CTECM) which, similar to the model in Liu and Timmermann (2013), consists of tradable risk factors as well as stocks. It is assumed that the stock prices follow a factor model on the long run (i.e. in the equilibrium state), but, short-term deviations from the factor model may occur and are captured by cointegrating relations between the stocks. We provide a thorough discussion on the CTECM market setting. We also introduce various assumptions on the market parameters along with their interpretation and justification. We also formalise the portfolio choice model, which is the Merton investment problem with general utility.

As in the bivariate case of Chapter 1, it turns out that the Merton problem with power utility is closely related to a specific second order PDE. This PDE is studied in Section 2.3, which generalizes the results of Appendix 1.A. We show that the solution of the second order PDE, if it exists, can be expressed in terms of the solution to a particular matrix Riccati differential equation (RDE). The RDE is, in turn, shown to be related to an algebraic Riccati differential equation (ARE). In particular, we show that the RDE has a stabilizing solution (i.e. its limit as time tends to infinity exists and is bounded) if and only if the ARE has a positive definite solution, (cf. Proposition 2.8). Furthermore, we prove the necessary and sufficient condition for the existence of a positive positive definite solution for the ARE, (cf. Theorem 2.11). This seems to be a new result for Riccati equations and can be of independent interest.

Theorem 2.17 then exploits the results obtained for the RDE and the ARE in order to solve the second order PDE. In particular, a sufficient condition for well-posedness of the second order PDE is obtained, and the behaviour of the solution, if the sufficient condition fails, is characterised.

In Section 2.4, we solve the Merton investment problem for power utility. In particular, we introduce Condition 2.20 which generalizes Condition 1.11, i.e. the well-posedness condition for the bivariate case. We show that this condition is a sufficient condition for well-posedness of the Merton problem for all power utilities, and investigate what happens if Condition 2.20 fails. Finally, Section 2.5 provides the main result of this chapter, i.e. Theorem 2.24, which shows the central role of Condition 2.20 in the CTECM market setting and with general terminal utility functions.

2.1 Preliminaries

We use the following notations. By $(H \cdot R)$, we mean the stochastic integral $(\int H_s dR_s)$, where (R_t) is a semimartingale and (H_t) is an R-integrable process. $\mathcal{E}(H)$ denotes the usual stochastic exponential. If X denotes a vector, then X^i is its i-th element. $\mathbb{R}^{n,m}$ is the set of $m \times n$ real matrices and $\mathcal{S}^n$ is the set of symmetric $n \times n$ real matrices. $\mathrm{tr}(M)$ is the trace of a square matrix M, and $|M|$ is its determinant. For an integer $r > 0$, the $r \times r$ matrix of zeros is denoted by 0_r.

2.2 CTECM market setting

The market consists of a riskless asset that pays no interest (otherwise, we work with discounted units), k tradable indices with prices $(I_t) = (I_t^1, \ldots, I_t^k)^\top$ that are proxies for the possibly non-tradable market risk-factors, and n stocks with prices $(S_t) = (S_t^1, \ldots, S_t^n)^\top$.

The **accumulated return processes** of the indices and stocks, denoted by $(R_{I,t}) = (R_{I,t}^1, \ldots, R_{I,t}^k)^\top$ and $(R_{S,t}) = (R_{S,t}^1, \ldots, R_{S,t}^n)^\top$, is defined as the *stochastic logarithm* of the prices, i.e. the unique processes satisfying

$$I_t^i = I_0^i + \int_0^t I_{s-}^i dR_{I,t}^i, \quad R_{I,0}^i = \log I_0^i, \quad i \in \{1, \ldots, k\}, \tag{2.5}$$

and

$$S_t^j = S_0^j + \int_0^t S_{s-}^j dR_{S,t}^j, \quad R_{S,0}^j = \log S_0^j, \quad j \in \{1, \ldots, n\}. \tag{2.6}$$

Remark 2.2. *One may argue that there is no practicality in models based on stochastic logarithm of the prices, because the stochastic logarithms are not observable and, therefore, such a models can not be calibrated. In response, we point out that in most practical scenarios, stochastic logarithm can be obtained through logarithmic prices, which are observable. For example, assuming that the prices are continuous and positive (which is the assumption taken herein, see (2.8) and (2.9)), applying Itô's lemma yields:*

$$\log I_t^i = R_{I,t}^i - \frac{1}{2}\int_0^t \sigma_{i,I,s}^2 ds, \qquad \text{and} \qquad \log S_t^j = R_{S,t}^j - \frac{1}{2}\int_0^t \sigma_{j,S,s}^2 ds, \tag{2.7}$$

where $(\sigma_{i,I,t})$ and $(\sigma_{j,S,t})$ are the volatility of the i-th index and the j-th stock, respectively. In many practical scenarios, the integrated volatility, i.e. the so-called ***realised volatility****, can be estimated, see the seminal work of Barndorff-Nielsen and Shephard (2002). Therefore, the return process can be obtained by adjusting the observable logarithmic prices.*

Next, we consider an $\mathbb{R}^{k+n}$-valued standard Brownian motion $(W_t)_{t\geq 0}$ in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where $(\mathcal{F}_t)$ is the augmented Brownian filtration. Assume that the Brownian motion W is decomposed into the $\mathbb{R}^k$-valued Brownian motion $(W_{I,t})_{t\geq 0}$ and the $\mathbb{R}^n$-valued Brownian motion $(W_{S,t})_{t\geq 0}$, i.e.

$$W_t^\top = \begin{pmatrix} W_{I,t} \\ W_{S,t} \end{pmatrix}, \quad t \geq 0.$$

In particular, we note that $W_I \perp W_S$.

Assumption 2.3. *The indices (I_t) are geometric Brownian motion (GBM), i.e.*

$$R_{I,t} = \log I_0 + \mu_I t + \Sigma_I W_{I,t}, \tag{2.8}$$

and the stock returns follow the CTECM

$$\begin{aligned} dR_{S,t} &= \alpha\eta^\top R_{S,t}dt + \beta dR_{I,t} + \Sigma_S dW_{S,t} \\ &= \left(\alpha\eta^\top R_{S,t} + \beta\mu_I\right) dt + (\beta\Sigma_I, \Sigma_S)\, dW_t. \end{aligned} \tag{2.9}$$

Here, $\mu_I \in \mathbb{R}^k$, $\Sigma_I \in \mathbb{R}^{k,k}$, $\beta \in \mathbb{R}^{n,k}$, $\Sigma_S \in \mathbb{R}^{n,n}$, and $\eta, \alpha \in \mathbb{R}^{n,r}$, for $r \in \{1, \ldots, n-1\}$. Further assumptions will be added as the standing assumptions below.

Equation (2.9) is interpreted as follows: the (instantaneous) stock returns $dR_{S,t}$ are described by a *factor model*, with the indices as the *risk factors* that influence the stock returns according to an $n \times k$ *loading matrix* β. Furthermore, stock prices are allowed to deviate from the factor model through the *error-correction* component $\alpha\eta^\top R_{S,t}dt$ which, as will be discussed below, imposes the *cointegrating relations* $(\eta^\top R_{S,t})$ between the stocks.

The following assumptions are standing throughout and will not be cited in subsequent theorems.

Standing assumption. *The following conditions hold:*

(i) Σ_I *and* Σ_S *are invertible.*

(ii) η *and* α *are of full column rank* $r \in \{1, 2, \ldots, n-1\}$*;* r *is called the* ***cointegration rank****. Furthermore, the* $r \times r$ *matrix* $(\eta^\top \alpha)$ *is* ***c-stable****, i.e. all of its eigenvalues have negative real parts.*

(iii) $(\eta^\top \log S_0)$ *is an* r*-variate Gaussian random vector with mean zero and covariance matrix*

$$V_Z := \int_0^\infty e^{x\eta^\top \alpha}\, \Omega_z\, e^{x\alpha^\top \eta} dx, \tag{2.10}$$

where $\Omega_z := \eta^\top \Sigma_S \Sigma_S^\top \eta$*. Note that the improper integral in (2.10) is convergent because of (ii), c.f. the argument after equation (4.2.5) on page 266 of Doob (1944). It is also assumed that* $(\eta^\top \log S_0)$ *is independent of* $(W_t)_{t\geq 0}$*.*

(iv) $\eta^\top \beta = 0$*.*

For ease of presentation, we introduce the stacked returns $(R_t^\top) := (R_{I,t}^\top,\, R_{S,t}^\top)$, satisfying

$$dR_t = \mu(\eta^\top R_{S,t})dt + \Sigma dW_t, \tag{2.11}$$

where

$$\Sigma := \begin{pmatrix} \Sigma_I & \mathbf{0} \\ \beta\Sigma_I & \Sigma_S \end{pmatrix} \quad \text{and} \quad \mu(z) := \begin{pmatrix} \mu_I \\ \alpha z + \beta \mu_I \end{pmatrix}, \quad z \in \mathbb{R}^r. \tag{2.12}$$

The purpose of assumption (i) is to guarantee, from one hand, the existence of the *market price of risk* process $(\lambda(\eta^\top R_{S,t}))_{t\geq 0}$, where

$$\lambda(z) := \begin{pmatrix} \Sigma_I & \mathbf{0} \\ \beta\Sigma_I & \Sigma_S \end{pmatrix}^{-1} \begin{pmatrix} \mu_I \\ \beta\mu_I + \alpha z \end{pmatrix} = \begin{pmatrix} \Sigma_I^{-1}\mu_I \\ \Sigma_S^{-1}\alpha z \end{pmatrix}, \quad z \in \mathbb{R}^r, \tag{2.13}$$

and, from the other hand, the existence of *state price density* $(Y_t)_{t\geq 0}$ given by

$$Y := \mathcal{E}\left(\lambda\left(\eta^\top R_S\right) \cdot W\right). \tag{2.14}$$

The role of the standing assumptions (ii)-(iv) and the main feature of the CTECM (2.9), is to ensure that the r-variate process $(Z_t)_{t\geq 0}$, given by

$$Z_t := \eta^\top R_{S,t}, \tag{2.15}$$

is a **zero-mean stationary process**. The following proposition proves this fact.

Proposition 2.4. *The process $(Z_t) := (\eta^\top R_{S,t})_{t\geq 0}$ is stationary r-variate Gaussian with mean zero and covariance function*

$$\mathbb{E}\left(Z_s Z_t^\top\right) = \begin{cases} V_Z e^{(t-s)\alpha^\top \eta}, & 0 \leq s \leq t, \\ e^{(s-t)\eta^\top \alpha} V_Z, & 0 \leq t \leq s, \end{cases} \tag{2.16}$$

where V_Z is the stationary covariance matrix of (Z_t) and is given by (2.10).

Proof. This proposition is essentially Theorem 1 of Kessler and Rahbek (2001), but, for the sake of completeness, we provide the proof. Premultiplying (2.9) by $\eta^\top$ yields that Z solves

$$dZ_t = \eta^\top \beta \mu_I + \eta^\top \alpha Z_t dt + \eta^\top \left(\beta \Sigma_I, \Sigma_S\right) dW_t$$

Due to assumption (iv), this equation becomes

$$dZ_t = \eta^\top \alpha Z_t dt + \eta^\top \Sigma_S dW_{S,t}. \tag{2.17}$$

Under assumptions (ii) and (iii), one can apply (Karatzas and Shreve, 1991, Theorem 6.7, p. 357) to obtain the result. $\square$

Proposition 2.4 implies that the accumulated stock returns, $(R_{S,t})_{t\geq 0}$, **cointegrate** with **cointegrating relations** $(Z_t)_{t\geq 0}$. Recall that:

Definition 2.5. *A non-stationary n-variate process (X_t)* ***cointegrates****, or is* ***cointegrated****, if, for some $r \in \{1, 2, \ldots, n-1\}$, there exists an $n \times r$ matrix η such that the r-variate linear combination $(\eta^\top X_t)$ is a stationary process. In this case, η is called the* ***cointegrating matrix****, the r elements of $(\eta^\top X_t)$ are called the* ***cointegrating relations****, and r is called the* ***cointegration rank****.*

The interested reader is referred to Johansen (1995) for further details on cointegration.

With a slight abuse of terminology, we refer to the process (Z_t) as the stocks' **log spreads**, since its components can be obtained by adjusting the linear combination of log-prices. Indeed, by (2.7), one has

$$Z_t := \eta^\top R_{S,t} = \eta^\top \begin{pmatrix} \log S_t^1 \\ \vdots \\ \log S_t^n \end{pmatrix} + \frac{1}{2}\eta^\top \begin{pmatrix} \sigma_{1,S}^2 \\ \vdots \\ \sigma_{n,S}^2 \end{pmatrix} t,$$

where $\sigma_{i,S}$ is the volatility of the i-th stock.

It must be emphasised that assumption (iv), i.e. $\eta^\top \beta = 0$, is not necessary for cointegration, for without it the accumulated stocks returns are still cointegrated. Nonetheless, the assumption is economically significant. It is needed to make the stationary mean of (Z_t) to be zero as well as making (Z_t) independent of the indices (I_t). This means that any deviation of the prices from the factor model:

$$dR_{S,t} = \beta dR_{I,t} + \Sigma_S dW_{S,t},$$

is temporary and is zero on average (i.e. in terms of expected value). In other words, the factor model is the equilibrium state of the market and provides the *long-term risk premium* for the investors; while (αZ_t) represents short-lived deviations from the equilibrium state which provide *short-term risk premium*. Interestingly, we find that the investor's optimal strategy is decomposed accordingly: it consists of two terms. The first represents an investment in the indices to capture the long-term risk premium (e.g. through a mutual fund), while the second is a market-neutral active portfolio strategy involving the stocks to capture the short-term risk premium (e.g. through a hedge fund). See (2.62) and (2.63).

Assumption (iv) also makes sense from a statistical point of view, as estimation methods for cointegrated systems always return demeaned cointegrating relations. One can see this, for example, in the Engle-Granger two-step method or the Johansen procedure. Finally, we emphasise that the standing assumption (iv) is also assumed by Liu and Timmermann (2013), since therein $\eta^\top = (1, -1)$, and $\beta^\top = (b, b)$.

Next, we formalise the portfolio choice model. Consider an agent who invests in the above market with the initial wealth $x_0 > 0$ and over a finite time horizon $t \in [0, T]$. An *admissible investment strategy* is represented by a predictable R_I-integrable $\mathbb{R}^k$-valued process $(\pi_{I,t})_{t\in[0,T]}$ which contains the *proportions* of agent's wealth invested in the indices, and a predictable R_S-integrable $\mathbb{R}^n$-valued process $(\pi_{S,t})_{t\in[0,T]}$ which contains the *proportions* of her wealth invested in the stocks. The set of all admissible strategies is denoted by $\mathcal{A}$.

Later on, we might parametrise the set of admissible strategies based on specific parameters, for example $\mathcal{A}(T)$ represents the set of admissible strategies for investment period $[0, T]$, and so forth.

For an admissible strategy $(\pi_t^\top)_{t\in[0,T]} = (\pi_{I,t}^\top, \pi_{S,t}^\top)_{t\in[0,T]}$, the agent's wealth process $(X_t^\pi)_{t\in[0,T]}$ is given by the stochastic exponential:

$$X^\pi := x_0\, \mathcal{E}\left(\pi \cdot R\right) = x_0\, \mathcal{E}\left(\pi_I \cdot R_I + \pi_S \cdot R_S\right). \tag{2.18}$$

In particular, the continuity of the return processes R_I and R_S implies that the wealth process is positive, and there is no need to include non-negativity of wealth processes in the definition of admissible strategies. For future reference, we point out that the wealth process $(X_t^\pi)_{t\in[0,T]}$ satisfies the SDE

$$dX_t^\pi = X_t^\pi \pi_t^\top \Sigma \Big(\lambda(Z_t)dt + dW_t\Big), \tag{2.19}$$

with $X_0^\pi = x_0$.

Next, we consider the Merton investment problem in this market setting:

Assumption 2.6. *We assume that the agent seeks to maximise her expected utility of wealth at T, and tries to implement her optimal strategy defined as:*

$$(\pi_t^\star)_{t\in[0,T]} := \arg\max_{\pi\in\mathcal{A}} \mathbb{E}\left(U\left(X_T^\pi\right)\right), \tag{2.20}$$

assuming that the maximum is finite and is attained, otherwise the optimal strategy does not exist. Here, the utility function $U : \mathbb{R}^+ \to \mathbb{R}$ is strictly increasing, strictly concave, continuously differentiable, and satisfies the Inada conditions:

$$\lim_{x\downarrow 0} U'\left(x\right) = \infty, \quad \text{and} \quad \lim_{x\to\infty} U'\left(x\right) = 0. \tag{2.21}$$

Without loss of generality, we also assume $U(1) = 0$ and $U'(1) = 1$. Finally, we assume the asymptotic elasticity of U to be less than 1, i.e.

$$\limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1. \tag{2.22}$$

All the assumptions on the utility function U are standard except (2.22) which was originally introduced by Kramkov and Schachermayer (1999) to handle market incompleteness. Regardless of the completeness of our market setting, (2.22) is adapted herein solely to provide the following result, which is a modified version of (Kramkov and Schachermayer, 1999, Lemma 6.5) and facilitates the proof of Theorem 2.24.

Lemma 2.7. *A utility function U satisfying the conditions of Assumption 2.6, can be dominated by a power utility, i.e. there exists $p < 1$ such that*

$$U(x) \leq \frac{x^p - 1}{p}, \quad \textit{for} \quad x > 0. \tag{2.23}$$

Proof. The proof is rather long and is included in Appendix 2.A. □

2.3 Solution of an auxiliary second order linear PDE

As it will be shown later, the special case of power utility plays a central role in solving the Merton problem (2.20) for which the general utility is assumed. The Merton problem with power utility is, in turn, closely related to a specific second order linear PDE. To facilitate the exposition, we chose to study this PDE first. We note that some of the results below are, to the best of our knowledge, new, and generalize some of the one-dimensional results of Appendix 1.A.

To this end, consider the equation

$$\varphi_t + z^\top \alpha^\top \eta \varphi_z + \frac{1}{2}\mathrm{tr}\left(\eta^\top \Sigma \Sigma^\top \eta \varphi_{zz}\right) + \frac{\xi}{2}\left(d^2 + z^\top \alpha^\top (\Sigma\Sigma^\top)^{-1} \alpha z\right)\varphi = 0; \tag{2.24}$$

$(t, z) \in [0, T) \times \mathbb{R}^r$, with terminal condition $\varphi(T, z) = 1$.

It is assumed that $\alpha, \eta \in \mathbb{R}^{n,r}$, $n > r$, $\mathrm{rank}(\alpha) = \mathrm{rank}(\eta) = \mathrm{rank}(\eta^\top \alpha) = r$, $\eta^\top \alpha$ is c-stable (i.e. all its eigenvalues have negative real parts), $\Sigma \in \mathbb{R}^{n,n}$, $|\Sigma| \neq 0$, $\xi \in (-\infty, 1]$, $d \in \mathbb{R}$, and $T \in (0, \infty)$.

As it is argued in the proof of Theorem 2.17 below, the solution of (2.24) is closely related to the solution of the *matrix Riccati differential equation (RDE)*:

$$H'(t) = H(t)\eta^\top \Sigma\Sigma^\top \eta H(t) + H(t)\eta^\top \alpha + \alpha^\top \eta H(t) + \xi \alpha^\top (\Sigma\Sigma^\top)^{-1}\alpha, \qquad (2.25)$$

with the initial condition $H(0) = H_0 \in \mathcal{S}^r$. It is known, by classical existence results for ordinary differential equations, that the above problem (2.25) has a unique local solution for $t \in [0, \varepsilon]$, for some $\varepsilon > 0$. Henceforth, we denote the unique solution of (2.25), by $H(\cdot, H_0)$. We only consider the case $H_0 = 0_r$. Furthermore, our main interest is the existence of a **stabilising** solution, i.e. when $H(\cdot, 0_r)$ is defined over $[0, \infty)$ and $\lim_{t\uparrow\infty} H(t, 0_r)$ exists.

If $\xi \leq 0$, then the constant term $\xi\alpha^\top(\Sigma\Sigma^\top)^{-1}\alpha \leq 0$, while the coefficient of the quadratic term $\eta^\top\Sigma\Sigma^\top\eta > 0$. This is a well-studied case of *linear-quadratic regulator problems*. It has been shown in various studies that the solution $H(\cdot, 0_r)$ is stabilising, see, for example, Theorem 16.4.3 and Corollary 16.4.5 on pp. 361-362 of Lancaster and Rodman (1995). The *stabilisability* condition on the pair $(\eta^\top\alpha, \eta^\top\Sigma\Sigma^\top\eta)$ which is required for applying those results is established in the proof of Lemma 2.32 below.

Note, however, that these well known results are *not* enough for our arguments in the next sections. In particular, we need to consider the case $\xi \in (0, 1]$. For this case, the stabilisability property of $H(\cdot, 0_r)$ is more delicate in comparison to the well behaved case $\xi \leq 0$. Indeed, it is well known that RDE (2.25) may have **finite escape time**, i.e. its solution may exist only up to a finite time. See, for example, Martin (1981) and Sasagawa (1982).

Proposition 2.8 below provides the necessary and sufficient condition for the stabilisation property of the solution of (2.25), when $\xi \in (0, 1]$. Specifically, it is shown that such solution exists if and only if the **algebraic Riccati equation (ARE)**:

$$H\eta^\top\Sigma\Sigma^\top\eta H + H\eta^\top\alpha + \alpha^\top\eta H + \xi\alpha^\top(\Sigma\Sigma^\top)^{-1}\alpha = 0, \qquad (2.26)$$

has a positive semidefinite solution. Henceforth, we denote by $\mathcal{R}(\xi)$ the set of symmetric solutions of (2.26), i.e.

$$\mathcal{R}(\xi) := \Big\{ H \in \mathcal{S}^r : H\eta^\top\Sigma\Sigma^\top\eta H + H\eta^\top\alpha + \alpha^\top\eta H + \xi\alpha^\top(\Sigma\Sigma^\top)^{-1}\alpha = 0 \Big\}. \quad (2.27)$$

Proposition 2.8. *For $\xi \in (0,1]$, the solution $H(\cdot, 0_r)$ is stabilising, i.e.*

$$H(\infty, 0_r) := \lim_{t\uparrow\infty} H(t, 0_r)$$

exists, if and only if there exists $H_0 \in \mathcal{R}(\xi)$ such that $H_0 \geq 0$.

The following two comparison lemmas are needed for the proof of Proposition 2.8. We present them without proof.

Lemma 2.9. *(Abou-Kandil et al., 1994, Lemma 1, pp. 1632)*
Suppose $H(t, H_0)$ exists for $t \in [0,T]$. Then, the inequality

$$H_0\eta^\top\Sigma\Sigma^\top\eta H_0 + H_0\eta^\top\alpha + \alpha^\top\eta H_0 + \xi\alpha^\top(\Sigma\Sigma^\top)^{-1}\alpha \geq 0, \quad (resp.\ \leq 0),$$

implies $H'(t, H_0) \geq 0$ (resp. $H'(t, H_0) \leq 0$) for all $t \in [0,T]$.

Lemma 2.10. *(Freiling et al., 1996, Theorem 2.1, pp. 293)*
For $i \in \{1,2\}$, assume that the matrix functions $A_i, Q_i, S_i : [0,T] \to \mathbb{R}^{r,r}$ are integrable, $S_i(t), Q_i(t) \in \mathcal{S}^r$ for $t \in [0,T]$,

$$\begin{pmatrix} Q_1(t) & A_1^\top(t) \\ A_1(t) & S_1(t) \end{pmatrix} \leq \begin{pmatrix} Q_2(t) & A_2^\top(t) \\ A_2(t) & S_2(t) \end{pmatrix}, \quad \textit{for } t \in [0,T],$$

and that there exists $K_i : [0,T] \to \mathbb{R}^{r,r}$, $i \in \{1,2\}$, such that:

$$K_i'(t) = K_i(t)S_i(t)K_i(t) + K_i(t)A_i(t) + A_i^\top(t)K_i(t) + Q_i(t), \quad t \in [0,T].$$

Then, $K_1(0) \leq K_2(0)$ (resp. $<$) implies that $K_1(t) \leq K_2(t)$ (resp. $<$) for all $t \in [0,T]$.

We are now ready to provide the proof of Proposition 2.8.

Proof of Proposition 2.8. We first show the sufficiency part, namely, that

$$(\exists H_0 \in \mathcal{R}(\xi) : H_0 \geq 0) \Rightarrow \exists H(\infty, 0_r).$$

By the existence results for ordinary differential equations, $H(t, 0_r)$ exists locally over some interval $[0, \varepsilon]$, $\varepsilon > 0$. Let $\delta \in (0, \infty]$ be the supremum of such ε. It is known that the solution of RDE (2.25) may escape to infinity at a finite time, which is known as the finite escape time phenomenon for the Riccati differential equation.

We show that the there is no finite escape time, i.e. $\delta = \infty$. We argue by contradiction, assuming that $\delta < \infty$. Since both $H(\cdot, H_0)$ and $H(t, 0_r)$ satisfy RDE (2.25) and $H(0, 0_r) := 0_r \leq H_0 =: H(0, H_0)$, Lemma 2.10 yields that $H(t, 0_r) \leq H(t, H_0) = H_0$, for $t \in [0, \delta)$. On the other hand, since $H'(0, 0_r) = \xi\alpha^\top(\Sigma\Sigma^\top)^{-1}\alpha > 0$, Lemma 2.9 yields that $H'(t, 0_r) \geq 0$, $t \in [0, \delta)$. It then follows that for all $x \in \mathbb{R}^r$, the function $f(t) := x^\top H(t, 0_r)x$ is increasing and bounded above by $x^\top H_0 x$, and therefore $\lim_{t\uparrow\delta} x^\top H(t, 0_r)x$ exists. Note that $H(\cdot, 0_r)$ must be symmetric. Indeed, if $H(\cdot, 0_r)$ satisfies (2.25) over the interval $[0, \delta)$, then so does $H(\cdot, 0_r)^\top$, and by the uniqueness of the solution of (2.25) one has $H(\cdot, 0_r)^\top = H(\cdot, 0_r)$. Now, since $\lim_{t\uparrow\delta} x^\top H(t, 0_r)x$ exists and $H(\cdot, 0_r)$ is symmetric, it follows that the elementwise limit $\lim_{t\uparrow\delta} H_{i,j}(t, 0_r)$ exists. We then may define $H(\delta, 0_r) = \lim_{t\uparrow\delta} H(t, 0_r)$. But, then, the RDE with the initial value $H(\delta, 0_r)$ at δ has a unique solution over some interval $[\delta, \delta + \varepsilon)$, $\varepsilon > 0$. If follows that one may extend the solution to $[0, \delta + \varepsilon)$, which contradicts the maximality of δ. Thus, there is no finite escape time.

Finally, since the solution $H(\cdot, 0_r)$ defined over the interval $[0, \infty)$ is monotone (in the positive-definite sense) and bounded above by H_0, one may repeat the limit argument to show the existence of $H(\infty, 0_r)$.

Next, we show the necessity part, i.e. that

$$\exists H(\infty, 0_r) \Rightarrow (\exists H_0 \in \mathcal{R}(\xi) : H_0 \geq 0).$$

As shown in the previous step, $H(\cdot, 0_r)$ must be symmetric. Furthermore, Lemma 2.9 still applies and $H'(t, 0_r) \geq 0$, $t \in [0, \infty)$. Therefore, $H(\infty, 0_r) \geq 0$. Finally, taking the limit as $t \to \infty$ in RDE (2.25) yields that $H(\infty, 0_r) \in \mathcal{R}(\xi)$. $\square$

The following theorem provides our main result on the existence of symmetric solutions for ARE (2.26), and is, to our knowledge, new.

Theorem 2.11. *Consider ARE (2.26) and recall that $\mathcal{R}(\xi)$ is defined by (2.27). The following statements are true:*

(i) There exists $\xi_0 \in [0,1]$ such that $\mathcal{R}(\xi) \neq \emptyset$ for all $\xi \leq \xi_0$, and $\mathcal{R}(\xi) = \emptyset$ for all $\xi \in (\xi_0, 1]$.

(ii) $\mathcal{R}(\xi) \neq \emptyset$ for all $\xi \leq 1$ (i.e. $\xi_0 = 1$), if and only if there exists a real and symmetric matrix $H > 0$ such that $\alpha = -\Sigma\Sigma^\top \eta H$.

Remark 2.12. *By considering RDE (2.25), we have essentially studied the corresponding **deterministic linear quadratic (LQ) control problem***

$$\begin{aligned} &\textit{Minimize:} \quad J\Big(u(.)\Big) := \int_0^T \Big[x(t)^\top Q x(t) + u(t)^\top R u(t)\Big] dt, \\ &\textit{subject to:} \quad \dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0 \in \mathbb{R}^r, \end{aligned} \tag{2.28}$$

where

$$A = \eta^\top \alpha, \quad B = \eta^\top, \quad R = -(\Sigma\Sigma^\top)^{-1}, \quad \textit{and} \quad Q = \xi \alpha^\top (\Sigma\Sigma^\top)^{-1}\alpha. \tag{2.29}$$

*In particular, we have shown that this family of LQ control problems, **which has negative definite control weighting matrix in the cost function** (i.e. $R < 0$), can be well-posed. Indeed, by combining Proposition 2.8 and Theorem 2.11, RDE (2.25) and, hence, the LQ control problem (2.28) are well-posed if and only if there exists a real and symmetric matrix $H > 0$ such that $\alpha = -\Sigma\Sigma^\top \eta H$. To the best of our knowledge, such example of well-posed deterministic LQ control problems with $R < 0$ has not been provided in the literature before.*

To be more specific, consider the generalized Riccati differential equation (GRDE):

$$\begin{cases} \dot{P}(t) + A(t)^\top P(t) + P(t)A(t) \\ \qquad + C(t)^\top P(t) C(t) + Q(t) \\ \qquad - [P(t)B(t) + C(t)^\top P(t) D(t) + L(t)] \\ \qquad \times [R(t) + D(t)^\top P(t) D(t)]^{-1} \\ \qquad \times [P(t)B(t) + C(t)^\top P(t) D(t) + L(t)]^\top = 0, \quad t \in [0,T), \\ P(T) = M, \end{cases} \tag{2.30}$$

with the extra constraint

$$R(t) + D(t)^\top P(t) D(t) > 0, \quad a.e.\, t \in [0,T]. \tag{2.31}$$

As shown in Ait Rami et al. (2001) and Ait Rami et al. (2002), the GRDE corresponds to the stochastic linear quadratic (LQ) control problem:
Minimize

$$\begin{aligned} J\Big(u(.)\Big) := \mathbb{E} \int_0^T \Big[x(t)^\top Q(t) x(t) + 2x(t)^\top L(t) u(t) + u(t)^\top R(t) u(t)\Big] dt \\ + \mathbb{E}\Big[x(T)^\top M x(T)\Big], \end{aligned}$$

subject to

$$\begin{aligned} dx(t) &= \Big[A(t)x(t) + B(t)u(t)\Big]dt + \Big[C(t)x(t) + D(t)u(t)\Big] dW_t, \\ x(0) &= x_0 \in \mathbb{R}^r. \end{aligned}$$

The Riccati differential equation (2.25) is a special case[1] *of GRDE (2.30) with parameters given by (2.29) and*

$$M = 0, \quad C \equiv 0, \quad D \equiv 0, \quad L \equiv 0.$$

Therefore, RDE (2.25) corresponds to the LQ control problem (2.28).

As mentioned before, the control weighting matrix in the cost function of the LQ control problem (2.28) is negative definite, i.e. $R = -(\Sigma\Sigma^\top)^{-1} < 0$. *On the other hand, the assumption* $R > 0$ *has been taken for granted in all of the studies on deterministic LQ models (see, among others, Wonham (1968), Bensoussan (1982) and Davis (1977)). It is well-known that if* $R > 0$ *fails, then the LQ control problem can be meaningless, in that "the-larger-the-better" policy applies. Nonetheless, to the best of our knowledge, there is no extensive study for well-posedness of deterministic LQ control problem when* $R < 0$ *and* $D \equiv 0$.

More recently, Chen et al. (1998) and Ait Rami et al. (2001) studied the stochastic LQ problems with $D \neq 0$. *They have provided the necessary and sufficient condition for well-posedness of GRDE (2.30)-(2.31) when there is no restriction on the definiteness of* $R(t)$. *Note that their analysis, although quite related to*

[1]Note that (2.25) is forward in time, while (2.30) is backward.

our problem, does not directly apply to RDE (2.25). Indeed, their results depend on the extra constraint (2.31), which for the deterministic LQ problem (i.e. $D = 0$) boils down to $R > 0$. In particular, their well-posedness conditions do not apply to the LQ control problem (2.28) where $D \equiv 0$ and $R < 0$, c.f. (Chen et al., 1998, Theorem 4.6, p. 1696) and (Ait Rami et al., 2001, Theorem 3.3, p. 433).

The proof of Theorem 2.11 will be given after establishing three preliminary results.

Lemma 2.13. *(Comparison) $\mathcal{R}(\xi) \neq \emptyset$ if and only if the algebraic Riccati inequality:*

$$\underline{H}\eta^\top \Sigma\Sigma^\top \eta \underline{H} + \underline{H}\eta^\top \alpha + \alpha^\top \eta \underline{H} + \xi \alpha^\top (\Sigma\Sigma^\top)^{-1}\alpha \leq 0, \tag{2.32}$$

has a real and symmetric solution.

Proof. This lemma follows from Theorem 9.1.1, p. 232 of Lancaster and Rodman (1995), modified according to section 9.5 therein to accommodate the real case. In particular, note that the required assumption of **c-stabilisability** of $(\eta^\top \alpha, \eta^\top \Sigma\Sigma^\top \eta)$ is satisfied: $(\eta^\top \alpha, \eta^\top \Sigma\Sigma^\top \eta)$ is c-stabilisable if it is **controllable**[2] (Lancaster and Rodman, 1995, Theorem 4.4.2, page 90), and $(\eta^\top \alpha, \eta^\top \Sigma\Sigma^\top \eta)$ is controllable if and only if $(\eta^\top \alpha, \eta^\top)$ is controllable (Lancaster and Rodman, 1995, corollary 4.1.3, page 85). Finally, the controllability of $(\eta^\top \alpha, \eta^\top)$ follows from:

$$\operatorname{rank}\Big(\big[\eta^\top \,|\, \eta^\top \alpha \eta^\top \,|\, \dots \,|\, (\eta^\top \alpha)^{r-1}\eta^\top\big]\Big) = \operatorname{rank}(\eta^\top) = r. \qquad \square$$

Lemma 2.14. *(Perturbation) If there exist a sequence $\{\xi_m\}$ such that $\xi_m \to \xi$ and $\mathcal{R}(\xi_m) \neq \emptyset$ for all m, then $\mathcal{R}(\xi) \neq \emptyset$.*

Proof. The lemma is obtained by applying (Lancaster and Rodman, 1995, Theorem 11.1.1, page 257) to the ARE (2.26), taking into account the modification for the real coefficients as outlined in Section 11.4 therein. When applying the

[2] A pair of matrices $(A, B) \in \mathbb{R}^{r,r} \times \mathbb{R}^{r,n}$ is **controllable** if $\operatorname{rank}\Big(\big[B \,|\, AB \,|\, A^2B \,|\, \dots \,|\, A^{r-1}B\big]\Big) = r$. For further details on controllability consult Chapter 4 of Lancaster and Rodman (1995).

Theorem, note that required **sign-controllability** of the pair $(\eta^\top \alpha, \eta^\top \Sigma\Sigma^\top \eta)$ follows from its controllability. The latter property is shown in the proof of Lemma 2.13. $\square$

Lemma 2.15. *For any $r \times r$ matrices A and B, the non-zero eigenvalues of AB and BA coincide.*

Proof. If v is an eigenvector AB corresponding to an eigenvalue $\lambda \neq 0$, then $Bv \neq \mathbf{0}$ and

$$\lambda(Bv) = B(ABv) = BA(Bv).$$

Therefore, λ is an eigenvalue of BA with the corresponding eigenvector Bv. $\square$

We are now ready to provide the proof of Theorem 2.11.

Proof of Theorem 2.11.(i). Define the constant

$$\xi_0 := \sup\Big\{\xi \leq 1 : \mathcal{R}(\xi) \neq \emptyset\Big\},$$

and note that $\xi_0 \in [0, 1]$. Indeed, $\xi_0 \geq 0$ since $0_r \in \mathcal{R}(0)$, and $\xi \leq 1$ by definition.

We must show $\mathcal{R}(\xi) \neq \emptyset$ for all $\xi < \xi_0$, and that $\mathcal{R}(\xi_0) \neq \emptyset$. To show the former, it suffices to establish that if there exists $\xi' \leq 1$ such that $\mathcal{R}(\xi') \neq \emptyset$, then $\mathcal{R}(\xi) \neq \emptyset$ for all $\xi < \xi'$. Indeed, if $H \in \mathcal{R}(\xi')$, then

$$H\eta^\top \Sigma\Sigma^\top \eta H + H\eta^\top \alpha + \alpha^\top \eta H + \xi\alpha^\top (\Sigma\Sigma^\top)^{-1}\alpha \leq 0.$$

Therefore, by the comparison Lemma 2.13, $\mathcal{R}(\xi) \neq \emptyset$.

To show that $\mathcal{R}(\xi_0) \neq \emptyset$, we first observe that by the definition of ξ_0, there exists a sequence $\{\xi_m\}$ such that $\xi_m \to \xi_0$ and $\mathcal{R}(\xi_m) \neq \emptyset$ for all m. It then follows from the perturbation Lemma 2.14 that $\mathcal{R}(\xi_0) \neq \emptyset$. $\square$

Proof of Theorem 2.11.(ii). Observe that $\xi_0 = 1$ is equivalent to $\mathcal{R}(1) \neq \emptyset$, and that $\mathcal{R}(1)$ is the same as the set of **real and symmetric** solutions of

$$\Big(H\eta^\top \Sigma + \alpha^\top \Sigma^{-1\top}\Big)\Big(H\eta^\top \Sigma + \alpha^\top \Sigma^{-1\top}\Big)^\top = 0.$$

Therefore, $\xi_0 = 1$ if and only if there exists an $H \in \mathcal{S}^r$ that satisfies $\alpha = -\Sigma\Sigma^\top \eta H$. It only remains to show that such H is positive definite.

Define $\sigma := (\eta^\top \Sigma\Sigma^\top \eta)^{1/2} > 0$, which exists because $\eta^\top \Sigma\Sigma^\top \eta > 0$. Moreover, note that $H > 0$ if and only if $\sigma H \sigma > 0$. Therefore, we only need to show that the eigenvalues of $\sigma H \sigma$ are positive. Premultiplying by $\eta^\top$ yields:

$$-\eta^\top \alpha = \eta^\top \Sigma\Sigma^\top \eta H = \sigma\sigma H.$$

It follows that $\sigma H \sigma$ and $\sigma\sigma H$ are both nonsingular, since $|H| = |(\eta^\top \Sigma\Sigma^\top \eta)^{-1}|\,| - \eta^\top \alpha| \neq 0$. Therefore, Lemma 2.15 yields that the eigenvalues of $\sigma H \sigma$ are the same as the eigenvalues of $\sigma\sigma H = -\eta^\top \alpha$. Finally, by the c-stabilisability of $\eta^\top \alpha$, the eigenvalues of $\sigma H \sigma$ must all be positive. □

Remark 2.16. *In the proof of Theorem 2.11.(ii), we also showed the following result: if there exists $H \in \mathcal{R}(1)$, then the eigenvalues of $\eta^\top \alpha$ must be real. This result can be obtained by an alternative approach. Indeed, if such an H exists, then*

$$\big| -\eta^\top \alpha - \lambda I\big| = \big|\eta^\top \Sigma\Sigma^\top \eta\big| \left| H - \lambda(\eta^\top \Sigma\Sigma^\top \eta)^{-1}\right|,$$

which means that the eigenvalues of $-\eta^\top \alpha$ coincide with the eigenvalues of the matrix pencil $H - \lambda(\eta^\top \Sigma\Sigma^\top \eta)^{-1}$. Now, since H is Hermitian and $(\eta^\top \Sigma\Sigma^\top \eta)^{-1}$ is positive definite, the matrix pencil is definite and, hence, its eigenvalues are all real, see, for example, Gantmacher (1959)[Theorem 8, pp. 310].

We are now ready to provide the main result of this section, i.e. the solution of PDE (2.24).

Theorem 2.17. *Assume that the matrix function $H : [0, T] \to \mathcal{S}^r$ satisfies RDE (2.25). Then, the solution of PDE (2.24) exists and is given by:*

$$\varphi(t, z) = \exp\Big(g(T - t) + \frac{1}{2} z^\top H(T - t) z\Big), \tag{2.33}$$

where

$$g(t) = \frac{1}{2} tr\Big(\eta^\top \Sigma\Sigma^\top \eta \int_0^t H(s) ds\Big) + \frac{\xi d^2}{2} t. \tag{2.34}$$

Furthermore, there exists $\xi_0 \in [0, 1]$ such that:

*(i) (**ill-posed** case) If $\xi \in (\xi_0, 1]$, then RDE (2.25) does not have a stabilising solution. In particular, it is either the case that RDE (2.25) and PDE (2.24) have **finite escape time**, i.e. H (and therefore φ) exists only up to $T_{esc} < \infty$, or that $\lim_{t\to\infty} H(t)$ does not exists.*

(ii) $\xi_0 = 1$, RDE (2.25) has a unique stabilising solution and PDE (2.24) is well-posed for all $(T, \xi) \in (0, \infty) \times (-\infty, 1]$ if and only if there exists a real symmetric matrix $H > 0$ such that

$$\alpha = -\Sigma\Sigma^\top \eta H. \tag{2.35}$$

Proof. Substituting the ansatz (2.33), with unknown functions $g : [0, T] \to \mathbb{R}$ and $H : [0, T] \to \mathcal{S}^r$, into (2.24) yields:

$$\begin{aligned}
\varphi(t,z)\Big\{ &- g'(T-t) - \frac{1}{2} z^\top H'(T-t) z \\
&+ \frac{1}{2} z^\top \left(\alpha^\top \eta H(T-t) + H(T-t)\eta^\top \alpha\right) z \\
&+ \frac{1}{2}\mathrm{tr}\Big[\eta^\top \Sigma\Sigma^\top \eta\Big(H(T-t)zz^\top H(T-t) + H(T-t)\Big)\Big] \\
&+ \frac{\xi}{2}\left(d^2 + z^\top \alpha^\top (\Sigma\Sigma^\top)^{-1} z\right)\Big\} = \\
\varphi(t,z)\Big\{ &\frac{1}{2} z^\top \Big[- H'(T-t) + \alpha^\top \eta H(T-t) + H(T-t)\eta^\top \alpha \\
&\qquad + H(T-t)\eta^\top \Sigma\Sigma^\top \eta H(T-t) + \xi \alpha^\top (\Sigma\Sigma^\top)\alpha\Big] z \\
&- g'(T-t) + \frac{1}{2}\mathrm{tr}\Big[\eta^\top \Sigma\Sigma^\top \eta H(T-t)\Big] + \frac{\xi d^2}{2}\Big\} = 0; \quad (t,z) \in [0,T) \times \mathbb{R}^r.
\end{aligned}$$

Therefore, H must satisfy (2.25) and g must be given by (2.34).

Next, we utilise the connection between RDE (2.25) and ARE (2.26). In particular, by Theorem 2.11.(i), there exists $\xi_0 \in [0, 1]$ such that for $\xi \in (\xi_0, 1]$ ARE (2.26) does not have a symmetric solution. Statement (i) then follows from Proposition 2.8. Similarly, by Theorem 2.11.(ii), $\xi_0 = 1$ if and only if (2.35) holds, and statement (ii) follows. $\square$

Remark 2.18. *It must be mentioned that Theorem 2.17 provides weaker results when compare to the bivariate case, i.e. Propositions 1.18 and 1.19 in Appendix*

1.A. In particular, the bivariate results provide the necessary and sufficient condition for well-posedness of the PDE, while Theorem 2.17 only provides the sufficient condition. The main difference between the two cases is that, in the bivariate case, the scalar Riccati equation (1.72) was shown to either has a stabilizing solution (cf. Lemma 1.16) or a finite escape time (cf. Lemma 1.17). We strongly believe that this dichotomy also holds for the multivariate case, i.e.

Conjecture 2.19. *RDE (2.25) either has a stabilising solution, or a finite escape time.*

We were unable to prove this result, or to give a counterexample. Note that, if Conjecture 2.19 holds, then by Theorem 2.17.(i), condition (2.35) is also necessary for well-posedness of PDE (2.24) with $(T, \xi) \in (0, \infty) \times (-\infty, 1]$. In other words, we are only one result shy of identifying the necessary and sufficient conditions for well-posedness of PDE (2.24).

2.4 Optimal strategy for power utility

In this section, we solve the Merton investment problem for power utility in CTECM market setting, i.e.

$$(\pi_t^{\star,(p)})_{t\in[0,T]} := \arg\max_{\pi\in\mathcal{A}} \mathbb{E}\left(\frac{\left(X_T^{\pi}\right)^p}{p}\right), \tag{2.36}$$

for $p \in (-\infty, 0) \cup (0, 1)$. The results of this section generalise those obtained by Liu and Timmermann (2013). We also provide new well-posedness conditions.

When comparing the results with Liu and Timmermann (2013), the reader should take into consideration that therein it is assumed that $n = 2$, $k = 1$, $\beta = (b, b)$, $\eta^\top = (1, -1)$, $\alpha^\top = (-\lambda_1, \lambda_2)$, and

$$\Sigma_S = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}.$$

Furthermore, they parametrised the power utility with the relative risk aversion parameter $\gamma := 1 - p$.

We start by introducing the value function. For $(\pi, s, x, z) \in \mathcal{A} \times [0, T] \times \mathbb{R}^+ \times \mathbb{R}^r$, let us define $(X_t^{\pi,s,x,z})_{t\in[s,T]}$ as the wealth process of the agent if she follows an admissible strategy π from s to T and, at s her wealth is x and the log-spread is z. Similarly, define $(Z_t^{s,z})_{t\in[s,T]}$ as the log-spread from s to T if at s the log-spread is z. In other words, $(X_t^{\pi,s,x,z})$ and $(Z_t^{s,z})$ are the solution of the equations

$$dX_t^{\pi,s,x,z} = X_t^{\pi,s,x,z}\pi_t^\top \Sigma\Big(\lambda(Z_t^{s,z})dt + dW_t\Big), \tag{2.37}$$

and

$$dZ_t^{s,z} = \eta^\top \alpha Z_t^{s,z} dt + \eta^\top \Sigma_S dW_{S,t}, \tag{2.38}$$

with the initial conditions $X_s^{\pi,s,x,z} = x$, and $Z_s^{s,z} = z$. Then, the value function associated with the Merton problem (2.20) is defined as:

$$u(t, x, z; T) := \sup_{\pi\in\mathcal{A}} \mathbb{E}\left(U\left(X_T^{\pi,t,x,z}\right)\right), \quad (t, x, z) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^r. \tag{2.39}$$

For the special case of power utility, we define the value function:

$$u(t, x, z; T, p) := \sup_{\pi\in\mathcal{A}} \mathbb{E}\left(\frac{\left(X_T^{\pi,t,x,z}\right)^p}{p}\right), \quad (t, x, z) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^r. \tag{2.40}$$

The following condition is the generalisation of Condition 1.11 to multivariate case.

Condition 2.20. *There exists a real symmetric $r \times r$ matrix $H_0 > 0$ such that*

$$\alpha = -\Sigma_S \Sigma_S^\top \eta H_0. \tag{2.41}$$

Our main goal for the rest of this chapter is to show that Condition 2.20 plays the same central role as its bivariate counterpart. The first of such results is the next theorem which shows that Condition 2.20 is sufficient for the well-posedness of value function (2.40) with $(T, p) \in (0, \infty) \times (-\infty, 1]$, and provides the optimal strategy when the condition holds.

Theorem 2.21. *(Well-posed case) Assume that Condition 2.20 holds. Then, for all $p \in (-\infty, 1)$ and $T \in (0, \infty)$, the Merton problem (2.36) is well-posed. Furthermore, the value function $u(t, x, z; T, p)$, is given by*

$$u(t, x, z; T, p) = \frac{x^p}{p}\left(e^{g(T-t)+\frac{1}{2}z^\top H(T-t)z}\right)^{1-p}, \tag{2.42}$$

the optimal investment strategy in the indices is:

$$\pi^{\star}_{I,t} := \frac{1}{1-p}(\Sigma_I \Sigma_I^{\top})^{-1}\mu_I, \tag{2.43}$$

and the optimal investment strategy in the stocks is:

$$\pi^{\star}_{S,t} := \eta\left(H(T-t) - \frac{1}{1-p}H_0\right)Z_t. \tag{2.44}$$

Here, the matrix function H is the stabilising solution of the matrix Riccati differential equation:

$$H'(t) = H(t)\eta^{\top}\Sigma_S\Sigma_S^{\top}\eta H(t) + H(t)\frac{\eta^{\top}\alpha}{1-p} + \frac{\alpha^{\top}\eta}{1-p}H(t) + \frac{p}{(1-p)^2}\alpha^{\top}(\Sigma_S\Sigma_S^{\top})^{-1}\alpha, \tag{2.45}$$

with the initial condition $H(0) = 0$, and $g : [0,T] \to \mathbb{R}$ is given by:

$$g(t) = \frac{1}{2}tr\Big(\eta^{\top}\Sigma_S\Sigma_S^{\top}\eta\int_0^t H(s)ds\Big) + \frac{p}{2}\left(\frac{\mu_I^{\top}\Omega_I^{-1}\mu_I}{(1-p)^2}\right)t. \tag{2.46}$$

Proof. The proof is similar to the proof of Theorem 1.8 and is divided into three steps. The first step provides an upper bound for the value function. In the second step, the upper bound is proved to be bounded. In the third step, an admissible strategy is obtained which attains the upper bound found in the first step, hence verifying that the upper bound is the value function.

Step 1: Let $(Y_t^{s,z})_{t\in[s,T]}$ be the state price density if $Z_s = z$, i.e.

$$Y_t^{s,z} = \mathcal{E}\big(-\lambda(Z^{s,z})\cdot W\big)_t, \quad t \in [s,T]. \tag{2.47}$$

A duality argument similar to Step 1 of the proof of Theorem 1.8, yields an upper bound for the value function, namely,

$$u(t,x,z;T,p) \leq \frac{x^p}{p}\Big(\mathbb{E}\big((Y_T^{t,z})^{\frac{p}{p-1}}\big)\Big)^{1-p}. \tag{2.48}$$

Step 2: Define:

$$\psi(t,y,z) := \mathbb{E}\big((Y_T^{t,y,z})^{\frac{p}{p-1}}\big), \tag{2.49}$$

where,

$$Y_s^{t,y,z} := y\mathcal{E}\big(-\lambda(Z^{t,z})\cdot W\big)_s, \quad s \in [t,T]. \tag{2.50}$$

We consider the related Cauchy problem:

$$\begin{aligned}\psi_t + \tfrac{1}{2}y^2\|\lambda(z)\|^2\psi_{yy} + z^\top\alpha^\top\eta\,\psi_z \\ + \tfrac{1}{2}\mathrm{tr}\big(\eta^\top\Sigma_S\Sigma_S^\top\eta\,\psi_{zz}\big) - y\lambda(z)^\top(\mathbf{0}_{n\times k},\Sigma_S)^\top\eta\psi_{zy} = 0;\end{aligned} \tag{2.51}$$

$(t, y, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}^r$, with terminal condition $\psi(T, y, z) = y^{\frac{p}{p-1}}$. If (2.51) has a (classical) solution, then a suitable version of the Feynman-Kac formula (e.g. Janson and Tysk (2006)), yields the stochastic representation (2.49). Substituting the ansatz

$$\psi(t, y, z) = y^{\frac{p}{p-1}}\varphi(t, z),$$

yields the following PDE for the unknown function φ:

$$\varphi_t + \frac{1}{1-p}z^\top\alpha^\top\eta\,\varphi_z + \frac{1}{2}\mathrm{tr}\big(\eta^\top\Sigma_S\Sigma_S^\top\eta\,\varphi_{zz}\big) + \frac{p}{2(p-1)^2}\|\lambda(z)\|^2\varphi = 0; \tag{2.52}$$

$(t, z) \in [0, T) \times \mathbb{R}^r$, with the terminal condition $\varphi(T, z) = 1$. This PDE is studied in Section 2.3. In particular, PDE (2.24) becomes (2.52) if one substitutes α, Σ, ξ, and d^2 with $\alpha/(1-p)$, Σ_S, p, and $\frac{\mu_I^\top\Omega_I^{-1}\mu_I}{(1-p)^2}$, respectively. Since condition (2.35) holds by Condition 2.20, Theorem 2.17.(ii) yields that PDE (2.52) is well-posed for all $p \in (-\infty, 1)$ and $T \in (0, \infty)$, and its unique solution is given by

$$\varphi(t, z) = \exp\Big(g(T-t) + \frac{1}{2}z^\top H(T-t)z\Big), \tag{2.53}$$

where H is the unique stabilising solution of (2.45) and g is given by (2.46). It then follows that $\psi(t, y, z)$ of (2.49) is bounded for all $(T, p) \in (0, \infty) \times (-\infty, 1)$ and, in particular,

$$\mathbb{E}\Big((Y_T^{t,z})^{\frac{p}{p-1}}\Big) = \exp\Big(g(T-t) + \frac{1}{2}z^\top H(T-t)z\Big); \quad (t, z) \in [0, T] \times \mathbb{R}^r. \tag{2.54}$$

Step 3: To construct the optimal strategy, we take the classical stochastic control approach through the Hamilton-Jacobi-Bellman (HJB) equation:

$$u_t + H(x, z, u_x, u_{xx}, u_z, u_{zz}, u_{xz}) = 0, \quad (t, x, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}^r, \tag{2.55}$$

with the terminal condition $u(T, x, z; T, \gamma) = \frac{x^p}{p}$. Here, the **Hamiltonian** is

$$\begin{aligned}H(x, z, u_x, u_{xx}, u_z, u_{zz}, u_{xz}) = z^\top\alpha^\top\eta u_z + \tfrac{1}{2}\,\mathrm{tr}\,\big(\eta^\top\Sigma_S\Sigma_S^\top\eta\,u_{zz}\big) \\ + \sup_{\pi\in\mathbb{R}^{k+n}}\Big\{x\pi^\top\Big(\begin{pmatrix}\mu_I \\ \beta\mu_I+\alpha z\end{pmatrix}u_x + \begin{pmatrix}0_{k\times r} \\ \Sigma_S\Sigma_S^\top\eta\end{pmatrix}u_{zx}\Big) + \tfrac{1}{2}x^2\pi^\top\Sigma\Sigma^\top\pi u_{xx}\Big\}.\end{aligned} \tag{2.56}$$

Optimising the right side of this equation yields the candidate optimal strategy:

$$\pi^\star := (\Sigma\Sigma^\top)^{-1} \begin{pmatrix} \mu_I \\ \beta\mu_I+\alpha z \end{pmatrix} \frac{-u_x}{xu_{xx}} + (\Sigma\Sigma^\top)^{-1} \begin{pmatrix} 0_{k\times r} \\ \Sigma_S\Sigma_S^\top\eta \end{pmatrix} \frac{-u_{zx}}{xu_{xx}}.$$

Recall that, by (2.12):

$$(\Sigma\Sigma^\top)^{-1} = \begin{pmatrix} (\Sigma_I\Sigma_I^\top)^{-1}+\beta^\top(\Sigma_S\Sigma_S^\top)^{-1} & -\beta^\top(\Sigma_S\Sigma_S^\top)^{-1} \\ -(\Sigma_S\Sigma_S^\top)^{-1}\beta & (\Sigma_S\Sigma_S^\top)^{-1} \end{pmatrix},$$

and, by the standing assumption (iv), we have that $\eta^\top\beta = 0$. By substituting these into the candidate optimal strategy, one obtains:

$$\pi^\star := \begin{pmatrix} (\Sigma_I\Sigma_I^\top)^{-1}\mu_I-\beta^\top(\Sigma_S\Sigma_S^\top)^{-1}\alpha z \\ (\Sigma_S\Sigma_S^\top)^{-1}\alpha z \end{pmatrix} \frac{-u_x}{xu_{xx}} + \begin{pmatrix} 0_{k\times r} \\ \eta \end{pmatrix} \frac{-u_{zx}}{xu_{xx}}. \tag{2.57}$$

Substituting $\pi^\star$ into (2.56) and then into (2.55) yields

$$\begin{aligned} &u_t + z^\top\alpha^\top\eta u_z + \frac{1}{2}\,\mathrm{tr}\left(\eta^\top\Sigma_S\Sigma_S^\top\eta\, u_{zz}\right) \\ &-\frac{1}{2}\|\lambda(z)\|^2\frac{u_x^2}{u_{xx}} - \frac{1}{2u_{xx}}u_{xz}\eta^\top\Sigma_S\Sigma_S^\top\eta u_{zx} - z^\top\alpha^\top\eta\frac{u_xu_{zx}}{u_{xx}} = 0; \end{aligned} \tag{2.58}$$

$(t,x,z) \in [0,T)\times\mathbb{R}^+\times\mathbb{R}^r$, with the terminal condition $u(T,x,z;T,\gamma) = \frac{x^p}{p}$ as before.

Next, we guess that the upper bound in (2.48) is the value function. This suggests the ansatz

$$u(t,x,z;T,\gamma) = \frac{x^p}{p}\big(\varphi(t,z)\big)^{1-p}$$

with the unknown function $\varphi(\cdot,\cdot)$. Substituting this ansatz into (2.58) yields that φ must satisfy (2.52) and, hence, is given by (2.53). It follows that the solution of the HJB equation (2.58) is (2.42). Substituting u into (2.57) and imposing Condition 2.20, i.e. $\alpha = -\Sigma_S\Sigma_S^\top\eta H_0$, yields the optimal strategies (2.43) and (2.44). Finally, the expected terminal utility corresponding to strategies (2.43) and (2.44) is the solution of (2.58), which coincides with the upper bound (2.48) found in Step 1. This verifies that the function given in (2.42) is indeed the value function. □

Theorem 2.21 has two main implications. Firstly, it proves that Condition 2.20 is a sufficient conditions for well-posedness of the Merton problem for **all** power utilities, or, equivalently, a sufficient condition for the absence of nirvana strategies for investors with power utility. We will not show that Condition 2.20 is also a necessary condition, since, as mentioned in Remark 2.18, we still lack sharp results for well-posedness of PDE (2.52). In particular, if Conjecture 2.19 is correct, then the necessity of Condition 2.20 will immediately follow.

Secondly, Theorem 2.21 provides an intuitive decomposition for the investor's optimal portfolio. Indeed, if Condition 2.20 holds, then the investor only needs two types of market traders: **indexers** and **market-neutral agents**. Indexers provide portfolio component (2.43). Like mutual-funds, their goal is generally to be long the market by taking risk-adjusted positions in market indices. The market-neutral agents, on the other hand, provide portfolio component (2.44), which is beta-neutral as well as market-neutral in the sense of Definition 2.1. Hence, like hedge funds, market-neutral agents seek returns which are uncorrelated with the market and are obtained by frequent rebalancing of their portfolio.

We end this section by investigating what happens if Condition 2.20 fails. Let us introduce the notion of **non-stabilising** trading strategies.

Definition 2.22. *Let $\{\pi^{(T)}\}_{T\in\mathbb{R}^+}$ be a family of admissible strategies parametrised by the investment horizon T. In other words, for each $T>0$, $\pi^{(T)}=(\pi_t^{(T)})_{t\in[0,T]}\in\mathcal{A}(T)$. The family of admissible strategies is called **stabilising**, if $\lim_{T\to\infty}\pi_0^{(T)}$ exists and is finite. Otherwise, it is called **non-stabilising**.*

Non-stabilisability of optimal strategies is a weaker notion than existence of nirvana strategies. In particular, if the value function explodes in finite time, then the family of optimal strategies parametrised by investment horizon is non-stabilising. The converse is not true in general. For example, it may be the case that the value function is bounded for all investment horizon T, but the optimal initial position approaches infinity as $T\to\infty$.

The following theorem identifies what happens to the Merton problem if Condition 2.20 fails.

Theorem 2.23. *(Ill-posed case) Assume that Condition 2.20 fails. Then, there exists $T_{esc}\in\mathbb{R}^+\cup\{\infty\}$, such that (2.45) has a solution for $T<T_{esc}$. For such T, the value function is still given by (2.42) with g as in (2.46). But now, the optimal*

investment strategies in the indices and stocks are

$$\pi^{\star}_{I,t} := \frac{1}{1-p}\Big[(\Sigma_I\Sigma_I^{\top})^{-1}\mu_I - \beta^{\top}(\Sigma_S\Sigma_S^{\top})^{-1}\alpha Z_t\Big] \tag{2.59}$$

and

$$\pi^{\star}_{S,t} := \left[\frac{1}{1-p}(\Sigma_S\Sigma_S^{\top})^{-1}\alpha + \eta H(T-t)\right] Z_t, \tag{2.60}$$

respectively. Furthermore, there exists $p_0 \in [0,1)$ *such that, for* $p \in (p_0, 1)$, *the Riccati equation (2.45) does not have a stabilising solution, and the family of optimal strategies (2.60), parametrised by the investment horizon, are non-stabilising.*

Proof. The proof is parallel to that of Theorem 2.21. The only difference is that since Condition 2.20 fails, there is no guarantee that the PDE (2.52) has a solution, unless it is assumed that the matrix Riccati equation (2.45) has a solution. By the existence results for ordinary differential equations, the solution of (2.45) exists locally over some interval $[0, T_{\text{esc}})$, $T_{\text{esc}} \in (0,\infty]$. The existence and properties of p_0 follow by Theorem 2.17.(i). In particular, $p_0 < 1$ since Condition 2.20 fails. The rest of the arguments are similar to the well-posed case and are, thus, omitted. □

The main implication of Theorem 2.23 is that, when Condition 2.20 fails, there is no guarantee that the Merton problem with power utility is well-posed. In particular, there exists power utilities for which the optimal strategy is not well-behaved as the time to maturity increases, i.e. the optimal strategy is non-stabilizing. If Conjecture 2.19 is true, these strategies are shown to be the nirvana strategies and the Merton problem will be ill-posed for these power utilities.

Another implication of Theorem 2.23 is that the intuitive decomposition of agent's strategy into indexers' and market-neutral components, is no longer true if Condition 2.20 fails. In particular, as (2.59) indicates, there is an active investment strategy involving the indices. Furthermore, the optimal investment in the stocks, i.e. (2.60), is no longer market-neutral, cf. Definition 2.1. Note that these observations are consistent with what is documented by Liu and Timmermann (2013) about the optimal convergence trading with power utilities.

2.5 Market-neutrality, well-posedness, and no-arbitrage for general utility functions

The following theorem is the main result of this chapter. It states the central role of Condition 2.20 in the CTECM market setting and with general terminal utility functions.

Theorem 2.24. *Suppose Condition 2.20 holds. Then,*

(i) for all $T \in (0, \infty)$, the Novikov condition holds, i.e.

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \|\lambda(Z_s)\|^2 ds\Big)\Big] < \infty, \quad \forall T \in (0, \infty). \tag{2.61}$$

Furthermore, consider the Merton problem (2.20) with any utility function U satisfying the conditions stated in Assumption 2.6. Then:

(ii) The Merton problem (2.20) is well posed for any investment horizon $T > 0$.

(iii) The optimal investment in the stocks is market-neutral. In particular, the optimal investment in the stocks is given by:

$$\pi^\star_{S,t} := \eta\left[\frac{u_x}{xu_{xx}}(t, X_t, Z_t)H_0 Z_t - \frac{u_{zx}}{xu_{xx}}(t, X_t, Z_t)\right], \tag{2.62}$$

and the optimal investment in the indices is given by:

$$\pi^\star_{I,t} := \frac{-u_x}{xu_{xx}}(t, X_t, Z_t)(\Sigma_I\Sigma_I^\top)^{-1}\mu_I. \tag{2.63}$$

Proof. **(i)**: Define the function $\varphi : [0, T] \times \mathbb{R}^r \to \mathbb{R}$ by

$$\varphi(t, z) := \mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_t^T \|\lambda(Z_s^{t,z})\|^2 ds\Big)\Big], \tag{2.64}$$

and consider the Cauchy problem:

$$\varphi_t + z^\top\alpha^\top\eta\,\varphi_z + \frac{1}{2}\mathrm{tr}\big(\eta^\top\Sigma_S\Sigma_S^\top\eta\,\varphi_{zz}\big) + \frac{1}{2}\|\lambda(z)\|^2\varphi = 0; \tag{2.65}$$

$(t, z) \in [0, T) \times \mathbb{R}^r$, with the terminal conditions $\varphi(T, z) = 1$. By a suitable version of the Feynman-Kac formula, e.g. Janson and Tysk (2006), φ is the (classical) solution of (2.65) if and only if it satisfies the stochastic representation (2.64). PDE (2.65) is a special case of the PDE (2.24) solved in Section 2.3, with $\Sigma = \Sigma_S$, $\xi = 1$, and $d^2 = \mu_I^\top(\Sigma_I\Sigma_I^\top)^{-1}\mu_I$. Finally, since Condition 2.20 holds, Theorem 2.17.(ii) yields that (2.65) has a solution for all T.

(ii) and **(iii):** Since the market is complete, one may apply Theorem 2.0 in Kramkov and Schachermayer (1999) to show the regularity of the value function (2.39). But first, one must check the validity of conditions (2.2), (2.4) and (2.5) therein. Condition (2.4) is the Inada condition (2.21) and Condition (2.2), i.e. existence of a risk-neutral measure, follows from the Novikov condition (2.61). To show condition (2.5), namely, that $u(0,x,z;T) < \infty$ for $(x,z) \in \mathbb{R}^+ \times \mathbb{R}^r$, note that, by Lemma 2.7, there exist $p < 1$ such that, for any $T > 0$ and $(x,z) \in \mathbb{R}^+ \times \mathbb{R}^r$:

$$u(0,x,z;T) \leq \sup_{\pi\in\mathcal{A}}\mathbb{E}\left(\frac{\left(X_T^{\pi,0,x,z}\right)^p - 1}{p}\right) = u(0,x,z;T,p) - \frac{1}{p}.$$

Since Condition 2.20 holds, it follows from Theorem 2.21 that the right side is bounded. Hence, condition (2.5) in Kramkov and Schachermayer (1999) is also valid. It then follows from Theorem 2.0 therein that the value function (2.39) and its dual are bounded and smooth, and that the dual value function is given by

$$v(t,y,z;T) = \mathbb{E}\left(V\left(Y_T^{t,y,z}\right)\right), \tag{2.66}$$

where V is the convex conjugate of U, and $(Y_s^{t,y,z})_{s\in[t,T]}$ is given by (2.50).

It only remains to find the optimal strategy. The HJB equation associated with the value function (2.39) is the same as (2.55) with the terminal condition $u(T,x,z;T) = U(x)$. The same argument as in the Step 3 of the proof of Theorem 2.21, yields the candidate optimal strategy

$$\pi^\star := \left(\begin{smallmatrix}(\Sigma_I\Sigma_I^\top)^{-1}\mu_I - \beta^\top(\Sigma_S\Sigma_S^\top)^{-1}\alpha z\\ (\Sigma_S\Sigma_S^\top)^{-1}\alpha z\end{smallmatrix}\right)\frac{-u_x}{xu_{xx}} + \left(\begin{smallmatrix}0_{k\times r}\\ \eta\end{smallmatrix}\right)\frac{-u_{zx}}{xu_{xx}}, \tag{2.67}$$

which, in turn, simplifies the HJB equation to

$$\begin{aligned}&u_t + z^\top\alpha^\top\eta u_z + \frac{1}{2}\operatorname{tr}\left(\eta^\top\Sigma_S\Sigma_S^\top\eta\, u_{zz}\right)\\ &-\frac{1}{2}\|\lambda(z)\|^2\frac{u_x^2}{u_{xx}} - \frac{1}{2u_{xx}}u_{xz}\eta^\top\Sigma_S\Sigma_S^\top\eta u_{zx} - z^\top\alpha^\top\eta\frac{u_xu_{zx}}{u_{xx}} = 0;\end{aligned} \tag{2.68}$$

$(t, x, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}^r$, with the terminal condition $u(T, x, z; T) = U(x)$. Applying the Legendre transform, i.e.

$$v(t, y, z) = \sup_x \Big\{ u(t, x, z; T) - xy \Big\},$$

to the simplified HJB equation, yields the dual HJB equation

$$v_t + z^\top \alpha^\top \eta v_z + \frac{1}{2} \operatorname{tr}\left(\eta^\top \Sigma_S \Sigma_S^\top \eta \, v_{zz} \right) + \frac{1}{2} \|\lambda(z)\|^2 y^2 v_{yy} - y z^\top \alpha^\top \eta v_{zy} = 0; \quad (2.69)$$

$(t, y, z) \in [0, T) \times \mathbb{R}^+ \times \mathbb{R}^r$, with the terminal condition $v(T, y, z) = V(y)$. By the Feynman-Kac theorem, the solution of (2.69) is the dual value function (2.66). Therefore, the solution of (2.68) is indeed the value function, and hence, (2.67) is the optimal portfolio strategy in feedback form. Assertions (2.62) and (2.63) follow by applying Condition 2.20 to (2.67). In particular, the optimal investment in the stocks is market-neutral. □

Theorem 2.24 highlights a connection between the following seemingly unrelated properties of the CTECM market setting and the associated Merton problem with general utility functions:

- Market-neutrality of the optimal convergence strategies in the stocks.
- Well-posedness of the Merton problem and absence of nirvana strategies.
- Absence of arbitrage.

Therefore, the theorem provides theoretical justification for the market-neutrality assumption.

Furthermore, Theorem 2.24 shows that the intuitive decomposition of the investors optimal portfolio, which was found in the power utility case, also holds for the case of general utility. Indeed, assume that there are "indexer" agents who provide the constant weight investment strategy $(\Sigma_I \Sigma_I^\top)^{-1} \mu_I$ in the indices. Like mutual-funds, their goal is to generally take a long risk-adjusted positions in market indices. According to (2.63), the optimal strategy of an investor in the indices is to simply adjust her investment in an indexer's portfolio according to her relative risk aversion $-u_x/(x u_{xx})$. "Market-neutral agents", on the other hand, provide the optimal investment in the stocks according to (2.62), which is beta-neutral as well as market-neutral in the sense of Definition 2.1. Hence, like hedge funds, market-neutral agents seek returns which are uncorrelated with the market and are obtained by frequent rebalancing of their portfolio.

Theorem 2.24 is weaker than its bivariate counterpart, i.e. Theorem 1.12. In particular, Theorem 2.24 does not provide the necessity of Condition 2.20. The only reason for this difference is the lack of sharp results for well-posedness of PDE (2.65) and well-posedness of the Merton problem with power utilities. Indeed, if Conjecture 2.19 is correct, then Condition 2.20 is shown to also be necessary for the Novikov condition and well-posedness of the Merton problem for all utility functions, and the generality of the bivariate case will be recovered.

We end our discussion by identifying what can be said if Condition 2.20 does not hold. Note that if Condition 2.20 fails, there is no guarantee that the optimal strategy (2.20) even exists. Furthermore, by Theorem 2.23, there exists utility functions (i.e. power utilities with $p > p_0$) for which the family of the optimal strategies (2.20) parametrised by T is non-stabilising, cf. Definition 2.22. Assuming that the optimal strategy exists, the argument through the HJB equation yields that the optimal strategy must necessarily satisfy (2.67). This yields the candidate optimal investment in the stocks as follows:

$$\pi^{\star}_{S,t} := \frac{-u_x}{xu_{xx}}(\Sigma_S\Sigma_S^{\top})^{-1}\alpha z + \eta\frac{-u_{zx}}{xu_{xx}}.$$

Note that, in general, this strategy is not market-neutral.

2.6 Conclusion

In this chapter, we generalized, to a great extent, the results obtained for pairs-trading in Chapter 1 to the more realistic scenario of convergence trading. In particular, we considered a rather general CTECM market setting which consists of tradable risk factors as well as multiple cointegrated stocks and solved the Merton investment problem with general utility functions. We provided Condition 2.20 which is the counterpart of well-posedness condition 1.11 found in the previous chapter, and showed its connection with market-neutrality of the optimal investment strategy in the stocks, well-posedness of the Merton problem, and absence of arbitrage strategies. In the process of proving these results, we also obtained some well-posedness conditions for matrix Riccati differential equations which are, to the best of our knowledge, new.

On the downside, the results shown in this chapter are not as strong as their counterparts for the bivariate case of Chapter 1. The main reason for this discrepancy is that the well-posedness result for the auxiliary PDE for the power utility case, which are studied in Section 2.3, are not as sharp as their bivariate counterparts in Appendix 1.A. In particular, we identified a single result, i.e. Conjecture 2.19 which, if it is true, will make the multivariate results of this chapter as strong as those of Chapter 1.

2.A Proof of Lemma 2.7

First we consider a bounded domain, that is, we show that for any $x_0 > 1$, there exists a $\tilde{p} < 1$ such that

$$U(x) \leq \frac{x^{\tilde{p}} - 1}{\tilde{p}}, \quad \text{for} \quad x \in (0, x_0). \tag{2.70}$$

One way to see this, is as follows. Since $U''(1) < 0$, one may choose a constant p_1 such that

$$1 + U''(1) < p_1 < 1. \tag{2.71}$$

Let

$$\hat{U}_1(x) = \frac{x^{p_1} - 1}{p_1}.$$

Then, one has $U(1) = \hat{U}_1(1)$, $U'(1) = \hat{U}_1'(1)$, and $U''(1) < p_1 - 1 = \hat{U}_1''(1)$. Hence, there exist $\overline{\epsilon}, \underline{\epsilon} > 0$ such that

$$U(x) \leq \hat{U}_1(x), \quad \text{for} \quad x \in (x_0 - \underline{\epsilon}, x_0 + \overline{\epsilon}).$$

Define

$$\overline{x} := \inf\left\{x \in (1, x_0) : U(x) = \hat{U}_1(x)\right\} \tag{2.72}$$

and

$$\underline{x} := \sup\left\{x \in (0, 1) : U(x) = \hat{U}_1(x)\right\}.$$

If $\overline{x} \geq x_0$ and $\underline{x} = 0$, then (2.70) holds for $\tilde{p} = p_1$. Otherwise, assume $\overline{x} < x_0$. Since $U'(\overline{x}) < U'(1) = 1$ and $x_0 > 1$, one may choose a constant p_2 such that

$$p_2 := 1 + \frac{\log U'(\overline{x})}{\log(x_0)} < 1. \tag{2.73}$$

Let

$$\hat{U}_2(x) = \frac{x^{p_2} - 1}{p_2}.$$

Then, one has

$$\overline{x}^{p_1 - 1} = \hat{U}_1'(\overline{x}) \leq U'(\overline{x}) = x_0^{p_2 - 1} < \overline{x}^{p_2 - 1},$$

where the first and the second equalities come from (2.72) and (2.73), respectively. It follows that $p_1 < p_2$, and therefore

$$U(x) \leq \hat{U}_1(x) < \hat{U}_2(x), \quad \text{for} \quad x \in (1, \overline{x}). \tag{2.74}$$

On the other hand, let f be the tangent line to U at $x = \overline{x}$, i.e. $f(x) := U(\overline{x}) + U'(\overline{x})(x - \overline{x})$. Then,

$$f'(x) = U'(\overline{x}) = x_0^{p_2 - 1} = \hat{U}_2'(x_0) < \hat{U}_2'(x), \quad \text{for} \quad x \in (\overline{x}, x_0).$$

Besides, $f(\overline{x}) = U(\overline{x}) = \hat{U}_1(\overline{x}) < \hat{U}_2(\overline{x})$. It then follows that

$$U(x) \leq f(x) < \hat{U}_2(x), \quad \text{for} \quad x \in (\overline{x}, x_0). \tag{2.75}$$

Combining (2.74) and (2.75) yields

$$U(x) < \hat{U}_2(x), \quad \text{for} \quad x \in (1, x_0).$$

Similarly, for the case $\underline{x} > 0$, define

$$p_3 := \frac{-1}{U(\underline{x}) - \underline{x}U'(\underline{x})} < 1, \tag{2.76}$$

to obtain

$$U(x) < \frac{x^{p_3} - 1}{p_3}, \quad \text{for} \quad x \in (0, 1).$$

Therefore, inequality (2.70) holds for

$$\tilde{p} = \max\{p_1, p_2, p_3\}.$$

It only remains to address the unbounded domain (i.e. $x \in (0, \infty)$). Lemma 6.5 in Kramkov and Schachermayer (1999) (assertion $(ii) \Rightarrow (iii)$) yields that if (2.22) holds, then there exist $p_4 < 1$ and $\tilde{x}_0 > 0$ such that

$$U(x) \leq \frac{x^{p_4} - 1}{p_4}, \quad \text{for} \quad x \in (\tilde{x}_0, \infty).$$

Let $x_0 = \min\{2, \tilde{x}_0\}$, and define p_1, p_2, and p_3 by (2.71), (2.73), and (2.76), respectively. Then, for

$$p = \max\{p_1, p_2, p_3, p_4\},$$

(2.23) hold, and we easily conclude.

Chapter 3

Convergence-Trading with Two Futures and Proportional Transaction Costs

Transaction costs are the main obstacle for implementing active trading strategies, such as convergence trading, that require frequent rebalancing of the portfolio. The literature of portfolio choice problems with proportional transaction costs[1] and static opportunity sets, i.e. when the market price of risk is deterministic, is quite extensive. It starts with seminal works of Magill and Constantinides (1976) who provided fundamental insight on the structure optimal strategy. Taksar et al. (1988) and Dumas and Luciano (1991) considered the optimal long-run growth rate problem with transaction costs and provided non-standard arguments which got renewed attention recently, see Guasoni and Muhle-Karbe (2013). Davis and Norman (1990) provided the first mathematically rigorous treatment of the problem. Shreve and Soner (1994) highlighted the connection to viscosity theory for partial differential equations. Consult Zariphopoulou (1999), Bichuch (2010), Guasoni and Muhle-Karbe (2013) and Bichuch and Shreve (2013) for further literature on proportional transaction costs.

Models with static opportunity set, such as the widely assumed geometric Brownian motion, are not appropriate for convergence trading scenarios when

[1] There are two main categories of models with transaction costs: fixed cost models which lead to impulse control problems, and proportional or linear cost models which lead to singular control problems. We only consider the proportional case, and refer the interested reader to Morton and Pliska (1995) and Korn (1998) for further details on fixed cost models.

the market price of risk is explicitly stochastic. This point is well articulated by Bielecki and Pliska (2000):

> There are at least two reasons why it is desirable to explicitly include factor processes in the optimization model. First, factors are often used to make forecasts of asset returns, so their inclusion facilitates understanding of the statistical issues and estimation difficulties. Second, the optimal strategies that are obtained when the factor processes are included are often different from, and thus superior to, those obtained with the certainty equivalence approach. In other words, the naive approach of first computing statistical estimates of asset drift and diffusion coefficients by conducting, say, linear regressions of returns against factor levels and then substituting these statistical estimates in formulas that emerge from conventional optimization models will lead to strategies that are not optimal. This difference is sometimes called the "hedging effect" in the financial economics literature.

On the other hand, the literature on portfolio choice with transaction costs and stochastic opportunity set is rather thin. Bielecki and Pliska (2000) considered a market setting with multiple risky assets whose prices follow a general factor model. They considered fixed transaction costs and solved the long-run growth rate problem. Soner and Touzi (2012) and Possamaï et al. (2012) assumed a similar market setting and considered the Merton investment and consumption problem over infinite time horizon with general utility functions and proportional transaction costs. They provided asymptotic expansions for the value function as the transaction costs becomes smaller. Jang et al. (2007) considered Merton's problem with proportional transaction costs and a regime-switching market model where each regime has static opportunity set. They characterised the optimal consumption/investment in closed form. Finally, Martin and Schoneborn (2011) considered a single Ornstein-Uhlenbeck asset and proportional transaction costs, and provided asymptotic expansion for the infinite horizon problem. Of these studies, only the last one is directly applicable to convergence trading. But, to make such a connection, one must assume that the optimal convergence trading strategy is market-neutral. This assumption is yet to be established in the presence of transaction costs.

Portfolio choice problems with linear transaction costs are known to be notoriously difficult. Indeed, to the best of our knowledge, there is no non-trivial closed-form solution for a Merton investment problem with linear transaction costs, despite being studied for more than three decades. One trivial case is exponential utility with static opportunity set where Merton's optimal strategy is to buy-and-hold. There are also cases were a closed form solution exists for a variation of the problem. One interesting example is Liu and Loewenstein (2002) who provide a closed form for the problem with stochastic investment horizon. The lack of closed form solutions has left the researchers with three choices: 1) Analytical studies confined to general and qualitative properties of the solution, with no quantitative results. 2) Asymptotic analysis when the transaction costs approaches zero, which is pertinent to scenarios with small transaction costs. 3) Numerical approximations. Although the first two options are very important and often lead to insightful results, only the third can be implemented in a practical set up.

In this chapter, we consider optimal pairs-trading with two cointegrated futures assets and assuming proportional transaction costs. In Section 3.1, we discuss the technicalities of trading futures and set up a Merton problem that is appropriate for portfolio managers and traders as apposed to individual investors. Section 3.2 provides the market setting and formalises the portfolio choice problem. The model possesses three characteristics whose combination makes it different from the existing literature of proportional transaction costs: 1) finite time horizon, 2) Multiple risky assets 3) stochastic opportunity set.

There are three main analytical approaches to singular control problems. The more recent approaches are through Pontryagin's maximum principle which leads to a backward stochastic differential equation (BSDE), c.f. Pham (2005); or through the use of convex optimization techniques and the dual value function, c.f. Kallsen and Muhle-Karbe (2010) and Choi et al. (2012). In Section 3.3, we adapt the more classical approach through the (primal) value function. By exploiting Bellman's dynamic programing principle, we prove that the value function is a **viscosity** solution, c.f. Crandall et al. (1992), to a so-called Hamilton-Jacobi-Bellman (HJB) variational inequality. The uniqueness and continuity of the value function are then obtained as a consequence of the **comparison principle** which holds for HJB variational inequalities.

In Section 3.4, we devise a numerical scheme to approximate the solution of the HJB variational inequality. There are two main numerical approaches that have been successfully applied to multi asset portfolio choice problems with proportional transaction costs (and static opportunity set). In the **moving boundary method**, introduced by Muthuraman and Kumar (2006), the variational inequality is replaced with a sequence of fixed-boundary problems. At each iteration, the fixed boundaries are changed according to a boundary update criterion which is devised such that the solution of the fixed boundary problems are guaranteed to converge to the solution of the original problem. The second approach, which we adapt herein, is called the **penalty method**. It was introduced by Forsyth and Vetzal (2002) in the context of pricing of American options, and has been applied to Merton's problem with linear cost (and static opportunity set) in Dai and Zhong (2010). In this approach, the variational inequality is converted to a penalized nonlinear PDE which is solved by a suitable numerical scheme.

3.1 A portfolio choice criterion for trading futures

The market setting traditionally used in mathematical finance is based on trading equities. Here, we consider a market setting suitable for trading futures continuously, which means that the investor "rollover" her positions to the next available futures contract before the current one expires, and is a common practice among futures traders.

Consider an investor who trades n futures assets with generic price processes $P_t = (P_t^1, \dots, P_t^n)$, $t \in \{0, 1, 2, \dots\}$.[2] To obtain the generic prices, it is assumed that the investor switches to the next available contract at a certain date before the expiry of the currently traded futures contract. Trading each futures market also involves maintaining a minimum deposit in a margin account to cover for losses. As in the case of taking short positions in equities, these margin accounts

[2] In this section, we use a discrete-time setting to keep the arguments simple. From the next section, all the argument and results will be presented in continuous-time.

are assumed to be a risk-less investment. We will not model the margin accounts separately. Instead, we add up all the agent's margin accounts and her other riskless investment and refer to the sum as agent's margin account denoted by the process (X_t), $t \in \{0, 1, 2, \dots\}$. It is assumed that the margin account has a short rate r.

Let $\pi_t = (\pi_t^1, \dots, \pi_t^n)^\top$ be the number of contracts held in the futures assets during the period t to $t+1$. Ignoring market imperfections like transaction costs and slippage, the market clearing mechanism in the futures markets implies the budget constraint:

$$X_t = (1+r)X_{t-1} + \pi_{t-1} \cdot \Delta P_t = (1+r)^t X_0 + \sum_{s=1}^{t} (1+r)^{t-s} \pi_{s-1} \cdot \Delta P_s.$$

By defining the discounted margin account

$$\tilde{X}_t := X_t/(1+r)^t$$

and the discounted prices

$$\tilde{P}_t := P_0 + \sum_{s=1}^{t} \Delta P_s/(1+r)^s,$$

the budget constraint can be written as:

$$\tilde{X}_t = X_0 + \sum_{s=1}^{t} \pi_{s-1} \Delta \tilde{P}_s.$$

From now on, we assume $r = 0$, i.e. that the prices (P_t) and the margin account (X_t) are already discounted and the budget constraint is given by

$$X_t = X_{t-1} + \pi_{t-1} \cdot \Delta P_t = X_0 + \sum_{s=1}^{t} \pi_{s-1} \cdot \Delta P_s. \tag{3.1}$$

In practice, instead of working with the original prices, one usually considers the volatility normalized prices

$$F_t := \sum_{s=1}^{t} \frac{P_s^i - P_{s-1}^i}{\hat{\sigma}_{s-1}^i}.$$

Here $\hat{\sigma}_t^i$ is a non-forward looking measure of volatility:

$$\hat{\sigma}_t^i := \sqrt{\mathbb{E}\left(\left(P_{t+1}^i - P_t^i\right)^2 | \mathcal{F}_t\right)},$$

where $\mathcal{F}_t$, is the σ-field generated by the random variables $(P_0, \ldots, P_t)$. Applying such normalization make the prices dimensionless with unit volatility, c.f. Martin and Zou (2012). Let the **geared positions** $\theta_t^\top = (\theta_t^1, \ldots, \theta_t^n)$ be given by:

$$\theta_t^i := \hat{\sigma}_t^i \pi_t^i, \quad i \in \{1, \ldots, n\}.$$

Then, the budget constraint 3.1 can be expressed in terms of (F_t) and (θ_t), as follows:

$$X_t = X_0 + \sum_{s=1}^{t} \theta_{s-1} \cdot \Delta F_s. \tag{3.2}$$

Henceforth, we will neither refer to the actual prices (P_t) nor the actual positions (π_t), and, for simplicity, refer to (F_t) and (θ_t) as the prices and the positions, respectively.

Next, we introduce the portfolio choice criterion. The majority of classical portfolio choice criteria originates from financial economics and are based on *terminal wealth* and *intermediate consumption*. These notions are mainly relevant to individual investors rather than traders and portfolio managers who invest money on behalf of their clients. On the other hand, portfolio managers are mainly evaluated by their historical **profit and loss (P/L)**, which is announced periodically, say, every quarter or year. The most popular indicators for performance of a portfolio manager are based on distributional properties of P/L, e.g. the ratio of the first two moments (Sharpe ratio) or higher moments.

A natural criterion along this line is the mean-variance criterion pioneered by Markowitz (1952). But, one must be careful when setting up and solving a mean-variance problem. One source of difficulty is the information asymmetry between the portfolio managers and their clients. Indeed, managers have access to various data sets, proprietary forecasting models, sophisticated trading strategies, and superior data processing infrastructure which are out of reach of their clientele. This information asymmetry means that there must be two information sets when setting up the portfolio choice model: a smaller set for the portfolio choice criterion (e.g. various moments of P/L) and a larger set for investment strategies (i.e. the admissibility set). This class of mean-variance models is referred to as *unconditional mean-variance models using conditioning information*, and are more difficult to solve. See, Ferson and Siegel (2001) for further discussion.

Our portfolio choice criterion is formulated as a classical Merton problem with exponential utility. In particular, let the period $[0, T]$ be a "representative" announcement period, say one quarter, and assume that the portfolio is rebalanced at times $\{0, 1, \ldots, T-1\}$. We define the optimal strategy, $(\theta^\star)$, as follows:

$$\theta^\star := \arg\max_{\theta\in\mathcal{A}} \mathbb{E}\left(-e^{-\gamma\left(X_0+\sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s\right)}\right), \quad \gamma > 0, \tag{3.3}$$

where $\mathcal{A}$ is an appropriately defined set of *admissible strategies*. By a Tylor's expansion, one obtains

$$\begin{aligned}\theta^\star = \arg\max_{\theta\in\mathcal{A}}\Bigg\{&\mathbb{E}\left(\sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s\right) - \frac{\gamma}{2!}\mathbb{E}\left(\left(\sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s\right)^2\right)\\ &+\frac{\gamma^2}{3!}\mathbb{E}\left(\left(\sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s\right)^3\right) - \frac{\gamma^3}{4!}\mathbb{E}\left(\left(\sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s\right)^4\right) + \cdots\Bigg\},\end{aligned}$$

That is, the optimality criterion in (3.3) penalises the unconditional even moments of $P/L := \sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s$ and rewards the unconditional odd moments.

We conclude this section by a discussion on our choice of utility function. Indeed, exponential utility is deemed to be inappropriate for portfolio choice models. The reason is that its main characteristic, i.e. constant risk aversion, is at odds with investors' behaviour. But, this argument does not concern us because we do not model investor's wealth. Instead, as discussed above, we interpret the expected utility framework as a weighted sum of moments of portfolio P/L.

There are other reasons that exponential utility is a good choice for our portfolio choice problem. Firstly, the margin account $X_0 + \sum_{s=1}^{T}\theta_{s-1}\cdot\Delta F_s$ may become negative. Secondly, we seek a criterion which only depend on the change in the portfolio value (i.e. P/L) rather than its level. Exponential utility is known for this property, i.e. the corresponding optimal strategy is independent of the initial value of the portfolio. Indeed, this is the main reason for popularity of exponential utility in derivative pricing, see Carmona (2009).

3.2 Market setting

The market consists of a margin account which is a risk-less asset that pays no interest and two *continuously traded*[3] futures contracts. It is assumed that the volatility normalized prices of the futures, denoted by $(F_t^\top) = (F_t^1, F_t^2)_{t\geq 0}$, follow a *continuous time error correction model*:

$$dF_t^1 = \alpha_1(F_t^1 - cF_t^2)dt + dW_{1,t}, \tag{3.4}$$

$$dF_t^2 = \alpha_2(F_t^1 - cF_t^2)dt + \rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}. \tag{3.5}$$

These equations can also be expressed in the compact vector form:

$$dF_t = \alpha\beta^\top F_t dt + \Sigma dW_t, \tag{3.6}$$

where $\alpha^\top = (\alpha_1, \alpha_2)$, $\beta^\top = (1, -c)$, and

$$\Sigma = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}. \tag{3.7}$$

Here, the coefficients α_1, α_2, c, ρ are constant and $(W_t^\top) = (W_{1,t}, W_{2,t})_{t\geq 0}$ is a 2-dimensional standard Brownian motion in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\geq 0}$ is the augmented Brownian filtration. The following standing assumptions are always in force and will not be mentioned in the theorems.

Standing assumption. *It is assumed that:*

(i) $|\rho| < 1$.

(ii) $c\alpha_2 - \alpha_1 > 0$.

(iii) $(F_0^1 - cF_0^2)$ *is a Gaussian random variable with mean zero and variance*

$$\sigma_z^2 := 1 + c^2 - 2\rho c. \tag{3.8}$$

It is also independent of $(W_t)_{t\geq 0}$.

[3] As discussed in the previous section, this means that the investor rollover her positions to the next available futures contract before the current one expires.

The standing assumptions (ii) and (iii), as well as the price dynamics (3.4) and (3.5) yield that the *spread* $(z_t)_{t\geq 0}$ given by the process

$$z_t := F_t^1 - cF_t^2, \tag{3.9}$$

is a stationary processing, and hence the futures prices are **cointegrated**. We state this result below. The proof is similar to Proposition 1.2 and is, thus, omitted.

Proposition 3.1. *The spread* $(z_t) := (F_t^1 - cF_t^2)_{t\geq 0}$ *satisfies*

$$dz_t = -\kappa z_t dt + \sigma_z dW_{z,t}, \tag{3.10}$$

where

$$\kappa := c\alpha_2 - \alpha_1, \quad \sigma_z^2 := 1 + c^2 - 2c\rho, \tag{3.11}$$

and

$$W_{z,t} := \frac{1}{\sigma_z}\left\{(1-c\rho)\, W_{1,t} - c\sqrt{1-\rho^2}W_{2,t}\right\}. \tag{3.12}$$

In particular, $(z_t)_{t\geq 0}$ *is a stationary Ornstein-Uhlenbeck process given by*

$$z_t = e^{-\kappa t}\left(z_0 + \sigma_z \int_0^t e^{\kappa s} dW_{z,s}\right), \tag{3.13}$$

which is a Gaussian process with the following moments

$$\mathbb{E}\left(z_t\right) = 0, \quad \mathbb{E}\left(z_t z_s\right) = \frac{\sigma_z^2}{2\kappa} e^{-\kappa|t-s|}, \quad t, s \geq 0. \tag{3.14}$$

We consider an agent who invests in the market for a fixed investment horizon $t \in [0, T]$. It is assume that the agent is subject to *linear transaction costs*, i.e. when trading the i-th futures, $i \in \{1, 2\}$, each lot of long position is opened at a higher price, $F_t^i + \eta_i$, and each lot of short position is opened at a lower price, $F_t^i - \eta_i$, where $\eta_i \geq 0$ is a fixed constant. This means that buying or selling one contract of the i-th futures will cost the investor the deterministic amount η_i. The main reason for choosing this model of transaction costs is tractability. This assumption is more or less justified if the bid-ask spread is constant through time and the size of trades are not large enough to move the prices. In general, the trading cost is a convex function of the size of trade. It should also be mentioned that we only consider transaction costs resulting from *slippage*, and we ignore other sources of transaction costs, such as fixed costs and the costs for rolling over the futures contracts before they mature.

Let a predictable F-integrable $\mathbb{R}^2$-valued process $(\theta_t^\top) = (\theta_t^1, \theta_t^2)_{t\in[0,T]}$ be the position taken in the futures. The margin account $(X_t)_{t\in[0,T]}$ is then given by

$$X_t = X_0 + \int_0^t \theta_s^\top dF_s - \sum_{i=1}^{2} \eta_i V_0^t\left(\theta^i\right).$$

Here, $V_0^t\left(\theta^i\right)$ is the total variation of the positions taken in the i-th futures $(\theta_s^i)_{s\in[0,t]}$:

$$V_0^t\left(\theta^i\right) := \sup\left\{\sum_{j=1}^{p}\left|\theta_{t_j}^i - \theta_{t_{j-1}}^i\right| : 0 = t_0 < t_1 < \cdots < t_p = t\right\},$$

which represents the number of contracts of the i-th futures bought or sold from 0 to time t. It then follows that (θ_t^i) must be of *bounded variation* or the transaction cost will be unbounded. It is well known that for any process (θ_t^i) of bounded variation, there exists non-negative and non-decreasing processes (L_t^i) and (M_t^i) such that $(\theta_t^i) = (L_t^i - M_t^i)$.

We define **investment strategies** by $\mathbb{R}^2$-valued processes $(L_t) = (L_t^1, L_t^2)_{t\in[0,T]}$ and $(M_t) = (M_t^1, M_t^2)_{t\in[0,T]}$ which are non-negative, non-decreasing, $(\mathcal{F}_t)$-adapted and right continuous with left limit (RCLL); the *cumulative number of futures contracts* bought and sold, respectively. The margin account $(X_t^{L,M})_{t\in[0,T]}$ of an investment strategy (L, M) is then given by

$$\begin{aligned} X_t^{L,M} &= X_0 + \int_0^t \theta_s^\top dF_s - \sum_{i=1}^{2} \eta_i\left(L_t^i + M_t^i\right) \\ &= X_0 + \int_0^t \theta_s^\top \alpha z_s ds + \int_0^t \theta_s^\top \Sigma dW_s - \sum_{i=1}^{2} \eta_i\left(L_t^i + M_t^i\right), \end{aligned}$$

where $(\theta_t) = (L_t - M_t)$.

Definition 3.2. *An investment strategy $(L_t, M_t)_{t\in[0,T]}$ is **admissible** if its margin account is uniformly bounded from below, i.e. there exists $b \in \mathbb{R}$ such that: $X_t^{L,M} \geq b$, $\mathbb{P} - a.s$, $t \in [0, T]$. The set of all admissible strategies is denoted by $\mathcal{A}$.*

Next, we consider the Merton investment problem with exponential utility, which is motivated in the previous section.

Definition 3.3. *The agent's optimal strategy is given by:*

$$(L_t^\star, M_t^\star)_{t\in[0,T]} := \underset{(L,M)\in\mathcal{A}}{\arg\max}\,\mathbb{E}\Big[-\exp\Big(-\gamma X_T^{L,M} + \gamma \sum_{i=1}^{2} \eta_i |L_T^i - M_T^i|\Big)\Big] \tag{3.15}$$

assuming that the maximum is finite and is attained, otherwise the optimal strategy does not exist. Here, $\gamma > 0$ is the absolute risk aversion coefficient.

Remark 3.4. *The portfolio choice criterion widely used for finite horizon problems is:*

$$\sup_{\mathcal{A}} \mathbb{E}[U(X_T)], \tag{3.16}$$

where U is the terminal utility function, see, for example, Davis et al. (1993), Gennotte and Jung (1994) and Bichuch (2011). Note, however, that this model specification is inappropriate for our setting, as it does not take into account the cost of closing the positions at T. Indeed, under (3.16), the investor becomes increasingly more reluctant to trade the assets as she approaches the end of the investment horizon, since such trades have less opportunity to "pay off" and would probably mainly incur transaction costs. In particular, any open position at T will not be closed, because doing so would only reduce the utility function. This explains why in other studies, such as Gennotte and Jung (1994), the no-trade region is found to "widen to infinity" as the time to maturity tends to zero.
Therefore, we need to directly account for the cost of closing the positions at the end of the investment horizon. One approach is to prohibit any open position at T by restricting the admissibility set. Another approach, which we adapted here, is to explicitly include the cost of closing the positions, i.e. $\sum_{i=1}^{2} \eta_i |L_T^i - M_T^i|$, in the preference criterion.

3.3 The value function and the HJB equation

We take the classical approach for solving the stochastic control problem (3.15), i.e. through the (primal) value function and using the dynamic programing principle (DPP) to relate the value function to the associated Hamilton-Jacobi-Bellman

(HJB) equation.[4] Rigorous derivation of the HJB equation is our main goal in this section. In the next section, we concentrate on solving the HJB equation via numerical approximation.

Let us start with the definition of the value function. Given $(s, x, z, \theta) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$ and a control $(L, M) \in \mathcal{A}$, the process $(X_t^{s,x,z,\theta})_{t\in[s,T]}$ is defined as the margin account if one starts at s with the initial account x, initial position $\theta^\top = (\theta_1, \theta_2)$, and spread z and follows the strategy (L, M) from s to T. In other words, $(X_t^{s,x,z,\theta})_{t\in[s,T]}$ satisfies:

$$dX_t^{s,x,z,\theta} = \theta_t^\top \alpha z_t \, dt + \theta_t^\top \Sigma \, dW_t - \eta^\top (dL_t + dM_t), \tag{3.17}$$

where (z_t) and (θ_t) solve

$$dz_t = -\kappa z_t \, dt + \beta^\top \Sigma \, dW_t \tag{3.18}$$

and

$$d\theta_t = dL_t - dM_t, \tag{3.19}$$

and the initial conditions are $X_s^{s,x,z,\theta} = x$, $z_s = z$, and $\theta_{s^-} = \theta$. Here, we omit the dependence of the margin account $X^{s,x,z,\theta}$ on the control (L, M) to ease the notation.

Definition 3.5. *The value function $u : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ associated with the agent's portfolio choice problem (3.15) is defined as:*

$$u(t, x, z, \theta; T, \gamma) := \sup_{(L,M)\in\mathcal{A}} \mathbb{E}\Big[- \exp\Big(- \gamma X_T^{t,x,z,\theta} + \gamma \sum_{i=1}^{2} \eta_i |\theta_T^i| \Big)\Big]. \tag{3.20}$$

The following lemmas provide elementary properties of the value function, which will be useful for subsequent arguments.

Lemma 3.6. *The value function is locally bounded. In particular,*

$$-e^{-\gamma(x-\eta_1|\theta_1|-\eta_2|\theta_2|)} \leq u(t, x, z, \theta; T, \gamma) < 0, \tag{3.21}$$

for $(t, x, z, (\theta_1, \theta_2)) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$.

[4] Other main approaches for solving singular control problems are through Pontryagin's maximum principle which leads to a backward stochastic differential equation (BSDE), see Pham (2005); or by considering the dual value function and using convex optimization techniques, see Kallsen and Muhle-Karbe (2010) and Choi et al. (2012).

Proof. The upper bound follows from the upper bound of the exponential utility. The lower bound follows from the sub-optimality of the strategy that closes the position at t. $\square$

Lemma 3.7. *The value function has the following scaling property:*

$$u(t,x,z,\theta;T,\gamma) = e^{-\gamma x}u(t,0,z,\theta;T,\gamma), \tag{3.22}$$

for $(t,x,z,(\theta_1,\theta_2)) \in [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$.

The scaling property (3.22) is parallel to the *homothetic property* of the power utility in Magill and Constantinides (1976) and Davis and Norman (1990). It implies that to identify the value function, one only needs to find the **scaled value function** $v : [0,T]\times\mathbb{R}\times\mathbb{R}^2 \to \mathbb{R}$ given by:

$$v(t,z,\theta;T,\gamma) := u(t,0,z,\theta;T,\gamma) = \sup_{(L,M)\in\mathcal{A}} \mathbb{E}\Big[-e^{-\gamma X_T^{t,0,z,\theta}+\gamma\eta_1|\theta_T^1|+\gamma\eta_2|\theta_T^2|}\Big]. \tag{3.23}$$

Therefore, the number of arguments is reduced from 5 to 4.

Let $\tilde{\mathcal{L}}$ be the infinitesimal operator associated with the diffusions (3.17) and (3.18), i.e.

$$\begin{aligned}\tilde{\mathcal{L}}\,u(t,x,z,\theta) = u_t &+ \theta^\top\alpha z u_x + \tfrac{1}{2}\theta^\top\Sigma\Sigma^\top\theta u_{xx} + \theta^\top\Sigma\Sigma^\top\beta u_{xz} \\ &+ \beta^\top\alpha z u_z + \tfrac{1}{2}\beta^\top\Sigma\Sigma^\top\beta u_{zz}.\end{aligned} \tag{3.24}$$

Define

$$u_i\big(t,x,z,(\theta^1,\theta^2)\big) := \frac{\partial u}{\partial\theta_i}\big(t,x,z,(\theta^1,\theta^2)\big), \quad i\in\{1,2\}.$$

The scaling property (3.22) then yields

$$\begin{aligned} u_x = -\gamma e^{-\gamma x}v, \quad u_{xx} = \gamma^2 e^{-\gamma x}v, \quad u_{xz} = -\gamma e^{-\gamma x}v_z, \\ u_t = e^{-\gamma x}v_t, \quad u_z = e^{-\gamma x}v_z, \quad u_{zz} = e^{-\gamma x}v_{zz}, \\ u_i = e^{-\gamma x}v_i, \quad i\in\{1,2\}. \end{aligned} \tag{3.25}$$

Therefore, by defining the infinitesimal operator $\mathcal{L}$ for the scaled value function v as

$$\begin{aligned}\mathcal{L}\,v(t,z,\theta;T,\gamma) = v_t &+ \beta^\top(\alpha z - \gamma\Sigma\Sigma^\top\theta)v_z \\ &+ \tfrac{1}{2}\beta^\top\Sigma\Sigma^\top\beta v_{zz} + (\tfrac{1}{2}\gamma^2\theta^\top\Sigma\Sigma^\top\theta - \gamma\theta^\top\alpha z)v,\end{aligned} \tag{3.26}$$

one obtains:

$$\tilde{\mathcal{L}}\, u(t, x, z, \theta; T, \gamma) = e^{-\gamma x} \mathcal{L}\, v(t, z, \theta; T, \gamma). \tag{3.27}$$

The rest of this section is devoted to prove the following result:

Theorem 3.8. *The scaled value function (3.23) is the unique continuous viscosity solution of the HJB variational inequality:*

$$\max_{i \in \{1,2\}} \Big\{ \mathcal{L}v,\ \gamma\eta_i v \pm v_i \Big\} = 0; \quad (t, z, \theta) \in [0, T) \times \mathbb{R} \times \mathbb{R}^2, \tag{3.28}$$

with the terminal condition:

$$v(T, z, \theta) = -e^{\gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|}, \quad (z, \theta) \in \mathbb{R} \times \mathbb{R}^2. \tag{3.29}$$

Theorem 3.8 is interpreted as follows. The 4-dimensional space of the state variable $(t, z, (\theta^1, \theta^2))$ is divided into nine regions: a *no-trade region* where $\mathcal{L}v = 0$, and eight *trading regions* where at least one of the following holds:

(B_1) $\gamma\eta_1 v + v_1 = 0$, buy the first futures.

(B_2) $\gamma\eta_2 v + v_2 = 0$, buy the second futures.

(S_1) $\gamma\eta_1 v - v_1 = 0$, sell the first futures.

(S_2) $\gamma\eta_2 v - v_2 = 0$, sell the second futures.

Note that it is never optimal to buy and sell the same futures, therefore, there are only 8 possible combinations: 1) only B_1, 2) only B_2, 3) only S_1, 4) only S_2, 5) B_1 and B_2, 6) S_1 and S_2, 7) B_1 and S_2, and 8) S_1 and B_2. Whilst inside the no-trade region, it is not optimal for the agent to trade the futures. Upon hitting its boundaries, the agent must trade instantaneously to remain in the no-trade region.

The proof of Theorem 3.8 is divided into three steps:

(1) Proving the viscosity property of the value function on the intermediate interval $[0, T)$, (c.f. Theorem 3.10).

(2) Proving the viscosity property of the value function at T, (c.f. Theorem 3.12).

(3) Proving the uniqueness and continuity of the value function based on its viscosity properties and the comparison principle for the corresponding variational inequalities, (c.f. Theorem 3.13 and the discussion that follows it).

Remark 3.9. *We pay special attention to the continuity of the value function. In the Merton consumption problem with **infinite** time-horizon and proportional transaction costs, e.g. in Shreve and Soner (1994) and Bichuch and Shreve (2013), it is possible to show the continuity of the value function before introducing the HJB equation. Indeed, the time-invariant value function can be easily shown to be concave, a property inherited from the utility function, and a concave function is always continuous in the interior of its domain (Rockafellar, 1970, theorem 10.1).*

*The matter of continuity is more delicate in a **finite time** setting as ours. Since the value function is time-variant, the aforementioned concavity argument will only result in the continuity of the value function over its spacial arguments at a fixed time, and the continuity of the value function is yet to be established. Indeed, there are examples of finite horizon singular control problems where the value function is discontinuous at T, c.f. Broadie et al. (1998) and Cvitanić et al. (1999). Therefore, we opted to to work with the weaker definition of viscosity solutions using lower and upper semicontinuous envelopes which do not rely on the ex-ante continuity of the value function. The advantage of this approach is that a standard argument, based on the comparison principle for the HJB equation, yields the continuity of the value function as a byproduct of its viscosity property.*

Theorem 3.10 below is our first main result in this section and relates the (scaled) value function to the HJB equation. Its proof is based on the following version of DPP which, despite its intuitive nature, requires a technical proof beyond the scope of our presentation. Therefore, we present it without a proof and refer the interested reader to Fleming and Soner (2006) and Pham (2005) for a discussion on DPP and the related literature.

Dynamic Programming Principle (DPP): Denote by $\mathcal{T}[t,T]$ the set of all the stopping times in $[t,T]$. Furthermore, for $(t,\theta,(L,M)) \in [0,T] \times \mathbb{R}^2 \times \mathcal{A}$ define $(\theta_s^{t,\theta})_{s\in[t,T]}$ as the positions from t to T if one trades according to (L,M) with the initial positions θ at t, i.e.

$$\theta_s^{t,\theta} = (L_s - L_{t^-}) - (M_s - M_{t^-}) + \theta, \quad s \in [t,T]. \tag{3.30}$$

Then, for $(t, x, z, \theta) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$, both of the following statements are true:

(i) For any $(L, M) \in \mathcal{A}$ and $\tau \in \mathcal{T}[t, T]$:

$$u(t, x, z, \theta; T, \gamma) \geq \mathbb{E}\left[u(\tau, X_\tau^{t,x,z,\theta}, Z_\tau^{t,z}, \theta_\tau^{t,\theta}; T, \gamma)\right]. \tag{3.31}$$

(ii) For any $\varepsilon > 0$, there exists $(L^\varepsilon, M^\varepsilon) \in \mathcal{A}$ such that for all $\tau \in \mathcal{T}[t, T]$:

$$u(t, x, z, \theta; T, \gamma) - \varepsilon \leq \mathbb{E}\left[u(\tau, X_\tau^{t,x,z,\theta}, Z_\tau^{t,z}, \theta_\tau^{t,\theta}; T, \gamma)\right]. \tag{3.32}$$

Theorem 3.10. *The scaled value function (3.23) is a viscosity solution of the HJB variational inequality:*

$$\max\{\mathcal{L}v,\, \gamma\eta_1 v + v_1,\, \gamma\eta_1 v - v_1, \gamma\eta_2 v + v_2,\, \gamma\eta_2 v - v_2\} = 0, \tag{3.33}$$

for $(t, z, \theta) \in [0, T) \times \mathbb{R} \times \mathbb{R}^2$. In other words, let $v_\star$ (resp. $v^\star$) be the ***lower*** *(resp.* ***upper****)* ***semicontinuous envelop*** *of v, i.e.*

$$\begin{aligned} v_\star(t, z, \theta) &:= \liminf_{(t',z',\theta') \to (t,z,\theta)} v(t', z', \theta'), \\ v^\star(t, z, \theta) &:= \limsup_{(t',z',\theta') \to (t,z,\theta)} v(t', z', \theta'). \end{aligned} \tag{3.34}$$

Then, for any $(t, z, \theta) \in [0, T) \times \mathbb{R} \times \mathbb{R}^2$, the following statements are true:

(i) v is a viscosity ***supersolution*** *of (3.33), i.e. for any smooth* ***test function*** *$\varphi \in \mathcal{C}^{1,2,1}([0, T) \times \mathbb{R} \times \mathbb{R}^2)$ that satisfies $\varphi(t, z, \theta) = v_\star(t, z, \theta; T, \gamma)$, the inequality $\varphi \leq v_\star$ implies*

$$\max\{\mathcal{L}\varphi,\, \gamma\eta_1\varphi + \varphi_1,\, \gamma\eta_1\varphi - \varphi_1, \gamma\eta_2\varphi + \varphi_2,\, \gamma\eta_2\varphi - \varphi_2\} \leq 0. \tag{3.35}$$

(ii) v is a viscosity ***subsolution*** *of (3.33), i.e. for any smooth* ***test function*** *$\varphi \in \mathcal{C}^{1,2,1}([0, T) \times \mathbb{R} \times \mathbb{R}^2)$ that satisfies $\varphi(t, z, \theta) = v^\star(t, z, \theta; T, \gamma)$, the strict inequality:*

$$\varphi(t', z', \theta') > v^\star(t', z', \theta'), \quad (t', z', \theta') \in ([0, T) \times \mathbb{R} \times \mathbb{R}^2) \backslash \{(t, z, \theta)\},$$

implies

$$\max\{\mathcal{L}\varphi,\, \gamma\eta_1\varphi + \varphi_1,\, \gamma\eta_1\varphi - \varphi_1, \gamma\eta_2\varphi + \varphi_2,\, \gamma\eta_2\varphi - \varphi_2\} \geq 0. \tag{3.36}$$

Proof. We prove that the value function u is a viscosity solution of the variational inequality

$$\max\left\{\tilde{\mathcal{L}}u,\ u_1-\eta_1 u_x,\ -u_1-\eta_1 u_x,\ u_2-\eta_2 u_x,\ -u_2-\eta_2 u_x\right\}=0, \tag{3.37}$$

for $(t,x,z,\theta)\in[0,T)\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$. By the scaling property of the value function, specifically (3.25) and (3.27), it then follows that the scaled value function v satisfies (3.33).

The proof is divided into two steps. In the first step, the supersolution property is proved based on DPP.(i) and a standard argument in viscosity solution theory that can be traced back to Lions (1983); and in the second step, the subsolution property is derived from DPP.(ii) using the argument presented in (Pham, 2005, section 3.2.2, page 526), which itself is originated from the results of Soner and Touzi (2002).

Step 1: supersolution property
Let $(t,x,z,\theta)\in[0,T)\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$ and consider an arbitrary smooth test function $\varphi\in\mathcal{C}^{1,2,2,1}([0,T)\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2)$ satisfying $\varphi(t,x,z,\theta)=u_\star(t,x,z,\theta;T,\gamma)$ and:

$$\varphi(t',x',z',\theta')\le u_\star(t',x',z',\theta');\quad \forall(t',x',z',\theta')\in[0,T)\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2. \tag{3.38}$$

To prove the supersolution property, we need to show that

$$\tilde{\mathcal{L}}\varphi(t,x,z,\theta)\le 0, \tag{3.39}$$

$$\varphi_i(t,x,z,\theta)-\eta_i\varphi_x(t,x,z,\theta)\le 0,\quad i\in\{1,2\}, \tag{3.40}$$

and

$$-\varphi_i(t,x,z,\theta)-\eta_i\varphi_x(t,x,z,\theta)\le 0,\quad i\in\{1,2\}. \tag{3.41}$$

We start with (3.39). By the definition of $u_\star$, there exists a sequence $\{(t_m,x_m,z_m,\theta_m)\}_{m\in\mathbb{N}}$ such that

$$\lim_{m\to\infty}(t_m,x_m,z_m,\theta_m)=(t,x,z,\theta)\quad\text{and}\quad \lim_{m\to\infty}u(t_m,x_m,z_m,\theta_m)=u_\star(t,x,z,\theta).$$

Furthermore, by the continuity of φ, the sequence $\{\gamma_m\}:=u(t_m,x_m,z_m,\theta_m)-\varphi(t_m,x_m,z_m,\theta_m)\}_{m\in\mathbb{N}}$ satisfies $\lim\limits_{m\to\infty}\gamma_m=0$. We will also need an arbitrary positive sequence $\{h_m\}_{m\in\mathbb{N}}$ such that

$$\lim_{m\to\infty}h_m=\lim_{m\to\infty}\frac{\gamma_m}{h_m}=0.$$

Let $(\tilde{X}^m_s) := (\tilde{X}^{t_m,z_m,x_m,\theta_m}_s)_{s\geq t_m}$ be the margin account if one starts at t_m with the initial account x_m, initial position θ_m, and spread z_m, and does not trade afterwards. For a fixed constant $b > 0$, define the stopping time τ_m as

$$\tau_m := \min\Big\{T, \inf\big\{s \geq t_m : |\tilde{X}^m_s - x_m| \geq b\big\}\Big\}, \tag{3.42}$$

i.e. τ_m is either T or the first time that the margin account $\tilde{X}^m$ changes by at least b, whichever comes first. Then, the strategy (L^m, M^m) that unwinds the initial position θ_m at τ_m, i.e

$$L^{m,i}_s := \max\{-\theta^i_m, 0\}\mathbb{1}_{\{s\geq\tau_m\}}, \quad M^{m,i}_s := \max\{\theta^i_m, 0\}\mathbb{1}_{\{s\geq\tau_m\}};$$

for $i \in \{1, 2\}$ and $s \in [t_m, T]$, is admissible (c.f. Definition 3.2).

Let $X^m := X^{t_m,z_m,x_m,\theta_m}$ be the margin account associated with (L^m, M^m). Applying DDP.(i), with the stopping time $\hat{\tau}_m := \min\{\tau_m, t_m + h_m\}$ yields:

$$u(t_m, x_m, z_m, \theta_m; T, \gamma) \geq \mathbb{E}\left[u(\hat{\tau}_m, X^m_{\hat{\tau}_m}, Z^{t_m,z_m}_{\hat{\tau}_m}, \theta^{t_m,\theta_m}_{\hat{\tau}_m}; T, \gamma)\right].$$

From (3.38) and the definition of $\{\gamma_m\}$, it follows that:

$$\varphi(t_m, x_m, z_m, \theta_m) + \gamma_m \geq \mathbb{E}\left[\varphi(\hat{\tau}_m, X^m_{\hat{\tau}_m}, Z^{t_m,z_m}_{\hat{\tau}_m}, \theta^{t_m,\theta_m}_{\hat{\tau}_m})\right].$$

Applying Itô's formula to $\varphi(s, X^m_s, Z^{t_m,z_m}_s, \theta^{t_m,\theta_m}_s)$ in $(t_m, \hat{\tau}_m)$ and noting that the expectation of the stochastic integral vanishes because of bounded integrand yield:

$$\begin{aligned}\tfrac{1}{h_m}\mathbb{E}\Big[&-\gamma_m + \textstyle\int_{t_m}^{\hat{\tau}_m} \tilde{\mathcal{L}}\varphi(s, X^m_s, Z^{t_m,z_m}_s, \theta_m)ds \\ &+ \varphi(\hat{\tau}_m, X^m_{\hat{\tau}_m}, Z^{t_m,z_m}_{\hat{\tau}_m}, \theta^{t_m,\theta_m}_{\hat{\tau}_m}) - \varphi(\hat{\tau}_m, X^m_{\hat{\tau}_m^-}, Z^{t_m,z_m}_{\hat{\tau}_m}, \theta_m)\Big] \leq 0.\end{aligned} \tag{3.43}$$

Since the trajectory $\tilde{X}^m$ is almost surely continuous, one has $\hat{\tau}_m = t_m + h_m$, almost surely, for sufficiently large values of m, say $m \geq N(\omega)$. Therefore,

$$\begin{aligned}&\lim_{m\to\infty}\tfrac{1}{h_m}\Big(-\gamma_m + \textstyle\int_{t_m}^{\hat{\tau}_m} \tilde{\mathcal{L}}\varphi(s, X^m_s, Z^{t_m,z_m}_s, \theta_m)ds \\ &\qquad + \varphi(\hat{\tau}_m, X^m_{\hat{\tau}_m}, Z^{t_m,z_m}_{\hat{\tau}_m}, \theta^{t_m,\theta_m}_{\hat{\tau}_m}) - \varphi(\hat{\tau}_m, X^m_{\hat{\tau}_m^-}, Z^{t_m,z_m}_{\hat{\tau}_m}, \theta_m)\Big) \\ &= \lim_{m\to\infty}\tfrac{1}{h_m}\textstyle\int_{t_m}^{t_m+h_m} \tilde{\mathcal{L}}\varphi(s, X^m_s, Z^{t_m,z_m}_s, \theta_m)ds = \tilde{\mathcal{L}}\varphi(t, x, z, \theta),\end{aligned} \tag{3.44}$$

where the limits are interpreted almost surely and the last step follows from the mean value theorem. Finally, applying the Dominated Convergence Theorem to (3.43) yields (3.39).

It remains to prove (3.40) and (3.41). Their proofs are essentially a simpler version of the proof of (3.39), in that they do not involve the time argument and expectation. To prove (3.40), consider the sequence $\{(x_m, z_m, \theta_m)\}_{m\in\mathbb{N}}$ such that

$$\lim_{m\to\infty}(x_m, z_m, \theta_m) = (x, z, \theta) \quad \text{and} \quad \lim_{m\to\infty} u(t, x_m, z_m, \theta_m) = u_\star(t, x, z, \theta).$$

Consider also the sequence $\{\gamma_m\} := u(t, x_m, z_m, \theta_m) - \varphi(t, x_m, z_m, \theta_m)\}_{m\in\mathbb{N}}$ which satisfies $\lim_{m\to\infty}\gamma_m = 0$, and an arbitrary positive sequence $\{h_m\}_{m\in\mathbb{N}}$ such that

$$\lim_{m\to\infty} h_m = \lim_{m\to\infty}\frac{\gamma_m}{h_m} = 0.$$

Next, let $(\varepsilon, i) \in \mathbb{R}^+ \times \{1, 2\}$, and consider the admissible strategy $L^{t,\varepsilon,i}$ that opens ε amount of long positions in the i-th futures at t, and unwinds the position at the stopping time which is either T, or the first time that the margin account becomes less than or equal some fixed lower bound b, whichever comes first. Let $(X_s^{t,\varepsilon,i})_{s\in[t,T]}$ be the margin account if one starts at t with the initial account x, initial position θ, and spread z, and follows the strategy $L^{t,\varepsilon,i}$. Applying DPP.(i), with the trivial stopping time t yields:

$$u(t, x_m, z_m, \theta_m; T, \gamma) \geq u(t, x_m - \eta_i\varepsilon, z_m, \theta_m + \varepsilon\mathbf{e}_i; T, \gamma),$$

where we used the unit vectors $\mathbf{e}_1^\top = (1, 0)$, $\mathbf{e}_2^\top = (0, 1)$. From (3.38) and the definition of $\{\gamma_m\}$, it follows that

$$\varphi(t, x_m, z_m, \theta_m) + \gamma_m \geq \varphi(t, x_m - \eta_i\varepsilon, z_m, \theta_m + \varepsilon\mathbf{e}_i; T, \gamma).$$

Differentiating with respect to ε, and letting $\varepsilon \to 0$ and $m \to \infty$ yields (3.40). Finally, the proof of (3.41) is parallel to that of (3.40); one only needs to substitute the strategy $L^{t,\varepsilon,i}$ with the strategy $M^{t,\varepsilon,i}$ which opens ε amount of short positions in the i-th futures at t, and unwinds the position at the same stopping time as $L^{t,\varepsilon,i}$.

Step 2: subsolution property

Let $(t, x, z, \theta) \in [0, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$ and consider an arbitrary smooth test function $\varphi \in \mathcal{C}^{1,2,2,1}([0, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2)$ such that $\varphi(t, x, z, \theta) = u^\star(t, x, z, \theta; T, \gamma)$ and

$$\varphi(t', x', z', \theta') > u^\star(t', x', z', \theta'), \tag{3.45}$$

$(t', x', z', \theta') \in ([0, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2) \backslash \{(t, x, z, \theta)\}$. We are going to show that

$$\begin{aligned}\max \Big\{ & \tilde{\mathcal{L}}\varphi(t, x, z, \theta), \\ & \varphi_1(t, x, z, \theta) - \eta_1 \varphi_x(t, x, z, \theta), \\ & -\varphi_1(t, x, z, \theta) - \eta_1 \varphi_x(t, x, z, \theta), \\ & \varphi_2(t, x, z, \theta) - \eta_2 \varphi_x(t, x, z, \theta), \\ & -\varphi_2(t, x, z, \theta) - \eta_2 \varphi_x(t, x, z, \theta) \Big\} \geq 0.\end{aligned} \tag{3.46}$$

Define the contraposition set

$$\begin{aligned}\mathcal{M}(\varphi) := \Big\{ & (t', x', z', \theta') \in [0, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2 : \tilde{\mathcal{L}}\varphi(t', x', z', \theta') < 0, \\ & \varphi_i(t', x', z', \theta') - \eta_i \varphi_x(t', x', z', \theta') < 0; \quad i \in \{1, 2\}, \\ & -\varphi_i(t', x', z', \theta') - \eta_i \varphi_x(t', x', z', \theta') < 0; \quad i \in \{1, 2\} \Big\},\end{aligned} \tag{3.47}$$

and note that the subsolution property (3.46) holds if and only if $(t, x, z, \theta) \notin \mathcal{M}(\varphi)$. We argue by contradiction, assuming that $(t, x, z, \theta) \in \mathcal{M}(\varphi)$. The set $\mathcal{M}(\varphi)$ is open, since φ is smooth. Therefore, there exist $t_2 \in (t, T]$ and $\xi > 0$ such that

$$\mathfrak{B} := [t, t_2] \times \bar{B}_\xi(x, z, \theta) \subset \mathcal{M}(\varphi),$$

where $\bar{B}_\xi(x, z, \theta)$ is the closed ball of radius ξ and center (x, z, θ) in $\mathbb{R}^4$. It then follows from Lemma 3.11 in the next page that:

$$\sup_{\partial_P(\mathfrak{B})} (u - \varphi) = \max_{\mathfrak{B}} (u^\star - \varphi),$$

where $\partial_P(\mathfrak{B})$ is the forward parabolic boundary of $\mathfrak{B}$ given by

$$\partial_P(\mathfrak{B}) := [t, t_2] \times \partial \bar{B}_\xi(x, z, \theta) \cup t_2 \times \bar{B}_\xi(x, z, \theta).$$

In particular,

$$\sup_{\partial_P(\mathfrak{B})} (u^\star - \varphi) \geq 0,$$

which implies that the upper semicontinuous function $u^\star - \varphi$ attains a non-negative maximum at some point (t', x', z', θ') in the compact set $\partial_P(\mathfrak{B})$, i.e.

$$\varphi(t', x', z', \theta') \leq u^\star(t', x', z', \theta'). \tag{3.48}$$

Since by the definition of $\mathfrak{B}$, (t,x,z,θ) is not a limit point of $\partial_P(\mathfrak{B})$, one has $(t',x',z',\theta') \neq (t,x,z,\theta)$. Therefore, (3.48) is in contradiction with (3.45).

The only remaining part of the proof is the following lemma, which will be also used for deriving the viscosity property at T in Theorem 3.12 to follow.

Lemma 3.11. *Let $\varphi \in \mathcal{C}^{1,2,2,1}([0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2)$, and assume that there exist $t_1 < t_2 \le T$, $(\bar{x},\bar{z},\bar{\theta}) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$, and $\xi > 0$ such that:*

$$\mathfrak{B} := [t_1,t_2]\times\bar{B}_\xi(\bar{x},\bar{z},\bar{\theta}) \subset \mathcal{M}(\varphi),$$

where $\mathcal{M}(\varphi)$ is given by (3.47) and $\bar{B}_\xi(x,z,\theta)$ is the closed ball of radius ξ and center (x,z,θ) in $\mathbb{R}^4$. Then,

$$\sup_{\partial_P(\mathfrak{B})}(u-\varphi) = \max_{\mathfrak{B}}(u^\star - \varphi).$$

Proof of Lemma 3.11. First, observe that by the definition of $\mathcal{M}(\varphi)$, we have

$$\begin{aligned} \tilde{\mathcal{L}}\varphi(t'',x'',z'',\theta'') &< 0, \\ (\varphi_i - \eta_i\varphi_x)(t'',x'',z'',\theta'') &< 0; \quad i \in \{1,2\}, \\ (-\varphi_i - \eta_i\varphi_x)(t'',x'',z'',\theta'') &< 0; \quad i \in \{1,2\}, \end{aligned} \tag{3.49}$$

for all $(t'',x'',z'',\theta'') \in \mathfrak{B}$. We argue by contradiction, assuming that:

$$2\delta := \max_{\mathfrak{B}}(u^\star-\varphi) - \sup_{\partial_P(\mathfrak{B})}(u-\varphi) > 0. \tag{3.50}$$

Let $B_\xi(\bar{x},\bar{z},\bar{\theta})$ be the **open** ball of radius ξ and center $(\bar{x},\bar{z},\bar{\theta})$ in $\mathbb{R}^4$. Then, there exists $(t',x',z',\theta') \in [t_1,t_2)\times B_\xi(\bar{x},\bar{z},\bar{\theta})$ such that

$$\max_{\mathfrak{B}}(u^\star-\varphi) - (u-\varphi)(t',x',z',\theta') \le \delta.$$

By (3.50), it follows that

$$(u-\varphi)(t',x',z',\theta') \ge \delta + \sup_{\partial_P(\mathfrak{B})}(u-\varphi).$$

In particular, for $(t'',x'',z'',\theta'') \in \partial_P(\mathfrak{B})$, one has:

$$u(t'',x'',z'',\theta'') \le \varphi(t'',x'',z'',\theta'') + (u-\varphi)(t',x',z',\theta') - \delta. \tag{3.51}$$

By DPP.(ii), for $\varepsilon := \delta/2$ and the stopping time

$$\tau := \inf\left\{s \geq t' : \left(s, X_s^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_s^{t',\theta'}\right) \notin \mathfrak{B}\right\},$$

we deduce that there exists a policy $(L^\varepsilon, M^\varepsilon) \in \mathcal{A}$ such that:

$$u(t', x', z', \theta'; T, \gamma) - \varepsilon \leq \mathbb{E}\left[u(\tau, X_\tau^{t',x',z',\theta'}, Z_\tau^{t',z'}, \theta_\tau^{t',\theta'}; T, \gamma)\right]. \tag{3.52}$$

Since the jumps in the process $(s, X_s^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_s^{t',\theta'})_{s\in[t',T]}$ only happen by instantaneous trades in the futures, and such trades do not increase the value function, without loss of generality we may assume that $\left(\tau, X_\tau^{t',x',z',\theta'}, Z_\tau^{t',z'}, \theta_\tau^{t',\theta'}\right) \in \partial_P(\mathfrak{B})$. Therefore, we may apply (3.51) to (3.52) to obtain:

$$\mathbb{E}\left[\varphi(\tau, X_\tau^{t',x',z',\theta'}, Z_\tau^{t',z'}, \theta_\tau^{t',\theta'}) - \varphi(t', x', z', \theta')\right] \geq \delta - \varepsilon = \frac{\delta}{2}.$$

Applying Itô's formula to the left side, and noting that the stochastic integral vanishes because of bounded integrand, yields

$$\begin{aligned}
&\mathbb{E}\Big[\int_{t'}^{\tau} \tilde{\mathcal{L}}\varphi(s, X_s^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_s^{t',\theta'})ds\Big] \\
&+ \mathbb{E}\Big[\sum_{i=1}^{2}\Big(\int_{t'}^{\tau}(\varphi_i - \eta_i\varphi_x)(s^-, X_{s^-}^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_{s^-}^{t',\theta'})d(L_s^i - \Delta L_s^i) \\
&\qquad\qquad + \int_{t'}^{\tau}(-\varphi_i - \eta_i\varphi_x)(s^-, X_{s^-}^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_{s^-}^{t',\theta'})d(M_s^i - \Delta M_s^i)\Big)\Big] \\
&+ \sum_{s\in(t',\tau]} \mathbb{E}\Big[\varphi(s, X_s^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_s^{t',\theta'}) - \varphi(s^-, X_{s^-}^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_{s^-}^{t',\theta'})\Big] \geq \frac{\delta}{2}.
\end{aligned} \tag{3.53}$$

Next, we show that the term involving the jumps cannot be positive. To simplify the presentation, we drop the superscripts and define $X := X^{t',x',z',\theta'}$, $Z := Z^{t',z'}$, and $\theta := \theta^{t',\theta'}$. Furthermore, since we argue for a fixed time s, we slightly abuse notation and define $\varphi(x, \theta) = \varphi(s, x, Z_s, \theta)$.

Consider an arbitrary $s \in (t', \tau]$. Without loss of generality, we may assume that $\Delta L_s^i \Delta M_s^i = 0$, $i \in \{1, 2\}$, since it is never optimal to simultaneously buy and sell the same futures. Therefore, there are eight possible trades, four involving only one futures and the other four involving both. Since the same argument applies for all these cases, we only show the argument for the case of buying the first futures and selling the second, i.e. $\Delta L_s^1, \Delta M_s^2 > 0$ and $\Delta M_s^1 = \Delta L_s^2 = 0$. To this end, consider the line

$$\mathfrak{L} := \Big\{ \big(s, X_{s^-} - \xi(\eta_1 \Delta L_s^1 + \eta_2 \Delta M_s^2), Z_s, \theta_{s^-}^1 + \xi \Delta L_s^1, \theta_{s^-}^2 - \xi \Delta M_s^2\big) : \xi \in [0,1] \Big\}$$

Applying the Gradient theorem along $\mathfrak{L}$, we then easily deduce that

$$\begin{aligned}
&\varphi(X_s, \theta_s) - \varphi(X_{s^-}, \theta_{s^-}) = \\
&\Delta L_s^1 \textstyle\int_0^1 (\varphi_1 - \eta_1 \varphi_x)(X_{s^-} - \xi(\eta_1 \Delta L_s^1 + \eta_2 \Delta M_s^2), \theta_{s^-}^1 + \xi \Delta L_s^1, \theta_{s^-}^2 - \xi \Delta M_s^2) d\xi + \\
&\Delta M_s^2 \textstyle\int_0^1 (-\varphi_2 - \eta_2 \varphi_x)(X_{s^-} - \xi(\eta_1 \Delta L_s^1 + \eta_2 \Delta M_s^2), \theta_{s^-}^1 + \xi \Delta L_s^1, \theta_{s^-}^2 - \xi \Delta M_s^2) d\xi.
\end{aligned}$$

On the other hand, by the definition of τ, the right continuity of X and θ, and the closedness of $\mathfrak{B}$, we obtain that the end points of $\mathfrak{L}$, i.e. (s, X_s, Z_s, θ_s) and $(s, X_{s^-}, Z_s, \theta_{s^-})$, are in $\mathfrak{B}$. Therefore, $\mathfrak{L} \subset \mathfrak{B}$ (since $\mathfrak{B}$ is convex), and we may apply (3.49) to the integrands on the right side of the last equation to obtain $\varphi(X_s, \theta_s) - \varphi(X_{s^-}, \theta_{s^-}) \leq 0$. Repeating this argument for the remaining possible trades yields:

$$\varphi(s, X_s^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_s^{t',\theta'}) - \varphi(s^-, X_{s^-}^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_{s^-}^{t',\theta'}) \leq 0, \quad s \in (t', \tau],$$

which, along with (3.53), implies

$$\begin{aligned}
&\mathbb{E}\Big[\int_{t'}^{\tau} \tilde{\mathcal{L}}\varphi(s, X_s^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_s^{t',\theta'}) ds\Big] \\
&+ \mathbb{E}\Big[\sum_{i=1}^{2} \Big(\int_{t'}^{\tau} (\varphi_i - \eta_i \varphi_x)(s^-, X_{s^-}^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_{s^-}^{t',\theta'}) d(L_s^i - \Delta L_s^i) \\
&\qquad + \int_{t'}^{\tau} (-\varphi_i - \eta_i \varphi_x)(s^-, X_{s^-}^{t',x',z',\theta'}, Z_s^{t',z'}, \theta_{s^-}^{t',\theta'}) d(M_s^i - \Delta M_s^i)\Big)\Big] \geq \frac{\delta}{2}.
\end{aligned}$$

Finally, since $(s^-, X_{s^-}^{t',x',z',\theta'}, Z_{s^-}^{t',z'}, \theta_{s^-}^{t',\theta'}) \in \mathfrak{B}$ for $s \in [t', \tau]$, we may apply (3.49) to the integrands on the left side to obtain the required contradiction $0 \geq \frac{\delta}{2}$. □

Next, we turn our attention to identifying the terminal condition which, as in the case all parabolic PDEs, is crucial for obtaining the uniqueness result. The natural candidate is obtained by the very definition of the value function:

$$u\big(T, x, z, (\theta_1, \theta_2); T, \gamma\big) = -e^{-\gamma x + \gamma \eta_1 |\theta_1| + \gamma \eta_2 |\theta_2|}, \quad (x, z, \theta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2. \quad (3.54)$$

But, as mentioned earlier, the value function is not necessarily continuous at T. Therefore, the relevant terminal condition for obtaining the uniqueness result, as well as the convergence theorems for numerical procedures in the next section, should be identified for the left limits:

$$\underline{v}(z,\theta) := \liminf_{t\uparrow T,\ (z',\theta')\to(z,\theta)} v(t,z',\theta') \quad \text{and} \quad \overline{v}(z,\theta) := \limsup_{t\uparrow T,\ (z',\theta')\to(z,\theta)} v(t,z',\theta').$$

The next theorem establishes the viscosity properties of these limits. These properties will be later used to prove the continuity of the value function and the uniqueness of the solution of the HJB variational inequality.

Theorem 3.12. *Consider the variational inequality*

$$\max\left\{ -e^{\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|} - \hat{v},\ \max_{i\in\{1,2\}} \left\{ \gamma\eta_i\hat{v} + \frac{\partial\hat{v}}{\partial\theta_i},\ \gamma\eta_i\hat{v} - \frac{\partial\hat{v}}{\partial\theta_i} \right\} \right\} = 0. \tag{3.55}$$

Then:

(i) $\underline{v}(z,\theta) := \liminf_{t\uparrow T,\ (z',\theta')\to(z,\theta)} v(t,z',\theta')$, $(z,\theta)\in\mathbb{R}\times\mathbb{R}^2$, *is a viscosity supersolution of (3.55).*

(ii) $\overline{v}(z,\theta) := \limsup_{t\uparrow T,\ (z',\theta')\to(z,\theta)} v(t,z',\theta')$, $(z,\theta)\in\mathbb{R}\times\mathbb{R}^2$, *is a viscosity subsolution of (3.55).*

Proof. Define:

$$\underline{u}(x,z,\theta) := \liminf_{t\uparrow T,\ (x',z',\theta')\to(x,z,\theta)} u(t,x',z',\theta'),$$

and

$$\overline{u}(x,z,\theta) := \limsup_{t\uparrow T,\ (x',z',\theta')\to(x,z,\theta)} u(t,x',z',\theta').$$

We prove that the function $\underline{u}$ (resp. $\overline{u}$) is a viscosity supersolution (resp. subsolution) of

$$\max\left\{ -e^{-\gamma x+\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|} - \hat{u},\ \max_{i\in\{1,2\}} \left\{ \frac{\partial\hat{u}}{\partial\theta_i} - \eta_i\frac{\partial\hat{u}}{\partial x},\ -\frac{\partial\hat{u}}{\partial\theta_i} - \eta_i\frac{\partial\hat{u}}{\partial x} \right\} \right\} = 0, \tag{3.56}$$

for $(x,z,\theta)\in\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$. Then, the corresponding viscosity properties for $\underline{v}$ and $\overline{v}$ follow from the scaling properties $\underline{u}(x,z,\theta) = e^{-\gamma x}\underline{v}(z,\theta)$ and $\overline{u}(x,z,\theta) =$

$e^{-\gamma x}\overline{v}(z,\theta)$ which are inherited from the value function. The rest of the proof is the modified version of the arguments in (Pham, 2005, section 3.2.3, p. 528).

(i): $\underline{u}$ is a supersolution of (3.56)
Let $(x,z,\theta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$ and consider an arbitrary test function $\varphi \in \mathcal{C}^{2,2,1}(\mathbb{R}^4)$ such that $\varphi(x,z,\theta) = \underline{u}(x,z,\theta)$ and

$$\varphi(x',z',\theta') \leq \underline{u}(x',z',\theta'), \quad (x',z',\theta') \in \mathbb{R}^4. \tag{3.57}$$

We need to show that

$$\max\left\{ -e^{-\gamma x+\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|} - \varphi,\ \max_{i\in\{1,2\}}\left\{\frac{\partial\varphi}{\partial\theta_i} - \eta_i\frac{\partial\varphi}{\partial x},\ -\frac{\partial\varphi}{\partial\theta_i} - \eta_i\frac{\partial\varphi}{\partial x}\right\}\right\} \leq 0.$$

Therefore, we must prove:

$$\varphi(x,z,\theta) \geq -e^{-\gamma x+\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|}, \tag{3.58}$$

$$\left(\frac{\partial\varphi}{\partial\theta_i} - \eta_i\frac{\partial\varphi}{\partial x}\right)(x,z,\theta) \leq 0, \quad i \in \{1,2\}, \tag{3.59}$$

and

$$\left(-\frac{\partial\varphi}{\partial\theta_i} - \eta_i\frac{\partial\varphi}{\partial x}\right)(x,z,\theta) \leq 0, \quad i \in \{1,2\}. \tag{3.60}$$

<u>Proof of (3.58):</u>
Since $\varphi(x,z,\theta) = \underline{u}(x,z,\theta)$, (3.58) is equivalent to

$$\underline{u}\big(x,z,(\theta_1,\theta_2)\big) \geq -e^{-\gamma x+\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|}. \tag{3.61}$$

Take an arbitrary sequence $(t_m,x_m,z_m,\theta_m) \to (T,x,z,\theta)$ where $t_m < T$, and define $(\tilde{X}^m_s)_{s\geq t_m}$ as the margin account if the agent starts at t_m with the initial state (x_m,z_m,θ_m) and does not trade afterwards. Next, fix a constant $b > 0$, and define the stopping time τ_m as the first time that the margin account $\tilde{X}^m$ changes by at least b. Then, the strategy (L^m,M^m) that unwinds the initial position θ_m at τ_m is admissible. Let $(X^m_s)_{s\in[t_m,\infty]}$ be the corresponding margin account. The definition of the value function implies:

$$u(t_m,x_m,z_m,\theta_m) \geq \mathbb{E}\Big[-\exp\Big(-\gamma X^m_T + \mathbb{1}_{\{T<\tau_m\}}(\gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|)\Big)\Big]. \tag{3.62}$$

By the admissibility of (L^m, M^m), the random variable

$$-\exp\Big(-\gamma X_T^m + \mathbb{1}_{\{T<\tau_m\}}(\gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|)\Big)$$

is bounded from below. Therefore, Fatou's lemma yields that

$$\begin{aligned}\underline{u}(x,z,\theta) &:= \liminf_{m\to\infty} u(t_m, x_m, z_m, \theta_m)\\ &\geq \mathbb{E}\Big[\liminf_{m\to\infty}\Big\{-\exp\Big(-\gamma X_T^m + \mathbb{1}_{\{T<\tau_m\}}(\gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|)\Big)\Big\}\Big]\\ &= -e^{-\gamma x + \gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|},\end{aligned}$$

where in the last step we used that $\lim_{m\to\infty} X_T^m = x$ and that $\lim_{m\to\infty} \mathbb{1}_{\{T<\tau_m\}} = 1$ which, in turn, hold because of the continuity of the path of $\tilde{X}^m$.

Proof of (3.59) and (3.60):

Let $u_\star$ be the lower semicontinuous envelop of the value function u, c.f. (3.34). Then, by definition, there exists a sequence $\{(t'_m, x'_m, z'_m, \theta'_m)\}$ converging to (T, x, z, θ) such that $t_m < T$ and

$$\lim_{m\to\infty} u_\star(t'_m, x'_m, z'_m, \theta'_m) = \underline{u}(x,z,\theta). \tag{3.63}$$

Consider the auxiliary test functions

$$\varphi_m(t', x', z', \theta') := \varphi(x', z', \theta') + \frac{T-t'}{(T-t'_m)^2} - \|(x'-x, z'-z, \theta'-\theta)\|^4, \tag{3.64}$$

for $(t', x', z', \theta') \in [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$, and define the sequence $\{(t_m, x_m, z_m, \theta_m)\}$ as follows:

$$\begin{aligned}(t_m, x_m, z_m, \theta_m) := \arg\min\Big\{&(u_\star - \varphi_m)(t', x', z', \theta') :\\ &(t', x', z', \theta') \in [t'_m, T]\times \bar{B}_1(x,z,\theta)\Big\},\end{aligned} \tag{3.65}$$

where $\bar{B}_1(x,z,\theta)$ is the closed ball of radius 1 and center (x,z,θ) in $\mathbb{R}^4$. We will obtain (3.59) and (3.60) from the supersolution property of u at $(t_m, x_m, z_m, \theta_m)$ and then taking limit as $m\to\infty$. To do so, we need to show that:

(a) The supersolution property holds for u at $(t_m, x_m, z_m, \theta_m)$. In particular, for $i \in \{1,2\}$,

$$\Big(\frac{\partial\varphi_m}{\partial\theta_i} - \eta_i\frac{\partial\varphi_m}{\partial x}\Big)(t_m, x_m, z_m, \theta_m) \leq 0, \tag{3.66}$$

and

$$(-\frac{\partial \varphi_m}{\partial \theta_i} - \eta_i \frac{\partial \varphi_m}{\partial x})(t_m, x_m, z_m, \theta_m) \leq 0. \tag{3.67}$$

(b) Up to a subsequence, $(x_m, z_m, \theta_m) \to (x, z, \theta)$.

Let us assume for now that we have proved (a) and (b). From (3.64), we have that

$$\frac{\partial \varphi_m}{\partial \theta'_i} = \frac{\partial \varphi}{\partial \theta'_i} - 4(\theta'_i - \theta_i)\|(x' - x, z' - z, \theta' - \theta)\|^2,$$

and

$$\frac{\partial \varphi_m}{\partial x'} = \frac{\partial \varphi}{\partial x'_i} - 4(x' - x)\|(x' - x, z' - z, \theta' - \theta)\|^2.$$

Then, the inequalities (3.59) and (3.60) are obtained by passing to the limit in (3.66) and (3.66), having restricted ourselves to the subsequence in (b).

It then only remains to prove (a) and (b). To show (a), it suffices to show that $t_m < T$ for sufficiently large m. Then, the supersolution property would follow from Theorem 3.10. Note that by (3.64), for all m one has:

$$\varphi(x'_m, z'_m, \theta'_m) - \varphi_m(t'_m, x'_m, z'_m, \theta'_m) - \|(x'_m - x, z'_m - z, \theta'_m - \theta)\|^4 = -\frac{1}{T - t'_m}.$$

As $m \to \infty$, the right side of this equality tends to $-\infty$ while, because of (3.63) and that *inequalities*$\varphi(x, z, \theta) = \underline{u}(x, z, \theta)$, the left side approaches $(u_\star - \varphi_m)(t'_m, x'_m, z'_m, \theta'_m)$. Therefore, for sufficiently large m, one has

$$(u_\star - \varphi_m)(t'_m, x'_m, z'_m, \theta'_m) < 0. \tag{3.68}$$

On the other hand, for any $(x', z', \theta') \in \mathbb{R}^4$, (3.64) yields:

$$\begin{aligned}(u_\star - \varphi_m)(T, x', z', \theta') &=: (\underline{u} - \varphi)(x', z', \theta') + \|(x' - x, z' - z, \theta' - \theta)\|^4 \\ &\geq (\underline{u} - \varphi)(x', z', \theta') \geq 0,\end{aligned} \tag{3.69}$$

where the last inequality is because of (3.57). Finally, since t_m is defined to be the time when the minimum of $u_\star - \varphi_m$ happens, the inequalities (3.68) and (3.69) imply that $t_m < T$ for sufficiently large m.

Lastly, to show (b), observe that since $\{(x_m, z_m, \theta_m)\} \in \bar{B}_1(x, z, \theta)$, there exists a subsequence which converges to $(x_0, z_0, \theta_0) \in \bar{B}_1(x, z, \theta)$. With a slight abuse of notation, we also denote the convergent sequence by $\{(x_m, z_m, \theta_m)\}$. It then follows from (3.57) and (3.65) that:

$$
\begin{aligned}
0 &\le (\underline{u} - \varphi)(x_0, z_0, \theta_0) - (\underline{u} - \varphi)(x, z, \theta) \\
&= \liminf_{m\to\infty} \Big[(\underline{u} - \varphi)(x_m, z_m, \theta_m) - (\underline{u} - \varphi)(x'_m, z'_m, \theta'_m)\Big] \\
&=: \liminf_{m\to\infty} \Big[(u_\star - \varphi_m)(t_m, x_m, z_m, \theta_m) - (u_\star - \varphi_m)(t'_m, x'_m, z'_m, \theta'_m) \\
&\qquad\qquad - \big\|(x_m - x, z_m - z, \theta_m - \theta)\big\|^4 - \frac{t_m - t'_m}{(T - t'_m)^2}\Big] \\
&\le \liminf_{m\to\infty} \Big[- \big\|(x_m - x, z_m - z, \theta_m - \theta)\big\|^4\Big] = -\big\|(x_0 - x, z_0 - z, \theta_0 - \theta)\big\|^4,
\end{aligned}
$$

which proves that $x_0 = x$, $z_0 = z$, and $\theta_0 = \theta$.

(ii): $\overline{u}$ is a subsolution of (3.56)

Let $(x, z, \theta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$ and consider an arbitrary test function $\varphi \in \mathcal{C}^{2,2,1}(\mathbb{R}^4)$ such that $\varphi(x, z, \theta) = \overline{u}(x, z, \theta)$ and

$$
\varphi(x', z', \theta') \ge \overline{u}(x', z', \theta'), \quad (x', z', \theta') \in \mathbb{R}^4. \tag{3.70}
$$

We are going to show that

$$
\max\left\{ - e^{-\gamma x + \gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|} - \varphi,\ \max_{i\in\{1,2\}} \left\{ \frac{\partial\varphi}{\partial\theta_i} - \eta_i \frac{\partial\varphi}{\partial x}, -\frac{\partial\varphi}{\partial\theta_i} - \eta_i \frac{\partial\varphi}{\partial x} \right\} \right\} \ge 0,
$$

Note that by (3.61), one has

$$
\overline{u}(x, z, \theta) \ge \underline{u}(x, z, \theta) \ge -e^{-\gamma x + \gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|}. \tag{3.71}
$$

Therefore, we only need to show that if

$$
\varphi(x, z, \theta) = \overline{u}(x, z, \theta) > -e^{-\gamma x + \gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|}, \tag{3.72}
$$

then

$$
\max_{i\in\{1,2\}} \left\{ \frac{\partial\varphi}{\partial\theta_i} - \eta_i \frac{\partial\varphi}{\partial x}, -\frac{\partial\varphi}{\partial\theta_i} - \eta_i \frac{\partial\varphi}{\partial x} \right\} \ge 0.
$$

We argue by contradiction, and assume that (3.72) holds while

$$
\max_{i\in\{1,2\}} \left\{ \frac{\partial\varphi}{\partial\theta_i} - \eta_i \frac{\partial\varphi}{\partial x}, -\frac{\partial\varphi}{\partial\theta_i} - \eta_i \frac{\partial\varphi}{\partial x} \right\} < 0. \tag{3.73}
$$

For $m \in \mathbb{N}$, define the auxiliary test functions

$$\varphi_m(t', x', z', \theta') = \varphi(x', z', \theta') + m(T - t) + \|(x' - x, z' - z, \theta' - \theta)\|^4. \quad (3.74)$$

As $(x', z', \theta') \to (x, z, \theta)$, one has

$$\frac{\partial \varphi_m}{\partial x'}(t', x', z', \theta') = \frac{\partial \varphi}{\partial x'} + 4(x' - x)\|(x' - x, z' - z, \theta' - \theta)\|^2 \to \frac{\partial \varphi}{\partial x'}(x, z, \theta)$$

and

$$\frac{\partial \varphi_m}{\partial \theta'_i}(t', x', z', \theta') = \frac{\partial \varphi}{\partial \theta'_i} + 4(\theta'_i - \theta_i)\|(x' - x, z' - z, \theta' - \theta)\|^2 \to \frac{\partial \varphi}{\partial \theta'_i}(x, z, \theta).$$

Therefore, by (3.73) and the smoothness of φ, there exist $t_0 < T$ and $\xi > 0$ such that

$$\max_{i \in \{1,2\}} \left\{ \Big(\frac{\partial \varphi_m}{\partial \theta'_i} - \eta_i \frac{\partial \varphi_m}{\partial x'} \Big)(t', x', z', \theta'),\ \Big(-\frac{\partial \varphi_m}{\partial \theta'_i} - \eta_i \frac{\partial \varphi_m}{\partial x'} \Big)(t', x', z', \theta') \right\} < 0, \quad (3.75)$$

for all $(t', x', z', \theta') \in [t_0, T] \times \bar{B}_\xi(x, z, \theta)$ and $m \in \mathbb{N}$.

In the remaining part of the proof, we obtain the contradiction by showing that the opposite of (3.75) must be true, i.e. that there exists a point $(t', x', z', \theta') \in [t_0, T] \times \bar{B}_\xi(x, z, \theta)$ and $m \in \mathbb{N}$ such that

$$\max_{i \in \{1,2\}} \left\{ \Big(\frac{\partial \varphi_m}{\partial \theta'_i} - \eta_i \frac{\partial \varphi_m}{\partial x'} \Big)(t', x', z', \theta'),\ \Big(-\frac{\partial \varphi_m}{\partial \theta'_i} - \eta_i \frac{\partial \varphi_m}{\partial x'} \Big)(t', x', z', \theta') \right\} \geq 0. \quad (3.76)$$

The arguments are divided into four steps:

Step 1:

We show that

$$\limsup_{m \to \infty} \left(\sup \Big\{ (u - \varphi_m)(T, x', z', \theta') : (x', z', \theta') \in B_\xi(x, z, \theta) \Big\} \right) < 0. \quad (3.77)$$

First, observe that $\sup \Big\{ (u - \varphi_m)(T, x', z', \theta') : (x', z', \theta') \in B_\xi(x, z, \theta) \Big\} \leq 0$, since, by (3.71) and (3.70):

$$\begin{aligned}(u - \varphi_m)(T, x', z', \theta') &= -e^{-\gamma x' + \gamma \eta_1 |\theta'_1| + \gamma \eta_2 |\theta'_2|} - \varphi(x', z', \theta') - \|(x' - x, z' - z, \theta' - \theta)\|^4 \\ &\leq (\overline{u} - \varphi)(x', z', \theta') - \|(x' - x, z' - z, \theta' - \theta)\|^4 \\ &\leq -\|(x' - x, z' - z, \theta' - \theta)\|^4 \leq 0. \quad (3.78)\end{aligned}$$

To show (3.77), assume, by contradiction, that

$$\limsup_{m\to\infty} \sup\left\{(u-\varphi_m)(T,x',z',\theta') : (x',z',\theta') \in B_\xi(x,z,\theta)\right\} = 0.$$

Then, there exists a subsequence of $\{\varphi_m\}$ (which by a slight abuse of notation is also denoted by $\{\varphi_m\}$) such that

$$\lim_{m\to\infty} \sup\left\{(u-\varphi_m)(T,x',z',\theta') : (x',z',\theta') \in B_\xi(x,z,\theta)\right\} = 0.$$

For each m, let $\{(x_{m,k}, z_{m,k}, \theta_{m,k})\}_k$ be a maximizing sequence for $(u-\varphi_m)(T,\cdot)$ in $B_\xi(x,z,\theta)$, i.e.

$$\lim_{k\to\infty}(u-\varphi_m)(x_{m,k}, z_{m,k}, \theta_{m,k}) = \sup\left\{(u-\varphi_m)(T,x',z',\theta') : (x',z',\theta') \in B_\xi(x,z,\theta)\right\}.$$

Then,

$$\lim_{m\to\infty}\lim_{k\to\infty}(u-\varphi_m)(T, x_{m,k}, z_{m,k}, \theta_{m,k}) = 0,$$

and it follows that

$$\lim_{m\to\infty}\lim_{k\to\infty}(x_{m,k}, z_{m,k}, \theta_{m,k}) = (x,z,\theta),$$

because by (3.78), one has

$$(u-\varphi_m)(T, x_{m,k}, z_{m,k}, \theta_{m,k}) \leq -\|(x_{m,k}-x, z_{m,k}-z, \theta_{m,k}-\theta)\|^4.$$

Moreover,

$$\begin{aligned}
0 &= \lim_{m\to\infty}\lim_{k\to\infty}(u-\varphi_m)(T, x_{m,k}, z_{m,k}, \theta_{m,k}) \\
&= -e^{-\gamma x+\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|} - \varphi(x,z,\theta) < (\overline{u}-\varphi)(x,z,\theta),
\end{aligned}$$

where the last inequality follows from by (3.72). This contradicts that

$$(\overline{u}-\varphi)(x,z,\theta) = 0.$$

<u>Step 2:</u>
We show that there exists a sequence $\{t_m\} \to T$ such that $t_0 \leq t_m < T$ and

there exists a subsequence of $\{\varphi_m\}$ (also denoted by $\{\varphi_m\}$ with a slight abuse of notation) such that

$$\limsup_{m\to\infty}\Big(\sup_{[t_m,T]\times\partial\bar{B}_\xi(x,z,\theta)}(u-\varphi_m)\Big)<0. \tag{3.79}$$

Consider an arbitrary sequence $\{t''_m\}\to T$ such that $t_0\le t''_m<T$, and let $\{(t'_m,x'_m,z'_m,\theta'_m)\}$ be a sequence of maximum points of $u^\star-\varphi_m$ on $[t''_m,T]\times\partial\bar{B}_\xi(x,z,\theta)$, i.e.

$$(t'_m,x'_m,z'_m,\theta'_m):=\operatorname*{arg\,max}_{[t''_m,T]\times\partial\bar{B}_\xi(x,z,\theta)}(u^\star-\varphi_m).$$

Then, because of the compactness of $\partial\bar{B}_\xi(x,z,\theta)$, there exists a subsequence of $\{(t'_m,x'_m,z'_m,\theta'_m)\}$ (also denoted by $\{(t'_m,x'_m,z'_m,\theta'_m)\}$) such that

$$\{(t'_m,x'_m,z'_m,\theta'_m)\}\to(T,x_0,z_0,\theta_0),$$

for some $(x_0,z_0,\theta_0)\in\partial\bar{B}_\xi(x,z,\theta)$. Again, with a slight abuse of notation, let $\{\varphi_m\}$ and $\{t_m\}$ be the corresponding subsequences of $\{\varphi_m\}$ and $\{t''_m\}$, respectively. Finally, (3.79) is obtained as follows:

$$\begin{aligned}
&\limsup\nolimits_{m\to\infty}\Big(\sup\nolimits_{[t_m,T]\times\partial\bar{B}_\xi(x,z,\theta)}(u-\varphi_m)\Big)\\
&\le\limsup\nolimits_{m\to\infty}\Big(\sup\nolimits_{[t_m,T]\times\partial\bar{B}_\xi(x,z,\theta)}(u^\star-\varphi_m)\Big)\\
&\le\limsup\nolimits_{m\to\infty}\big(u^\star(t'_m,x'_m,z'_m,\theta'_m)-\varphi(x'_m,z'_m,\theta'_m)\big)-\xi^4\\
&=(\overline{u}-\varphi)(x_0,z_0,\theta_0)-\xi^4\le-\xi^4<0,
\end{aligned}$$

where we used (3.74), the definition of $\overline{u}$, and (3.70) in the last two lines.

Step 3:

Let $\{t_m\}$ and $\{\varphi_m\}$ be the sequences found in Step 2. We show that

$$[t_m,T]\times\bar{B}_\xi(x,z,\theta)\nsubseteq\mathcal{M}(\varphi_m), \tag{3.80}$$

where $\mathcal{M}(\varphi_m)$ is given by (3.47). Indeed, (3.77) and (3.79) yield that, for sufficiently large m,

$$\sup_{\partial_P\big([t_m,T]\times\bar{B}_\xi(x,z,\theta)\big)}(u-\varphi_m)<0, \tag{3.81}$$

where $\partial_P([t_m, T] \times \bar{B}_\xi(x, z, \theta)) := [t_m, T] \times \partial \bar{B}_\xi(x, z, \theta) \cup T \times \bar{B}_\xi(x, z, \theta)$. On the other hand,

$$\max_{[t_m,T]\times\bar{B}_\xi(x,z,\theta)} (u^\star - \varphi_m) \geq (u^\star - \varphi_m)(T, x, z, \theta) := (\overline{u} - \varphi)(x, z, \theta) = 0. \quad (3.82)$$

Finally, combining (3.81) and (3.82) yields:

$$\sup_{\partial_P\left([t_m,T]\times\bar{B}_\xi(x,z,\theta)\right)} (u - \varphi_m) < \max_{[t_m,T]\times\bar{B}_\xi(x,z,\theta)} (u^\star - \varphi_m)$$

which, in turn, implies (3.80) by Lemma 3.11.

Step 4:
Let $\{t_m\}$ and $\{\varphi_m\}$ be the sequences in Step 2. By (3.74), we have

$$\tilde{\mathcal{L}}\varphi_m(t', x', z', \theta') = \tilde{\mathcal{L}}\Big(\varphi(x', z', \theta') + \|(x' - x, z' - z, \theta' - \theta)\|^4\Big) - m.$$

Moreover, by the smoothness of φ, the function $\mathcal{L}\Big(\varphi(x', z', \theta') + \|(x' - x, z' - z, \theta' - \theta)\|^4\Big)$ is bounded on the compact set $[t_0, T] \times \bar{B}_\xi(x, z, \theta)$. Therefore,

$$\tilde{\mathcal{L}}\varphi_m(t', x', z', \theta') \leq h_0 - m, \quad \forall (t', x', z', \theta') \in [t_0, T] \times \bar{B}_\xi(x, z, \theta), \quad (3.83)$$

for some constant $h_0 < \infty$ independent of m. It follows that for sufficiently large m,

$$\tilde{\mathcal{L}}\varphi_m(t', x', z', \theta') < 0, \qquad \forall (t', x', z', \theta') \in [t_m, T] \times \bar{B}_\xi(x, z, \theta).$$

Hence, by (3.80) and the definition $\mathcal{M}$, i.e. (3.47), it follows that there must be $(t', x', z', \theta') \in [t_m, T] \times \bar{B}_\xi(x, z, \theta)$ such that (3.76) holds. $\square$

Finally, we show the continuity of the value function and the uniqueness of the solution to the HJB equation. We start by proving the continuity of the value function at T and characterizing the terminal data as the natural candidate in (3.54).

Theorem 3.13. *Let $\underline{v}$ and $\overline{v}$ be as given in Theorem 3.12. Then,*

$$\underline{v}(z, \theta) = \overline{v}(z, \theta) = \lim_{t\uparrow T,\, (z',\theta')\to(z,\theta)} v(t, z', \theta') = -e^{\gamma\eta_1|\theta_1| + \gamma\eta_2|\theta_2|}, \quad (z, \theta) \in \mathbb{R}^3. \quad (3.84)$$

Proof. Define

$$F\big(\mathbf{x}, r, \mathbf{p}\big) := \min\Big\{e^{\gamma\eta_1|x_1|+\gamma\eta_2|x_2|-\gamma x_3} + r,\ \min_{i\in\{1,2\}}\big\{\eta_i p_3 - p_i,\ \eta_i p_3 + p_i\big\}\Big\}, \tag{3.85}$$

and note that inequality (3.56) can be written as

$$F\Big((x,\theta_1,\theta_2), \hat{u}, (\frac{\partial \hat{u}}{\partial \theta_1}, \frac{\partial \hat{u}}{\partial \theta_2}, \frac{\partial \hat{u}}{\partial x})\Big) = 0.$$

There are quite general comparison results for this first order PDE which state that any upper-semicontinuous sub-solution is less than or equal to a lower-semicontinuous supersolution, see, for example, Crandall et al. (1987). Combining the comparison result with Theorem 3.12 yields that $\overline{u} \leq \underline{u}$. On the other hand, by definition, $\overline{u} \geq \underline{u}$. Hence, $\overline{u} = \underline{u}$ and the limit

$$\overline{v}(z,\theta) = \lim_{t\uparrow T,\ (z',\theta')\to(z,\theta)} v(t,z',\theta')$$

exists and is the unique viscosity solution of (3.56). Finally, a direct substitution yields that the function $-e^{-\gamma x+\gamma\eta_1|\theta_1|+\gamma\eta_2|\theta_2|}$ is a classical solution of (3.56). $\square$

Theorems (3.10) and (3.13) yield that the scaled value function satisfies the Cauchy problem (3.28) and (3.29). Similar to the elliptic case, there is a comparison principle which asserts that any upper-semicontinuous subsolution of (3.28) is less than or equal to a lower-semicontinuous supersolution, c.f. Crandall et al. (1992). Hence, $v^\star \leq v_\star$ on $[0,T)\times\mathbb{R}^3$.

On the other hand, $v^\star \geq v_\star$ by definition (3.34). Therefore, $v^\star = v_\star$ and v is the unique and continuous viscosity solution of (3.28) on the intermediate domain $[0,T)\times\mathbb{R}^3$. This concludes the proof of the main result of this section, i.e. Theorem 3.8.

By applying a simple exponential transformation, the variational inequality (3.28) can be simplified even further. To the best of our knowledge, the first use of exponential transformation (3.86) for simplifying an HJB equation was in Pham (2002), see also Dai and Yi (2009) for its application in singular control case. Both of those studies assumed power utility functions, and, it is quite interesting that the transformation also works for the exponential case. Apart from simplifying the equations, the exponential transformation plays a more significant role. Indeed, the transformed variational inequality (3.88) is numerically more stable than (3.28), see Dai and Zhong (2010).

Proposition 3.14. *Let the functions* $v : [0,T] \times \mathbb{R}^3 \to \mathbb{R}^-$ *and* $w : [0,T] \times \mathbb{R}^3 \to \mathbb{R}$ *be such that*

$$v(t, z, \theta_1, \theta_2) = -e^{w(t,z,\gamma\theta_1,\gamma\theta_2)}. \tag{3.86}$$

and consider the operator

$$\begin{aligned}\mathfrak{L}\, w(t, z, \xi) := \beta^\top(\alpha z - \Sigma\Sigma^\top \xi) w_z + \tfrac{1}{2}\beta^\top \Sigma\Sigma^\top \beta (w_z^2 + w_{zz}) \\ + \tfrac{1}{2}\xi^\top \Sigma\Sigma^\top \xi - \xi^\top \alpha z.\end{aligned} \tag{3.87}$$

Then, v is a viscosity solution of (3.28) and (3.29) if and only if w is a viscosity solution of the variational inequality

$$\min_{i\in\{1,2\}} \Big\{ w_t + \mathfrak{L} w,\ \eta_i \pm w_i \Big\} = 0; \quad (t, z, \xi_1, \xi_2) \in [0,T) \times \mathbb{R}^3, \tag{3.88}$$

with the terminal condition $w(T, z, \xi_1, \xi_2) = \eta_1|\xi_1| + \eta_2|\xi_2|$, $(z, \xi_1, \xi_2) \in \mathbb{R}^3$.

Proof. let $\varphi, \psi \in \mathcal{C}^{1,2,1,1}([0,T) \times \mathbb{R}^3)$ be two test functions such that

$$\varphi(t, z, \theta_1, \theta_2) = -e^{\psi(t,z,\gamma\theta_1,\gamma\theta_2)}. \tag{3.89}$$

Then,

$$\begin{gathered}\varphi_t = \varphi\psi_t, \quad \varphi_z = \varphi\psi_z, \quad \varphi_{zz} = \varphi(\psi_z^2 + \psi_{zz}), \\ \varphi_i = \gamma\varphi\psi_i, \quad i \in \{1,2\},\end{gathered}$$

and

$$\mathcal{L}\varphi(t, z, \theta_1, \theta_2) = \varphi(t, z, \theta_1, \theta_2)\Big(\psi_t + \mathfrak{L}\psi\Big)(t, z, \gamma\theta_1, \gamma\theta_2).$$

In particular, because $\varphi < 0$, we have

$$\max_{i\in\{1,2\}} \Big\{ \mathcal{L}\varphi,\ \gamma\eta_i\varphi \pm \varphi_i \Big\} = \varphi \min_{i\in\{1,2\}} \Big\{ \psi_t + \mathfrak{L}\psi,\ \gamma(\eta_i \pm \psi_i) \Big\}.$$

Furthermore,

$$\min_{i\in\{1,2\}} \Big\{ \psi_t + \mathfrak{L}\psi,\ \gamma(\eta_i \pm \psi_i) \Big\} = 0,$$

if and only if

$$\min_{i\in\{1,2\}} \Big\{ \psi_t + \mathfrak{L}\psi,\ \eta_i \pm \psi_i \Big\} = 0.$$

It then follows that v is a viscosity sub (resp. super) solution of (3.28) if and only if w is a viscosity sub (resp. super) solution of (3.88), and the statement follows.

Note that, in the definition of viscosity solution, we may assume the properties of the test functions to be local, see, for example, (Pham, 2005, Remark 3.1.(2), pp. 522). Therefore, for the statement

$$w \text{ is a sub solution of (3.88)} \quad \Rightarrow \quad v \text{ is a sub solution of (3.28)},$$

it suffices to consider only test functions such that $v^\star \leq \varphi < 0$, such that (3.89) can be true. □

It is worth mentioning that the transformed variational inequality (3.88) does not depend on the risk aversion parameter γ, and the reliance of the scaled value function v on γ is explicit in the exponential transformation (3.86). This can be quite useful when analyzing the dependence of the value function or optimal policies on the risk aversion of the agent, since the variational inequality needs to be solved only once instead for every value of γ.

In light of Theorem 3.8, the transformed variational inequality (3.88) has a unique continuous viscosity solution which is related to the (scaled) value function through (3.86). The following result summarizes our findings in this section.

Corollary 3.15. *The value function (3.20) is given by*

$$u(t, x, z, \theta; T, \gamma) = -e^{-\gamma x + w(t,z,\gamma\theta;T)} \tag{3.90}$$

where w is the unique continuous viscosity solution of the variational inequality (3.88).

3.4 Penalty method for the HJB equation

As mentioned in the introduction, we employ the so-called penalty method, c.f. Forsyth and Vetzal (2002), to numerically approximate the solution of the variational inequality (3.88). The idea behind the penalty method is as follows. First observe that w is a (classical) solution of the variational inequality (3.88) if and only if the following conditions hold:

(i) $w_t + \mathfrak{L}w \geq 0$.

(ii) $-\eta_i \leq w_i \leq \eta_i$, $i \in \{1, 2\}$.

(iii) If $-\eta_i < w_i < \eta_i$, $i \in \{1, 2\}$, then $\mathfrak{L}w = 0$.

In the penalty method, one replaces (3.88) with the **penalized** PDE:

$$\begin{cases} w_t + \mathfrak{L}w = \dfrac{1}{\varepsilon^2} \displaystyle\sum_{i=1}^{2} (|w_i| - \eta_i + \varepsilon)^+; \quad (t, z, \xi_1, \xi_2) \in [0, T) \times \mathbb{R}^3, \\ w_t + \mathfrak{L}w \leq \varepsilon^{-1}; \quad (t, z, \xi_1, \xi_2) \in [0, T) \times \mathbb{R}^3, \\ w(T, z, \xi_1, \xi_2) = \eta_1 |\xi_1| + \eta_2 |\xi_2|; \quad (z, \xi_1, \xi_2) \in \mathbb{R}^3. \end{cases} \tag{3.91}$$

for some **tolerance** $\varepsilon > 0$. It is easily shown that a solution w^ε of (3.91) satisfies:

$(i)'$ $0 \leq w_t^\varepsilon + \mathfrak{L}w^\varepsilon \leq \varepsilon^{-1}$.

$(ii)'$ $-\eta_i \leq w_i^\varepsilon \leq \eta_i$, $i \in \{1, 2\}$.

$(iii)'$ If $-\eta_i + \varepsilon \leq w_i^\varepsilon \leq \eta_i - \varepsilon$, $i \in \{1, 2\}$, then $\mathfrak{L}w^\varepsilon = 0$.

Now, if $\varepsilon \downarrow 0$, then conditions $(i)'$–$(iii)'$ become conditions (i)–(iii). Therefore, $w^\varepsilon \to w$, with w satisfying the original variational inequality (3.88). By applying this simple argument to test functions and noting that the viscosity solutions are quite stable to pass to limits, see, for example, (Crandall et al., 1992, Section 6), one obtains the following result.

Lemma 3.16. *If the viscosity solutions of a family of penalized PDEs (3.91), parametrized by ε, converge as $\varepsilon \downarrow 0$, then the limit is the viscosity solution of the variational inequality (3.88).*

Therefore, the problem has been reduced to the numerical approximation of the nonlinear PDE (3.91) for some small value of ε.

3.4.1 Truncating the domain and the boundary conditions

To implement any numerical scheme, one needs to restrict the unbounded domain of the value function. To this end, we choose the bounds $\bar{z}, \bar{\xi}_1, \bar{\xi}_2 > 0$, and define the **truncated domain** $\mathcal{O}$ and its closure $\overline{\mathcal{O}}$ by:

$$\mathcal{O} := [-\bar{z}, \bar{z}] \times (-\bar{\xi}_1, \bar{\xi}_1) \times (-\bar{\xi}_2, \bar{\xi}_2), \quad \text{and} \quad \overline{\mathcal{O}} := [-\bar{z}, \bar{z}] \times [-\bar{\xi}_1, \bar{\xi}_1] \times [-\bar{\xi}_2, \bar{\xi}_2].$$

Then, the truncated version of (3.91) is:

$$\begin{aligned} &w_t = -\mathfrak{L}w + \frac{1}{\varepsilon^2}\sum_{i=1}^{2}(|w_i| - \eta_i + \varepsilon)^+;\ (t, z, \xi_1, \xi_2) \in [0, T) \times \mathcal{O}, \\ &w_t + \mathfrak{L}w \leq \varepsilon^{-1};\ (t, z, \xi_1, \xi_2) \in [0, T) \times \mathcal{O}. \end{aligned} \tag{3.92}$$

To obtain the boundary conditions at $z = \pm\bar{z}$, we assume

$$w_{zz}(t, \pm\bar{z}, \xi_1, \xi_2) = 0;\ (t, \xi_1, \xi_2) \in [0, T) \times (-\bar{\xi}_1, \bar{\xi}_1) \times (-\bar{\xi}_2, \bar{\xi}_2),$$

which is a common assumption for truncating unbounded domains, see Wilmott (1998), Tavella and Randall (2000), and Windcliff et al. (2004). This assumption results in the boundary conditions:

$$\begin{aligned} &w_t = -\mathfrak{L}_{\pm\bar{z}}w + \frac{1}{\varepsilon^2}\sum_{i=1}^{2}(|w_i| - \eta_i + \varepsilon)^+, \\ &w_t + \mathfrak{L}_{\pm\bar{z}}w \leq \varepsilon^{-1}; \end{aligned} \tag{3.93}$$

on $[0, T) \times \{\pm\bar{z}\} \times (-\bar{\xi}_1, \bar{\xi}_1) \times (-\bar{\xi}_2, \bar{\xi}_2)$, where:

$$\mathfrak{L}_{\pm\bar{z}}\, w := \beta^\top(\alpha z - \Sigma\Sigma^\top\xi)w_z + \frac{1}{2}\beta^\top\Sigma\Sigma^\top\beta w_z^2 + \frac{1}{2}\xi^\top\Sigma\Sigma^\top\xi - \xi^\top\alpha z. \tag{3.94}$$

The boundary conditions at $\xi = \pm\bar{\xi}$ are obtained by assuming that excessive long (resp. short) positions are to be reduced by selling (resp. buying) more futures, which is justified if $\bar{\xi}$ is sufficiently large. recall that in (3.88), $\eta_i - w_i = 0$ (resp. $\eta_i + w_i = 0$) corresponds to $\gamma\eta_i v - v_i = 0$ (resp. $\gamma\eta_i v + v_i = 0$) in (3.28) which, in turn, corresponds to the sell region S_i (resp. buy region B_i). Therefore, one obtains the boundary conditions:

$$\begin{aligned} &w_1(t, z, \pm\bar{\xi}, \xi_2) = \pm\eta_1;\ (t, z, \xi_2) \in [0, T) \times [-\bar{z}, \bar{z}] \times [-\bar{\xi}_2, \bar{\xi}_2], \\ &w_2(t, z, \xi_1, \pm\bar{\xi}) = \pm\eta_2;\ (t, z, \xi_1) \in [0, T) \times [-\bar{z}, \bar{z}] \times (-\bar{\xi}_1, \bar{\xi}_1). \end{aligned} \tag{3.95}$$

Finally, the terminal condition is as before:

$$w(T, z, \xi_1, \xi_2) = \eta_1|\xi_1| + \eta_2|\xi_2|;\ (z, \xi_1, \xi_2) \in \overline{\mathcal{O}}. \tag{3.96}$$

3.4.2 The finite difference scheme

We intend to discretize (3.92)-(3.96) thorough a finite difference scheme. We need to introduce some notation first. Fix the integers $N, M, I, J \in \mathbb{N}$ and let $\Delta_t := T/N$, $\Delta_z := \bar{z}/M$, $\Delta_1 := \bar{\xi}_1/I$, and $\Delta_2 := \bar{\xi}_2/J$. Denote the set $\{n, n+1, \ldots, m\}$ by $\mathbb{Z}_n^m$, and consider the discrete times and states:

$$\begin{aligned} t_n &:= n\Delta_t, \quad z_m := m\Delta_z, \\ \xi_{1,i} &:= i\Delta_1, \quad \xi_{2,j} := j\Delta_2, \end{aligned} \qquad \text{for} \quad (n, m, i, j) \in \mathbb{Z}_0^N \times \mathbb{Z}_{-M}^M \times \mathbb{Z}_{-I}^I \times \mathbb{Z}_{-J}^J,$$

and the 3-dimensional regular grid:

$$\mathcal{G} := \Big\{(m\Delta_z, i\Delta_1, j\Delta_2) : m \in \mathbb{Z}_{-M}^M, i \in \mathbb{Z}_{-I}^I, j \in \mathbb{Z}_{-J}^J\Big\}.$$

Furthermore, let $w_{m,i,j}^n := w(t_n, z_m, \xi_{1,i}, \xi_{2,j})$, and $w^n := w(t_n, \mathcal{G})$ and, with a slight abuse of notation, use w as a general 3-dimensional array of size w^n, e.g. $f(w)$. This use of w should not be confused with the function w as in $\mathfrak{L}w$. Finally, define

$$\begin{aligned} \hat{a}_{m,i,j} &:= \beta^\top\big(\alpha z_m - \Sigma\Sigma^\top(\xi_{1,i}, \xi_{2,j})^\top\big), \quad \hat{b} := \tfrac{1}{2}\beta^\top\Sigma\Sigma^\top\beta, \\ \hat{c}_{m,i,j} &:= \tfrac{1}{2}(\xi_{1,i}, \xi_{2,j})\Sigma\Sigma^\top(\xi_{1,i}, \xi_{2,j})^\top - (\xi_{1,i}, \xi_{2,j})\alpha z_m. \end{aligned} \tag{3.97}$$

Then, the fully implicit discretization of (3.92) and (3.93) is given by:

$$\frac{w_{m,i,j}^{n+1} - w_{m,i,j}^n}{\Delta t} + \Big(\mathfrak{L}_{m,i,j} + \mathcal{P}_{m,i,j}\Big)(w^n) = 0; \tag{3.98}$$

for $(n, m, i, j) \in \mathbb{Z}_0^{N-1} \times \mathbb{Z}_{-M}^M \times \mathbb{Z}_{-I+1}^{I-1} \times \mathbb{Z}_{-J+1}^{J-1}$, where the functions $\mathfrak{L}_{m,i,j}$ and $\mathcal{P}_{m,i,j}$ are given in Appendix 3.A.

The boundary conditions (3.95) can only be discretized via backward or forward differencing, which yield

$$\begin{aligned} w_{m,\pm I,j}^n - w_{m,\pm I\mp 1,j}^n &= \Delta_1\eta_1, \\ w_{m,i,\pm J}^n - w_{m,i,\pm J\mp 1}^n &= \Delta_2\eta_2, \end{aligned} \tag{3.99}$$

for $(n, m, i, j) \in \mathbb{Z}_0^{N-1} \times \mathbb{Z}_{-M}^M \times \mathbb{Z}_{-I+1}^{I-1} \times \mathbb{Z}_{-J}^J$. Finally, the terminal condition (3.96) yields:

$$w_{m,i,j}^N = \eta_1\Delta_1|i| + \eta_2\Delta_2|j|, \quad (m, i, j) \in \mathbb{Z}_{-M}^M \times \mathbb{Z}_{-I+1}^{I-1} \times \mathbb{Z}_{-J+1}^{J-1}. \tag{3.100}$$

3.4.3 Nonlinearities and Newton's sub-iteration

The discretized equations (3.98)-(3.100) constitute an algebraic system which, in theory, can be solved recursively to yield the approximations w^n, $n = \{N-1, N-2, \ldots, 0\}$. On the other hand, this system of equations is non-linear because of the quadratic terms in $\mathfrak{L}_{m,i,j}(\cdot)$ as well as the non-smooth penalty function $\mathcal{P}_{m,i,j}(\cdot)$. There are two main approaches for solving the non-linear algebraic systems that originate from HJB equations. The classical approach is to use **relaxation schemes** based on Markov chain approximations, c.f. Kushner and Dupuis (2001). However, these methods are prone to time step limitations due to stability considerations, c.f. (Forsyth and Labahn, 2007, section 6.1). A second approach, which has been successfully implemented in similar settings as ours by Forsyth and Vetzal (2002) and Dai and Zhong (2010), is to treat the nonlinearities in (3.98) through a **Newton's iteration** scheme which takes into account the non-smoothness of the penalty terms $\mathcal{P}_{m,i,j}(\cdot)$. We opted to use the Newton iteration method.

Let $w^{n,(k)}_{m,i,j}$ be the approximation of $w^n_{m,i,j}$ obtained from the k-th Newton's iteration, and set $\Delta w^{n,(k+1)}_{m,i,j} := w^{n,(k+1)}_{m,i,j} - w^{n,(k)}_{m,i,j}$. Applying Newton's method to (3.98) and (3.99) yields the iteration:

$$\begin{aligned}\Delta w^{n,(k+1)}_{m,i,j} &- \Delta t \sum_{m',i',j'} \frac{\partial \mathfrak{L}_{m,i,j}}{\partial w_{m',i',j'}}\big(w^{n,(k)}\big)\Delta w^{n,(k+1)}_{m',i',j'} \\ &- \Delta t \sum_{m',i',j'} \frac{\partial \mathcal{P}_{m,i,j}}{\partial w_{m',i',j'}}\big(w^{n,(k)}\big)\Delta w^{n,(k+1)}_{m',i',j'} \\ = w^{n+1}_{m,i,j} &- w^{n,(k)}_{m,i,j} + \Delta t\Big(\mathfrak{L}_{m,i,j} + \mathcal{P}_{m,i,j}\Big)\big(w^{n,(k)}\big);\end{aligned} \tag{3.101}$$

for $(n,m,i,j) \in \mathbb{Z}_0^{N-1} \times \mathbb{Z}_{-M}^{M} \times \mathbb{Z}_{-I+1}^{I-1} \times \mathbb{Z}_{-J+1}^{J-1}$, and

$$\Delta w^{n,(k+1)}_{m,\pm I,j} - \Delta w^{n,(k+1)}_{m,\pm I\mp 1,j} = \Delta_1 \eta_1 - w^{n,(k)}_{m,\pm I,j} + w^{n,(k)}_{m,\pm I\mp 1,j}, \tag{3.102}$$

$$\Delta w^{n,(k+1)}_{m,i,\pm J} - \Delta w^{n,(k+1)}_{m,i,\pm J\mp 1} = \Delta_2 \eta_2 - w^{n,(k)}_{m,i,\pm J} + w^{n,(k)}_{m,i,\pm J\mp 1}; \tag{3.103}$$

for $(n,m,i,j) \in \mathbb{Z}_0^{N-1} \times \mathbb{Z}_{-M}^{M} \times \mathbb{Z}_{-I+1}^{I-1} \times \mathbb{Z}_{-J}^{J}$. Note that only a few terms on the left side of (3.101) are nonzero. Indeed, one has:

$$\begin{aligned}
-\Delta t \sum_{m',i',j'} \frac{\partial \mathfrak{L}_{m,i,j}}{\partial w_{m',i',j'}}\big(w^{n,(k)}\big)\Delta w^{n,(k+1)}_{m',i',j'} &= \sum_{\tau=-1}^{1} A_{m,i,j,\tau}(w^{n,(k)})\Delta w^{n,(k+1)}_{m+\tau,i,j} \\
-\Delta t \sum_{m',i',j'} \frac{\partial \mathcal{P}_{m,i,j}}{\partial w_{m',i',j'}}\big(w^{n,(k)}\big)\Delta w^{n,(k+1)}_{m',i',j'} &= \sum_{\tau=-1}^{1} P^{1,\text{P1-dis.}}_{m,i,j,\tau}(w^{n,(k)})\Delta w^{n,(k+1)}_{m,i+\tau,j} \\
&\quad + \sum_{\tau=-1}^{1} P^{2,\text{P2-dis.}}_{m,i,j,\tau}(w^{n,(k)})\Delta w^{n,(k+1)}_{m,i,j+\tau}\,.
\end{aligned} \tag{3.104}$$

The coefficients $A_{m,i,j,\tau}(\cdot)$, $P^{1,\text{P1-dis.}}_{m,i,j,\tau}(\cdot)$, and $P^{2,\text{P2-dis.}}_{m,i,j,\tau}(\cdot)$ are given in Appendix 3.B. In particular, $A_{m,i,j,\tau}(\cdot)$ is given by (3.111) and (3.112), while the discretization choices of the penalty term, represented by P1-dis., and P2-dis. are determined by the following rule:

The upstream scheme for the penalty terms:
Assume $w^{n,(k)}$ to be known. For iteration $k+1$, P1-dis. and P2-dis. in (3.104) at node $(n,m,i,j) \in \mathbb{Z}_0^{N-1} \times \mathbb{Z}_{-M}^{M} \times \mathbb{Z}_{-I+1}^{I-1} \times \mathbb{Z}_{-J+1}^{J-1}$ are determined as follows:

(i) If $w^{n,(k)}_{m,i+1,j} - w^{n,(k)}_{m,i,j} \le \Delta_1(\eta_1 - \varepsilon)$ (resp. $w^{n,(k)}_{m,i,j+1} - w^{n,(k)}_{m,i,j} \le \Delta_2(\eta_2 - \varepsilon)$), then P1-dis. = F (resp. P2-dis. = F). The coefficients $P^{1,\text{F}}_{m,i,j,\tau}(\cdot)$ and $P^{2,\text{F}}_{m,i,j,\tau}(\cdot)$ are given by (3.113).

(ii) If $w^{n,(k)}_{m,i+1,j} - w^{n,(k)}_{m,i,j} > \Delta_1(\eta_1 - \varepsilon)$ and $w^{n,(k)}_{m,i,j} - w^{n,(k)}_{m,i-1,j} \ge -\Delta_1(\eta_1 - \varepsilon)$ (resp. $w^{n,(k)}_{m,i,j+1} - w^{n,(k)}_{m,i,j} > \Delta_2(\eta_2 - \varepsilon)$ and $w^{n,(k)}_{m,i,j} - w^{n,(k)}_{m,i,j-1} \ge -\Delta_2(\eta_2 - \varepsilon)$), then P1-dis. = B (resp. P2-dis. = B). The coefficients $P^{1,\text{B}}_{m,i,j,\tau}(\cdot)$ and $P^{2,\text{B}}_{m,i,j,\tau}(\cdot)$ are given by (3.114).

Figure 3.1 illustrates the numerical scheme for approximating the solution of (3.92)–(3.96). The constant $0 < \text{tol} \ll 1$ in the algorithm determines whether Newton's iteration has converged.

Set w^N according to (3.100);
for $n \leftarrow N-1, N-2, \ldots, 0$ **do**
  $w^{n,(0)} \leftarrow w^{n+1}$;
  $k \leftarrow 0$;
  Set up (3.101)–(3.103) based on the upstream rule, and solve the system to obtain $\Delta w^{n,(1)}$;
  $w^{n,(1)} \leftarrow w^{n,(0)} + \Delta w^{n,(1)}$;
  while $\max_{m,i,j}\left\{\frac{\left|\Delta w^{n,(k+1)}_{m,i,j}\right|}{\max\left(1,\left|w^{n,(k+1)}_{m,i,j}\right|\right)}\right\} \geq \mathrm{tol}$ **do** `// Newton iteration`
    $k \leftarrow k+1$;
    Set up (3.101)–(3.103) based on the upstream rule, and solve the system to obtain $\Delta w^{n,(k+1)}$;
    $w^{n,(k+1)} \leftarrow w^{n,(k)} + \Delta w^{n,(k+1)}$;
  end
  $w^n \leftarrow w^{n,(k+1)}$
end

Figure 3.1: Finite difference scheme with Newton's iteration

3.5 Numerical Results

We have implemented the algorithm of Figure 3.1 in Python. The code is included in Appendix 3.C.

Table 3.1 shows the values of the market parameters that we use in this section. Note that it is customary to normalize the prices such that the standard deviation of daily price change equals 0.01. This is roughly equivalent to standard deviation of $1 (= 0.01\sqrt{250 \times 40})$ per 40 years. Hence, we assumed that T is expressed in units of 40 years. Also, note that by Proposition 3.1, $\sigma_z^2 = 1$ and Z_t is a standard Gaussian random variable.

Table 3.1: Values of the market parameters

c	ρ	α_1	α_2	T	$\eta_1 = \eta_2$
1.0	0.5	-1.0	1.0	0.1*	0.05

* The units of T is 40 yrs.

Table 3.2 shows the values of the parameters that we use for the numerical scheme. We will shortly elaborate on the choice of the penalization parameter ε.

Table 3.2: Values of the discretization parameters

$\bar{z}$	$\bar{\xi}_1 = \bar{\xi}_2$	N	M	I	J	ε	tol
7.0	20.0	12	7	40	40	10^{-6}	10^{-4}

Figure 3.2 shows the estimated trading policy at $t = t_{N-1}$. As mentioned in Section 3.3 (see the discussion after Theorem 3.8), the trading policy is as follows. At each time t, the 3-dimensional space (z, θ^1, θ^2) is divided into nine regions: a *no-trade region* and eight *trading regions*. Each graph in Figure 3.2 is a cross-section for a fixed value of z, i.e. a plane parallel to the $\theta^1\theta^2$-plane. The levels of z correspond to multiples of standard deviation of Z_t (recall that Z_t is a standard Gaussian random variable). In each graph, we differentiate between each region by using different colors, as follows:

- **Blue:** represents the buy region. Dark blue means buying both futures, while light blue means buying only one asset.
- **Red:** represents the sell region. Dark red means selling both futures, while light red means selling only one asset.
- **Green:** represents the buy-sell regions, i.e. when, simultaneously, one asset is bought while the other is sold.
- **White:** represents the no-trade region.
- **Black:** represents the points in the space where the approximated solution does not satisfy the variational inequality.

Figure 3.2 suggests that the size of the no-trade region increases as z increases in size, i.e. as the prices diverge.

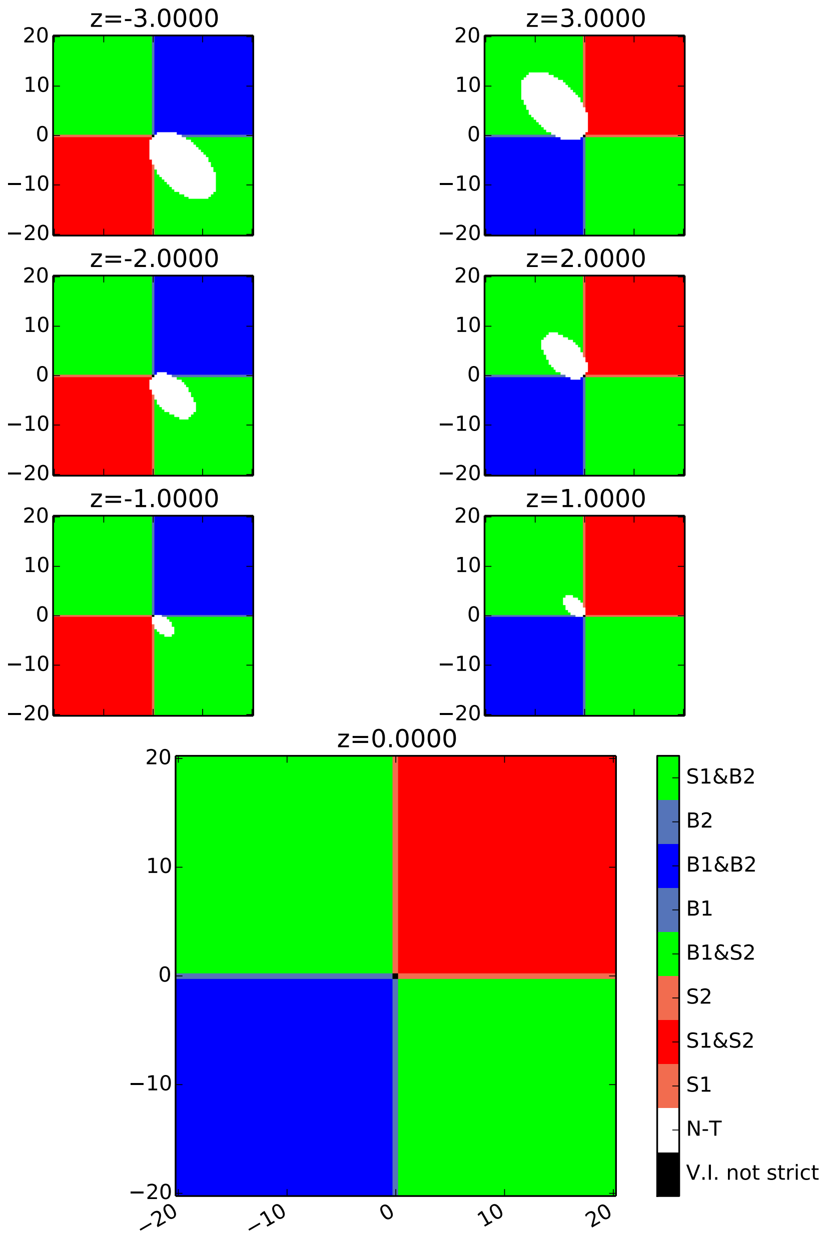


Figure 3.2: The approximated trading regions at $t = T_{N-1}$. Each graph represents different level of z. The horizontal (resp. vertical) axis corresponds to $\xi_1 := \gamma\theta_1$ (resp. $\xi_2 := \gamma\theta_2$).

Figure 3.3 shows the estimated trading policy at $t = 0$, which suggests the same pattern as in Figure 3.2, i.e. the size of the no-trade region increases as the prices diverge. Also, comparing the no-trade regions between Figures 3.2 and 3.3

suggests that the size of no-trade region increases as one approaches the terminal time.

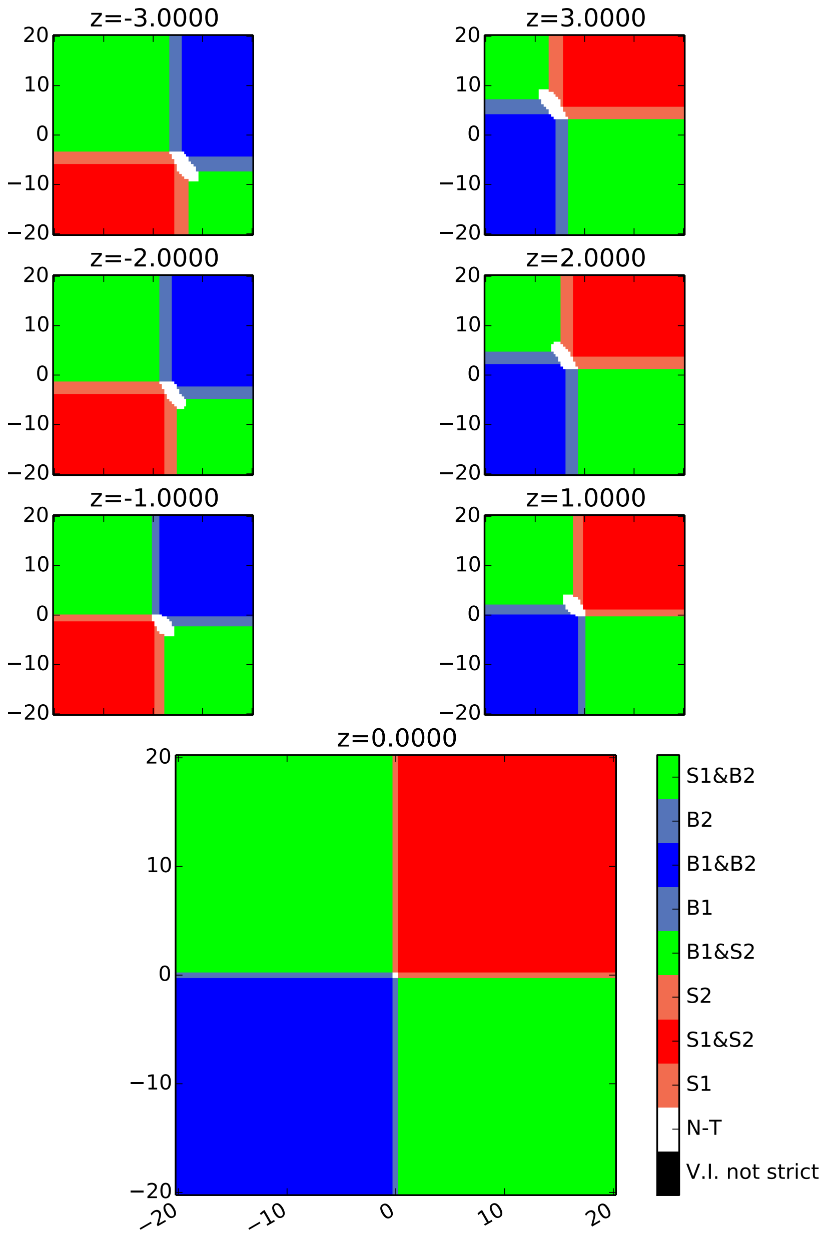


Figure 3.3: The approximated trading regions at $t = 0$. Each graph represents different level of z. The horizontal (resp. vertical) axis corresponds to $\xi_1 := \gamma\theta_1$ (resp. $\xi_2 := \gamma\theta_2$).

Figure 3.4 shows a specific cross-section (i.e. $z = 3.0$) at different times. It illustrates more clearly the change in the no-trade region as the terminal time is approached.

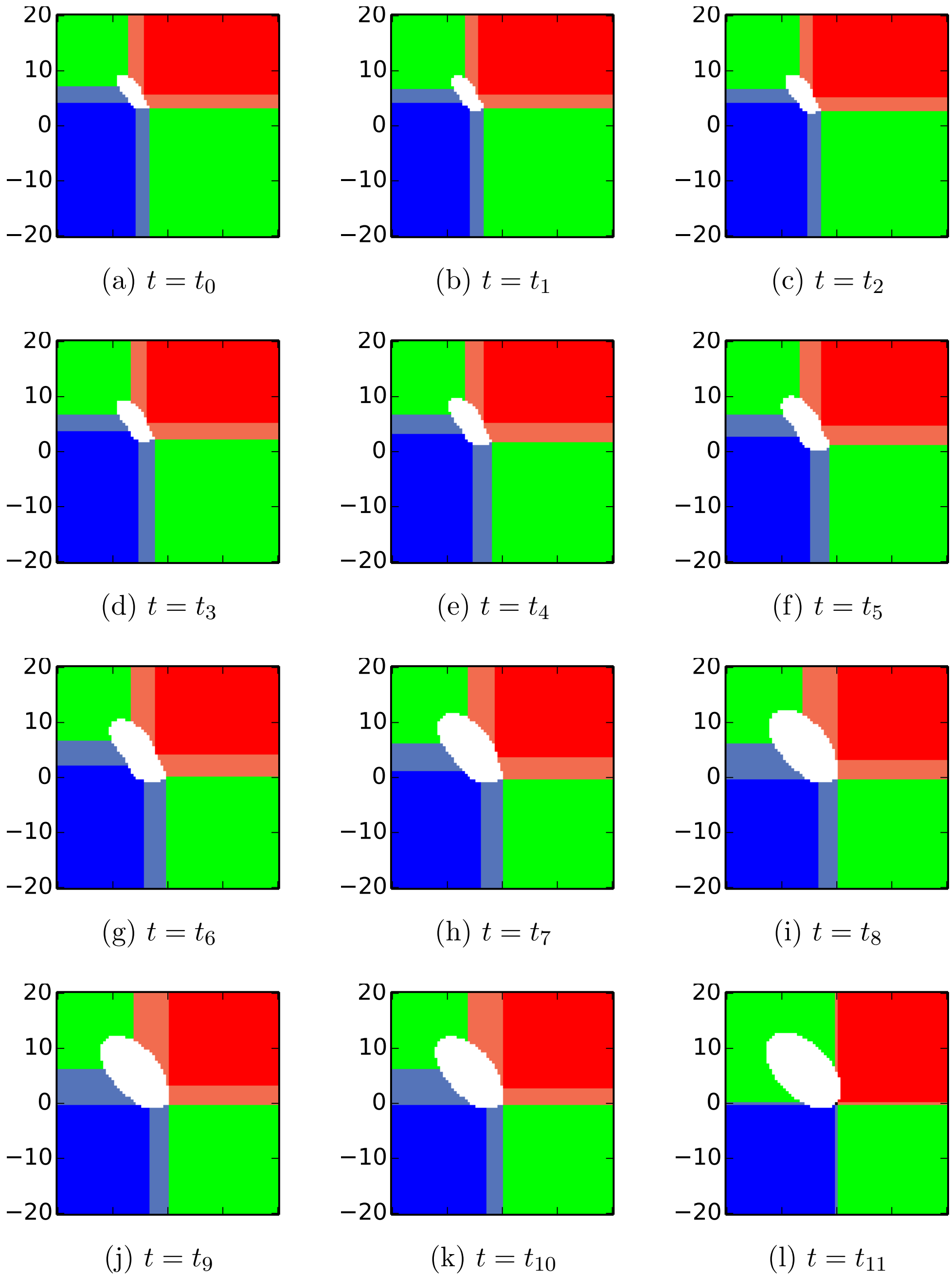
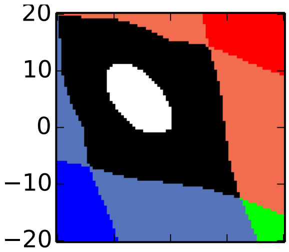
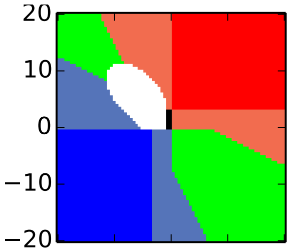


Figure 3.4: Term structure of trading regions at the cross-section $z = 3.0$.

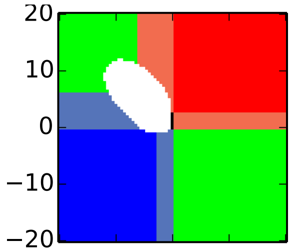
Finally, Figure 3.5 shows how the approximate solution changes as the penalization parameter ε of (3.91) decreases. The figure suggests that for $\varepsilon \leq 10^{-6}$, the solution is not affected by the change in ε.



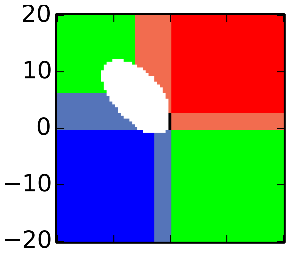
(a) $\varepsilon = 10^{-2}$



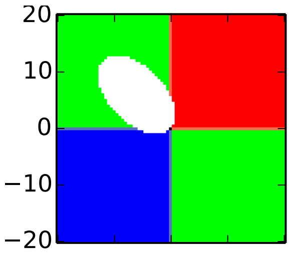
(b) $\varepsilon = 10^{-3}$



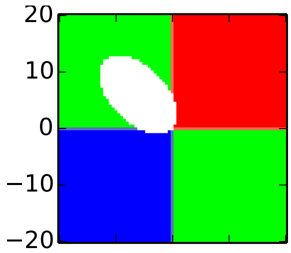
(c) $\varepsilon = 10^{-4}$



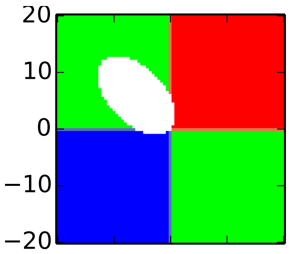
(d) $\varepsilon = 10^{-5}$



(e) $\varepsilon = 10^{-6}$



(f) $\varepsilon = 10^{-7}$



(g) $\varepsilon = 10^{-8}$

Figure 3.5: The effect of different choices of the penalization parameter ε on the cross-section $z = 3.0$.

3.6 Conclusion

In this chapter, we considered a Merton investment problem with proportional transaction cost, two risky assets, stochastic opportunity set, and finite horizon. We motivated this model by relating it to spread-trading with two futures assets and assuming linear transaction costs. We then introduced the HJB equation and provided rigorous arguments showing that the value function corresponding to the Merton problem is the viscosity solution of the HJB equation. We then proceeded by devising a numerical scheme, based on the penalty method of Forsyth and Vetzal (2002), to approximate the viscosity solution of the HJB equation. We concluded by a numerical example.

We point out the following two topics as future research topics which are very much unfinished parts of this chapter. Firstly, we must provide convergence results for the numerical scheme of Figure 3.1. Secondly, we must provide numerical example showing the efficiency and practicality of the approximation procedure.

3.A Differencing schemes

There are three differencing schemes to choose from: central, forward and backward, c.f. Fornberg (1988). The most accurate is the central differencing:

$$f''(x_0) = \frac{f(x_{-1}) + f(x_1) - 2f(x_0)}{\Delta_x^2} + O(\Delta_x^2),$$
$$f'(x_0) = \frac{f(x_1) - f(x_{-1})}{2\Delta_x} + O(\Delta_x^2),$$

Forward and backward differencing are less accurate:

$$f''(x_0) = \frac{f(x_{-2}) + f(x_0) - 2f(x_{-1})}{\Delta_x^2} + O(\Delta_x) = \frac{f(x_2) + f(x_0) - 2f(x_1)}{\Delta_x^2} + O(\Delta_x)$$
$$f'(x_0) = \frac{f(x_0) - f(x_{-1})}{\Delta_x} + O(\Delta_x) = \frac{f(x_1) - f(x_0)}{\Delta_x} + O(\Delta_x).$$

First consider discretizing $\mathfrak{L}w$ of (3.87) and $\mathfrak{L}_{\pm\bar{z}}w$ of (3.94), which by using (3.97) become

$$\begin{aligned} \mathfrak{L}w &= \hat{a}_{m,i,j}w_z + \hat{b}w_z^2 + \hat{b}w_{zz} + \hat{c}_{m,i,j}, \\ \mathfrak{L}_{\pm\bar{z}}w &= \hat{a}_{m,i,j}w_z + \hat{b}w_z^2 + \hat{c}_{m,i,j}. \end{aligned} \tag{3.105}$$

For discretization of $\mathfrak{L}_{\pm\bar{z}}w$ on the boundary $z=\bar{z}$ (resp. $z=-\bar{z}$), only backward (resp. forward) discretization is possible. The forward scheme is:

$$\begin{aligned}\mathfrak{L}_{-M,i,j}(w) := &-\tfrac{\hat{a}_{-M,i,j}}{\Delta_z}w_{-M,i,j}+\tfrac{\hat{a}_{-M,i,j}}{\Delta_z}w_{-M+1,i,j}+\tfrac{\hat{b}}{\Delta_z^2}(w_{-M+1,i,j})^2\\ &+\tfrac{\hat{b}}{\Delta_z^2}(w_{-M,i,j})^2-\tfrac{2\hat{b}}{\Delta_z^2}w_{-M+1,i,j}w_{-M,i,j}+\hat{c}_{-M,i,j},\end{aligned} \tag{3.106}$$

and the backward counterpart is:

$$\begin{aligned}\mathfrak{L}_{M,i,j}(w) = &-\tfrac{\hat{a}_{M,i,j}}{\Delta_z}w_{M-1,i,j}+\tfrac{\hat{a}_{M,i,j}}{\Delta_z}w_{M,i,j}\\ &+\tfrac{\hat{b}}{\Delta_z^2}(w_{M,i,j})^2+\tfrac{\hat{b}}{\Delta_z^2}(w_{M-1,i,j})^2-\tfrac{2\hat{b}}{\Delta_z^2}w_{M,i,j}w_{M-1,i,j}+\hat{c}_{M,i,j}.\end{aligned} \tag{3.107}$$

To discretize $\mathfrak{L}w$, we use central differencing for all the derivatives, which leads to:

$$\begin{aligned}\mathfrak{L}_{m,i,j}(w) = &\Big(\tfrac{\hat{b}}{\Delta_z^2}-\tfrac{\hat{a}_{m,i,j}}{2\Delta_z}\Big)w_{m-1,i,j}-\tfrac{2\hat{b}}{\Delta_z^2}w_{m,i,j}+\Big(\tfrac{\hat{b}}{\Delta_z^2}+\tfrac{\hat{a}_{m,i,j}}{2\Delta_z}\Big)w_{m+1,i,j}\\ &+\tfrac{\hat{b}}{4\Delta_z^2}(w_{m+1,i,j})^2+\tfrac{\hat{b}}{4\Delta_z^2}(w_{m-1,i,j})^2-\tfrac{\hat{b}}{2\Delta_z^2}w_{m+1,i,j}w_{m-1,i,j}+\hat{c}_{m,i,j}.\end{aligned} \tag{3.108}$$

Finally, to discretize the penalty term $-\frac{1}{\varepsilon^2}\sum_{i=1}^{2}(|w_i|-\eta_i+\varepsilon)^+$, we only consider forward and backward differencing. Forward discretization of the penalty term is

$$\begin{aligned}\mathcal{P}^{\mathrm{F}}_{m,i,j}(w) := &-\tfrac{1}{\Delta_1\varepsilon^2}\Big(\big|w_{m,i+1,j}-w_{m,i,j}\big|-\Delta_1(\eta_1-\varepsilon)\Big)^+\\ &-\tfrac{1}{\Delta_2\varepsilon^2}\Big(\big|w_{m,i,j+1}-w_{m,i,j}\big|-\Delta_2(\eta_2-\varepsilon)\Big)^+,\end{aligned} \tag{3.109}$$

while the backward counterpart is:

$$\begin{aligned}\mathcal{P}^{\mathrm{B}}_{m,i,j}(w) := &-\tfrac{1}{\Delta_1\varepsilon^2}\Big(\big|w_{m,i,j}-w_{m,i-1,j}\big|-\Delta_1(\eta_1-\varepsilon)\Big)^+\\ &-\tfrac{1}{\Delta_2\varepsilon^2}\Big(\big|w_{m,i,j}-w_{m,i,j-1}\big|-\Delta_2(\eta_2-\varepsilon)\Big)^+.\end{aligned} \tag{3.110}$$

3.B Coefficients of Newton's iteration

The coefficients $A^{\text{L-disc}}_{m,i,j,\tau}(\cdot)$, $P^{1,\text{P1-disc}}_{m,i,j,\tau}(\cdot)$, and $P^{2,\text{P2-disc}}_{m,i,j,\tau}(\cdot)$ in (3.104), depending on the discretization used, are as follows. At the boundaries $z=\pm\bar{z}$, when the forward scheme (3.106) or the backward scheme (3.107) is used, the coefficients of the Newton's iteration are:

$$\begin{aligned}A_{-M,i,j,0}(w) &= \tfrac{\Delta t}{\Delta_z^2}\Big(\hat{a}_{-M,i,j}\Delta_z+2\hat{b}\big(w_{-M+1,i,j}-w_{-M,i,j}\big)\Big),\\ A_{-M,i,j,1}(w) &= -\tfrac{\Delta t}{\Delta_z^2}\Big(\hat{a}_{-M,i,j}\Delta_z+2\hat{b}\big(w_{-M+1,i,j}-w_{-M,i,j}\big)\Big),\\ A_{M,i,j,0}(w) &= \tfrac{\Delta t}{\Delta_z^2}\Big(-\hat{a}_{M,i,j}\Delta_z-2\hat{b}\big(w_{M,i,j}-w_{M-1,i,j}\big)\Big),\\ A_{M,i,j,-1}(w) &= -\tfrac{\Delta t}{\Delta_z^2}\Big(-\hat{a}_{M,i,j}\Delta_z-2\hat{b}\big(w_{M,i,j}-w_{M-1,i,j}\big)\Big).\end{aligned} \tag{3.111}$$

For the central differencing scheme (3.108), the coefficients of the Newton's iteration are:

$$\begin{aligned}
A_{m,i,j,-1}(w) &= -\tfrac{\Delta t}{2\Delta_z^2}\Big(-\hat{a}_{m,i,j}\Delta_z - \hat{b}\big(w_{m+1,i,j} - w_{m-1,i,j} - 2\big)\Big),\\
A_{m,i,j,0}(w) &= \tfrac{2\Delta t\hat{b}}{\Delta_z^2},\\
A_{m,i,j,1}(w) &= -\tfrac{\Delta t}{2\Delta_z^2}\Big(\hat{a}_{m,i,j}\Delta_z + \hat{b}\big(w_{m+1,i,j} - w_{m-1,i,j} + 2\big)\Big).
\end{aligned} \tag{3.112}$$

The Newton's iteration coefficients for the forward scheme (3.109) are

$$\begin{aligned}
P^{1,\mathrm{F}}_{m,i,j,1}(w) = -P^{1,\mathrm{F}}_{m,i,j,0}(w) &= \begin{cases} \frac{\Delta t}{\Delta_1 \varepsilon^2}; & w_{m,i+1,j} - w_{m,i,j} > \Delta_1(\eta_1 - \varepsilon),\\ \frac{-\Delta t}{\Delta_1 \varepsilon^2}; & w_{m,i+1,j} - w_{m,i,j} < -\Delta_1(\eta_1 - \varepsilon),\\ 0; & \text{otherwise}, \end{cases}\\
P^{2,\mathrm{F}}_{m,i,j,1}(w) = -P^{2,\mathrm{F}}_{m,i,j,0}(w) &= \begin{cases} \frac{\Delta t}{\Delta_2 \varepsilon^2}; & w_{m,i,j+1} - w_{m,i,j} > \Delta_2(\eta_2 - \varepsilon),\\ \frac{-\Delta t}{\Delta_2 \varepsilon^2}; & w_{m,i,j+1} - w_{m,i,j} < -\Delta_2(\eta_2 - \varepsilon),\\ 0; & \text{otherwise}, \end{cases}\\
P^{1,\mathrm{F}}_{m,i,j,-1}(w) = P^{2,\mathrm{F}}_{m,i,j,-1}(w) &= 0.
\end{aligned} \tag{3.113}$$

Finally, the Newton's iteration coefficients for the backward differencing of the penalty term, i.e. (3.110), are:

$$\begin{aligned}
P^{1,\mathrm{B}}_{m,i,j,0}(w) = -P^{1,\mathrm{B}}_{m,i,j,-1}(w) &= \begin{cases} \frac{\Delta t}{\Delta_1 \varepsilon^2}; & w_{m,i,j} - w_{m,i,j-1} > \Delta_1(\eta_1 - \varepsilon),\\ \frac{-\Delta t}{\Delta_1 \varepsilon^2}; & w_{m,i,j} - w_{m,i,j-1} < -\Delta_1(\eta_1 - \varepsilon),\\ 0; & \text{otherwise}, \end{cases}\\
P^{2,\mathrm{B}}_{m,i,j,0}(w) = -P^{2,\mathrm{B}}_{m,i,j,-1}(w) &= \begin{cases} \frac{\Delta t}{\Delta_2 \varepsilon^2}; & w_{m,i,j} - w_{m,i-1,j} > \Delta_2(\eta_2 - \varepsilon),\\ \frac{-\Delta t}{\Delta_2 \varepsilon^2}; & w_{m,i,j} - w_{m,i,j-1} < -\Delta_2(\eta_2 - \varepsilon),\\ 0; & \text{otherwise}, \end{cases}\\
P^{1,\mathrm{B}}_{m,i,j,1}(w) = P^{2,\mathrm{B}}_{m,i,j,1}(w) &= 0.
\end{aligned} \tag{3.114}$$

3.C Python Code

3.C.1 An example of using the functions

```
import numpy as np
import matplotlib
matplotlib.use("pdf",force=True)
```

```python
import matplotlib.pyplot as plt
matplotlib.get_backend()
plt.ioff()
import MyLib #IMPORTANT: import MyLib after choosing the backend for
    matplotlib

#===============================================================================
#                              Parameters
#===============================================================================
MarketParams = {'alpha': (-1.0,1.0),
               'c': 1.0,
               'rho': 0.5,
               'T': 1.0/44.4 * 4,# units in 44.4 yrs: 0.15 yearly
                  volatility = vol of 1 in 44.4 yrs
               'eta': (0.05,0.05)}
kappa = MarketParams['c'] * MarketParams['alpha'][1] -
    MarketParams['alpha'][0]
sigma_z2 = 1.0 + MarketParams['c']**2 - 2 * MarketParams['c'] *
    MarketParams['rho']
Var_z = kappa / 2.0 / sigma_z2
MeshParams = { 'zBar': 7, # 7 * Var_z
               'xiBar': (20.0,20.0),
               'N': 12,
               'M': 7,
               'I': 40,
               'J': 40}
tol = 1e-4
epsilon = 1e-6
Pol_error = 1e-3
RemoveCorners = np.nan # If not np.nan, the corners will be fixed to
    their terminal value + RemoveCorners
numprocesses=8
OutPutPath = '/Users/bahman/Archive/DPhil/work/TC/Figs'
Z_levels = 3 # Plots at z levels -Z_levels,...,Z_levels
Symmetries = True # if true, symmetries the no-trade plots around the
    origin
#===============================================================================
#===============================================================================
# Uncomment in windows
# if __name__ == '__main__':
#     freeze_support()
res = MyLib.TC(MarketParams, MeshParams, OutPutPath, tol=tol,
```

```
            epsilon=epsilon,
               Pol_error=Pol_error,RemoveCorners=RemoveCorners,
            numprocesses=numprocesses,
            Z_levels=Z_levels, Symmetries=Symmetries)
```

3.C.2 The functions library

```
import pandas as pd
import numpy as np
from scipy.sparse.linalg import spsolve
from scipy.sparse import csr_matrix
import sys
import matplotlib
import matplotlib.pyplot as plt
from matplotlib.backends.backend_pdf import PdfPages
import time
from multiprocessing import Pool

def _3ixToFlat(mij,Dims):
    # Auxiliary function for _AlgebSys
    M = Dims[0]
    I = Dims[1]
    return mij[0] + mij[1] * (2*M+1) + mij[2] * (2*M+1) * (2*I+1)

def _FlatTo3ix(ix,Dims):
    # Auxiliary function for _AlgebSys
    M = Dims[0]
    I = Dims[1]
    j = ix / ((2*M+1) * (2*I+1))
    i = (ix % ((2*M+1) * (2*I+1)) ) / (2*M+1)
    m = (ix % ((2*M+1) * (2*I+1)) ) % (2*M+1)
    return (m,i,j)

def _AlgebSys_Aux(args):
    # Auxiliary function for multiprocessing in _AlgebSys
    #inputs
    I = args['I']
    J = args['J']
    Dims = args['Dims']
    eta1 = args['eta1']
    eta2 = args['eta2']
```

```
Delta_t = args['Delta_t']
Delta_z = args['Delta_z']
Delta_1 = args['Delta_1']
Delta_2 = args['Delta_2']
ahat = args['ahat']
bhat = args['bhat']
chat = args['chat']
epsilon = args['epsilon']
w_old = args['w_old']
w_lastItr = args['w_lastItr']
m = args['m']
#Initializations
res={}
max_entries = 5 * (2*I-1) * (2*J-1)
row = np.array([np.nan for it in range(max_entries)]) #
    @UnusedVariable
col= np.array([np.nan for it in range(max_entries)]) #
    @UnusedVariable
data= np.array([np.nan for it in range(max_entries)]) #
    @UnusedVariable
RHS = np.zeros(shape=((2*I-1) * (2*J-1)),dtype=np.float64)
RHS_IXs = np.zeros(shape=((2*I-1) * (2*J-1)),dtype=np.int64)
entry_ix = 0
RHS_ix=0
for i in range(1,2*I):
    for j in range(1,2*J):
        # Calculating A and L operator
        A0 = 2.0 * Delta_t * bhat / Delta_z**2
        A1 = -A0 / 4.0 * (- ahat.ix[m,i,j] * Delta_z
                         - bhat *
                         (w_old.ix[m+1,i,j] - w_old.ix[m-1,i,j] - 2.0))
        Am1 = -A0 / 4.0 * (ahat.ix[m,i,j] * Delta_z
                         + bhat *
                         (w_old.ix[m+1,i,j] - w_old.ix[m-1,i,j] + 2.0))
        L_op = ( w_old.ix[m-1,i,j] * ( bhat / (Delta_z**2) -
            ahat.ix[m,i,j] / 2.0 / Delta_z)
                - w_old.ix[m,i,j] * 2.0 * bhat / (Delta_z**2)
                + w_old.ix[m+1,i,j] * ( bhat / (Delta_z**2) +
                    ahat.ix[m,i,j] / 2.0 / Delta_z)
                + bhat / 4.0 / (Delta_z**2)
                *(w_old.ix[m+1,i,j]**2 + w_old.ix[m-1,i,j]**2
                  - 2.0 * w_old.ix[m+1,i,j] * w_old.ix[m-1,i,j])
                + chat.ix[m,i,j])
```

```
# Calculating penalty terms
if((w_old.ix[m,i+1,j] - w_old.ix[m,i,j])
    < - Delta_1 * (eta1 - epsilon)):
    # Forward scheme
    P1 = Delta_t / Delta_1 / (epsilon**2)
    P1_IX = (m,i+1,j)
    P_op = - 1.0 / Delta_1 / (epsilon**2) *\
        ( - (w_old.ix[m,i+1,j] - w_old.ix[m,i,j])
          - Delta_1 * (eta1 - epsilon))
elif((w_old.ix[m,i,j] - w_old.ix[m,i-1,j])
    > Delta_1 * (eta1 - epsilon)):
    # Backward scheme
    P1 = Delta_t / Delta_1 / (epsilon**2)
    P1_IX = (m,i-1,j)
    P_op = - 1.0 / Delta_1 / (epsilon**2) *\
        ( w_old.ix[m,i,j] - w_old.ix[m,i-1,j]
          - Delta_1 * (eta1 - epsilon))
else:
    P1 = 0.0
    P1_IX = ()
    P_op = 0.0
if((w_old.ix[m,i,j+1] - w_old.ix[m,i,j])
    < - Delta_2 * (eta2 - epsilon)):
    # Forward scheme
    P2 = Delta_t / Delta_2 / (epsilon**2)
    P2_IX = (m,i,j+1)
    P_op += - 1.0 / Delta_2 / (epsilon**2) *\
        ( - (w_old.ix[m,i,j+1] - w_old.ix[m,i,j])
          - Delta_2 * (eta2 - epsilon))
elif((w_old.ix[m,i,j] - w_old.ix[m,i,j-1])
    > Delta_2 * (eta2 - epsilon)):
    # Backward scheme
    P2 = Delta_t / Delta_2 / (epsilon**2)
    P2_IX = (m,i,j-1)
    P_op += - 1.0 / Delta_2 / (epsilon**2) *\
        ( w_old.ix[m,i,j] - w_old.ix[m,i,j-1]
          - Delta_2 * (eta2 - epsilon))
else:
    P2 = 0.0
    P2_IX = ()
    P_op += 0.0
# Filling the Matrix
row[entry_ix] = _3ixToFlat((m,i,j),Dims)
```

```
            row[entry_ix+1] = _3ixToFlat((m,i,j),Dims)
            row[entry_ix+2] = _3ixToFlat((m,i,j),Dims)
            col[entry_ix]  = _3ixToFlat((m,i,j),Dims)
            col[entry_ix+1] = _3ixToFlat((m-1,i,j),Dims)
            col[entry_ix+2] = _3ixToFlat((m+1,i,j),Dims)
            data[entry_ix] = 1.0 + A0 + P1 + P2
            data[entry_ix+1]= Am1
            data[entry_ix+2]= A1
            entry_ix += 3
            if (P1 != 0.0):
                row[entry_ix] = _3ixToFlat((m,i,j),Dims)
                col[entry_ix] = _3ixToFlat(P1_IX,Dims)
                data[entry_ix] = -P1
                entry_ix += 1
            if (P2 != 0.0):
                row[entry_ix] = _3ixToFlat((m,i,j),Dims)
                col[entry_ix] = _3ixToFlat(P2_IX,Dims)
                data[entry_ix] = -P2
                entry_ix += 1
            RHS[RHS_ix] =\
                (w_lastItr.ix[m,i,j] - w_old.ix[m,i,j]
                 + Delta_t * (L_op + P_op))
            RHS_IXs[RHS_ix] = _3ixToFlat((m,i,j),Dims)
            RHS_ix +=1
    if (entry_ix < max_entries):
        row = row[:entry_ix]
        col = col[:entry_ix]
        data = data[:entry_ix]
    else:
        sys.stdout.write("\nRare event: number of max entries reached
            (nothing wrong, ignore the message)\n")
    res['row'] = row
    res['col'] = col
    res['data'] = data
    res['RHS'] = RHS
    res['RHS_IXs'] = RHS_IXs
    return res

def _AlgebSys(MarketParams,MeshParams,w_old,w_lastItr,
              ahat,bhat,chat,epsilon,RemoveCorners=np.nan,
              numprocesses=8):
    #===========================================================================
    # Sets up the algebraic system for Newton's sub iteration in TC.
```

```
# Inputs:
#   MarketParams: Dictionary with the following keys:
#       'alpha': pair of floats (alpha1,alpha2), factor loadings.
#       'c': float, cointegration coefficient
#       'rho': float, correlation
#       'T': float, investment horizon
#       'eta': pair of floats, proportional transaction costs in
    each market
#   MeshParams: Dictionary with the following keys:
#       'zBar': positive float, z is assumed to take values in
    (-zBar, zBar)
#       'xiBar': pair of floats (xiBar_1, xiBar_2). xi_i takes value
    in (-xiBar_i, xiBar_i)
#       'N': integer, number of time steps
#       'M': integer, number of z steps
#       'I': integer, number of xi_1 steps
#       'J': integer, number of xi_2 steps
#   w_old: the output of the previous Newton's iteration
#   w_lastItr: the solution for the previous time
#   ahat, bhat, chat: coefficients, ahat and chat are the same
    dimension as
#       w_old and w_lastItr, while bhat is a scalar.
#   epsilon: float, small number used for penalty term
#   RemoveCorners: float, if not np.nan, the corners will be fixed
    to their terminal value + RemoveCorners
#   numprocesses: integer, if greater that one, multiprocessing is
    used
# Outputs:
#   (Matrix, RHS), where Matrix is a sparse matrix.
#============================================================================
# Initializations
T = MarketParams['T']
eta1, eta2 = MarketParams['eta']
zBar = MeshParams['zBar']
N = MeshParams['N']
M = MeshParams['M']
I = MeshParams['I']
J = MeshParams['J']
Dims = (M,I)
xiBar = MeshParams['xiBar']
Delta_t = T / np.float64(N)
Delta_z = zBar / np.float64(M)
Delta_1 = xiBar[0] / np.float64(I)
```

```
Delta_2 = xiBar[1] / np.float64(J)
max_entries = 4 * (2*M+1) * (2*J+1) + 4 * (2*M+1) * (2*I-1) +\
             6 * (2*I-1) * (2*J-1) + 5 * (2*M-1) * (2*I-1) * (2*J-1)
row = np.array([np.nan for it in range(max_entries)]) #
    @UnusedVariable
col= np.array([np.nan for it in range(max_entries)]) #
    @UnusedVariable
data= np.array([np.nan for it in range(max_entries)]) #
    @UnusedVariable
entry_ix = 0
RHS = np.zeros((2*M+1)*(2*I+1)*(2*J+1))
# xi1-boundaries
sys.stdout.write("\n        xi1 boundaries... ")
start = time.time()
for m in range(2*M+1):
    for j in range(2*J+1):
        if((not np.isnan(RemoveCorners)) and
           ((j==0) or (j==2*J))):#fix the corners
            # I
            row[entry_ix] = _3ixToFlat((m,2*I,j),Dims)
            col[entry_ix] = _3ixToFlat((m,2*I,j),Dims)
            data[entry_ix] = 1.0
            entry_ix += 1
            RHS[_3ixToFlat((m,2*I,j),Dims)] = RemoveCorners
            # -I
            row[entry_ix]  = _3ixToFlat((m,0,j),Dims)
            col[entry_ix]  = _3ixToFlat((m,0,j),Dims)
            data[entry_ix] = 1.0
            entry_ix += 1
            RHS[_3ixToFlat((m,0,j),Dims)]  = RemoveCorners
        else:
            # I
            row[entry_ix] = _3ixToFlat((m,2*I,j),Dims)
            row[entry_ix+1] = _3ixToFlat((m,2*I,j),Dims)
            col[entry_ix] = _3ixToFlat((m,2*I,j),Dims)
            col[entry_ix+1] = _3ixToFlat((m,2*I-1,j),Dims)
            data[entry_ix] = 1.0
            data[entry_ix+1] = -1.0
            entry_ix += 2
            RHS[_3ixToFlat((m,2*I,j),Dims)] = \
                (Delta_1 * eta1 - w_old.ix[m,2*I,j] +
                 w_old.ix[m,2*I-1,j])
            # -I
```

```
            row[entry_ix]  = _3ixToFlat((m,0,j),Dims)
            row[entry_ix+1] = _3ixToFlat((m,0,j),Dims)
            col[entry_ix]  = _3ixToFlat((m,0,j),Dims)
            col[entry_ix+1] = _3ixToFlat((m,1,j),Dims)
            data[entry_ix] = 1.0
            data[entry_ix+1] = -1.0
            entry_ix += 2
            RHS[_3ixToFlat((m,0,j),Dims)] = \
                (Delta_1 * eta1 - w_old.ix[m,0,j] +
                 w_old.ix[m,1,j])
duration = time.time()-start
sys.stdout.write("duration: %i hr, %i min, %.2f sec."%
                 (int(duration)/3600, int(duration%3600)/60,
                    (duration%3600)%60))
# xi2-boundaries, i.e. j in {J,-J} and i ne I,-I
sys.stdout.write("\n        xi2 boundaries... ")
start = time.time()
for m in range(2*M+1):
    for i in range(1,2*I):
        # J
        row[entry_ix]  = _3ixToFlat((m,i,2*J),Dims)
        row[entry_ix+1] = _3ixToFlat((m,i,2*J),Dims)
        col[entry_ix]  = _3ixToFlat((m,i,2*J),Dims)
        col[entry_ix+1] = _3ixToFlat((m,i,2*J-1),Dims)
        data[entry_ix] = 1.0
        data[entry_ix+1]= -1.0
        entry_ix += 2
        RHS[_3ixToFlat((m,i,2*J),Dims)] = \
            (Delta_2 * eta2 - w_old.ix[m,i,2*J] +
             w_old.ix[m,i,2*J-1])
        # -J
        row[entry_ix]  = _3ixToFlat((m,i,0),Dims)
        row[entry_ix+1] = _3ixToFlat((m,i,0),Dims)
        col[entry_ix]  = _3ixToFlat((m,i,0),Dims)
        col[entry_ix+1] = _3ixToFlat((m,i,1),Dims)
        data[entry_ix] = 1.0
        data[entry_ix+1]= -1.0
        entry_ix += 2
        RHS[_3ixToFlat((m,i,2*J),Dims)] = \
            (Delta_2 * eta2 - w_old.ix[m,i,0] +
             w_old.ix[m,i,1])
duration = time.time()-start
sys.stdout.write("duration: %i hr, %i min, %.2f sec."%
```

```
              (int(duration)/3600, int(duration%3600)/60,
                (duration%3600)%60))
# z-boundaries, i.e. m in {M,-M}, i ne I,-I and j ne J,-J
sys.stdout.write("\n        z boundaries... ")
start = time.time()
for i in range(1,2*I):
    for j in range(1,2*J):
        # M
        # Calculating A and L operator
        A = Delta_t * (- ahat.ix[2*M,i,j] / Delta_z
                       - 2 * bhat / Delta_z**2
                       * (w_old.ix[2*M,i,j]
                          - w_old.ix[2*M-1,i,j]))
        L_op = (ahat.ix[2*M,i,j] / Delta_z * (w_old.ix[2*M,i,j] -
            w_old.ix[2*M-1,i,j])
                + bhat / (Delta_z**2) * (w_old.ix[2*M,i,j]**2
                                         + w_old.ix[2*M-1,i,j]**2
                                         - 2 * w_old.ix[2*M,i,j]
                                         * w_old.ix[2*M-1,i,j])
                + chat.ix[2*M,i,j])
        # Calculating penalty terms
        m = 2*M
        if((w_old.ix[m,i+1,j] - w_old.ix[m,i,j]) < - Delta_1 * (eta1
            - epsilon)):
            # Forward scheme
            P1 = Delta_t / Delta_1 / (epsilon**2)
            P1_IX = (m,i+1,j)
            P_op = - 1.0 / Delta_1 / (epsilon**2) *\
                ( - (w_old.ix[m,i+1,j] - w_old.ix[m,i,j])
                  - Delta_1 * (eta1 - epsilon))
        elif((w_old.ix[m,i,j] - w_old.ix[m,i-1,j])
            > Delta_1 * (eta1 - epsilon)):
            # Backward scheme
            P1 = Delta_t / Delta_1 / (epsilon**2)
            P1_IX = (m,i-1,j)
            P_op = - 1.0 / Delta_1 / (epsilon**2) *\
                ( w_old.ix[m,i,j] - w_old.ix[m,i-1,j]
                  - Delta_1 * (eta1 - epsilon))
        else:
            P1 = 0.0
            P1_IX = ()
            P_op = 0.0
        if((w_old.ix[m,i,j+1] - w_old.ix[m,i,j])
```

```
    < - Delta_2 * (eta2 - epsilon)):
    # Forward scheme
    P2 = Delta_t / Delta_2 / (epsilon**2)
    P2_IX = (m,i,j+1)
    P_op += - 1.0 / Delta_2 / (epsilon**2) *\
        ( - (w_old.ix[m,i,j+1] - w_old.ix[m,i,j])
          - Delta_2 * (eta2 - epsilon))
elif((w_old.ix[m,i,j] - w_old.ix[m,i,j-1])
    > Delta_2 * (eta2 - epsilon)):
    # Backward scheme
    P2 = Delta_t / Delta_2 / (epsilon**2)
    P2_IX = (m,i,j-1)
    P_op += - 1.0 / Delta_2 / (epsilon**2) *\
        ( w_old.ix[m,i,j] - w_old.ix[m,i,j-1]
          - Delta_2 * (eta2 - epsilon))
else:
    P2 = 0.0
    P2_IX = ()
    P_op += 0.0
# Filling the Matrix
row[entry_ix]  = _3ixToFlat((2*M,i,j),Dims)
row[entry_ix+1] = _3ixToFlat((2*M,i,j),Dims)
col[entry_ix]  = _3ixToFlat((2*M,i,j),Dims)
col[entry_ix+1] = _3ixToFlat((2*M-1,i,j),Dims)
data[entry_ix] = 1.0 + A + P1 + P2
data[entry_ix+1] = -A
entry_ix += 2
if (P1 != 0.0):
    row[entry_ix] = _3ixToFlat((2*M,i,j),Dims)
    col[entry_ix] = _3ixToFlat(P1_IX,Dims)
    data[entry_ix] = -P1
    entry_ix += 1
if (P2 != 0.0):
    row[entry_ix] = _3ixToFlat((2*M,i,j),Dims)
    col[entry_ix] = _3ixToFlat(P2_IX,Dims)
    data[entry_ix] = -P2
    entry_ix += 1
RHS[_3ixToFlat((2*M,i,j),Dims)] = \
    (w_lastItr.ix[2*M,i,j] - w_old.ix[2*M,i,j]
     + Delta_t * (L_op + P_op))
# -M
# Calculating A and L operator
A = Delta_t * (ahat.ix[0,i,j] / Delta_z
```

```
                + 2 * bhat / Delta_z**2
                * (w_old.ix[1,i,j]
                   - w_old.ix[0,i,j]))
L_op = (ahat.ix[0,i,j] / Delta_z * (w_old.ix[1,i,j] -
    w_old.ix[0,i,j])
        + bhat / (Delta_z**2) * (w_old.ix[0,i,j]**2
                                + w_old.ix[1,i,j]**2
                                - 2 * w_old.ix[0,i,j]
                                * w_old.ix[1,i,j])
        + chat.ix[0,i,j])
# Calculating penalty terms
m = 0
if((w_old.ix[m,i+1,j] - w_old.ix[m,i,j])
    < - Delta_1 * (eta1 - epsilon)):
    # Forward scheme
    P1 = Delta_t / Delta_1 / (epsilon**2)
    P1_IX = (m,i+1,j)
    P_op = - 1.0 / Delta_1 / (epsilon**2) *\
        ( - (w_old.ix[m,i+1,j] - w_old.ix[m,i,j])
          - Delta_1 * (eta1 - epsilon))
elif((w_old.ix[m,i,j] - w_old.ix[m,i-1,j])
    > Delta_1 * (eta1 - epsilon)):
    # Backward scheme
    P1 = Delta_t / Delta_1 / (epsilon**2)
    P1_IX = (m,i-1,j)
    P_op = - 1.0 / Delta_1 / (epsilon**2) *\
        ( w_old.ix[m,i,j] - w_old.ix[m,i-1,j]
          - Delta_1 * (eta1 - epsilon))
else:
    P1 = 0.0
    P1_IX = ()
    P_op = 0.0
if((w_old.ix[m,i,j+1] - w_old.ix[m,i,j])
    < - Delta_2 * (eta2 - epsilon)):
    # Forward scheme
    P2 = Delta_t / Delta_2 / (epsilon**2)
    P2_IX = (m,i,j+1)
    P_op += - 1.0 / Delta_2 / (epsilon**2) *\
        ( - (w_old.ix[m,i,j+1] - w_old.ix[m,i,j])
          - Delta_2 * (eta2 - epsilon))
elif((w_old.ix[m,i,j] - w_old.ix[m,i,j-1])
    > Delta_2 * (eta2 - epsilon)):
    # Backward scheme
```

```
            P2 = Delta_t / Delta_2 / (epsilon**2)
            P2_IX = (m,i,j-1)
            P_op += - 1.0 / Delta_2 / (epsilon**2) *\
                ( w_old.ix[m,i,j] - w_old.ix[m,i,j-1]
                  - Delta_2 * (eta2 - epsilon))
        else:
            P2 = 0.0
            P2_IX = ()
            P_op += 0.0
        # Filling the Matrix
        row[entry_ix]  = _3ixToFlat((0,i,j),Dims)
        row[entry_ix+1] = _3ixToFlat((0,i,j),Dims)
        col[entry_ix]  = _3ixToFlat((0,i,j),Dims)
        col[entry_ix+1] = _3ixToFlat((1,i,j),Dims)
        data[entry_ix] = 1.0 + A + P1 + P2
        data[entry_ix+1] = -A
        entry_ix += 2
        if (P1 != 0.0):
            row[entry_ix] = _3ixToFlat((0,i,j),Dims)
            col[entry_ix] = _3ixToFlat(P1_IX,Dims)
            data[entry_ix] = -P1
            entry_ix += 1
        if (P2 != 0.0):
            row[entry_ix] = _3ixToFlat((0,i,j),Dims)
            col[entry_ix] = _3ixToFlat(P2_IX,Dims)
            data[entry_ix] = -P2
            entry_ix += 1
        RHS[_3ixToFlat((0,i,j),Dims)] = \
            (w_lastItr.ix[0,i,j] - w_old.ix[0,i,j]
             + Delta_t * (L_op + P_op))
duration = time.time()-start
sys.stdout.write("duration: %i hr, %i min, %.2f sec."%
                 (int(duration)/3600, int(duration%3600)/60,
                     (duration%3600)%60))
# The interior of the grid
# argument lists
args = [{'I':I, 'J':J,'Dims':Dims,
         'eta1':eta1, 'eta2':eta2,
         'Delta_t':Delta_t, 'Delta_z':Delta_z,
         'Delta_1':Delta_1, 'Delta_2':Delta_2,
         'ahat':ahat, 'bhat':bhat, 'chat':chat,
         'epsilon':epsilon, 'w_old':w_old, 'w_lastItr':w_lastItr,
         'm':m} for m in range(1,2*M)]
```

```
    # terminate row, column and data
    row = row[:entry_ix]
    col = col[:entry_ix]
    data = data[:entry_ix]
    sys.stdout.write("\n        the interior... ")
    start = time.time()
    if (numprocesses > 1): # multiprocessing
        pool = Pool(processes=numprocesses)
        res = pool.map(_AlgebSys_Aux, args)
        pool.close()
        pool.join()
        for m in range(1,2*M):
            row = np.append(row, res[m-1]['row'])
            col = np.append(col, res[m-1]['col'])
            data = np.append(data, res[m-1]['data'])
            for RHS_ix in range(len(res[m-1]['RHS_IXs'])):
                RHS[res[m-1]['RHS_IXs'][RHS_ix]] = res[m-1]['RHS'][RHS_ix]
    else: #single processing
        for m in range(1,2*M):
            res = _AlgebSys_Aux(args[m-1])
            row = np.append(row, res['row'])
            col = np.append(col, res['col'])
            data = np.append(data, res['data'])
            for RHS_ix in range(len(res['RHS_IXs'])):
                RHS[res['RHS_IXs'][RHS_ix]] = res['RHS'][RHS_ix]
    duration = time.time()-start
    sys.stdout.write("duration: %i hr, %i min, %.2f sec."%
                (int(duration)/3600, int(duration%3600)/60,
                    (duration%3600)%60))
    FullDim = (2*M+1)*(2*I+1)*(2*J+1)
    Matrix = csr_matrix((data,(row,col)),shape=(FullDim,FullDim))
    RHS = np.array(RHS)
    return (Matrix, RHS)


def _PolVI(LwPLUSwt, w1, w2, eta1, eta2, error_tol):
    #============================================================================
    # Auxiliary function used in _TradingPolicy. Given w_t, w_1, w_2 and
        Lw, determine
    # the policy and whether V.I. is satisfied
    # Output:
    #   pol:
    #                                 0.5 (S2)
```

```
#
#                      0.75 (B1&S2)        0.25 (S1&S2)
#
#                  1.0 (B1)      -0.25 (NT)        0.0 (S1)
#
#                      1.25 (B1&B2)        1.75 (S1&B2)
#
#                                 1.5 (B2)
#     vi:
#           2.0: V.I. is satisfies
#           -0.5, 0.0, 0.5, 1.0, and 1.5: The corresponding
   inequality does not hold
#           -1.0: Non of the inequalities are strict
#============================================================================
vi = 2.0
if (LwPLUSwt <= error_tol):# no trade region
    pol = -0.25
    if(LwPLUSwt < -error_tol):
        vi = -0.5
elif(eta1 + w1 <= error_tol): # B1
    if(eta2 + w2 <= error_tol): # B1B2
        pol = 1.25
    elif(eta2 - w2 <= error_tol): # B1S2
        pol = 0.75
    else:
        pol = 1.0 #B1
    if(eta1 + w1 < -error_tol):
        vi = 1.0
elif(eta1 - w1 <= error_tol): # S1
    if(eta2 + w2 <= error_tol): # S1B2
        pol = 1.75
    elif(eta2 - w2 <= error_tol): # S1S2
        pol = 0.25
    else:
        pol = 0.0
    if(eta1 - w1 < -error_tol):
        vi = 0.0
elif(eta2 + w2 <= error_tol): # B2
    pol = 1.5
    if(eta2 + w2 <= error_tol):
        vi = 1.5
elif(eta2 - w2 <= error_tol): # S2
    pol = 0.5
```

```
        if(eta2 - w2 < -error_tol):
            vi= 0.5
    else:
        pol = -0.5
        vi = -1.0
    return pol, vi


def _TradingPolicy_Aux(arg):
    #==============================================================================
    # Given log(-scaled value function), identifies the trading policy
        and checks
    # that the associated variational inequality is satisfies.
    # Inputs:
    #   arg is a dictionary with the following keys:
    #   w: dictionary of Panels. The key of the dictionary is time
        steps 0,1,... N.
    #       w[n] contains the value of the function w = log(-scaled
        value function) at t_n.
    #       the items are z_m, the major axis is xi_i and the minor axis
        is xi_2.
    #   n: integer in w.keys(), the time at which the calculations are
        done
    #   MarketParams: Dictionary with the following keys:
    #       'alpha': pair of floats (alpha1,alpha2), factor loadings.
    #       'c': float, cointegration coefficient
    #       'rho': float, correlation
    #       'T': float, investment horizon
    #       'eta': pair of floats, proportional transaction costs in
        each market
    #   MeshParams: Dictionary with the following keys:
    #       'zBar': positive float, z is assumed to take values in
        (-zBar, zBar)
    #       'xiBar': pair of floats (xiBar_1, xiBar_2). xi_i takes value
        in (-xiBar_i, xiBar_i)
    #       'N': integer, number of time steps
    #       'M': integer, number of z steps
    #       'I': integer, number of xi_1 steps
    #       'J': integer, number of xi_2 steps
    #   ahat, bhat, chat: coefficients, ahat and chat are the same
        dimension as
    #       w[n], while bhat is a scalar.
```

```
#   Pol_error: float, the tolerance used when determining the
    policy and status
#       of the variational inequality
#
# Outputs:
#   Pol: Panel, contains the trading policy at t_n. The items are
    z_m, the major
#       axis is xi_i and the minor axis is xi_2. Each entry is
    labeled as follows:
#                               0.5 (S2)
#
#                    0.75 (B1&S2)        0.25 (S1&S2)
#
#                 1.0 (B1)      -0.25 (NT)       0.0 (S1)
#
#                    1.25 (B1&B2)        1.75 (S1&B2)
#
#                               1.5 (B2)
#   VI: Panel, determines whether the variational inequality at t_n
    is satisfied or not.
#       The items are z_m, the major axis is xi_i and the minor axis
    is xi_2.
#       Each entry is labeled as follows:
#           2.0: V.I. is satisfies
#           -0.5, 0.0, 0.5, 1.0, and 1.5: The corresponding
    inequality does not hold
#           -1.0: Non of the inequalities are strict
#============================================================================
# Initializations
w = arg['w']
n = arg['n']
MarketParams = arg['MarketParams']
MeshParams = arg['MeshParams']
ahat = arg['ahat']
bhat = arg['bhat']
chat = arg['chat']
Pol_error = arg['Pol_error']
T = MarketParams['T']
eta1, eta2 = MarketParams['eta']
zBar = MeshParams['zBar']
N = MeshParams['N']
M = MeshParams['M']
I = MeshParams['I']
```

```
J = MeshParams['J']
xiBar = MeshParams['xiBar']
Delta_t = T / np.float64(N)
Delta_z = zBar / np.float64(M)
Delta_1 = xiBar[0] / np.float64(I)
Delta_2 = xiBar[1] / np.float64(J)
Pol = pd.Panel(items=w[n].items,
              major_axis=w[n].major_axis,
              minor_axis=w[n].minor_axis)
VI = Pol.copy()
# identifying the policy and V.I. status
for m in range(2*M+1):
    for i in range(2*I+1):
        for j in range(2*J+1):
            # Calculate w_t
            wt = (w[n+1].ix[m,2*I,j] - w[n].ix[m,2*I,j]) / Delta_t
            # Calculate w_1
            if(i==2*I):
                w1 = (w[n].ix[m,2*I,j] - w[n].ix[m,2*I-1,j]) / Delta_1
            elif(i==0):
                w1 = (w[n].ix[m,1,j] - w[n].ix[m,0,j]) / Delta_1
            else:
                w1 = (w[n].ix[m,i+1,j] - w[n].ix[m,i-1,j]) / 2
                    /Delta_1
            # Calculate w_2
            if(j==2*J):
                w2 = (w[n].ix[m,i,2*J] - w[n].ix[m,i,2*J-1]) / Delta_2
            elif(j==0):
                w2 = (w[n].ix[m,i,1] - w[n].ix[m,i,0]) / Delta_2
            else:
                w2 = (w[n].ix[m,i,j+1] - w[n].ix[m,i,j-1]) / 2
                    /Delta_2
            # Calculate Lw
            if(m==2*M):
                Lw = (ahat.ix[2*M,i,j] / Delta_z * (w[n].ix[2*M,i,j]
                    - w[n].ix[2*M-1,i,j])
                      + bhat / (Delta_z**2) * (w[n].ix[2*M,i,j]**2
                                              + w[n].ix[2*M-1,i,j]**2
                                              - 2 * w[n].ix[2*M,i,j]
                                              * w[n].ix[2*M-1,i,j])
                      + chat.ix[2*M,i,j])
            elif(m==0):
```

```
                Lw = (ahat.ix[0,i,j] / Delta_z * (w[n].ix[1,i,j] -
                    w[n].ix[0,i,j])
                    + bhat / (Delta_z**2) * (w[n].ix[0,i,j]**2
                                             + w[n].ix[1,i,j]**2
                                             - 2 * w[n].ix[0,i,j]
                                             * w[n].ix[1,i,j])
                    + chat.ix[0,i,j])
            else:
                Lw = ( w[n].ix[m-1,i,j] * ( bhat / (Delta_z**2) -
                    ahat.ix[m,i,j] / 2.0 / Delta_z)
                           - w[n].ix[m,i,j] * 2.0 * bhat /
                               (Delta_z**2)
                           + w[n].ix[m+1,i,j] * ( bhat / (Delta_z**2)
                               + ahat.ix[m,i,j] / 2.0 / Delta_z)
                           + bhat / 4.0 / (Delta_z**2)
                           *(w[n].ix[m+1,i,j]**2 + w[n].ix[m-1,i,j]**2
                             - 2.0 * w[n].ix[m+1,i,j] *
                                 w[n].ix[m-1,i,j])
                           + chat.ix[m,i,j])
            # Determine the policy and whether V.I. is satisfied
            Pol.ix[m,i,j], VI.ix[m,i,j] = _PolVI((Lw+wt), w1, w2,
                eta1, eta2, Pol_error)
    return (Pol, VI)


def _TradingPolicy(w,MarketParams,MeshParams, ahat, bhat, chat,
    Pol_error, numprocesses=8):
    #=======================================================================
    # Given log(-scaled value function), identifies the trading policy
        and checks
    # that the associated variational inequality is satisfies.
    # Inputs:
    #    w:  dictionary of Panels. The key of the dictionary is time
        steps 0,1,... N.
    #       w[n] contains the value of the function w = log(-scaled
        value function) at t_n.
    #       the items are z_m, the major axis is xi_i and the minor axis
        is xi_2.
    #    MarketParams: Dictionary with the following keys:
    #       'alpha': pair of floats (alpha1,alpha2), factor loadings.
    #       'c': float, cointegration coefficient
    #       'rho': float, correlation
    #       'T': float, investment horizon
```

```
#       'eta': pair of floats, proportional transaction costs in
   each market
#    MeshParams: Dictionary with the following keys:
#       'zBar': positive float, z is assumed to take values in
   (-zBar, zBar)
#       'xiBar': pair of floats (xiBar_1, xiBar_2). xi_i takes value
   in (-xiBar_i, xiBar_i)
#       'N': integer, number of time steps
#       'M': integer, number of z steps
#       'I': integer, number of xi_1 steps
#       'J': integer, number of xi_2 steps
#    ahat, bhat, chat: coefficients, ahat and chat are the same
   dimension as
#       w[n], while bhat is a scalar.
#    Pol_error: float, the tolerance used when determining the
   policy and status
#       of the variational inequality
#    numprocesses: integer, if greater that one, multiprocessing is
   used
# Outputs:
#    Pol: dictionary of Panels. The key of the dictionary is time
   steps 0,1,... N-1.
#       Pol[n] contains the trading policy at t_n. The items are
   z_m, the major
#       axis is xi_1 and the minor axis is xi_2. Each entry is
   labeled as follows:
#                               0.5 (S2)
#
#                     0.75 (B1&S2)        0.25 (S1&S2)
#
#                  1.0 (B1)      -0.25 (NT)        0.0 (S1)
#
#                     1.25 (B1&B2)        1.75 (S1&B2)
#
#                               1.5 (B2)
#    VI: dictionary of Panels. The key of the dictionary is time
   steps 0,1,... N-1.
#       VI[n] determines whether the variational inequality is
   satisfied or not.
#       The items are z_m, the major axis is xi_1 and the minor axis
   is xi_2.
#       Each entry is labeled as follows:
#          2.0: V.I. is satisfies
```

```
    #           -0.5, 0.0, 0.5, 1.0, and 1.5: The corresponding
        inequality does not hold
    #           -1.0: Non of the inequalities are strict
    #============================================================================
    # Initializations
    Pol = {}
    VI = {}
    N = MeshParams['N']
    args = [{'w':w, 'n': n, 'MarketParams': MarketParams, 'MeshParams':
        MeshParams,
                'ahat':ahat, 'bhat':bhat, 'chat':chat,
                    'Pol_error':Pol_error} for n in range(N)]
    if (numprocesses > 1): # multiprocessing
        pool = Pool(processes=numprocesses)
        res = pool.map(_TradingPolicy_Aux, args)
        pool.close()
        pool.join()
        for n in range(N):
            Pol[n], VI[n] = res[n]
    else: #single processing
        for n in range(N):
            Pol[n], VI[n] = _TradingPolicy_Aux(args[n])
    return (Pol, VI)


def TC(MarketParams,MeshParams,OutPutPath,tol=1e-5,
       epsilon=1e-5,Pol_error=1e-4,RemoveCorners=np.nan,
       numprocesses=1,Z_levels=-1.0,Symmetries=False):
    #============================================================================
    # Finite deference scheme to find the no trade region.
    # Inputs:
    #    MarketParams: Dictionary with the following keys:
    #        'alpha': pair of floats (alpha1,alpha2), factor loadings.
    #        'c': float, cointegration coefficient
    #        'rho': float, correlation
    #        'T': float, investment horizon
    #        'eta': pair of floats, proportional transaction costs in
        each market
    #    MeshParams: Dictionary with the following keys:
    #        'zBar': positive float, z is assumed to take values in
        (-zBar, zBar)
    #        'xiBar': pair of floats (xiBar_1, xiBar_2). xi_i takes value
        in (-xiBar_i, xiBar_i)
```

```
#       'N': integer, number of time steps
#       'M': integer, number of z steps
#       'I': integer, number of xi_1 steps
#       'J': integer, number of xi_2 steps
#    OutPutPath: string, the path to the folder in which the outputs
   will be created.
#    tol: float, the tolerance for convergence of Newton's iteration
#    epsilon: float, small number used for penalty term
#    Pol_error: float, the tolerance used when determining the
   policy and status of variational inequality
#    RemoveCorners: float, if not np.nan, the corners will be fixed
   to their terminal value + RemoveCorners
#    numprocesses: integer, if greater that one, multiprocessing is
   used
#    Z_levels: integer, if >= 1, plots at z levels
   -Z_levels,...,Z_levels, if =-1, plots all
#    Symmetries: boolean, if true, symmetries the no-trade plots
   around the origin
# Outputs:
#    w:  dictionary of Panels. The key of the dictionary is time
   steps 0,1,... N.
#       w[n] contains the value of the function w = log(-scaled
   value function) at t_n.
#       the items are z_m, the major axis is xi_1 and the minor axis
   is xi_2.
#
#    Pol: dictionary of Panels. The key of the dictionary is time
   steps 0,1,... N-1.
#       Pol[n] contains the trading policy at t_n. The items are
   z_m, the major
#       axis is xi_1 and the minor axis is xi_2. Each entry is
   labeled as follows:
#                               0.5 (S2)
#
#                     0.75 (B1&S2)        0.25 (S1&S2)
#
#                  1.0 (B1)      -0.25 (NT)        0.0 (S1)
#
#                     1.25 (B1&B2)        1.75 (S1&B2)
#
#                               1.5 (B2)
#    VI: dictionary of Panels. The key of the dictionary is time
   steps 0,1,... N-1.
```

```
#       VI[n] determines whether the variational inequality is
    satisfied or not.
#       The items are z_m, the major axis is xi_1 and the minor axis
    is xi_2.
#       Each entry is labeled as follows:
#           2.0: V.I. is satisfies
#           -0.5, 0.0, 0.5, 1.0, and 1.5: The corresponding
    inequality does not hold
#           -1.0: Non of the inequalities are strict
#============================================================================
#Initialization
Grand_start = time.time()
alpha1,alpha2 = MarketParams['alpha']
c = MarketParams['c']
rho = MarketParams['rho']
T = MarketParams['T']
eta1, eta2 = MarketParams['eta']
zBar = MeshParams['zBar']
xiBar = MeshParams['xiBar']
N = MeshParams['N']
M = MeshParams['M']
I = MeshParams['I']
J = MeshParams['J']
Delta_1 = xiBar[0] / np.float64(I)
Delta_2 = xiBar[1] / np.float64(J)
time_ix = range(N,-1,-1)
z_ix = range(-M,M+1,1)
xi1_ix = range(-I,I+1,1)
xi2_ix = range(-J,J+1,1)
GridTemplate = pd.Panel(items=['%i'%(it) for it in z_ix],
                     major_axis=['%i'%(it) for it in xi1_ix],
                     minor_axis=['%i'%(it) for it in xi2_ix])
x,y = np.meshgrid(np.linspace(-xiBar[0],xiBar[0], num=2*I+1),
                np.linspace(-xiBar[1],xiBar[1],
                    num=2*J+1),indexing='ij')
z = np.linspace(-zBar,zBar, num=2*M+1)
# Calculating the coefficients
sys.stdout.write('calculating the coefficients...')
ahat = GridTemplate.copy()
chat = GridTemplate.copy()
for ix, m in enumerate(GridTemplate.items):
    ahat.ix[m] = (alpha1 - c * alpha2) * z[ix] - (1 - c * rho) * x -
        (rho - c) * y
```

```
    chat.ix[m] = 0.5 * x**2 + 0.5 * y**2 + rho * x * y - z[ix] *
        (alpha1 * x + alpha2 * y)
bhat = 0.5 * (1 + c**2 - 2 * rho * c)
sys.stdout.write('done\n')
# Set w[N]
w = {}
w[N] = GridTemplate.copy()
sys.stdout.write('Solving for time T = %i ...'%(N))
for m in GridTemplate.items:
    w[N].ix[m] = eta1 * np.abs(x) + eta2 * np.abs(y)
sys.stdout.write('done\n')
for n in time_ix[1:]:
    sys.stdout.write('Solving for time T = %i ...\n'%(n))
    w_old = w[n+1].copy()
    # Setup the algebraic system
    Nitr = 0
    sys.stdout.write('  Newton iteration = %i,\n     Setting up the
        algebraic system...'%(Nitr))
    CoefMat, RHS = _AlgebSys(MarketParams,MeshParams,
                             w_old,w[n+1],ahat,
                             bhat,chat,epsilon,RemoveCorners=RemoveCorners,
                             numprocesses=numprocesses)
    sys.stdout.write('done,\n     Solving the system...')
    # Solve the algebraic system
    Dw_new_flatten = spsolve(CoefMat,RHS)
    sys.stdout.write('done,\n     storing the solution...')
    Dw_new = GridTemplate.copy()
    w_new = GridTemplate.copy()
    for ix in range(len(Dw_new_flatten)):
        m,i,j = _FlatTo3ix(ix,(M,I))
        Dw_new.ix[m,i,j] = Dw_new_flatten[ix]
        w_new.ix[m,i,j] = w_old.ix[m,i,j] + Dw_new_flatten[ix]
    sys.stdout.write('done, ')
    max_err =
        np.max(np.abs(Dw_new.values)/np.maximum(np.abs(w_new.values),1.0))
    reshape_error = np.max(np.abs(Dw_new.values - w_new.values +
        w_old.values))
    sys.stdout.write('\n     reshape error = %f, '%(reshape_error))
    sys.stdout.write('error = %f\n'%(max_err))
    while (max_err >= tol):
        Nitr += 1
        sys.stdout.write('  Newton iteration = %i:\n     Setting up
            the algebraic system...'%(Nitr))
```

```
        w_old = w_new.copy()
        # Setup the algebraic system
        CoefMat, RHS = _AlgebSys(MarketParams,MeshParams,
                              w_old,w[n+1],ahat,
                              bhat,chat,epsilon,RemoveCorners=RemoveCorners,
                              numprocesses=numprocesses)
        sys.stdout.write('done, \n     Solving the system...')
        # Solve the algebraic system
        Dw_new_flatten = spsolve(CoefMat,RHS)
        sys.stdout.write('done, \n     storing the solution...')
        Dw_new = GridTemplate.copy()
        w_new = GridTemplate.copy()
        for ix in range(len(Dw_new_flatten)):
            m,i,j = _FlatTo3ix(ix,(M,I))
            Dw_new.ix[m,i,j] = Dw_new_flatten[ix]
            w_new.ix[m,i,j] = w_old.ix[m,i,j] + Dw_new_flatten[ix]
        sys.stdout.write('done, ')
        max_err =
            np.max(np.abs(Dw_new.values)/np.maximum(np.abs(w_new.values),1.0))
        reshape_error = np.max(np.abs(Dw_new.values - w_new.values +
            w_old.values))
        sys.stdout.write('\n  reshape error = %f, '%(reshape_error))
        sys.stdout.write('error = %f\n'%(max_err))
    sys.stdout.write('============================================\nConvergence
        achieved, storing the solution...')
    w[n] = w_new.copy()
    sys.stdout.write('done\n============================================\n')
# Verifying the solution and identifying the trading strategy
sys.stdout.write('Verifying the solution and identifying the trading
    strategy...')
Pol, VI = _TradingPolicy(w,MarketParams,MeshParams, ahat, bhat,
    chat, Pol_error,numprocesses=numprocesses)
sys.stdout.write('done\n============================================\n')
#============================================================================
#                              plots
#============================================================================
# Initializations
file_name = ('(%i,%i,%i,%i)_'
            '(%.2f,%.2f,%.4f,%.2f,%.2f)_'
            '(%.0e,%.0e,%.2e)'%(N,M,I,J,
                              c,rho,T,eta1,eta2,
                              tol,epsilon,Pol_error))
pdf_path = '%s/%s.pdf'%(OutPutPath,file_name)
```

```
PDFfile = PdfPages(pdf_path)
plt.Figure()
plt.suptitle('Market parameters:\n'
             '   alpha: (%.2f,%.2f)\n'
             '   c: %.2f\n'
             '   rho: %.2f\n'
             '   T: %.4f\n'
             '   eta: (%.2f,%.2f)\n'
             'Mesh parameters:\n'
             '   zBar: %.2f\n'
             '   xiBar: (%.2f,%.2f)\n'
             '   N: %i\n'
             '   M: %i\n'
             '   I: %i\n'
             '   J: %i\n'
             'Other parameters:\n'
             '   tol: %.2e\n'
             '   epsilon: %.2e\n'
             '   Pol_error: %.2e\n'
             '   numprocesses: %i'%(alpha1,alpha2,c,rho,T,eta1,eta2,
                                     zBar,xiBar[0],xiBar[1],N,M,I,J,
                                     tol,epsilon,Pol_error,numprocesses),
             x=0.1,horizontalalignment='left')
plt.gcf().set_size_inches(8.3,11.7)
plt.gcf().set_dpi(100)
plt.savefig(PDFfile, orientation='portrait', format='pdf',dpi=200)
plt.close()
if (Z_levels<=0):
    Z_levels = M
Z_range = range(M-Z_levels,M+Z_levels+1,1)
# 2D Plots of w
sys.stdout.write('********************************************\n')
sys.stdout.write('********************************************\n')
sys.stdout.write('plotting w...')
for n in range(N,-1,-1):
    plt.Figure()
    plt.suptitle('log scaled value function for period %i/%i'%(n,N))
    gs = matplotlib.gridspec.GridSpec(Z_levels+2,2,left=0.1,
        right=0.9,
                                      top=0.9, bottom=0.1,
                                      hspace=0.7,wspace=0.2)
    for ix, m in enumerate(Z_range):
        if (m<M):
```

```
            plt.subplot(gs[ix,0])
        elif(m == M):
            plt.subplot(gs[ix:,0:])
        else:
            plt.subplot(gs[2*Z_levels-ix,1])
        xi1_bnds = np.linspace(-xiBar[0] - Delta_1/2,
                               xiBar[0] + Delta_1/2,
                               num=2*I+2)
        xi2_bnds = np.linspace(-xiBar[1] - Delta_2/2.0,
                               xiBar[1] + Delta_2/2.0,
                               num=2*J+2)
        VF_vals = w[n].ix[m].values.T # NOTE that xi1 is the major
            axis, translates to vertical axis, hence .T
        levels =
            matplotlib.ticker.MaxNLocator(nbins=15).tick_values(VF_vals.min(),
            VF_vals.max())
        # pick the desired colormap, sensible levels, and define a
            normalization.
        cmap = plt.get_cmap('PiYG')
        norm = matplotlib.colors.BoundaryNorm(levels,
            ncolors=cmap.N, clip=True)
        plt.pcolormesh(xi1_bnds, xi2_bnds, VF_vals, cmap=cmap,
            norm=norm,rasterized=True)
        plt.colorbar()
        # set the limits of the plot to the limits of the data
        plt.axis([xi1_bnds.min(), xi1_bnds.max(), xi2_bnds.min(),
            xi2_bnds.max()])
        plt.title('z=%.4f'%(z[m]))
    #saving to pdf file
    plt.gcf().autofmt_xdate()
    plt.gcf().set_size_inches(8.3,11.7)
    plt.gcf().set_dpi(100)
    plt.savefig(PDFfile, orientation='portrait',
        format='pdf',dpi=200)
    plt.close()
sys.stdout.write('done\n')
# 2D Plots of no-trade region
sys.stdout.write('********************************************\n')
sys.stdout.write('********************************************\n')
sys.stdout.write('plotting the no-trade regions...')
for n in range(N-1,-1,-1):
    plt.Figure()
    plt.suptitle('No-trade region for period %i/%i'%(n,N))
```

```
gs = matplotlib.gridspec.GridSpec(Z_levels+2,2,left=0.1,
    right=0.9,
                                 top=0.9, bottom=0.1,
                                 hspace=0.2,wspace=0.1)
for ix, m in enumerate(Z_range):
    if (m<M):
        plt.subplot(gs[ix,0], aspect='equal')
    elif(m == M):
        plt.subplot(gs[ix:,0:], aspect='equal')
    else:
        plt.subplot(gs[2*Z_levels-ix,1], aspect='equal')
    xi1_bnds = np.linspace(-xiBar[0] - Delta_1/2,
                           xiBar[0] + Delta_1/2,
                           num=2*I+2)
    xi2_bnds = np.linspace(-xiBar[1] - Delta_2/2.0,
                           xiBar[1] + Delta_2/2.0,
                           num=2*J+2)
    if Symmetries:
        Pol_vals = (Pol[n].ix[m].values.T if(m>=M)
                    else
                        np.fliplr(np.rot90(Pol[n].ix[2*M-m].values,1)))
    else:
        Pol_vals = Pol[n].ix[m].values.T # xi1 is the major_axis
            which translates to vertical axis, hence .T
    # pick the desired colormap, sensible levels, and define a
        normalization.
    levels = [-0.5, -0.25, 0.0, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5,
        1.75]
    bounds = (np.asarray(levels+[2.0]) - 0.125).tolist()
    cbar_labels = ['V.I. not strict', 'N-T', 'S1', 'S1&S2', 'S2',
                   'B1&S2', 'B1', 'B1&B2', 'B2', 'S1&B2']
    pol_colors = ["#000000", # V.I. not strict, black
                  "#FFFFFF", # no-trade, grey
                  "#F26C4F", # S1, light red
                  "#FF0000", # S1&S2, dark red
                  "#F26C4F", # S2, light red
                  "#00FF00", # B1&S2, green
                  "#5574B9", # B1, light blue
                  "#0000FF", # B1&B2, dark blue
                  "#5574B9", # B2, light blue
                  "#00FF00"] # S1&B2, green
    cmap = matplotlib.colors.ListedColormap(pol_colors)
```

```
        norm = matplotlib.colors.BoundaryNorm(bounds,
            ncolors=cmap.N, clip=True)
        plt.pcolormesh(xi1_bnds, xi2_bnds, Pol_vals, cmap=cmap,
            norm=norm, rasterized=True)
        # set the limits of the plot to the limits of the data
        plt.axis([xi1_bnds.min(), xi1_bnds.max(), xi2_bnds.min(),
            xi2_bnds.max()])
        if(m==M):
            cbar = plt.colorbar(ticks=levels)
            cbar.set_ticklabels(cbar_labels)
        plt.title('z=%.4f'%(z[m]))
    #saving to pdf file
    plt.gcf().autofmt_xdate()
    plt.gcf().set_size_inches(8.3,11.7)
    plt.gcf().set_dpi(100)
    plt.savefig(PDFfile, orientation='portrait',
        format='pdf',dpi=200)
    plt.close()
sys.stdout.write('done\n')
# 2D Plots of V.I status
sys.stdout.write('********************************************\n')
sys.stdout.write('********************************************\n')
sys.stdout.write('plotting V.I. status...')
for n in range(N-1,-1,-1):
    plt.Figure()
    plt.suptitle('Status of the variational inequality for period
        %i/%i'%(n,N))
    gs = matplotlib.gridspec.GridSpec(Z_levels+2,2,left=0.1,
        right=0.9,
                                  top=0.9, bottom=0.1,
                                  hspace=0.2,wspace=0.1)
    for ix, m in enumerate(Z_range):
        if (m<M):
            plt.subplot(gs[ix,0])
        elif(m == M):
            plt.subplot(gs[ix:,0:])
        else:
            plt.subplot(gs[2*Z_levels-ix,1])
        xi1_bnds = np.linspace(-xiBar[0] - Delta_1/2,
                             xiBar[0] + Delta_1/2,
                             num=2*I+2)
        xi2_bnds = np.linspace(-xiBar[1] - Delta_2/2.0,
                             xiBar[1] + Delta_2/2.0,
```

```
                                num=2*J+2)
        VI_vals = VI[n].ix[m].values.T # xi1 is the major_axis which
           translates to vertical axis, hence .T
        # pick the desired colormap, sensible levels, and define a
           normalization.
        levels = [-1.0, -0.5, 0.0, 0.5, 1.0, 1.5, 2.0]
        bounds = (np.asarray(levels+[2.5]) - 0.25).tolist()
        cbar_labels = ['V.I. is not strict', r'$w_t+\mathfrak{L}w <
           0$', '$w_1 > \eta_1$',
                       r'$w_2 > \eta_2$', r'$w_1 < -\eta_1$', r'$w_2 <
                          -\eta_2$',
                       'V.I. is satisfied']
        pol_colors = ["#000000", # V.I. is not strict, black
                      "#FFFFFF", # no-trade
                      "#F26C4F", # S1, light red
                      "#F26C4F", # S2, light red
                      "#5574B9", # B1, light blue
                      "#5574B9", # B2, light blue
                      "#00FF00"] # V.I. is satisfied
        cmap = matplotlib.colors.ListedColormap(pol_colors)
        norm = matplotlib.colors.BoundaryNorm(bounds,
           ncolors=cmap.N, clip=True)
        plt.pcolormesh(xi1_bnds, xi2_bnds, VI_vals, cmap=cmap,
           norm=norm, rasterized=True)
        # set the limits of the plot to the limits of the data
        plt.axis([xi1_bnds.min(), xi1_bnds.max(), xi2_bnds.min(),
           xi2_bnds.max()])
        if(m==M):
            cbar = plt.colorbar(ticks=levels)
            cbar.set_ticklabels(cbar_labels)
        plt.title('z=%.4f'%(z[m]))
    #saving to pdf file
    plt.gcf().autofmt_xdate()
    plt.gcf().set_size_inches(8.3,11.7)
    plt.gcf().set_dpi(100)
    plt.savefig(PDFfile, orientation='portrait',
        format='pdf',dpi=200)
    plt.close()
sys.stdout.write('done\n')
#closing the pdf file
PDFfile.close()
sys.stdout.write('*********************************************\n')
sys.stdout.write('*********************************************\n')
```

```
Total_duration = time.time()-Grand_start
sys.stdout.write("Total duration: %i hr, %i min, %.2f sec."%
                 (int(Total_duration)/3600,
                     int(Total_duration%3600)/60,
                  (Total_duration%3600)%60))
return {'w':w, 'Pol':Pol, 'VI':VI}
```

Bibliography

Abou-Kandil, H., G. Freiling, and G. Jank: Solution and asymptotic behavior of coupled riccati equations in jump linear systems. *Automatic Control, IEEE Transactions on*, volume 39, no. 8: pp. 1631–1636 (1994).

Ait Rami, M., X. Chen, J. B. Moore, and X. Y. Zhou: Solvability and asymptotic behavior of generalized riccati equations arising in indefinite stochastic lq controls. *IEEE Transactions on Automatic Control*, volume 46, no. 3: pp. 428–440 (2001).

Ait Rami, M., J. B. Moore, and X. Y. Zhou: Indefinite stochastic linear quadratic control and generalized differential riccati equation. *SIAM Journal on Control and Optimization*, volume 40, no. 4: pp. 1296–1311 (2002).

Angoshtari, B.: *On utility of wealth maximization.* Master's thesis, University of Twente (2009).

Avellaneda, M. and J. Lee: Statistical arbitrage in the US equities market. *Quantitative Finance*, volume 10, no. 7: pp. 761–778 (2010).

Baillie, R. and T. Bollerslev: Common stochastic trends in a system of exchange rates. *The Journal of Finance*, volume 44, no. 1: pp. 167–181 (1989).

Barndorff-Nielsen, O. and N. Shephard: Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, volume 64, no. 2: pp. 253–280 (2002).

Bensoussan, A.: Lectures on stochastic control. In *Nonlinear Filtering and Stochastic Control*, pp. 1–62. Springer (1982).

Benth, F. and K. Karlsen: A note on Merton's portfolio selection problem for the Schwartz mean-reversion model. *Stochastic Analysis and Applications*, volume 23: pp. 687–704 (2005).

Bichuch, M.: *Asymptotic analysis for optimal investment with transaction costs in finite time with two correlated futures contracts.* Ph.D. thesis, Carnegie Mellon University (2010).

——— Pricing a contingent claim liability with transaction costs using asymptotic analysis for optimal investment. *arXiv preprint arXiv:1112.3012* (2011).

Bichuch, M. and S. Shreve: Utility maximization trading two futures with transaction costs. *SIAM Journal on Financial Mathematics*, volume 4, no. 1: pp. 26–85 (2013).

Bielecki, T. R. and S. R. Pliska: Risk sensitive asset management with transaction costs. *Finance and Stochastics*, volume 4, no. 1: pp. 1–33 (2000).

Brenner, R. and K. Kroner: Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *Journal of Financial and Quantitative Analysis*, volume 30, no. 1: pp. 23–42 (1995).

Broadie, M., J. Cvitanic, and H. M. Soner: Optimal replication of contingent claims under portfolio constraints. *Review of Financial Studies*, volume 11, no. 1: pp. 59–79 (1998).

Carmona, R.: *Indifference Pricing: Theory and Applications.* Princeton Series in Financial Engineering. Princeton University Press (2009).

Cerchi, M. and A. Havenner: Cointegration and stock prices: the random walk on wall street revisited. *Journal of Economic Dynamics and Control*, volume 12: pp. 333–346 (1988).

Chen, S., X. Li, and X. Y. Zhou: Stochastic linear quadratic regulators with indefinite control weight costs. *SIAM Journal on Control and Optimization*, volume 36, no. 5: pp. 1685–1702 (1998).

Chiu, M. and H. Wong: Mean-variance portfolio selection of cointegrated assets. *Journal of Economic Dynamics and Control*, volume 35: pp. 1369–1385 (2011).

Choi, J., M. Sirbu, and G. Zitkovic: Shadow prices and well-posedness in the problem of optimal investment and consumption with transaction costs. *arXiv preprint arXiv:1204.0305* (2012).

Comte, F.: Discrete and continuous time cointegration. *Journal of Econometrics*, volume 88: pp. 207–226 (1999).

Crandall, M., H. Ishii, and P.-L. Lions: Uniqueness of viscosity solutions of hamilton-jacobi equations revisited. *Journal of the Mathematical Society of Japan*, volume 39, no. 4: pp. 581–596 (1987).

——— User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, volume 27, no. 1: pp. 1–67 (1992).

Cvitanić, J., H. Pham, and N. Touzi: Super-replication in stochastic volatility models under portfolio constraints. *Journal of Applied Probability*, pp. 523–545 (1999).

Dai, M. and F. Yi: Finite-horizon optimal investment with transaction costs: a parabolic double obstacle problem. *Journal of Differential Equations*, volume 246, no. 4: pp. 1445–1469 (2009).

Dai, M. and Y. Zhong: Penalty methods for continuous-time portfolio selection with proportional transaction costs. *J. Comp. Finance*, volume 13: pp. 1–31 (2010).

Davis, M.: Linear estimation and stochastic control. *Chapman and Hall mathematics series* (1977).

Davis, M. and A. Norman: Portfolio selection with transaction costs. *Mathematics of Operations Research*, volume 15, no. 4: pp. 676–713 (1990).

Davis, M. H. A., V. G. Panas, and T. Zariphopoulou: European option pricing with transaction costs. *SIAM Journal on Control and Optimization*, volume 31, no. 2: pp. 470–493 (1993).

Do, B. and R. Faff: Does simple pairs trading still work? *Financial Analysts Journal*, volume 66, no. 4: pp. 83–95 (2010).

Doob, J. L.: The elementary gaussian processes. *The Annals of Mathematical Statistics*, volume 15, no. 3: pp. 229–282 (1944).

Duan, J.-C. and S. Pliska: Option valuation with co-integrated asset prices. *Journal of Economic Dynamics and Control*, volume 28: pp. 727–754 (2004).

Dumas, B. and E. Luciano: An exact solution to a dynamic portfolio choice problem under transactions costs. *The Journal of Finance*, volume 46, no. 2: pp. 577–595 (1991).

Ehrman, D.: *The handbook of pairs trading*. John Wiley and Sons (2006).

Engelberg, J. P., P. Gao, and R. Jagannathan: An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity (2009). Working paper, University of North Carolina at Chapel Hill, University of Notre Dame, and Northwestern University (February).

Engle, R. and C. Granger: Co-integration and error correction, representation, estimation and testing. *Econometrica*, volume 55: pp. 251–276 (1987).

Ferson, W. E. and A. F. Siegel: The efficient use of conditioning information in portfolios. *The Journal of Finance*, volume 56, no. 3: pp. 967–982 (2001).

Fleming, W. H. and H. M. Soner: *Controlled Markov processes and viscosity solutions*, volume 25. springer New York (2006).

Fornberg, B.: Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of computation*, volume 51, no. 184: pp. 699–706 (1988).

Forsyth, P. and G. Labahn: Numerical methods for controlled hamilton-jacobi-bellman pdes in finance. *Journal of Computational Finance*, volume 11, no. 2: p. 1 (2007).

Forsyth, P. and K. R. Vetzal: Quadratic convergence of a penalty method for valuing american options. *SIAM Journal on Scientific Computation*, volume 23: pp. 2096–2123 (2002).

Freiling, G., G. Jank, and H. Abou-Kandil: Generalized riccati difference and differential equations. *Linear algebra and its applications*, volume 241: pp. 291–303 (1996).

Galenko, A., E. Popova, and I. Popova: Trading in the presence of cointegration. *The Journal of Alternative Investments*, volume 15, no. 1: pp. 85–97 (2012).

Gantmacher, F.: The theory of matrices, vol. i. *Chelsea, New York*, volume 220 (1959).

Gatev, E. W., W. Goetzmann, and K. Rouwenhorst: Performance of a relative value arbitrage rule (1999). Working paper, Yale School of Management.

——— Performance of a relative–value arbitrage rule. *Review of Financial Studies*, volume 19, no. 3: pp. 797–827 (2006).

Gennotte, G. and A. Jung: Investment strategies under transaction costs: the finite horizon case. *Management Science*, volume 40, no. 3: pp. 385–404 (1994).

Goll, T. and J. Kallsen: A complete explicit solution to the log-optimal portfolio problem. *The Annals of Applied Probability*, volume 13, no. 2: pp. 774–799 (2003).

Granger, C. W. J.: Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, volume 16: pp. 121–130 (1981).

Guasoni, P. and J. Muhle-Karbe: Portfolio choice with transaction costs: A users guide. In *Paris-Princeton Lectures on Mathematical Finance 2013*, volume 2081 of *Lecture Notes in Mathematics*, pp. 169–201. Springer International Publishing (2013). ISBN 978-3-319-00412-9.

Hall, A., H. Anderson, and C. Granger: A cointegration analysis of treasury bill yields. *Review of Economics and Statistics*, volume 74: pp. 116–126 (1992).

Hamilton, J.: *Time Series Analysis*. Princeton University Press (1994).

Jang, B.-G., H. Keun Koo, H. Liu, and M. Loewenstein: Liquidity premia and transaction costs. *The Journal of Finance*, volume 62, no. 5: pp. 2329–2366 (2007).

Janson, S. and J. Tysk: Feynman–Kac formulas for Black–Scholes–type operators. *Bulletin of the London Mathematical Society*, volume 38, no. 2: pp. 269–282 (2006).

Johansen, S.: *Likelihood-based inference in cointegrated vector autoregressive models*. Advanced Texts in Econometrics. Oxford University Press (1995).

Jurek, J. W. and L. M. Viceira: Optimal value and growth tilts in long-horizon portfolios. *Review of Finance*, volume 15, no. 1: pp. 29–74 (2011).

Jurek, J. W. and H. Yang: Dynamic portfolio selection in arbitrage (2007). Working paper.

Juselius, K.: *The cointegrated VAR model: methodology and applications*. Advanced Texts in Econometrics. Oxford University Press (2006).

Kallsen, J. and J. Muhle-Karbe: On using shadow prices in portfolio optimization with transaction costs. *The Annals of Applied Probability*, volume 20, no. 4: pp. 1341–1358 (2010).

Karatzas, I. and C. Kardaras: The numéraire portfolio in semimartingale financial models. *Finance Stoch.*, volume 11: pp. 447–493 (2007).

Karatzas, I. and J. Ruf: Distribution of the time to explosion for one-dimensional diffusions. *arXiv preprint arXiv:1303.5899* (2013).

Karatzas, I. and S. Shreve: *Brownian motion and stochastic calculus*. Springer-Verlag (1991).

Kessler, M. and A. Rahbek: Asymptotic likelihood based inference for co–integrated homogenous gaussian diffusions. *Scandinavian Journal of Statistics*, volume 28, no. 3: pp. 455–470 (2001).

——— Identification and inference for multivariate cointegrated and ergodic gaussian diffusions. *Statistical Inference for Stochastic Processes*, volume 7: pp. 137–151 (2004).

Khandani, A. and A. Lo: What happened to the quants in August 2007? *Journal of Investment Management*, volume 5, no. 4: pp. 5–54 (2007).

Kim, S. and E. Omberg: Dynamic nonmyopic portfolio behavior. *The Review of Financial Studies*, volume 9, no. 1: pp. 141–161 (1996).

Korn, R.: Portfolio optimisation with strictly positive transaction costs and impulse control. *Finance and Stochastics*, volume 2, no. 2: pp. 85–114 (1998).

Korn, R. and H. Kraft: On the stability of continuous-time portfolio problems with stochastic opportunity set. *Mathematical Finance*, volume 14, no. 3: pp. 403–414 (2004).

Kramkov, D. and W. Schachermayer: The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *The Annals of Applied Probability*, volume 9, no. 3: pp. 904–950 (1999).

Kushner, H. and P. Dupuis: *Numerical Methods for Stochastic Control Problems in Continuous Time*. Applications of mathematics. Springer (2001).

Lancaster, P. and L. Rodman: *Algebraic riccati equations*. Oxford University Press (1995).

Lions, P.-L.: Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations part 2: viscosity solutions and uniqueness. *Communications in partial differential equations*, volume 8, no. 11: pp. 1229–1276 (1983).

Liu, H. and M. Loewenstein: Optimal portfolio selection with transaction costs and finite horizons. *Review of Financial Studies*, volume 15, no. 3: pp. 805–835 (2002).

Liu, J.: Portfolio selection in stochastic environments. *The Review of Financial Studies*, volume 20, no. 1: pp. 1–39 (2007).

Liu, J. and F. A. Longstaff: Losing money on arbitrage: optimal dynamic portfolio choice in markets with arbitrage opportunities. *The Review of Financial Studies*, volume 17, no. 3: pp. 611–641 (2004).

Liu, J. and A. Timmermann: Optimal convergence trade strategies. *Review of Financial Studies*, volume 26, no. 4: pp. 1048–1086 (2013).

MacDonald, R. and M. Taylor: Metals prices, efficiency and cointegration: some evidence from the london metal exchange. *Bulletin of Economic Research*, volume 40: pp. 235–239 (1988).

Magill, M. and G. Constantinides: Portfolio selection with transaction costs. *Journal of Economic Theory*, volume 13, no. 2: pp. 245–263 (1976).

Markowitz, H.: Portfolio selection. *The Journal of Finance*, volume 7, no. 1: pp. 77–91 (1952).

Martin, C.: Finite escape time for riccati differential equations. *Systems & Control Letters*, volume 1, no. 2: pp. 127–131 (1981).

Martin, R. and T. Schoneborn: Mean reversion pays, but costs. *Risk Magazine*, volume 24, no. 2 , Febraury: pp. 84–89 (2011).

Martin, R. and D. Zou: Momentum trading: skews me. *Risk Magazine*, volume 25, no. 8, August (2012).

Morton, A. J. and S. R. Pliska: Optimal portfolio management with fixed transaction costs. *Mathematical Finance*, volume 5, no. 4: pp. 337–356 (1995).

Mossin, J.: Optimal multiperiod portfolio policies. *The Journal of Business*, volume 41, no. 2: pp. 215–229 (1968).

Mudchanatongsuk, S., J. Primbs, and W. Wong: Optimal pairs trading: a stochastic control approach. In *Proceedings of the American Control Conference*, pp. 1035–1039. Seattle, Washington (2008).

Muthuraman, K. and S. Kumar: Multidimensional portfolio optimization with proportional transaction costs. *Mathematical Finance*, volume 16, no. 2: pp. 301–335 (2006).

Nutz, M.: *Optimal Consumption and Investment with Power Utility*. Ph.D. thesis, ETH Zurich (2010). Diss. ETH No. 19272.

Papadakis, G. and P. Wysocki: Pairs trading and accounting information (2007). Working paper, Boston University School of Management and MIT Sloan School of Management.

Pham, H.: Smooth solutions to optimal investment models with stochastic volatilities and portfolio constraints. *Applied Mathematics and Optimization*, volume 46, no. 1: pp. 55–78 (2002).

——— On some recent aspects of stochastic control and their applications. *Probability Surveys*, volume 2: pp. 506–549 (2005).

Phillips, P. and S. Ouliaris: Asymptotic properties of residual based tests for cointegration. *Econometrica*, volume 58: pp. 165–193 (1990).

Phillips, P. C. B.: Error correction and long–run equilibrium in continuous time. *Econometrica*, volume 59, no. 4: pp. 967–980 (1991).

Possamaï, D., H. M. Soner, and N. Touzi: Homogenization and asymptotics for small transaction costs: the multidimensional case. *arXiv preprint arXiv:1212.6275* (2012).

Rockafellar, R.: *Convex Analysis*. Princeton mathematical series. Princeton University Press (1970).

Sasagawa, T.: On the finite escape phenomena for matrix riccati equations. *Automatic Control, IEEE Transactions on*, volume 27, no. 4: pp. 977–979 (1982).

Schwartz, E.: The stochastic behaviour of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, volume 52, no. 3: pp. 923–973 (1997).

Shreve, S. E. and H. M. Soner: Optimal investment and consumption with transaction costs. *The Annals of Applied Probability*, pp. 609–692 (1994).

Soner, H. M. and N. Touzi: Stochastic target problems, dynamic programming, and viscosity solutions. *SIAM Journal on Control and Optimization*, volume 41, no. 2: pp. 404–424 (2002).

——— Homogenization and asymptotics for small transaction costs. *Swiss Finance Institute Research Paper*, , no. 12-13 (2012).

Taksar, M., M. Klass, and D. Assaf: A diffusion model for optimal portfolio selection in the presence of brokerage fees. *Mathematics of Operations Research*, volume 13, no. 2: pp. 277–294 (1988).

Tavella, D. and C. Randall: *Pricing Financial Instruments: The Finite Difference Method*. Wiley Series in Financial Engineering. Wiley (2000).

Taylor, M. and I. Tonks: The internationalisation of stock markets and the abolition of u.k. exchange control. *The Review of Economics and Statistics*, volume 71, no. 2: pp. 332–336 (1989).

Vidyamurthy, G.: *Pairs trading: quantitative methods and analysis*. John Wiley and Sons (2004).

Wachter, J. A.: Portfolio and consumption decisions under mean-reverting returns: an exact solution for complete markets. *The Journal of Financial and Quantitative Analysis*, volume 37, no. 1: pp. 63–91 (2002).

Whistler, M.: *Trading pairs: capturing profits and hedging risk with statistical arbitrage strategies*. John Wiley and Sons (2004).

Wilmott, P.: *Derivatives: the theory and practice of financial engineering*. Wiley Frontiers in Finance Series. J. Wiley (1998).

Windcliff, H., P. A. Forsyth, and K. R. Vetzal: Analysis of the stability of the linear boundary condition for the black-scholes equation. *Journal of Computational Finance*, volume 8: pp. 65–92 (2004).

Wonham, W. M.: On a matrix riccati equation of stochastic control. *SIAM Journal on Control*, volume 6, no. 4: pp. 681–697 (1968).

Xiong, W.: Convergence trading with wealth effects: An amplification mechanism in financial markets. *Journal of Financial Economics*, volume 62: pp. 247–292 (2001).

Zariphopoulou, T.: Transaction costs in portfolio management and derivative pricing. In *Proceedings of symposia in applied mathematics*, volume 57, pp. 101–164 (1999).