Dynamic Optimal Portfolios for Multiple Co-Integrated Assets

T. N. Lǐ* and A. Papanicolaou†

August 7, 2019

Abstract

In this paper we construct and analyse a multi-asset model to be used for long-term statistical arbitrage strategies. A key feature of the model is that all assets have *co-integration*, which, if sustained, allows for long-term positive profits with low probability of losses. Optimal portfolios are found by solving a Hamilton-Jacobi-Bellman equation, to which we can introduce portfolio constraints such as market neutral or dollar neutral. Under specific conditions of the parameters, we can prove there is long-term stability for an optimal portfolio with stable growth rate. Historical prices of the S&P500 constituents can be tested for co-integration and our model calibrated for analysis, from which we find that co-integration strategies require a terminal investment horizon sufficiently far into the future in order for the optimal portfolios to gain from co-integration. The data also demonstrates that statistical arbitrage portfolios will have improved in-sample Sharpe ratios compared to multivariate Merton portfolios, and that statistical arbitrage portfolios are naturally immune to market fluctuations.

Keywords: co-integrated assets, market-neutral portfolios, statistical arbitrage, stochastic optimal control, matrix Riccati equations.

AMS Subject Codes: 62P05, 91B28, 93E20.

1 Introduction

Statistical arbitrage strategies can be constructed by trading among pairs of assets having co-integrated prices. The essential idea is that a pair of co-integrated prices will have a difference that is mean reverting. This mean-reverting difference is referred to as a *spread*. For an arbitrageur with the financial means, a possible strategy is to long the cheaper asset, short the expensive asset, and then to wait for the spread to converge, at which point the position can be closed for a profit. This is an example of *statistical arbitrage* because, while it may seem like a sure profit, there is no finite time by when a spread will have almost-surely converged. Instead there is only a high probability of the spread converging before reaching a fixed, finite investment horizon.

In this paper we construct a model for trading of multiple assets where each asset is co-integrated with respect to a common benchmark index. Co-integration is found utilising the method of [9] where an asset's log price is regressed against the log of the benchmark, and co-integration conclusively identified if a unit-root hypothesis is rejected for the residual time series. As assumed in [19, 20], these residuals form a stationary Ornstein-Uhlenbeck (OU) process that henceforth is

*Department of Mathematics, New York University, 251 Mercer Street, New York, NY, 10012. *thomli@cims.nyu.edu*

†Department of Finance and Risk Engineering, New York University, 1 MetroTech Center, Brooklyn NY, 11201. *ap1345@nyu.edu*

referred to as the spread. Our construction is similar to [4] wherein the spreads are relative to a set of factors given by the sector exchange-traded funds (ETFs). However, ETFs themselves have co-integration, and so we would like our model to be able to statistically arbitrage spreads between ETFs that track the same index. Hence we take the common benchmark index to be a non-tradeable target, and by doing so we take into account (1) the ETF tracking-error spreads and any mean reversion that they may have, and (2) any co-integrated pair trades among the multiple ETFs tracking the common target. For example, there is substantial tracking error among ETFs in commodities, high-yield bonds, municipal bonds and foreign indices, with two or more ETFs tracking a common index in each sector.

The model we propose is an one-factor version of the multiple-pairs statistical arbitrage model developed in [4]. We formulate a stochastic control and optimisation problem with the set of state variables comprised of the wealth process W_t, and the spreads $Z_t^i = a_i + b_i t + \log\left(S_t^0\right) + \beta_i \log\left(S_t^i\right)$ for $i = 1,\ 2,\ \cdots,\ d$, where S_t^0 is the benchmark index, S_t^i is the price of the i^{th} traded assets, and $(a_i,\ b_i,\ \beta_i)$ are the regression coefficients returned by the Engle-Granger test making Z_t^i stationary. The optimal portfolio is the solution to a Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE), which in the case of power utility we are able to reduce to a system of ordinary differential equations (ODEs) that includes a matrix Riccati equation and a pair of linear equations. We perform long-term stability analysis of these ODEs, which gives us an indication of the model's soundness for financial application. In particular, finiteness of the ODEs for any finite-time investment horizon indicates that the model has an absence of arbitrage; if there was an arbitrage then the ODEs would have a singularity. In addition, if the solution of the HJB converges to an equilibrium as the investment horizon tends toward infinity, then there is a long-term statistical arbitrage portfolio that earns positive profits with probability close to one. Our analysis shows that how optimal strategies depend on the co-integration spreads, mean-reversion speeds, risk aversion, and different portfolio constraints.

We also perform empirical analysis based on historical financial market data. The purpose of these studies is to show that indeed, market data agree with our theoretical conditions for long-term stability. These studies also give us a sense of the effectiveness of implementing the optimal strategies of these models. For our studies we select the subset of S&P500 constituents that are co-integrated with the S&P500 Index itself during the period of 2012-01-01 and 2019-05-15, and then look at the long-term statistics of a portfolio of the 10 stocks whose spreads have the fastest rate of mean reversion. We find 106 co-integrated constituents, for which we then compute mean and covariance matrices of their returns processes, and then allocating among the top 10 fastest mean reverters we look at long-term Sharpe ratio for a few different optimal strategies. These studies show (1) that statistical arbitrage strategies require a sufficiently-long investment horizon in order for the optimal portfolios to have a high probability of gaining from co-integration, and (2) that there is in-sample out-performance of co-integration-based portfolios in comparison to trading strategies based on constant-parameters with no consideration given to spreads (in other words, a multivariate Merton portfolio), and that statistical arbitrage portfolios have some natural immunity to market fluctuations such as the Chinese currency devaluation of August 2015. We stress that these data findings are in-sample tests, and merely serve to provide intuition and analysis for statistical arbitrage opportunities; in-sample tests suffer from over-fitting and there are major hurdles to implement out-of-sample tests because portfolio performance is significantly effected by parameter-estimation error.

1.1 Overview of Related Work

A test for co-integration of financial time series was designed in [9], namely, the Engle-Granger test. An application of [9] and some basic examples of co-integration-based trading rules were shown in [21], trading of co-integrated pairs alongside methods for filtering and parameter estimation to handle latency was studied in [8], and an in-depth statistical analysis of the performance of pairs trading strategies was done in [13]. Principal component analysis of large number of assets co-integrated through common factors was the topic in [4], and is the basis for the model in this paper; empirical testing of pairs trading, including out-of-sample experiments with changing parameters, was completed in [12]. Analysis showing significance of short-term reversal and momentum factors on returns of pairs trading portfolios was finished in [5].

The literature on stochastic control and optimisation for co-integration models also has some depth. Stochastic optimal control for pairs trading with OU spreads was done in [19]. A stochastic control for optimal trading of co-integrated pairs is posed and solved in [3, 20], and stochastic optimal control for pairs trading with a local-volatility model was analysed by [18]. Optimal trading of spreads with transaction costs and stop-loss criterion were solved and analysed in [17, 16]. Related to the stability analysis presented in this paper, there was the long-term stability analysis for matrix Riccati equations of multi-asset models done in [7], and the matrix Riccati equations analysis for a single co-integrated pair done in [15]. There were also machine learning approaches taken to statistical arbitrage, such as reinforcement learning and boosting applied to co-integrated constituents in the S&P500 were finished in [10]. An HJB equation for an optimal portfolio constrained to be 100% long was completed in [2] with application toward comparing active and passive fund management.

1.2 Structure of this Paper

In this paper, we presents and solve the stochastic control and optimisation problems and then analyse the solutions – both for the cases of unconstrained and constrained portfolios. Later on in the paper we perform some empirical studies on historical data. The organisation of the paper is as follows: Section 2 contains the definitions for the model along with analysis of the HJB equations, with Section 2.2 presenting the solution of the HJB for the unconstrained portfolio via an exponential-affine ansatz, with Section 2.3 presenting the stability analysis, and then with Section 2.4 presenting the HJB and stability analysis for optimisation with market neutral and dollar-neutral constraints; the empirical analysis of historical data comes in Section 3 with preliminary data analysis and parameter estimation presented in Section 3.1 and analysis of portfolio performance (e.g., Sharpe ratios) presented in Section 3.2; Section 4 is the conclusion.

2 Model Construction & Optimisation for Multiple Co-Integrated Assets

This section introduces the stochastic model for multiple co-integrated assets, derives the HJB equations for optimal portfolios, and shows the stability analysis for the solutions.

2.1 Model with Co-Integration

Let $\left(\Omega,\, \mathcal{F},\, (\mathcal{F}_t)_{t\geq 0}\,,\, \mathbb{P}\right)$ be a filtered probability space. Suppose S_t^0 is a benchmark for the financial market or an industrial sector with no uncertainty in expected returns. The stochastic differential equation for its dynamic is

$$d\log S_t^0 = \left(\mu_0 - \frac{\sigma_0^2}{2}\right) dt + \sigma_0 dB_t^0 \ , \tag{2.1}$$

where μ_0 and $\sigma_0 > 0$ are constants, B_t^0 is an one dimensional Standard Brownian Motion on the filtered probability space. Suppose S_t^i, $i = 1,\ 2,\ \cdots,\ d \in \mathbb{N}$, is the price of an individual firm in the financial market or the industrial sector, whose dynamic is modelled by the stochastic differential equation

$$d\log S_t^i = \left(\mu_i - \frac{\sigma_i^2}{2} + \delta_i Z_t^i\right) dt + \sigma_i dB_t^i \ , \tag{2.2}$$

where μ_i, δ_i, and $\sigma_i > 0$ are constants, B_t^i is an element of the vector $\boldsymbol{B}_t = \left[B_t^1,\ B_t^2,\ \cdots,\ B_t^d\right]^\top \in \mathbb{R}^d$, which is a d-dimensional Standard Brownian Motions with correlation coefficient $\rho_{ij} \in (-1,\ 1)$ for $i,\ j = 1,\ 2,\ \cdots,\ d$ such that $dB_t^i dB_t^j = \rho_{ij} dt$, and where $\left[B_t^0,\ B_t^i\right]$ is a 2-dimensional Standard Brownian Motions with correlation coefficient $\rho_i^0 \in (-1,\ 1)$ such that $dB_t^i dB_t^0 = \rho_i^0 dt$ for $i = 1,\ 2,\ \cdots,\ d$. Furthermore, the co-integrated process Z_t^i that appears in equation (2.2) is spread given by

$$Z_t^i = a_i + b_i t + \log S_t^0 + \beta_i \log S_t^i \ , \tag{2.3}$$

where typically $\beta_i < 0$ as equities generally move in the same direction over longer time periods (see [9]). Utilising equations (2.1) and (2.2), the dynamic of the co-integrated process Z_t^i is the differential of (2.3)

$$\begin{aligned} dZ_t^i &= \left(b_i + \mu_0 - \frac{\sigma_0^2}{2} + \beta_i\mu_i - \beta_i\frac{\sigma_i^2}{2} + \beta_i\delta_i Z_t^i\right) dt + \sigma_0 dB_t^0 + \beta_i\sigma_i dB_t^i \\ &= -\beta_i\delta_i \left(-\frac{b_i + \mu_0 + \beta_i\mu_i - \frac{\sigma_0^2}{2} - \beta_i\frac{\sigma_i^2}{2}}{\beta_i\delta_i} - Z_t^i\right) dt + \sigma_0 dB_t^0 + \beta_i\sigma_i dB_t^i \ . \end{aligned} \tag{2.4}$$

We assume that each element Z_t^i of the vector $\boldsymbol{Z}_t = \left[Z_t^1,\ Z_t^2, \cdots,\ Z_t^d\right]^\top \in \mathbb{R}^d$ is a one dimensional OU process, which holds true if $-\beta_i\delta_i > 0$, in other words if $\delta_i > 0$. We also assume that there is a risk-free asset, such as a money market account, with interest rate $r > 0$.

Proposition 2.1. *Given correlation matrix $\boldsymbol{\rho} = [\rho_{ij}] \in \mathbb{R}^{d\times d}$, $i,\ j = 1,\ 2,\ \cdots,\ d$, and correlation column vector $\vec{\boldsymbol{\rho}}_0 = \left[\rho_1^0,\ \rho_2^0,\ \cdots,\ \rho_d^0\right]^\top \in \mathbb{R}^d$, in order for them to combine to form a $(d+1)$-dimensional correlation structure, they must satisfy the condition of*

$$\vec{\boldsymbol{\rho}}_0^\top \boldsymbol{\rho}^{-1} \vec{\boldsymbol{\rho}}_0 \leq 1 \ .$$

Proof. Consider a linear representation of B_t^0

$$B_t^0 = \omega_0 \tilde{B}_t^0 + \sum_{i=1}^d \omega_i B_t^i \ ,$$

where $\omega_0 \in \mathbb{R}$ and $\omega_i \in \mathbb{R}$, $\tilde{B}_t^0$ and B_t^i are two independent Standard Brownian Motions. The quadratic variation of B_t^0 is

$$\begin{aligned} dt &= dB_t^0 dB_t^0 \\ &= \left(\omega_0 d\tilde{B}_t^0 + \sum_{i=1}^d \omega_i dB_t^i\right)^2 \\ &= \left(\omega_0^2 + \boldsymbol{\omega}^\top \boldsymbol{\rho}\boldsymbol{\omega}\right) dt \ , \end{aligned}$$

where $\boldsymbol{\omega} = [\omega_1,\ \omega_2, \dots,\ \omega_d]^\top \in \mathbb{R}^d$*. Thus, it must be that* $0 \le \boldsymbol{\omega}^\top \boldsymbol{\rho}\boldsymbol{\omega} = 1 - \omega_0^2 \le 1$*; and the cross variation of* B_t^0 *and* B_t^i *is*

$$\begin{aligned}\rho_i^0 dt &= dB_t^i dB_t^0 \\ &= dB_t^i \left(\omega_0 d\tilde{B}_t^0 + \sum_{j=1}^d \omega_j dB_t^j \right) \\ &= \sum_{j=1}^d \omega_j \rho_{ij} dt \ .\end{aligned}$$

Hence, $\boldsymbol{\rho}_0 = \boldsymbol{\rho}\boldsymbol{\omega}$*. Therefore,*

$$\begin{aligned}\vec{\boldsymbol{\rho}}_0^\top \boldsymbol{\rho}^{-1} \vec{\boldsymbol{\rho}}_0 &= (\boldsymbol{\rho}\boldsymbol{\omega})^\top \boldsymbol{\rho}^{-1} (\boldsymbol{\rho}\boldsymbol{\omega}) \\ &= \boldsymbol{\omega}^\top \boldsymbol{\rho}\boldsymbol{\omega} \\ &= 1 - \omega_0^2 \le 1 \ .\end{aligned}$$

This inequality describes the relationship between the correlations within the assets and the correlations between the benchmark and assets. □

Example 2.1. *Suppose that* $\boldsymbol{\rho} = \mathbf{I}$*, where* $\mathbf{I} \in \mathbb{R}^{d\times d}$ *is an identity matrix. Then it must be that* $\|\vec{\boldsymbol{\rho}}_0\| \le 1$ *to ensure* $\vec{\boldsymbol{\rho}}_0^\top \boldsymbol{\rho}^{-1} \vec{\boldsymbol{\rho}}_0 \le 1$*.*

Example 2.2. *Consider the d-dimensional case in which the correlation coefficients* $\vec{\boldsymbol{\rho}}_0$ *and* $\boldsymbol{\rho}$ *have the following structure*

$$\begin{aligned}\boldsymbol{\rho} &= (1-c)\,\mathbf{I} + c\mathbf{1}\mathbf{1}^\top \\ \vec{\boldsymbol{\rho}}_0 &= c_0 \mathbf{1} \ ,\end{aligned}$$

where $\mathbf{1} = [1,\ 1, \cdots,\ 1]^\top \in \mathbb{R}^d$*,* $c_0 \in \mathbb{R}$*, and* $c \in \mathbb{R}$ *with* $0 < c < 1$*. From the Sherman-Woodbury-Morrison formula, we have*

$$\boldsymbol{\rho}^{-1} = \frac{1}{1-c}\left(\mathbf{I} - \frac{c}{1+(d-1)\,c} \mathbf{1}\mathbf{1}^\top \right).$$

Therefore,

$$\begin{aligned}\vec{\boldsymbol{\rho}}_0^\top \boldsymbol{\rho}^{-1} \vec{\boldsymbol{\rho}}_0 &= c_0^2 \mathbf{1}^\top \boldsymbol{\rho}^{-1} \mathbf{1} \\ &= \frac{c_0^2}{1+(d-1)\,c} \le 1 \ ,\end{aligned}$$

which is the case if $c_0^2 \le \left(1 - \frac{1}{n}\right) c + \frac{1}{n}$*.*

The wealth process W_t of an arbitrageur is constructed in the following way. We denote by π_t^i the fractions of wealth invested in the i^{th} risky asset at time t, then $\left(1 - \sum_{i=1}^d \pi_t^i\right)$ is the fractions of wealth invested in the risk-free asset. Therefore, consider a wealth process that is among the class of self-financing strategies and hence, its evolution is given by

$$dW_t = \sum_{i=1}^d \pi_t^i W_t \frac{dS_t^i}{S_t^i} + r\left(1 - \sum_{i=1}^d \pi_t^i\right) W_t dt \ . \tag{2.5}$$

Please note that the benchmark S_t^0 is not contained in the portfolio, as you can observe from the equation (2.5).

Next, we define the value function $u(t,\ w,\ \mathbf{z})$ for any $(t, w, \mathbf{z}) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^d$ where $T < \infty$ is a terminal time. The arbitrageur seeks a control $\boldsymbol{\pi} : [0,\ T] \times \Omega \to \mathbb{R}^d$ where at each time the portfolio allocation is $\boldsymbol{\pi}_t = \left[\pi_t^1,\ \pi_t^2,\ \cdots,\ \pi_t^d\right]^\top \in \mathbb{R}^d$, to maximise the expectation of terminal utility,

$$u(t,\ w,\ \mathbf{z}) = \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}\left[U(W_T)\,\middle|\,W_t = w,\ \boldsymbol{Z}_t = \mathbf{z}\right]\ , \tag{2.6}$$

where w and $\mathbf{z} = [z_1,\ \cdots,\ z_d]^\top \in \mathbb{R}^d$ are the state variables, and $\boldsymbol{\pi}$ is selected from a set of admissible controls $\mathcal{A}$ defined as

$$\mathcal{A} = \left\{ \boldsymbol{\pi} : [0,\ T] \times \Omega \to \mathbb{R}^d \,\middle|\, \boldsymbol{\pi}_t \in \mathcal{F}_t\ ,\ \text{and} \quad \int_0^T \|\boldsymbol{\pi}_t W_t\|^2\, dt < \infty \ \text{ a.s.} \right\}. \tag{2.7}$$

For a mathematical primer on the theory of stochastic control and optimisation, we refer the reader to chapter 4 of [11].

We assume that the arbitrageur has a concave utility function $U(x)$ over power type:

$$U(w) = \frac{1}{\gamma} w^\gamma, \tag{2.8}$$

where $\gamma < 0$ is called the risk aversion parameter, which is a coefficient that is from the utility theory of economics. Risk aversion coefficient is utilised to measure risk preference of an arbitrageur, if γ approaches to zero, it indicates that the arbitrageur is more risk loving, and if γ tends to $-\infty$, it represents that the arbitrageur is more risk averse.

2.2 Hamilton-Jacobi-Bellman Equation

In order to have convenience in mathematical derivations for the rest sections of this paper, we introduce some notations initially. For $i,\ j = 1,\ 2,\ \cdots,\ d \in \mathbb{N}$, we denote the following vectors and matrices:

$$\begin{aligned}
&\vec{\boldsymbol{\mu}} = [\mu_i] \in \mathbb{R}^d, \qquad \boldsymbol{\mu} = [\mu_i - r] \in \mathbb{R}^d,\\
&\vec{\boldsymbol{\sigma}} = [\sigma_i] \in \mathbb{R}^d, \qquad \boldsymbol{\sigma} = \operatorname{diag}(\vec{\boldsymbol{\sigma}}) \in \mathbb{R}^{d\times d},\\
&\vec{\boldsymbol{\beta}} = [\beta_i] \in \mathbb{R}^d, \qquad \boldsymbol{\beta} = \operatorname{diag}\left(\vec{\boldsymbol{\beta}}\right) \in \mathbb{R}^{d\times d},\\
&\vec{\boldsymbol{\delta}} = [\delta_i] \in \mathbb{R}^d, \qquad \boldsymbol{\delta} = \operatorname{diag}\left(\vec{\boldsymbol{\delta}}\right) \in \mathbb{R}^{d\times d},\\
&\vec{\boldsymbol{\kappa}} = [-\beta_i\delta_i] \in \mathbb{R}^d, \qquad \boldsymbol{\kappa} = \operatorname{diag}(\vec{\boldsymbol{\kappa}}) \in \mathbb{R}^{d\times d},\\
&\boldsymbol{b} = [b_i] \in \mathbb{R}^d, \qquad \boldsymbol{\theta} = \left[-\frac{b_i + \mu_0 + \beta_i\mu_i - \frac{1}{2}\left(\sigma_0^2 + \beta_i\sigma_i^2\right)}{\beta_i\delta_i}\right] \in \mathbb{R}^d,\\
&\boldsymbol{\Sigma}_1 = [\sigma_i\sigma_j\rho_{ij}] \in \mathbb{R}^{d\times d}, \qquad \boldsymbol{\Sigma}_2 = \left[\sigma_0\sigma_i\rho_i^0 + \sigma_i\sigma_j\beta_j\rho_{ij}\right] \in \mathbb{R}^{d\times d},\\
&\boldsymbol{\Sigma}_3 = \left[\sigma_0^2 + \sigma_0\sigma_i\beta_i\rho_i^0 + \sigma_0\sigma_j\beta_j\rho_j^0 + \sigma_i\beta_i\sigma_j\beta_j\rho_{ij}\right] \in \mathbb{R}^{d\times d}\ .
\end{aligned} \tag{2.9}$$

Please note that we assume that $\boldsymbol{\Sigma}_1$ is positive definite and invertible in this paper. We can also observe that $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_3$ are symmetric matrices, $\boldsymbol{\Sigma}_2$, however, is not. Subsequently, we apply the standard stochastic control and optimisation techniques and expect the value function u defined in (2.6) to satisfy the following HJB equation:

$$
\begin{aligned}
& -u_t-(\boldsymbol{\theta}-\mathbf{z})^{\top} \boldsymbol{\kappa} \nabla_{\mathbf{z}} u-\frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Sigma}_3 \nabla_{\mathbf{z}}^2 u\right)-r w u_w \\
& -\sup _{\boldsymbol{\pi} \in \mathbb{R}^d}\left(\boldsymbol{\pi}^{\top}(\boldsymbol{\mu}+\boldsymbol{\delta} \mathbf{z}) w u_w+\boldsymbol{\pi}^{\top} \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}}\left(\nabla_w u\right) w+\frac{1}{2} \boldsymbol{\pi}^{\top} \boldsymbol{\Sigma}_1 \boldsymbol{\pi} w^2 u_{w w}\right)=0\ , \\
& u(T,\ w,\ \mathbf{z})=\frac{1}{\gamma} w^\gamma\ .
\end{aligned} \tag{2.10}
$$

The wealth variable w can be factored out of the solution by utilising the ansatz

$$
u(t,\ w,\ \mathbf{z})=\frac{1}{\gamma} w^\gamma g(t,\ \mathbf{z}), \tag{2.11}
$$

hence, we have

$$
\begin{aligned}
& u_t=\frac{1}{\gamma} w^\gamma g_t\ , \qquad u_{w w}=(\gamma-1) w^{\gamma-2} g\ , \qquad \nabla_{\mathbf{z}} u=\frac{1}{\gamma} w^\gamma \nabla_{\mathbf{z}} g\ , \\
& u_w=w^{\gamma-1} g\ , \qquad \nabla_{\mathbf{z}}^2 u=\frac{1}{\gamma} w^\gamma \nabla_{\mathbf{z}}^2 g\ , \qquad \nabla_{\mathbf{z}}\left(\nabla_w u\right)=w^{\gamma-1} \nabla_{\mathbf{z}} g\ .
\end{aligned}
$$

Therefore, the HJB equation (2.10) can be transformed into

$$
\begin{aligned}
& -g_t-(\boldsymbol{\theta}-\mathbf{z})^{\top} \boldsymbol{\kappa} \nabla_{\mathbf{z}} g-\frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Sigma}_3 \nabla_{\mathbf{z}}^2 g\right)-r \gamma g \\
& -\inf _{\boldsymbol{\pi} \in \mathbb{R}^d}\left(\boldsymbol{\pi}^{\top}(\boldsymbol{\mu}+\boldsymbol{\delta} \mathbf{z}) \gamma g+\boldsymbol{\pi}^{\top} \gamma \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} g+\frac{1}{2} \boldsymbol{\pi}^{\top} \boldsymbol{\Sigma}_1 \boldsymbol{\pi} g \gamma(\gamma-1)\right)=0\ , \\
& g(T,\ \mathbf{z})=1\ .
\end{aligned} \tag{2.12}
$$

We compute the optimal control variable $\boldsymbol{\pi}^*$ by solving the optimisation problem that is contained inside the HJB equation (2.12) in terms of g and its partial derivatives

$$
\boldsymbol{\pi}^*=\frac{1}{1-\gamma} \boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\mu}+\boldsymbol{\delta} \mathbf{z})+\frac{1}{1-\gamma} \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}}(\log g)\ . \tag{2.13}
$$

Inserting the optimal $\boldsymbol{\pi}^*$ back into the HJB equation (2.12) results in the following nonlinear partial differential equation

$$
\begin{aligned}
& g_t+(\boldsymbol{\theta}-\mathbf{z})^{\top} \boldsymbol{\kappa} \nabla_{\mathbf{z}} g+\frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Sigma}_3 \nabla_{\mathbf{z}}^2 g\right)+r \gamma g \\
& +\frac{\gamma g}{2(1-\gamma)}\left(\boldsymbol{\mu}+\boldsymbol{\delta} \mathbf{z}+\boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}}(\log g)\right)^{\top} \boldsymbol{\Sigma}_1^{-1}\left(\boldsymbol{\mu}+\boldsymbol{\delta} \mathbf{z}+\boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}}(\log g)\right)=0\ , \\
& g(T,\ \mathbf{z})=1\ .
\end{aligned} \tag{2.14}
$$

We approach solving PDE (2.14) with an exponential-affine ansatz for g,

$$
g(t,\ \mathbf{z})=\exp \left(a(t)+\boldsymbol{b}^{\top}(t) \mathbf{z}+\mathbf{z}^{\top} \boldsymbol{C}(t) \mathbf{z}\right), \tag{2.15}
$$

where $a(t) \in \mathbb{R}$ is a scalar, $\boldsymbol{b}(t)=\left[b_i(t)\right] \in \mathbb{R}^d$ is a column vector, and $\boldsymbol{C}(t)=\left[c_{ij}(t)\right] \in \mathbb{R}^{d \times d}$, $i,\ j=1,\ 2,\ \cdots,\ d$, is a $d \times d$ symmetric matrix. By utilising the ansatz (2.15), the nonlinearity in PDE (2.14) can be removed, and the equation can be transformed into a system of ODEs.

Proposition 2.2. *The PDE* (2.14) *is solved utilising the exponential-affine an-satz of* (2.15)*, and the functions* $a(t) \in \mathbb{R}$*,* $\boldsymbol{b}(t) \in \mathbb{R}^d$*, and* $\boldsymbol{C}(t) \in \mathbb{R}^{d\times d}$ *satisfy the following system of ODEs:*

$$\begin{aligned}\frac{\partial a(t)}{\partial t} &= -\boldsymbol{b}^\top(t)\left(\frac{\gamma}{2(1-\gamma)}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 + \frac{1}{2}\boldsymbol{\Sigma}_3\right)\boldsymbol{b}(t) \\ &\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{b}^\top(t)\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu} - \frac{\gamma}{2(1-\gamma)}\boldsymbol{\mu}^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2\boldsymbol{b}(t) \\ &\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{\mu}^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu} - r\gamma - \boldsymbol{\theta}^\top\boldsymbol{\kappa}\boldsymbol{b}(t) - tr[\boldsymbol{\Sigma}_3\boldsymbol{C}(t)] \ , \\ a(T) &= 0 \ ;\end{aligned} \tag{2.16}$$

$$\begin{aligned}\frac{\partial \boldsymbol{b}(t)}{\partial t} &= -2\boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 + \boldsymbol{\Sigma}_3\right)\boldsymbol{b}(t) \\ &\quad - \left(\frac{\gamma}{1-\gamma}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 - \boldsymbol{\kappa}\right)\boldsymbol{b}(t) \\ &\quad - 2\boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu} + \boldsymbol{\kappa}\boldsymbol{\theta}\right) - \frac{\gamma}{1-\gamma}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu} \ , \\ \boldsymbol{b}(T) &= \boldsymbol{0} \ ;\end{aligned} \tag{2.17}$$

$$\begin{aligned}\frac{\partial \boldsymbol{C}(t)}{\partial t} &= -\boldsymbol{C}(t)\left(\frac{2\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 + 2\boldsymbol{\Sigma}_3\right)\boldsymbol{C}(t) \\ &\quad - \boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\delta} - \boldsymbol{\kappa}\right) \\ &\quad - \left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\delta} - \boldsymbol{\kappa}\right)^\top\boldsymbol{C}(t) \\ &\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\delta} \ , \\ \boldsymbol{C}(T) &= \boldsymbol{0} \ .\end{aligned} \tag{2.18}$$

Proof. *By inserting the exponential-affine ansatz* (2.15) *into PDE* (2.14)*, and grouping terms as either quadratic in* $\mathbf{z}$*, linear in* $\mathbf{z}$ *or constant in* $\mathbf{z}$*, then equations* (2.16)*,* (2.17) *and* (2.18) *are respectively obtained.* □

We can observe that the three ODEs (2.16), (2.17) and (2.18) with respect to $a(t)$, $\boldsymbol{b}(t)$ and $\boldsymbol{C}(t)$ are coupled. The coupling is recursive in the sense that the equation with respect to $\boldsymbol{C}(t)$ is autonomous, while we can solve for $\boldsymbol{b}(t)$ given $(\boldsymbol{C}(t))_{u\geq t}$, and solve for $a(t)$ given $(\boldsymbol{b}(t))_{u\geq t}$ and $(\boldsymbol{C}(t))_{u\geq t}$. The type of the ODE that $\boldsymbol{C}(t)$ solves is a matrix Riccati equation. The gradient of the exponential-affine ansatz is $\nabla_{\mathbf{z}}(\log g) = 2\boldsymbol{C}(t)\mathbf{z} + \boldsymbol{b}(t)$ so that the optimal $\boldsymbol{\pi}$ from equation (2.13) can be expressed as

$$\boldsymbol{\pi}^* = \frac{1}{1-\gamma}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z}) + \frac{1}{1-\gamma}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2(2\boldsymbol{C}(t)\mathbf{z} + \boldsymbol{b}(t)) \ .$$

Verification that this is indeed an optimal strategy is straight-forward under the multiple-co-integrated model that we've constructed:

Proposition 2.3 (Verification)**.** *The optimal control variable $\boldsymbol{\pi}^*$ that is given by formula* (2.13) *utilising the exponential-affine ansatz that is proposed by* (2.15)*, where $\boldsymbol{C}(t)$, $\boldsymbol{b}(t)$ and $a(t)$ are respectively the solutions of the matrix Riccati equation* (2.18)*, the ODEs* (2.17) *and* (2.16)*, belongs to the set of admissible controls $\mathcal{A}$ that is described by formula* (2.7) *and maximises the expected utility function that is defined in formula* (2.6)*.*

Proof. The form of the model fits into the framework set forth in [6] and [7], and hence the same argument for verification applies here. □

2.3 Stability Analysis

Stability analysis of the matrix Riccati equation (2.18) and the linear ODE (2.17) informs us that whether our solution to PDE (2.14) blows up or not. We extend the time domain for the ODEs (2.18) and (2.17) to $(-\infty, T]$ for any finite T, and if the solution remains finite at all time then we have a stable system from which we can draw intuition about long-term investment strategies.

Our analysis of the matrix Riccati equation $\boldsymbol{C}(t)$ will use Theorem 2.1 from [23] to show that the solution of equation (2.18) exists, is bounded and unique for all $t \leq T$. Let us rewrite the matrix Riccati equation (2.18) as

$$\begin{aligned} \frac{\partial \boldsymbol{C}(t)}{\partial t} &= -\mathbf{A}_u^\top \boldsymbol{C}(t) - \boldsymbol{C}(t)\mathbf{A}_u - \boldsymbol{C}(t)\mathbf{Q}_u\boldsymbol{C}(t) - \mathbf{P}_u \ , \\ \boldsymbol{C}(T) &= \mathbf{0} \ , \end{aligned} \tag{2.19}$$

where

$$\begin{aligned} \mathbf{Q}_u &= \frac{2\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 + 2\boldsymbol{\Sigma}_3 \ , \\ \mathbf{A}_u &= \frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\delta} - \boldsymbol{\kappa} \ , \\ \mathbf{P}_u &= \frac{\gamma}{2(1-\gamma)}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\delta} \ . \end{aligned}$$

Proposition 2.4. *For $\gamma < 0$, the coefficient matrices of the quadratic term and the zero order term of the matrix Riccati equation $\boldsymbol{C}(t)$, namely, $\mathbf{Q}_u$ and $\mathbf{P}_u$ in equation* (2.19) *are symmetric positive definite and symmetric negative definite, respectively.*

Proof. From formula (2.9)*, we have*

$$\begin{aligned} \boldsymbol{\Sigma}_2 &= \sigma_0\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \boldsymbol{\Sigma}_1\boldsymbol{\beta} \ , \\ \boldsymbol{\Sigma}_3 &= \sigma_0^2\mathbf{1}\mathbf{1}^\top + \sigma_0\left(\boldsymbol{\beta}\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \left(\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top\right)^\top\boldsymbol{\beta}\boldsymbol{\sigma}\right) + \boldsymbol{\beta}\boldsymbol{\Sigma}_1\boldsymbol{\beta} \ . \end{aligned} \tag{2.20}$$

Hence, the coefficient matrix $\mathbf{Q}_u$ of the quadratic term of the matrix Riccati equation (2.19) *has the following decomposition,*

$$\begin{aligned} \mathbf{Q}_u &= \frac{2\gamma}{1-\gamma}\left(\sigma_0\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \boldsymbol{\Sigma}_1\boldsymbol{\beta}\right)^\top\boldsymbol{\Sigma}_1^{-1}\left(\sigma_0\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \boldsymbol{\Sigma}_1\boldsymbol{\beta}\right) \\ &\quad + 2\sigma_0\mathbf{1}\mathbf{1}^\top + 2\sigma_0\left(\boldsymbol{\beta}\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \left(\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top\right)^\top\boldsymbol{\beta}\boldsymbol{\sigma}\right) \\ &= \frac{2}{1-\gamma}\boldsymbol{\Sigma}_3 - \frac{2\gamma\sigma_0^2}{1-\gamma}\left(\mathbf{1}\mathbf{1}^\top - \frac{\gamma}{1-\gamma}\boldsymbol{\rho}_0^\top\boldsymbol{\rho}^{-1}\boldsymbol{\rho}_0\right) \ , \end{aligned} \tag{2.21}$$

where $\boldsymbol{\rho}_0 = \vec{\boldsymbol{\rho}}_0 \mathbf{1}^\top \in \mathbb{R}^{d\times d}$. The matrix $\boldsymbol{\Sigma}_3$ is symmetric positive definite by construction because it is the diffusion matrix of the OU processes (2.4)*, and it is straightforward to check that matrix $\mathbf{1}\mathbf{1}^\top - \frac{\gamma}{1-\gamma}\boldsymbol{\rho}_0^\top \boldsymbol{\rho}^{-1}\boldsymbol{\rho}_0$ is symmetric positive semidefinite for $\gamma < 0$. Hence, the third line of equation* (2.21) *is a symmetric positive definite matrix, and it follows that the coefficient matrix of the quadratic term of the matrix Riccati equation* (2.19) *is symmetric positive definite.*

Proving that $\mathbf{P}_u$ is symmetric negative definite is uncomplicated. We can observe that matrix $\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\delta}$ is symmetric positive definite. Consequently, for $\gamma < 0$, $-\mathbf{P}_u$ is symmetric positive definite, in other words $\mathbf{P}_u$ is symmetric negative definite. □

Given Proposition 2.4, the stability analysis from [23] applies directly, but first it will be useful to define the following properties:

Definition 2.1 (Controllability)**.** *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$ be constant matrices. The controllability matrix of $(\mathbf{A},\, \mathbf{B})$ is the $n \times mn$ matrix*

$$\boldsymbol{\Gamma}(\mathbf{A},\, \mathbf{B}) = \left[\mathbf{B},\, \mathbf{A}\mathbf{B},\, \cdots,\, \mathbf{A}^{n-1}\mathbf{B}\right].$$

The pair $(\mathbf{A},\, \mathbf{B})$ is controllable if the rank of $\boldsymbol{\Gamma}$ is n. If $(\mathbf{A},\, \mathbf{B})$ is controllable, so is $(\mathbf{A} - \mathbf{B}\mathbf{M},\, \mathbf{B})$ for every matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$.

Definition 2.2 (Observability)**.** *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{E} \in \mathbb{R}^{p\times n}$ be constant matrices. The pair $(\mathbf{E},\, \mathbf{A})$ is observable if the pair $\left(\mathbf{A}^\top,\, \mathbf{E}^\top\right)$ is controllable.*

Definition 2.3 (Stabilizability)**.** *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$ be constant matrices. The pair $(\mathbf{A},\, \mathbf{B})$ is stabilizable if there exists a constant matrix $\mathbf{M}$ such that all the eigenvalues of $\mathbf{A} - \mathbf{B}\mathbf{M}$ have negative real parts.*

With these definitions we have the following proposition for the matrix Riccati equation (2.18).

Proposition 2.5. *For $\gamma < 0$, the coefficient matrix $\mathbf{Q}_u$ of the quadratic term in the matrix Riccati equation* (2.19) *is symmetric positive definite. Consequently, there are matrices $\mathbf{B}_u$, $\mathbf{E}_u$, and $\mathbf{N}_u$ such that $\mathbf{Q}_u = \mathbf{B}_u\mathbf{N}_u^{-1}\mathbf{B}_u^\top$ and $-\mathbf{P}_u = \mathbf{E}_u^\top\mathbf{E}_u$, with the pair $(\mathbf{A}_u,\, \mathbf{B}_u)$ being stabilizable, and the pair $(\mathbf{E}_u,\, \mathbf{A}_u)$ being observable. Hence, there is an unique solution $\boldsymbol{C}(t)$ to matrix Riccati equation* (2.19) *that is negative semidefinite and bounded on $(-\infty,\, T]$, and there exists a unique limit $\bar{\mathbf{C}} = \lim_{t\to-\infty} \boldsymbol{C}(t)$.*

Proof. We first perform a change of variable. Define $\tilde{\boldsymbol{C}}(t) = -\boldsymbol{C}(t)$, so the matrix Riccati equation (2.19) *becomes*

$$\begin{aligned}&\frac{\partial \tilde{\boldsymbol{C}}(t)}{\partial t} + \mathbf{A}_u^\top \tilde{\boldsymbol{C}}(t) + \tilde{\boldsymbol{C}}(t)\mathbf{A}_u - \tilde{\boldsymbol{C}}(t)\mathbf{Q}_u\tilde{\boldsymbol{C}}(t) + (-\mathbf{P}_u) = 0\ ,\\ &\tilde{\boldsymbol{C}}(T) = 0\ .\end{aligned} \tag{2.22}$$

Proposition 2.4 has shown that matrix $-\mathbf{P}_u \in \mathbb{R}^{d\times d}$ is symmetric positive definite. So, all eigenvalues $\boldsymbol{\lambda}_{P_u}$ of $-\mathbf{P}_u$ are positive and there exists an orthonormal basis for $\mathbb{R}^d$ of their associated eigenvectors, in other words, there is an orthonormal matrix $\mathbf{O}_u$ such that $-\mathbf{P}_u = \mathbf{O}_u\mathbf{D}_u\mathbf{O}_u^\top$, where $\mathbf{D}_u = diag\left(\boldsymbol{\lambda}_{P_u}\right) \in \mathbb{R}^{d\times d}$ is a diagonal matrix with positive entries on the diagonal. Hence, we can write $-\mathbf{P}_u = \mathbf{E}_u^\top\mathbf{E}_u$, where $\mathbf{E}_u = \left(\mathbf{O}_u\sqrt{\mathbf{D}_u}\right)^\top$ is a real square matrix. The matrix $\mathbf{E}_u$ is invertible, and so the controllability matrix $\boldsymbol{\Gamma}\left(\mathbf{A}_u^\top,\, \mathbf{E}_u^\top\right) \in \mathbb{R}^{d\times d^2}$, as defined by Definition 2.1, has rank d. Consequently, the pair $\left(\mathbf{A}_u^\top,\, \mathbf{E}_u^\top\right)$ is controllable and the pair $(\mathbf{E}_u,\, \mathbf{A}_u)$ is observable as per Definition 2.2.

The symmetric positive definiteness of $\mathbf{Q}_u$ is proven in Proposition 2.4 as well. Thus, the matrix $\mathbf{Q}_u \in \mathbb{R}^{d\times d}$ also has diagonal decomposition, $\mathbf{Q}_u = \mathbf{B}_u \mathbf{N}_u^{-1} \mathbf{B}_u^\top$, where $\mathbf{B}_u$ is an orthogonal matrix, and $\mathbf{N}_u^{-1} = diag\left(\boldsymbol{\lambda}_{Q_u}\right) \in \mathbb{R}^{d\times d}$ where $\boldsymbol{\lambda}_{Q_u}$ are the positive eigenvalues of $\mathbf{Q}_u$. The matrix $\mathbf{B}_u$ is invertible, and therefore we can find a constant matrix $\mathbf{M}_u \in \mathbb{R}^{d\times d}$ such that all eigenvalues of $\mathbf{A}_u - \mathbf{B}_u\mathbf{M}_u$ have negative real parts, and therefore the pair $(\mathbf{A}_u, \mathbf{B}_u)$ stabilizable.

The above analysis of matrices $\mathbf{Q}_u$ and $-\mathbf{P}_u$ confirms that we can apply Theorem 2.1 from [23] to conclude that solution $\tilde{\boldsymbol{C}}(t)$ to the matrix Riccati equation (2.22) is unique, positive semidefinite, bounded on $(-\infty, T]$, and has unique limit as t tends toward $-\infty$. □

Example 2.3. *Suppose instead of utilising the power utility function (2.8), we consider the exponential utility function*

$$U(x) = -\exp(-\gamma x) \ ,$$

where $\gamma > 0$. Then the ansatz function (2.11) becomes

$$u(t,\, w,\, \mathbf{z}) = -\exp(-\gamma w)\, g(t,\, \mathbf{z}) \ .$$

In the case, the HJB equation is

$$\begin{aligned} & g_t + (\boldsymbol{\theta} - \mathbf{z})^\top \boldsymbol{\kappa} \nabla_{\mathbf{z}} g + \frac{1}{2} tr\left(\boldsymbol{\Sigma}_3 \nabla^2_{\mathbf{z}} g\right) + r\gamma g \\ & \quad - \frac{1}{2}\left(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}}(\log g)\right)^\top \boldsymbol{\Sigma}_1^{-1}\left(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}}(\log g)\right) = 0 \ , \\ & g(T,\, \mathbf{z}) = 1 \ . \end{aligned}$$

Therefore, the condition for long-term stability of the matrix Riccati equation with respect to $\boldsymbol{C}(t)$ is that the matrix $\boldsymbol{\Sigma}_2^\top \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2 + \boldsymbol{\Sigma}_3$ has to be symmetric positive definite. In other words exponential utility function has the same matrix Riccati equation as the power utility function with $\gamma \to -\infty$.

Remark 1. *The stability analysis of Proposition 2.5 is sufficient for there to be no arbitrage in the model proposed by equations (2.1) and (2.2). If there were an arbitrage then it would always be optimal to take additional positions in the arbitrage portfolio, hence causing the value function to reach a Nirvana (see [15]), but there will be singularity if there is stability of the matrix Riccati equation with respect to $\boldsymbol{C}(t)$.*

After solving the matrix Riccati equation (2.18), we then can study the stability of the solution to the ODE with respect to $\boldsymbol{b}(t)$, in other words equation (2.17). We start the analysis by introducing the following lemma (a theorem from [22]) in regard to the eigenvalues of matrices (general background for this lemma can be found in Chapter 1 of [14]).

Lemma 2.1 (Theorem from [22])**.** *Let $\mathbf{M}$ be a $d \times d$ matrix and define the field of values $\boldsymbol{S}(\mathbf{M}) := \left\{\mathbf{v}^\top \mathbf{M} \mathbf{v} \mid \mathbf{v} \text{ is a vector such that } \mathbf{v}^\top \mathbf{v} = 1\right\}$, which contains the eigenvalues of $\mathbf{M}$.*

(a) Suppose $\mathbf{M}_1$ and $\mathbf{M}_2$ are two $d \times d$ matrices. If λ is an eigenvalue of $\mathbf{M}_1 + \mathbf{M}_2$, then $\lambda \in \boldsymbol{S}(\mathbf{M}_1) + \boldsymbol{S}(\mathbf{M}_2) = \{\lambda_1 + \lambda_2 \mid \lambda_1 \in \boldsymbol{S}(\mathbf{M}_1),\ \lambda_2 \in \boldsymbol{S}(\mathbf{M}_2)\}$;

(b) Suppose $\mathbf{M}_1$ and $\mathbf{M}_2$ are two $d \times d$ matrices, $0 \notin \boldsymbol{S}(\mathbf{M}_2)$ and $\mathbf{M}_2^{-1}$ exits. If λ is an eigenvalues of $\mathbf{M}_2^{-1}\mathbf{M}_1$, then $\lambda \in {\boldsymbol{S}(\mathbf{M}_1)}/{\boldsymbol{S}(\mathbf{M}_2)} = \{{\lambda_1}/{\lambda_2} \mid \lambda_1 \in \boldsymbol{S}(\mathbf{M}_1),\ \lambda_2 \in \boldsymbol{S}(\mathbf{M}_2)\}$;

(c) Suppose $\mathbf{M}_1$ is an arbitrary $d \times d$ matrix, $\mathbf{M}_2$ is symmetric positive semidefinite matrix. If λ is an eigenvalues of $\mathbf{M}_1\mathbf{M}_2$, then $\lambda \in \boldsymbol{S}(\mathbf{M}_1)\boldsymbol{S}(\mathbf{M}_2) = \{\lambda_1\lambda_2 \mid \lambda_1 \in \boldsymbol{S}(\mathbf{M}_1),\ \lambda_2 \in \boldsymbol{S}(\mathbf{M}_2)\}$.

Proposition 2.6. *Let $\boldsymbol{R}_u(t)$ be the coefficient matrix of homogeneous part of equation (2.17). For $\gamma < 0$, if $\mathbf{1}^\top \vec{\boldsymbol{\rho}}_0 = \sum_i^d \rho_i^0 \geq 0$, then there exists $t^* > -\infty$ such that $\boldsymbol{R}_u(t)$ has all positive eigenvalues for $t < t^*$, and therefore the solution of ODE (2.17) has a finite steady state.*

Proof. By observing the ODE (2.17) and utilising the notations given by formulae (2.9) and (2.20), we have the coefficient matrix for the ODE,

$$\begin{aligned}\boldsymbol{R}_u(t) &= -2\boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2+\boldsymbol{\Sigma}_3\right)-\left(\frac{\gamma}{1-\gamma}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2-\boldsymbol{\kappa}\right) \\ &= -\boldsymbol{C}(t)\,\mathbf{Q}_u+\frac{1}{1-\gamma}\boldsymbol{\kappa}-\frac{\gamma\sigma_0}{1-\gamma}\boldsymbol{\delta}\boldsymbol{\sigma}^{-1}\boldsymbol{\rho}^{-1}\boldsymbol{\rho}_0\ .\end{aligned} \tag{2.23}$$

Let $\bar{\mathbf{C}}=\lim_{t\to-\infty}\boldsymbol{C}(t)$, then the limit of equation (2.23) is

$$\mathbf{R}_u=-\bar{\mathbf{C}}\mathbf{Q}_u+\frac{1}{1-\gamma}\boldsymbol{\kappa}-\frac{\gamma\sigma_0}{1-\gamma}\left(\boldsymbol{\rho}\boldsymbol{\sigma}\boldsymbol{\delta}^{-1}\right)^{-1}\boldsymbol{\rho}_0\ . \tag{2.24}$$

We can observe that matrices $\boldsymbol{\kappa}$, $\boldsymbol{\delta}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\rho}$ are symmetric positive definite. Proposition 2.4 proves that matrix $\mathbf{Q}_u$ is symmetric positive definite, and proposition 2.5 implies that $-\bar{\mathbf{C}}$ is symmetric positive semidefinite. Hence, both $\boldsymbol{\kappa}$ and $\mathbf{Q}_u$ are symmetric positive definite and so all the elements of sets $\boldsymbol{S}(\boldsymbol{\kappa})$ and $\boldsymbol{S}(\mathbf{Q}_u)$ are positive, and $-\bar{\mathbf{C}}$ is symmetric positive semidefinite and so all the elements of set $\boldsymbol{S}\left(-\bar{\mathbf{C}}\right)$ are non-negative. By the symmetric positive definiteness of those matrices, their product $\boldsymbol{\rho}\boldsymbol{\sigma}\boldsymbol{\delta}^{-1}$ is a positive definite matrix as well. Also, because $\boldsymbol{\rho}_0=\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top$, so $\operatorname{rank}\left(\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top\right)=1$, and

$$\begin{aligned}\boldsymbol{\rho}_0\vec{\boldsymbol{\rho}}_0 &= \vec{\boldsymbol{\rho}}_0\mathbf{1}^\top\vec{\boldsymbol{\rho}}_0 \\ &= \left(\mathbf{1}^\top\vec{\boldsymbol{\rho}}_0\right)\vec{\boldsymbol{\rho}}_0\ ,\end{aligned}$$

it follows that $\mathbf{1}^\top\vec{\boldsymbol{\rho}}_0$ is an eigenvalue of $\boldsymbol{\rho}_0$. Therefore, if $\mathbf{1}^\top\vec{\boldsymbol{\rho}}_0=\sum_i^d\rho_i^0\geq 0$, then $\boldsymbol{\rho}_0$ has non-negative eigenvalues.

We then apply Lemma 2.1 on the matrix $\mathbf{R}_u$ that is given by equation (2.24). By item (c), we see that matrix $-\bar{\mathbf{C}}\mathbf{Q}_u$ has non-negative eigenvalues. By item (b), we see that matrix $-\frac{\gamma\sigma_0}{1-\gamma}\left(\boldsymbol{\rho}\boldsymbol{\sigma}\boldsymbol{\delta}^{-1}\right)^{-1}\boldsymbol{\rho}_0$ has non-negative eigenvalues, and then by item (a) we see that $\frac{1}{1-\gamma}\boldsymbol{\kappa}-\frac{\gamma\sigma_0}{1-\gamma}\left(\boldsymbol{\rho}\boldsymbol{\sigma}\boldsymbol{\delta}^{-1}\right)^{-1}\boldsymbol{\rho}_0$ has positive eigenvalues. Hence, by item (a), we have that matrix $\mathbf{R}_u$ has positive eigenvalues. Therefore, there exists a $t^*>-\infty$ such that $\boldsymbol{R}_u(t)$ has positive eigenvalues for all $t<t^*$, and the solution $\boldsymbol{b}(t)$ of the equation (2.17) has a finite steady state. ☐

Proposition 2.7. *The long-term growth rate of the certainty equivalent is proportional to the solution $a(t)$ of the ODE (2.16).*

Proof. Given the utility function (2.8), the value function (2.11), and the expo- nential-affine ansatz for g that is defined by formula (2.15), the certainty equivalent is defined by

$$\begin{aligned}U^{-1}(u(t,\ w,\ \mathbf{z})) &= U^{-1}\left(\frac{1}{\gamma}w^\gamma g(t,\ \mathbf{z})\right) \\ &= w\exp\left(\frac{1}{\gamma}\left(a(t)+\boldsymbol{b}^\top(t)\,\mathbf{z}+\mathbf{z}^\top\boldsymbol{C}(t)\,\mathbf{z}\right)\right)\ .\end{aligned}$$

Hence, under the optimal control variable $\boldsymbol{\pi}^*$, the long-term growth rate is

$$\begin{aligned}\frac{\ln\left(U^{-1}(u(t,\ w,\ \mathbf{z}))\right)}{T-t} &= \frac{1}{T-t}\log\left(w\exp\left(\frac{1}{\gamma}\left(a(t)+\boldsymbol{b}^\top(t)\,\mathbf{z}+\mathbf{z}^\top\boldsymbol{C}(t)\,\mathbf{z}\right)\right)\right) \\ &= \frac{1}{\gamma(T-t)}\left(\gamma\log(w)+\boldsymbol{b}^\top(t)\,\mathbf{z}+\mathbf{z}^\top\boldsymbol{C}\mathbf{z}+a(t)\right)\ .\end{aligned}$$

From the analysis in Propositions 2.5 and 2.6, both solutions $\boldsymbol{C}(t)$ and $\boldsymbol{b}(t)$ to the ODEs (2.18) *and* (2.17) *have finite limits as t tends to $-\infty$, and so it follows that $a(t)$ is asymptotically linear as t tends to $-\infty$, and therefore,*

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T-t}\log\left(U^{-1}\left(u\left(t,\ w,\ \mathbf{z}\right)\right)\right) \qquad (2.25)\\
&=\lim_{t\to-\infty}\frac{1}{\gamma\left(T-t\right)}\left(\gamma\log\left(w\right)+\boldsymbol{b}^{\top}\left(t\right)\mathbf{z}\ +\mathbf{z}^{\top}\boldsymbol{C}\left(t\right)\mathbf{z}+a\left(t\right)\right)\\
&=\lim_{t\to-\infty}\frac{1}{\gamma\left(T-t\right)}a\left(t\right)\ .
\end{aligned}$$

Furthermore, denote by $L(t)$ the right hand side of the ODE with respect to $a(t)$, in other words equation (2.16)*. As T tends toward infinity we have*

$$\begin{aligned}
\lim_{t\to-\infty}L\left(t\right)=&-\bar{\boldsymbol{b}}^{\top}\left(\frac{\gamma}{2\left(1-\gamma\right)}\boldsymbol{\Sigma}_2^{\top}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2+\frac{1}{2}\boldsymbol{\Sigma}_3\right)\bar{\boldsymbol{b}}-\boldsymbol{\theta}^{\top}\boldsymbol{\kappa}\bar{\boldsymbol{b}}-tr\left(\boldsymbol{\Sigma}_3\bar{\boldsymbol{C}}\right)\\
&-\frac{\gamma}{2\left(1-\gamma\right)}\bar{\boldsymbol{b}}^{\top}\boldsymbol{\Sigma}_2^{\top}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}-\frac{\gamma}{2\left(1-\gamma\right)}\boldsymbol{\mu}^{\top}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2\bar{\boldsymbol{b}}\\
&-\frac{\gamma}{2\left(1-\gamma\right)}\boldsymbol{\mu}^{\top}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}-r\gamma\\
=&\ \bar{L}\ ,
\end{aligned}$$

where $\bar{\boldsymbol{b}}=\lim_{t\to-\infty}\boldsymbol{b}(t)$ and $\bar{\mathbf{C}}=\lim_{t\to-\infty}\boldsymbol{C}(t)$. Hence, as t tends toward negative infinity, equation (2.16) *relaxes and we have*

$$\lim_{t\to-\infty}\frac{1}{T-t}\int_t^T\frac{\partial a\left(s\right)}{\partial s}ds=\bar{L}\ ,$$

and therefore the limit in equation (2.25) *is*

$$\lim_{T\to\infty}\frac{1}{T-t}\log\left(U^{-1}\left(u\left(t,\ w,\ \mathbf{z}\right)\right)\right)=\frac{1}{\gamma}\bar{L}\ ,$$

and the long-term growth rate is a constant. □

2.4 Portfolio Constraints

It is common practice to seek statistical arbitrage strategies that have market neutrality. Market-neutrality generally means that the returns of a portfolio are impacted only by the idiosyncratic returns of the assets contained in the portfolio, and are uncorrelated with the returns of a benchmark or market factors. Hence, under the condition of market neutrality, if we can diversify with a large number of co-integrated assets, then there is a very high probability that the portfolio can maintain steady growth and low volatility. Market neutrality was discussed at length in the principal component analysis of [4]. We will consider a portfolio to be market neutral if $\frac{dW_t}{W_t}\frac{dS_t^0}{S_t^0}=0$. This happens as follows,

$$\begin{aligned}
\frac{dW_t}{W_t}\frac{dS^0_t}{S^0_t} &= \sum_{i=1}^{d} \pi^i_t \frac{dS^i_t}{S^i_t}\frac{dS^0_t}{S^0_t} \\
&= \sum_{i=1}^{d} \pi^i_t d\log(S^i_t)\left(dZ^i_t - \beta_i \log S^i_t\right) \\
&= \sum_{i=1}^{d} \pi^i_t \left(\mathbf{\Sigma}^{ii}_2 - \boldsymbol{\beta}\mathbf{\Sigma}^{ii}_1\right) dt \\
&= \sigma_0 \sum_{i=1}^{d} \pi^i_t \sigma_i \rho^0_i dt \\
&= 0 \ .
\end{aligned}$$

where $\mathbf{\Sigma}^{ii}_1$ and $\mathbf{\Sigma}^{ii}_2$ are the i^{th} elements of the diagonals of matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ respectively.

In matrix/vector notation, market neutrality is $\boldsymbol{\pi}^\top \boldsymbol{s} = 0$, where $\boldsymbol{s} = \sigma_0[\sigma_1\rho^0_1, \sigma_2\rho^0_2, \cdots, \sigma_d\rho^0_d]^\top \in \mathbb{R}^d$. If $\boldsymbol{\pi}^\top \boldsymbol{s} = 1$, then the portfolio has a unit-loading on the benchmark return. Dollar neutrality is $\boldsymbol{\pi}^\top \mathbf{1} = 0$, where $\mathbf{1} = [1,\ 1, \cdots,\ 1]^\top \in \mathbb{R}^d$, in other words the dollar amount invested in the risky assets sums to zero. If $\boldsymbol{\pi}^\top \mathbf{1} = 1$ then the dollar amount invested in the riskless asset is zero so that the portfolio is 100% long. Consequently, we obtain the following four different equality constraints:

$$\begin{aligned}
\boldsymbol{\pi}^\top \boldsymbol{s} = 0 &: \text{ Constraint for market neutral,} \\
\boldsymbol{\pi}^\top \boldsymbol{s} = 1 &: \text{ Constraint for unit loading on market portfolio's returns,} \\
\boldsymbol{\pi}^\top \mathbf{1} = 0 &: \text{ Constraint for dollar neutral,} \\
\boldsymbol{\pi}^\top \mathbf{1} = 1 &: \text{ Constraint for 100\% long.}
\end{aligned} \tag{2.26}$$

With the constraints of (2.26) in mind, we now reformulate the optimal portfolio that is studied in Section 2.1 and Section 2.2. We shall start from a general situation, in other words, we want to integrate an equality constraint of the control variables $\pi_1 s_1 + \pi_2 s_2 + \cdots + \pi_d s_d = \boldsymbol{\pi}^\top \mathbf{s} = s_\pi$ into the stochastic control and optimisation problem, which can include all the four different cases that are given by formula (2.26), where $\mathbf{s} = [s_1,\ s_2,\ \cdots,\ s_d] \in \mathbb{R}^d$ is column vector and $s_\pi \in \mathbb{R}$ is a constraint coefficient. Therefore, the HJB equation (2.10) becomes

$$\begin{aligned}
&- u_t - (\boldsymbol{\theta} - \mathbf{z})^\top \boldsymbol{\kappa} \nabla_{\mathbf{z}} u - \frac{1}{2}\text{tr}\left(\mathbf{\Sigma}_3 \nabla^2_{\mathbf{z}} u\right) - rwu_w \\
&- \sup_{\substack{\boldsymbol{\pi}\in\mathbb{R}^d \\ \boldsymbol{\pi}^\top \mathbf{s} = s_\pi}} \left(\boldsymbol{\pi}^\top (\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z})\, wu_w + \boldsymbol{\pi}^\top \mathbf{\Sigma}_2 \nabla_{\mathbf{z}} (\nabla_w u)\, w + \frac{1}{2}\boldsymbol{\pi}^\top \mathbf{\Sigma}_1 \boldsymbol{\pi} w^2 u_{ww}\right) = 0 \ . \\
&u(T,\ w,\ \mathbf{z}) = \frac{1}{\gamma} w^\gamma \ .
\end{aligned} \tag{2.27}$$

Hence the transformed HJB equation (2.12) becomes

$$\begin{aligned}
&- g_t - (\boldsymbol{\theta} - \mathbf{z})^\top \boldsymbol{\kappa} \nabla_{\mathbf{z}} g - \frac{1}{2}\text{tr}\left(\mathbf{\Sigma}_3 \nabla^2_{\mathbf{z}} g\right) - r\gamma g \\
&- \inf_{\substack{\boldsymbol{\pi}\in\mathbb{R}^d \\ \boldsymbol{\pi}^\top \mathbf{s} = s_\pi}} \left(\boldsymbol{\pi}^\top (\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z})\, \gamma g + \boldsymbol{\pi}^\top \gamma \mathbf{\Sigma}_2 \nabla_{\mathbf{z}} g + \frac{1}{2}\boldsymbol{\pi}^\top \mathbf{\Sigma}_1 \boldsymbol{\pi} g\gamma (\gamma - 1)\right) = 0 \ , \\
&g(T,\ \mathbf{z}) = 1 \ .
\end{aligned} \tag{2.28}$$

Letting λ denote a Lagrange multiplier, we define the Lagrangian function for the constrained control model

$$
\begin{aligned}
L(\boldsymbol{\pi},\, \lambda) = &\boldsymbol{\pi}^\top (\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z})\, \gamma g + \boldsymbol{\pi}^\top \gamma \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} g \\
&+ \frac{1}{2}\boldsymbol{\pi}^\top \boldsymbol{\Sigma}_1 \boldsymbol{\pi} g\gamma (\gamma - 1) - \lambda \left(\boldsymbol{\pi}^\top \mathbf{s} - s_\pi\right) .
\end{aligned}
$$

Therefore, by the first order condition $\nabla_{\boldsymbol{\pi}} L = \mathbf{0}$, we can get the optimal values for the control variables

$$
\begin{aligned}
\boldsymbol{\pi}^* =& \frac{1}{g\gamma(1-\gamma)} \boldsymbol{\Sigma}_1^{-1} \left(g\gamma\boldsymbol{\mu} + g\gamma\boldsymbol{\delta}\mathbf{z} + \gamma\boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} g - \lambda \mathbf{s}\right) \\
=& \frac{1}{1-\gamma} \boldsymbol{\Sigma}_1^{-1} \left(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} (\log g) - \frac{\lambda}{g\gamma}\mathbf{s}\right) .
\end{aligned} \tag{2.29}
$$

We then solve it for the Lagrange multiplier λ to get its optimal value with respect to the condition of constraint $s_\pi = (\boldsymbol{\pi}^*)^\top \mathbf{s} = \mathbf{s}^\top \boldsymbol{\pi}^*$

$$
\lambda^* = \frac{1}{\mathbf{s}^\top \boldsymbol{\Sigma}_1^{-1} \mathbf{s}} \left(\gamma \mathbf{s}^\top \boldsymbol{\Sigma}_1^{-1} \left(g\boldsymbol{\mu} + g\boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} g\right) - g\gamma(1-\gamma)\, s_\pi\right) . \tag{2.30}
$$

Inserting λ^* that is given by formula (2.30) back into equation (2.29), we can get the optimal control variables $\boldsymbol{\pi}^*$ with respect to the constraint coefficient s_π

$$
\begin{aligned}
\boldsymbol{\pi}^* = &\frac{s_\pi}{d_c} \boldsymbol{\Sigma}_1^{-1} \mathbf{s} + \frac{1}{1-\gamma}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z}) \\
&+ \frac{1}{1-\gamma}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right) \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} (\ln g) \ ,
\end{aligned} \tag{2.31}
$$

where $d_c = \mathbf{s}^\top \boldsymbol{\Sigma}_1^{-1} \mathbf{s} \in \mathbb{R}$ is a scalar, $\boldsymbol{\Sigma}_c = \boldsymbol{\Sigma}_1^{-1} \mathbf{s}\mathbf{s}^\top \boldsymbol{\Sigma}_1^{-1} \in \mathbb{R}^{d\times d}$ is a symmetric matrix. We then insert the optimal control variable $\boldsymbol{\pi}^*$ that is given by formula (2.31) back into the HJB equation (2.28) and get the following non-linear partial differential equation

$$
\begin{aligned}
&g_t + (\boldsymbol{\theta} - \mathbf{z})^\top \boldsymbol{\kappa} \nabla_{\mathbf{z}} g + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_3 \nabla_{\mathbf{z}}^2 g\right) + r\gamma g \\
&\quad + \frac{g\gamma(\gamma-1)\, s_\pi^2}{2d_c} + \frac{\gamma g s_\pi}{d_c} \mathbf{s}^\top \boldsymbol{\Sigma}_1^{-1} \left(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} (\ln g)\right) \\
&\quad + \frac{\gamma g}{2(1-\gamma)} \left(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} (\ln g)\right)^\top \left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right) \\
&\quad \times \left(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z} + \boldsymbol{\Sigma}_2 \nabla_{\mathbf{z}} (\ln g)\right) = 0 \\
&g(T,\, \mathbf{z}) = 1 \ .
\end{aligned} \tag{2.32}
$$

Similar to the unconstrained stochastic control problem of Section 2.2, we approach removing the nonlinearity in PDE (2.32) by utilising the exponential-affine ansatz of (2.15), from which the optimal portfolio can be expressed as

$$
\begin{aligned}
\boldsymbol{\pi}^* = &\frac{s_\pi}{d_c} \boldsymbol{\Sigma}_1^{-1} \mathbf{s} + \frac{1}{1-\gamma}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)(\boldsymbol{\mu} + \boldsymbol{\delta}\mathbf{z}) \\
&+ \frac{1}{1-\gamma}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right) \boldsymbol{\Sigma}_2 \left(2\boldsymbol{C}(t)\mathbf{z} + \boldsymbol{b}(t)\right) \ ,
\end{aligned}
$$

where $a(t)$, $\boldsymbol{b}(t)$ and $\boldsymbol{C}(t)$ are the solutions to the following system of ODEs:

Proposition 2.8. *The PDE* (2.32) *is solved with the exponential-affine ansatz of* (2.15)*, and the functions* $a(t) \in \mathbb{R}$, $\boldsymbol{b}(t) \in \mathbb{R}^d$, *and* $\boldsymbol{C}(t) \in \mathbb{R}^{d\times d}$ *satisfy the following system of ODEs:*

$$\begin{aligned}
\frac{\partial a(t)}{\partial t} &= -\boldsymbol{b}^\top(t)\left(\frac{\gamma}{2(1-\gamma)}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 + \frac{1}{2}\boldsymbol{\Sigma}_3\right)\boldsymbol{b}(t) \\
&\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{\mu}^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\mu} - tr[\boldsymbol{\Sigma}_3\boldsymbol{C}(t)] \\
&\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{b}^\top(t)\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\mu} - \frac{\gamma(\gamma-1)s_\pi^2}{2d_c} \\
&\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{\mu}^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2\boldsymbol{b}(t) - r\gamma \\
&\quad - \left(\boldsymbol{\theta}^\top\boldsymbol{\kappa} + \frac{\gamma s_\pi}{d_c}\mathbf{s}^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2\right)\boldsymbol{b}(t) - \frac{\gamma s_\pi}{d_c}\mathbf{s}^\top\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}, \\
a(T) &= 0\ ;
\end{aligned} \tag{2.33}$$

$$\begin{aligned}
\frac{\partial \boldsymbol{b}(t)}{\partial t} &= -2\boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 + \boldsymbol{\Sigma}_3\right)\boldsymbol{b}(t) \\
&\quad - \left(\frac{\gamma}{1-\gamma}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 - \boldsymbol{\kappa}\right)\boldsymbol{b}(t) \\
&\quad - 2\boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\mu} + \boldsymbol{\kappa}\boldsymbol{\theta}\right) \\
&\quad - \frac{\gamma}{1-\gamma}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\mu} - \frac{\gamma s_\pi}{d_c}\left[2\boldsymbol{C}(t)\boldsymbol{\Sigma}_2^\top + \boldsymbol{\delta}\right]\boldsymbol{\Sigma}_1^{-1}\mathbf{s}, \\
\boldsymbol{b}(T) &= \mathbf{0}\ ;
\end{aligned} \tag{2.34}$$

$$\begin{aligned}
\frac{\partial \boldsymbol{C}(t)}{\partial t} &= -\boldsymbol{C}(t)\left(\frac{2\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 + 2\boldsymbol{\Sigma}_3\right)\boldsymbol{C}(t) \\
&\quad - \boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\delta} - \boldsymbol{\kappa}\right) \\
&\quad - \left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\delta} - \boldsymbol{\kappa}\right)^\top\boldsymbol{C}(t) \\
&\quad - \frac{\gamma}{2(1-\gamma)}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\delta}, \\
\boldsymbol{C}(T) &= \mathbf{0}\ .
\end{aligned} \tag{2.35}$$

Proof. *The proof is the same as that for Proposition 2.2.* □

Proposition 2.9 (Verification)**.** *The optimal control variable* $\boldsymbol{\pi}^*$ *that is given by formula* (2.13) *utilising the exponential-affine ansatz that is proposed by formula* (2.15)*, where* $\boldsymbol{C}(t)$, $\boldsymbol{b}(t)$ *and* $a(t)$ *are respectively the solutions of matrix Riccati equation* (2.35)*, ODEs* (2.34) *and* (2.33)*, belongs to the set of admissible controls* $\mathcal{A}$ *that is described by formula* (2.7) *and maximises the expected utility function that is defined in formula* (2.6)*.*

Proof. *Similar to Proposition 2.3 for the unconstrained stochastic control problem, the verification method of [6] and [7] applies.* □

Similar to the unconstrained stochastic control problem of Sections 2.2 and 2.3, we perform the stability analysis for the solutions of the system of ODEs given by equations (2.33), (2.34), and (2.35). We first study the behaviour of the solution to the matrix Riccati equation $\boldsymbol{C}(t)$ of the constrained stochastic control problem, which is expressed by equation (2.35). We rewrite this matrix Riccati equation in the following way

$$\begin{aligned} \frac{\partial \boldsymbol{C}(t)}{\partial t} &= -\boldsymbol{C}(t)\,\mathbf{Q}_c\boldsymbol{C}(t) - \boldsymbol{C}(t)\,\mathbf{A}_c - \mathbf{A}_c^{\top}\boldsymbol{C}(t) - \mathbf{P}_c \ , \\ \boldsymbol{C}(T) &= \mathbf{0} \ , \end{aligned} \tag{2.36}$$

where

$$\begin{aligned} \mathbf{Q}_c &= \frac{2\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^{\top}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 + 2\boldsymbol{\Sigma}_3 \ , \\ \mathbf{A}_c &= \frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^{\top}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\delta} - \boldsymbol{\kappa} \ , \\ \mathbf{P}_c &= \frac{\gamma}{2(1-\gamma)}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\delta} \ . \end{aligned}$$

Similar to the unconstrained stochastic control problem, we prove that the coefficient matrix of its quadratic term is positive definite.

Proposition 2.10. *For $\gamma < 0$, the quadratic term $\mathbf{Q}_c$ of the matrix Riccati equation* (2.36) *is symmetric positive definite and the matrix $\mathbf{P}_c$ is symmetric negative semidefinite.*

Proof. Proposition 2.4 has shown that $\mathbf{Q}_u$ is positive definite. By utilising formula (2.20)*, $\mathbf{Q}_c$ has the following decomposition*

$$\begin{aligned} \mathbf{Q}_c &= 2\boldsymbol{\Sigma}_3 + \frac{2\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^{\top}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 \\ &= 2\boldsymbol{\Sigma}_3 + \frac{2\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^{\top}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\mathbf{s}^{\top}\boldsymbol{\Sigma}_1^{-1}\right)\boldsymbol{\Sigma}_2 \\ &= \mathbf{Q}_u - \frac{2\gamma}{d_c(1-\gamma)}\boldsymbol{\Sigma}_2^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\mathbf{s}^{\top}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 \ . \end{aligned}$$

Because $\boldsymbol{\Sigma}_1$ is symmetric positive definite, so its inverse $\boldsymbol{\Sigma}_1^{-1}$ is symmetric positive definite as well. Hence $d_c = \boldsymbol{s}^{\top}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{s} > 0$. Therefore, for $\gamma < 0$, $\mathbf{Q}_c$ is positive definite.

In order to prove $\mathbf{P}_c$ is symmetric negative semidefinite, we first examine the symmetric matrix $\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c$. Utilising the Cauchy-Schwarz inequality $(\mathbf{x}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 \leq (\mathbf{x}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{x})(\boldsymbol{s}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{s})$, we observe that for any $\mathbf{x} \in \mathbb{R}^d$,

$$\mathbf{x}^{\top}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\mathbf{x} = \mathbf{x}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{x} - \frac{\mathbf{x}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\mathbf{s}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{x}}{\mathbf{s}^{\top}\boldsymbol{\Sigma}_1^{-1}\mathbf{s}} \geq 0 \ ,$$

where the equality holds if and only if $\mathbf{x} = \boldsymbol{s}$. Hence, $\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c$ is positive semidefinite. Consequently, $\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\delta}$ is a symmetric positive semidefinite matrix with $\boldsymbol{\delta}^{-1}\mathbf{s}$ being the single null vector. Therefore, $-\mathbf{P}_c$ is symmetric positive semidefinite with null-space spanned by $\boldsymbol{\delta}^{-1}\mathbf{s}$. ☐

Similar to Proposition 2.5, because $\mathbf{Q}_c$ is symmetric positive definite, $-\mathbf{P}_c$ is symmetric positive semidefinite, theorem 2.1 in [23] applies directly, and the solution to matrix Riccati equation (2.35) exists, is bounded and is unique.

Proposition 2.11. *For $\gamma < 0$ and $\boldsymbol{\kappa} \not\propto \mathbf{I}$, where $\mathbf{I} \in \mathbb{R}^{d\times d}$ is the identity matrix, the coefficient matrix $\mathbf{Q}_c$ of the quadratic term in the matrix Riccati equation (2.36) of the constrained stochastic control problem is symmetric positive definite and $-\boldsymbol{P}_c$ is symmetric positive semidefinite. Consequently there are matrices $\mathbf{B}_c$, $\mathbf{E}_c$, and $\mathbf{N}_c$ such that $\mathbf{Q}_c = \mathbf{B}_c\mathbf{N}_c^{-1}\mathbf{B}_c^\top$ and $-\mathbf{P}_c = \mathbf{E}_c^\top\mathbf{E}_c$, with the pair $(\mathbf{A}_c, \mathbf{B}_c)$ being stabilizable, and the pair $(\mathbf{E}_c, \mathbf{A}_c)$ being observable. Hence, there is an unique solution $\boldsymbol{C}(t)$ to matrix Riccati equation (2.36) that is negative semidefinite and bounded on $(-\infty, T]$, and there exists a unique limit $\bar{\mathbf{C}} = \lim_{t\to-\infty}\boldsymbol{C}(t)$.*

Proof. The symmetric positive semidefiniteness of the matrix $-\mathbf{P}_c$ comes from Proposition 2.10, for which the proof shows that $\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\right)\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\right)^\top$ is symmetric positive semidefinite. This matrix is diagonalizable, $\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\right)\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\right)^\top = \mathbf{O}_c\mathbf{D}_c\mathbf{O}_c^\top$, where $\mathbf{O}_c \in \mathbb{R}^{d\times d}$ is orthonormal and $\mathbf{D}_c$ is a diagonal matrix with the non-negative eigenvalues of $\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\right)\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\right)^\top$ on its diagonal. Thus $-\mathbf{P}_c = \mathbf{E}_c^\top\mathbf{E}_c$, where $\mathbf{E}_c = \sqrt{\frac{-\gamma}{2(1-\gamma)}}\sqrt{\mathbf{D}_c}\mathbf{O}_c^\top\boldsymbol{\delta}$. As shown in the proof of Proposition 2.10, there is an one dimensional null space of $-\mathbf{P}_c$, spanned by $\boldsymbol{\delta}^{-1}\boldsymbol{s}$, and hence the rank of matrix $\mathbf{E}_c$ is $d-1$ also with null space spanned by $\boldsymbol{\delta}^{-1}\boldsymbol{s}$. However, the rank of the controllability matrix $\boldsymbol{\Gamma}\left(\mathbf{A}_c^\top, \mathbf{E}_c^\top\right) \in \mathbb{R}^{d\times d^2}$ given by Definition 2.1 is d if $\boldsymbol{k} \not\propto \boldsymbol{I}$. Indeed , recall the matrix $\mathbf{A}_c$ defined for the matrix Riccati equation (2.36), $\mathbf{A}_c^\top = \frac{\gamma}{1-\gamma}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)^\top\boldsymbol{\Sigma}_2 - \boldsymbol{\kappa} = 2\mathbf{E}_c^\top\mathbf{E}_c\boldsymbol{\delta}^{-1}\boldsymbol{\Sigma}_2 - \boldsymbol{\kappa}$, which when multiplied on the left-hand side by $\boldsymbol{\delta}^{-1}\boldsymbol{s}$ and on the right by $\boldsymbol{E}_c^\top$,

$$
\begin{aligned}
\left(\boldsymbol{\delta}^{-1}\mathbf{s}\right)^\top \mathbf{A}_c^\top\mathbf{E}_c^\top &= \left(\boldsymbol{\delta}^{-1}\mathbf{s}\right)^\top\left(2\mathbf{E}_c^\top\mathbf{E}_c\boldsymbol{\delta}^{-1}\boldsymbol{\Sigma}_2 - \boldsymbol{\kappa}\right)\mathbf{E}_c^\top \\
&= 2\left(\mathbf{E}_c\boldsymbol{\delta}^{-1}\mathbf{s}\right)^\top\mathbf{E}_c\boldsymbol{\delta}^{-1}\boldsymbol{\Sigma}_2\mathbf{E}_c^\top - \left(\boldsymbol{\delta}^{-1}\mathbf{s}\right)^\top\boldsymbol{\kappa}\mathbf{E}_c^\top \\
&= -\left(\mathbf{E}_c\boldsymbol{\kappa}\boldsymbol{\delta}^{-1}\mathbf{s}\right)^\top \\
&\neq 0, \quad \text{if} \quad \boldsymbol{\kappa} \not\propto \mathbf{I}\ .
\end{aligned}
$$

Therefore, if $\boldsymbol{\kappa} \not\propto \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, then $\boldsymbol{\Gamma}\left(\mathbf{A}_c^\top, \mathbf{E}_c^\top\right) \in \mathbb{R}^{d\times d^2}$ has full rank. Thus, the pair $\left(\mathbf{A}_u^\top, \mathbf{E}_u^\top\right)$ is controllable and the pair $(\mathbf{E}_u, \mathbf{A}_u)$ is observable as per Definition 2.2.

The symmetric positive definiteness of matrix $\mathbf{Q}_c$ is proven in Proposition 2.10 as well. Similar to Proposition 2.5, $\mathbf{Q}_c = \mathbf{B}_c\mathbf{N}_c^{-1}\mathbf{B}_c^\top$, where $\mathbf{B}_c$ is an orthogonal matrix, $\boldsymbol{\lambda}_{Q_c}$ are the eigenvalues of $\mathbf{Q}_c$, and $\mathbf{N}_c^{-1} = diag\left(\boldsymbol{\lambda}_{Q_c}\right) \in \mathbb{R}^{d\times d}$. The matrix $\mathbf{B}_c$ is invertible, therefore we can find a constant matrix $\mathbf{M}_c \in \mathbb{R}^{d\times d}$ such that all eigenvalues of matrix $\mathbf{A}_c - \mathbf{B}_c\mathbf{M}_c$ have negative real parts, and therefore the pair $(\mathbf{A}_c, \mathbf{B}_c)$ is stabilizable.

Finally, we let $\tilde{\boldsymbol{C}}(t) = -\boldsymbol{C}(t)$, so that the matrix Riccati equation (2.36) becomes

$$
\begin{aligned}
&\frac{\partial\tilde{\boldsymbol{C}}(t)}{\partial t} + \mathbf{A}_u^\top\tilde{\boldsymbol{C}}(t) + \tilde{\boldsymbol{C}}(t)\mathbf{A}_u - \tilde{\boldsymbol{C}}(t)\mathbf{B}_c\mathbf{N}_c^{-1}\mathbf{B}_c^\top\tilde{\boldsymbol{C}}(t) + \mathbf{E}_c^\top\mathbf{E}_c = 0\ , \\
&\tilde{\boldsymbol{C}}(T) = 0\ ,
\end{aligned}
\tag{2.37}
$$

and the above analysis of matrices $\mathbf{Q}_c$ and $-\mathbf{P}_c$ confirms that we can apply Theorem 2.1 from [23] again, to conclude that solution $\tilde{\boldsymbol{C}}(t)$ to equation (2.37) is unique, positive semidefinite, bounded on $(-\infty, T]$, and has unique limit as t tends toward $-\infty$. □

Next we analyse the behaviour of the solution for the ODE with respect to $\boldsymbol{b}(t)$ of the constrained stochastic control problem that is described by equation (2.34).

Proposition 2.12. *Let $\boldsymbol{R}_c(t)$ be the coefficient matrix for the homogeneous part of equation* (2.34)*. For $\gamma < 0$, if $\mathbf{1}^\top \vec{\boldsymbol{\rho}}_0 \geq 0$ then there exists $t^* > -\infty$ such that $\boldsymbol{R}_c(t)$ has all positive eigenvalues for $t < t^*$. Therefore, the solution of the equation* (2.34) *has a finite steady state.*

Proof. *By observing the equation* (2.34) *and following the notations that are denoted by formula* (2.9) *and formula* (2.20)*, we have*

$$\begin{aligned}
\boldsymbol{R}_c(t) &= -2\boldsymbol{C}(t)\left(\frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}_2^\top\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 + \boldsymbol{\Sigma}_3\right) \\
&\quad - \left(\frac{\gamma}{1-\gamma}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_c\right)\boldsymbol{\Sigma}_2 - \boldsymbol{\kappa}\right) \\
&= -\boldsymbol{C}(t)\mathbf{Q}_c + \boldsymbol{\kappa} - \frac{\gamma}{1-\gamma}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\left(\sigma_0\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \boldsymbol{\Sigma}_1\boldsymbol{\beta}\right) \\
&\quad + \frac{\gamma}{1-\gamma}\boldsymbol{\delta}\frac{1}{d_c}\boldsymbol{\Sigma}_c\left(\sigma_0\boldsymbol{\sigma}\vec{\boldsymbol{\rho}}_0\mathbf{1}^\top + \boldsymbol{\Sigma}_1\boldsymbol{\beta}\right) \\
&= -\boldsymbol{C}(t)\mathbf{Q}_c + \frac{1}{1-\gamma}\boldsymbol{\kappa} + \frac{1}{d_c(1-\gamma)}\boldsymbol{\delta}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{s}\boldsymbol{s}^\top\gamma\boldsymbol{\beta} \\
&\quad - \frac{\gamma\sigma_0}{1-\gamma}\boldsymbol{\delta}\left(\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{s}\boldsymbol{s}^\top\boldsymbol{\Sigma}_1^{-1}\right)\boldsymbol{\sigma}\boldsymbol{\rho}_0 \ .
\end{aligned}$$

Because of $\gamma < 0$ and the assumption for the co-integrated vector, in other words $\beta_i < 0$, the matrix $\gamma\boldsymbol{\beta}$ is positive definite. From Proposition 2.10, we know that matrix $\mathbf{Q}_c$ is positive definite, from Proposition 2.11, we have that matrix $-\boldsymbol{C}(t)$ is positive semidefinite for $t < t^*$ with $t^* > -\infty$, and by the assumption of the model, we have that matrices $\boldsymbol{\kappa}$ and $\boldsymbol{\delta}$ are positive definite. Note also that $\boldsymbol{\Sigma}_1^{-1} - \frac{1}{d_c}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{s}\boldsymbol{s}^\top\boldsymbol{\Sigma}_1^{-1}$ is symmetric positive semidefinite. Hence, with the attribute for the eigenvalue of matrix $\boldsymbol{\rho}_0$ that is proven in Proposition 2.6, if $\mathbf{1}^\top\vec{\boldsymbol{\rho}}_0 \geq 0$ then by Lemma 2.1, $\boldsymbol{R}_c(t)$ has all positive eigenvalues for $t < t^*$ with $t^* > -\infty$. Therefore, the solution of the equation (2.34) *is stable.* □

Remark 2. *The long-term growth rate of the certainty equivalent of the constrained stochastic control problem is proportional to the solution $a(t)$ of equation* (2.33)*. The proof of such a proposition is similar to the proof of Proposition 2.7.*

3 Parameter Estimation & Portfolio Perfomance

This section describes the data, parameter estimation and numerical methods of estimating parameters for solving the optimal controls. The results demonstrate how the formulae in Section 2 apply to real-life finance.

3.1 Data and Parameters

We utilise the Yahoo Finance as our data source. The data set that we utilise is the adjusted daily close stock prices of the S&P 500 constituents from 2012-01-03 to 2019-05-15, and also includes the SPY ETF among these traded assets. The non-traded benchmark S^0 is the S&P500 Index. We assume that the interest rate r is 0.02, and in every calendar year, there are 252 trading days. The three parameters for the OU process Z_t^i that is defined by formula (2.3) are $(a_i,\, b_i,\, \beta_i)$ for $i = 1,\, 2,\, \cdots,\, d = 501$, and are the outputs of the Engle-Granger co-integration test. By including the parameter a_i, the process Z_t^i has a stationary mean that is not significantly different from zero. After finding all the co-integrated pairs, we select 10 out of the 106 companies that are the fastest mean reverters, in other words such that their values of $-\beta_i\delta_i$ are the largest ones. In order to estimate the parameter δ_i for the co-integrated process that is described by equation (2.4), we estimate an order-one auto-regressive coefficient for the discrete time series of the Z_t^i process via least squares estimation, and then take logarithms, divide by $\Delta t = 1/252$, and multiply by $-\beta_i$ to obtain κ_i. Table 3.1 shows the 10 fastest mean reverters alongside their Sharpe ratios and estimated rates of mean reversion.

For the 10 selected assets for trading, it remains to calculate their parameters for the returns model. The expected rate of returns μ_i are estimated with the a sample mean/median, but tend to be higher than prior views would suggest, so we place a multiplier of 0.5 in order to keep them near the range of single-digit percentages, in other words, the estimator is $\hat{\mu}_i = 0.5 \times \text{mean}[\Delta S_t^i / S_t^i]$ is such that $1\% \leq \hat{\mu}_i \leq 20\%$. The covariance matrix $\mathbf{\Sigma}_1 = [\sigma_i\sigma_j\rho_{ij}] \in \mathbb{R}^{10\times 10}$ is estimated with a standard method of moments. Figure 3.1 shows the estimated Z_t^i processes.

Ticker	**Sharpe Ratio**	**Mean-Reversion Rate**
HST	0.338003	7.507944
URI	0.651762	7.125851
WDC	0.346716	7.624259
APH	1.011030	8.724717
MU	0.713108	9.123977
AMP	0.698346	9.321913
BWA	0.170926	9.305448
TEL	0.780984	7.016104
TIF	0.329176	7.433595
SPY	0.915641	43.844862

Table 3.1: Sharpe ratios and mean-reversion rates for the 10 fastest mean-reverting stocks. The Sharpe ratios are raw time-series estimates.

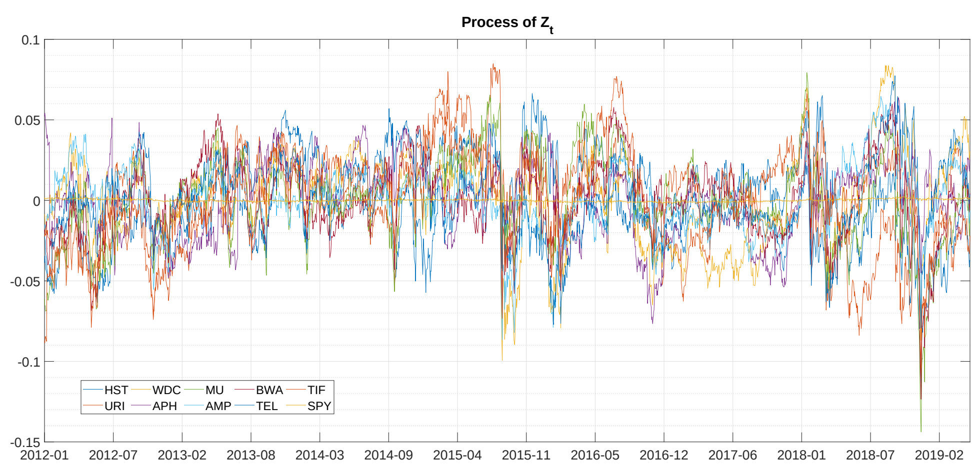

Figure 3.1: Co-integrated process Z_t for the 10 fastest mean reverters.

After estimating all parameters in the stochastic system, we then solve the matrix Riccati equation for $\boldsymbol{C}(t)$ by applying Radon's Lemma (see Chapter 3 of [1]). The method is a forward-time scheme, so we first perform the change of variable $\tau = T - t$, and then the matrix Riccati equation (2.22) for the unconstrained stochastic control problem becomes

$$\frac{\partial \tilde{\boldsymbol{C}}(\tau)}{\partial \tau} = \mathbf{A}_u^\top \tilde{\boldsymbol{C}}(\tau) + \tilde{\boldsymbol{C}}(\tau)\mathbf{A}_u - \tilde{\boldsymbol{C}}(\tau)\mathbf{Q}_u \tilde{\boldsymbol{C}}(\tau) - \mathbf{P}_u \ ,$$
$$\tilde{\boldsymbol{C}}(\tau) = 0 \ .$$

Then the solution is given by

$$\tilde{\boldsymbol{C}}(\tau) = \tilde{\boldsymbol{C}}_1(\tau)\tilde{\boldsymbol{C}}_2^{-1}(\tau) \ ,$$

where $\tilde{\boldsymbol{C}}_2(\tau)$ is a solution of the initial value problem

$$\frac{\partial \tilde{\boldsymbol{C}}_2(\tau)}{\partial \tau} = \left(-\mathbf{A}_u + \mathbf{Q}_u \tilde{\boldsymbol{C}}(\tau)\right)\tilde{\boldsymbol{C}}_2(\tau) \, ,$$
$$\tilde{\boldsymbol{C}}_2(\tau) = \tilde{\boldsymbol{C}}_2(0) \ ,$$

and $\left[\tilde{\boldsymbol{C}}_2(\tau)\, , \, \tilde{\boldsymbol{C}}_1(\tau)\right]^\top$ is a solution of the associated linear system of ODEs

$$\frac{d}{dt}\begin{bmatrix} \tilde{\boldsymbol{C}}_2(\tau) \\ \tilde{\boldsymbol{C}}_1(\tau) \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_u & \mathbf{Q}_u \\ -\mathbf{P}_u & \mathbf{A}_u^\top \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{C}}_2(\tau) \\ \tilde{\boldsymbol{C}}_1(\tau) \end{bmatrix} \ ,$$

where $\tilde{\boldsymbol{C}}_2(0)$ is the initial condition, and matrices $\mathbf{A}_u$, $\mathbf{P}_u$, and $\mathbf{Q}_u$ are given by formula (2.19). To solve the matrix Riccati equation (2.35) for the constrained stochastic control problem, we apply the same numerical method. Then we solve the ODEs with respect to $\boldsymbol{b}(t)$ and $a(t)$ by utilising the fourth-order Runge-Kutta method.

For $\gamma = -1$, the numerical solutions are illustrated in Figure 3.2, Figure 3.3, and Figure 3.4 (note that the time change-of-variable has been reversed so that the system is plotted backward in time as the original problem had been posed). As we can observe clearly from these figures, the solutions $\boldsymbol{C}(t)$ and $\boldsymbol{b}(t)$ to matrix Riccati equation (2.18) and ODE (2.17) are stable, as the sufficient condition stated in Propositions 2.1, 2.4, 2.5, 2.6, 2.10, 2.11, and 2.12 are satisfied so long as $\gamma < 0$.

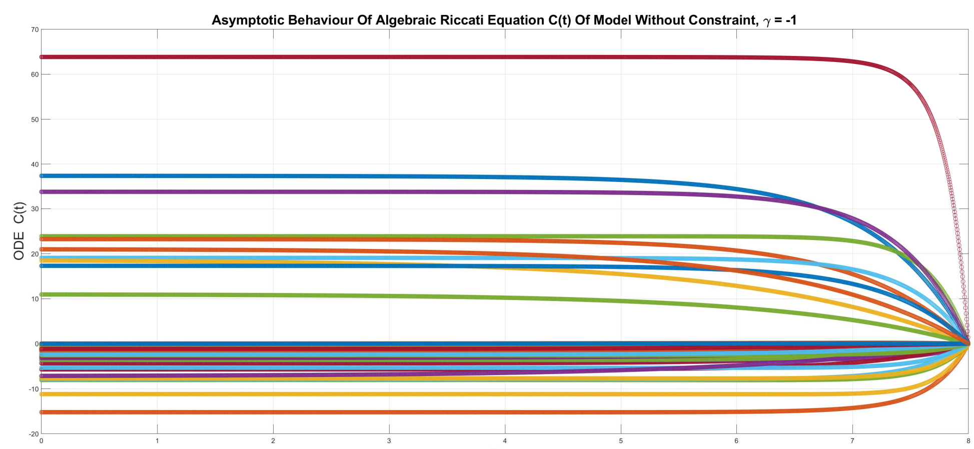

Figure 3.2: The numerical solution of the matrix Riccati equation (2.18) for $\boldsymbol{C}(t)$, for unconstrained stochastic control problem, utilising parameters estimated from the historical data. Notice the system starts to resemble its steady state if the trading time if there are 2 or more trading years. If the trading time is less, then there is significant probability that the spreads may not converge to zero by the terminal time. Hence, the statistical arbitrage strategy will have a noticeable marginal improvement if extra time is allotted.

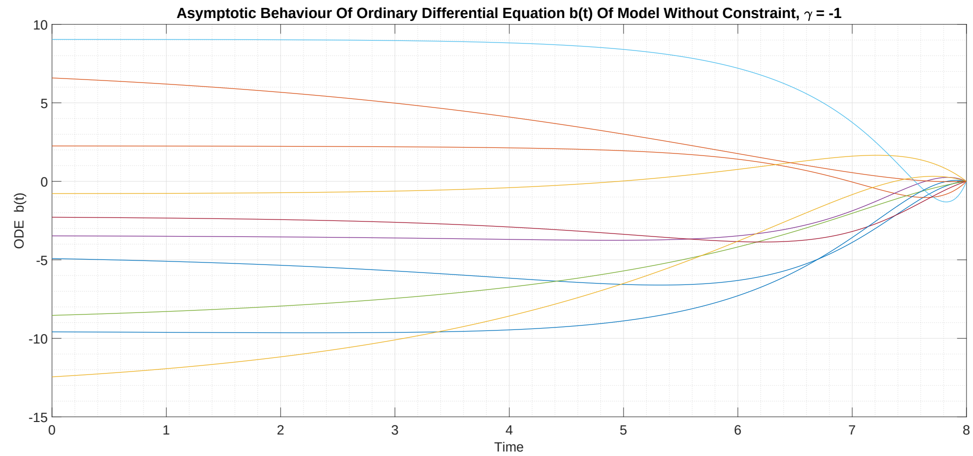

Figure 3.3: The numerical solution of linear ODE (2.17) for $\boldsymbol{b}(t)$, for the unconstrained stochastic control problem, utilising parameters estimated from the historical data. Notice the system starts to resemble its steady state if the trading time if there are 2 or more trading years. If the trading time is less, then there is significant probability that the spreads may not converge to zero by the terminal time. Hence, the statistical arbitrage strategy will have a noticeable marginal improvement if extra time is allotted.

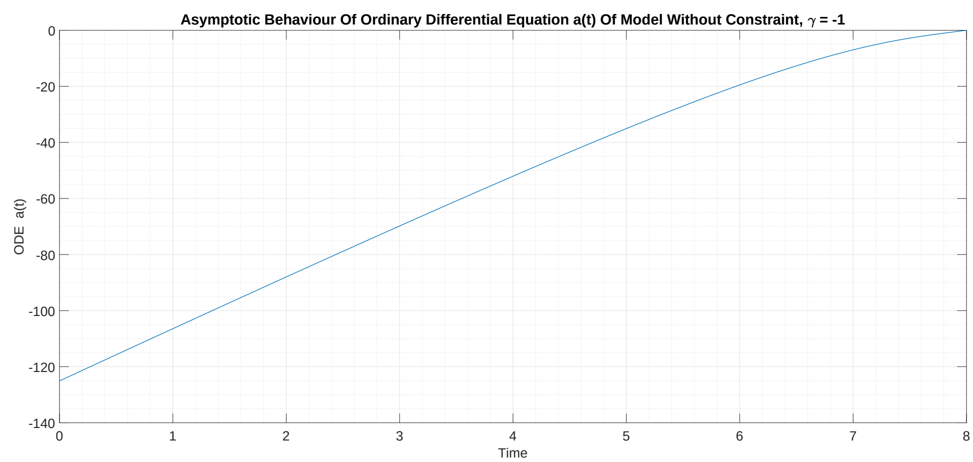

Figure 3.4: The numerical solution of linear ODE (2.18) for $a(t)$, for the unconstrained stochastic control problem, utilising parameters estimated from the historical data.

Figures 3.2, 3.3, and 3.4 give us some idea of the required time for a statistical arbitrage strategy to be effective. From the plots showing the solutions of the ODEs, we see that the co-integrated system's value function will need between one and two years before the marginal gain of extra trading time becomes log linear. If the allotted trading time for a statistical arbitrage is too short, then there is significant probability that spreads will not converge, and hence, for short time there will noticeable marginal improvement if more time is allowed.

3.2 Portfolio Performance

After we solve the system of ODEs for the unconstrained and constrained stochastic control problems, we then calculate historical wealth processes given by equation (2.5), and compare Sharpe ratios for different γ and different constraints. We work in the setting of very large terminal time T and consider a steady-state portfolio to demonstrate the results. In other words, we work with the limiting vector $\boldsymbol{\overline{\pi}}^{\,*}$ that is calculated utilising $\bar{\mathbf{C}} = \lim_{t\to-\infty} \boldsymbol{C}(t)$ and $\bar{\mathbf{b}} = \lim_{t\to-\infty} \boldsymbol{b}(t)$, e.g., for the unconstrained stochastic control problem we have

$$\boldsymbol{\overline{\pi}}^{\,*}(\mathbf{z}) = \frac{1}{1-\gamma}\boldsymbol{\Sigma}_1^{-1}\left(\boldsymbol{\mu} + \boldsymbol{\Sigma}_2\bar{\mathbf{b}} + (\boldsymbol{\delta} + 2\boldsymbol{\Sigma}_2\bar{\mathbf{C}})\mathbf{z}\right) \ .$$

Table 3.2 presents the annualised statistics for the optimal portfolio under the multivariate Merton model with no co-integration (i.e., $\boldsymbol{\pi}^{\,merton} = \frac{1}{1-\gamma}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}$), Table 3.3 presents annualised statistics for the unconstrained portfolios under the model with co-integration, and Table 3.4 presents annualised statistics for the constrained portfolios under the model with co-integration. From these tables we can see that as the risk aversion coefficient γ becomes more negative, in other words, as the arbitrageur becomes more risk averse, then there is a decrease in the expectation of excess return, the volatility of excess return, and the Sharpe ratio. Overall, the results of these numerical experiments demonstrate that the optimal portfolio with co-integration performs significantly better than a constant-coefficient Merton portfolio. We can also observe that for the same values of γ, the unconstrained portfolios in Table 3.3 have higher expected excess returns than the constrained portfolios in Table 3.4, but also have higher volatility. Note that when γ tends to zero, $\gamma \to 0$, this is the log-optimal case $\lim_{\gamma\to0}\mathbb{E}\left[\frac{1}{\gamma}(W_T^\gamma - 1)\middle|W_t = w,\ \boldsymbol{Z}_t = \mathbf{z}\right]$, but in the formula for the optimal $\boldsymbol{\pi}^*$ we can simply set $\gamma = 0$ to obtain the log optimal. Before moving on it is important to point out that the high Sharpe ratios shown in Tables 3.2, 3.3 and 3.4 are obtained from in-sample backtesting, and for real-life traders there are out-of-sample issues such as parameter error that have significant effect on portfolio performance; the data analysis shown here is merely to gain some intuition on statistical arbitrage through the lens of our multiple co-integrated model.

Finally, we see in Figure 3.5 the time series of the logarithm of the optimal wealth processes. The case considered is the model with co-integration, for the portfolio with risk aversion coefficient $\gamma = -1$, and the overall proportion of portfolio weight in the risky assets reduced to 0.5% of that prescribed by the optimal, in other words $\boldsymbol{\overline{\pi}}^{\,*} \leftarrow 0.005\times\boldsymbol{\overline{\pi}}^{\,*}$. Note that reducing the risky allocation does not affect the Sharpe ratio. From the figure we can observe that the optimal co-integration-based portfolio out-preforms the SPY ETF – even during the 4th quarter of 2015 when there was the devaluation of the Renminbi of China. The 2015 devaluation was a series of events executed surprisingly by the People's Bank of China, which drove stock markets in the United States, Europe, and Latin America into decline. The time series shown in Figure 3.5 demonstrates that the optimal portfolio under the co-integration model can have some immunity – or even generate profits – when the market has this type of downward fluctuation.

Statistics / Assets	**Optimal Wealth**				
	$\gamma \to 0$	$\gamma = -1$	$\gamma = -2$	$\gamma = -5$	$\gamma = -10$
$\mathbb{E}\left(\frac{\Delta W_t}{\Delta t W_t} - r\right)$	0.0084	0.0042	0.0028	0.0014	0.0008
$\mathbb{V}\left(\frac{\Delta W_t}{\sqrt{\Delta t} W_t}\right)$	0.0064	0.0032	0.0021	0.0011	0.0006
Sharpe Ratio	1.3139	1.3141	1.3142	1.3146	1.3152

Table 3.2: ***Merton Portfolio Statistics.*** *Annualised statistics of the multivariate Merton portfolio wealth processes without co-integration, where* $\mathbb{E}\left(\frac{\Delta W_t}{\Delta t W_t} - r\right)$ *is the expectation of excess return,* $\mathbb{V}\left(\frac{\Delta W_t}{\sqrt{\Delta t} W_t}\right)$ *is the volatility of excess return.*

Statistics / Assets	Optimal Wealth				
	$\gamma \to 0$	$\gamma = -1$	$\gamma = -2$	$\gamma = -5$	$\gamma = -10$
$\mathbb{E}\left(\frac{\Delta W_t}{\Delta t W_t} - r\right)$	0.1868	0.1544	0.1278	0.0888	0.0630
$\mathbb{V}\left(\frac{\Delta W_t}{\sqrt{\Delta t} W_t}\right)$	0.0612	0.0430	0.0348	0.0239	0.0170
Sharpe Ratio	3.0520	3.5886	3.6723	3.7223	3.7065

Table 3.3: ***Statistical-Arbitrage Portfolio Statistics (no constraints).*** *Annualised statistics of the wealth processes for unconstrained model, where* $\mathbb{E}\left(\frac{\Delta W_t}{\Delta t W_t} - r\right)$ *is the expectation of excess return,* $\mathbb{V}\left(\frac{\Delta W_t}{\sqrt{\Delta t} W_t}\right)$ *is the volatility of excess return.*

Statistics / Assets	SPY	Optimal Wealth	
$\gamma = -1$	ETF	$\boldsymbol{\pi}^\top \mathbf{1} = 0$	$\boldsymbol{\pi}^\top \boldsymbol{s} = 0$
$\mathbb{E}\left(\frac{\Delta W_t}{\Delta t W_t} - r\right)$	0.1176	0.1084	0.1449
$\mathbb{V}\left(\frac{\Delta W_t}{\sqrt{\Delta t} W_t}\right)$	0.1356	0.0354	0.0372
Sharpe Ratio	0.9156	3.0632	3.8948

Table 3.4: ***Statistical-Arbitrage Portfolio Statistics (with constraints).*** *Annualised statistics of the wealth processes for constrained model, where* $\mathbb{E}\left(\frac{\Delta W_t}{\Delta t W_t} - r\right)$ *is the expectation of excess return,* $\mathbb{V}\left(\frac{\Delta W_t}{\sqrt{\Delta t} W_t}\right)$ *is the volatility of excess return.*

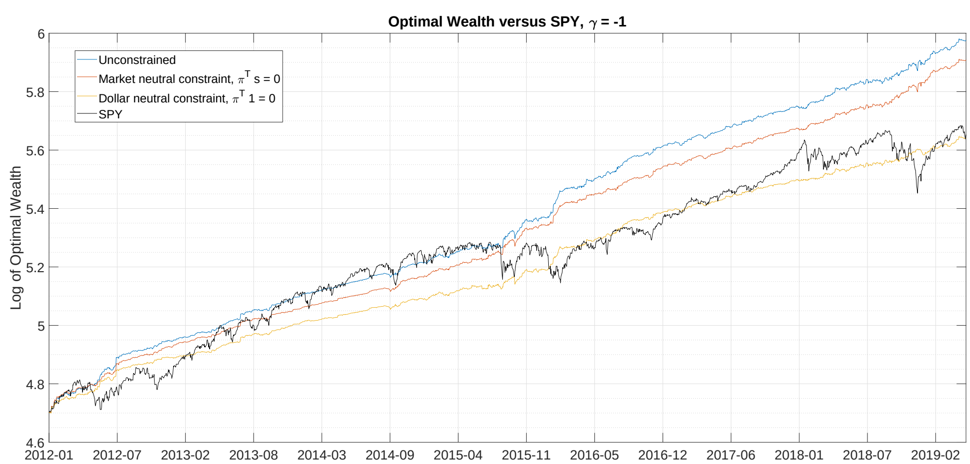

Figure 3.5: Wealth processes for different optimal portfolios.

4 Conclusion

We have presented a multi-asset model for co-integration that can be used to analyse statistical arbitrage opportunities. We compute optimal portfolios for both unconstrained stochastic control problem and constrained stochastic control problem. The optimal value functions are the solutions to Hamilton-Jacobi-Bellman equations, which for power utility function, can be solved with an exponential-affine ansatz that leads to a system of ordinary differential equations. This system consists of a matrix Riccati equation and two first-order linear ordinary differential equations. We present the stability analyses of the solutions to these differential equations for the constrained and unconstrained cases. We then apply the optimal formulae to historical data and estimate the model parameters, and find that stability in fact holds given real-life parameter estimates. We solve the ordinary differential equations and look at portfolios based on these historical parameters, from which we draw conclusions about implementation of statistical arbitrage strategies. Our first main conclusion is that, if a short trading time is allotted, then there is significant probability that the spreads will not converge, and hence the optimal value function will have noticeable marginal gain if more trading time is allotted. Our second conclusion is that there is significant in-sample outperformance of statistical arbitrage portfolios over standard multivariate Merton portfolios, but we stress parameter estimation error is likely to change this finding for out-of-sample tests.

References

[1] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank. *Matrix Riccati equations in control and systems theory*. Systems & Control: Foundations & Applications. Birkhäuser Basel, 2003.

[2] A. Al-Aradi and S. Jaimungal. Outperformance and tracking: Dynamic asset allocation for active and passive portfolio management. *Applied Mathematical Finance*, 25(3):268–294, 2018.

[3] B. Angoshtari. *Stochastic modeling and methods for portfolio management in cointegrated markets*. PhD thesis, University of Oxford, 2014.

[4] M. Avellaneda and J. H. Lee. Statistical arbitrage in the us equities market. *Quantitative Finance*, 10(7):761–782, August 2010.

[5] H. Chen, S. Chen, Z. Chen, and F. Li. Empirical investigation of an equity pairs trading strategy. *Management Science*, Articles in Advance:1–20, September 2017.

[6] M. A. Davis and S. Lleo. Risk-sensitive benchmarked asset management. *Quantitative Finance*, 8(4):415426, June 2008.

[7] M. A. Davis and S. Lleo. *Risk-sensitive investment management*, volume 19 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing, September 2014.

[8] R. J. Elliott, J. Van Der Hoek, and W. P. Malcolm. Pairs trading. *Quantitative Finance*, 5(5):271–276, June 2005.

[9] R. F. Engle and Clive W. J. Granger. Co-integration and error correction representation estimation and testing. *Econometrica*, 55(02):251–276, 1987.

[10] S. Fallahpour, H. Hakimian, K. Taheri, and E. Ramezanifar. Pairs trading strategy optimization using the reinforcement learning method: a cointegration approach. *Soft Computing*, 20(12):5051–5066, December 2016.

[11] W. Fleming and M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic modelling and applied probability*. Springer-Verlag, 2nd edition, 2006.

[12] A. Galenko, E. Popova, and I. Popova. Trading in the presence of cointegration. *Journal of Alternative Investments*, 15(1):85–97, 2012.

[13] E. G. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst. Pairs trading: performance of a relative average arbitrage rule. *The Review of Financial Studies*, 19(3):797–827, October 2006.

[14] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, New York, NY, 1991.

[15] S. Lee and A. Papanicolaou. Pairs trading of two assets with uncertainty in co-integration's level of mean reversion. *International Journal of Theoretical and Applied Finance*, 19(8), November 2016.

[16] Y. Lei and J. Xu. Costly arbitrage through pairs trading. *Journal of Economic Dynamics and Control*, 56:1–19, July 2015.

[17] T. Leung and X. Li. Optimal mean reversion trading with transaction costs and stoploss exit. *International Journal of Theoretical and Applied Finance*, 18(3), May 2015.

[18] T. N. Li and A. Tourin. Optimal pairs trading with time-varying volatility. *International Journal of Financial Engineering*, 03(02), 2016.

[19] S. Mudchanatongsuk, J. A. Primbs, and W. Wong. Optimal pairs trading: a stochastic control approach. In *Proceedings of 2008 American Control Conference*. Institute of Electrical and Electronics Engineers, June 2008.

[20] A. Tourin and R. Yan. Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control*, 37(10):1947–2156, 2013.

[21] G. Vidyamurthy. *Pairs trading quantitative methods and analysis*. Wiley Finance. John Wiley & Sons, Inc., Hoboken, New Jersey, U.S.A., August 2004.

[22] H. Wielandt. On the eigenvalues of a+b and ab. *Journal of Research of the National Bureau of Standards, Section B: Mathematical Sciences*, 77B(1 and 2):61–72, January 1973.

[23] W. M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM Journal on Control*, 6(4):681–697, 1968.