Fu_Yijun_02073515.pdf

by Yijun Fu

Imperial College London

Department of Mathematics

Dynamic Portfolio Optimisation under Multiscale Stochastic Structures

Author: Yijun Fu (CID: 02073515)

A thesis submitted for the degree of

MSc in Mathematics and Finance, 2021-2022

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Acknowledgements

The completion of this paper labels the end of my educational journey lasting over 17 years. The year I spent in London was unforgettable and rewarding, during which I challenged myself to comprehensively study math and work in a totally different environment, and made a lot of friends coming from all over the world. All those experiences, however, were based on supports from my family and my girlfriend's family in China, so I would like to say a special thank to them.

For this paper, my gratification is given towards my supervisor Dr. Engel John C. Dela Vega, who has been always friendly and super supportive, not only academically, but a great mentor in many aspects of life. I feel honored to have chosen you as my supervisor and this project as my thesis to farewell my graduate studies.

My appreciation also goes to my internship supervisor Peter Freeborough, and my colleagues including Leo Poisot, Ben Huang, Cheng Luo, Teppei Kubo, Giuseppe Argentieri, Akanksh Vashisth, Dmitry Proselkov, Alfred Besong, and Ivan Milokhin, at Barclays who have taught me a lot not only in terms of credit products, but also great deal of philosophy of working and living.

I would like to specially thank my best friend in London, Daivid Cai, and my best friend in Beijing, Wei He, for sharing my happiness and sadness together and always being supportive for my decisions.

My girlfriend Sumin Jiang, who studied at LSE, has been accompanying me for over four years. She always encourages me to pursue whatever I want and is always my source of power to continue to explore the world. Thank you Sumin for being in my life and I really look forward to the life with you in London in the upcoming several years and wherever on the earth in my rest of my life!

Lastly, I would like to extend my appreciation to my friends including but not limited to Mingyu Zhang, Shikang Du, Kaidi Xiang, Kexin Xie, and Junwei Yuan, for supporting me all the time and pushing me to be a better man.

Abstract

In this thesis, we propose a two-factor stock price model and construct a dynamic wealth allocation scheme to form a portfolio through the timeline which aims to minimise the quadratic risk consisting of the running cost and the variance of the terminal total wealth at a targeted level, where the two factors are slowly varying and fast varying. To solve the scheme from the setup, we apply four different methods: the primal HJB equation, the dual HJB equation, the primal FBSDEs, and the dual FBSDEs, and we also prove their equivalences theoretically. In addition, we also analyse the simplified cases theoretically with the assumption that the running cost is zero, including introducing the correlation structure of the Brownian motions and the convex cone constraint on the optimal controls or the amount allocated to each stock at any time spot. Furthermore, since the non-linear system can be barely solved analytically, we apply the finite difference scheme to solve the PDEs arising from the primal and dual HJB equations to verify their equivalence numerically, and we also transform the problem of solving the primal and dual FBSDEs into a problem solving two corresponding optimisation problems in order not to break the nonanticipativeness. However, due to the lack of computing powers and the high dimensional nature of our problem, we are unable to fully solve the transformed optimisation problems with satisfactory speed. Nevertheless, we verify the effectiveness of FBSDE methods by evaluating the optimisation objectives with the initial guesses produced by utilising the links between these two FBSDE methods, which supports the effectiveness of the methods. Lastly, we test the numerical schema on two settings with different scales to test the stability of our algorithm.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The factor investment idea can be dated back to Fama and French three-factor model. The nature of the specification is that we are assuming there exists a factor structure where the value of factors can impact the prices of stocks. Nevertheless, the factors themselves can vary over time, for example, the GDP growth rate, the inflation rate, the ROE ratio, the net profit rate, etc., and this incentivised us to parameterise their dynamics. The factors, generally, can be roughly divided into two groups: the slowly varying factors and the fast varying factors. Therefore, we simplify our discussion by only specifying two factors according to Fouque & Hu (2020): one is slowly varying and the other is fast varying. Meanwhile, the so called factor structure can be also embedded into the dynamics of prices of stocks by making the drift term and the local volatility term dependent on the factors, except for different functional forms of different stocks.

With the specification of the dynamics of the factors and the factor structures, we are able to make a dynamic decision of allocating the wealth into different stocks at each time spot to minimise the quadratic risk defined by Li & Zheng (2018): the sum of the running costs and the quadratic risk of the terminal total wealth. This objective is intuitive: the first component can be interpreted as one's preference to less exposure from risky assets during the time period, and according to Hu & Zhou (2006), the second component can be treated as the variance of the terminal total wealth with a targeted level. Clearly we prefer a stabler growth and thus a stabler closer-to-the-target terminal total wealth. Given the objective, we are able to establish the framework to solve it. Li & Zheng (2018) provide a thorough discussion about the way to establish the primal Hamilton-Jacobi-Bellman (HJB) equation, the dual HJB equation, the primal forward-backward stochastic differential equations (FBSDEs) and the dual FBSDEs and give the link between the two FBSDE methods. Ma et al. (2020) provide a more detailed description for the strong duality, and, more importantly, demonstrate a concrete example for the single factor dynamic portfolio optimisation through the dual HJB equation method, which gives us an anchor to construct our double factor dual process. Yong & Zhou (1999) and Rockafellar (1970) provide detailed theoretical backgrounds of the FBSDE method. Our work is based on those works, extending to the two factor dynamic portfolio optimisation context with the objective of minimising the quadratic risk and solving the problem both theoretically and numerically by the primal HJB equation, the dual HJB equation, the primal FBSDEs, and the dual FBSDEs. In addition, we analyse the simplified case where the running cost is set to zero, and under this setting, we further discuss the impact of introducing the correlation structure of the Brownian motions and the impact of introducing the convex cone to the optimal controls (i.e. the wealth we should allocate to each stock) theoretically.

However, it is generally unrealistic to fully solve the problem by either method by hand, which motivated us to dig into the numerical solution of HJB-type partial differential equations (PDEs) and FBSDEs, and simulations of the optimal total wealth paths. When solving the HJB-type PDEs arising from the primal and dual HJB equations by the finite difference scheme, we faced the divergence issue due to the lack of refinement of the time discretisation. LeVeque (1998) discusses a rule to mitigate the divergence issue by providing the rough number of time intervals we should make given the discretisations of the two factors dimensions of the value functions. In addition to the primal and dual HJB methods, Zhu & Zheng (2022) provide a detailed description of transforming the problem of solving the FBSDEs into a problem of solving an equivalent optimisation problem without breaking the nonanticipativeness of the solution. However, due to the restrictions of computing powers and the high dimensional nature of our problem, we are unable to solve the optimisation with relatively acceptable speed. Nevertheless, we try to verify FBSDE method's validity by evaluating the optimisation objective using the initial guesses by utilising the links mentioned in Li & Zheng (2018).

The rest of the thesis is organised as follows: Chapter 2 describes the details of the model setup and formulates the quadratic risk optimisation problem; Chapter 3 solves the primal HJB equation into three PDEs using the quadratic ansatz; Chapter 4 derives and solves the dual HJB equation also into three PDEs using the similar quadratic ansatz, and verifies the equivalence between the primal and the dual HJB equations theoretically; Chapter 5 discusses three special cases theoretically under the assumption that the running cost is zero including introducing the correlation structures of Brownian motion and the convex cone constraint; Chapter 6 elaborates the primal and dual FBSDEs and states the links between them; Chapter 7 establishes the numerical methods for solving the primal and dual HJB equation, transforms the problem of solving FBSDEs into a problem of solving optimisation problems, and presents the numerical results under two settings with different scale to test the equivalence and the stability of the algorithm.

Chapter 2

Model Setup

We are considering a portfolio optimisation problem driven by a multiscale factor framework inspired by Fouque & Hu (2020). More specifically, there are two factors, η_t and ζ_t, that drive the movement, including the drift term and the local volatility, of the prices of the stocks (N stocks in total), while these two factors are slowly varying and fast varying and follow two separate stochastic processes. We summarise them as follows:

$$\begin{array}{ll} \mathrm{d}S_{n,t} = \mu_n(\eta_t,\zeta_t)S_{n,t}\mathrm{d}t + \sum_{m=1}^N \sigma_{nm}(\eta_t,\zeta_t)S_{n,t}\mathrm{d}W_{m,t}, & S_{n,0} > 0, \\ \mathrm{d}\eta_t = \frac{1}{\epsilon}b(\eta_t)\mathrm{d}t + \frac{1}{\sqrt{\epsilon}}a(\eta_t)\mathrm{d}W_t^{\eta}, & \eta_0 > 0, \\ \mathrm{d}\zeta_t = \delta c(\zeta_t)\mathrm{d}t + \sqrt{\delta}l(\zeta_t)\mathrm{d}W_t^{\zeta}, & \zeta_0 > 0, \end{array} \tag{2.0.1}$$

where $W_{n,t}$, W_t^{η} and W_t^{ζ} are mutually independent. The slowly varying and the fast varying structures are controlled by two constants: ϵ and δ, both of which are sufficiently small. Thus, we can imagine that η_t is the fast varying factor as $1/\epsilon$ significantly scales up its infinitesimal changes, whereas ζ_t is the slowly varying factor as δ significantly scales down its infinitesimal changes. Intuitively, we can treat the slowly varying factor as a long-term (e.g. monthly) driver, and the fast varying factor as a short-term (e.g. daily) driver.

From now on, to simplify our discussion, we denote the drift terms by:

$$\boldsymbol{\mu} = [\mu_1, \cdots, \mu_N]',$$

the local volatility term by:

$$\boldsymbol{\sigma} = (\sigma_{nm})_{N\times N},$$

and the collection of the mutually independent Brownian motions that drive the stock prices by:

$$\boldsymbol{W}_t = [W_{1,t}, \cdots, W_{N,t}]'.$$

We base our discussion on the following settings: let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be the filtered probability space. $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the natural filtration generated by $(\boldsymbol{W}_t, W_t^{\eta}, W_t^{\zeta})$ and is augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$.

Assuming we invest $\boldsymbol{\pi}_t = [\pi_{1,t}, \cdots, \pi_{N,t}]'$ in the N stocks and the total wealth is denoted by X_t^{π} (so the amount invested in the bank account is $X_t^{\pi} - \boldsymbol{\pi}_t'\mathbf{1}$), we then can directly give the dynamics of the total wealth:

$$\begin{aligned} \mathrm{d}X_t^{\boldsymbol{\pi}} =& \mathrm{d}(X_t^{\pi} - \boldsymbol{\pi}_t'\mathbf{1}) + \mathrm{d}(\boldsymbol{\pi}_t'\mathbf{1}) \\ =& r_t(X_t^{\pi} - \boldsymbol{\pi}_t'\mathbf{1})\mathrm{d}t + \boldsymbol{\pi}_t'[\boldsymbol{\mu}(\eta_t,\zeta_t)\mathrm{d}t + \boldsymbol{\sigma}(\eta_t,\zeta_t)\mathrm{d}\boldsymbol{W}_t] \\ =& \left[X_t^{\boldsymbol{\pi}} r_t + \boldsymbol{\pi}_t'\left(\boldsymbol{\mu}(\eta_t,\zeta_t) - r_t\mathbf{1}\right)\right]\mathrm{d}t + \boldsymbol{\pi}_t'\boldsymbol{\sigma}(\eta_t,\zeta_t)\mathrm{d}\boldsymbol{W}_t, \quad X_0^{\pi} = x_0, \end{aligned} \tag{2.0.2}$$

where r_t is the short rate which can be varied by time. Our objective is to minimise the risk subject to the admissibility of $(X_t, \eta_t, \zeta_t, \boldsymbol{\pi}_t)$, which is equivalent to solving the value function $H(t,x,\eta,\zeta)$ (especially, $\boldsymbol{\pi}_t \in K$) defined as:

$$H(t,x,\eta,\zeta) = \min_{\boldsymbol{\pi}\in K} \mathbb{E}\left\{\int_t^T f(t, X_t^{\boldsymbol{\pi}}, \boldsymbol{\pi}_t)\mathrm{d}t + g(X_T^{\boldsymbol{\pi}}) | X_t^{\boldsymbol{\pi}} = x, \eta_t = \eta, \zeta_t = \zeta\right\}, \tag{2.0.3}$$

where:

$$f(t, x, \boldsymbol{\pi}) = \frac{1}{2} Q_t x^2 + \boldsymbol{S}_t' \boldsymbol{\pi} x + \frac{1}{2} \boldsymbol{\pi}' \boldsymbol{R}_t \boldsymbol{\pi},$$
$$g(x) = \frac{1}{2} A x^2 + B x,$$

$Q_t \in \mathbb{R}$, $\boldsymbol{S}_t \in \mathbb{R}^N$, and $\boldsymbol{R}_t \in \mathbb{R}^{N \times N}$ are $\mathcal{F}_t$-measurable functions such that

$$\begin{pmatrix} Q_t & \boldsymbol{S}_t' \\ \boldsymbol{S}_t & \boldsymbol{R}_t \end{pmatrix}$$

is non-negative definite for all $(\omega, t) \in \Omega \times [0, T]$, the matrix $\boldsymbol{R}_t$ is symmetric, A is a positive constant, and B is a constant.

The intuition behind the optimisation objective (or the risk metric) can be interpreted in two parts. The first part is the expectation of the integral, which is called the running cost of the strategy. The running cost, for example, can be treated as a penalty for large exposure to risky assets if we make Q and $\boldsymbol{S}$ equal to zero. The second part is the expectation of the quadratic function of the terminal value of the total wealth. Inspired by Hu & Zhou (2006, Section 6), it can be treated as a penalty for large variance of the terminal total wealth. More specifically, it can be derived from the classical mean-variance optimisation (i.e. given a level of the expectation of the terminal total wealth, minimising the variance of the terminal total wealth) by transforming it into an unconstrained problem parameterised by the Lagrange multiplier.

Chapter 3

Primal HJB Equation

With the formalised optimisation problem, we are able to derive the HJB equation directly from the dynamic programming principle:

$$\inf_{\boldsymbol{\pi}} \left\{ \frac{1}{2}\boldsymbol{\pi}'(\boldsymbol{R} + \partial_X^2 H \boldsymbol{\sigma}\boldsymbol{\sigma}')\boldsymbol{\pi} + \left[X\boldsymbol{S}' + \partial_X H(\boldsymbol{\mu} - r\mathbf{1})'\right]\boldsymbol{\pi} \right\}$$
$$+\partial_t H + Xr\partial_X H + \frac{1}{2}QX^2 + \mathcal{L}_\eta^\epsilon H + \mathcal{L}_\zeta^\delta H = 0,$$

with the terminal condition:

$$H(T, X, \eta, \zeta) = \frac{1}{2}AX^2 + BX,$$

where:

$$\mathcal{L}_\eta^\epsilon = \frac{b}{\epsilon}\partial_\eta + \frac{a^2}{2\epsilon}\partial_\eta^2,$$
$$\mathcal{L}_\zeta^\delta = \delta c \partial_\zeta + \frac{1}{2}\delta l^2 \partial_\zeta^2.$$

To simplify our discussion, we assume that $K = \mathbb{R}^N$ from now on. According to the first order condition, the optimal control $\boldsymbol{\pi}^*$ is obtained (assuming that $(\boldsymbol{R} + \partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}')$ is invertible):

$$\begin{aligned} &\left(\boldsymbol{R} + \partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)\boldsymbol{\pi}^* + \left[X\boldsymbol{S} + \partial_X H(\boldsymbol{\mu} - r\mathbf{1})\right] = 0 \\ \Longrightarrow \quad &\boldsymbol{\pi}^* = -\left(\boldsymbol{R} + \partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left[X\boldsymbol{S} + \partial_X H(\boldsymbol{\mu} - r\mathbf{1})\right], \end{aligned} \tag{3.0.1}$$

and therefore, the PDE for $H(t, x, \eta, \zeta)$ is:

$$\begin{aligned} &-\frac{1}{2}\left[X\boldsymbol{S} + \partial_X H(\boldsymbol{\mu} - r\mathbf{1})\right]'\left(\boldsymbol{R} + \partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left[X\boldsymbol{S} + \partial_X H(\boldsymbol{\mu} - r\mathbf{1})\right] \\ &\qquad +\partial_t H + Xr\partial_X H + \frac{1}{2}QX^2 + \mathcal{L}_\eta^\epsilon H + \mathcal{L}_\zeta^\delta H = 0, \end{aligned} \tag{3.0.2}$$

with the terminal condition:

$$H(T, X, \eta, \zeta) = \frac{1}{2}AX^2 + BX.$$

To solve it, we can use the ansatz:

$$H(t, X, \eta, \zeta) = v_0(t, \eta, \zeta) + v_1(t, \eta, \zeta)X + v_2(t, \eta, \zeta)X^2,$$

to establish a group of PDEs for v_0, v_1 and v_2 by substituting the ansatz into (3.0.2), grouping terms in order of power of X, and equating all the three coefficients of each power of X to zeors. Here is the results:

$$\begin{cases} \mathcal{D}v_2 + 2v_2 r + \frac{1}{2}Q - \frac{1}{2}\left[\boldsymbol{S} + 2v_2(\boldsymbol{\mu} - r\mathbf{1})\right]'\left(\boldsymbol{R} + 2v_2\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left[\boldsymbol{S} + 2v_2(\boldsymbol{\mu} - r\mathbf{1})\right] = 0, \\ \mathcal{D}v_1 + \left\{r - (\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{R} + 2v_2\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}[\boldsymbol{S} + 2v_2(\boldsymbol{\mu} - r\mathbf{1})]\right\}v_1 = 0, \\ \mathcal{D}v_0 - \frac{1}{2}(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{R} + 2v_2\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})v_1^2 = 0, \end{cases} \tag{3.0.3}$$

where

$$\mathcal{D} = \partial_t + \mathcal{L}_\eta^\epsilon + \mathcal{L}_\zeta^\delta,$$

and the terminal conditions are:

$$\begin{cases} v_0(T,\eta,\zeta) = 0, \\ v_1(T,\eta,\zeta) = B, \\ v_2(T,\eta,\zeta) = \dfrac{A}{2}. \end{cases}$$

Finally, by taking advantage of the ansatz above, the optimal control $\boldsymbol{\pi}^*$ can be expressed as follows:

$$\boldsymbol{\pi}^* = -(\boldsymbol{R} + 2v_2\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\left\{X[\boldsymbol{S} + 2v_2(\boldsymbol{\mu} - r\mathbf{1})] + v_1(\boldsymbol{\mu} - r\mathbf{1})\right\}. \tag{3.0.4}$$

It is noteworthy that the optimal control $\boldsymbol{\pi}^*$ does not depend on v_0, which means if we do not care about the precise form of the value function H and only want to know the optimal control $\boldsymbol{\pi}^*$ and the path of the total wealth, we do not need to solve v_0 explicitly. In fact, we will see in all following discussions, we normally do not have to solve the PDEs in full to obtain the optimal control(s) and the corresponding path(s) of state variable(s) in both primal HJB equation and dual HJB equation.

Chapter 4

Dual HJB Equation

Solving the primal HJB equation is sometimes not the easiest way, while using the duality may simplify the procedure. Although under this setting, the dual HJB equation is not obviously easier to solve, for the completeness of the theoretical analysis, we derive it in this section, and finally we will be able to verify its equivalence to the primal HJB equation method both theoretically and numerically.

4.1 Derivation of Dual HJB Equation

In this section, we will derive the dual HJB equation following the same logic in Li & Zheng (2018, Remark 1, page 1135) by constructing a supermartingale, utilising its property to derive a lower bound of the primal value function, and linking the primal and the dual value functions by the strong duality due to Ma et al. (2020), which is a more intuitive way to comprehend. We firstly define the dual process as follows:

$$\mathrm{d}Y_t = \alpha_{1t}\mathrm{d}t + \boldsymbol{\beta}_{1t}'\mathrm{d}\boldsymbol{W}_t + \gamma_t \mathrm{d}W_t^{\eta} + \xi_t \mathrm{d}W_t^{\zeta}, \quad Y_0 = y_0, \tag{4.1.1}$$

where the coefficients are chosen to make the stochastic process

$$X_t^{\pi} Y_t - \int_0^t (X_s^{\pi}\alpha_s + \boldsymbol{\pi}_s'\boldsymbol{\beta}_s)\mathrm{d}s$$

be a supermartingale. By Itô's formula, we have:

$$\begin{aligned}\mathrm{d}(X_t^{\pi}Y_t) &= \{(r_tY_t + \alpha_{1t})X_t^{\pi} + \boldsymbol{\pi}_t'[Y_t(\boldsymbol{\mu} - r_t\mathbf{1}) + \boldsymbol{\sigma}\boldsymbol{\beta}_{1t}]\}\,\mathrm{d}t + \textit{local martingale}\\ &= (X_t^{\pi}\alpha_t + \boldsymbol{\pi}_t'\boldsymbol{\beta}_t)\mathrm{d}t + \textit{local martingale},\end{aligned}$$

and we thus have:

$$\begin{aligned}\alpha_{1t} &= \alpha_t - r_tY_t,\\ \boldsymbol{\beta}_{1t} &= \boldsymbol{\sigma}^{-1}[\boldsymbol{\beta}_t - Y_t(\boldsymbol{\mu} - r_t\mathbf{1})].\end{aligned}$$

Furthermore, by the definition of supermartingale, we have:

$$\mathbb{E}\left\{X_T^{\pi}Y_T - \int_0^T (X_s^{\pi}\alpha_s + \boldsymbol{\pi}_s'\boldsymbol{\beta}_s)\mathrm{d}s\right\} \le X_0^{\pi}Y_0 = x_0y_0. \tag{4.1.2}$$

To internalise the constraint of the control, we define a penalty function:

$$\Psi_K(\boldsymbol{\pi}) = \begin{cases} +\infty, & \boldsymbol{\pi} \notin K,\\ 0, & \boldsymbol{\pi} \in K,\end{cases}$$

so we have an equivalence:

$$\begin{aligned}&\min_{\boldsymbol{\pi}\in K}\mathbb{E}\left\{\int_0^T f(t, X_t^{\pi}, \boldsymbol{\pi}_t)\mathrm{d}t + g(X_T^{\pi})\right\}\\ \Longleftrightarrow\ &\max_{\boldsymbol{\pi}}\mathbb{E}\left\{\int_0^T [-f(t, X_t^{\pi}, \boldsymbol{\pi}_t) - \Psi_K(\boldsymbol{\pi}_t)]\,\mathrm{d}t - g(X_T^{\pi})\right\}.\end{aligned}$$

In addition, we define:

$$\begin{aligned}\phi(t,\alpha,\boldsymbol{\beta}) &= \sup_{x,\boldsymbol{\pi}}\left\{-f(t,x,\boldsymbol{\pi})-\Psi_K(\boldsymbol{\pi})+x\alpha+\boldsymbol{\pi}'\boldsymbol{\beta}\right\},\\ m_T(y) &= \sup_x\left\{-g(x)-xy\right\}.\end{aligned}$$

Finally, using (4.1.2), we have:

$$\begin{aligned}&\max_{\boldsymbol{\pi}}\mathbb{E}\left\{\int_0^T\left[-f(t,X_t^{\pi},\boldsymbol{\pi}_t)-\Psi_K(\boldsymbol{\pi}_t)\right]\mathrm{d}t-g(X_T^{\pi})\right\}\\ \leq &\min_{y_0,\alpha,\boldsymbol{\beta},\gamma,\xi}\left\{x_0y_0+\mathbb{E}\left[\int_0^T\phi(t,\alpha_t,\boldsymbol{\beta}_t)\mathrm{d}t+m_T(Y_T)\right]\right\}.\end{aligned}$$

Given the above results, we are able to define the dual value function by treating the latter term in above expression as a function of y:

$$\tilde{H}(t,y,\eta,\zeta)=\min_{\alpha,\boldsymbol{\beta},\gamma,\xi}\mathbb{E}\left\{\int_t^T\phi(t,\alpha_t,\boldsymbol{\beta}_t)\mathrm{d}t+m_T(Y_T)|Y_t=y,\eta_t=\eta,\zeta_t=\zeta\right\}. \tag{4.1.3}$$

According to (Ma et al. 2020, Theorem 2.1, page 7), we have the strong duality and thus are able to link the primal and dual value functions as follows:

$$\begin{aligned}H(t,x,\eta,\zeta) &= -\min_y\left[\tilde{H}(t,y,\eta,\zeta)+xy\right]\\ &= \max_y\left[-\tilde{H}(t,y,\eta,\zeta)-xy\right].\end{aligned} \tag{4.1.4}$$

The optimal $y=y(t,x,\eta,\zeta)$ satisfies:

$$-x-\partial_y\tilde{H}(t,y(t,x,\eta,\zeta),\eta,\zeta)=0, \tag{4.1.5}$$

and thus, (4.1.4) can be expressed as:

$$H(t,x,\eta,\zeta)=-\tilde{H}(t,y(t,x,\eta,\zeta),\eta,\zeta)-xy(t,x,\eta,\zeta). \tag{4.1.6}$$

Following (4.1.5), we can have some further conclusions:

$$\begin{aligned}&-1-\partial_Y^2\tilde{H}\partial_X y=0\\ \Longrightarrow\ &\partial_X y=-\frac{1}{\partial_Y^2\tilde{H}},\\ &-\partial_Y\partial_\eta\tilde{H}-\partial_Y^2\tilde{H}\partial_\eta y=0\\ \Longrightarrow\ &\partial_\eta y=-\frac{\partial_Y\partial_\eta\tilde{H}}{\partial_Y^2\tilde{H}},\\ &-\partial_Y\partial_\zeta\tilde{H}-\partial_Y^2\tilde{H}\partial_\zeta y=0\\ \Longrightarrow\ &\partial_\zeta y=-\frac{\partial_Y\partial_\zeta\tilde{H}}{\partial_Y^2\tilde{H}}.\end{aligned}$$

It is straightforward to obtain the following relationship by differentiating (4.1.6) in terms of t, x, η, and ζ:

$$
\begin{aligned}
\partial_t H =& -\partial_t \tilde{H}, \\
\partial_X H =& -y, \\
\partial_X^2 H =& -\partial_X y, \\
\partial_\eta H =& -\partial_\eta \tilde{H}, \\
\partial_\eta^2 H =& -\partial_\eta^2 \tilde{H} - \partial_Y \partial_\eta \tilde{H} \partial_\eta y \\
=& -\partial_\eta^2 \tilde{H} + \frac{(\partial_Y \partial_\eta \tilde{H})^2}{\partial_Y^2 \tilde{H}}, \\
\partial_\zeta H =& -\partial_\zeta \tilde{H}, \\
\partial_\zeta^2 H =& -\partial_\zeta^2 \tilde{H} - \partial_Y \partial_\zeta \tilde{H} \partial_\zeta y \\
=& -\partial_\zeta^2 \tilde{H} + \frac{(\partial_Y \partial_\zeta \tilde{H})^2}{\partial_Y^2 \tilde{H}}.
\end{aligned}
\tag{4.1.7}
$$

Thus, we know that the optimal initial value of Y (i.e. y_0^*) is:

$$
\begin{aligned}
y_0^* =& -\partial_X H(0, x_0, \eta_0, \zeta_0) \\
=& -v_1(0, \eta_0, \zeta_0) - 2v_2(0, \eta_0, \zeta_0) x_0.
\end{aligned}
\tag{4.1.8}
$$

To simplify the discussion, we now make $K = \mathbb{R}^N$ and we obtain the HJB equation for $\tilde{H}$ by dynamic programming principle:

$$
\begin{aligned}
\inf_{\alpha,\boldsymbol{\beta},\gamma,\xi} \Bigg\{ & \frac{1}{2}\partial_Y^2 \tilde{H} \Big[[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})] + \gamma^2 + \xi^2 \Big] \\
& + (\alpha - rY)\partial_Y \tilde{H} + \sup_{x,\boldsymbol{\pi}} [-f(t, x, \boldsymbol{\pi}) + x\alpha + \boldsymbol{\pi}'\boldsymbol{\beta}] \\
& + \partial_Y \partial_\eta \tilde{H} \frac{1}{\sqrt{\epsilon}} a\gamma + \partial_Y \partial_\zeta \tilde{H} \sqrt{\delta} l \xi \Bigg\} + \partial_t \tilde{H} + \mathcal{L}_\eta^\epsilon \tilde{H} + \mathcal{L}_\zeta^\delta \tilde{H} = 0,
\end{aligned}
\tag{4.1.9}
$$

with the terminal condition:

$$
\tilde{H}(T, Y, \eta, \zeta) = m_T(Y).
$$

We firstly solve the maximisation inside the HJB equation by solving the first order condition:

$$
\begin{cases}
-Qx - \boldsymbol{S}'\boldsymbol{\pi} + \alpha = 0, \\
-x\boldsymbol{S} - \boldsymbol{R}\boldsymbol{\pi} + \boldsymbol{\beta} = 0,
\end{cases}
$$

which gives:

$$
\begin{aligned}
x =& \frac{\alpha - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}}, \\
\boldsymbol{\pi} =& \frac{\boldsymbol{R}^{-1}\boldsymbol{S}\boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta} - \alpha\boldsymbol{R}^{-1}\boldsymbol{S}}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \boldsymbol{R}^{-1}\boldsymbol{\beta}.
\end{aligned}
\tag{4.1.10}
$$

and:

$$
\phi(t, \alpha, \boldsymbol{\beta}) = \frac{1}{2}\frac{(\alpha - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta})^2}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \frac{1}{2}\boldsymbol{\beta}'\boldsymbol{R}^{-1}\boldsymbol{\beta}.
\tag{4.1.11}
$$

If $\partial_Y^2 \tilde{H} > 0$, the optimal η^* and ζ^* should be:

$$
\begin{aligned}
\eta^* =& -\frac{a}{\sqrt{\epsilon}} \frac{\partial_Y \partial_\eta \tilde{H}}{\partial_Y^2 \tilde{H}}, \\
\zeta^* =& -\sqrt{\delta} l \frac{\partial_Y \partial_\zeta \tilde{H}}{\partial_Y^2 \tilde{H}}.
\end{aligned}
\tag{4.1.12}
$$

It can be noted that $\partial_Y^2 \tilde{H}(T, Y, \eta, \zeta) > 0$, which serves as a necessary condition for the positiveness of the second order derivative of $\tilde{H}$ with respect to Y across the whole timeline, and it can be verified by the following:

$$
\begin{aligned}
\tilde{H}(T, y, \eta, \zeta) =& m_T(y) \\
=& \sup_x \left\{ -\frac{1}{2} A x^2 - Bx - xy \right\} \\
=& -\frac{1}{2} A \left(-\frac{B+y}{A} \right)^2 - (B+y) \left(-\frac{B+y}{A} \right) \\
=& \frac{y^2 + 2By + B^2}{2A}.
\end{aligned}
$$

By substituting (4.1.10) and (4.1.12) into the HJB equation for $\tilde{H}$, we have:

$$
\begin{aligned}
&\inf_{\alpha, \boldsymbol{\beta}} \Bigg\{ \frac{1}{2} \frac{\alpha^2}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \alpha \partial_Y \tilde{H} + \boldsymbol{\beta}' \Big[\frac{1}{2} \partial_Y^2 \tilde{H} (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} + \frac{1}{2} \frac{\boldsymbol{R}^{-1}\boldsymbol{S}\boldsymbol{S}'\boldsymbol{R}^{-1}}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \frac{1}{2}\boldsymbol{R}^{-1} \Big] \boldsymbol{\beta} \\
&\qquad - \Big[\alpha \frac{\boldsymbol{S}'\boldsymbol{R}^{-1}}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \partial_Y^2 \tilde{H} Y (\boldsymbol{\mu} - r\mathbf{1})' (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} \Big] \boldsymbol{\beta} \Bigg\} \\
&\qquad - \frac{1}{2} \frac{(\partial_Y \partial_\eta \tilde{H})^2 \frac{a^2}{\epsilon}}{\partial_Y^2 \tilde{H}} - \frac{1}{2} \frac{(\partial_Y \partial_\zeta \tilde{H})^2 l^2 \delta}{\partial_Y^2 \tilde{H}} \\
&- rY \partial_Y \tilde{H} + \frac{1}{2} \partial_Y^2 \tilde{H} Y^2 (\boldsymbol{\mu} - r\mathbf{1})' (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} (\boldsymbol{\mu} - r\mathbf{1}) + \partial_t \tilde{H} + \mathcal{L}_\eta^\epsilon \tilde{H} + \mathcal{L}_\zeta^\delta \tilde{H} = 0,
\end{aligned}
$$

with the terminal condition:

$$
\tilde{H}(T, Y, \eta, \zeta) = \frac{Y^2 + 2BY + B^2}{2A}.
$$

The way to compute the optimal α^* and $\boldsymbol{\beta}^*$ is theoretically simple but practically tedious, so we just present the results here:

$$
\begin{aligned}
\boldsymbol{\beta}^* =& \big[\boldsymbol{R}^{-1} + \partial_Y^2 \tilde{H} (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} \big]^{-1} [\partial_Y^2 \tilde{H} Y (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} (\boldsymbol{\mu} - r\mathbf{1}) - \partial_Y \tilde{H} \boldsymbol{R}^{-1} \boldsymbol{S}], \\
\alpha^* =& \boldsymbol{S}'\boldsymbol{R}^{-1} \big[\boldsymbol{R}^{-1} + \partial_Y^2 \tilde{H} (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} \big]^{-1} [\partial_Y^2 \tilde{H} Y (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} (\boldsymbol{\mu} - r\mathbf{1}) - \partial_Y \tilde{H} \boldsymbol{R}^{-1} \boldsymbol{S}] \\
&- \partial_Y \tilde{H} (Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}).
\end{aligned}
\tag{4.1.13}
$$

We can then substitute them into the HJB equation for $\tilde{H}$ and we get:

$$
\begin{aligned}
&-\frac{1}{2} \big[\partial_Y^2 \tilde{H} Y (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} (\boldsymbol{\mu} - r\mathbf{1}) - \boldsymbol{R}^{-1}\boldsymbol{S} \partial_Y \tilde{H} \big]' \big[\boldsymbol{R}^{-1} + \partial_Y^2 \tilde{H} (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} \big]^{-1} \\
&\quad \big[\partial_Y^2 \tilde{H} Y (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} (\boldsymbol{\mu} - r\mathbf{1}) - \boldsymbol{R}^{-1}\boldsymbol{S} \partial_Y \tilde{H} \big] - \frac{1}{2} (\partial_Y \tilde{H})^2 (Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}) \\
&\qquad - \frac{1}{2} \frac{(\partial_Y \partial_\eta \tilde{H})^2 \frac{a^2}{\epsilon}}{\partial_Y^2 \tilde{H}} - \frac{1}{2} \frac{(\partial_Y \partial_\zeta \tilde{H})^2 l^2 \delta}{\partial_Y^2 \tilde{H}} \\
&- rY \partial_Y \tilde{H} + \frac{1}{2} \partial_Y^2 \tilde{H} Y^2 (\boldsymbol{\mu} - r\mathbf{1})' (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} (\boldsymbol{\mu} - r\mathbf{1}) + \partial_t \tilde{H} + \mathcal{L}_\eta^\epsilon \tilde{H} + \mathcal{L}_\zeta^\delta \tilde{H} = 0,
\end{aligned}
\tag{4.1.14}
$$

with the terminal condition:

$$
\tilde{H}(T, Y, \eta, \zeta) = \frac{Y^2 + 2BY + B^2}{2A}.
$$

To solve (4.1.14), we apply the ansatz:

$$
\tilde{H}(t, Y, \eta, \zeta) = \tilde{v}_0(t, \eta, \zeta) + \tilde{v}_1(t, \eta, \zeta) Y + \tilde{v}_2(t, \eta, \zeta) Y^2.
$$

By substituting it into (4.1.14), grouping terms in order of power of Y, and equating all the three coefficients of each power of Y to zero. We have the following results:

$$\begin{cases} -2\tilde{v}_2^2[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\boldsymbol{R}^{-1}\boldsymbol{S}]'[\boldsymbol{R}^{-1}+\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}]^{-1} \\ \quad \big[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\boldsymbol{R}^{-1}\boldsymbol{S}\big]-2\tilde{v}_2^2(Q-\boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}) \\ \quad +\tilde{v}_2[(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-2r]-\dfrac{a^2(\partial_\eta\tilde{v}_2)^2}{\epsilon\tilde{v}_2}-\dfrac{l^2\delta(\partial_\zeta\tilde{v}_2)^2}{\tilde{v}_2}+\mathcal{D}\tilde{v}_2=0, \\ 2\tilde{v}_1\tilde{v}_2\{[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\boldsymbol{R}^{-1}\boldsymbol{S}]'[\boldsymbol{R}^{-1}+\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}]^{-1}\boldsymbol{R}^{-1}\boldsymbol{S}-Q+\boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}\} \\ \quad -r\tilde{v}_1-\dfrac{a^2\partial_\eta\tilde{v}_1\partial_\eta\tilde{v}_2}{\epsilon\tilde{v}_2}-\dfrac{l^2\delta\partial_\zeta\tilde{v}_1\partial_\zeta\tilde{v}_2}{\tilde{v}_2}+\mathcal{D}\tilde{v}_1=0, \\ -\dfrac{1}{2}\tilde{v}_1^2\{\boldsymbol{S}'\boldsymbol{R}^{-1}[\boldsymbol{R}^{-1}+\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}]^{-1}\boldsymbol{R}^{-1}\boldsymbol{S}+Q-\boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}\} \\ \quad -\dfrac{a^2(\partial_\eta\tilde{v}_1)^2}{4\epsilon\tilde{v}_2}-\dfrac{l^2\delta(\partial_\zeta\tilde{v}_1)^2}{4\tilde{v}_2}+\mathcal{D}\tilde{v}_0=0, \end{cases} \tag{4.1.15}$$

where the terminal conditions are:

$$\begin{cases} \tilde{v}_0(T,\eta,\zeta)=\dfrac{B^2}{2A}, \\ \tilde{v}_1(T,\eta,\zeta)=\dfrac{B}{A}, \\ \tilde{v}_2(T,\eta,\zeta)=\dfrac{1}{2A}. \end{cases}$$

With this ansatz, we can express all controls in terms of $\tilde{v}_0$, $\tilde{v}_1$, and $\tilde{v}_2$:

$$\begin{cases} \boldsymbol{\beta}^* = \big[\boldsymbol{R}^{-1}+2\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\big]^{-1}\left\{2\tilde{v}_2Y[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\boldsymbol{R}^{-1}\boldsymbol{S}]-\tilde{v}_1\boldsymbol{R}^{-1}\boldsymbol{S}\right\}, \\ \alpha^* = \boldsymbol{S}'\boldsymbol{R}^{-1}\big[\boldsymbol{R}^{-1}+2\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\big]^{-1}\left\{2\tilde{v}_2Y[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\boldsymbol{R}^{-1}\boldsymbol{S}]-\tilde{v}_1\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \\ \qquad -(\tilde{v}_1+\tilde{v}_2Y)(Q-\boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}), \\ \gamma^* = -\dfrac{a}{\sqrt{\epsilon}}\dfrac{\partial_\eta\tilde{v}_1+2\partial_\eta\tilde{v}_2Y}{2\tilde{v}_2}, \\ \xi^* = -l\sqrt{\delta}\dfrac{\partial_\zeta\tilde{v}_1+2\partial_\zeta\tilde{v}_2Y}{2\tilde{v}_2}, \\ y_0^* = -\dfrac{x_0+\tilde{v}_1(0,\eta_0,\zeta_0)}{2\tilde{v}_2(0,\eta_0,\zeta_0)}. \end{cases} \tag{4.1.16}$$

Again, we do not have to solve $\tilde{v}_0$ to obtain these optimal controls and the corresponding path of Y.

4.2 Theoretical Verification of the Equivalence between the Dual and Primal HJB Equations

We can verify, once we combine (4.1.7) and (4.1.14), that the PDEs (4.1.14) will be converted to the PDEs for H in (3.0.3):

$$
\begin{aligned}
LHS =& -\frac{1}{2}\left[-\frac{\partial_X H}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})+\boldsymbol{R}^{-1}\boldsymbol{S}X\right]'\left[\boldsymbol{R}^{-1}+\frac{1}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\right]^{-1}\\
&\left[-\frac{\partial_X H}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})+\boldsymbol{R}^{-1}\boldsymbol{S}X\right]-\frac{1}{2}X^2(Q-\boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S})-rX\partial_X H\\
&+\frac{1}{2}\frac{(\partial_X H)^2}{\partial_X^2 H}(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\partial_t H-\mathcal{L}_\eta^\epsilon H-\mathcal{L}_\zeta^\delta H\\
=&\frac{1}{2}X^2\boldsymbol{S}'\left\{\boldsymbol{I}-\left[\boldsymbol{I}+\frac{1}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\boldsymbol{R}\right]^{-1}\right\}\boldsymbol{R}^{-1}\boldsymbol{S}\\
&+\frac{1}{2\partial_X^2 H}(\partial_X H)^2(\boldsymbol{\mu}-r\mathbf{1})'\left\{\boldsymbol{I}-\left[\partial_X^2 H\boldsymbol{R}^{-1}\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{I}\right]^{-1}\right\}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})\\
&+X\partial_X H\boldsymbol{S}'(\partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{R})^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\frac{1}{2}QX^2-rX\partial_X H-\partial_t H-\mathcal{L}_\eta^\epsilon H-\mathcal{L}_\zeta^\delta H\\
=&\frac{1}{2}X^2\boldsymbol{S}'(\partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{R})^{-1}\boldsymbol{S}+\frac{1}{2}(\partial_X H)^2(\boldsymbol{\mu}-r\mathbf{1})'(\partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{R})^{-1}(\boldsymbol{\mu}-r\mathbf{1})\\
&+X\partial_X H\boldsymbol{S}'(\partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{R})^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\frac{1}{2}QX^2-rX\partial_X H-\partial_t H-\mathcal{L}_\eta^\epsilon H-\mathcal{L}_\zeta^\delta H\\
=&\frac{1}{2}[X\boldsymbol{S}+\partial_X H(\boldsymbol{\mu}-r\mathbf{1})]'(\partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{R})^{-1}[X\boldsymbol{S}+\partial_X H(\boldsymbol{\mu}-r\mathbf{1})]\\
&-\frac{1}{2}QX^2-rX\partial_X H-\partial_t H-\mathcal{L}_\eta^\epsilon H-\mathcal{L}_\zeta^\delta H,
\end{aligned}
$$

where the cross partial derivative terms are cancelled out by the same terms with opposite sign in the relationship between $(\partial_\eta^2\tilde{H},\partial_\zeta^2\tilde{H})$ and $(\partial_\eta^2 H,\partial_\zeta^2 H)$, and the second to the last step is obtained by using the fact that:

$$
\begin{aligned}
\boldsymbol{I}&=\left[\boldsymbol{I}+\frac{1}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\boldsymbol{R}\right]\left[\boldsymbol{I}+\frac{1}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\boldsymbol{R}\right]^{-1},\\
\boldsymbol{I}&=\left[\partial_X^2 H\boldsymbol{R}^{-1}\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{I}\right]\left[\partial_X^2 H\boldsymbol{R}^{-1}\boldsymbol{\sigma}\boldsymbol{\sigma}'+\boldsymbol{I}\right]^{-1}.
\end{aligned}
$$

Therefore, the dual HJB equation and the primal HJB equation are equivalent intrinsically.

In addition to the theoretical equivalence, we can obtain some useful relations between the dual HJB equation and the primal HJB equation. With the ansatzes of both primal and dual value functions, we can derive the relationship between (v_0,v_1,v_2) and $(\tilde{v}_0,\tilde{v}_1,\tilde{v}_2)$:

$$
\begin{aligned}
H(t,X,\eta,\zeta)&=\max_Y(-\tilde{H}-XY)\\
&=\max_Y(-\tilde{v}_0-\tilde{v}_1Y-\tilde{v}_2Y^2-XY)\\
&=\frac{1}{4\tilde{v}_2}(X+\tilde{v}_1)^2-\tilde{v}_0\\
&=\frac{1}{4\tilde{v}_2}X^2+\frac{\tilde{v}_1}{2\tilde{v}_2}X+\frac{\tilde{v}_1^2}{4\tilde{v}_2}-\tilde{v}_0\\
&=v_2X^2+v_1X+v_0,
\end{aligned}
$$

and thus we have:

$$
\begin{aligned}
\tilde{v}_0&=\frac{v_1^2}{4v_2}-v_0,\\
\tilde{v}_1&=\frac{v_1}{2v_2},\\
\tilde{v}_2&=\frac{1}{4v_2}.
\end{aligned}
\tag{4.2.1}
$$

By (4.2.1), we immediately know that the last equation of (5.2.12) is equivalent to (4.1.8). In addition, we can easily use the dual HJB equation to recover the the optimal control $\boldsymbol{\pi}^*$ and the corresponding primal state variable $X^{\boldsymbol{\pi}^*}$ by using (4.1.7), (4.1.10), (5.2.12), and (4.2.1). We start from (4.1.10) to express the optimal control $\boldsymbol{\pi}^*$ in terms of v_1, v_2, and $X^{\boldsymbol{\pi}^*}$:

$$
\begin{aligned}
\boldsymbol{\pi}^* =& \boldsymbol{R}^{-1}\boldsymbol{S}(\tilde{v}_1 + 2\tilde{v}_2 Y) \\
&+ \boldsymbol{R}^{-1}\big[\boldsymbol{R}^{-1} + 2\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\big]^{-1}\left\{2\tilde{v}_2 Y[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) - \boldsymbol{R}^{-1}\boldsymbol{S}] - \tilde{v}_1\boldsymbol{R}^{-1}\boldsymbol{S}\right\} \\
=& \boldsymbol{R}^{-1}\big[\boldsymbol{R}^{-1} + 2\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\big]^{-1}\Big\{\big[\boldsymbol{R}^{-1} + 2\tilde{v}_2(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\big]\boldsymbol{S}(\tilde{v}_1 + 2\tilde{v}_2 Y) \\
&+ 2\tilde{v}_2 Y[(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) - \boldsymbol{R}^{-1}\boldsymbol{S}] - \tilde{v}_1\boldsymbol{R}^{-1}\boldsymbol{S}\Big\} \\
=& \frac{1}{2v_2}\boldsymbol{R}^{-1}\left[\boldsymbol{R}^{-1} + \frac{1}{2v_2}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\right]^{-1}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\left[\frac{v_1}{2v_2}\boldsymbol{S} + \frac{1}{2v_2}Y\boldsymbol{S} + Y(\boldsymbol{\mu} - r\mathbf{1})\right] \\
=& (2v_2\boldsymbol{\sigma}\boldsymbol{\sigma}' + \boldsymbol{R})^{-1}\left\{\frac{v_1}{2v_2}\boldsymbol{S} - \left[\frac{1}{2v_2}\boldsymbol{S} + (\boldsymbol{\mu} - r\mathbf{1})\right](v_1 + 2v_2 X^{\boldsymbol{\pi}^*})\right\} \\
=& -(2v_2\boldsymbol{\sigma}\boldsymbol{\sigma}' + \boldsymbol{R})^{-1}\left\{X^{\boldsymbol{\pi}^*}[S + 2v_2(\boldsymbol{\mu} - r\mathbf{1})] + v_1(\boldsymbol{\mu} - r\mathbf{1})\right\},
\end{aligned}
$$

which coincides with (3.0.4). Furthermore, we can express the state variable X in terms of Y for the convenience of validating the equivalence numerically later:

$$
X^{\boldsymbol{\pi}^*} = -\tilde{v}_1 - 2\tilde{v}_2 Y^{\{y_0^*, \alpha^*, \boldsymbol{\beta}^*, \gamma^*, \xi^*\}}. \tag{4.2.2}
$$

Chapter 5

Simplified Case: $f = 0$

Suppose $Q = 0$, $\boldsymbol{S} = \mathbf{0}$, and $\boldsymbol{R} = \mathbf{0}$, or equivalently, $f = 0$. This makes the model more tractable. In addition, this also makes our optimisation objective purely focus on minimising the variance of the terminal total wealth. In this chapter, we will firstly discuss the simplified primal HJB equation and the dual HJB equation, and partly solve them using the Feynman-Kac theorem. Secondly, we will introduce the correlation structure of the Brownian motions and then try to see whether it is compatible with the uncorrelated structure. Lastly, we will introduce the constraint, specifically, a convex cone, on the control from the perspective of the dual HJB equation. We will only discuss the last two setups theoretically in this paper.

5.1 Primal and Dual HJB under $f = 0$

5.1.1 Primal HJB Equation

By directly plugging in the zero coefficients of the running cost function, the primal HJB equation (3.0.2) becomes:

$$\mathcal{D}H + Xr\partial_X H - \frac{(\partial_X H)^2}{2\partial_X^2 H}(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) = 0, \tag{5.1.1}$$

with the terminal condition:

$$H(T, X, \eta, \zeta) = \frac{1}{2}AX^2 + BX.$$

Using the same ansatz:

$$H(t, X, \eta, \zeta) = v_0(t, \eta, \zeta) + v_1(t, \eta, \zeta)X + v_2(t, \eta, \zeta)X^2,$$

we obtain the PDEs for v_0, v_1, and v_2:

$$\begin{cases} \mathcal{D}v_2 + [2r - (\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})]v_2 = 0, \\ \mathcal{D}v_1 + [r - (\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})]v_1 = 0, \\ \mathcal{D}v_0 + (\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})\dfrac{v_1^2}{4v_2} = 0, \end{cases} \tag{5.1.2}$$

with the terminal conditions:

$$\begin{cases} v_2(T, \eta, \zeta) = \dfrac{A}{2}, \\ v_1(T, \eta, \zeta) = B, \\ v_0(T, \eta, \zeta) = 0. \end{cases}$$

Noting that the structures of PDEs of v_1 and v_2 in (5.1.2) are both linear, we can use the multidimensional Feynman-Kac theorem to express v_1 and v_2 through the conditional expectation as follows:

$$\begin{aligned} v_2(t, \eta, \zeta) &= \mathbb{E}\left\{\frac{1}{2}Ae^{\int_t^T [2r - (\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})]\mathrm{d}s} | \eta_t = \eta, \zeta_t = \zeta\right\}, \\ v_1(t, \eta, \zeta) &= \mathbb{E}\left\{Be^{\int_t^T [r - (\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})]\mathrm{d}s} | \eta_t = \eta, \zeta_t = \zeta\right\}. \end{aligned} \tag{5.1.3}$$

In addition, with the aforementioned solved ansatz, the optimal control $\boldsymbol{\pi}^*$ can be expressed as follows:

$$\begin{aligned}\boldsymbol{\pi}^*(t,X,\eta,\zeta) =&(\partial_X^2 H\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}[\partial_X H(\boldsymbol{\mu}-r\mathbf{1})]\\ =&-\frac{v_1+2v_2X}{2v_2}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1}).\end{aligned} \tag{5.1.4}$$

5.1.2 Dual HJB Equation

From the general case of the dual HJB equation, it is nontrivial to derive the dual HJB equation by directly plugging in $f=0$ as $\boldsymbol{R}^{-1}$ will be hard to deal with. Instead, we notice that in the supremum calculation in (4.1.9), we can easily know that the optimal α^* and $\boldsymbol{\beta}^*$ should both be zero as otherwise x and $\boldsymbol{\pi}$ will explode the supremum. In addition, the optimisations of γ and ξ do not involve f so the optimisation results with respect to these two controls remain the same. Then the dual HJB equation becomes:

$$\mathcal{D}\tilde{H}-rY\partial_Y\tilde{H}+\frac{1}{2}\partial_Y^2\tilde{H}Y^2(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-\frac{1}{2}\frac{(\partial_Y\partial_\eta\tilde{H})^2\frac{a^2}{\epsilon}}{\partial_Y^2\tilde{H}}-\frac{1}{2}\frac{(\partial_Y\partial_\zeta\tilde{H})^2l^2\delta}{\partial_Y^2\tilde{H}}=0, \tag{5.1.5}$$

with the terminal condition:

$$\tilde{H}(T,Y,\eta,\zeta)=\frac{Y^2+2BY+B^2}{2A}.$$

By plugging in the same ansatz:

$$\tilde{H}(t,Y,\eta,\zeta)=\tilde{v}_0(t,\eta,\zeta)+\tilde{v}_1(t,\eta,\zeta)Y+\tilde{v}_2(t,\eta,\zeta)Y^2,$$

and expressing the dual HJB equation as a polynomial of Y, we obtain the PDEs for $\tilde{v}_0$, $\tilde{v}_1$, and $\tilde{v}_2$:

$$\begin{cases}\mathcal{D}\tilde{v}_2+[(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-2r]\tilde{v}_2-\dfrac{a^2(\partial_\eta\tilde{v}_2)^2}{\epsilon\tilde{v}_2}-\dfrac{l^2\delta(\partial_\zeta\tilde{v}_2)^2}{\tilde{v}_2}=0,\\ \mathcal{D}\tilde{v}_1-r\tilde{v}_1-\dfrac{a^2\partial_\eta\tilde{v}_1\partial_\eta\tilde{v}_2}{\epsilon\tilde{v}_2}-\dfrac{l^2\delta\partial_\zeta\tilde{v}_1\partial_\zeta\tilde{v}_2}{\tilde{v}_2}=0,\\ \mathcal{D}\tilde{v}_0-\dfrac{a^2(\partial_\eta\tilde{v}_1)^2}{4\epsilon\tilde{v}_2}-\dfrac{l^2\delta(\partial_\zeta\tilde{v}_1)^2}{4\tilde{v}_2}=0,\end{cases} \tag{5.1.6}$$

with the terminal conditions:

$$\begin{cases}\tilde{v}_2(T,\eta,\zeta)=\dfrac{1}{2A},\\ \tilde{v}_1(T,\eta,\zeta)=\dfrac{B}{A},\\ \tilde{v}_0(T,\eta,\zeta)=\dfrac{B^2}{2A}.\end{cases}$$

It is not difficult to spot that $\tilde{v}_0$ and $\tilde{v}_1$ are both independent of η and ζ due to their η- and ζ-independent terminal conditions. More specifically, we can see the point from two aspects starting from $\tilde{v}_1$. On the one hand, if we look a little step backward from the terminal time T, as $\tilde{v}_1(T)$ is independent of η and ζ, all derivatives with respect to η and ζ at T will be zero and thus, $\tilde{v}_1$ at a slightly earlier time should also be independent of η and ζ. By proceeding this logic backwardly, we will ultimately find that $\tilde{v}_1$ is independent of η and ζ throughout the timeline. On the other hand, if we assume that $\tilde{v}_1$ is independent of η and ζ before we solve its PDE given that $\tilde{v}_1(T)$ is independent of η and ζ, then we will be able to spot that the dynamics of $\tilde{v}_1$ becomes also independent of η and ζ:

$$\partial_t\tilde{v}_1-r\tilde{v}_1=0,$$

which also leads to an η- and ζ-independent structure of $\tilde{v}_1$ throughout the timeline. As for $\tilde{v}_0$, since we already know that $\tilde{v}_1$ is independent of η and ζ, all derivatives of $\tilde{v}_1$ with respect to η and ζ become zero and we can reason from the same two aspects aforementioned, given that $\tilde{v}_0(T)$ is independent of η and ζ, then we can conclude that $\tilde{v}_0$ is independent of η and ζ throughout the timeline. In addition, this fact makes the dynamics of $\tilde{v}_0$ become:

$$\partial_t\tilde{v}_0=0,$$

indicating $\tilde{v}_0$ is a constant. So we conclude the solution of them as follows:

$$
\begin{aligned}
\tilde{v}_1(t) &= \frac{B}{A}\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s}, \\
\tilde{v}_0(t) &= \frac{B^2}{2A}.
\end{aligned} \tag{5.1.7}
$$

With this ansatz, we can express all controls in terms of $\tilde{v}_0$, $\tilde{v}_1$, and $\tilde{v}_2$:

$$
\begin{cases}
\boldsymbol{\beta}^* = \mathbf{0}, \\
\alpha^* = 0, \\
\gamma^* = -\dfrac{a}{\sqrt{\epsilon}}\dfrac{\partial_\eta \tilde{v}_1 + 2\partial_\eta \tilde{v}_2 Y}{2\tilde{v}_2}, \\
\xi^* = -l\sqrt{\delta}\dfrac{\partial_\zeta \tilde{v}_1 + 2\partial_\zeta \tilde{v}_2 Y}{2\tilde{v}_2}, \\
y_0^* = -\dfrac{x_0 + \tilde{v}_1(0,\eta_0,\zeta_0)}{2\tilde{v}_2(0,\eta_0,\zeta_0)}.
\end{cases} \tag{5.1.8}
$$

5.2 Introducing the Correlation under $f = 0$

Previously, our discussion is restricted to the uncorrelated case. Now we introduce the correlation among $\boldsymbol{W}_t$, W_t^η, and W_t^ζ for the theoretical completeness and we will lastly verify that we can start from the correlated case to the uncorrelated case by making $\tilde{\boldsymbol{\rho}} = \boldsymbol{I}$. Let us assume that the correlation structure is in a general form:

$$
\tilde{\boldsymbol{\rho}} = \begin{pmatrix} \boldsymbol{\rho} & \boldsymbol{\rho}^\eta & \boldsymbol{\rho}^\zeta \\ \boldsymbol{\rho}^{\eta\prime} & 1 & \rho^{\eta\zeta} \\ \boldsymbol{\rho}^{\zeta\prime} & \rho^{\eta\zeta} & 1 \end{pmatrix}, \tag{5.2.1}
$$

where $\boldsymbol{\rho}$ is the correlation matrix of $\boldsymbol{W}_t$,

$$
\boldsymbol{\rho}^\eta = [\rho_1^\eta, \cdots, \rho_N^\eta]' \in \mathbb{R}^N
$$

is the correlation vector between $\boldsymbol{W}_t$ and W_t^η,

$$
\boldsymbol{\rho}^\zeta = [\rho_1^\zeta, \cdots, \rho_N^\zeta]' \in \mathbb{R}^N
$$

is the correlation vector between $\boldsymbol{W}_t$ and W_t^ζ, $\rho^{\eta\zeta}$ is the correlation between W_t^η and W_t^ζ, and $\tilde{\boldsymbol{\rho}} \in \mathbb{R}^{N+2}$ and positive semidefinite. To simplify the calculation, we still derive the primal HJB equation and the dual HJB equation under the assumption $f = 0$.

5.2.1 Primal HJB Equation

The effect of the appearance of the correlation structure is introducing three additional terms in the primal HJB equation arising from the cross partial derivatives, and by the dynamic programming principle, the primal HJB equation (5.1.1) becomes:

$$
\begin{aligned}
\inf_{\boldsymbol{\pi}} \left\{ \frac{1}{2}\partial_X H \boldsymbol{\pi}'\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}'\boldsymbol{\pi} + \left[\partial_X H(\boldsymbol{\mu} - r\mathbf{1})' + \partial_X\partial_\eta H \frac{a}{\sqrt{\epsilon}}\boldsymbol{\rho}^{\eta\prime}\boldsymbol{\sigma}' + \partial_X\partial_\zeta H\sqrt{\delta}l\boldsymbol{\rho}^{\zeta\prime}\boldsymbol{\sigma}'\right]\boldsymbol{\pi}\right\} \\
+\tilde{\mathcal{D}}H + Xr\partial_X H = 0,
\end{aligned}
$$

with the terminal condition:

$$
H(T, X, \eta, \zeta) = \frac{1}{2}AX^2 + BX,
$$

where:

$$
\tilde{\mathcal{D}} = \mathcal{D} + \sqrt{\frac{\delta}{\epsilon}}al\rho^{\eta\zeta}\partial_\eta\partial_\zeta.
$$

The optimisation in terms of π is in a quadratic form and thus the optimal control is:

$$
\boldsymbol{\pi}^* = -\frac{1}{\partial_X^2 H}(\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}')^{-1}\left[\partial_X H(\boldsymbol{\mu} - r\mathbf{1}) + \partial_X\partial_\eta H\frac{a}{\sqrt{\epsilon}}\boldsymbol{\sigma}\boldsymbol{\rho}^\eta + \partial_X\partial_\zeta H l\sqrt{\delta}\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]. \tag{5.2.2}
$$

By plugging in (5.2.2), we can obtain:

$$
\begin{aligned}
&-\frac{1}{2\partial_X^2 H}\left[\partial_X H(\boldsymbol{\mu}-r\mathbf{1})+\partial_X\partial_\eta H\frac{a}{\sqrt{\epsilon}}\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+\partial_X\partial_\zeta H l\sqrt{\delta}\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]'(\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}')^{-1}\\
&\left[\partial_X H(\boldsymbol{\mu}-r\mathbf{1})+\partial_X\partial_\eta H\frac{a}{\sqrt{\epsilon}}\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+\partial_X\partial_\zeta H l\sqrt{\delta}\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]+\tilde{\mathcal{D}}H+Xr\partial_X H=0.
\end{aligned}
\tag{5.2.3}
$$

with the terminal condition:

$$H(T,X,\eta,\zeta)=\frac{1}{2}AX^2+BX.$$

By applying the ansatz:

$$H(t,X,\eta,\zeta)=v_0(t,\eta,\zeta)+v_1(t,\eta,\zeta)X+v_2(t,\eta,\zeta)X^2,$$

we have the PDEs:

$$
\begin{cases}
-\dfrac{1}{v_2}\left[v_2(\boldsymbol{\mu}-r\mathbf{1})+\dfrac{a}{\sqrt{\epsilon}}\partial_\eta v_2\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+l\sqrt{\delta}\partial_\zeta v_2\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]'(\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}')^{-1}\\
\quad\left[v_2(\boldsymbol{\mu}-r\mathbf{1})+\dfrac{a}{\sqrt{\epsilon}}\partial_\eta v_2\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+l\sqrt{\delta}\partial_\zeta v_2\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]+2rv_2+\tilde{\mathcal{D}}v_2=0,\\
-\dfrac{1}{v_2}\left[v_2(\boldsymbol{\mu}-r\mathbf{1})+\dfrac{a}{\sqrt{\epsilon}}\partial_\eta v_2\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+l\sqrt{\delta}\partial_\zeta v_2\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]'(\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}')^{-1}\\
\quad\left[v_1(\boldsymbol{\mu}-r\mathbf{1})+\dfrac{a}{\sqrt{\epsilon}}\partial_\eta v_1\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+l\sqrt{\delta}\partial_\zeta v_1\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]+rv_1+\tilde{\mathcal{D}}v_1=0,\\
-\dfrac{1}{4v_2}\left[v_1(\boldsymbol{\mu}-r\mathbf{1})+\dfrac{a}{\sqrt{\epsilon}}\partial_\eta v_1\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+l\sqrt{\delta}\partial_\zeta v_1\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]'(\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}')^{-1}\\
\quad\left[v_1(\boldsymbol{\mu}-r\mathbf{1})+\dfrac{a}{\sqrt{\epsilon}}\partial_\eta v_1\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+l\sqrt{\delta}\partial_\zeta v_1\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]+\tilde{\mathcal{D}}v_0=0,
\end{cases}
\tag{5.2.4}
$$

with the terminal conditions:

$$
\begin{cases}
v_2(T,\eta,\zeta)=\dfrac{A}{2},\\
v_1(T,\eta,\zeta)=B,\\
v_0(T,\eta,\zeta)=0.
\end{cases}
$$

Therefore, we can express the optimal control by taking advantage of the ansatz once we solve them from (5.2.4) as follows:

$$
\begin{aligned}
\boldsymbol{\pi}^*=&-\frac{1}{2v_2}(\boldsymbol{\sigma}\boldsymbol{\rho}\boldsymbol{\sigma}')^{-1}\Bigg\{X\left[2v_2(\boldsymbol{\mu}-r\mathbf{1})+2\partial_\eta v_2\frac{a}{\sqrt{\epsilon}}\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+2\partial_\zeta v_2 l\sqrt{\delta}\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\right]\\
&+v_1(\boldsymbol{\mu}-r\mathbf{1})+\partial_\eta v_1\frac{a}{\sqrt{\epsilon}}\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+\partial_\zeta v_1 l\sqrt{\delta}\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta\Bigg\}.
\end{aligned}
\tag{5.2.5}
$$

5.2.2 Dual HJB Equation

We still assume the dynamics of Y follows (4.1.1), but take the correlation into consideration, and thus we have:

$$\mathrm{d}(XY)=[\alpha X+\boldsymbol{\pi}'(\boldsymbol{\beta}+\gamma\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+\xi\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta)]\mathrm{d}t+\textit{local martingale},$$

where:

$$
\begin{aligned}
\alpha_{1t}&=\alpha_t-r_tY_t,\\
\boldsymbol{\beta}_{1t}&=\boldsymbol{\sigma}^{-1}[\boldsymbol{\beta}_t-Y_t(\boldsymbol{\mu}-r_t\mathbf{1})].
\end{aligned}
$$

Notint that when $f=0$:

$$\phi(t,\alpha,\boldsymbol{\beta},\gamma,\xi)=\sup_{x,\boldsymbol{\pi}}\{x\alpha+\boldsymbol{\pi}'(\boldsymbol{\beta}+\gamma\boldsymbol{\sigma}\boldsymbol{\rho}^\eta+\xi\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta)\},$$

we have:

$$\begin{aligned}
&\phi = 0,\\
&\alpha = 0,\\
&\boldsymbol{\beta} + \gamma\boldsymbol{\sigma}\boldsymbol{\rho}^{\eta} + \xi\boldsymbol{\sigma}\boldsymbol{\rho}^{\zeta} = 0,
\end{aligned}$$

as otherwise ϕ should be positive infinity by selecting a wild value of x or $\boldsymbol{\beta}$. We then have the dual HJB equation by dynamic programming principle:

$$\begin{aligned}
\inf_{\gamma,\xi,\boldsymbol{\beta}+\gamma\boldsymbol{\sigma}\boldsymbol{\rho}^{\eta}+\xi\boldsymbol{\sigma}\boldsymbol{\rho}^{\zeta}=0} &\Bigg\{ \frac{1}{2}\partial_Y^2\tilde{H}\Big\{[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'(\boldsymbol{\sigma}\boldsymbol{\rho}^{-1}\boldsymbol{\sigma}')^{-1}[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})] + \gamma^2 + \xi^2 \\
&+ 2\gamma\xi\rho^{\eta\zeta} + 2[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'\boldsymbol{\sigma}'^{-1}(\gamma\boldsymbol{\rho}^{\eta} + \gamma\rho^{\zeta})\Big\} \\
&+\gamma(\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}} + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\rho^{\eta\zeta}) + \xi(\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\rho^{\eta\zeta} + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l) \\
&+[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'\boldsymbol{\sigma}'^{-1}(\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\boldsymbol{\rho}^{\eta} + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\boldsymbol{\rho}^{\zeta})\Bigg\} + \tilde{\mathcal{D}}\tilde{H} - rY\partial_Y\tilde{H} = 0.
\end{aligned}$$

with the terminal condition:

$$\tilde{H}(T, Y, \eta, \zeta) = \frac{Y^2 + 2BY + B^2}{2A}.$$

By plugging in the constraint of $\boldsymbol{\beta}$, we can obtain a quadratic optimisation with respect to γ and ξ:

$$\begin{aligned}
&\inf_{\gamma,\xi}\left\{\frac{1}{2}A_1\gamma^2 + \frac{1}{2}A_2\xi^2 + A_3\gamma\xi + A_4\gamma + A_5\xi\right\} \\
&+\tilde{\mathcal{D}}\tilde{H} + \frac{1}{2}\partial_Y^2\tilde{H}Y^2(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\rho}^{-1}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) \\
&-Y\left[r\partial_Y\tilde{H} + (\boldsymbol{\mu} - r\mathbf{1})'\boldsymbol{\sigma}'^{-1}(\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\boldsymbol{\rho}^{\eta} + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\boldsymbol{\rho}^{\zeta})\right] = 0,
\end{aligned} \tag{5.2.6}$$

with the terminal condition:

$$\tilde{H}(T, Y, \eta, \zeta) = \frac{Y^2 + 2BY + B^2}{2A},$$

where:

$$\begin{aligned}
A_1 =& \partial_Y^2\tilde{H}\left[\boldsymbol{\rho}^{\eta\prime}(\boldsymbol{\rho} - 2\boldsymbol{I})\boldsymbol{\rho}^{\eta} + 1\right], \\
A_2 =& \partial_Y^2\tilde{H}\left[\boldsymbol{\rho}^{\zeta\prime}(\boldsymbol{\rho} - 2\boldsymbol{I})\boldsymbol{\rho}^{\zeta} + 1\right], \\
A_3 =& \partial_Y^2\tilde{H}\left(\rho^{\eta\zeta} + \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}\boldsymbol{\rho}^{\zeta} - 2\boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^{\zeta}\right), \\
A_4 =& -\partial_Y^2\tilde{H}\boldsymbol{\rho}^{\eta\prime}(\boldsymbol{I} - \boldsymbol{\rho})\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})Y \\
&+ \partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\left(1 - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^{\zeta}\right) + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\left(\rho^{\eta\zeta} - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^{\zeta}\right), \\
A_5 =& -\partial_Y^2\tilde{H}\boldsymbol{\rho}^{\zeta\prime}(\boldsymbol{I} - \boldsymbol{\rho})\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})Y \\
&+ \partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\left(\rho^{\eta\zeta} - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^{\zeta}\right) + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\left(1 - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^{\zeta}\right).
\end{aligned} \tag{5.2.7}$$

The optimisers are:

$$\begin{aligned}
\gamma^* &= \frac{A_3A_5 - A_2A_4}{A_1A_2 - A_3^2}, \\
\xi^* &= \frac{A_3A_4 - A_1A_5}{A_1A_2 - A_3^2},
\end{aligned} \tag{5.2.8}$$

and the dual HJB equation becomes:

$$\begin{aligned}
&\frac{2A_3A_4A_5 - A_4^2A_2 - A_5^2A_1}{2(A_1A_2 - A_3^2)} + \tilde{\mathcal{D}}\tilde{H} + \frac{1}{2}\partial_Y^2\tilde{H}Y^2(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\rho}^{-1}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) \\
&-Y\left[r\partial_Y\tilde{H} + (\boldsymbol{\mu} - r\mathbf{1})'\boldsymbol{\sigma}'^{-1}(\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\boldsymbol{\rho}^{\eta} + \partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\boldsymbol{\rho}^{\zeta})\right] = 0,
\end{aligned} \tag{5.2.9}$$

with the terminal condition:

$$\tilde{H}(T,Y,\eta,\zeta) = \frac{Y^2+2BY+B^2}{2A}.$$

By plugging in the same ansatz:

$$\tilde{H}(t,Y,\eta,\zeta) = \tilde{v}_0(t,\eta,\zeta) + \tilde{v}_1(t,\eta,\zeta)Y + \tilde{v}_2(t,\eta,\zeta)Y^2,$$

and expressing the dual HJB equation as a polynomial of Y, we obtain the PDEs for $\tilde{v}_0$, $\tilde{v}_1$, and $\tilde{v}_2$:

$$\begin{cases} \dfrac{2A_3A_{41}A_{51} - A_2A_{41}^2 - A_1A_{51}^2}{2(A_1A_2 - A_3^2)} + \tilde{v}_2\left[(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\rho}^{-1}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) - 2r\right] \\ \qquad -2(\boldsymbol{\mu} - r\mathbf{1})'\boldsymbol{\sigma}'^{-1}\left(\dfrac{a}{\sqrt{\epsilon}}\partial_\eta\tilde{v}_2\boldsymbol{\rho}^\eta + l\sqrt{\delta}\partial_\zeta\tilde{v}_2\boldsymbol{\rho}^\zeta\right) + \tilde{\mathcal{D}}\tilde{v}_2 = 0, \\ \qquad \dfrac{A_3A_{42}A_{51} + A_3A_{41}A_{52} - A_2A_{41}A_{42} - A_1A_{51}A_{52}}{A_1A_2 - A_3^2} - r\tilde{v}_1 \\ \qquad -(\boldsymbol{\mu} - r\mathbf{1})'\boldsymbol{\sigma}'^{-1}\left(\dfrac{a}{\sqrt{\epsilon}}\partial_\eta\tilde{v}_1\boldsymbol{\rho}^\eta + l\sqrt{\delta}\partial_\zeta\tilde{v}_1\boldsymbol{\rho}^\zeta\right) + \tilde{\mathcal{D}}\tilde{v}_1 = 0, \\ \qquad \dfrac{2A_3A_{42}A_{52} - A_2A_{42}^2 - A_1A_{52}^2}{2(A_1A_2 - A_3^2)} + \tilde{\mathcal{D}}\tilde{v}_0 = 0, \end{cases} \tag{5.2.10}$$

with the terminal conditions:

$$\begin{cases} \tilde{v}_2(T,\eta,\zeta) = \dfrac{1}{2A}, \\ \tilde{v}_1(T,\eta,\zeta) = \dfrac{B}{A}, \\ \tilde{v}_0(T,\eta,\zeta) = \dfrac{B^2}{2A}, \end{cases}$$

where:

$$\begin{aligned} A_1 =& 2\tilde{v}_2\left[\boldsymbol{\rho}^{\eta\prime}(\boldsymbol{\rho} - 2\boldsymbol{I})\boldsymbol{\rho}^\eta + 1\right], \\ A_2 =& 2\tilde{v}_2\left[\boldsymbol{\rho}^{\zeta\prime}(\boldsymbol{\rho} - 2\boldsymbol{I})\boldsymbol{\rho}^\zeta + 1\right], \\ A_3 =& 2\tilde{v}_2\left(\rho^{\eta\zeta} + \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}\boldsymbol{\rho}^\zeta - 2\boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right), \\ A_4 =& A_{41}Y + A_{42}, \\ A_5 =& A_{51}Y + A_{52}, \\ A_{41} =& -2\tilde{v}_2\boldsymbol{\rho}^{\eta\prime}(\boldsymbol{I} - \boldsymbol{\rho})\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}) \\ & + 2\partial_\eta\tilde{v}_2\frac{a}{\sqrt{\epsilon}}\left(1 - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right) + 2\partial_\zeta\tilde{v}_2\sqrt{\delta}l\left(\rho^{\eta\zeta} - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right), \\ A_{42} =& \partial_\eta\tilde{v}_1\frac{a}{\sqrt{\epsilon}}\left(1 - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right) + \partial_\zeta\tilde{v}_1\sqrt{\delta}l\left(\rho^{\eta\zeta} - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right), \\ A_{51} =& -2\tilde{v}_2\boldsymbol{\rho}^{\zeta\prime}(\boldsymbol{I} - \boldsymbol{\rho})\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}) \\ & + 2\partial_\eta\tilde{v}_2\frac{a}{\sqrt{\epsilon}}\left(\rho^{\eta\zeta} - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right) + 2\partial_\zeta\tilde{v}_2\sqrt{\delta}l\left(1 - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right), \\ A_{52} =& \partial_\eta\tilde{v}_1\frac{a}{\sqrt{\epsilon}}\left(\rho^{\eta\zeta} - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right) + \partial_\zeta\tilde{v}_1\sqrt{\delta}l\left(1 - \boldsymbol{\rho}^{\eta\prime}\boldsymbol{\rho}^\zeta\right). \end{aligned} \tag{5.2.11}$$

With this ansatz, we can express all controls in terms of $\tilde{v}_0$, $\tilde{v}_1$, and $\tilde{v}_2$:

$$\begin{cases} \boldsymbol{\beta}^* = -\dfrac{(A_3A_5 - A_2A_4)\boldsymbol{\sigma}\boldsymbol{\rho}^\eta + (A_3A_4 - A_1A_5)\boldsymbol{\sigma}\boldsymbol{\rho}^\zeta}{A_1A_2 - A_3^2}, \\ \alpha^* = 0, \\ \gamma^* = \dfrac{A_3A_5 - A_2A_4}{A_1A_2 - A_3^2}, \\ \xi^* = \dfrac{A_3A_4 - A_1A_5}{A_1A_2 - A_3^2}, \\ y_0^* = -\dfrac{x_0 + \tilde{v}_1(0,\eta_0,\zeta_0)}{2\tilde{v}_2(0,\eta_0,\zeta_0)}. \end{cases} \tag{5.2.12}$$

Similar to the uncorrelated case under $f = 0$, we can also conclude that:

$$\begin{aligned}\tilde{v}_1(t) &= \frac{B}{A} \mathrm{e}^{-\int_t^T r(s)\mathrm{d}s}, \\ \tilde{v}_0(t) &= \frac{B^2}{2A}.\end{aligned} \tag{5.2.13}$$

Following the similar logic, starting from $\tilde{v}_1$, we can assume that $\tilde{v}_1$ is independent of η and ζ throughout the timeline due to the fact that $\tilde{v}_1(T)$ is independent of η and ζ, which makes A_{42} and S_{52} become zero and ultimately makes the dynamics of $\tilde{v}_1$ becomes:

$$\partial_t \tilde{v}_1 - r\tilde{v}_1 = 0.$$

As for $\tilde{v}_0$, it is more trivial to see that:

$$\partial_t \tilde{v}_0 = 0,$$

given A_{42} and A_{52} are zero.

5.2.3 Verification: $\tilde{\boldsymbol{\rho}} = \boldsymbol{I}_{(N+2)\times(N+2)}$

Lastly, we let $\boldsymbol{\rho}$ to be identity matrix, $\boldsymbol{\rho}^{\eta}$ and $\boldsymbol{\rho}^{\zeta}$ to be zero vectors, and $\rho^{\eta\zeta}$ to be zero ('no correlation condition' for short) to verify the correlated case can go back to the uncorrelated case (under the assumption $f = 0$). We can immediately infer that $\tilde{\mathcal{D}} = \mathcal{D}$.

For the primal HJB equation, when we apply the no correlation condition to (5.2.2), we immediately get:

$$\boldsymbol{\pi}^* = -\frac{1}{\partial_X^2 H} (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} \partial_X H (\boldsymbol{\mu} - r\mathbf{1}),$$

which is identical to (5.1.4). Then, (5.2.3) becomes:

$$-\frac{1}{2\partial_X^2 H} \left[\partial_X H(\boldsymbol{\mu} - r\mathbf{1})\right]' (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} \left[\partial_X H(\boldsymbol{\mu} - r\mathbf{1})\right] + \mathcal{D}H + Xr\partial_X H = 0.$$

with the terminal condition:

$$H(T, X, \eta, \zeta) = \frac{1}{2} A X^2 + BX,$$

and (5.2.4) becomes:

$$\begin{cases} -v_2(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) + 2rv_2 + \mathcal{D}v_2 = 0, \\ -v_1(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) + rv_1 + \mathcal{D}v_1 = 0, \\ -\dfrac{v_1^2}{4v_2}(\boldsymbol{\mu} - r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1}) + \mathcal{D}v_0 = 0. \end{cases}$$

with the terminal conditions:

$$\begin{cases} v_2(T, \eta, \zeta) = \dfrac{A}{2}, \\ v_1(T, \eta, \zeta) = B, \\ v_0(T, \eta, \zeta) = 0, \end{cases}$$

which are both identical to (5.1.1) and (5.1.2) respectively.

For the dual HJB equation, when we apply the no correlation condition, firstly the constraint on $\boldsymbol{\beta}^*$ becomes $\boldsymbol{\beta}^* = \mathbf{0}$, and (5.2.7) becomes:

$$\begin{aligned} A_1 &= \partial_Y^2 \tilde{H}, \\ A_2 &= \partial_Y^2 \tilde{H}, \\ A_3 &= 0, \\ A_4 &= \partial_Y \partial_\eta \tilde{H} \frac{a}{\sqrt{\epsilon}}, \\ A_5 &= \partial_Y \partial_\zeta \tilde{H} \sqrt{\delta} l. \end{aligned}$$

Given this, (5.2.6) becomes:

$$\inf_{\gamma,\xi}\left\{\frac{1}{2}\partial_Y^2\tilde{H}\gamma^2+\frac{1}{2}\partial_Y^2\tilde{H}\xi^2+\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}\gamma+\partial_Y\partial_\zeta\tilde{H}\sqrt{\delta}l\xi\right\}$$
$$+\mathcal{D}\tilde{H}+\frac{1}{2}\partial_Y^2\tilde{H}Y^2(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-rY\partial_Y\tilde{H}=0,$$

with the terminal condition:

$$\tilde{H}(T,Y,\eta,\zeta)=\frac{Y^2+2BY+B^2}{2A},$$

which is a quadratic optimisation with respect to γ and ξ. Now (5.2.8) becomes:

$$\gamma^*=-\frac{\partial_Y\partial_\eta\tilde{H}\frac{a}{\sqrt{\epsilon}}}{\partial_Y^2\tilde{H}},$$
$$\xi^*=-\frac{\partial_Y\partial_\zeta\tilde{H}l\sqrt{\delta}}{\partial_Y^2\tilde{H}},$$

and thus (5.2.9) becomes:

$$\frac{-(\partial_Y\partial_\eta\tilde{H})^2\frac{a^2}{\epsilon}-(\partial_Y\partial_\zeta\tilde{H})^2l^2\delta}{2\partial_Y^2\tilde{H}}+\mathcal{D}\tilde{H}+\frac{1}{2}\partial_Y^2\tilde{H}Y^2(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-rY\partial_Y\tilde{H}=0,$$

with the terminal condition:

$$\tilde{H}(T,Y,\eta,\zeta)=\frac{Y^2+2BY+B^2}{2A},$$

which is identical to (5.1.5). Furthermore, after plugging the ansatz:

$$\tilde{H}(t,Y,\eta,\zeta)=\tilde{v}_0(t,\eta,\zeta)+\tilde{v}_1(t,\eta,\zeta)Y+\tilde{v}_2(t,\eta,\zeta)Y^2,$$

(5.2.11) becomes:

$$\begin{aligned}
A_1&=2\tilde{v}_2,\\
A_2&=2\tilde{v}_2,\\
A_3&=0,\\
A_4&=A_{41}Y+A_{42},\\
A_5&=A_{51}Y+A_{52},\\
A_{41}&=2\partial_\eta\tilde{v}_2\frac{a}{\sqrt{\epsilon}},\\
A_{42}&=\partial_\eta\tilde{v}_1\frac{a}{\sqrt{\epsilon}},\\
A_{51}&=2\partial_\zeta\tilde{v}_2\sqrt{\delta}l,\\
A_{52}&=\partial_\zeta\tilde{v}_1\sqrt{\delta}l,
\end{aligned}$$

and (5.2.10) becomes:

$$\begin{cases}
\dfrac{-(\partial_\eta\tilde{v}_2)^2\frac{a^2}{\epsilon}-(\partial_\zeta\tilde{v}_2)^2l^2\delta}{\tilde{v}_2}+\tilde{v}_2\left[(\boldsymbol{\mu}-r\mathbf{1})'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu}-r\mathbf{1})-2r\right]+\mathcal{D}\tilde{v}_2=0,\\
\dfrac{-\partial_\eta\tilde{v}_1\partial_\eta\tilde{v}_2\frac{a^2}{\epsilon}-\partial_\zeta\tilde{v}_1\partial_\zeta\tilde{v}_2l^2\delta}{\tilde{v}_2}-r\tilde{v}_1+\mathcal{D}\tilde{v}_1=0,\\
\dfrac{-(\partial_\eta\tilde{v}_1)^2\frac{a^2}{\epsilon}-(\partial_\zeta\tilde{v}_1)^2l^2\delta}{4\tilde{v}_2}+\mathcal{D}\tilde{v}_0=0,
\end{cases}$$

with the terminal conditions:

$$\begin{cases}
\tilde{v}_2(T,\eta,\zeta)=\dfrac{1}{2A},\\
\tilde{v}_1(T,\eta,\zeta)=\dfrac{B}{A},\\
\tilde{v}_0(T,\eta,\zeta)=\dfrac{B^2}{2A},
\end{cases}$$

which is identical to (5.1.6). In summary, the correlation setting is compatible with the no correlation setting.

5.3 Introducing the Convex Cone under $f = 0$

Previously we assumed that $K = \mathbb{R}^N$. In this section, we assume K is a closed convex cone and only consider the uncorrelated case. In addition, since $f = 0$, α should be zero as otherwise we can always choose an arbitrary X to explode ϕ, and according to Li & Zheng (2018, Section 4.2, page 1144), we further assume $\boldsymbol{\beta} \in K^0$ where:

$$K^0 = \{\boldsymbol{\beta} : \boldsymbol{\beta}'\boldsymbol{\pi} \leq 0,\ \forall \boldsymbol{\pi} \in K\}$$

is the polar cone of K. Given these assumptions, we can conclude that the optimal ϕ should be zero, and thus, by the dynamic programming principle, the dual HJB equation becomes:

$$\begin{aligned}\inf_{\boldsymbol{\beta},\gamma,\xi} \Bigg\{ \frac{1}{2}\partial_Y^2 \tilde{H}\Big[[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})] + \gamma^2 + \xi^2\Big] \\ + \partial_Y\partial_\eta \tilde{H}\frac{1}{\sqrt{\epsilon}}a\gamma + \partial_Y\partial_\zeta \tilde{H}\sqrt{\delta}l\xi \Bigg\} \\ -rY\partial_Y\tilde{H} + \mathcal{D}\tilde{H} = 0,\end{aligned}$$

with the terminal condition:

$$\tilde{H}(T, Y, \eta, \zeta) = \frac{Y^2 + 2BY + B^2}{2A},$$

The optimisations with respect to γ and ξ are the same as the previous results as they are unconstrained:

$$\gamma^* = -\frac{\partial_Y\partial_\eta \tilde{H}\frac{a}{\sqrt{\epsilon}}}{\partial_Y^2 \tilde{H}},$$
$$\xi^* = -\frac{\partial_Y\partial_\zeta \tilde{H}\sqrt{\delta}l}{\partial_Y^2 \tilde{H}}.$$

However, the optimisation with respect to $\boldsymbol{\beta}$ is nontrivial and it is dependent of the value of Y. We derive the optimisation in three cases:

1. If $Y = 0$, it is obvious that $\boldsymbol{\beta} = \mathbf{0}$ is in K^0. Thus, the infimum term of $\boldsymbol{\beta}$ is zero and is obtained when $\boldsymbol{\beta}^* = \mathbf{0}$.

2. If $Y > 0$:

$$\begin{aligned}\inf_{\boldsymbol{\beta}\in K^0} \left\|\boldsymbol{\sigma}^{-1}\boldsymbol{\beta} - Y\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2 &= \inf_{\boldsymbol{\beta}\in K^0} Y^2 \left\|\boldsymbol{\sigma}^{-1}\frac{\boldsymbol{\beta}}{Y} - \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2 \\ &= Y^2 \inf_{\frac{\boldsymbol{\beta}}{Y}\in K^0} \left\|\boldsymbol{\sigma}^{-1}\frac{\boldsymbol{\beta}}{Y} - \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2 \\ &= Y^2 \left\|\boldsymbol{\sigma}^{-1}\boldsymbol{\beta}_+ - \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2,\end{aligned}$$

where:

$$\boldsymbol{\beta}_+ = \arg\min_{\tilde{\boldsymbol{\beta}}\in K^0} \|\boldsymbol{\sigma}^{-1}\tilde{\boldsymbol{\beta}} - \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\|, \tag{5.3.1}$$

and the second equality is due to $Y > 0$ and $\boldsymbol{\beta}/Y$ is still in K^0.

3. If $Y < 0$:

$$\begin{aligned}\inf_{\boldsymbol{\beta}\in K^0} \left\|\boldsymbol{\sigma}^{-1}\boldsymbol{\beta} - Y\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2 &= \inf_{\boldsymbol{\beta}\in K^0} Y^2 \left\|\boldsymbol{\sigma}^{-1}\frac{\boldsymbol{\beta}}{Y} - \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2 \\ &= Y^2 \inf_{\frac{\boldsymbol{\beta}}{-Y}\in K^0} \left\|\boldsymbol{\sigma}^{-1}\frac{\boldsymbol{\beta}}{-Y} + \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2 \\ &= Y^2 \left\|\boldsymbol{\sigma}^{-1}\boldsymbol{\beta}_- + \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\right\|^2,\end{aligned}$$

where:

$$\boldsymbol{\beta}_{-} = \arg \min_{\tilde{\boldsymbol{\beta}} \in K^0} \|\boldsymbol{\sigma}^{-1}\tilde{\boldsymbol{\beta}} + \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})\|, \tag{5.3.2}$$

and the second equality is due to $Y < 0$ and $\boldsymbol{\beta}/(-Y)$ is still in K^0.

Therefore, we can define:

$$\boldsymbol{\Theta}(t, Y) = \begin{cases} \mathbf{0} & \text{if } Y = 0, \\ \boldsymbol{\sigma}^{-1}\boldsymbol{\beta}_{-} + \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}) & \text{if } Y < 0, \\ \boldsymbol{\sigma}^{-1}\boldsymbol{\beta}_{+} - \boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}) & \text{if } Y > 0, \end{cases} \tag{5.3.3}$$

and the dual HJB equation becomes:

$$\mathcal{D}\tilde{H} - rY\partial_Y\tilde{H} + \frac{1}{2}Y^2\|\boldsymbol{\Theta}\|^2\partial_Y^2\tilde{H} - \frac{1}{2}\frac{(\partial_Y\partial_\eta\tilde{H})^2\frac{a^2}{\epsilon}}{\partial_Y^2\tilde{H}} - \frac{1}{2}\frac{(\partial_Y\partial_\zeta\tilde{H})^2 l^2\delta}{\partial_Y^2\tilde{H}} = 0, \tag{5.3.4}$$

with the terminal condition:

$$\tilde{H}(T, Y, \eta, \zeta) = \frac{Y^2 + 2BY + B^2}{2A}.$$

Furthermore, by plugging the ansatz:

$$\tilde{H}(t, Y, \eta, \zeta) = \tilde{v}_0(t, \eta, \zeta) + \tilde{v}_1(t, \eta, \zeta)Y + \tilde{v}_2(t, \eta, \zeta)Y^2,$$

and expressing the dual HJB equation as a polynomial of Y, we are able to get the similar PDEs for $\tilde{v}_0$, $\tilde{v}_1$, and $\tilde{v}_2$ as follows:

$$\begin{cases} \mathcal{D}\tilde{v}_2 + (\|\boldsymbol{\Theta}\|^2 - 2r)\tilde{v}_2 - \dfrac{a^2(\partial_\eta\tilde{v}_2)^2}{\epsilon\tilde{v}_2} - \dfrac{l^2\delta(\partial_\zeta\tilde{v}_2)^2}{\tilde{v}_2} = 0, \\ \mathcal{D}\tilde{v}_1 - r\tilde{v}_1 - \dfrac{a^2\partial_\eta\tilde{v}_1\partial_\eta\tilde{v}_2}{\epsilon\tilde{v}_2} - \dfrac{l^2\delta\partial_\zeta\tilde{v}_1\partial_\zeta\tilde{v}_2}{\tilde{v}_2} = 0, \\ \mathcal{D}\tilde{v}_0 - \dfrac{a^2(\partial_\eta\tilde{v}_1)^2}{4\epsilon\tilde{v}_2} - \dfrac{l^2\delta(\partial_\zeta\tilde{v}_1)^2}{4\tilde{v}_2} = 0, \end{cases} \tag{5.3.5}$$

with the terminal conditions:

$$\begin{cases} \tilde{v}_2(T, \eta, \zeta) = \dfrac{1}{2A}, \\ \tilde{v}_1(T, \eta, \zeta) = \dfrac{B}{A}, \\ \tilde{v}_0(T, \eta, \zeta) = \dfrac{B^2}{2A}. \end{cases}$$

Following the similar logic, it is not hard to observe that:

$$\begin{aligned} \tilde{v}_1(t) &= \frac{B}{A}\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s}, \\ \tilde{v}_0(t) &= \frac{B^2}{2A}. \end{aligned} \tag{5.3.6}$$

Chapter 6

FBSDE Method

With the FBSDE method, we are able to transform the problem of solving a stochastic control problem into a problem of solving a stochastic Hamiltonian system, which may be more tractable. In this section, we will derive both primal and dual FBSDE systems for the general uncorrelated setting, and we will also see the equivalence of the primal and dual FBSDE systems.

6.1 Primal Problem and the Associated FBSDE

In the previous primal problem setting, we have three state variables: X, η, and ζ. Hence, the associated adjoint process $\boldsymbol{p}_1(t) \in \mathbb{R}^3$ and the associated control $\boldsymbol{q}_1(t) \in \mathbb{R}^{3\times(N+2)}$. Before we write down the dynamic of $\boldsymbol{p}_1(t)$, we firstly simplify our discussion by setting no correlation structure and $K = \mathbb{R}^N$ and reformalise the dynamics of all the three state variables in a matrix form as follows:

$$
\begin{aligned}
&\mathrm{d}\begin{pmatrix} X^{\boldsymbol{\pi}} \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} X^{\boldsymbol{\pi}} r + \boldsymbol{\pi}'(\boldsymbol{\mu} - r\mathbf{1}) \\ \frac{b}{\epsilon} \\ c\delta \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \boldsymbol{\pi}'\boldsymbol{\sigma} & 0 & 0 \\ \mathbf{0} & \frac{a}{\sqrt{\epsilon}} & 0 \\ \mathbf{0} & 0 & \sqrt{\delta} l \end{pmatrix} \mathrm{d}\begin{pmatrix} \boldsymbol{W}_t \\ W_t^{\eta} \\ W_t^{\zeta} \end{pmatrix}, \\
&\begin{pmatrix} X^{\boldsymbol{\pi}} \\ \eta \\ \zeta \end{pmatrix}_0 = \begin{pmatrix} x_0 \\ \eta_0 \\ \zeta_0 \end{pmatrix}.
\end{aligned} \tag{6.1.1}
$$

According to Yong & Zhou (1999, Section 3.1, page 115), the following dynamics of $\boldsymbol{p}_1(t)$ admit a unique $\mathcal{F}_t$-adapted solution pair $(\boldsymbol{p}_1, \boldsymbol{q}_1)$:

$$
\begin{aligned}
\mathrm{d}\boldsymbol{p}_1(t) =& \Bigg\{ - \left[\frac{\partial\left(X^{\boldsymbol{\pi}} r + \boldsymbol{\pi}'(\boldsymbol{\mu} - r\mathbf{1}), \frac{b}{\epsilon}, c\delta\right)}{\partial(X^{\boldsymbol{\pi}}, \eta, \zeta)} \right]' \boldsymbol{p}_1(t) - \sum_{j=1}^{N} \left[\frac{\partial\left(\boldsymbol{\pi}'\boldsymbol{\sigma}^j, 0, 0\right)}{\partial(X^{\boldsymbol{\pi}}, \eta, \zeta)} \right]' \boldsymbol{q}_1^j(t) \\
& - \left[\frac{\partial\left(0, \frac{a}{\sqrt{\epsilon}}, 0\right)}{\partial(X^{\boldsymbol{\pi}}, \eta, \zeta)} \right]' \boldsymbol{q}_1^{N+1}(t) - \left[\frac{\partial\left(0, 0, \sqrt{\delta} l\right)}{\partial(X^{\boldsymbol{\pi}}, \eta, \zeta)} \right]' \boldsymbol{q}_1^{N+2}(t) \\
& + \nabla_{(X^{\boldsymbol{\pi}}, \eta, \zeta)} f(t, X^{\boldsymbol{\pi}}, \boldsymbol{\pi}) \Bigg\}\Bigg|_{\boldsymbol{\pi} = \boldsymbol{\pi}^*} \mathrm{d}t + \boldsymbol{q}_1(t) \mathrm{d}\begin{pmatrix} \boldsymbol{W}_t \\ W_t^{\eta} \\ W_t^{\zeta} \end{pmatrix} \\
=& \Bigg\{ - \begin{pmatrix} r & \boldsymbol{\pi}^{*\prime}\partial_\eta \boldsymbol{\mu} & \boldsymbol{\pi}^{*\prime}\partial_\zeta \boldsymbol{\mu} \\ 0 & \frac{\mathrm{d}_\eta b}{\epsilon} & 0 \\ 0 & 0 & \mathrm{d}_\zeta c\delta \end{pmatrix} \boldsymbol{p}_1(t) - \sum_{j=1}^{N} \begin{pmatrix} 0 & \boldsymbol{\pi}^{*\prime}\partial_\eta \boldsymbol{\sigma}^j & \boldsymbol{\pi}^{*\prime}\partial_\zeta \boldsymbol{\sigma}^j \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{q}_1^j(t) \\
& - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{\mathrm{d}_\eta a}{\sqrt{\epsilon}} & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{q}_1^{N+1}(t) - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathrm{d}_\zeta l\sqrt{\delta} \end{pmatrix} \boldsymbol{q}_1^{N+2}(t) \\
& + \begin{pmatrix} Q X^{\boldsymbol{\pi}^*} + \boldsymbol{S}'\boldsymbol{\pi}^* \\ 0 \\ 0 \end{pmatrix} \Bigg\} \mathrm{d}t + \boldsymbol{q}_1(t) \mathrm{d}\begin{pmatrix} \boldsymbol{W}_t \\ W_t^{\eta} \\ W_t^{\zeta} \end{pmatrix},
\end{aligned} \tag{6.1.2}
$$

with the terminal condition:

$$
\begin{aligned}
\boldsymbol{p}_1(T) =& -\nabla_{(X^{\boldsymbol{\pi}},\eta,\zeta)}\ g(X^{\boldsymbol{\pi}}(T))\big|_{\boldsymbol{\pi}=\boldsymbol{\pi}^*}\\
=& -\begin{pmatrix} AX^{\boldsymbol{\pi}^*}(T)+B \\ 0 \\ 0 \end{pmatrix},
\end{aligned}
$$

where:

$$
\begin{aligned}
\boldsymbol{\sigma} &= \left[\boldsymbol{\sigma}^1,\cdots,\boldsymbol{\sigma}^N\right],\\
\boldsymbol{q}_1(t) &= \left[\boldsymbol{q}_1^1(t),\cdots,\boldsymbol{q}_1^{N+2}(t)\right],
\end{aligned}
$$

and $\boldsymbol{\pi}^*$ is the optimal control. We now define the Hamiltonian using Yong & Zhou (1999, (3.10), page 116) to assist to write down the maximum principle:

$$
\begin{aligned}
\mathcal{H}(t,X^{\boldsymbol{\pi}},\boldsymbol{\pi},\boldsymbol{p}_1,\boldsymbol{q}_1) =& \left\langle \boldsymbol{p}_1, \begin{pmatrix} X^{\boldsymbol{\pi}}r+\boldsymbol{\pi}'(\boldsymbol{\mu}-r\boldsymbol{1}) \\ \frac{b}{\epsilon} \\ c\delta \end{pmatrix}\right\rangle + \operatorname{tr}\left[\boldsymbol{q}_1'\begin{pmatrix} \boldsymbol{\pi}'\boldsymbol{\sigma} & 0 & 0 \\ \boldsymbol{0} & \frac{a}{\sqrt{\epsilon}} & 0 \\ \boldsymbol{0} & 0 & \sqrt{\delta}l \end{pmatrix}\right]\\
&- f(t,X^{\boldsymbol{\pi}},\boldsymbol{\pi})\\
=&\boldsymbol{p}_1'\begin{pmatrix} X^{\boldsymbol{\pi}}r+\boldsymbol{\pi}'(\boldsymbol{\mu}-r\boldsymbol{1}) \\ \frac{b}{\epsilon} \\ c\delta \end{pmatrix} + \sum_{j=1}^{N}\boldsymbol{\pi}'\boldsymbol{\sigma}^j q_1^{j,1}\\
&+\frac{a}{\sqrt{\epsilon}}q_1^{(N+1),2}+l\sqrt{\delta}q_1^{(N+2),3}-\frac{1}{2}Q\left(X^{\boldsymbol{\pi}}\right)^2-\boldsymbol{S}'\boldsymbol{\pi}X^{\boldsymbol{\pi}}-\frac{1}{2}\boldsymbol{\pi}'\boldsymbol{R}\boldsymbol{\pi},
\end{aligned}
\tag{6.1.3}
$$

where $q_1^{j,i},\ j=1,\cdots,N+2,\ i=1,2,3$ is the element of $\boldsymbol{q}_1(t)$ at j^{th} column and i^{th} row. According to the local form of the maximum principle (Yong & Zhou 1999, (3.26), page 120), we have the relationship among $(X^{\boldsymbol{\pi}^*},\boldsymbol{\pi}^*,\boldsymbol{p}_1,\boldsymbol{q}_1)$:

$$
\begin{aligned}
&(\boldsymbol{\pi}-\boldsymbol{\pi}^*)'\nabla_{\boldsymbol{\pi}}\mathcal{H}(t,X^{\boldsymbol{\pi}^*},\boldsymbol{\pi}^*,\boldsymbol{p}_1,\boldsymbol{q}_1)\le 0,\\
\Longrightarrow&(\boldsymbol{\pi}^*-\boldsymbol{\pi})'\left[p_1^1(\boldsymbol{\mu}-r\boldsymbol{1})+\sum_{j=1}^{N}q_1^{j,1}\boldsymbol{\sigma}^j-X^{\boldsymbol{\pi}^*}\boldsymbol{S}-\boldsymbol{R}\boldsymbol{\pi}^*\right]\ge 0,
\end{aligned}
\tag{6.1.4}
$$

where:

$$
\boldsymbol{p}_1(t)=\begin{pmatrix} p_1^1 \\ p_1^2 \\ p_1^3 \end{pmatrix},
$$

and $\boldsymbol{\pi}\in K$ is an arbitrary control. As we previously assumed that $K=\mathbb{R}^N$, (6.1.4) can be simplified further as follows:

$$
p_1^1(\boldsymbol{\mu}-r\boldsymbol{1})+\sum_{j=1}^{N}q_1^{j,1}\boldsymbol{\sigma}^j-X^{\boldsymbol{\pi}^*}\boldsymbol{S}-\boldsymbol{R}\boldsymbol{\pi}^*=0. \tag{6.1.5}
$$

In fact, we can simplify the primal FBDSE system by getting rid of the last two dimensions of $\boldsymbol{p}_1$. From (6.1.2), we immediately know that:

$$
\begin{aligned}
&q_1^{j,i}=0,\forall j=1,\cdots,N+2,\ i=2,3,\\
&p_1^2=0,\\
&p_1^3=0,
\end{aligned}
$$

for all $t\in[0,T]$ satisfy (6.1.2) for the last two dimensions, which is then guaranteed by the uniqueness of the solution under some technical conditions according to Yong & Zhou (1999, page 116). More specifically, the uniqueness of the solution is guaranteed if the following assumptions are satisfied:

Assumption 6.1.1. *$(K,\tilde{d})$ is a separable metric space and $T>0$.*

Assumption 6.1.2. *The maps:*

$$\begin{aligned}
\tilde{\boldsymbol{\mu}}_1(t,x,\eta,\zeta,\boldsymbol{\pi}) &= \begin{pmatrix} xr(t)+\boldsymbol{\pi}'[\boldsymbol{\mu}(\eta,\zeta)-r(t)\mathbf{1}] \\ \frac{b(\eta)}{\epsilon} \\ c(\zeta)\delta \end{pmatrix}, \\
\tilde{\boldsymbol{\sigma}}_1(t,x,\eta,\zeta,\boldsymbol{\pi}) &= \begin{pmatrix} \boldsymbol{\pi}'\boldsymbol{\sigma}(\eta,\zeta) & 0 & 0 \\ \mathbf{0} & \frac{a(\eta)}{\sqrt{\epsilon}} & 0 \\ \mathbf{0} & 0 & \sqrt{\delta}l(\zeta) \end{pmatrix}, \\
f(t,x,\eta,\zeta,\boldsymbol{\pi}) &= \frac{1}{2}Qx^2+\boldsymbol{S}'\boldsymbol{\pi}x+\frac{1}{2}\boldsymbol{\pi}'\boldsymbol{R}\boldsymbol{\pi}, \\
g(t,x,\eta,\zeta,\boldsymbol{\pi}) &= \frac{1}{2}Ax^2+Bx,
\end{aligned}$$

are measurable, and there exists a constant $L>0$ and a modulus of continuity $\bar{\omega}:[0,\infty)\longrightarrow[0,\infty)$ such that for $\psi(t,x,\eta,\zeta,\boldsymbol{\pi})\in\{\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\sigma}},f,g\}$, we have:

$$\begin{cases} |\psi(t,x,\eta,\zeta,\boldsymbol{\pi})-\psi(t,\hat{x},\hat{\eta},\hat{\zeta},\hat{\boldsymbol{\pi}})| \le L\left|\begin{pmatrix} x\\ \eta\\ \zeta\end{pmatrix}-\begin{pmatrix}\hat{x}\\ \hat{\eta}\\ \hat{\zeta}\end{pmatrix}\right| + \bar{\omega}(\tilde{d}(\boldsymbol{\pi},\hat{\boldsymbol{\pi}})), \\ |\psi(t,0,0,0,\boldsymbol{\pi})|\le L, \end{cases}$$

for all $t\in[0,T]$, $x,\hat{x}\in\mathbb{R}$, $\eta,\hat{\eta}\in\mathbb{R}$, $\zeta,\hat{\zeta}\in\mathbb{R}$, and $\boldsymbol{\pi},\hat{\boldsymbol{\pi}}\in\mathbb{R}^N$.

Assumption 6.1.3. *The maps $\tilde{\boldsymbol{\pi}}$, $\tilde{\boldsymbol{\sigma}}$, f, and g are C^2 in $(x,\eta,\zeta)'$. Moreover, there exists a constant $L>0$ and a modulus of continuity $\omega:[0,\infty)\longrightarrow[0,\infty)$ such that for $\psi(t,x,\eta,\zeta,\boldsymbol{\pi})\in\{\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\sigma}},f,g\}$, we have:*

$$\begin{cases} |\nabla_{(x,\eta,\zeta)}\psi(t,x,\eta,\zeta,\boldsymbol{\pi})-\nabla_{(x,\eta,\zeta)}\psi(t,\hat{x},\hat{\eta},\hat{\zeta},\hat{\boldsymbol{\pi}})| \le L\left|\begin{pmatrix} x\\ \eta\\ \zeta\end{pmatrix}-\begin{pmatrix}\hat{x}\\ \hat{\eta}\\ \hat{\zeta}\end{pmatrix}\right| + \bar{\omega}(\tilde{d}(\boldsymbol{\pi},\hat{\boldsymbol{\pi}})), \\ |\nabla\nabla_{(x,\eta,\zeta)}\psi(t,x,\eta,\zeta,\boldsymbol{\pi})-\nabla\nabla_{(x,\eta,\zeta)}\psi(t,\hat{x},\hat{\eta},\hat{\zeta},\hat{\boldsymbol{\pi}})| \le \bar{\omega}\left(\left|\begin{pmatrix} x\\ \eta\\ \zeta\end{pmatrix}-\begin{pmatrix}\hat{x}\\ \hat{\eta}\\ \hat{\zeta}\end{pmatrix}\right| + \tilde{d}(\boldsymbol{\pi},\hat{\boldsymbol{\pi}})\right), \end{cases}$$

for all $t\in[0,T]$, $x,\hat{x}\in\mathbb{R}$, $\eta,\hat{\eta}\in\mathbb{R}$, $\zeta,\hat{\zeta}\in\mathbb{R}$, and $\boldsymbol{\pi},\hat{\boldsymbol{\pi}}\in\mathbb{R}^N$.

We can thus simplify the system into the general version and the unrestricted version by removing η and ζ:

1. The general version:

$$\begin{cases} \mathrm{d}X^{\boldsymbol{\pi}^*} = [X^{\boldsymbol{\pi}^*}r+\boldsymbol{\pi}^{*\prime}(\boldsymbol{\mu}-r\mathbf{1})]\mathrm{d}t+\boldsymbol{\pi}^{*\prime}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{W}_t, \\ X^{\boldsymbol{\pi}^*}(0)=x_0, \\ \mathrm{d}\hat{p}_1 = [-r\hat{p}_1+QX^{\boldsymbol{\pi}^*}+\boldsymbol{S}'\boldsymbol{\pi}^*]\mathrm{d}t+\hat{\boldsymbol{q}}_1'\mathrm{d}[\boldsymbol{W}_t,W_t^{\eta},W_t^{\zeta}]', \\ \hat{p}_1(T) = -AX^{\boldsymbol{\pi}^*}(T)-B, \end{cases} \tag{6.1.6}$$

   and the maximum principle:

$$(\boldsymbol{\pi}^*-\boldsymbol{\pi})'\left\{\hat{p}_1(\boldsymbol{\mu}-r\mathbf{1})+\boldsymbol{\sigma}[\boldsymbol{I}_{N\times N},\ \mathbf{0}_{N\times 2}]\hat{\boldsymbol{q}}_1-X^{\boldsymbol{\pi}^*}\boldsymbol{S}-\boldsymbol{R}\boldsymbol{\pi}^*\right\}\ge 0, \tag{6.1.7}$$

   where:

$$\begin{aligned} \hat{p}_1(t) &= p_1^1(t)\in\mathbb{R}, \\ \hat{\boldsymbol{q}}_1(t) &= [q_1^{1,1},\cdots,q_1^{N+2,1}]'(t)\in\mathbb{R}^{N+2}, \end{aligned}$$

   and $\boldsymbol{\pi}\in K$ is an arbitrary control.

2. The unrestricted version ($K=\mathbb{R}^N$): the dynamics of the system is the same as the general version, and the only difference is the maximum principle:

$$\hat{p}_1(\boldsymbol{\mu}-r\mathbf{1})+\boldsymbol{\sigma}[\boldsymbol{I}_{N\times N},\ \mathbf{0}_{N\times 2}]\hat{\boldsymbol{q}}_1-X^{\boldsymbol{\pi}^*}\boldsymbol{S}-\boldsymbol{R}\boldsymbol{\pi}^*=0. \tag{6.1.8}$$

Before we move on to the dual problem, let us further discuss the maximum principle under the assumption $K = \mathbb{R}^N$. It is clear that the maximum principle provides some information among the state variables, the optimal primal control, the associated adjoint process, and the associated control. More importantly, if we further assume $\boldsymbol{S} = \boldsymbol{0}$, then we can solve $\boldsymbol{\pi}^*$ from (6.1.8) explicitly:

$$\boldsymbol{\pi}^* = \boldsymbol{R}^{-1} \left\{ \hat{p}_1(\boldsymbol{\mu} - r\boldsymbol{1}) + \boldsymbol{\sigma}[\boldsymbol{I}_{N\times N},\ \boldsymbol{0}_{N\times 2}]\hat{\boldsymbol{q}}_1 \right\},$$

which directly links the associated adjoint process and control to the optimal primal control. However, one may ask what will happen if $\boldsymbol{S}$ and $\boldsymbol{R}$ are set to zero simultaneously? It seems that (6.1.8) no longer provides such a link aforementioned, and thus the system seems to be decoupled. Nevertheless, we should notice that the last condition in (6.1.6) reminds us that we still need to choose a control $\boldsymbol{\pi}$ to make sure that the terminal value of $\hat{p}_1$ and X matched in the specific way, and therefore, the system is actually not decoupled even when we set $\boldsymbol{S}$ and $\boldsymbol{R}$ to be zero simultaneously.

6.2 Dual Problem and the Associated FBSDE

Similar to the primal problem, we firstly reformalise the dynamics of all the three state variables as follows:

$$\begin{aligned}
\mathrm{d}\begin{pmatrix} Y \\ \eta \\ \zeta \end{pmatrix} &= \begin{pmatrix} \alpha - rY \\ \frac{b}{\epsilon} \\ c\delta \end{pmatrix} \mathrm{d}t + \begin{pmatrix} [\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\boldsymbol{1})]'\boldsymbol{\sigma}'^{-1} & \gamma & \xi \\ \boldsymbol{0} & \frac{a}{\sqrt{\epsilon}} & 0 \\ \boldsymbol{0} & 0 & \sqrt{\delta} l \end{pmatrix} \mathrm{d}\begin{pmatrix} \boldsymbol{W}_t \\ W_t^\eta \\ W_t^\zeta \end{pmatrix}, \\
\begin{pmatrix} Y \\ \eta \\ \zeta \end{pmatrix}_0 &= \begin{pmatrix} y_0 \\ \eta_0 \\ \zeta_0 \end{pmatrix},
\end{aligned} \tag{6.2.1}$$

and analogously, the associated adjoint process of the dual problem (i.e. $\boldsymbol{p}_2(t)$) should be three dimensional:

$$\begin{aligned}
\mathrm{d}\boldsymbol{p}_2(t) = &\Bigg\{ - \left[\frac{\partial\left(\alpha - rY, \frac{b}{\epsilon}, c\delta\right)}{\partial(Y,\eta,\zeta)} \right]' \boldsymbol{p}_2(t) - \sum_{j=1}^N \left[\frac{\partial\left([\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\boldsymbol{1})]'(\boldsymbol{\sigma}'^{-1})^j, 0, 0\right)}{\partial(Y,\eta,\zeta)} \right]' \boldsymbol{q}_2^j(t) \\
&- \left[\frac{\partial\left(\gamma, \frac{a}{\sqrt{\epsilon}}, 0\right)}{\partial(Y,\eta,\zeta)} \right]' \boldsymbol{q}_2^{N+1}(t) - \left[\frac{\partial\left(\gamma, 0, \sqrt{\delta} l\right)}{\partial(Y,\eta,\zeta)} \right]' \boldsymbol{q}_2^{N+2}(t) \\
&+ \nabla_{(Y,\eta,\zeta)}\phi(t,\alpha,\boldsymbol{\beta}) \Bigg\}\Bigg|_{y_0=y_0^*,\xi=\xi^*,\gamma=\gamma^*,\boldsymbol{\beta}=\boldsymbol{\beta}^*,\alpha=\alpha^*} \mathrm{d}t + \boldsymbol{q}_2(t)\mathrm{d}\begin{pmatrix} \boldsymbol{W}_t \\ W_t^\eta \\ W_t^\zeta \end{pmatrix} \\
= &\Bigg\{ - \begin{pmatrix} -r & 0 & 0 \\ 0 & \frac{\mathrm{d}_\eta b}{\epsilon} & 0 \\ 0 & 0 & \mathrm{d}_\zeta c\delta \end{pmatrix} \boldsymbol{p}_2(t) \\
&- \sum_{j=1}^N \begin{pmatrix} -(\boldsymbol{\mu} - r\boldsymbol{1})'(\boldsymbol{\sigma}'^{-1})^j & -Y(\partial_\eta\boldsymbol{\mu})'(\boldsymbol{\sigma}'^{-1})^j & -Y(\partial_\zeta\boldsymbol{\mu})'(\boldsymbol{\sigma}'^{-1})^j \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{q}_2^j(t) \\
&- \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{\mathrm{d}_\eta a}{\sqrt{\epsilon}} & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{q}_2^{N+1}(t) - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathrm{d}_\zeta l\sqrt{\delta} \end{pmatrix} \boldsymbol{q}_2^{N+2}(t) \\
&+ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Bigg\} \mathrm{d}t + \boldsymbol{q}_2(t)\mathrm{d}\begin{pmatrix} \boldsymbol{W}_t \\ W_t^\eta \\ W_t^\zeta \end{pmatrix},
\end{aligned} \tag{6.2.2}$$

with the terminal condition:

$$\begin{aligned}
\boldsymbol{p}_2(T) &= -\nabla_{(Y,\eta,\zeta)}\ m_T(Y(T))\big|_{y_0=y_0^*,\xi=\xi^*,\gamma=\gamma^*,\boldsymbol{\beta}=\boldsymbol{\beta}^*,\alpha=\alpha^*} \\
&= -\begin{pmatrix} \frac{Y^*(T)+B}{A} \\ 0 \\ 0 \end{pmatrix},
\end{aligned}$$

where:

$$\boldsymbol{\sigma}'^{-1} = \left[(\boldsymbol{\sigma}'^{-1})^1, \cdots, (\boldsymbol{\sigma}'^{-1})^N\right],$$
$$\boldsymbol{q}_2(t) = \left[\boldsymbol{q}_2^1(t), \cdots, \boldsymbol{q}_2^{N+2}(t)\right].$$

It is straightforward to see that:

$$\begin{aligned}
&q_2^{j,i} = 0, \forall j = 1, \cdots, N+2,\ i = 2, 3,\\
&q_2^{N+1,1} = 0,\\
&q_2^{N+2,1} = 0,\\
&p_2^2 = 0,\\
&p_2^3 = 0,
\end{aligned}$$

for all $t \in [0, T]$ satisfy (6.2.2) for the last two dimensions, which is then guaranteed by the uniqueness of the solution under Assumption 6.1.1 to Assumption 6.1.3 where the state variables become (Y, η, ζ), controls become $(\alpha, \boldsymbol{\beta}, \gamma, \xi, y_0)$, and the maps become:

$$\begin{aligned}
\tilde{\boldsymbol{\mu}}_2(t, y, \eta, \zeta, \alpha, \boldsymbol{\beta}, \gamma, \xi, y_0) &= \begin{pmatrix} \alpha - r(t)y \\ \frac{b(\eta)}{\epsilon} \\ c(\zeta)\delta \end{pmatrix},\\
\tilde{\boldsymbol{\sigma}}_2(t, y, \eta, \zeta, \alpha, \boldsymbol{\beta}, \gamma, \xi, y_0) &= \begin{pmatrix} \{\boldsymbol{\beta} - Y[\boldsymbol{\mu}(\eta,\zeta) - r(t)\mathbf{1}]\}'\boldsymbol{\sigma}'^{-1}(\eta,\zeta) & \gamma & \xi \\ \mathbf{0} & \frac{a(\eta)}{\sqrt{\epsilon}} & 0 \\ \mathbf{0} & 0 & \sqrt{\delta}l(\zeta) \end{pmatrix},\\
\phi(t, y, \eta, \zeta, \alpha, \boldsymbol{\beta}, \gamma, \xi, y_0) &= \frac{1}{2}\frac{(\alpha - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta})^2}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \frac{1}{2}\boldsymbol{\beta}'\boldsymbol{R}^{-1}\boldsymbol{\beta},\\
m_T(t, y, \eta, \zeta, \alpha, \boldsymbol{\beta}, \gamma, \xi, y_0) &= \frac{y^2 + 2By + B^2}{2A}.
\end{aligned}$$

Therefore, we can simplify the system by removing η and ζ and adapt to the similar notation as previous discussion:

$$\begin{cases}
\mathrm{d}Y^* = (\alpha^* - rY^*)\mathrm{d}t + [\boldsymbol{\beta}^* - Y^*(\boldsymbol{\mu} - r\mathbf{1})]'\boldsymbol{\sigma}'^{-1}\mathrm{d}\boldsymbol{W}_t + \gamma^*\mathrm{d}W_t^\eta + \xi^*\mathrm{d}W_t^\zeta,\\
Y^*(0) = y_0^*,\\
\mathrm{d}\hat{p}_2 = [r\hat{p}_2 + \hat{\boldsymbol{q}}_2'\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})]\mathrm{d}t + \hat{\boldsymbol{q}}_2'\mathrm{d}\boldsymbol{W}_t,\\
\hat{p}_2(T) = -\dfrac{Y^*(T) + B}{A},
\end{cases} \tag{6.2.3}$$

where:

$$\hat{p}_2(t) = p_2^1(t) \in \mathbb{R},$$
$$\hat{\boldsymbol{q}}_2(t) = [q_2^{1,1}, \cdots, q_2^{N,1}]'(t) \in \mathbb{R}^N,$$

and:

$$Y^* = Y^{(y_0^*, \alpha^*, \boldsymbol{\beta}^*, \gamma^*, \xi^*)},$$

is the state variable under the optimal dual controls. In this case, the dimensionality of the $\hat{\boldsymbol{q}}_2$ is different from $\hat{\boldsymbol{q}}_1(t)$.

To proceed to the optimality of Y^*, we need to make sure that the following Assumption holds:

Assumption 6.2.1. *Let $(\hat{\alpha}, \hat{\boldsymbol{\beta}})$ be given, and let α, $\boldsymbol{\beta}$ be any admissible controls. Then there exists a stochastic process Z, which it is a real valued progressively measurable process on $[0, T] \times \Omega$, satisfying:*

$$\mathbb{E}\left[\int_0^T |Z(t)|\mathrm{d}t\right] < \infty,$$

and:

$$Z(t) \geq \frac{\phi(t, \hat{\alpha}(t) + \epsilon^*\alpha(t), \hat{\boldsymbol{\beta}}(t) + \epsilon^*\boldsymbol{\beta}(t)) - \phi(t, \hat{\alpha}(t), \hat{\boldsymbol{\beta}})}{\epsilon^*},$$

for $(\mathbb{P} \otimes Leb)$-a.e. $(\omega, t) \in \Omega \times [0, T]$, and $\epsilon^* \in (0, 1]$.

The optimality of Y^* is then guaranteed by the following sufficient and necessary condition, which serves as the 'maximum principle', according to Li & Zheng (2018, Theorem 7, page 1138) and Rockafellar (1970, Theorem 23.5, page 218):

$$\begin{cases} \hat{p}_2(0) = x_0, \\ \boldsymbol{\sigma}'^{-1}\hat{\boldsymbol{q}}_2 \in K, \\ (\hat{p}_2, \boldsymbol{\sigma}'^{-1}\hat{\boldsymbol{q}}_2) \in \partial\phi(\alpha^*, \boldsymbol{\beta}^*), \end{cases} \tag{6.2.4}$$

where $\partial\phi(\alpha^*, \boldsymbol{\beta}^*)$ is the set of generalised gradients evaluated at $(\alpha^*, \boldsymbol{\beta}^*)$.

Remark 6.2.2. *If $K = \mathbb{R}^N$, we can express* (6.2.4) *more explicitly. From* (4.1.11)*, we know the xplicit form of ϕ:*

$$\phi(t, \alpha, \boldsymbol{\beta}) = \frac{1}{2}\frac{(\alpha - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta})^2}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \frac{1}{2}\boldsymbol{\beta}'\boldsymbol{R}^{-1}\boldsymbol{\beta},$$

and thus, we obtain two equations as follows using the last condition in (6.2.4)*, which indicates that $(\hat{p}_2, \boldsymbol{\sigma}'^{-1}\hat{\boldsymbol{q}}_2)'$ is the gradient of ϕ evaluated at $(\alpha^*, \boldsymbol{\beta}^*)$:*

$$\begin{cases} \hat{p}_2 = \dfrac{\alpha^* - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}^*}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}}, \\ \boldsymbol{\sigma}'\hat{\boldsymbol{q}}_2 = \boldsymbol{R}^{-1}\boldsymbol{\beta}^* - \dfrac{\boldsymbol{R}^{-1}\boldsymbol{S}(\alpha^* - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}^*)}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}}. \end{cases} \tag{6.2.5}$$

Remark 6.2.3. *The above explicit conditions* (6.2.5) *can be also derived from the maximum principle similar to the previous argument using the Hamiltonian. To verify this, we define the Hamiltonian as follows:*

$$\begin{aligned} \mathcal{H}(t, Y, \alpha, \boldsymbol{\beta}, \gamma, \xi, \hat{p}_2, \hat{\boldsymbol{q}}_2) =& \langle \hat{p}_2, \alpha - rY \rangle + \mathrm{tr}\left[\hat{\boldsymbol{q}}_2[\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'\boldsymbol{\sigma}'^{-1}\right] \\ &- \phi(t, \alpha, \boldsymbol{\beta}) \\ =&(\alpha - rY)\hat{p}_2 + [\boldsymbol{\beta} - Y(\boldsymbol{\mu} - r\mathbf{1})]'\boldsymbol{\sigma}'^{-1}\hat{\boldsymbol{q}}_2 \\ &- \frac{1}{2}\frac{(\alpha - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta})^2}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} + \frac{1}{2}\boldsymbol{\beta}'\boldsymbol{R}^{-1}\boldsymbol{\beta}. \end{aligned} \tag{6.2.6}$$

We differentiate it with respect to α and $\boldsymbol{\beta}$ (given that $\mathcal{H}$ does not explicitly depend on γ and ξ) at $(\alpha^*, \boldsymbol{\beta}^*)$:

$$\nabla_{(\alpha,\boldsymbol{\beta})}\mathcal{H}(t, Y, \alpha^*, \boldsymbol{\beta}^*, \hat{p}_2, \hat{\boldsymbol{q}}_2) = \begin{pmatrix} \hat{p}_2 - \frac{\alpha^* - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}^*}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} \\ \boldsymbol{\sigma}'\hat{\boldsymbol{q}}_2 - \boldsymbol{R}^{-1}\boldsymbol{\beta}^* - \frac{\boldsymbol{R}^{-1}\boldsymbol{S}(\alpha^* - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}^*)}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}} \end{pmatrix},$$

and equate them to zero (when $K = \mathbb{R}^N$), and will get the exact results shown in (6.2.5)*. Following the similar argument as the previous section, although γ^* and ξ^* do not appear in the above equations (or more specifically, in $\mathcal{H}$), it does not necessarily mean that these two controls have no effect or can be chosen arbitrarily. Instead, their values will affect $Y^*(T)$ and then consequently the terminal condition of $\hat{p}_2$.*

6.3 Links between the primal and dual FBSDEs

From (6.2.5), we can observe that the right hand sides of these equations exactly match with those of (4.1.10). It is also not difficult to figure out that the structure of the drift term of the dynamics of $\hat{p}_2$ is quite similar to that of $X^{\boldsymbol{\pi}^*}$, and the structure of the drift term of the dynamics of Y^* is quite similar to that of $\hat{p}_1$. Indeed, the relationship is guaranteed by Li & Zheng (2018, Theorem 9, page 1139) and Li & Zheng (2018, Theorem 11, page 1139). More specifically, we have the following two theorems:

Theorem 6.3.1 (from dual problem to primal problem)**.** *Suppose that $(y_0^*, \alpha^*, \boldsymbol{\beta}^*, \gamma^*, \xi^*)$ is optimal for the dual problem. Let $(Y^*, \hat{p}_2, \hat{\boldsymbol{q}}_2)$ be the associated process that satisfies the FBSDE* (6.2.3) *and condition* (6.2.4)*. Define:*

$$\boldsymbol{\pi}^* = \boldsymbol{\sigma}'^{-1}\hat{\boldsymbol{q}}_2. \tag{6.3.1}$$

Then $\boldsymbol{\pi}^*$ *is the optimal control for the primal problem with the initial total wealth* x_0, *and the optimal total wealth process* $X^{\boldsymbol{\pi}^*}$ *and the associated adjoint processes* $(\hat{p}_1, \hat{\boldsymbol{q}}_2)$ *are given by the following equations:*

$$\begin{cases} X^{\boldsymbol{\pi}^*} = \hat{p}_2, \\ \hat{p}_1 = Y^*, \\ \hat{\boldsymbol{q}}_1 = \left[[\boldsymbol{\beta}^* - Y^*(\boldsymbol{\mu} - r\mathbf{1})]' \boldsymbol{\sigma}'^{-1}, \gamma^*, \xi^* \right]'. \end{cases} \tag{6.3.2}$$

Theorem 6.3.2 (from primal problem to dual problem)**.** *Suppose that* $\boldsymbol{\pi}^* \in K$ *is optimal for the primal problem with the initial total wealth* x_0. *Let* $(X^{\boldsymbol{\pi}^*}, \hat{p}_1, \hat{\boldsymbol{q}}_1)$ *be the associated process that satisfies the FBSDE* (6.1.6) *and condition* (6.1.7). *Define:*

$$\begin{cases} y_0^* = \hat{p}_1(0), \\ \alpha^* = QX^{\boldsymbol{\pi}^*} + \boldsymbol{S}'\boldsymbol{\pi}^*, \\ \boldsymbol{\beta}^* = \boldsymbol{\sigma}[\boldsymbol{I}_{N\times N},\ \mathbf{0}_{N\times 2}]\hat{\boldsymbol{q}}_1 + Y^*(\boldsymbol{\mu} - r\mathbf{1}), \\ \gamma^* = q_1^{N+1,1}, \\ \xi^* = q_1^{N+2,1}. \end{cases} \tag{6.3.3}$$

Then $(y_0^*, \alpha^*, \boldsymbol{\beta}^*, \gamma^*, \xi^*)$*: is the optimal control for the dual problem. The optimal dual state process* Y^* *and associated adjoint processes* $(\hat{p}_2, \hat{\boldsymbol{q}}_2)$ *are given by:*

$$\begin{cases} Y^* = \hat{p}_1, \\ \hat{p}_2 = X^{\boldsymbol{\pi}^*}, \\ \hat{\boldsymbol{q}}_2 = \boldsymbol{\sigma}'\boldsymbol{\pi}^*. \end{cases} \tag{6.3.4}$$

Chapter 7

Numerical Methods for HJB equations and FBSDEs

According to the previous discussion, our problem can bearly be solved analytically, even after making the functionals in the setup explicitly, for example, $\boldsymbol{\mu}(\eta, \zeta)$, $\boldsymbol{\sigma}(\eta, \zeta)$, $a(\eta)$, $b(\eta)$, $c(\zeta)$, $l(\zeta)$, etc. Thus, we have to resort to the numerical methods to solve the problem. In this chapter, we will walk through some basic ideas of discretising the PDEs and the FBSDE systems, and then focus on the finite difference method for solving PDEs for the HJB equations and transforming solving FBSDEs into solving optimisations. We will only discuss the general case without introducing the correlation structure. Lastly, we will also specify a setup for our problem, solve it through the primal HJB equation and the dual HJB equation, and verify the equivalence of all the four methods we discussed before.

7.1 Solving the Primal and Dual HJB equations Numerically

This section is purely based on the finite difference method to solve PDEs numerically, which is an intuitive and computationally efficient way to solve PDEs. We will discuss the primal HJB equation and the dual HJB equation separately.

7.1.1 Primal HJB Equation

For the primal HJB equation (3.0.3), we discretise the PDEs by t, η, and ζ, namely, by making the three functions to be solved into three separate three-dimensional grids and initialising the last two-dimensional grids (i.e. at $t = T$) by each function's terminal conditions. One grid should be formed by $\Delta t * \Delta \eta * \Delta \zeta$, where:

$$
\begin{aligned}
\Delta t &= \frac{T}{N_t}, \\
\Delta \eta &= \frac{\eta_{max} - \eta_{min}}{N_\eta}, \\
\Delta \zeta &= \frac{\zeta_{max} - \zeta_{min}}{N_\zeta}.
\end{aligned}
\tag{7.1.1}
$$

And we propagate the values through the PDEs backwardly as follows:

$$\begin{cases} v_2^{i,j,k-1} = & v_2^{i,j,k} + \dfrac{b^i}{\epsilon}\dfrac{v_2^{i+1,j,k} - v_2^{i,j,k}}{\Delta\eta} + \dfrac{(a^i)^2}{2\epsilon}\dfrac{v_2^{i+1,j,k} + v_2^{i-1,j,k} - 2v_2^{i,j,k}}{\Delta\eta^2} \\ & + \delta c^j \dfrac{v_2^{i,j+1,k} - v_2^{i,j,k}}{\Delta\zeta} + \dfrac{1}{2}\delta(l^j)^2 \dfrac{v_2^{i,j+1,k} + v_2^{i,j-1,k} - 2v_2^{i,j,k}}{\Delta\zeta^2} \\ & + 2v_2^{i,j,k} r^k + \dfrac{1}{2}Q^k - \dfrac{1}{2}\left[\boldsymbol{S}^k + 2v_2^{i,j,k}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})\right]' \\ & \left[\boldsymbol{R}^k + 2v_2^{i,j,k}\boldsymbol{\sigma}^{i,j,k}(\boldsymbol{\sigma}^{i,j,k})'\right]^{-1}\left[\boldsymbol{S}^k + 2v_2^{i,j,k}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})\right], \\ v_1^{i,j,k-1} = & v_1^{i,j,k} + \dfrac{b^i}{\epsilon}\dfrac{v_1^{i+1,j,k} - v_1^{i,j,k}}{\Delta\eta} + \dfrac{(a^i)^2}{2\epsilon}\dfrac{v_1^{i+1,j,k} + v_1^{i-1,j,k} - 2v_1^{i,j,k}}{\Delta\eta^2} \\ & + \delta c^j \dfrac{v_1^{i,j+1,k} - v_1^{i,j,k}}{\Delta\zeta} + \dfrac{1}{2}\delta(l^j)^2 \dfrac{v_1^{i,j+1,k} + v_1^{i,j-1,k} - 2v_1^{i,j,k}}{\Delta\zeta^2} \\ & + \left\{ r^k - (\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})'\left[\boldsymbol{R}^k + 2v_2^{i,j,k}\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1}\right. \\ & \left.\left[\boldsymbol{S}^k + 2v_2^{i,j,k}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})\right]\right\} v_1^{i,j,k}, \\ v_0^{i,j,k-1} = & v_0^{i,j,k} + \dfrac{b^i}{\epsilon}\dfrac{v_0^{i+1,j,k} - v_0^{i,j,k}}{\Delta\eta} + \dfrac{(a^i)^2}{2\epsilon}\dfrac{v_0^{i+1,j,k} + v_0^{i-1,j,k} - 2v_0^{i,j,k}}{\Delta\eta^2} \\ & + \delta c^j \dfrac{v_0^{i,j+1,k} - v_0^{i,j,k}}{\Delta\zeta} + \dfrac{1}{2}\delta(l^j)^2 \dfrac{v_0^{i,j+1,k} + v_0^{i,j-1,k} - 2v_0^{i,j,k}}{\Delta\zeta^2} \\ & - \dfrac{1}{2}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})'\left[\boldsymbol{R}^k + 2v_2^{i,j,k}\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})(v_1^{i,j,k})^2, \end{cases} \tag{7.1.2}$$

with the terminal conditions:

$$\begin{cases} v_0(t_{N_t}, \eta_i, \zeta_j) = 0, \\ v_1(t_{N_t}, \eta_i, \zeta_j) = B, \\ v_2(t_{N_t}, \eta_i, \zeta_j) = \dfrac{1}{2}A, \end{cases}$$
$$\forall\ i = 0, \cdots, N_\eta,\ j = 0, \cdots, N_\zeta,$$

where:

$$\begin{aligned} & v_2^{i,j,k} = v_2(t_k, \eta_i, \zeta_j), \\ & v_1^{i,j,k} = v_1(t_k, \eta_i, \zeta_j), \\ & v_0^{i,j,k} = v_0(t_k, \eta_i, \zeta_j), \\ & Q^k = Q(t_k); \\ & \boldsymbol{S}^k = \boldsymbol{S}(t_k), \\ & \boldsymbol{R}^k = \boldsymbol{R}(t_k), \\ & r^k = r(t_k), \\ & \boldsymbol{\mu}^{i,j} = \boldsymbol{\mu}(\eta_i, \zeta_j), \\ & \boldsymbol{\sigma}^{i,j} = \boldsymbol{\sigma}(\eta_i, \zeta_j), \\ & a^i = a(\eta_i), \\ & b^i = b(\eta_i), \\ & c^i = c(\zeta_j), \\ & l^i = l(\zeta_j), \\ & \eta_i = \eta_{min} + i\Delta\eta,\ \zeta_j = \zeta_{min} + j\Delta\zeta,\ t_k = k\Delta t, \\ & i = 0, \cdots, N_\eta,\ j = 0, \cdots, N_\zeta,\ k = 0, \cdots, N_t. \end{aligned}$$

Here are more detailed descriptions of the procedure:

1. It is not difficult to spot that if we only care about the optimal control $\boldsymbol{\pi}^*$ and the optimal total wealth path $X^{\boldsymbol{\pi}^*}$, we only need to solve v_2 and v_1 from (7.1.2). And we will only focus on v_2 and v_1 in the following discussion.

2. As the finite difference scheme involves computational errors and tends to accumulate them through the timeline, we need to ensure the stability of the numerical calculation, and a rough rule is to make:
$$\Delta t \sim \Delta\eta * \Delta\zeta, \tag{7.1.3}$$
according to LeVeque (1998, Section 12.3, page 157).

3. We need to know η_{min}, η_{max}, ζ_{min}, and ζ_{max}. Since the dynamics of η and ζ are independent of controls, we can directly simulate them $N_{simulation}$ times using the following discretisation:
$$\begin{aligned} \eta_{k+1}^m &= \eta_k^m + \frac{b(\eta_k^m)}{\epsilon}\Delta t + \frac{a(\eta_k^m)}{\sqrt{\epsilon}}\sqrt{\Delta t}Z_{k,m}^{\eta}, \\ \zeta_{k+1}^m &= \zeta_k^m + \delta c(\zeta_k^m)\Delta t + \sqrt{\delta}l(\zeta_k^m)\sqrt{\Delta t}Z_{k,m}^{\zeta}, \end{aligned} \tag{7.1.4}$$
where $Z_{k,m}^{\eta}$ and $Z_{k,m}^{\zeta}$ are the k^{th} realisation of the standard normal distribution independently in the m^{th} simulation, and η_k^m and ζ_k^m are the values of η and ζ at t_k in the m^{th} simulation. By (7.1.4), we can extract the range of η and ζ.

4. At each t_k, we will walk through all possible combinations of (η_i, ζ_j) to fulfill $v_1^{i,j,k}$ and $v_2^{i,j,k}$ based on the later time spot (i.e. $v_1^{i,j,k+1}$, $v_1^{i+1,j,k+1}$, $v_1^{i-1,j,k+1}$, $v_1^{i,j+1,k+1}$, $v_1^{i,j-1,k+1}$, $v_2^{i,j,k+1}$, $v_2^{i+1,j,k+1}$, $v_2^{i-1,j,k+1}$, $v_2^{i,j+1,k+1}$, $v_2^{i,j-1,k+1}$). When it comes to the upper boundaries of the grid (i.e. when $i = N_\eta$ or $j = N_\zeta$), we are no longer able to compute the partial derivatives with respect to η and ζ using the 'a-step-backward' approximation demonstrated in (7.1.2). Instead, we need to approximate the value of the partial derivatives using the values a steps away from the boundaries:
$$\begin{aligned} \partial_\eta v_1(t_k,\eta_i,\zeta_j) &= \frac{v_1^{min\{i+1,N_\eta\},j,k} - v_1^{min\{i,N_\eta-1\},j,k}}{\Delta\eta}, \\ \partial_\eta v_2(t_k,\eta_i,\zeta_j) &= \frac{v_2^{min\{i+1,N_\eta\},j,k} - v_2^{min\{i,N_\eta-1\},j,k}}{\Delta\eta}, \\ \partial_\zeta v_1(t_k,\eta_i,\zeta_j) &= \frac{v_1^{i,min\{j+1,N_\zeta\},k} - v_1^{i,min\{j,N_\zeta-1\},k}}{\Delta\zeta}, \\ \partial_\zeta v_2(t_k,\eta_i,\zeta_j) &= \frac{v_2^{i,min\{j+1,N_\zeta\},k} - v_2^{i,min\{j,N_\zeta-1\},k}}{\Delta\zeta}, \\ \partial_\eta^2 v_1(t_k,\eta_i,\zeta_j) &= \frac{v_1^{min\{i+1,N_\eta\},j,k} + v_1^{min\{i-1,N_\eta-2\},j,k} - 2v_1^{min\{i,N_\eta-1\},j,k}}{\Delta\eta^2}, \\ \partial_\eta^2 v_2(t_k,\eta_i,\zeta_j) &= \frac{v_2^{min\{i+1,N_\eta\},j,k} + v_2^{min\{i-1,N_\eta-2\},j,k} - 2v_2^{min\{i,N_\eta-1\},j,k}}{\Delta\eta^2}, \\ \partial_\zeta^2 v_1(t_k,\eta_i,\zeta_j) &= \frac{v_1^{i,min\{j+1,N_\zeta\},k} + v_1^{i,min\{j-1,N_\zeta-2\},k} - 2v_1^{i,min\{j,N_\zeta-1\},k}}{\Delta\zeta^2}, \\ \partial_\zeta^2 v_2(t_k,\eta_i,\zeta_j) &= \frac{v_2^{i,min\{j+1,N_\zeta\},k} + v_2^{i,min\{j-1,N_\zeta-2\},k} - 2v_2^{i,min\{j,N_\zeta-1\},k}}{\Delta\zeta^2}. \end{aligned}$$

5. It is not difficult to spot that some terms are time-independent or v-independent, and if we compute them at each t_k, then we simply repeat N_t+1 times' computation, which significantly slows down the computing speed. Therefore, we can compute those time- or v-independent components including, for instance, $\boldsymbol{\mu}^{i,j}$, $\boldsymbol{\sigma}^{i,j}$, a^i, Q^k, before we start solving the PDEs and this will not incur additional interpolation costs as all computations during solving PDEs are on the pre-defined grids. More specifically, we can compute $\boldsymbol{\mu}^{i,j}$ as an $N_\eta * N_\zeta * N * 1$ array, S^k as an $N_t * N * 1$ array, and a^i as an $N_\eta * 1$ array. Another advantage of doing so is that we can vectorise some of the computations. For example, after we solve v_2 and when

we are about to solve v_1, the coefficient of $v_1^{i,j,k}$ is independent of v_1 itself, and this leads us to vectorise the multiplication between the coeficient and $v_1^{i,j,k}$, which can further save the computing costs.

After solving the PDEs above numerically, we are able to get the optimal control $\boldsymbol{\pi}^*$ and the optimal total wealth path $X^{\boldsymbol{\pi}^*}$. For a single simulation, we need to firstly get a series of realisations of independent $(N+2)$ dimensional standard normal random variables:

$$\begin{pmatrix} Z_0^1 & Z_0^2 & \dots & Z_0^N & Z_0^\eta & Z_0^\zeta \\ Z_1^1 & Z_1^2 & \dots & Z_1^N & Z_1^\eta & Z_1^\zeta \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ Z_{N_t}^1 & Z_{N_t}^2 & \dots & Z_{N_t}^N & Z_{N_t}^\eta & Z_{N_t}^\zeta \end{pmatrix}, \tag{7.1.5}$$

to be used as the simulated Brownian motions. And then we use the last two columns to simulate a path of (η, ζ):

$$\begin{aligned} \eta_{k+1} &= \eta_k + \frac{b(\eta_k)}{\epsilon}\Delta t + \frac{a(\eta_k)}{\sqrt{\epsilon}}\sqrt{\Delta t} Z_k^\eta, \\ \zeta_{k+1} &= \zeta_k + \delta c(\zeta_k)\Delta t + \sqrt{\delta} l(\zeta_k)\sqrt{\Delta t} Z_k^\zeta. \end{aligned} \tag{7.1.6}$$

Note that the subscription represents at t_k and has nothing to do with the grid we defined previously. Before we move onto the next step, we need to ensure that the simulated path of both state variables should be bounded by their ranges calculated previously, as otherwise we are unable to do the extrapolation for the v's. If the values of these two state variables at any time spot fall outside their ranges, we need to redo the simulation until all values are bounded by their ranges. Then we simulate the optimal total wealth path $X^{\boldsymbol{\pi}^*}$ and calculate the optimal control $\boldsymbol{\pi}^*$ simultaneously as follows:

$$\begin{aligned} \boldsymbol{\pi}_k^* =& - \left[\boldsymbol{R}^k + 2v_2(t_k, \eta_k, \zeta_k)\boldsymbol{\sigma}(\eta_k, \zeta_k)\boldsymbol{\sigma}(\eta_k, \zeta_k)'\right]^{-1} \\ & \left\{X_k^{\boldsymbol{\pi}^*}\left\{\boldsymbol{S}^k + 2v_2(t_k, \eta_k, \zeta_k)[\boldsymbol{\mu}(\eta_k, \zeta_k) - r^k\mathbf{1}]\right\} + v_1(t_k, \eta_k, \zeta_k)[\boldsymbol{\mu}(\eta_k, \zeta_k) - r^k\mathbf{1}]\right\}, \\ X_{k+1}^{\boldsymbol{\pi}^*} =& X_k^{\boldsymbol{\pi}^*} + \left\{X_k^{\boldsymbol{\pi}^*} r^k + \boldsymbol{\pi}_k^{*\prime}[\boldsymbol{\mu}(\eta_k, \zeta_k) - r^k\mathbf{1}]\right\}\Delta t + \boldsymbol{\pi}_k^{*\prime}\boldsymbol{\sigma}(\eta_k, \zeta_k)\boldsymbol{Z}_k'\sqrt{\Delta t}, \end{aligned} \tag{7.1.7}$$

with the initial condition:

$$X_0^{\boldsymbol{\pi}^*} = x_0,$$

where:

$$\boldsymbol{Z}_k = (Z_k^1, \cdots, Z_k^N).$$

In (7.1.7), we need a interpolation scheme to compute values of functions of η and ζ evaluated at non-grid points. The scheme is a multidimensional linear scheme. For example, if $\eta_k \in [\eta_i, \eta_{i+1}]$ and $\zeta_k \in [\zeta_j, \zeta_{j+1}]$ (k represents the time while i and j represent the grid) and we are going to evaluate $v_2(t_k, \eta_k, \zeta_k)$ (noting that t_k is always on the grid point), it can be interpolated as follows:

$$\begin{aligned} v_2(t_k, \eta_k, \zeta_k) =& \frac{\eta_{i+1} - \eta_k}{\Delta\eta}\frac{\zeta_{j+1} - \zeta_k}{\Delta\zeta}v_2^{i,j,k} + \frac{\eta_{i+1} - \eta_k}{\Delta\eta}\frac{\zeta_k - \zeta_j}{\Delta\zeta}v_2^{i,j+1,k} \\ &+ \frac{\eta_k - \eta_i}{\Delta\eta}\frac{\zeta_{j+1} - \zeta_k}{\Delta\zeta}v_2^{i+1,j,k} + \frac{\eta_k - \eta_i}{\Delta\eta}\frac{\zeta_k - \zeta_j}{\Delta\zeta}v_2^{i+1,j+1,k}. \end{aligned} \tag{7.1.8}$$

7.1.2 Dual HJB Equation

For the dual HJB equation (4.1.15), adapting to the previous notation, we discretise the PDEs as follows:

$$
\begin{cases}
\tilde{v}_2^{i,j,k-1} = & \tilde{v}_2^{i,j,k} + \frac{b^i}{\epsilon}\frac{\tilde{v}_2^{i+1,j,k} - \tilde{v}_2^{i,j,k}}{\Delta\eta} + \frac{(a^i)^2}{2\epsilon}\frac{\tilde{v}_2^{i+1,j,k} + \tilde{v}_2^{i-1,j,k} - 2\tilde{v}_2^{i,j,k}}{\Delta\eta^2} \\
& + \delta c^j \frac{\tilde{v}_2^{i,j+1,k} - \tilde{v}_2^{i,j,k}}{\Delta\zeta} + \frac{1}{2}\delta(l^j)^2\frac{\tilde{v}_2^{i,j+1,k} + \tilde{v}_2^{i,j-1,k} - 2\tilde{v}_2^{i,j,k}}{\Delta\zeta^2} \\
& - 2(\tilde{v}_2^{i,j,k})^2 \left\{ \left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1}) - (\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \right\} \\
& \left\{ (\boldsymbol{R}^k)^{-1} + \tilde{v}_2^{i,j,k}\left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1} \right\}^{-1} \\
& \left\{ \left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1}) - (\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \right\} \\
& - 2(\tilde{v}_2^{i,j,k})^2 \left[Q^k - (\boldsymbol{S}^k)'(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \right] \\
& + \tilde{v}_2^{i,j,k}\left\{ (\boldsymbol{\mu}^{i,j} - r^k\mathbf{1})'\left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1}) - 2r^k \right\} \\
& - \frac{(a^i)^2}{\epsilon\tilde{v}_2^{i,j,k}}\left(\frac{\tilde{v}_2^{i+1,j,k} - \tilde{v}_2^{i,j,k}}{\Delta\eta}\right)^2 - \frac{(l^j)^2\delta}{\tilde{v}_2^{i,j,k}}\left(\frac{\tilde{v}_2^{i,j+1,k} - \tilde{v}_2^{i,j,k}}{\Delta\zeta}\right)^2, \\
\tilde{v}_1^{i,j,k-1} = & \tilde{v}_1^{i,j,k} + \frac{b^i}{\epsilon}\frac{\tilde{v}_1^{i+1,j,k} - \tilde{v}_1^{i,j,k}}{\Delta\eta} + \frac{(a^i)^2}{2\epsilon}\frac{\tilde{v}_1^{i+1,j,k} + \tilde{v}_1^{i-1,j,k} - 2\tilde{v}_1^{i,j,k}}{\Delta\eta^2} \\
& + \delta c^j \frac{\tilde{v}_1^{i,j+1,k} - \tilde{v}_1^{i,j,k}}{\Delta\zeta} + \frac{1}{2}\delta(l^j)^2\frac{\tilde{v}_1^{i,j+1,k} + \tilde{v}_1^{i,j-1,k} - 2\tilde{v}_1^{i,j,k}}{\Delta\zeta^2} \\
& + 2\tilde{v}_1^{i,j,k}\tilde{v}_2^{i,j,k}\Bigg\{ \left\{ \left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1}(\boldsymbol{\mu}^{i,j} - r^k\mathbf{1}) - (\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \right\}' \\
& \left\{ (\boldsymbol{R}^k)^{-1} + \tilde{v}_2^{i,j,k}\left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1} \right\}^{-1}(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \\
& - Q^k + (\boldsymbol{S}^k)'(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \Bigg\} \\
& - r^k\tilde{v}_1^{i,j,k} - \frac{(a^i)^2}{\epsilon\tilde{v}_2^{i,j,k}}\frac{(\tilde{v}_1^{i+1,j,k} - \tilde{v}_1^{i,j,k})(\tilde{v}_2^{i+1,j,k} - \tilde{v}_2^{i,j,k})}{\Delta\eta^2} \\
& - \frac{(l^j)^2\delta}{\tilde{v}_2^{i,j,k}}\frac{(\tilde{v}_1^{i,j+1,k} - \tilde{v}_1^{i,j,k})(\tilde{v}_2^{i,j+1,k} - \tilde{v}_2^{i,j,k})}{\Delta\zeta^2}, \\
\tilde{v}_0^{i,j,k-1} = & \tilde{v}_0^{i,j,k} + \frac{b^i}{\epsilon}\frac{\tilde{v}_0^{i+1,j,k} - \tilde{v}_0^{i,j,k}}{\Delta\eta} + \frac{(a^i)^2}{2\epsilon}\frac{\tilde{v}_0^{i+1,j,k} + \tilde{v}_0^{i-1,j,k} - 2\tilde{v}_0^{i,j,k}}{\Delta\eta^2} \\
& + \delta c^j \frac{\tilde{v}_0^{i,j+1,k} - \tilde{v}_0^{i,j,k}}{\Delta\zeta} + \frac{1}{2}\delta(l^j)^2\frac{\tilde{v}_0^{i,j+1,k} + \tilde{v}_0^{i,j-1,k} - 2\tilde{v}_0^{i,j,k}}{\Delta\zeta^2} \\
& - \frac{1}{2}(\tilde{v}_1^{i,j,k})^2\Bigg\{ (\boldsymbol{S}^k)'(\boldsymbol{R}^k)^{-1} \\
& \left\{ (\boldsymbol{R}^k)^{-1} + \tilde{v}_2^{i,j,k}\left[\boldsymbol{\sigma}^{i,j}(\boldsymbol{\sigma}^{i,j})'\right]^{-1} \right\}^{-1}(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \\
& + Q^k - (\boldsymbol{S}^k)'(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k \Bigg\} \\
& - \frac{(a^i)^2}{4\epsilon\tilde{v}_2^{i,j,k}}\left(\frac{\tilde{v}_1^{i+1,j,k} - \tilde{v}_1^{i,j,k}}{\Delta\eta^2}\right)^2 - \frac{(l^j)^2\delta}{4\tilde{v}_2^{i,j,k}}\left(\frac{\tilde{v}_1^{i,j+1,k} - \tilde{v}_1^{i,j,k}}{\Delta\zeta^2}\right)^2,
\end{cases}
\tag{7.1.9}
$$

with the terminal conditions:

$$\begin{cases} \tilde{v}_0(t_{N_t}, \eta_i, \zeta_j) = \dfrac{B^2}{2A}, \\ \tilde{v}_1(t_{N_t}, \eta_i, \zeta_j) = \dfrac{B}{A}, \\ \tilde{v}_2(t_{N_t}, \eta_i, \zeta_j) = \dfrac{1}{2A}, \end{cases}$$
$$\forall\ i = 0, \cdots, N_\eta,\ j = 0, \cdots, N_\zeta,$$

where:

$$\begin{aligned} \tilde{v}_2^{i,j,k} &= \tilde{v}_2(t_k, \eta_i, \zeta_j), \\ \tilde{v}_1^{i,j,k} &= \tilde{v}_1(t_k, \eta_i, \zeta_j), \\ \tilde{v}_0^{i,j,k} &= \tilde{v}_0(t_k, \eta_i, \zeta_j). \end{aligned}$$

Analogous to the previous analysis, we have several points to mention:

1. As we only need to care about the optimal controls y_0^*, α^*, $\boldsymbol{\beta}^*$, γ^*, and η^*, and the optimal dual state variable path Y^*, it suffices to only solve $\tilde{v}_1$ and $\tilde{v}_2$. Thus, we will only discuss solving $\tilde{v}_1$ and $\tilde{v}_2$ from now on.

2. We will use the same grids as those of the primal problem and it will automatically satisfy the stability condition.

3. To make the results from the dual problem and the primal problem comparable, we simply use the same simulated η and ζ obtained during solving the primal problem.

4. The way to fulfill $\tilde{v}_1^{i,j,k}$ and $\tilde{v}_2^{i,j,k}$ is the same as the previous one: based on the later time spot (i.e. $\tilde{v}_1^{i,j,k+1}$, $\tilde{v}_1^{i+1,j,k+1}$, $\tilde{v}_1^{i-1,j,k+1}$, $\tilde{v}_1^{i,j+1,k+1}$, $\tilde{v}_1^{i,j-1,k+1}$, $\tilde{v}_2^{i,j,k+1}$, $\tilde{v}_2^{i+1,j,k+1}$, $\tilde{v}_2^{i-1,j,k+1}$, $\tilde{v}_2^{i,j+1,k+1}$, $\tilde{v}_2^{i,j-1,k+1}$). Also, we will have to deal with evaluations at boundaries and we simply use the same approximation technique aforementioned.

5. Likewise, we can speed up the computations by pre-computing some time-independent or $\tilde{v}$-independent components, which are basically the same as those of the primal problem. Meanwhile, we can also vectorise our computations in solving $\tilde{v}_1$ so that we can further save the computing costs.

After solving the PDEs above numerically, we are able to get the optimal controls y_0^*, α^*, $\boldsymbol{\beta}^*$, γ^*, and ξ^*, and the optimal dual state variable path Y^*. For a single simulation, we can use the same realisations of independent $(N+2)$ dimensional standard normal random variables in (7.1.5) to construct the Brownian motions. Furthermore, we can also use the same single simulations of η and ζ as in solving the primal problem so that we are making both dual and primal problems comparable and circumventing the out-of-range issue of η and ζ. Then we simulate the optimal state variable path Y^* and calculate the optimal controls y_0^*, α^*, $\boldsymbol{\beta}^*$, γ^*, and ξ^* simultaneously as follows:

$$
\begin{aligned}
\boldsymbol{\beta}_k^* =& \left\{(\boldsymbol{R}^k)^{-1} + 2\tilde{v}_2(t_k,\eta_k,\zeta_k)\left[\boldsymbol{\sigma}(\eta_k,\zeta_k)\boldsymbol{\sigma}(\eta_k,\zeta_k)'\right]^{-1}\right\}^{-1} \\
& \left\{2\tilde{v}_2(t_k,\eta_k,\zeta_k)Y_k^*\left\{\left[\boldsymbol{\sigma}(\eta_k,\zeta_k)\boldsymbol{\sigma}(\eta_k,\zeta_k)'\right]^{-1}\left[\boldsymbol{\mu}(\eta_k,\zeta_k) - r^k\mathbf{1}\right] - (\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k\right\}\right. \\
& \left. - \tilde{v}_1(t_k,\eta_k,\zeta_k)(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k\right\}, \\
\alpha_k^* =& (\boldsymbol{S}^k)'(\boldsymbol{R}^k)^{-1}\left\{(\boldsymbol{R}^k)^{-1} + 2\tilde{v}_2(t_k,\eta_k,\zeta_k)\left[\boldsymbol{\sigma}(\eta_k,\zeta_k)\boldsymbol{\sigma}(\eta_k,\zeta_k)'\right]^{-1}\right\}^{-1} \\
& \left\{2\tilde{v}_2(t_k,\eta_k,\zeta_k)Y_k^*\left\{\left[\boldsymbol{\sigma}(\eta_k,\zeta_k)\boldsymbol{\sigma}(\eta_k,\zeta_k)'\right]^{-1}\left[\boldsymbol{\mu}(\eta_k,\zeta_k) - r^k\mathbf{1}\right] - (\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k\right\}\right. \\
& \left. - \tilde{v}_1(t_k,\eta_k,\zeta_k)(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k\right\} \\
& - \left[\tilde{v}_1(t_k,\eta_k,\zeta_k) + \tilde{v}_2(t_k,\eta_k,\zeta_k)Y_k^*\right]\left(Q^k - (\boldsymbol{S}^k)'(\boldsymbol{R}^k)^{-1}\boldsymbol{S}^k\right), \\
\gamma_k^* =& -\frac{a(\eta_k)}{\sqrt{\epsilon}}\frac{\partial_\eta\tilde{v}_1(t_k,\eta_k,\zeta_k) + 2\partial_\eta\tilde{v}_2(t_k,\eta_k,\zeta_k)Y_k^*}{2\tilde{v}_2(t_k,\eta_k,\zeta_k)}, \\
\xi_k^* =& -l(\zeta_k)\sqrt{\delta}\frac{\partial_\zeta\tilde{v}_1(t_k,\eta_k,\zeta_k) + 2\partial_\zeta\tilde{v}_2(t_k,\eta_k,\zeta_k)Y_k^*}{2\tilde{v}_2(t_k,\eta_k,\zeta_k)}, \\
Y_{k+1}^* =& Y_k^* + (\alpha_k^* - r^kY_k^*)\Delta t + \left\{\boldsymbol{\beta}_k^* - Y_k^*\left[\boldsymbol{\mu}(\eta_k,\zeta_k) - r^k\mathbf{1}\right]\right\}'\boldsymbol{\sigma}(\eta_k,\zeta_k)'^{-1}\boldsymbol{Z}_k'\sqrt{\Delta t} \\
& + \gamma_k^*Z_k^\eta\sqrt{\Delta t} + \xi_k^*Z_k^\zeta\sqrt{\Delta t}.
\end{aligned} \tag{7.1.10}
$$

with the initial condition:

$$
y_0^* = -\frac{x_0 + \tilde{v}_1(0,\eta_0,\zeta_0)}{2\tilde{v}_2(0,\eta_0,\zeta_0)}. \tag{7.1.11}
$$

Analogously, we still need the same interpolation scheme for $\tilde{v}_1$ and $\tilde{v}_2$. In addition, we will also need to do the interpolations for $\partial_\eta\tilde{v}_1$, $\partial_\eta\tilde{v}_2$, $\partial_\zeta\tilde{v}_1$, and $\partial_\zeta\tilde{v}_2$ to compute γ_k^* and ξ_k^*. Technically, we firstly need to compute $N_t * N_\eta * N_\zeta$ arrays for each partial derivative. Let us take $\partial_\eta\tilde{v}_1$ as an example:

$$
\partial_\eta\tilde{v}_1^{i,j,k} = \frac{\tilde{v}_1^{i+1,j,k} - \tilde{v}_1^{i,j,k}}{\Delta\eta}. \tag{7.1.12}
$$

where we need to take care of the out-of-range issue of η and ζ by using the 'a-step-backward' approximation. Then finally, we are able to apply the same interpolation scheme to compute γ_k^* and ξ_k^*.

After solving the dual problem numerically, we are able to recover the optimal total wealth path $X^{\boldsymbol{\pi}^*}$ using Y^*, $\tilde{v}_1$ and $\tilde{v}_2$ by (4.2.2).

7.2 Solving the Primal and Dual FBSDEs Numerically

This section is inspired by Zhu & Zheng (2022, Section 4). It is not difficult to spot that we are given the terminal conditions of SDEs, for example, (6.2.3). If it is an ODE, then we can start from the terminal time to progress backwardly without any severe issue. However, if it is an SDE, then solve it backwardly given the terminal condition will break the nonanticipativeness of the solution. Thus, we need to come up with a scheme to circumvent this issue and the solution is to transform the problem of solving the FBSDE into a problem of solving an optimisation problem.

7.2.1 Primal FBSDE Method

For the primal FBSDE method, we can solve it by transforming it into a minimisation problem. To solve $\hat{p}_1$ from (6.1.6) backwardly is hard, as we need to know $X^{\boldsymbol{\pi}^*}$ firstly. However, we can treat the terminal condition as follows:

$$\begin{aligned} &\hat{p}_1(T) + AX^{\boldsymbol{\pi}^*}(T) + B = 0 \\ \iff \quad &\mathbb{E}\left[(\hat{p}_1(T) + AX^{\boldsymbol{\pi}^*}(T) + B)^2\right] = 0. \end{aligned}$$

Thus we can consider this to be an error minimisation problem as follows:

$$\min_{\hat{p}_1(0), \hat{\boldsymbol{q}}_1} \mathbb{E}\left[(\hat{p}_1(T) + AX^{\boldsymbol{\pi}^*}(T) + B)^2\right]. \tag{7.2.1}$$

Therefore, we can discretise the time horizon into m intervals and then make the optimal controls $\boldsymbol{\pi}^*$ and $\hat{\boldsymbol{q}}_1$ piecewise constant. More specifically, we are going to solve the following optimisation problem:

$$\min_{\hat{p}_{1,0}, \hat{\boldsymbol{q}}_{1,0}, \cdots, \hat{\boldsymbol{q}}_{1,m-1}} \mathbb{E}\left[\left(\hat{p}_1(t_m) + AX^{\boldsymbol{\pi}^*}(t_m) + B\right)^2\right],$$

subject to:

$$\begin{cases} X^{\boldsymbol{\pi}^*}(t_{i+1}) = X^{\boldsymbol{\pi}^*}(t_i) + [X^{\boldsymbol{\pi}^*}(t_i)r + \boldsymbol{\pi}_i^{*\prime}(\boldsymbol{\mu} - r\mathbf{1})]\Delta t + \boldsymbol{\pi}_i^{*\prime}\boldsymbol{\sigma}\boldsymbol{M}_i\sqrt{\Delta t}, \\ X^{\boldsymbol{\pi}^*}(0) = x_0, \\ \hat{p}_1(t_{i+1}) = \hat{p}_1(t_i) + [-r\hat{p}_1(t_i) + QX^{\boldsymbol{\pi}^*}(t_i) + \boldsymbol{S}'\boldsymbol{\pi}_i^*]\Delta t + \hat{\boldsymbol{q}}_{1,i}'[\boldsymbol{M}_i, M_i^{\eta}, M_i^{\zeta}]'\sqrt{\Delta t}, \\ \hat{p}_1(0) = \hat{p}_{1,0}, \end{cases}$$

where $t_i = (iT)/m$, $\Delta t = T/m$, $\boldsymbol{M}_i$ is the i^{th} i.i.d realisation of the N-dimensional standard normal random variable for $X^{\boldsymbol{\pi}^*}$, M_i^{η} is the i^{th} i.i.d realisation of the standard normal random variable for η, and M_i^{ζ} is the i^{th} i.i.d realisation of the standard normal random variable for ζ, and:

$$\hat{p}_1(t_i)(\boldsymbol{\mu} - r\mathbf{1}) + \boldsymbol{\sigma}[\boldsymbol{I}_{N\times N},\ \mathbf{0}_{N\times 2}]\hat{\boldsymbol{q}}_{1,i} - X^{\boldsymbol{\pi}^*}(t_i)\boldsymbol{S} - \boldsymbol{R}\boldsymbol{\pi}_i^* = 0.$$

if we take $K = \mathbb{R}^N$.

To ensure that numerical results of the primal FBSDEs are comparable with the primal HJB equation, we need to make $m = N_t$, $\boldsymbol{M}_k = \boldsymbol{Z}_k$, $M_k^{\eta} = Z_k^{\eta}$, and $M_k^{\zeta} = Z_k^{\zeta}$, and then we can extract the simulated path of $X^{\boldsymbol{\pi}^*}$ from the FBSDEs. However, we should notice that the extraction may not be precisely matched with the results produced by the primal HJB equation or recovered by the dual HJB equation. Although theoretically we can treat $\hat{\boldsymbol{q}}_1$ to be optimised as an $N_t * N_\eta * N_\zeta * (N+2) * 1$ array, in practice, $N_t * (N+2) * 1$ is generally a sufficiently large number which is already computationally demanding for the optimisation algorithm. Therefore, we at most can approximate it as an $N_t * N * 1$ array, indicating that $\hat{\boldsymbol{q}}_1$ to be optimised is approximated to be independent of η and ζ and is the source of the errors.

7.2.2 Dual FBSDE Method

For the dual FBSDE method, the logic is roughly the same. We firstly transform the problem into the following optimisation:

$$\min_{y_0, \gamma, \xi, \hat{\boldsymbol{q}}_2} \mathbb{E}\left[\left(\hat{p}_2(T) + \frac{Y^*(T) + B}{A}\right)^2\right], \tag{7.2.2}$$

Then we discretise the time horizon into m intervals and make the optimal controls α^*, $\boldsymbol{\beta}^*$, γ^*, ξ^*, and $\hat{\boldsymbol{q}}_2$ piecewise constant. More specifically, we are going to solve the following optimisation problem:

$$\min_{y_0, \gamma_0, \cdots, \gamma_{m-1}, \xi_0, \cdots, \xi_{m-1}, \hat{\boldsymbol{q}}_{2,0}, \cdots, \hat{\boldsymbol{q}}_{2,m-1}} \mathbb{E}\left[\left(\hat{p}_2(t_m) + \frac{Y^*(t_m) + B}{A}\right)^2\right],$$

subject to:

$$\begin{cases} Y^*(t_{i+1}) = & Y^*(t_i) + (\alpha_i^* - rY^*(t_i))\Delta t \\ & + \{[\boldsymbol{\beta}_i^* - Y^*(t_i)(\boldsymbol{\mu} - r\mathbf{1})]'\boldsymbol{\sigma}'^{-1}\boldsymbol{M}_i + \gamma M_i^{\eta} + \xi M_i^{\zeta}\}\sqrt{\Delta t}, \\ Y^*(0) = & y_0, \\ \hat{p}_2(t_{i+1}) = & \hat{p}_2(t_i) + [r\hat{p}_2(t_i) + \hat{\boldsymbol{q}}_{2,i}'\boldsymbol{\sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})]\Delta t + \hat{\boldsymbol{q}}_{2,i}'\boldsymbol{M}_i\sqrt{\Delta t}, \\ \hat{p}_2(0) = & x_0, \end{cases}$$

where the time interval length and the simulated Brownian motions are the same as those of the primal FBSDEs, and:

$$\begin{cases} \hat{p}_2(t_i) = \dfrac{\alpha_i^* - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}_i^*}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}}, \\ \boldsymbol{\sigma}'\hat{\boldsymbol{q}}_{2,i} = \boldsymbol{R}^{-1}\boldsymbol{\beta}_i^* - \dfrac{\boldsymbol{R}^{-1}\boldsymbol{S}(\alpha_i^* - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{\beta}_i^*)}{Q - \boldsymbol{S}'\boldsymbol{R}^{-1}\boldsymbol{S}}, \end{cases}$$

if we take $K = \mathbb{R}^N$.

Likewise, we are unable to fully parameterise the controls γ, ξ, and $\hat{\boldsymbol{q}}_2$ due to the unrealistic high dimensionalities. Instead, we approximate them as only time-dependent: γ and ξ are approximated as two $N_t * 1$ arrays and $\hat{\boldsymbol{q}}_2$ is approximated as an $N_t * N * 1$ array, which contribute to the matching errors.

7.3 Numerical Results of Solutions of PDEs and Simulations for FBSDEs

In this section, we are going to present the results produced by the previously mentioned numerical methods. We will specify our settings (one with a smaller scale and another with a larger scale) and will mainly focus on comparing results produced by different methods and testing the stability of the algorithm.

7.3.1 Smaller Scaled Setting

For the smaller scaled setting, we simplify the only-time-dependent components by making them constant throughout the time since in practice we may not be always able to precisely describe the time structure of them and this setting will not impair the generality:

$$\begin{aligned} Q(t) &= 0, \\ \boldsymbol{S}(t) &= \boldsymbol{1}, \\ \boldsymbol{R}(t) &= \boldsymbol{I}, \\ r(t) &= 0.001. \end{aligned} \tag{7.3.1}$$

For the constants in our model, we specify them as follows:

$$\begin{aligned} N &= 2, \\ T &= 1, \\ A &= 2, \\ B &= -22, \\ \epsilon &= 0.01, \\ \delta &= 0.01, \\ x_0 &= 10, \\ \eta_0 &= 0.01, \\ \zeta_0 &= 0.01. \end{aligned} \tag{7.3.2}$$

The rationale to choose $B = -22$ is that we want to centre the terminal total wealth around 11, a number slightly higher than $x_0 \mathrm{e}^{rT}$, as $-B/A$ represents the mean of $X_T^{\boldsymbol{\pi}^*}$.

We also want η and ζ to be able to reach negative numbers. In order to avoid over complications, we simply parameterise their dynamics as follows:

$$\begin{aligned} a(\eta) &= 0.1, \\ b(\eta) &= 0.1, \\ c(\zeta) &= 0.1, \\ l(\zeta) &= 0.1. \end{aligned} \tag{7.3.3}$$

For the specification of the dynamics of the stock prices, we use the following parameterisations:

$$
\begin{aligned}
\boldsymbol{\mu}(\eta,\zeta) &= \begin{pmatrix} \frac{1}{2}(\zeta-\eta) \\ \frac{1}{5}\eta + \frac{2}{5}\zeta \end{pmatrix}, \\
\boldsymbol{\sigma}(\eta,\zeta) &= \begin{pmatrix} \frac{1}{2}\eta + \frac{3}{5} & 0 \\ \frac{1}{20}\zeta + \frac{1}{10} & \frac{1}{2}\eta + \frac{3}{5} \end{pmatrix}.
\end{aligned} \tag{7.3.4}
$$

Finally, we specify the grid parameters as follows:

$$
\begin{aligned}
N_t &= 625, \\
N_\eta &= 25, \\
N_\zeta &= 25.
\end{aligned} \tag{7.3.5}
$$

Firstly, after we solve both primal and dual HJB equations, we have the grids of v_1, v_2, $\tilde{v}_1$, and $\tilde{v}_2$ in hand, and we are able to verify the correctness by using (4.2.1) and plotting the errors (but we are unable to visualise the values since they are all three dimensional arrays). Here are the histograms of the errors:

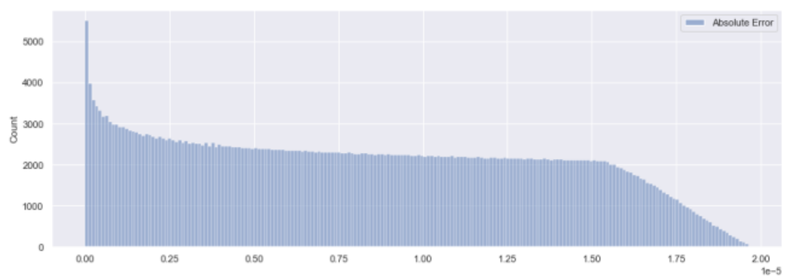

Figure 7.1: Histogram of absolute errors produced by $|\tilde{v}_1 - \frac{v_1}{2v_2}|$.

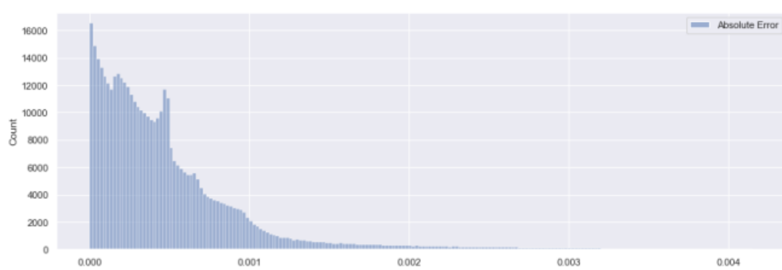

Figure 7.2: Histogram of absolute errors produced by $|\tilde{v}_2 - \frac{1}{4v_2}|$.

where $|\tilde{v}_1|$ is on 10^1 magnitude and $|\tilde{v}_2|$ is on 10^{-1} magnitude. Although the second errors seem to be significantly larger than the first, they are still acceptable as relatively they are on 1% magnitude.

Next, we can simulate $X^{\boldsymbol{\pi}^*}$ and Y^* using the solved components, and recover $X^{\boldsymbol{\pi}^*}$ form Y^* to check the equivalence. Here are the single simulation results:

Figure 7.3: Single simulation result of $X^{\boldsymbol{\pi}^*}$.

Figure 7.4: Single simulation result of Y^*.

Through (4.2.2), we are able to recover $X^{\boldsymbol{\pi}^*}$ from Y^*. We demonstrate the recovered path and the errors as follows:

Figure 7.5: $X^{\boldsymbol{\pi}^*}$ vs. recovered $X^{\boldsymbol{\pi}^*}$ and their errors produced by $X^{\boldsymbol{\pi}^*} - (-\tilde{v}_1 - 2\tilde{v}_2 Y^*)$.

Again, the errors are relatively small as they are on the relative 0.1% magnitude.

In addition to the single simulation, we also conduct the batch simulations (i.e. simulate the paths 100 times) for both $X^{\boldsymbol{\pi}^*}$ and Y^*. To have a better idea about the distribution of $X^{\boldsymbol{\pi}^*}$ across the timeline, we plot the mean, 25% percentiles and 75% percentiles in one figure as follows:

Figure 7.6: Batch simulation results of $X^{\boldsymbol{\pi}^*}$.

We also compute the recovered $X^{\boldsymbol{\pi}^*}$ from Y^* in each simulation and their absolute errors across the timeline. We average them and plot them against the timeline as follows:

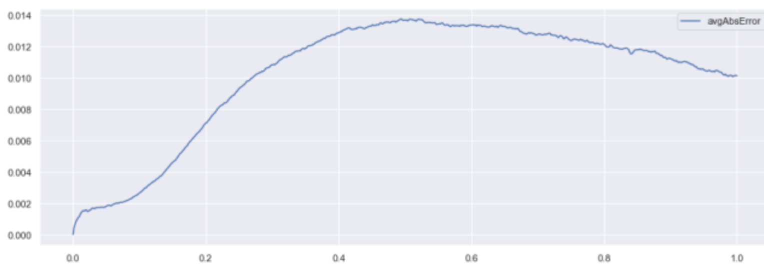

Figure 7.7: Averaged absolute errors produced by $|X^{\boldsymbol{\pi}^*} - (-\tilde{v}_1 - 2\tilde{v}_2 Y^*)|$.

The averaged errors are on the relative 0.1% magnitude, which is a strong evidence corroborating the equivalence of the primal and dual HJB equations.

As for the verification of the FBSDE method, unfortunately, we are unable to finish the optimisation due to the exceptionally high dimension. We can, however, produce an initial guess for both optimisations (i.e. (7.2.1) and (7.2.2)) by "cheating" via Theorem 6.3.1 and Theorem 6.3.2 and averaging the primal and dual optimal controls that we calculated during the batch simulations of $X^{\boldsymbol{\pi}^*}$ and Y^* at each time spot. We then evaluate the equivalence by comparing both optimisation objectives (index i represents the i^{th} simulation):

$$\frac{1}{100}\sum_{i=1}^{100}\Big[\hat{p}_1^i(t_{N_t}) + A(X^{\boldsymbol{\pi}^*})^i(t_{N_t}) + B\Big]^2,$$

and:

$$\frac{1}{100}\sum_{i=1}^{100}\left[\hat{p}_2^i(t_{N_t}) + \frac{(Y^*)^i(t_{N_t}) + B}{A}\right]^2,$$

with their corresponding initial guesses produced by "cheating" and other arbitrary initial guesses (i.e. by initialising all controls at all time spots across all dimensions by 1). Here are the results:

	Primal Obj	Dual Obj
"Cheating"	2668.75	0.93
Arbitrary	7017.89	6.17

Table 7.1: Values of optimisation objectives

where the smaller figure is more appreciated. It is clear that the figure achieved by "cheating" is smaller than the arbitrary initialisation, although "cheating" does not perform critically well in the primal objective evaluation. Overall it is a good evidence to support the equivalence.

7.3.2 Larger Scaled Setting

To test the stability of the algorithm, we redo the procedure on a larger scale. The grid parameters become:

$$\begin{aligned} N_t &= 2500, \\ N_\eta &= 50, \\ N_\zeta &= 50. \end{aligned} \tag{7.3.6}$$

And we also increase the number of stocks to be considered to $N = 5$ with the updated dynamics specifications:

$$\begin{aligned} \boldsymbol{\mu}(\eta,\zeta) &= \begin{pmatrix} \frac{1}{2}(\zeta-\eta) \\ \frac{1}{5}\eta + \frac{2}{5}\zeta \\ \frac{3}{5}\eta - \frac{3}{10}\zeta \\ \frac{3}{50}\eta - \frac{3}{10}\zeta \\ \frac{16}{25}\zeta - \frac{4}{25}\eta \end{pmatrix}, \\ \boldsymbol{\sigma}(\eta,\zeta) &= \begin{pmatrix} \frac{1}{2}\eta + \frac{3}{5} & 0 & 0 & 0 & 0 \\ \frac{1}{20}\zeta + \frac{1}{10} & \frac{1}{2}\eta + \frac{3}{5} & 0 & 0 & 0 \\ \frac{1}{25}\eta + \frac{1}{10} & \frac{1}{40}\zeta + \frac{1}{10} & \frac{1}{2}\eta + 14\zeta + \frac{3}{5} & 0 & 0 \\ \frac{1}{25}\zeta + \frac{1}{10} & \frac{1}{40}\eta + \frac{1}{10} & \frac{1}{40}\eta + \frac{1}{25}\zeta + \frac{1}{10} & \frac{1}{4}\eta + \frac{1}{4}\zeta + \frac{3}{5} & 0 \\ \frac{1}{40}\eta + \frac{1}{10} & \frac{1}{25}\zeta + \frac{1}{10} & \frac{1}{20}\eta + \frac{1}{40}\zeta + \frac{1}{10} & \frac{1}{20}\eta + \frac{3}{40}\zeta + \frac{1}{10} & \frac{1}{2}\eta + \frac{2}{5}\zeta + \frac{3}{5} \end{pmatrix}. \end{aligned} \tag{7.3.7}$$

Following the same logic as the previous section, we simply plot all figures as follows:

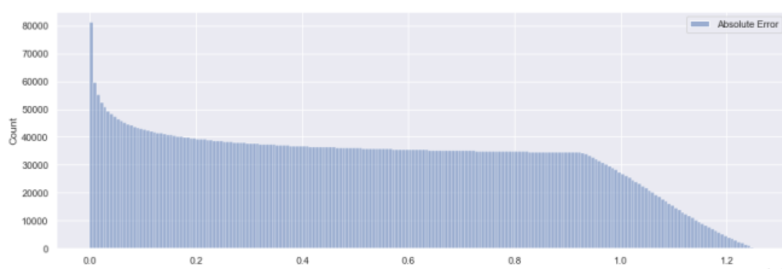

Figure 7.8: Histogram of absolute errors produced by $|\tilde{v}_1 - \frac{v_1}{2v_2}|$.

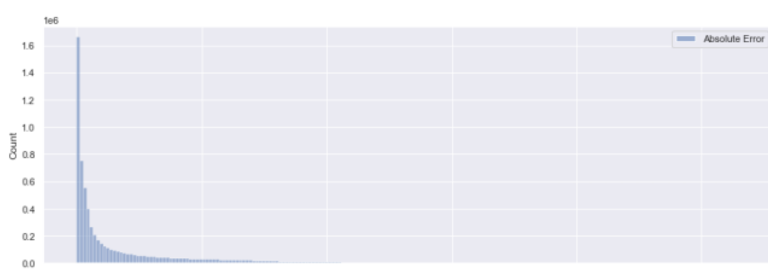

Figure 7.9: Histogram of absolute errors produced by $|\tilde{v}_2 - \frac{1}{4v_2}|$.

Figure 7.10: Single simulation result of X^{π^*}.

Figure 7.11: Single simulation result of Y^*.

Figure 7.12: X^{π^*} vs. recovered X^{π^*} and their errors produced by $X^{\pi^*} - (-\tilde{v}_1 - 2\tilde{v}_2 Y^*)$.

Figure 7.13: Batch simulation results of $X^{\boldsymbol{\pi}^*}$.

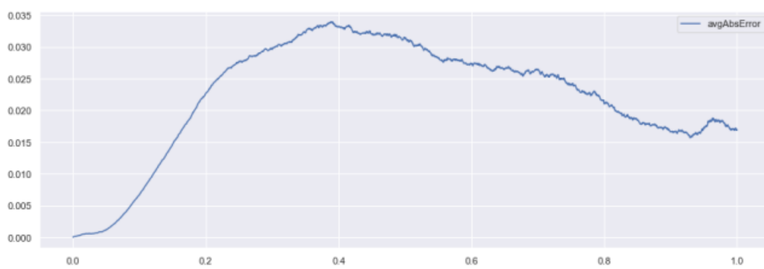

Figure 7.14: Averaged absolute errors produced by $|X^{\boldsymbol{\pi}^*} - (-\tilde{v}_1 - 2\tilde{v}_2 Y^*)|$.

and also the performance table similar to Table 7.3.1 as follows:

	Primal Obj	Dual Obj
"Cheating"	2165.68	2.19
Arbitrary	54052.11	23.88

Table 7.2: Values of optimisation objectives

From the above results, we know that the distribution of errors produced by $|\tilde{v}_1 - \frac{v_1}{2v_2}|$ becomes more fat-tailed, but still relatively acceptable. Also the averaged absolute errors $|X^{\boldsymbol{\pi}^*} - (-\tilde{v}_1 - 2\tilde{v}_2 Y^*)|$ are relatively acceptable, although they are almost doubled than that of the smaller scaled setting. In the performance table, we know that the figure achieved by "cheating" in the primal objective evaluation is improved relatively significantly, although that in the dual objective evaluation is slightly worsen off. In summary, the results of the larger scaled setting support the stability of our algorithm.

Chapter 8

Conclusion

In this paper, we discuss a continuous-time dynamic portfolio optimisation problem under multi-scale stochastic structures with the optimisation objective of minimising the quadratic risk function in the total wealth process and the portfolio strategy. We introduce the two-factor stochastic structure for the stock price dynamics and allow the number of stock to be an arbitrary number. We then applied four methods to solve the problem including the primal HJB eqaution, the dual HJB equation, the primal FBSDE method, and the dual FBSDE method and focus on their theoretical and numerical equivalences. The numerical results strongly support the equivalence between the primal and dual HJB equations. Nevertheless, subject to the speed of computing and the high dimensional nature, we are unable to solve the optimisations arising from the numerical method of solving the primal and dual FBSDEs. We instead check the effectiveness of FBSDE methods by evaluating the terminal variances using the initial guesses produced by "cheating" with Theorem 6.3.1 and Theorem 6.3.2, which also supports the equivalence among the four methods.

In addition to the general case, we briefly discuss the simplified case by assuming $f = 0$, the correlated case under the simplified setting, and the convex cone constraint without the correlation structure, and verify the compatibility between the correlated case and the uncorrelated case.

There are still a lot of possible extensions to the results discussed in this thesis. The most important thing on the to-do list is to numerically solve the optimisation problems arising from the FBSDE methods so that we are able to verify their equivalences more directly. On the other hand, if we have sufficiently strong computing powers, then we may try to refine the optimisation problems by fully parameterising controls with the dependence structure of η and ζ. Lastly, we can do more numerical experiments to discover the impact of the parameters in this model to the optimal total wealth process.

Bibliography

Fouque, J.-P. & Hu, R. (2020), 'Multiscale asymptotic analysis for portfolio optimization under stochastic environment', *Multiscale Modeling Simulation* **18**(3), 1318–1342.

Hu, Y. & Zhou, X. Y. (2006), 'Constrained stochastic LQ control with random coefficients, and application to portfolio selection', *SIAM journal on control and optimization* **44**(2), 444–466.

LeVeque, R. J. (1998), 'Finite difference methods for differential equations', *Draft version for use in AMath* **585**(6), 112.

Li, Y. & Zheng, H. (2018), 'Constrained quadratic risk minimization via forward and backward stochastic differential equations', *SIAM Journal on Control and Optimization* **56**(2), 1130–1153.

Ma, J., Li, W. & Zheng, H. (2020), 'Dual control Monte-Carlo method for tight bounds of value function under Heston stochastic volatility model', *European Journal of Operational Research* **280**(2), 428–440.

Rockafellar, R. T. (1970), *Convex Analysis*, Vol. 18, Princeton university press.

Yong, J. & Zhou, X. Y. (1999), *Stochastic controls: Hamiltonian systems and HJB equations*, Vol. 43, Springer Science Business Media.

Zhu, D. & Zheng, H. (2022), 'Effective approximation methods for constrained utility maximization with drift uncertainty', *Journal of optimization theory and applications* **194**(1), 191–219.