Optimal Measure Transformations and Optimal Trading

Renjie Wang

A Thesis
in the Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy (Mathematics) at
Concordia University
Montréal, Québec, Canada

August 2017

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Renjie Wang

Entitled: Optimal Measure Transformations and Optimal Trading

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____________________________________ Chair
Dr. Pascale Biron

____________________________________ External Examiner
Dr. Genevieve Gauthier

____________________________________ External to Program
Dr. Lorne Switzer

____________________________________ Examiner
Dr. Xiaowen Zhou

____________________________________ Examiner
Dr. Lea Popovic

____________________________________ Thesis Co-Supervisor
Dr. Cody Hyndman

____________________________________ Thesis Co-Supervisor
Dr. Wei Sun

Approved by ____________________________________
Dr. Arusharka Sen, Graduate Program Director

Monday, August 28, 2017 ____________________________________
Dr. André Roy, Dean
Faculty of Arts and Science

ABSTRACT

Optimal Measure Transformations and Optimal Trading

Renjie Wang, Ph.D.

Concordia University, 2017

We first associate the bond price with an optimal measure transformation problem which is closely related to decoupled nonlinear forward-backward stochastic differential equation (FBSDE).[1] The measure which solves the optimal measure transformation problem is the forward measure. These connections explain why the forward measure transformation employed in the FBSDE approach of Hyndman (Math. Financ. Econ. 2(2):107-128, 2009) is effective. We obtain explicit solutions to FBSDEs with jumps in affine term structure models and quadratic term structure models, which extend Hyndman (Math. Financ. Econ. 2(2):107-128, 2009). From the optimal measure transformation problem for defaultable bonds, we derive FBSDEs with random terminal condition to which we give a partially explicit solution. In the second part we consider trading against a hedge fund or large trader that must liquidate a large position in a risky asset if the market price of the asset crosses a certain threshold.[2] Liquidation occurs in a disorderly manner and negatively impacts the market price of the asset. We consider the perspective of small investors whose trades do not induce market impact and who possess different levels of information about the liquidation trigger mechanism and the market impact. We classify these market participants into three types: fully informed, partially informed and uninformed investors. We consider the portfolio optimization problems and compare the optimal trading and wealth processes for the three classes of investors theoretically and by numerical illustrations. Finally we study the portfolio optimization problems with risk constraints and make comparison with the results without risk constraints.

[1]Based on the paper with Cody Hyndman.

[2]Based on the paper with Caroline Hillairet, Cody Hyndman and Ying Jiao.

Acknowledgments

Firstly, I would like to express my sincere gratitude to my co-advisors Prof. Cody Hyndman and Prof. Sun Wei for their continuous support of my Ph.D study and related research. Their guidance helped me in all the time of research and writing of this thesis. I would also like to thank Prof. Caroline Hillairet (Université Paris Saclay, France) and Prof. Ying Jiao (Université Lyon, France), who generously share their knowledge and thoughts in our cooperative work on optimal trading problem with asymmetric information.

Secondly, I would like to thank my fellow graduate students in Mathematics and Statistics Department of Concordia University, Anastasis Kratsios, Xiang Gao and Chengrong Xie, for their feedback, cooperation and of course friendship.

Finally, I must express my very profound gratitude to my parents and to my wife for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

List of Figures

List of Tables

Introduction

The pricing problem for zero-coupon bonds based on an underlying short term interest rate process $r(t) \in \mathbb{R}^+$ is a fundamental and important topic in financial mathematics. Various models for $r(t)$ have been proposed under the risk neutral measure. One-factor models use the instantaneous spot rate $r(t)$ as the basic state variable, such as Vasicek [63] and Cox et al. [20]. Multi-factor models in which the short rate depends on a multidimensional factor process include the models of Longstaff and Schwartz [55], Hull and White [37], and Duffie and Kan [24]. There are several ways to characterize the bond price. In an arbitrage free market the bond price can be viewed as a solution to a partial differential equation called the term structure equation (see Björk [15, Proposition 21.2]) or, linked through the Feynman-Kac formula, by using risk neutral valuation (see Björk [15, Proposition 21.3]). Recently alternative approaches have been studied including the stochastic flow approach (see Elliott and van der Hoek [29], Hyndman and Zhou [38], and Hyndman [40]), a forward-backward stochastic differential equation approach (see Hyndman [39, 40] and Hyndman and Zhou [38]), and an optimal stochastic control approach of Gombani and Runggaldier [35].

Gombani and Runggaldier [35] associate the pricing problem of default-free bonds with an optimal stochastic control (OSC) problem by transforming the term structure equation to an equivalent Hamilton-Jacobi-Bellman equation. Inspired by Gombani and Runggaldier [35] and the notion of relative entropy we develop an optimal measure transformation (OMT) problem whose value function is connected with the price of bonds. We explore the equivalence between the OMT problem and OSC problem. One advantage of the OMT problem compared to the OSC problem is the straightforward extension to models with jumps or even to models for defaultable bonds. The OMT problem also provides a financial interpretation of the pricing problem in terms of maximization of returns subject to an entropy penalty term that quantifies financial risk.

We show that the optimal measure and the value process of the OMT problem can be completely characterized by a forward-backward stochastic differential equation (FBSDE). In addition, the optimal measure transformation has an explicit expression provided that the related FBSDE admits an explicit solution. From the explicit representation of the optimal measure transformation we note that the measure which solves the OMT problem coincides with the martingale measure using bond price as numéraire or the forward measure. These connections provide some insight into why the forward measure transformation employed in the FBSDE approach of Hyndman [40] is effective. Under the framework of affine term structure models (ATSMs) and QTSMs, Hyndman [40] and Hyndman and Zhou [38] presented explicit solutions for the related FBSDE.

We extend the OMT problem to include jumps and give explicit solutions of the related FBSDE with jumps under ATSMs and QTSMs extending the results of Hyndman [40] and Hyndman and Zhou [38]. Optimal measure transformation problems for futures and forward prices are also considered. Finally, we study the OMT problem for defaultable bonds. Due to the random payoff of defaultable bonds the related FBSDE terminal value depends generally on the default time and recovery amount. We obtain a partially explicit solution for the FBSDE whose solution relies on a Riccati equation and another simpler BSDE that incorporates the default variables.

In the second part of the thesis we are concerned with the optimal trading problem against a disorderly liquidation impact under asymmetric information. There is a large amount of literature on insider trading, asymmetric information, and market manipulation trading strategies including seminal works by [51, 11, 42, 43, 3]. These works generally assume that an insider is attempting to influence a price by, or profit from, the release of, potentially false, information known to the insider. These studies also generally break market participants into noise traders, standard informational traders, and informed traders. The existence of arbitrage strategies, price equilibrium, or specific market manipulation strategies are the primary concerns of these early works. Other papers dealing with insider information which quantify the value of insider information through the maximization of agents wealth or utility include [61, 28, 7, 8].

More recently liquidity modeling has become an intense area of study. Market micro-structure and limit order books present one approach to modelling liquidity based on trading mechanisms. Models that specify the price impact of trades as exogenously determined and depending on the size of a transactions constitute another strand of the literature. Both approaches treat problems associated with the fact that trading large positions impacts market prices. A good overview of liquidity models can be found in [34]. The modeling of market micro-structure and the optimal liquidation of large positions has also been studied extensively and an overview of these topics can be found in [1]. To the best of our knowledge, among works dealing with asymmetric information, only few papers concern the market impact of liquidation risk. In particular, [9] studies optimal liquidation problems of an insider.

In contrast to the existing literature we are concerned with disorderly, rather than optimal, liquidation and the point of view of market participants other than the large trader or hedge fund liquidating the position. In particular, we are interested in the following question: is it possible for a market participant to profit from the knowledge that another market participant, with large positions in a stock or derivative, will be forced to liquidate some or all of its position if the price crosses a certain threshold? There is ample evidence from financial markets concerning the importance of liquidity risks. For example, consider a hedge fund with a large position in natural gas futures contracts, such as Amaranth Advisors LLC in 2006, and macro-economic or weather events contribute to an unexpected adverse change in the price. In this case the fund may be forced to unwind its positions in a disorderly fashion, which would have a further market impact on the price. Other examples include the case of Long Term Capital Management L.P. (LTCM) in 1998 and numerous firms during the financial crisis of 2007-2008.

We assume that liquidation occurs immediately when the market price hits the liquidation trigger level and has a temporary impact on the asset price, whereby the market price is depressed away from the fundamental value, and gradually dissipates. We model the temporary market impact by a function with parameters that control the impact speed and magnitude. Other market participants may have different levels of information about the liquidation trigger mechanism and the liquidation impact. We aim to find the optimal trading strategy that maximizes an investor's terminal utility of wealth under different types of information that are accessible to particular market participants. In the standard information case an uninformed market participant is not aware of the liquidation trigger mechanism. They believe and act, erroneously when liquidation occurs, as if the market price is equal to the fundamental asset price. In the partial information case an insider or informed market participant knows the level at which the hedge fund will be forced to liquidate the position but does not have information about the liquidation volume which determines the price impact. In the full information case the insider has complete information about the liquidation threshold and the price impact. Certain market participants may have access to this type of information owing to their position, counter-party status, technology, or knowledge of the market. The fully informed investor's perfect information represents one extreme which may be unobtainable in practice. However, we shall show numerically in the power-utility case that the optimal strategy for the partially informed investor is quite close to that of the fully informed investor.

The remainder of the thesis is organized as follows. We review some classical results in Chapter 1. Chapter 2 presents the results of optimal measure transformation. Chapter 3 discusses the optimal trading problem and Chapter 4 extends to the optimal trading problem with risk constraints. Chapter 5 summarizes our results and discusses some future work and an appendix contains technical results and proofs.

Chapter 1

Preliminary Results

In this section we review some preliminary results about bond pricing and optimal portfolio investment.

1.1 PDE approach to bond pricing

Following Lemke [52] we model the financial market on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, 0 \leq t \leq T\}, \mathbb{P}^0)$, where $\mathcal{F}_t$ satisfies the usual conditions. The source of randomness is a standard $\mathbb{P}^0$-Brownian motion $W_t^{\mathbb{P}^0}$ which is adapted to $\mathcal{F}_t$. At this moment we suppose $\mathbb{P}^0$ to be real-world measure. We are mainly concerned with the pricing problem of a default-free zero-coupon bond with maturity T or T-bond for short.

1.1.1 One-factor models

We denote by r_t the instantaneous rate of interest and the money account process B_t is thus given by

$$B_t = \exp\left(\int_0^t r_s ds\right). \tag{1.1.1}$$

We denote the T-bond price at time t by $P(t,T)$. With the no arbitrage argument, there exists a probability measure $\mathbb{P}$ equivalent to $\mathbb{P}^0$ such that the discounted bond price process using the money account as numeraire is a martingale. Hence we have the following bond pricing formula

$$P(t,T) = \mathbb{E}^{\mathbb{P}}[\exp(-\int_t^T r_s ds)|\mathcal{F}_t], \tag{1.1.2}$$

where $\mathbb{E}^{\mathbb{P}}$ stands for the expectation with respect to the measure $\mathbb{P}$. Clearly the bond price $P(t,T)$ depends upon the behavior of the short rate of interest over the interval $[t,T]$. Let us model the short rate r_t as the solution of an SDE of the form

$$dr_t = \mu(t,r_t)dt + \sigma(t,r_t)dW_t^{\mathbb{P}^0} \tag{1.1.3}$$

under the real-world measure $\mathbb{P}^0$. We assume sufficient conditions on μ, σ to ensure existence and uniqueness of a strong solution to equation (1.1.3). This issue is discussed in Appendix A.1, however, we have yet to make any modelling assumptions.

The interest rate model in (1.1.3) is referred to as the one-factor model. The most prominent one-factor models are those by Vasicek [63], Cox et al. [20] and Hull and White [37]. Notice that the dynamics of the short rate r_t are given under $\mathbb{P}^0$ whereas the pricing formula (1.1.2) involves the martingale measure $\mathbb{P}$. To connect $\mathbb{P}$ with $\mathbb{P}^0$ mathematically, we assume that the measure transformation from $\mathbb{P}^0$ to $\mathbb{P}$ can be constructed by the Radon-Nikodym derivative

$$\left.\frac{d\mathbb{P}}{d\mathbb{P}^0}\right|_{\mathcal{F}_t} = \exp\left(-\int_0^t \lambda_s dW^{\mathbb{P}_t^0} - \frac{1}{2}\int_0^t |\lambda_s|^2 ds\right)$$

where λ_t satisfies

$$\mathbb{E}^{\mathbb{P}^0}[\exp\left(\frac{1}{2}\int_0^t |\lambda_s|^2 ds\right)] < \infty,$$

which is known as Novikov's condition. Then by Girsanov's theorem we know that

$$W^{\mathbb{P}} = W^{\mathbb{P}^0} + \int_0^t \lambda_s ds \tag{1.1.4}$$

is a standard $\mathbb{P}$-Brownian motion.

With the relation given by (1.1.4) we rewrite (1.1.1)

$$dr_t = (\mu(t,r_t) - \lambda_t \sigma(t,r_t))\, dt + \sigma(t,r_t)dW_t^{\mathbb{P}} \tag{1.1.5}$$

under the martingale measure $\mathbb{P}$. One can possibly solve the SDE (1.1.5) and then evaluate the expectation of the integral in (1.1.2). For example, the Vasicek model is defined by the dynamics

$$dr_t = k[\theta - r_t]dt + \sigma dW_t^{\mathbb{P}}.$$

The above SDE is linear and can be solved explicitly. Therefore the bond price given by (1.1.2) can be computed as an explicit expression only depending on k, θ, σ and r_t. However this explicit computation is not always feasible for a general interest rate model.

An alternative approach is to characterize the bond price by a partial differential equation (PDE). We may think of the bond price as a smooth function of two variables: the time t and the interest rate r_t and write the bond price as

$$P(t,T) = F^T(t, r_t)$$

where the superscript T is regarded as a parameter. Then the function $F^T(\cdot,\cdot)$ satisfies the PDE

$$\frac{\partial F^T(t,r)}{\partial t} + (\mu - \sigma\lambda_t)\frac{\partial F^T(t,r)}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 F^T(t,r)}{\partial r^2} - rF^T(t,r) = 0 \tag{1.1.6}$$

with the boundary condition $F^T(T,r) = 1$ for any r. This connection is a direct application of the Feynman-Kac formula.

On the other hand, we apply Itô's formula to $F^T(t,r_t)$ to find that

$$dF^T(t,r_t) = (\frac{\partial F^T(t,r_t)}{\partial t} + \mu\frac{\partial F^T(t,r_t)}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 F^T(t,r_t)}{\partial r^2})dt + \sigma\frac{\partial F^T(t,r_t)}{\partial r}dW_t^{\mathbb{P}^0}. \tag{1.1.7}$$

By substituting (1.1.6) into (1.1.7) we rewrite (1.1.7) as

$$\frac{dF^T(t,r_t)}{F^T(t,r_t)} = \left(r_t + \lambda_t\frac{\sigma}{F^T(t,r_t)}\frac{\partial F^T(t,r_t)}{\partial t}\right)dt + \frac{\sigma}{F^T(t,r_t)}\frac{\partial F^T(t,r_t)}{\partial t}dW_t^{\mathbb{P}^0}.$$

From the above SDE we can identify the drift and volatility of the return on the T-bond as

$$m_t = r_t + \lambda_t\frac{\sigma}{F^T(t,r_t)}\frac{\partial F^T(t,r_t)}{\partial t}, \tag{1.1.8}$$

$$s_t = \frac{\sigma}{F^T(t,r_t)}\frac{\partial F^T(t,r_t)}{\partial t}. \tag{1.1.9}$$

By inserting (1.1.9) into (1.1.8) it is easy to find the relation

$$\lambda_t = \frac{m_t - r_t}{s_t}.$$

λ_t is also referred to as the market price of risk. As pointed out by Björk [15, Chapter 23], the process λ_t is determined by the financial market, or in other words, the corresponding martingale measure $\mathbb{P}$ is chosen by the market. Notice that the term $\mu - \lambda_t \sigma$ in the PDE (1.1.6) is exactly the drift term in the dynamics of the short rate r_t in (1.1.1) under the martingale measure $\mathbb{P}$. It means that the bond price is completely determined by the dynamics of the short rate r_t under the martingale measure. Instead of specifying μ and λ under the objective probability measure, it is more convenient to directly model the short rate under the martingale measure $\mathbb{P}$, which is known as martingale modeling. We assume that r_t under $\mathbb{P}$ has dynamics given by

$$dr(t) = \tilde{\mu}(t, r_t)dt + \sigma(t, r_t)dW_t^{\mathbb{P}}.$$

Then the PDE (1.1.6) becomes

$$\frac{\partial F^T(t,r)}{\partial t} + \tilde{\mu}\frac{\partial F^T(t,r)}{\partial r} + \frac{1}{2}\sigma^2 \frac{\partial^2 F^T(t,r)}{\partial r^2} - rF^T(t,r) = 0,$$

which is referred to as the term structure equation.

1.1.2 Multifactor models of the short rate

In the previous one-factor model, the short rate dynamics is driven by one single Brownian motion. Multifactor models extends this model by incorporating more than one source of randomness to drive the short rate process. An n-dimensional factor process X is defined by the SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t^{\mathbb{P}} \tag{1.1.10}$$

where $W_t^{\mathbb{P}}$ is a d-dimensional $\mathbb{P}$-Brownian motion. $\mu(\cdot, \cdot)$ is a function from $\mathbf{R}^+ \times \mathbf{R}^n$ to $\mathbf{R}^n$ and $\sigma(\cdot, \cdot)$ is a function from $\mathbf{R}^+ \times \mathbf{R}^n$ to $\mathbf{R}^{n\times n}$. We assume sufficient conditions on μ, σ to ensure existence and uniqueness of a solution to equation (1.1.3). This issue is discussed in Appendix A.1, however, we have yet to make any modelling assumptions.

The short rate r_t is modeled as a function of the factor process

$$r_t = r(X_t)$$

for some function $r(\cdot)$ from $\mathbf{R}^n$ to $\mathbf{R}$.

Following a similar procedure as in the previous section, we obtain the term structure equation for the T-bond price

$$\frac{\partial}{\partial t}F^T(t,r) + (\mu(t,r))'\nabla_r F^T(t,r) + \frac{1}{2}\mathrm{tr}\left\{(\sigma(t,r))'\nabla_{rr}F^T(t,r)\sigma(t,r)\right\} - F^T(t,r)r = 0$$
$$F^T(T,r) = 1$$

with subindices r and t denoting partial derivatives.

With particular specifications on $\mu(\cdot,\cdot)$, $\sigma(\cdot,\cdot)$ and $r(\cdot)$, the term structure equation has explicit solution. In the following two subsections, we introduce two classes of term structure models: affine term structure models and quadratic term structure models.

1.1.3 Affine term structure

The so called affine term structure model (see Björk [15, Chapter 24]) is the framework under which the bond $P(t,T)$ is exponential affine in the short rate r_t, i.e.

$$P(t,T) = F^T(t,r_t) = e^{A(t,T)-B(t,T)r_t} \tag{1.1.11}$$

where A and B are deterministic functions. We next explore heuristically the proper choice of μ and σ under $\mathbb{P}$ for r_t so that the affine term structure holds. In the first place, we consider the one-factor model. Using the bond price $P(t,T)$ given in (1.1.11) above, we may easily compute the various partial derivatives of $F^T(t,r)$ and substitute into (1.1.6) to obtain

$$A_t(t,T) - (1 + B_t(t,T))r - \mu(t,r)B(t,T) + \frac{1}{2}\sigma^2(t,r)B^2(t,T) = 0, \tag{1.1.12}$$

where $A_t(t,T)$ and $B_t(t,T)$ denote the derivatives with respect to t. The boundary condition

$$P(T,r;T) = 1$$

implies

$$A(T,T) = 0,$$

$$B(T,T) = 0.$$

The idea is to separate $A(t,T)$ and $B(t,T)$ into two equations. Assume that μ and σ have the form

$$\mu(t,r) = \alpha(t)r + \beta(t),$$
$$\sigma(t,r) = \sqrt{\gamma(t)r + \delta(t)}.$$

Then we reorgnize (1.1.12) as

$$\begin{aligned} &A_t(t,T) - \beta(t)B(t,T) + \frac{1}{2}\delta(t)B^2(t,T) \\ &\qquad - (1 + B_t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B^2(t,T))r = 0. \end{aligned}$$

Since the equation holds for any r the coefficient of r must be equal to be zero. Hence we find the two equations as below

$$B_t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B^2(t,T) + 1 = 0 \tag{1.1.13}$$
$$A_t(t,T) - \beta(t)B(t,T) + \frac{1}{2}\delta(t)B^2(t,T) = 0. \tag{1.1.14}$$

We may first solve equation (1.1.13) for $B(t,T)$ and then insert it into equation (1.1.14) to find $A(t,T)$. Based on the discussion above we formulate the following result.

Proposition 1.1.1. *Assume that μ and σ have the form*

$$\mu(t,r) = \alpha(t)r + \beta(t),$$
$$\sigma(t,r) = \sqrt{\gamma(t)r + \delta(t)}.$$

Then the model admits an affine term structure of the form in (1.1.12)*, where A and B satisfy the system*

$$B_t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B^2(t,T) = -1, \tag{1.1.15}$$
$$A_t(t,T) = \beta(t)B(t,T) - \frac{1}{2}\delta(t)B^2(t,T), \tag{1.1.16}$$
$$B(T,T) = 0, \quad A(T,T) = 0.$$

Note that (1.1.15) is a Riccati equation for the determination of B which does not involve A. Having solved (1.1.16) we may then insert the solution B into (1.1.16) and simply integrate in order to obtain A.

In case of multifactor models, we model the short rate r_t as a linear function of the factors, i.e.

$$r_t = (U_t)' X_t + P_t$$

where U_t is a d-dimensional vector and P_t is a scalar. Then an similar result to (1.1.1) could be derived (see Duffie et al. [27], Duffie and Kan [24], Dai and Singleton [21]). Tractability is the main advantage of affine term structure models (ATSMs). However, ATSMs fail to capture some empirically observed nonlinearities as shown by Dai and Singleton [21].

1.1.4 Quadratic term structure

Under the framework of quadratic term structure models (see Ahn et al. [2]), the factor process is traditionally modeled by the SDE of the form

$$dX_t = (F_t X_t + H_t)dt + G_t dW_t^{\mathbb{P}}$$

and the short rate is quadratic function of the factors

$$r_t = (X_t)' Q_t X_t + (U_t)' + P_t,$$

where P_t is a scalar, H_t, U_t are n-dimensional vectors, and Q_t, F_t are $n \times n$ symmetric matrices and G_t is a $n \times d$ matrix. Then the bond price has the exponential quadratic form

$$P(t,T) = \exp\{-A(t,T) - (X_t)' B(t,T) + (X_t)' C(t,T) X_t\}$$

where $A(t,T), B(t,T), C(t,T)$ satisfy the following system of ODEs

$$\begin{aligned}
&\frac{\partial}{\partial t} C(t,T) + 2F_t(t,T) - 2(t,T) G_t (G_t)' C(t,T) + Q_t = 0 \\
&\frac{\partial}{\partial t} B(t,T) + (F_t)' B(t,T) + 2C(t,T) H_t - 2(G_t)' G_t C(t,T) B(t,T) + U_t = 0 \\
&\frac{\partial}{\partial t} A(t,T) + (B(t,T))' H_t - \frac{1}{2}(B(t,T))' G_t (G_t)' B(t,T) + \mathrm{tr}\left((G_t)' C(t,T) G_t\right) + P_t = 0 \\
&A(T,T) = 0, \quad B(T,T) = 0, \quad C(T,T) = 0.
\end{aligned}$$

Compared with ATSMs, QTSMs have more flexibility to characterize the term structure of bond price. Ahn et al. [2] showed the conditions for QTSMs which guarantee closed form solutions for the bond price.

1.2 FBSDE approach to bond pricing

Under the framework of ATSMs, Hyndman [40] characterized the bond price via a forward-backward stochastic differential equation (FBSDE). For simplicity of notations, we consider the one-factor model which is a special case of Hyndman [40]. The factor process is given by the SDE

$$dX_t = (AX_t + B)dt + (\sqrt{\alpha + \beta X_t})dW_t^{\mathbb{P}} \tag{1.2.1}$$

and the interest rate is modeled as linear function of the factor

$$r_t = RX_t + k$$

where $A, B, \alpha, \beta, R, k$ are all scalars.

Define $H_t = \exp\left(-\int_0^t r_s ds\right)$ and $V_t = \mathbb{E}[\exp\left(-\int_0^T r_s ds\right)|\mathcal{F}_t]$. Simple computation gives us

$$dH_t = -r_t H_t dt.$$

As a martingale V_t has the following representation

$$V_t = V_0 + \int_0^t J_s dW_s^{\mathbb{P}} \tag{1.2.2}$$

where J is a progressively measurable process.

Recall that the bond price is given by

$$P(t,T) = \mathbb{E}^{\mathbb{P}}[\exp\left(\int_t^T r_s ds\right)|\mathcal{F}_t],$$

and notice that $P(t,T) = \frac{V_t}{H_t}$. Let $Y_t = P(t,T)$ and using Itô's formula we find that

$$dY_t = r_t Y_t dt + \frac{J_t}{H_t} dW_t^{\mathbb{P}}. \tag{1.2.3}$$

Define $Z_t = \frac{J_t}{H_t}$ and recall $r_t = RX_t + k$ to rewrite (1.2.3) as

$$dY_t = (RX_t + k)Y_t dt + Z_t dW_t^{\mathbb{P}}. \tag{1.2.4}$$

The fact $P(T,T) = 1$ implies the boundary condition $Y_T = 1$. Combining (1.2.4) with (1.2.1) we obtain the decoupled FBSDE

$$dX_t = (AX_t + B)dt + (\sqrt{\alpha + \beta X_t})dW_t^{\mathbb{P}} \tag{1.2.5}$$

$$dY_t = (RX_t + k)Y_t dt + Z_t dW_t^{\mathbb{P}}. \tag{1.2.6}$$

If we use the T-bond as numeraire, the corresponding martingale measure, or the so called forward measure, $\mathbb{Q}$ is defined via the Radon-Nikodym derivative

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \Lambda_T = \{P(0,T)\}^{-1} \exp\left(-\int_0^T r_s ds\right).$$

Define $\Lambda_t = \mathbb{E}[\Lambda_T|\mathcal{F}_t]$. We observe the fact that $\Lambda_t = \frac{V_t}{V_0}$. Recall (1.2.2) to find that

$$\Lambda_t = 1 + \int_0^t \frac{Z_s}{Y_s}\Lambda_s dW_s^{\mathbb{P}}.$$

Then from Girsanov theorem the $\mathbb{Q}$-Brownian motion is given by

$$W_t^{\mathbb{Q}} = W_t^{\mathbb{P}} - \int_0^t \frac{Z_s}{Y_s} ds.$$

Under the forward measure $\mathbb{Q}$ the FBSDE (1.2.5)-(1.2.6) is rewritten as

$$dX_t = \left\{AX_t + B + (\sqrt{\alpha + \beta X_t})\frac{Z_t}{Y_t}\right\} dt + (\sqrt{\alpha + \beta X_t})dW_t^{\mathbb{Q}} \tag{1.2.7}$$

$$dY_t = \left\{(RX_t + k)Y_t + \frac{Z_t^2}{Y_t}\right\} dt + Z_t dW_t^{\mathbb{Q}}. \tag{1.2.8}$$

By adapting a technique for linear FBSDEs from Ma and Yong [57], the main result of Hyndman [40] below can be proven.

Theorem 1.2.1. *If the Riccati equation*

$$\dot{U}_t + AU_t + \frac{1}{2}\beta U_t^2 - \beta U_t - R = 0, \quad U_T = 0$$

admits a unique solution over the interval $[0,T]$ *then the FBSDE (1.2.7)-(1.2.8) admits a unique adapted solution* (X, Y, Z) *with explicit expression given by*

$$\begin{aligned} dX_t &= (AX_t + B + (\alpha + \beta)U_t X_t)\, dt + S\, diag(\sqrt{\alpha + \beta X_t})dW_t^{\mathbb{P}} \\ \ln Y_t &= U_t X_t + p_t, \quad \textit{and} \\ Z_t &= U_t(\sqrt{\alpha + \beta X_t})Y_t, \end{aligned}$$

where

$$p_t = -\int_t^T \left(k + \alpha U_s - BU_s - \frac{1}{2}\alpha U_s^2\right) ds.$$

Hyndman [40] applied the FBSDE approach to ATSMs while Hyndman and Zhou [38] applied the FBSDE approach to QTSMs and obtained similar results to Theorem 1.2.1. In this thesis, we will extend the results of Hyndman [40] and Hyndman and Zhou [38] to factor processes with jumps as well as to defaultable bonds. Further, by developing the new theory of optimal measure change, we shall give a new interpretation of the FBSDE approach.

1.3 Optimal portfolio investment

In this section we consider the optimal portfolio investment problem and present some classical results following Björk [15]. The financial market under consideration consists of one risky asset S_t given by

$$dS_t = S_t(\alpha_t dt + \sigma_t dW_t^{\mathbb{P}^0}) \tag{1.3.1}$$

and one riskless asset B_t given by

$$dB_t = rB_t dt.$$

Denote the relative portfolio weights on the risky asset by π_t, then the wealth process X_t is given by

$$dX_t = X_t\{\pi_t\alpha_t dt + (1-\pi_t)rdt + \pi_t\sigma_t dW_t^{\mathbb{P}^0}\}$$

Let us consider an investor with initial capital x and a utility function U for terminal wealth. The utility function is assumed to satisfy the Inada condition:

- $U(x)$ is twice differentiable on $(0, \infty)$,
- $U'(x) > 0$ and $U''(x) < 0$ for each $0 < x < \infty$
- $U'(0) = \infty$ and $\lim_{x\to\infty} U'(x) = 0$.

The investors' objective is to maximize the expected utility

$$V_0 = \sup_{\pi\in\mathcal{A}} \mathbb{E}^{\mathbb{P}^0}[U(X_T)], \tag{1.3.2}$$

where $\mathcal{A}$ is the set of admissible portfolio strategies.

Instead of solving the optimization problem (1.3.2) directly, we first consider the following static problem

$$\max_{X_T \in \mathcal{K}_T} \mathbb{E}^{\mathbb{P}^0}[U(X_T)] \tag{1.3.3}$$

where $\mathcal{K}_T$ represents the set of contingent T-claims which can be replicated by a self-financing portfolio with initial capital x. In this formulation, our main concern is not on the optimal portfolio strategy but instead on the terminal wealth X_T. Once we find the optimal wealth $\hat{X}_T$ we can compute the corresponding generating portfolio using martingale representation results.

Denote by $\mathbb{P}$ the martingale measure and L_t the likelihood process between $\mathbb{P}^0$ and $\mathbb{P}$, i.e.

$$L_t = \left.\frac{d\mathbb{P}}{d\mathbb{P}^0}\right|_{\mathcal{F}_t}.$$

From the price dynamics (1.3.1) and Girsanov's theorem it is easily seen that L is given by

$$\begin{aligned} dL_t &= (\sigma_t^{-1}(r - \alpha_t))dW_t^{\mathbb{P}^0}, \\ L_0 &= 1. \end{aligned}$$

Since the discounted wealth process $e^{-rt}X_t$ is a $\mathbb{P}$-martingale, the optimization problem (1.3.3) is subject to the budget constraint

$$e^{-rT}\mathbb{E}^{\mathbb{P}}[U(X_T)] = x. \tag{1.3.4}$$

By rewriting the budget constraint (1.3.4) as

$$e^{-rT}\mathbb{E}^{\mathbb{P}^0}[X_T L_T] = x,$$

we can solve the optimization problem (1.3.3) with the constraint (1.3.4) using the method of Lagrange multipliers. We only need to solve the following unconstrained optimization problem

$$\begin{aligned} W_0 &= \sup_{X_T \in \mathcal{K}_T} \mathbb{E}^{\mathbb{P}^0}[U(X_T)] - \lambda(e^{-rT}\mathbb{P}[X_T L_T] - x) \\ &= \sup_{X_T \in \mathcal{K}_T} \int_\Omega \left\{U(X_T(\omega)) - \lambda[e^{-rT}L_T(\omega)X_T(\omega) - x]\right\} d\mathbb{P}(\omega). \end{aligned}$$

The optimal solution is given by

$$\hat{X}_T = I(\lambda e^{-rT} L_T)$$

where $I(\cdot)$ is the inverse of the derivative of the utility function U, i.e. $I(x) = (U')^{-1}$.

It remains to determine the optimal portfolio strategy which generates $\hat{X}_T$. Define $\tilde{X}_t = e^{-rt}\hat{X}_t$ and denote by $\hat{\pi}_t$ the optimal strategy. From the Itô formula we have

$$d\tilde{X}_t = \tilde{X}_t \hat{\pi}_t \{(\alpha_t - r)dt + \sigma_t dW_t^{\mathbb{P}^0}\}. \tag{1.3.5}$$

We rewrite (1.3.5) under the martingale measure $\mathbb{P}$ as

$$d\tilde{X}_t = \tilde{X}_t \hat{\pi}_t \sigma_t dW_t^{\mathbb{P}}. \tag{1.3.6}$$

On the other hand, we have

$$\tilde{X}_t = \mathbb{E}^{\mathbb{Q}}[e^{-rT}\hat{X}_T | \mathcal{F}_t]. \tag{1.3.7}$$

By the martingale representation theorem $\tilde{X}$ has dynamics of the form

$$d\tilde{X}_t = \xi_t dW_t^{\mathbb{Q}}, \tag{1.3.8}$$

for some adapted process ξ_t. Comparing (1.3.8)-(1.3.6) we determine the optimal portfolio strategy

$$\hat{\pi}_t = (\tilde{X}_t \sigma_t)^{-1} \xi_t.$$

We need to point out that the martingale representation theorem is an existence theorem, that means we only know the existence of ξ_t in (1.3.7) but do not have closed form expression. However we are able to find explicit solution for some particular utility functions. We skip the detailed derivation and present the solutions for the optimization problem in cases of power utility and logarithmic utility respectively (refer to Björk [15]).

Proposition 1.3.1. *We consider the utility function is of the form*

$$U(x) = \frac{x^p}{p}$$

for some non-zero $p < 1$. Then the optimal strategy is given by

$$\hat{\pi}_t = \frac{\alpha_t - r}{(1-p)\sigma_t^2}$$

and the optimal utility is

$$V_0 = e^{rpT}\frac{x^p}{p}\left(\mathbb{E}^{\mathbb{P}}[L_T^{-p}]\right)^{1-p}.$$

Proposition 1.3.2. *We consider the utility function is of the form*

$$U(x) = \log(x)$$

Then the optimal strategy is given by

$$\hat{\pi}_t = \frac{\alpha_t - r}{\sigma_t^2}.$$

and the optimal utility is

$$V_0 = e^{rpT}\frac{x^p}{p}\left(\mathbb{E}^{\mathbb{P}}[L_T^{-p}]\right)^{1-p}.$$

In the second part of this thesis, we will study a new optimal trading problem which incorporates disorderly market liquidation and different levels of information accessible to market participants.

Chapter 2

Optimal measure transformation problems

2.1 Default-free bonds

We set up our model on a filtered probability space $(\Omega, \mathcal{A}, \{\mathcal{F}_s, 0 \leq s \leq T\}, \mathbb{P})$, where T is the investment horizon and $\mathbb{P}$ is a martingale measure using the money market account as numéraire. Suppose X_s is an $\mathbf{R}^n$-valued, $\mathcal{F}_s$-adapted factor process satisfying

$$dX_s = f(s, X_s)ds + g(s, X_s)dW_s^{\mathbb{P}} \tag{2.1.1}$$

where $W^{\mathbb{P}}$ is an n-dimensional $(\mathcal{F}, \mathbb{P})$-Brownian motion. Denote the short term interest rate by r_s, which can be characterized as a function of factors, $r_s = r(X_s)$, for some function $r(\cdot) : \mathbf{R}^n \to \mathbf{R}$. The price of default-free zero-coupon bonds at time $t \in [0, T]$ is then given by

$$P(t, T) = E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} | \mathcal{F}_t]. \tag{2.1.2}$$

In the following subsection we associate the bond price with an optimal measure transformation (OMT) problem.

2.1.1 The optimal measure transformation problem

Let $\mathcal{P}(\Omega)$ be the set of probability measures on $(\Omega, \mathcal{F})$. The following definitions generalize the classic definitions of the free energy and the relative entropy given in Dai Pra et al. [22] to the aggregate or dynamic version that incorporates the presence of a filtration $\mathcal{F}_s$.

Definition 2.1.1. *For $\mathbb{P} \in \mathcal{P}(\Omega)$ and φ an $\mathcal{F}_T$-measurable random variable, the aggregate free energy of φ with respect to $\mathbb{P}$, $\varepsilon_{t,T}(\varphi)$, is defined by*

$$\varepsilon_{t,T}(\varphi) = \ln(E_{\mathbb{P}}[e^{\varphi}|\mathcal{F}_t]), \quad t \in [0, T]. \tag{2.1.3}$$

Definition 2.1.2. *Consider, in addition to $\mathbb{P}$, another $\mathbb{Q} \in \mathcal{P}(\Omega)$. Suppose the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ is*

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_s} = \Gamma_s, \quad 0 \leq s \leq T. \tag{2.1.4}$$

Then, for $t \in [0, T]$, the aggregate relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as

$$H_{t,T}(\mathbb{Q}|\mathbb{P}) = \begin{cases} E_{\mathbb{Q}}[\ln(\frac{\Gamma_T}{\Gamma_t})|\mathcal{F}_t] & \textit{if} \quad \ln(\frac{\Gamma_T}{\Gamma_s}) \in L^1(\mathbb{P}), \\ +\infty & \textit{otherwise.} \end{cases} \tag{2.1.5}$$

For $t \in [0, T]$ we define a family of probability measures $\mathcal{P}_t(\Omega) \subseteq \mathcal{P}(\Omega)$ which are equivalent to $\mathbb{P}$ on $\mathcal{F}_t$ as

$$\mathcal{P}_t(\Omega) = \{ \ \mathbb{Q} \in \mathcal{P}(\Omega) \mid \mathbb{Q} \ll \mathbb{P} \quad \text{and} \quad \mathbb{Q}|_{\mathcal{F}_t} = \mathbb{P}|_{\mathcal{F}_t} \} \tag{2.1.6}$$

where $\mathbb{Q} \ll \mathbb{P}$ means that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$. Similar to Dai Pra et al. [22], the following proposition reveals the duality relationship between the aggregate free energy and the aggregate relative entropy.

Proposition 2.1.1. *For $t \in [0, T]$ and any $\mathcal{F}_T$-measurable random variable φ*

$$-\varepsilon_{t,T}(\varphi) = \inf_{\mathbb{Q} \in \mathcal{P}_t(\Omega)} \{E_{\mathbb{Q}}[-\varphi|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P})\}. \tag{2.1.7}$$

The infimum is attained at $\mathbb{Q}^\star$ determined by the Radon-Nikodym derivative

$$\left.\frac{d\mathbb{Q}^\star}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \frac{e^{\varphi}}{E_{\mathbb{P}}[e^{\varphi}|\mathcal{F}_t]}. \tag{2.1.8}$$

Proof. As in equation (2.1.4) we suppose

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_s} = \Gamma_s, \quad 0 \leq s \leq T.$$

Using the abstract Bayes' formula (see Shreve [62, Lemma 5.2.2]) we find

$$-\varepsilon_{t,T}(\varphi) = \ln(E_{\mathbb{P}}[e^{\varphi}|\mathcal{F}_t]) = \ln(E_{\mathbb{Q}}\Big[e^{\varphi}\frac{\Gamma_t}{\Gamma_T}\Big|\mathcal{F}_t\Big]). \tag{2.1.9}$$

Since for any $\mathbb{Q} \in \mathcal{P}_t(\Omega)$, we have

$$\mathbb{Q}|_{\mathcal{F}_t} = \mathbb{P}|_{\mathcal{F}_t},$$

that is

$$\Gamma_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_t} = 1.$$

Thus we may simplify equation (2.1.9) as

$$-\varepsilon_{t,T}(\varphi) = -\ln(E_{\mathbb{Q}}\Big[\frac{e^{\varphi}}{\Gamma_T}\Big|\mathcal{F}_t\Big]). \tag{2.1.10}$$

Recall Jensen's inequality gives that for a convex function f and a random variable X that

$$f(E[X|\mathcal{F}_t]) \leq E[f(X)|\mathcal{F}_t].$$

Since $-\ln(\cdot)$ is a convex function by Jensen's inequality we have that

$$-\ln(E_{\mathbb{Q}}\Big[\frac{e^{\varphi}}{\Gamma_T}\Big|\mathcal{F}_t\Big]) \leq E_{\mathbb{Q}}[-\varphi|\mathcal{F}_t] + E_{\mathbb{Q}}\Big[\ln(\Gamma_T)\Big|\mathcal{F}_t\Big]. \tag{2.1.11}$$

That is

$$-\varepsilon_{t,T}(\varphi) \leq E_{\mathbb{Q}}[-\varphi|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P}).$$

It is easy to check that the equality holds in equation (2.1.11) if we set

$$\Gamma_T = \frac{e^{\varphi}}{E_{\mathbb{P}}[e^{\varphi}|\mathcal{F}_t]}.$$

which completes the proof. □

Remark 2.1.1. *Suppose a zero coupon bond $P(t,T)$ pays \$1 at maturity date T. Then the yield of the bond over the interval $[t,T]$, denoted by γ, is*

$$\gamma_{t,T} = \ln\frac{1}{P(t,T)} = -\ln\Big\{E_{\mathbb{P}}[e^{-\int_t^T r(X_v)dv}|\mathcal{F}_t]\Big\}. \tag{2.1.12}$$

If we set $\varphi = -\int_t^T r(X_v)dv$ in equation (2.1.7), the left hand side of equation (2.1.7) is equal to the yield $\gamma_{t,T}$ given by equation (2.1.12). Then, by Proposition 2.1.1, we can construct the following optimal measure transformation problem.

Problem 2.1.1. *On a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_s, 0 \leq s \leq T\}, \mathbb{P})$ suppose that the factor process $(X_s, 0 \leq s \leq T)$ is given by*

$$dX_s = f(s, X_s)dt + g(s, X_s)dW_s^{\mathbb{P}}.$$

For $t \in [0, T]$ and any $\mathbb{Q} \in \mathcal{P}_t(\Omega)$, the performance criterion $J_{t,T}(\mathbb{Q})$ is defined as

$$J_{t,T}(\mathbb{Q}) = E_{\mathbb{Q}}\Big[\int_t^T r(X_v)dv\Big|\mathcal{F}_t\Big] + H_{t,T}(\mathbb{Q}|\mathbb{P}). \tag{2.1.13}$$

The optimal measure transformation problem for the default-free zero coupon bond is

$$V_{t,T} = \inf_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} J_{t,T}(\mathbb{Q}). \tag{2.1.14}$$

By Proposition 2.1.1 the solution of the OMT Problem 2.1.1 is given by the optimal value process

$$V_{t,T} = -\ln\Big\{E_{\mathbb{P}}[e^{-\int_t^T r_v dv}|\mathcal{F}_t]\Big\} \tag{2.1.15}$$

and the optimal measure $\mathbb{Q}^\star$ determined by

$$\frac{d\mathbb{Q}^\star}{d\mathbb{P}}\bigg|_{\mathcal{F}_T} = \frac{e^{-\int_t^T r(X_v)dv}}{E_{\mathbb{P}}[e^{-\int_t^T r(X_v)dv}|\mathcal{F}_t]}. \tag{2.1.16}$$

Comparing equations (2.1.2) and (2.1.15) we find the connection between the value function and the bond price

$$V_{t,T} = -\ln P(t, T).$$

Remark 2.1.2. *Note that the optimal measure $\mathbb{Q}^\star$ in equation (2.1.16) is actually the martingale measure using the bond price as numéraire (refer to Björk [15, Section 26.3]), which is also called the forward measure. The measure $\mathbb{Q}^\star$ in equation (2.1.16) is slightly different from the forward measure defined in Hyndman [40] as follows*

$$\frac{d\mathbb{Q}^T}{d\mathbb{P}}\bigg|_{\mathcal{F}_T} = \frac{e^{-\int_0^T r(X_v)dv}}{E_{\mathbb{P}}[e^{-\int_0^T r(X_v)dv}]}. \tag{2.1.17}$$

The measure $\mathbb{Q}^\star$ in equation (2.1.16) corresponds to the numéraire change over $[t, T]$, whereas the measure $\mathbb{Q}^T$ in equation (2.1.17) corresponds the numéraire change over the time period $[0, T]$.

We have from Hyndman [40] that equation (2.1.15) can be characterized in terms of a decoupled FBSDE under $\mathbb{P}$. In the next subsection we explore the connection between the optimal measure transformation problem and the FBSDE approach of Hyndman [40].

2.1.2 FBSDE characterization

Following Hyndman [40] we construct a decoupled FBSDE which characterizes the OMT Problem 2.1.1. The derivation presented here differs from that in Hyndman [40] since the backward process in our formulation represents the negative logarithm of the bond price rather than the bond price.

Let $Y_s = V_{s,T}$ for $s \in [t, T]$, then by equation (2.1.15) we have

$$e^{-Y_s} = E_{\mathbb{P}}[e^{-\int_s^T r(X_v)dv}|\mathcal{F}_s]. \tag{2.1.18}$$

Multiply both sides of equation (2.1.18) by $e^{-\int_0^s r(X_v)dv}$ to obtain

$$e^{-Y_s-\int_0^s r(X_v)dv} = E_{\mathbb{P}}[e^{-\int_0^T r(X_v)dv}|\mathcal{F}_s]. \tag{2.1.19}$$

Define

$$\eta_s = E_{\mathbb{P}}[e^{-\int_0^T r(X_v)dv}|\mathcal{F}_s].$$

Then, by the martingale representation theorem, there exists an $\mathcal{F}$-predictable $(1\times n)$-vector process J such that

$$\eta_s = \eta_0 + \int_0^s J_v dW_v^{\mathbb{P}}. \tag{2.1.20}$$

Since η_s is positive almost surely, we define $Z_s = \frac{J_s}{\eta_s}$ to rewrite equation (2.1.20) as

$$\eta_s = \eta_0 + \int_0^s \eta_v Z_v dW_v^{\mathbb{P}}.$$

From equation (2.1.19), we have

$$Y_s = -\ln \eta_s - \int_0^s r(X_v)dv.$$

Then, by Itô's formula, Y_s satisfies the following BSDE

$$Y_s = \int_s^T [r(X_v) - \frac{1}{2}Z_v Z_v']dv + \int_s^T Z_v dW_v^{\mathbb{P}}. \tag{2.1.21}$$

Combining the BSDE (2.1.21) and the SDE (2.1.1) we form the decoupled FBSDE

$$X_s = X_t + \int_t^s f(v, X_v)dv + \int_t^s g(v, X_v)dW_v^{\mathbb{P}} \tag{2.1.22}$$

$$Y_s = \int_s^T [r(X_v) - \frac{1}{2}Z_v Z_v']dv + \int_s^T Z_v dW_v^{\mathbb{P}} \tag{2.1.23}$$

for $s \in [t, T]$. Clearly we have

$$Y_s = -\ln \left\{ E_{\mathbb{P}}[e^{-\int_s^T r(X_v)dv} | \mathcal{F}_s] \right\},$$

and from equation (2.1.21) we find that

$$-\ln \left\{ E_{\mathbb{P}}[e^{-\int_s^T r(X_v)dv} | \mathcal{F}_s] \right\} - \int_s^T r(X_v)dv = -\int_s^T \frac{1}{2}Z_v Z_v' dv + \int_s^T Z_v dW_v^{\mathbb{P}}. \tag{2.1.24}$$

Taking the exponential of both sides of equation (2.1.24) gives

$$\frac{e^{-\int_s^T r(X_v)dv}}{E_{\mathbb{P}}[e^{-\int_s^T r(X_v)dv} | \mathcal{F}_s]} = e^{-\int_s^T \frac{1}{2}Z_v Z_v' dv + \int_s^T Z_v dW_v^{\mathbb{P}}}. \tag{2.1.25}$$

Simply let $s = t$ in equation (2.1.25) and compare with equation (2.1.16) to find

$$\left. \frac{d\mathbb{Q}^\star}{d\mathbb{P}} \right|_{\mathcal{F}_T} = e^{-\int_t^T \frac{1}{2}Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}}}. \tag{2.1.26}$$

The solution of the OMT Problem 2.1.1 is completely characterized by the FBSDE (2.1.22)-(2.1.23). If the FBSDE (2.1.22)-(2.1.23) admits a solution triple (X, Y, Z), the value function and the optimal measure for the OMT Problem 2.1.1 are characterized as follows

$$V_{t,T} = Y_t, \tag{2.1.27}$$

$$\left. \frac{d\mathbb{Q}^\star}{d\mathbb{P}} \right|_{\mathcal{F}_T} = e^{-\int_t^T \frac{1}{2}Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}}}. \tag{2.1.28}$$

We assume that the coefficients f and g satisfy conditions such that the forward SDE (2.1.22) admits a unique solution. In particular the examples of ATSMs and QTSMs we shall consider satisfy such conditions. Further, by Kobylanski [50, Theorem 2.3], the BSDE (2.1.23) with quadratic growth also admits a unique solution. Therefore, it is not difficult to show that the decoupled FBSDE (2.1.22)-(2.1.23) admits a unique solution with appropriate assumptions on the coefficients. However, it is not always possible to obtain an explicit representation of (Y, Z) in terms of X.

Hyndman [40] considered the FBSDEs which characterizes the bond price, under the T-forward measure $\mathbb{Q}^T$, and gave explicit solutions in the framework of ATSMs. Similar to Hyndman [40] we find that, by Girsanov's theorem, the process $W^{\mathbb{Q}^\star}$ defined as

$$W_s^{\mathbb{Q}^\star} = W_s^{\mathbb{P}} - \int_t^s Z_v dv, \quad t \leq s \leq T$$

is a Brownian motion under $\mathbb{Q}^\star$. Then we may rewrite the FBSDE (2.1.22)-(2.1.23) as a nonlinear coupled FBSDE

$$X_s = X_t + \int_t^s \big(f(v, X_v) + g(v, X_v)Z_v'\big)dv + \int_t^s g(v, X_v)dW_v^{\mathbb{Q}^\star}, \tag{2.1.29}$$

$$Y_s = \int_s^T \big(r(X_v) + \frac{1}{2}Z_v Z_v'\big)dv + \int_s^T Z_v dW_v^{\mathbb{Q}^\star}. \tag{2.1.30}$$

under $\mathbb{Q}^\star$, for $s \in [t, T]$.

We calculate the aggregate relative entropy of $\mathbb{Q}^\star$ with respect to $\mathbb{P}$ explicitly in terms of Z_s as

$$\begin{aligned}
H_{t,T}(\mathbb{Q}^\star|\mathbb{P}) &= E_{\mathbb{Q}^\star}[\ln(\frac{d\mathbb{Q}^\star}{d\mathbb{P}})|\mathcal{F}_t] \\
&= E_{\mathbb{Q}^\star}[\Big(-\int_t^T \frac{1}{2}Z_s Z_s' ds + \int_t^T Z_s dW_s^{\mathbb{P}}\Big)|\mathcal{F}_t] \\
&= E_{\mathbb{Q}^\star}[\Big(\int_t^T \frac{1}{2}Z_s Z_s' ds + \int_t^T Z_s dW_s^{\mathbb{Q}^\star}\Big)|\mathcal{F}_t] \\
&= E_{\mathbb{Q}^\star}[\int_t^T \frac{1}{2}Z_s Z_s' ds|\mathcal{F}_t] + E_{\mathbb{Q}^\star}[\int_t^T Z_s dW_s^{\mathbb{Q}^\star}|\mathcal{F}_t] \\
&= E_{\mathbb{Q}^\star}[\int_t^T \frac{1}{2}Z_s Z_s' ds|\mathcal{F}_t].
\end{aligned} \tag{2.1.31}$$

Remark 2.1.3. *Though they are defined with respect to different measures, it is equivalent to consider the FBSDEs* (2.1.22)-(2.1.23) *and* (2.1.29)-(2.1.30). *The dependence of (Y, Z) on X is invariant under a change of measure. That is, if $(Y_t, Z_t) = (\Phi(X_t), \Psi(X_t))$ for some functions Φ and Ψ under the optimal measure $\mathbb{Q}^\star$, then this representation also holds under $\mathbb{P}$. In Hyndman [40], it is natural to consider the analogue of the FBSDE* (2.1.29)-(2.1.30) *under the T-forward measure $\mathbb{Q}^T$ because the stochastic flow method, which motivated the development of the FBSDE method, and the exponential form of the discount function naturally lead to the use of the forward measure.*

Hyndman [40] studied the FBSDE (2.1.29)-(2.1.30) for ATSMs which are characterized by specifying

(i) $f(s,x) = Ax + B$

(ii) $g(s,x) = S\text{diag}\sqrt{\alpha_i + \beta_i x}$

(iii) $r(x) = R'x + k$

where A is an $(n \times n)$-matrix of scalars, B, R is an $(n \times 1)$-vector of scalars, for each $i \in \{1, \dots, n\}$ the α_i are scalars, for each $i \in \{1, \dots, n\}$ the $\beta_i = (\beta_{i1}, \dots, \beta_{in})$ are $(1 \times n)$-vectors, S is a non-singular $(n \times n)$-matrix, k is a scalar. Then the FBSDE (2.1.29)-(2.1.30) becomes

$$X_s = X_t + \int_t^s \left(AX_v + B + S\text{diag}\sqrt{\alpha_i + \beta_i X_v} Z_v'\right) dv + \int_t^s S\text{diag}\sqrt{\alpha_i + \beta_i X_v} dW_v^{\mathbb{Q}^\star}, \tag{2.1.32}$$

$$Y_s = \int_s^T (R'X_v + k + \frac{1}{2} Z_v Z_v') dv + \int_s^T Z_v dW_v^{\mathbb{Q}^\star}. \tag{2.1.33}$$

From Hyndman [40], we know the solution (Y, Z) to the BSDE (2.1.33) has explicit representation, in terms of the forward process X, as follows

$$Y_s = -U_s X_s - p_s \tag{2.1.34}$$

$$Z_s = U_s S\text{diag}(\sqrt{\alpha_i + \beta_i X_s}\ ) \tag{2.1.35}$$

where U_s and p_s are both deterministic process determined by a system of Riccati-type ordinary differential equations.

Remark 2.1.4. *Note that $S\,diag(\sqrt{\alpha_i + \beta_i X}\ )$ is the volatility of the factor process X. From equation (2.1.34) we have that U is the sensitivity of Y with respect to X. This interpretation can be made more precise by considering the associated flows indexed by the starting value (t, x) and, similar to Hyndman [40, Corollary 4.2], showing that $\frac{\partial}{\partial x} Y_s^{t,x}\big|_{x=X_t} = -U_s$. In that sense the aggregate relative entropy defined in* (2.1.31) *can be interpreted as the expected aggregate sensitivity of the log bond price to the factors weighted by volatility.*

Remark 2.1.5. *Suppose a financial agent pays* c *to buy one unit of the bond at time* t*, and receives a payoff of* 1 *at maturity* T*. The (logarithmic) rate of return on the investment over the time period* $[t,T]$ *is*

$$\gamma_{t,T} = \ln \frac{1}{c}.$$

The excess return over the risk-free rate, $\tilde{\gamma}$*, is given by*

$$\tilde{\gamma}_{t,T} = \gamma_{t,T} - \int_t^T r(X_v)dv,$$

which measures the investment performance. Note that equation (2.1.14) is equivalent to

$$\begin{aligned} \ln \frac{P(t,T)}{c} &= - \inf_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \big\{ E_{\mathbb{Q}}[-\tilde{\gamma}_{t,T}|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P}) \big\} \\ &= \sup_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \big\{ E_{\mathbb{Q}}[\tilde{\gamma}_{t,T}|\mathcal{F}_t] - H_{t,T}(\mathbb{Q}|\mathbb{P}) \big\}. \end{aligned} \tag{2.1.36}$$

The aggregate relative entropy $H_{t,T}(\mathbb{Q}|\mathbb{P})$ *in equation (2.1.36) can be interpreted as a penalty for removing the financial risk (the volatility risk of the factor process which drives the interest rate). The right-hand side of equation (2.1.36) maximizes the excess (risk-adjusted) return on the investment, which is equal to the equivalent instantaneous return given by the left-hand side of equation (2.1.36). Note that when* $c = P(t,T)$ *the instantaneous return is equal to zero, which is the equilibrium state.*

In the next section we compare the optimal measure transformation problem with the optimal stochastic control problem proposed by Gombani and Runggaldier [35]. We find that there exists an equivalence relationship between these two approaches.

2.1.3 Equivalence between the OMT problem and the OSC problem

Gombani and Runggaldier [35] considered the bond pricing problem under the same general framework as we set up in previous section. To avoid confusion, we denote the factor process by $\tilde{X}_s$ in the context of Gombani and Runggaldier [35]. Additionally,

$\tilde{X}_s$ is assumed to be Markovian with $X_t = x$ so that the price of default-free bond, denoted by $P(t,T,x)$, at time t is given by

$$P(t,T,x) = E_{\mathbb{P}}[e^{-\int_t^T r_v dv}|\mathcal{F}_t] = E_{\mathbb{P}}[e^{-\int_t^T r_v dv}|X_t = x].$$

Assuming $P(t,T,x) \in \mathcal{C}^{1,2}$, a sufficient condition for the term structure induced by $P(t,T,x)$ to be arbitrage-free is that $P(t,T,x)$ satisfies the following partial differential equation (see Björk [15, Proposition 21.2])

$$\begin{cases} \frac{\partial}{\partial t}P(t,T,x) + f'(t,x)\nabla_x P(t,T,x) + \frac{1}{2}\mathrm{tr}\big(g'(t,x)\nabla_{xx}P(t,T,x)g(t,x)\big) \\ \qquad\qquad\qquad\qquad\qquad\qquad - P(t,T,x)r(t,x) = 0 \\ P(T,T,x) = 1. \end{cases} \tag{2.1.37}$$

Gombani and Runggaldier [35] transform equation (2.1.37) to an equivalent HJB equation which corresponds to the following optimal stochastic control (OSC) problem.

Problem 2.1.2. *On a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_s, 0 \le s \le T\}, \mathbb{P})$, *with a Markovian process* $\tilde{X}_s$ *given by*

$$d\tilde{X}_s = \big[f(s,\tilde{X}_s) + g(s,\tilde{X}_s)u_s'\big]ds + g(s,\tilde{X}_s)dW_s^{\mathbb{P}}. \tag{2.1.38}$$

Let $\mathcal{U}$ *be the admissible control set, then for any control* $u \in \mathcal{U}$ *and* $t \in [0,T]$, *consider a performance criterion* $\tilde{J}_{t,T}(u)$ *of the form*

$$\tilde{J}_{t,T}(u) = E_{\mathbb{P}}^{t,x}\big[\int_t^T \big(\frac{1}{2}u_v u_v' + r(\tilde{X}_v)\big)dv\big],$$

where $E_{\mathbb{P}}^{t,x}$ *denotes the conditional expectation given* $\tilde{X}_t = x$. *The optimal control problem is*

$$U_{t,T} = \inf_{u\in\mathcal{U}} \tilde{J}_{t,T}(u).$$

Gombani and Runggaldier [35] established a connection between the price of default-free bonds and the OSC Problem 2.1.2 by showing that

$$P(t,T,x) = e^{-U_{t,T}(x)}.$$

We next explore an equivalence relationship between the OMT problem and the OSC problem. For any $\mathbb{Q} \in \mathcal{P}_t(\Omega)$, the Radon-Nikodym derivative process is of the following form

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_s} = \begin{cases} 1, & 0 \leq s \leq t. \\ \Lambda_s, & t < s \leq T. \end{cases}$$

where Λ_s is an $(\mathcal{F}, \mathbb{P})$-martingale from t to T. Since Λ_s is positive almost surely, by the martingale representation theorem, there exists an $\mathcal{F}$-predictable $(1 \times n)$-vector process u such that

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_s} = e^{-\int_t^s \frac{1}{2} u_v u_v' dv + \int_t^s u_v dW_v^{\mathbb{P}}}, \quad t < s \leq T \tag{2.1.39}$$

where u is an $\mathcal{F}$-predictable $(1 \times n)$-vector process. In the remaining part of this section we denote by $\mathbb{Q}^u$ the probability measure associated with the density process in equation (2.1.39). Then, by Girsanov's theorem, the process $W^{\mathbb{Q}^u}$ defined as

$$W_s^{\mathbb{Q}^u} = W_s^{\mathbb{P}} - \int_t^s u_v' dv, \quad t < s \leq T$$

is a Brownian motion under $\mathbb{Q}^u$. Then we calculate the relative entropy of $\mathbb{Q}^u$ with respect to $\mathbb{P}$ explicitly in terms of u as follows

$$\begin{aligned} H_{t,T}(\mathbb{Q}^u|\mathbb{P}) &= E_{\mathbb{Q}^u}[\ln(\frac{d\mathbb{Q}^u}{d\mathbb{P}})|\mathcal{F}_t] \\ &= E_{\mathbb{Q}^u}[\Big(-\int_t^T \frac{1}{2} u_v u_v' dv + \int_t^T u_v dW_v^{\mathbb{P}}\Big)|\mathcal{F}_t] \\ &= E_{\mathbb{Q}^u}[\Big(\int_t^T \frac{1}{2} u_v u_v' dv + \int_t^T Z_v dW_v^{\mathbb{Q}^u}\Big)|\mathcal{F}_t] \\ &= E_{\mathbb{Q}^u}[\int_t^T \frac{1}{2} u_v u_v' dv|\mathcal{F}_t]. \end{aligned} \tag{2.1.40}$$

Substituting the explicit expression of the relative entropy in equation (2.1.40) into equation (2.1.13) we restate the OMT Problem 2.1.1 as follows

Problem 2.1.3. *On a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_s, 0 \leq s \leq T\}, \mathbb{P})$, suppose that the factor process $(X_s, 0 \leq s \leq T)$ is given by*

$$dX_s = f(s, X_s)ds + g(s, X_s)dW_s^{\mathbb{P}}. \tag{2.1.41}$$

Find the optimal measure $\mathbb{Q}^\star \in \mathcal{P}_t(\Omega)$ such that

$$V_{t,T} = J_{t,T}(\mathbb{Q}^\star) = \inf_{\mathbb{Q}^u \in \mathcal{P}_t(\Omega)} E_{\mathbb{Q}^u}\Big[\int_t^T \big(r(X_v) + \frac{1}{2}u_v u_v'\big)dv\Big]. \tag{2.1.42}$$

In the OSC Problem 2.1.2, the distribution of $\tilde{X}_s$ is changed by the control process u. In the OMT Problem 2.1.3, the distribution of X_s is subject to the measure transformation from $\mathbb{P}$ to $\mathbb{Q}^u$. Note that $\tilde{X}_s$ in equation (2.1.38) and X_s in equation (2.1.41) follow SDEs of the same form under different measures, in other words, the u controlled process $\tilde{X}_s$ has the same distribution under $\mathbb{P}$ as the process X_s does under $\mathbb{Q}^u$. Hence for each admissible control u in the OSC problem with performance functional $\tilde{J}_{t,T}(u)$, there exists a corresponding measure $\mathbb{Q}^u$ in the OMT problem with performance functional $J_{t,T}(\mathbb{Q}^u)$, and $\tilde{J}_{t,T}(u) = J_{t,T}(\mathbb{Q}^u)$. So the optimal control $u^\star$ also corresponds to the optimal measure transformation $\mathbb{Q}^\star = \mathbb{Q}^{u^\star}$. In that sense, the OSC problem is equivalent to the OMT problem.

Example 2.1.1. *Now we compare the OMT problem and the OSC problem under the framework of QTSMs with specifications*

(i) $f(s,x) = Ax + B$

(ii) $g(s,x) = \Sigma$

(iii) $r(x) = x'Qx + R'x + k$

where A is an $(n\times n)$-matrix of scalars, B and R are $(n\times 1)$-column vectors, Q and Σ are $n\times n$ symmetric positive semidefinite matrices, k is a scalar. The OSC Problem 2.1.2 becomes

$$\begin{cases} d\tilde{X}_s = \big(A\tilde{X}_s + B + \Sigma u_s'\big)ds + \Sigma dW_s^{\mathbb{P}}, \\ V_{t,T} = \inf_{u\in\mathcal{U}} \tilde{J}_{t,T}(u) = \inf_{u\in\mathcal{U}} E_{t,x}\big[\int_t^T \big(\tilde{X}_v'Q\tilde{X}_v + R'\tilde{X}_v + k + \frac{1}{2}u_v u_v'\big)dv\big]. \end{cases} \tag{2.1.43}$$

The OSC Problem 2.1.43 is actually a linear-quadratic-Gaussian (LQG) control problem, whose optimal control $u_s^\star$ is of feedback form (see Gombani and Runggaldier [35, Proposition 3.4])

$$u_s^\star = u^\star(s,\tilde{X}_s) = \Big(X_s'(q_s + q_s') + v_s\Big)\Sigma, \quad t \le s \le T \tag{2.1.44}$$

with the value function $W_{t,T}(x)$ given by

$$W_{t,T}(x) = x'q_t x + v_t x + p_t, \tag{2.1.45}$$

where q_s, v_s, p_s satisfy the following ODE system

$$\begin{cases} \dot{q}_s + A'q_s + q_s A - 2q_s\Sigma\Sigma' q_s + Q = 0 \\ \dot{v}_s + v_s A + 2B'q_s' - 2v_s\Sigma\Sigma' q_s' + R = 0 \\ \dot{p}_s + v_s B + tr(\Sigma' q_s \Sigma) - \frac{1}{2}v\Sigma\Sigma' v_s' + k = 0 \\ q_T = 0, \quad v_T = 0, \quad p_T = 0. \end{cases} \tag{2.1.46}$$

Under the framework of QTSMs the OMT Problem 2.1.1 is specified as

$$\begin{cases} dX_s = \big(AX_s + B\big)ds + \Sigma dW_s^{\mathbb{P}}, \\ V_{t,T} = \inf\limits_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} E_{\mathbb{Q}}[\int_t^T \big(X_v'QX_v + R'X_v + k\big)dv|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P}). \end{cases} \tag{2.1.47}$$

From Section 2.1.2, we know the OMT Problem (2.1.47) is completely characterized via the related FBSDE

$$X_s = X_t + \int_t^s (AX_v + B + \Sigma Z_v')\, dv + \int_t^s \Sigma dW_v^{\mathbb{P}} \tag{2.1.48}$$

$$Y_s = \int_s^T (X_v'QX_v + R'X_v + k - \frac{1}{2}Z_v'Z_v)dv + \int_s^T Z_v dW_v^{\mathbb{P}}. \tag{2.1.49}$$

The value function is given by

$$V_{t,T} = Y_t, \tag{2.1.50}$$

and the optimal measure transformation is determined by

$$\left.\frac{d\mathbb{Q}^\star}{d\mathbb{P}}\right|_{\mathcal{F}_T} = e^{-\int_t^T \frac{1}{2}Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}}}. \tag{2.1.51}$$

Hyndman and Zhou [38] proved that the FBSDE (2.1.48)-(2.1.49) admits a unique solution (X, Y, Z), and (Y, Z) has explicit expressions in terms of X

$$Y_s = X_s'q_s X_s + v_s X_s + p_s,$$
$$Z_s = \Big(X_s'(q_s + q_s') + v_s\Big)\Sigma,$$

where q_s, v_s, p_s satisfy the same ODE system (2.1.46). Not surprisingly, the Girsanov kernel Z_s for the transition from $\mathbb{P}$ to $\mathbb{Q}^\star$ is the same as the optimal control $u^\star$, i.e. $Z_s = u_s^\star$, and they give the same value function $V_{t,T} = W_{t,T}$.

In next section we consider the OMT problems for futures and forward prices.

2.2 Futures and forward prices

Suppose the factor process X_s given by equation (2.1.1) drives not only the short rate but also a risky asset price. We assume that the risky asset price is a function of factors, $S_s = S(s, X_s)$, for some function $S(\cdot,\cdot) : [0,\infty) \times \mathbf{R}^n \to (0,\infty)$. For instance, $S(\cdot,\cdot)$ can be specified by

$$S(s,x) = e^{A_s' x + h_s},$$

which we refer to as an affine price model (APM), or

$$S(s,x) = e^{x' B_s x + A_s' x + h_s}$$

which we refer to as a quadratic price model (QPM), where $B_s : [0,T] \to \mathbf{R}^{n\times n}$, $A_s : [0,T] \to \mathbf{R}^n$, $h_s : [0,T] \to \mathbf{R}$.

We next consider futures and forward contracts on the risky asset S and associate the futures prices and forward prices with OMT problems.

2.2.1 Futures prices

The futures price of the risky asset S is given by

$$G(t,T) = E_{\mathbb{P}}[S(T, X_T)|\mathcal{F}_t], \tag{2.2.1}$$

at time t for maturity T. Similar to the derivation of the OMT Problem in Section 2.1, we let $\varphi = \ln S(T, X_T)$ and associate the futures price with the following OMT problem

$$\begin{cases} dX_s = f(s,X_s)ds + g(s,X_s)dW_s^{\mathbb{P}} \\ V_{t,T}^G = \inf\limits_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \big\{E_{\mathbb{Q}}[-\ln S(T,X_T)|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P})\big\}. \end{cases} \tag{2.2.2}$$

By Proposition 2.1.1 the solution of the OMT Problem (2.2.2) is given by the optimal measure $\mathbb{Q}^{G^\star}$, that is determined by

$$\left.\frac{d\mathbb{Q}^{G^\star}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \frac{S(T,X_T)}{E_{\mathbb{P}}[S(T,X_T)|\mathcal{F}_t]}, \tag{2.2.3}$$

and the optimal value function is given by

$$V_{t,T}^{G} = -\ln\{E_{\mathbb{P}}[S(T, X_T)|\mathcal{F}_t]\}. \tag{2.2.4}$$

Equation (2.2.4) connects the OMT Problem (2.2.2) with the futures price as

$$V_{t,T}^{G} = -\ln G(t, T)$$

and this relationship allows us to give the following financial interpretation similar to Remark 2.1.5 for the bond price.

Remark 2.2.1. *Suppose a financial agent enters a long position in a futures contract on the risky asset S at time t with a futures price c dollars. From time t to settlement date T the futures contract is marked to market daily where the cash flows are transfered through the margin account and the total transfers are $S(T, X_T) - c$. Since the marking to market mechanism eliminates the default risk, and also by the assumption that there is no interest paid on the margin account, the gain from a long position in the futures contract held from time t and closed at time T can be characterized by the quantity*

$$\gamma_{t,T} = \ln \frac{S(T, X_T)}{c}.$$

Then the OMT Problem (2.2.2) is equivalent to

$$\ln \frac{G(t,T)}{c} = \sup_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \big\{E_{\mathbb{Q}}[\gamma_{t,T}|\mathcal{F}_t] - H_{t,T}(\mathbb{Q}|\mathbb{P})\big\}. \tag{2.2.5}$$

The right-hand side of equation (2.2.5) maximizes the expectation of $\gamma_{t,T}$ under $\mathbb{Q}^{G^*}$ with an entropy penalty term for removing the market risk of the futures contract caused by the volatility risk of underlying risky asset. Note that equation (2.2.5) attains the equilibrium state where the supremum is equal to zero if the pre-specified future price c is equal to the fair future price $G(t, T)$.

Similar to the procedure in Section 2.1, we characterize the OMT Problem 2.2.2 by the FBSDE

$$X_s = X_t + \int_t^s f(v, X_v)dv + \int_t^s g(v, X_v)dW_v^{\mathbb{P}}, \tag{2.2.6}$$

$$Y_s = -\ln[S(T, X_T)] - \int_s^T \frac{1}{2} Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}}. \tag{2.2.7}$$

If the above FBSDE admits a solution triple (X, Y, Z), then the value function and the optimal measure to the OMT Problem 2.1.1 have expressions

$$V_{t,T}^{G} = Y_t,$$
$$\left.\frac{d\mathbb{Q}^{G^\star}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = e^{-\int_t^T \frac{1}{2} Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}}}.$$

Hyndman [40] and Hyndman and Zhou [38] studied the the FBSDE (2.2.6)-(2.2.7) in the framework of ATSMs and QTSMs, respectively, and gave explicit solutions.

We next consider a forward contract on the risky asset.

2.2.2 Forward prices

The forward price of the risky asset S is given by

$$F(t, T) = \frac{E_{\mathbb{P}}[e^{-\int_t^T r(X_v)dv} S(T, X_T)|\mathcal{F}_t]}{P(t, T)}, \tag{2.2.8}$$

at time t for maturity T. To ensure that the forward price is not simply equal to the futures price we assume that the interest rate process is stochastic and the factors influencing the interest rate are not independent of the factors influencing the underlying asset price. Further, to preclude the case where the numerator of equation (2.2.8) reduces to the underlying asset price at time t we suppose that the asset pays a stochastic dividend or convenience yield.

Similar to the derivation of the OMT Problem in Section 2.1 we let

$$\varphi = (\ln S(T, X_T) - \int_t^T r_v dv)$$

and associate the forward price with the following OMT problem

$$\begin{cases} dX_s = f(s, X_s)ds + g(s, X_s)dW_s^{\mathbb{P}} \\ V_{t,T}^{F} = \inf_{\mathbb{Q} \in \mathcal{P}_t(\Omega)} \big\{ E_{\mathbb{Q}}\big[-\ln S(T, X_T) + \int_t^T r_v dv | \mathcal{F}_t \big] + H_{t,T}(\mathbb{Q}|\mathbb{P}) \big\}. \end{cases} \tag{2.2.9}$$

By Proposition 2.1.1 the solution to the OMT Problem (2.2.9) is given by the optimal measure $\mathbb{Q}^{F^\star}$, that is determined by

$$\left.\frac{d\mathbb{Q}^{F^\star}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \frac{S(T, X_T)e^{-\int_t^T r_v dv}}{E_{\mathbb{P}}[S(T, X_T)e^{-\int_t^T r_v dv}|\mathcal{F}_t]}, \tag{2.2.10}$$

and the optimal value function given by

$$V^F_{t,T} = -\ln\left(E_{\mathbb{P}}[e^{-\int_t^T r(X_v)dv}S(T,X_T)|\mathcal{F}_t]\right). \tag{2.2.11}$$

Equation (2.2.11) connects the OMT Problem (2.2.9) with the forward price as

$$V^F_{t,T} = -\ln\left(F(t,T)P(t,T)\right).$$

Therefore, we have the following financial interpretation of the OMT Problem (2.2.9).

Remark 2.2.2. *Suppose a financial agent enters into a forward agreement at time t on the risky asset S with forward price c dollars. At the settlement date T the agent pays c dollars and receives the underlying asset worth $S(T,X_T)$. The logarithmic return over the period $[t,T]$ is*

$$\gamma_{t,T} = \ln\frac{S(T,X_T)}{c}.$$

The excess return over the risk-free rate, $\tilde{\gamma}_{t,T}$, is given by

$$\tilde{\gamma}_{t,T} = \gamma_{t,T} - \int_t^T r(X_v)dv.$$

Then the OMT Problem (2.2.9) is equivalent to

$$\ln\frac{F(t,T)P(t,T)}{c} = \sup_{\mathbb{Q}\in\mathcal{P}_t(\Omega)}\left\{E_{\mathbb{Q}}[\tilde{\gamma}_{t,T}|\mathcal{F}_t] - H_{t,T}(\mathbb{Q}|\mathbb{P})\right\}. \tag{2.2.12}$$

Similar to the financial interpretation of the OMT Problem for the bond and futures contract the right hand side of equation (2.2.12) maximizes the excess return $\tilde{\gamma}_{t,T}$ under $\mathbb{Q}^{F^*}$ with an entropy penalty term for removing the market risk of the value of the forward commitment due to the volatility risk of the factor process that determines both the interest rate and underlying asset volatilities. Note that equation (2.2.12) attains the equilibrium state where the supremum is equal to zero if the pre-specified forward price c is equal to the present value of the fair future price $F(t,T)P(t,T)$.

Similar to the procedure in Section 2.1, we characterize the OMT Problem 2.2.2 by the FBSDE

$$X_s = X_t + \int_t^s f(v,X_v)dv + \int_t^s g(v,X_v)dW^{\mathbb{P}}_v \tag{2.2.13}$$

$$Y_s = -\ln[S(T,X_T)] + \int_s^T [r(X_v) - \frac{1}{2}Z_vZ'_v]dv + \int_s^T Z_v dW^{\mathbb{P}}_v \tag{2.2.14}$$

If the above FBSDE admits a solution triple (X, Y, Z), then the value function and the optimal measure to the OMT Problem 2.1.1 have expressions

$$V^F_{t,T} = Y_t,$$
$$\left.\frac{d\mathbb{Q}^{F^\star}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = e^{-\int_t^T \frac{1}{2} Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}}}.$$

Hyndman [40] and Hyndman and Zhou [38] also studied the the FBSDE (2.2.13)-(2.2.14) in the framework of ATSMs and QTSMs, respectively, and gave explicit solutions.

The OMT approach seems to be more flexible with respect to the dynamics of the factors process than the OSC approach. In next section we extend the OMT approach to include jumps in the factors which would be difficult to incorporate using the OSC approach.

2.3 Models with jumps

In order to model sudden and unexpected jumps of the driving factor process, we add a jump component to the factor process X_t as follows

$$dX_s = f(s, X_{s-})ds + g(s, X_{s-})dW_s^{\mathbb{P}} + \int_{\mathbf{R}^n} z\tilde{N}^{\mathbb{P}}(ds, dz) \tag{2.3.1}$$

where $\tilde{N}^{\mathbb{P}}(\cdot, \cdot)$ is an $\mathbf{R}^n$-valued compensated random measure (refer to Delong [23, Section 2.1]) with the compensator

$$\eta(ds, dz) = v(dz)\lambda(X_{s-})ds$$

where $v(\cdot)$ is a measure on $\mathbf{R}^n$ and $\lambda(\cdot)$ is a function to be specified from $\mathbf{R}^n$ to $\mathbf{R}$.

Following a similar procedure as in Section 2.1, we associate the bond price with the following OMT Problem with jumps.

$$\begin{cases} dX_s = f(s, X_{s-})ds + g(s, X_{s-})dW_s^{\mathbb{P}} + \displaystyle\int_{\mathbf{R}^n} z\tilde{N}^{\mathbb{P}}(ds, dz), \\ V_{t,T} = \displaystyle\inf_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} E_{\mathbb{Q}}[\int_t^T r(X_s)ds|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P}). \end{cases} \tag{2.3.2}$$

Then the OMT Problem (2.3.2) is completely characterized by the following FBSDE with jumps

$$X_s = X_t + \int_t^s f(v, X_{v-})ds + \int_t^s g(v, X_{v-})dW_v^{\mathbb{P}} + \int_t^s \int_{\mathbf{R}^n} z\tilde{N}^{\mathbb{P}}(dv, dz), \tag{2.3.3}$$

$$\begin{aligned} Y_s = & \int_s^T \Big\{ r(X_{v-}) - \lambda(X_{v-}) \int_{\mathbf{R}^n} e^{G(v,z)} v(dz) - \frac{1}{2} Z_v Z_v' \Big\} dv \\ & + \int_s^T Z_v dW_v^{\mathbb{P}} + \int_s^T \int_{\mathbf{R}^n} G(v, z)\tilde{N}^{\mathbb{P}}(dv, dz). \end{aligned} \tag{2.3.4}$$

If the FBSDE (2.3.3)-(2.3.4) admits a solution (X, Y, Z, G), then the value function and the optimal measure $\mathbb{Q}^\star$ for the OMT Problem (2.3.2) are characterized by (Y, Z, G) respectively as

$$V_{t,T} = Y_t, \tag{2.3.5}$$

$$\begin{aligned} \frac{d\mathbb{Q}^\star}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = & \exp \Big\{ - \int_t^T \frac{1}{2} Z_v Z_v' dv + \int_t^T Z_v dW_v^{\mathbb{P}} - \int_t^T \int_{\mathbf{R}^n} \lambda(X_{v-}) e^{G(v,z)} v(dz)dv \\ & + \int_t^T \int_{\mathbf{R}^n} G(v, z)\tilde{N}^{\mathbb{P}}(dv, dz) \Big\}. \end{aligned} \tag{2.3.6}$$

The aggregate relative entropy of $\mathbb{Q}^\star$ with respect to $\mathbb{P}$ is

$$\begin{aligned} & H_{t,T}(\mathbb{Q}^\star|\mathbb{P}) \\ = & E_{\mathbb{Q}^\star}[\ln(\frac{d\mathbb{Q}^\star}{d\mathbb{P}})|\mathcal{F}_t] \\ = & E_{\mathbb{Q}^\star}[\Big(- \int_t^T \frac{1}{2} Z_s Z_s' ds + \int_t^T Z_s dW_s^{\mathbb{P}} - \int_t^T \int_{\mathbf{R}^n} \lambda(X_{s-}) e^{G(s,z)} v(dz)ds \\ & + \int_t^T \int_{\mathbf{R}^n} G(s, z)\tilde{N}^{\mathbb{P}}(ds, dz) \Big) |\mathcal{F}_t] \\ = & E_{\mathbb{Q}^\star}[\Big(\int_t^T \frac{1}{2} Z_s Z_s' ds + \int_t^T \int_{\mathbf{R}^n} \lambda(X_{s-})(G(s, z)e^{G(s,z)} - e^{G(s,z)})v(dz)ds \\ & + \int_t^T Z_s dW_s^{\mathbb{Q}^\star} + \int_t^T \int_{\mathbf{R}^n} G(s, z)\tilde{N}^{\mathbb{Q}^\star}(ds, dz) \Big) |\mathcal{F}_t] \\ = & E_{\mathbb{Q}^\star}[\int_t^T \big(\frac{1}{2} Z_s Z_s' + \lambda(X_{s-}) \int_{\mathbf{R}^n} (G(s, z)e^{G(s,z)} - e^{G(s,z)})v(dz) \big) ds|\mathcal{F}_t]. \end{aligned}$$

In the following two subsections we give explicit solutions to the FBSDE (2.3.3)-(2.3.4) under ATSMs and QTSMs, respectively, with jumps.

2.3.1 ATSMs with jumps

In the framework of ATSMs with jumps, we make the following specifications on the coefficients of FBSDE (2.3.3)-(2.3.4) as follows

(i) $f(s,x) = Ax + B$

(ii) $g(s,x) = S\text{diag}\sqrt{\alpha_i + \beta_i x}$

(iii) $r(x) = R'x + k$

(iv) $\lambda(x) = L'x + l$

where A is an $(n \times n)$-matrix of scalars, B, R and L are $(n \times 1)$-vectors, for each $i \in \{1, \ldots, n\}$ the α_i are scalars, for each $i \in \{1, \ldots, n\}$ the $\beta_i = (\beta_{i1}, \ldots, \beta_{in})$ are $(1 \times n)$-vectors, S is a non-singular $(n \times n)$-matrix, k and l are scalars.

Remark 2.3.1. *As in Duffie et al. [26] the intensity process is assumed to be an affine function of the factors to preserve the affine term structure.*

The FBSDE (2.3.3)-(2.3.4) becomes

$$X_s = X_t + \int_t^s (AX_{v-} + B)\,dv + \int_t^s S\text{diag}\sqrt{\alpha_i + \beta_i X_{v-}}dW_v^{\mathbb{P}} + \int_t^s \int_{\mathbf{R}^n} z\tilde{N}^{\mathbb{P}}(dv, dz), \tag{2.3.7}$$

$$\begin{aligned} Y_s = \int_s^T & \Big\{ R'X_{v-} + k - (L'X_{v-} + l) \int_{\mathbf{R}^n} e^{G(v,z)} v(dz) - \frac{1}{2} Z_v Z_v' \Big\} dv \\ & + \int_s^T Z_v dW_v^{\mathbb{P}} + \int_s^T \int_{\mathbf{R}^n} G(v,z) \tilde{N}^{\mathbb{P}}(dv, dz). \end{aligned} \tag{2.3.8}$$

We will give the explicit solution to FBSDE (2.3.7)-(2.3.8) by applying a similar technique to Hyndman [40] which extends the approach for linear FBSDEs from Ma and Yong [57]. In the statement of the following proposition, as in Hyndman [40], we shall adopt the notation of Björk and Landén [16] to write

$$S\text{diag}(\alpha_i + \beta_i x)S' = k_0 + \sum_{j=1}^n k_j x_j$$

for symmetric $(n \times n)$ matrices k_j, where x_j is the jth element of a vector $x \in D$. Define the $(n^2 \times n)$ matrix K and, given a $(1 \times n)$ row vector $\underline{\mathbf{y}}$, the $n \times n^2$ matrix $\beta(\underline{\mathbf{y}})$ by

$$K = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} \quad \text{and} \quad \beta(\underline{y}) = \begin{bmatrix} \underline{y} & 0_{1\times n} & \cdots & 0_{1\times n} \\ 0_{1\times n} & \underline{y} & & \\ \vdots & & \ddots & \vdots \\ 0_{1\times n} & \cdots & & \underline{y} \end{bmatrix}$$

respectively.

Theorem 2.3.1. *If the Riccati equation*

$$\dot{U}_s + U_s A + \frac{1}{2} U_s K'[\beta(U_s)] + \Big[\int_{\mathbf{R}^n} e^{U_s z} v(dz) \Big] L' - R' = 0, \quad t \in [0,T] \tag{2.3.9}$$

$$U_T = 0 \tag{2.3.10}$$

admits a unique bounded solution $U(\cdot)$ over the interval $[0,T]$, then the FBSDE (2.3.7)-(2.3.8) admits a unique solution and (Y, Z, G) has explicit expression in terms of X as follows

$$Y_s = -(U_s X_s + p_s), \tag{2.3.11}$$

$$Z_s = U_s S\, diag(\sqrt{\alpha_i + \beta_i X_{s-}}), \quad \textit{and} \tag{2.3.12}$$

$$G(s,z) = U_s z, \tag{2.3.13}$$

where p_s is given by

$$p_s = -\int_s^T \left(k - l \int_{\mathbf{R}^n} e^{U_v z} v(dz) - \frac{1}{2} U_v k_0 U_v' - U_v B \right) dv \tag{2.3.14}$$

Proof. We first prove the decoupled FBSDE (2.3.7)-(2.3.8) admits a unique solution (X, Y, Z, G). The SDE (2.3.7) admits a unique solution. As X_s is known, we consider the single BSDE (2.3.8). If we let

$$\tilde{Y}_s = e^{-Y_s},$$

$$\tilde{Z}_s = -\tilde{Y}_s \cdot Z_s,$$

$$\tilde{G}(z,s) = -\tilde{Y}_s e^G(s,z)$$

BSDE (2.3.8) becomes

$$\tilde{Y}_s = 1 + \int_s^T \left\{ \left(\int_{\mathbf{R}^n} e^{G(v,z)} v(dz) L' - R' \right) X_{v-} - \left(\int_{\mathbf{R}^n} e^{G(v,z)} v(dz) \right) l \right.$$
$$\left. + k \right\} \tilde{Y}_v dv + \int_s^T \tilde{Z}_v dW_v^{\mathbb{P}} + \int_s^T \int_{\mathbf{R}^n} \tilde{G}(v,z) \tilde{N}^{\mathbb{P}}(dv, dz). \tag{2.3.15}$$

By Delong [23, Theorem 3.1.1], we know that the BSDE (2.3.15) admits a unique solution $(\tilde{Y}, \tilde{Z}, \tilde{G})$. Therefore, the FBSDE (2.3.7)-(2.3.8) admits a unique solution (X, Y, Z, G).

To prove the explicit expression of (Y, Z, G), we need to show that (Y, Z, G) given by equations (2.3.11)-(2.3.13) satisfies the BSDE (2.3.8). Apply Itô's formula to the function $\phi(s, x) = -(U_s x + p_s)$, where U_s is the solution to the Riccati equation (2.3.9) and p_s satisfies equation (2.3.14). Let $Y_s = \phi(s, X_s)$, where X_s is given by equation (2.3.7), then we have

$$\begin{aligned}
& Y_T - Y_s \\
= & -\int_s^T \left(\dot{U}_v X_{v-} + U_v(AX_{v-} + B) + k_0 U_v' + K'[\beta(U_v)]' X_{v-} \right) dv \\
& - \int_s^T U_v S \mathrm{diag}(\sqrt{\alpha_i + \beta_i X_v}) dW_v^{\mathbb{P}} - \int_s^T \left(k - l \int_{\mathbf{R}^n} e^{U_v z} v(dz) \right. \\
& \left. - \frac{1}{2} U_v k_0 U_v' - U_v B \right) dv - \int_s^T \int_{\mathbf{R}^n} U_v z \tilde{N}^{\mathbb{P}}(dv, dz) \\
= & -\int_s^T \left\{ \left(\dot{U}_v + U_v A + \frac{1}{2} U_v K'[\beta(U_v)] + \left(\int_{\mathbf{R}^n} e^{U_v z} v(dz) \right) L' - R' \right) X_v \right. \\
& \left. + \left(R' X_{v-} + k + \frac{1}{2} \left(U_v K'[\beta(U_v)]' X_{v-} + U_v k_0 U_v' \right) \right) \right\} dv \\
& + \int_s^T \left(\left[\int_{\mathbf{R}^n} e^{U_v z} v(dz) \right] (L' X_{v-} + l) \right) dv \\
& - \left\{ U_v S \mathrm{diag}(\sqrt{\alpha_i + \beta_i X_{v-}}) \right\} dW_v^{\mathbb{P}} - \int_s^T \int_{\mathbf{R}^n} U_v z \tilde{N}^{\mathbb{P}}(dv, dz)
\end{aligned} \tag{2.3.16}$$

Substituting equations (2.3.11)-(2.3.13) into equation (2.3.16) we have

$$\begin{aligned}
Y_s = & Y_T + \int_s^T (R' X_{v-} + k - \frac{1}{2} Z_v Z_v') dv - \int_s^T \int_{\mathbf{R}^n} (L' X_{v-} + l) e^{G(v,z)} v(dz) dv \\
& + \int_s^T Z_v dW_v^{\mathbb{P}} + \int_s^T \int_{\mathbf{R}^n} G(s,z) \tilde{N}^{\mathbb{P}}(ds, dz)
\end{aligned}$$

By the boundary condition of (2.3.10) and (2.3.14) we have

$$Y_T = -(U_T X_T + p_T) = 0.$$

Therefore,

$$\begin{aligned} Y_s = &\int_s^T (R' X_{v-} + k - \frac{1}{2} Z_v Z_v') dv - \int_s^T \int_{\mathbf{R}^n} (L' X_{v-} + l) e^{G(v,z)} v(dz) dv \\ &+ \int_s^T Z_v dW_v^{\mathbb{P}} + \int_s^T \int_{\mathbf{R}^n} G(v,z) \tilde{N}^{\mathbb{P}}(dv, dz) \end{aligned}$$

Hence (Y, Z, G) given by equations (2.3.11)-(2.3.13) satisfy BSDE (2.3.8). □

Remark 2.3.2. *The complete discussion on the Riccati equation of the form as in* (2.3.9) *can be found in Duffie et al. [27, Section 6].*

2.3.2 QTSMs with jumps

In the framework of QTSMs with jumps, we make the following specifications

(i) $f(s, x) = Ax + B$

(ii) $g(s, x) = \Sigma$

(iii) $r(x) = x' Q x + R' x + k$

(iv) $\lambda(x) = x' L_2 x + L_1' x + l$

where A is an $(n \times n)$-matrix of scalars, B, R and L_1 are $(n \times 1)$-column vectors, Q, Σ and L_2 are $n \times n$ symmetric positive semidefinite matrices, k and l are scalars. Then the FBSDE (2.3.3)-(2.3.4) becomes

$$X_s = X_t + \int_t^s \big(A X_{v-} + B\big) dv + \int_t^s \Sigma dW_v^{\mathbb{P}} + \int_t^s \int_{\mathbf{R}^n} z \tilde{N}^{\mathbb{P}}(dv, dz) \tag{2.3.17}$$

$$\begin{aligned} Y_s = &\int_s^T (X_{v-}' Q X_{v-} + R' X_{v-} + k - \frac{1}{2} Z_v Z_v') dv + \int_s^T Z_v dW_v^{\mathbb{P}} - \int_s^T \int_{\mathbf{R}^n} (X_{v-}' L_2 X_{v-} \\ &+ L_1' X_{v-} + L_0) e^{G(v,z)} v(dz) dv + \int_s^T \int_{\mathbf{R}^n} G(v,z) \tilde{N}^{\mathbb{P}}(dv, dz). \end{aligned} \tag{2.3.18}$$

Similar to the result in ATSMs with jumps we obtain the following explicit solution of the FBSDE (2.3.17)-(2.3.18).

Theorem 2.3.2. *If the Riccati equation*

$$\dot{q}_s + q_s A + A' q_s + \frac{(q_s' + q_s)\Sigma\Sigma'(q_s' + q_s)}{2} + \Big[\int_{\mathbf{R}^n} (e^{z' q_s z + u_s z}) v(dz)\Big] L_2' - Q = 0_{n\times n},$$

$$\dot{u}_s + u_s A + B'(q_s' + q_s) + u_s \Sigma\Sigma'(q_s' + q_s) + \Big[\int_{\mathbf{R}^n} (e^{z' q_s z + u_s z}) v(dz)\Big] L_1' - R' = 0_{1\times n},$$

$$q_T = 0, \quad u_T = 0$$

admits unique bounded solutions $q(\cdot)$, $u(\cdot)$ over the interval $[0, T]$, then the FBSDE (2.3.17)-(2.3.18) *admits a unique solution and (Y, Z, G) has explicit expression in terms of X as follows*

$$\begin{aligned} Y_s &= -(X_s' q_s X_s + u_s X_s + p_s), \\ Z_s &= \Big(X_{t-}'(q_s + q_s') + u_s\Big)\Sigma, \quad \textit{and} \\ G(s,z) &= z' q_s z + u_s z, \end{aligned}$$

where p_s is given by

$$p_s = -\int_s^T \left(k - L_0\left(\int_{\mathbf{R}^n} (e^{z' q_v z + u_v z}) v(dz)\right) - u_v B - \frac{1}{2} tr\Big((q_v + q_v')\Sigma\Sigma'\Big) \right.$$
$$\left. - \frac{1}{2} u_v \Sigma\Sigma' u_v' \right) dv.$$

We omit the proof of Theorem 2.3.2 as it is similar to the proof of Theorem 2.3.1. In the special case where $\lambda(x) \equiv 0$, we discuss the solvability of the Riccati equation in Appendix B.1.

In next section we further discuss the OMT problem associated with defaultable bond price and give a partially explicit solution for the related FBSDE with random terminal condition.

2.4 Defaultable bonds

We consider a defaultable zero coupon bond with the promised payoff of \$1 at maturity, and denote the price at time $t \in [0, T]$ by $D(t, T)$. Unlike default-free bonds, the issuer of defaultable bonds, such as corporate bonds, may default before the maturity in which case the bondholders will not receive the promised payment in full but a recovery payment. There are different recovery schemes if default occurs before the bond's maturity according to the timing and the amount of recovery payment (see Bielecki and Rutkowski [13, Section 1.1.1] and Altman et al. [5]). For instance, if a fixed fraction of the bond's face value is paid to the bondholder at maturity T in case of default, then the bond has the random payoff at maturity

$$C_T = \mathbf{1}_{\{\tau > T\}} + \eta \mathbf{1}_{\{\tau \leq T\}}$$

where τ is the default time. If a fixed fraction of the pre-default market value of the bond value is paid at time of default, then the equivalent random payoff of the bond is

$$C_T = \mathbf{1}_{\{\tau > T\}} + \eta P(\tau-, T) e^{\int_\tau^T r_v dv} \mathbf{1}_{\{\tau \leq T\}}.$$

The time of default τ is also modelled differently. Under the structural credit risk models originating with Merton [58] the default of corporate bonds occurs when the value of the firm reaches a certain lower threshold. Reduced form credit risk models, such as Duffie and Singleton [25], assume that default is driven by an exogenous default process.

Since the recovery scheme is not our main concern in this paper, we will in general represent the equivalent payoff of defaultable bonds with a random payoff C_T, and assume that the price is given by (see Duffie and Singleton [25])

$$D(t, T) = E_{\mathbb{P}}[e^{-\int_t^T r(X_v) dv} C_T | \mathcal{F}_t], \tag{2.4.1}$$

where C_T is an $\mathcal{F}_T$-measurable random variable valued in $[0, 1]$.

In the extreme situation $C_T = 0$ of a complete default, in which the bondholders receive no recovery payment in the event of default, the bonds become worthless. In this paper, we exclude the occurrence of complete default by assuming $\mathbb{P}(C_T = 0) = 0$. We will later explain why we have to make this technical assumption. The other extreme case of default-free bonds is included in our model if we assume $\mathbb{P}(C_T = 1) = 1$.

Following the same ideas as in formulation of the OMT Problem for the default-free bond in Section 2.1 we let

$$\varphi = \big(- \int_t^T r(X_v)dv + \ln C_T \big)$$

in equation (2.1.7) so that the defaultable bond price $D(t,T)$ is characterized by the following OMT problem

$$\begin{cases} dX_s = f(s,X_s)ds + g(s,X_s)dW_s^{\mathbb{P}}, \\ V_{t,T} = \inf_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \big\{ E_{\mathbb{Q}}\big[\int_t^T r(X_v)dv - \ln C_T | \mathcal{F}_t \big] + H_{t,T}(\mathbb{Q}|\mathbb{P}) \big\}. \end{cases} \tag{2.4.2}$$

By Proposition 2.1.1 the solution to the OMT Problem (2.4.2) is given by the optimal measure $\mathbb{Q}^\star$ that is determined by

$$\frac{d\mathbb{Q}^\star}{d\mathbb{P}}\bigg|_{\mathcal{F}_T} = \frac{e^{-\int_t^T r(X_v)dv} C_T}{E_{\mathbb{P}}[e^{-\int_t^T r(X_v)dv} C_T | \mathcal{F}_t]} \tag{2.4.3}$$

and the optimal value function given by

$$V_{t,T} = -\ln \Big\{ E_{\mathbb{P}}[e^{-\int_t^T r_s ds} C_T | \mathcal{F}_t] \Big\}. \tag{2.4.4}$$

Equation (2.4.4) connects the value function and the defaultable bond price through

$$V_{t,T} = -\ln D(t,T).$$

On the other hand, we notice, from equation (2.4.3), that the optimal measure which solves the OMT Problem (2.4.2) is essentially the martingale measure using defaultable bond price $D(t,T)$ as numéraire. This is why we require $\mathbb{P}(C_T = 0) = 0$ to guarantee the defaultable bond price to be positive almost surely.

Remark 2.4.1. *Suppose a financial agent pays c to buy one unit of the bond at time t, and receives a payoff of C_T at maturity T. The internal logarithmic return on the investment over the time period $[t,T]$ is*

$$\gamma_{t,T} = \ln \frac{C_T}{c}.$$

The excess return over the risk-free rate, $\tilde{\gamma}$, is given by

$$\tilde{\gamma}_{t,T} = \gamma_{t,T} - \int_t^T r(X_v)dv,$$

which measures the investment performance. Note that the OMT Problem (2.4.2) is equivalent to

$$\ln \frac{D(t,T)}{c} = -\inf_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \left\{ E_{\mathbb{Q}}[-\tilde{\gamma}_{t,T}|\mathcal{F}_t] + H_{t,T}(\mathbb{Q}|\mathbb{P}) \right\}$$
$$= \sup_{\mathbb{Q}\in\mathcal{P}_t(\Omega)} \left\{ E_{\mathbb{Q}}[\tilde{\gamma}_{t,T}|\mathcal{F}_t] - H_{t,T}(\mathbb{Q}|\mathbb{P}) \right\}. \tag{2.4.5}$$

The aggregate relative entropy $H_{t,T}(\mathbb{Q}|\mathbb{P})$ in equation (2.4.5) can be interpreted as penalty for removing financial risk composed of market risk (volatility risk) and credit risk in the framework of our model. The right-hand side of equation (2.4.5) maximizes the excess (risk-adjusted) return on the investment, which is equal to the equivalent instantaneous return given by left-hand side of equation (2.4.5).

Similar to Section 2.1.2 we relate the OMT Problem 2.4.2 to a decoupled FBSDE with random terminal condition

$$X_s = X_t + \int_t^s f(v, X_v)dv + \int_t^s g(v, X_v)dW_v^{\mathbb{P}}, \tag{2.4.6}$$
$$Y_s = -\ln C_T + \int_s^T [r(X_v) - \frac{1}{2}Z_v Z_v']dv + \int_s^T Z_v dW_v^{\mathbb{P}}. \tag{2.4.7}$$

If the above decoupled FBSDE admits a solution triple (X, Y, Z), then the value function and the optimal measure $\mathbb{Q}^\star$ for the OMT Problem 2.4.2 are characterized by (Y, Z) respectively as

$$V_{t,T} = Y_t, \tag{2.4.8}$$
$$\left.\frac{d\mathbb{Q}^\star}{d\mathbb{P}}\right|_{\mathcal{F}_T} = e^{-\int_t^T \frac{1}{2} Z_s Z_s' ds + \int_t^T Z_s dW_s^{\mathbb{P}}}. \tag{2.4.9}$$

We will discuss the explicit solution to the FBSDE (2.4.6)-(2.4.7) in the case of ATSMs and QTSMs respectively. The possibility of default leads to solutions with an extra component compared to those considered previously

2.4.1 ATSMs

Under the framework of ATSMs the FBSDE (2.4.6)-(2.4.7) becomes

$$X_s = X_t + \int_t^s (AX_v + B)\, dv + \int_t^s S\mathrm{diag}\sqrt{\alpha_i + \beta_i X_v} dW_v^{\mathbb{P}} \tag{2.4.10}$$

$$Y_s = -\ln C_T + \int_s^T (R'X_v + k - \frac{1}{2} Z_v Z_v')dv + \int_s^T Z_v dW_v^{\mathbb{P}} \tag{2.4.11}$$

The following result can be seen as a generalization of Hyndman [40, Theorem 3.2] by incorporating a random terminal condition representing the recovery amount in the case of default.

Theorem 2.4.1. *If the Riccati equation*

$$\dot{U}_s + U_s A + \frac{1}{2} U_s K'[\beta(U_s)] - R' = 0, \quad s \in [0, T] \tag{2.4.12}$$

$$U_T = 0 \tag{2.4.13}$$

admits a unique bounded solution $U(\cdot) \in \mathbf{R}^n$ over the interval $[0, T]$, then FBSDE (2.4.10)-(2.4.11) *admits a unique solution and the solution (Y, Z) has explicit expression in terms of X*

$$Y_s = -(U_s X_s + p_s), \quad \textit{and} \tag{2.4.14}$$

$$Z_s = U_s S\, diag(\sqrt{\alpha_i + \beta_i X_s}) + z_s, \tag{2.4.15}$$

where (p_s, z_s) solves the following BSDE

$$p_s = -\ln C_T - \int_s^T \left(k - \frac{1}{2} U_v k_0 U_v' - U_v B + \frac{1}{2} z_v z_v' \right) dv - \int_s^T z_v dW_v^{\mathbb{P}}. \tag{2.4.16}$$

Proof. We first prove the decoupled FBSDE (2.4.10)-(2.4.11) admits a unique solution (X, Y, Z). Under our assumptions the SDE (2.4.10) admits a unique solution. Given X_s, we consider the BSDE (2.4.11). If we let

$$\tilde{Y}_s = e^{-Y_s},$$

$$\tilde{Z}_s = -\tilde{Y}_s \cdot Z_s$$

the BSDE (2.4.11) becomes

$$\tilde{Y}_t = C_T + \int_t^T \left[R'X_s + k \right] \tilde{Y}_s ds + \int_t^T \tilde{Z}_s dW_s^{\mathbb{P}}. \tag{2.4.17}$$

Clearly the BSDE (2.4.11) admits a unique solution $(\tilde{Y}, \tilde{Z})$ so the FBSDE (2.4.10)-(2.4.11) admits a unique solution (X, Y, Z). Using the same technique, we can also prove BSDE (2.4.16) admits a unique solution (p, z).

To prove the explicit representation of (Y, Z), we need to show (Y, Z) given by equations (2.4.14)-(2.4.15) satisfies the BSDE (2.4.11). Apply Itô's formula to the function $\phi(s, x, p) = -(U_s x + p)$ where U_s is the solution to (2.4.12). Let $Y_s = \phi(s, X_s, p_s)$ where X_s is given by (2.4.10) and p_s satisfies (2.4.16). Then we have

$$\begin{aligned} dY_s = & -\Big(\dot{U}_s X_s + U_s(AX_s + B + k_0 U_s' + K'[\beta(U_s)]' X_s + S\mathrm{diag}(\sqrt{\alpha_i + \beta_i X_s}) z_s'\Big) ds \\ & - U_s S \mathrm{diag}(\sqrt{\alpha_i + \beta_i X_s}) dW_s^{\mathbb{P}} - \Big(k - \frac{1}{2} U_s k_0 U_s' - U_s B + \frac{1}{2} z_s z_s'\Big) ds - z_s dW_s^{\mathbb{P}} \\ = & -\Big\{\Big(\dot{U}_s + U_s A + \frac{1}{2} U_s K'[\beta(U_s)] - R'\Big) X_s + (R' X_s + k \\ & + \frac{1}{2}\Big(U_s K'[\beta(U_s)]' X_s + U_s k_0 U_s' + 2U_s S \mathrm{diag}(\sqrt{\alpha_i + \beta_i X_s}) z_s' + z_s z_s'\Big)\Big)\Big\} ds \\ & - \Big\{U_s S \mathrm{diag}(\sqrt{\alpha_i + \beta_i X_s}) + z_s\Big\} dW_s^{\mathbb{P}}. \end{aligned} \tag{2.4.18}$$

Substituting equations (2.4.12) and (2.4.15) into equation (2.4.18) we have

$$dY_s = -(R' X_s + k - \frac{1}{2} Z_s Z_s') ds - Z_s dW_s^{\mathbb{P}}$$

Thus (Y_s, Z_s) defined by equations (2.4.14)-(2.4.15) satisfies

$$Y_s = Y_T + \int_s^T (R' X_v + k - \frac{1}{2} Z_v Z_v') dv + \int_s^T Z_v dW_v^{\mathbb{P}}$$

By the boundary conditions in equations (2.4.13) and (2.4.16) we have

$$Y_T = -\ln C_T$$

Therefore,

$$Y_s = -\ln C_T + \int_s^T (R' X_v + k - \frac{1}{2} Z_v Z_v') dv + \int_s^T Z_v dW_v^{\mathbb{P}}.$$

□

Remark 2.4.2. *The existence and uniqueness of the solution to the Riccati equation* (2.4.12) *is shown in Duffie et al. [27, Section 6] where a class of generalized Riccati equations has been considered.*

Note that the representation of (Y, Z) of the FBSDE (2.4.14)-(2.4.15) is not completely explicit, since the term z_t is to be determined by the quadratic BSDE (2.4.16). Fortunately we can convert the quadratic BSDE (2.4.16) into a linear BSDE by letting

$$\tilde{p}_t = e^{-p_t},$$

$$\tilde{z}_t = \tilde{p}_t \cdot z_t$$

then the BSDE (2.4.16) becomes

$$\tilde{p}_t = \frac{1}{C_T} + \int_t^T (k - \frac{1}{2} U_s k_0 U_s' - U_s B)\tilde{p}_s ds + \int_t^T \tilde{z}_s dW_s^{\mathbb{P}}. \tag{2.4.19}$$

In the excluded case that $P(C_T = 0) > 0$ then (2.4.19) would be a BSDE with singular terminal condition.

With further specification of C_T through a specific a default mechanism and recovery scheme the linear BSDE (2.4.19) can either be solved analytically or numerically. There is an extensive literature focused on the numerical solution schemes for BSDEs which we shall not discuss. Nevertheless, Theorem 2.4.1 simplifies the procedure to solve the coupled nonlinear FBSDE (2.4.10)-(2.4.11) to the solution of the Riccati equation (2.4.12) and the linear BSDE (2.4.19).

2.4.2 QTSMs

In the framework of QTSMs the FBSDE (2.4.6)-(2.4.7) becomes

$$X_s = X_t + \int_t^s \big(AX_v + B\big)dv + \int_t^s \Sigma dW_v^{\mathbb{P}} \tag{2.4.20}$$

$$Y_s = -\ln C_T + \int_s^T (X_v' Q X_v + R' X_v + k - \frac{1}{2} Z_v' Z_v)dv + \int_s^T Z_v dW_v^{\mathbb{P}}. \tag{2.4.21}$$

As in the case of ATSMs we obtain the partially explicit solutions to the FBSDE (2.4.20)-(2.4.21) stated in the following theorem.

Theorem 2.4.2. *If the Riccati equations*

$$\dot{q}_s + q_s A + A' q_s + \frac{1}{2}(q_s' + q_s)\Sigma\Sigma'(q_s' + q_s) - Q = 0_{n\times n}, \quad s \in [0, T] \tag{2.4.22}$$

$$\dot{u}_s + u_s A + B'(q_s' + q_s) + u_s \Sigma\Sigma'(q_s' + q_s) - R' = 0_{1\times n}, \quad s \in [0, T] \tag{2.4.23}$$

$$q_T = 0_{n\times n}, \quad u_T = 0_{1\times n} \tag{2.4.24}$$

admit unique bounded solutions $q(\cdot)$, $u(\cdot)$ *over the interval* $[0,T]$, *then the FBSDE* (2.4.20)-(2.4.21) *admits a unique solution and* (Y,Z) *has explicit expression in terms of* X *as follows*

$$Y_s = -(X_s' q_s X_s + u_t X_s + p_s), \tag{2.4.25}$$

$$Z_s = \Big(X_s'(q_s + q_s') + u_s\Big)\Sigma + z_s, \tag{2.4.26}$$

where (p_s, z_s) *solves the following BSDE*

$$p_s = -\ln C_T - \int_s^T \left(k - u_v B - \frac{1}{2}tr\left((q_v + q_v')\Sigma\Sigma'\right) - \frac{1}{2}u_v\Sigma\Sigma' u_v' + \frac{1}{2}z_v z_v'\right) dv \\ - \int_s^T z_v dW_v^{\mathbb{P}}. \tag{2.4.27}$$

By the same technique as in the ATSM case we make the change of variables

$$\tilde{p}_s = e^{-p_s},$$

$$\tilde{z}_s = \tilde{p}_s \cdot z_s$$

so that the BSDE (2.4.27) to obtain the linear BSDE

$$\tilde{p}_s = \frac{1}{C_T} + \int_s^T (k - u_v B - \frac{1}{2}\mathrm{tr}\big((q_v + q_v')\Sigma\Sigma'\big)\tilde{p}_v dv + \int_s^T \tilde{z}_v dW_v^{\mathbb{P}}. \tag{2.4.28}$$

The above BSDE is of the same form as BSDE (2.4.19), which can also be solved either analytically or numerically.

Remark 2.4.3. *The decoupled Riccati equations* (2.4.22)-(2.4.24) *are closely related to the LQ control problem. The existence and uniqueness of solutions to the Riccati equations* (2.4.22)-(2.4.24) *have been discussed in Hyndman and Zhou [38]. We provide a similar proof in the appendix based on the results of Gombani and Runggaldier [35].*

2.5 Numerical illustration

We consider a one dimensional factor process X satisfying

$$dX_t = (aX_t + b)dt + \sigma\sqrt{\alpha + \beta X_v}dW_t^{\mathbb{P}}.$$

The interest rate is given by

$$r(X_t) = RX_t + k.$$

We suppose the underlying company value V satisfies

$$V_t = V_0 \exp\{\int_0^t (r(X_v) - \frac{1}{2}\sigma_V^2)dv + \sigma_V W_t^{\mathbb{P}}\}.$$

Default is triggered if the value process V crosses below a certain level κV_0, i.e.

$$\tau := \inf\{t \geq 0, V_t \leq \kappa V_0\}. \tag{2.5.1}$$

Then the random payoff C_T is given by

$$C_T = \xi \cdot 1_{\tau \leq T} + 1_{\tau > T}$$

where ξ is the recovery rate in case of default.

The price of the defautable bond is given by

$$D(t,T) = E_{\mathbb{P}}[e^{-\int_t^T (RX_v + k)dv} \cdot C_T | \mathcal{F}_t].$$

The solution to the associated OMT problem is characterized by the FBSDE

$$X_t = X_0 + \int_0^t (aX_v + b)\, dv + \int_0^t \sigma\sqrt{\alpha + \beta X_v} dW_v^{\mathbb{P}} \tag{2.5.2}$$

$$Y_t = -\ln C_T + \int_t^T (RX_v + k - \frac{1}{2}Z_v^2)dv + \int_t^T Z_v dW_v^{\mathbb{P}}. \tag{2.5.3}$$

We have explicit expression for the solution to FBSDE (2.5.2)-(2.5.3)

$$Y_t = -(U_t X_t + p_t), \tag{2.5.4}$$

$$Z_t = \sigma U_t(\sqrt{\alpha + \beta X_t}) + q_t, \tag{2.5.5}$$

where U_s satisfies the Riccati equation

$$\dot{U}_t + aU_t + \frac{\beta}{2}\sigma^2 U_t^2 - R = 0, \quad t \in [0,T] \tag{2.5.6}$$

$$U_T = 0, \tag{2.5.7}$$

and (p,q) solves the BSDE

$$p_t = -\ln C_T - \int_t^T \left(k - \frac{\alpha}{2}\sigma^2 U_v^2 - bU_v - \frac{1}{2}q_v^2\right) dv - \int_t^T q_v dW_v^{\mathbb{P}}. \tag{2.5.8}$$

The defaultable bond price can be expressed as

$$D(t,T) = \exp\{-Y_t\}. \tag{2.5.9}$$

The aggregate relative entropy of the optimal measure $\mathbb{Q}^\star$ with respect to $\mathbb{P}$ is given by

$$H_{t,T}(\mathbb{Q}^\star|\mathbb{P}) = \mathbb{E}^{\mathbb{Q}^\star}[\int_t^T \frac{1}{2} Z_v^2 dv|\mathcal{F}_t].$$

We introduce the following proposition which gives explicit solution to a special type of quadratic BSDEs.

Proposition 2.5.1. *On a probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P}$, *consider the following BSDE*

$$y_t = \xi - \int_t^T (\frac{1}{2} z_s' z_s + g_s) ds - \int_t^T z_s dW_s^{\mathbb{P}},$$

where $(y_t, z_t) \in \mathbf{R} \times \mathbf{R}^n$, ξ *is real-valued* $\mathcal{F}_T$*-measurable random variable,* g_t *is real-valued* $\mathcal{F}_t$*-adapted process satisfying* $E_{\mathbb{P}}[\sup_{0 \leq t \leq T} |g_t|^2] < \infty$. *Then* y_t *can be expressed explicitly as*

$$y_t = -\ln\{E_{\mathbb{P}}[e^{-\xi}|\mathcal{F}_t]\} - \int_0^t g_s ds.$$

Proof. Make the exponential transformation $\tilde{y}_t = e^{-y_t}$, by Itô's formula $\tilde{y}_t$ satisfies

$$\tilde{y}_t = e^{-\xi} + \int_t^T g_s \tilde{y}_s ds + \int_t^T \tilde{y}_s z_s dW_s^{\mathbb{P}}.$$

Define the adjoint process

$$x_s = e^{\int_t^s g_u du}, \quad s \geq t.$$

Notice that $x_t = 1$, and apply Itô formula to $x_s \cdot \tilde{y}_s$ from t to T, to find

$$\begin{aligned} \tilde{y}_t &= x_T e^{-\xi} + \int_t^T \tilde{y}_s x_s z_s dW_s^{\mathbb{P}} \\ &= e^{\int_t^T g_s ds - \xi} + \int_t^T \tilde{y}_s e^{\int_t^s g_u du} z_s dW_s^{\mathbb{P}}. \end{aligned} \tag{2.5.10}$$

Take conditional expectation on $\mathcal{F}_t$ of both sides of (2.5.10), we obtain

$$\begin{aligned} \tilde{y}_t &= E_{\mathbb{P}^T}[e^{\int_t^T g_s ds - \xi}|\mathcal{F}_t] \\ &= e^{\int_t^T g_s ds} E_{\mathbb{P}}[e^{-\xi}|\mathcal{F}_t]. \end{aligned}$$

Finally we have

$$\begin{aligned} y_t &= -\ln \tilde{y}_t \\ &= -\ln\{E_{\mathbb{P}}[e^{-\xi}|\mathcal{F}_t]\} - \int_0^t g_s ds. \end{aligned}$$

$\square$

Remark 2.5.1. *The existence and uniqueness of the solution to general quadratic BSDEs was proven by Kobylanski [50]. Proposition* (2.5.1) *is only a special case in which we can give the explicit solution.*

Applying Proposition 2.5.1 to the BSDE (2.5.8), we may express p_t explicitly as

$$p_t = -\ln\{E_{\mathbb{P}}[\frac{1}{C_T}|\mathcal{F}_t]\} - \int_0^t \left(k - \frac{1}{2}U_v k_0 U_v' - U_v B\right) dv. \tag{2.5.11}$$

However, we do not have an explicit expression for the process q_t. Alternatively we can solve BSDE (2.5.8) numerically. Actually we can transform the quadratic BSDE (2.5.8) into an equivalent linear BSDE by letting

$$\tilde{p}_t = e^{-p_t}, \quad \tilde{q}_t = \tilde{p}_t \cdot q_t,$$

then the BSDE (2.5.8) becomes

$$\tilde{p}_t = \frac{1}{C_T} + \int_t^T (k - \frac{\alpha}{2}\sigma^2 U_s^2 - bU_s)\tilde{p}_s ds + \int_t^T \tilde{q}_s dW_s. \tag{2.5.12}$$

We approximate the BSDE (2.5.12) by the following discretized BSDE

$$\begin{aligned} \tilde{p}_{t_{m+1}} &= \tilde{p}_{t_m} - (k - \frac{\alpha}{2}\sigma^2 U_{t_m}^2 - bU_{t_m})\tilde{p}_{t_m}\Delta t - \tilde{q}_{t_m}\Delta W_{t_m}^{\mathbb{P}}, \quad t_0 \le t_m \le t_M, \\ \tilde{p}_{t_M} &= \frac{1}{C_T}. \end{aligned}$$

The discretized BSDE can be solved using the following recursive scheme (see [33])

$$\begin{aligned} \tilde{q}_{t_m} &= \frac{1}{\Delta t}\mathbb{E}[p_{t_{m+1}}\Delta W_{t_m}^{\mathbb{P}}|\mathcal{F}_{t_m}], \\ \tilde{p}_{t_m} &= \frac{\mathbb{E}[\tilde{p}_{t_{m+1}}|\mathcal{F}_{t_m}]}{1-(k-\frac{\alpha}{2}\sigma^2 U_{t_m}^2 - bU_{t_m})\Delta t}. \end{aligned}$$

We estimate the conditional expectation by the Monte-Carlo regression approach proposed by [33]. With a time discretization over $[0,T]$ we use the Euler scheme to generate the paths of the forward process X_t in (2.5.2), approximated by X_{t_m}. We denote by U_{t_m} the numerical solution to the Riccati equation (2.5.6). Then the defautable bond price is estimated as

$$D(t_{t_m}, T) \approx \exp(U_{t_m} X_{t_m} + p_{t_m}).$$

The aggregate relative entropy of the optimal measure $\mathbb{Q}^\star$ with respect to $\mathbb{P}$ is estimated as

$$H_{t_m,T}(\mathbb{Q}^\star|\mathbb{P}) = \mathbb{E}^{\mathbb{Q}^\star}[\sum_{t\le t_m\le T} \frac{1}{2}\left(\sigma U_{t_m}\left(\sqrt{\alpha+\beta X_{t_m}} + q_{t_m}\right)\right)^2 \Delta t|\mathcal{F}_t].$$

We now specify the parameters $a = -1\times 10^{-2}, b = 1\times 10^{-5}, \sigma = 7.4\times 10^{-3}, R = 1, k = 0, T = 1, V_0 = 20, \sigma_V = 0.2, \kappa = 0.8$ and ξ (recovery rate) is a uniform random variable on $[0.4, 0.6]$.

Figure 2.1 shows one sample path of the realized interest rate process. Figure 2.2 present the case where default occurs before the maturity T as the value process crosses the default barrier. Figure 2.2 also shows the evolution of the defaultable bond price. The defaultable bond price has much more fluctuations before the default time, which is affected not only by the distance between the value process and the default barrier but also the time to maturity. The defaultable bond price after default time is almost constant which is determined by the recovery rate. Lastly Figure 2.2 illustrates the the aggregate relative entropy process $H(t,T)$. Similar to the price process, the aggregate relative entropy process has much more fluctuations before default due to uncertainty of default timing. After default, the aggregate relative entropy decreases to zero almost linearly since the major uncertainty after default comes from the interest rate process which is negligible compared with default risk. Figure 2.3 illustrates the case where default does not occur before maturity. The default bond price fluctuates strongly in the early period of horizon $[0,T]$ and then converges to 1 as time approaches maturity without occurrence of default.

In this chapter we considered the pricing problem for default-free bonds from a new perspective by formulating an optimal measure transformation problem. The solution of these problems consists of the optimal measure transformation and the

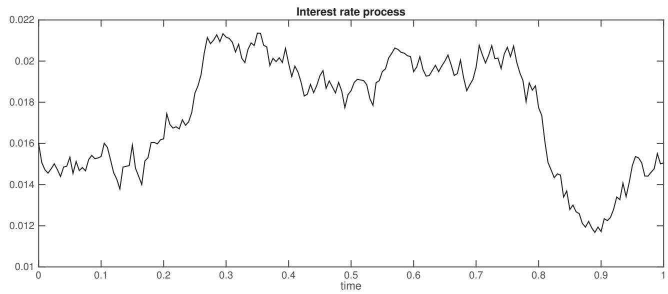

Figure 2.1: Interest rate process

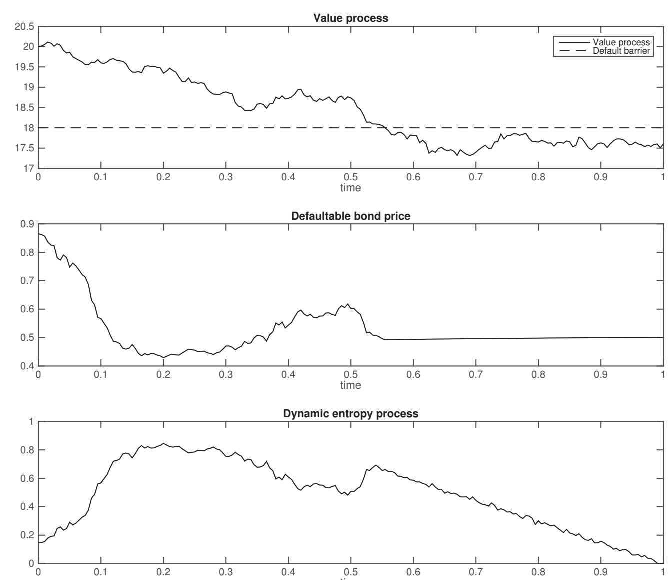

Figure 2.2: Realization with default

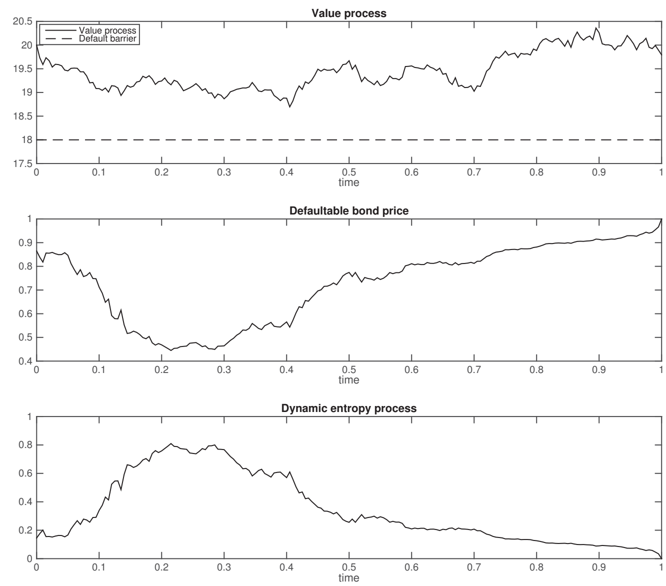

Figure 2.3: Realization without default

value process and these are characterized by the solution of a decoupled nonlinear FBSDE. The explicit solutions to FBSDEs under ATSMs and QTSMs can be found in Hyndman [40] and Hyndman and Zhou [38]. We provide an equivalence relationship between the optimal control approach in Gombani and Runggaldier [35] and the optimal measure transformation approach. We also extend the OMT problem to include jumps. We give explicit solutions to the related FBSDEs with jumps, which generalizes Hyndman [40] and Hyndman and Zhou [38]. Finally we form the OMT problem for defaultable bonds, in which case the related FBSDE generally does not have completely explicit solution due to the dependence on the general specification of the default time and recovery amount of the random terminal value of the BSDE. However, the partially explicit solution still simplifies the problem of solving the nonlinear FBSDE.

In next chapter, we will study the optimal trading problem for a small investor trading against the disorderly liquidation of a large position that impacts the market price. Further, we assume various market participants have different levels of information.

Chapter 3

Portfolio optimization under asymmetric information and market impact

3.1 The Market Model

3.1.1 Asset price and liquidation impact

Fix a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ equipped with a reference filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, with $(W_t, t \geq 0)$ an $(\mathbb{F}, \mathbb{P})$-Brownian motion. Let $T > 0$ be a finite horizon time. In our model, we assume that market participants may invest in a riskless asset and a risky asset. Without loss of generality we suppose that the interest rate of the riskless asset is zero. We assume that the fundamental value of the risky asset is modelled by a Black-Scholes diffusion:

$$dS_t = S_t(\mu dt + \sigma dW_t), \quad 0 \leq t \leq T, \tag{3.1.1}$$

where μ and σ are supposed to be constants, and $\sigma > 0$.

We consider a hedge fund which holds a large long position in the risky asset over the investment horizon $[0, T]$. In normal circumstances, this position could be held until time T. However, according to risk management policies, exchange rules, or regulatory requirements, the long position must be liquidated in certain circumstances. In this paper, we assume that the liquidation will be triggered when the market price of the risky asset passes below a pre-determined level. Before liquidation, the market price, denoted by S^M, is equal to the fundamental value S. So the liquidation time τ is defined as the first passage time of a fixed constant threshold αS_0 where $\alpha \in (0, 1)$, by the market price process S^M, i.e.,

$$\tau := \inf\{t \geq 0, S_t^M \leq \alpha S_0\} = \inf\{t \geq 0, S_t \leq \alpha S_0\} \tag{3.1.2}$$

with the convention $\inf \varnothing = \infty$. We note that τ is an $\mathbb{F}$-stopping time. In the simplest case the scenario described corresponds to a margin call that cannot be covered resulting in the liquidation, in full or in part, of the position.

The market price of the risky asset will be influenced by liquidation. Since the number of shares of the risky asset to be sold is very large in comparison to the average volume traded in a short time period, immediate liquidation would have a temporary impact on the market price which would be driven down away from the fundamental price after liquidation. We denote by $S_t^I(u)$ the market price of the risky asset at time t after the liquidation time $\tau = u$. Suppose that it is given as

$$S_t^I(u) = g(t - u; \Theta, K) S_t, \quad u \leq t \leq T. \tag{3.1.3}$$

where g is an impact function and Θ and K are parameters which will be made precise later. We note that the mathematical characterization of market impact is a very complicated problem, and we refer the interested reader to [49] for details. In this paper, inspired by [53], we characterize the temporary influence of liquidation on market by the impact function g of the form

$$g(t; \Theta, K) = 1 - \frac{Kt}{\Theta} e^{1 - \frac{t}{\Theta}} \tag{3.1.4}$$

where Θ and K are random variables with Θ controlling the speed of the market impact and K representing the magnitude of the market impact. In particular, we assume that Θ is a positive random variable and K is a random variable valued in $[0, 1]$, both of which are independent of $\mathbb{F}$ and with joint probability density function $\varphi(\cdot, \cdot)$, i.e. $\mathbb{P}(\Theta \in d\theta, K \in dk) = \varphi(\theta, k) d\theta dk$.

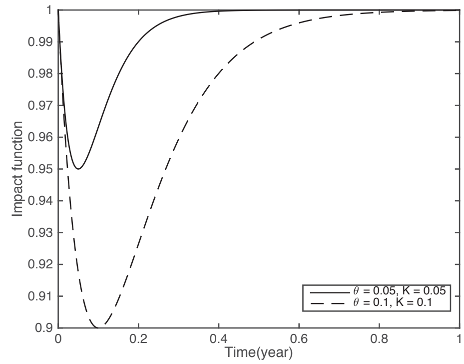

Figure 3.1: Impact function with 2 parameters

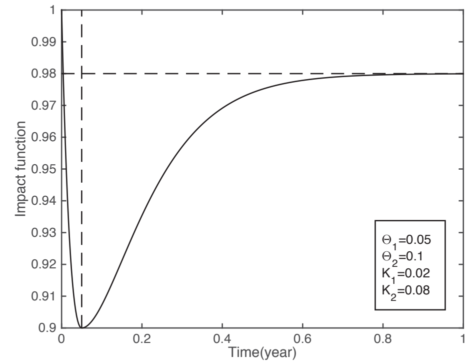

Figure 3.2: Impact function with 4 parameters

Figure 3.1 illustrates the impact function (3.1.4) with $K = 0.1$ and two different realized values of Θ. Clearly the shape of the impact function with $\Theta = 0.05$ is steeper than with $\Theta = 0.1$. We note that for each fixed scenario ω, the function g attains its minimum value $1 - K(\omega)$ at $t = \Theta(\omega)$. Also, we observe that the function g first declines from 1 and then rises back and converges to 1, which characterizes the market impact of liquidation with time evolution. For realized values $K = 0.1$ and $\Theta = 0.1$ it would take 0.1 year, which is approximately 25 trading days, for the asset price to reach the minimum value $(1 - K) * S_0$ after liquidation occurs. The market impact in the first trading day after liquidation is $1 - g(\frac{1}{250}; 0.1, 0.1) \approx 1\%$. Therefore, the parameter Θ needs to be small to more accurately reflect the impact of disorderly liquidation. In Section 3.5 we present some numerical results which use a rather large Θ that guarantees better accuracy of the numerical results, but these could be improved by applying other numerical techniques for smaller values of Θ.

Remark 3.1.1. *It is natural to consider a jump effect for the price impact of liquidation. In our model, by (3.1.3), the price before and just after liquidation satisfies the relation* $S_t^I(t) = S_t$. *However, we can approximate downward jumps of asset prices after liquidation by choosing small values of* Θ *in the smooth function* g. *Further, our model allows us to consider the situation that liquidation by the large trader may have no long-term informational content. The temporary impact on the market price decays as liquidity providers return to the market and other market participants real-*

ize that there may be no information about the fundamental value of the risky asset conveyed by the hedge fund's disorderly liquidation.

A possible extension is to consider a modified impact function with additional parameters and flexibility. For example, let

$$g(t;\Theta_1,\Theta_2,K_1,K_2) = \begin{cases} 1-\frac{(K_1+K_2)t}{\Theta_1}e^{1-\frac{t}{\Theta_1}} & 0 \le t < \Theta_1, \\ 1-K_1-\frac{K_2(t+\Theta_2-\Theta_1)}{\Theta_2}e^{1-\frac{t+\Theta_2-\Theta_1}{\Theta_2}} & \Theta_1 \le t. \end{cases} \tag{3.1.5}$$

The impact function given by (3.1.5) incorporates both permanent and temporary market impacts with K_1 and K_2 controlling the magnitude of permanent and temporary market impacts respectively. The parameters Θ_1 and Θ_2 determine both the deviation and reversal speed (see Figure 3.2). Moreover, at long-term time scale, the impact function can come back to a different level other than 1. For simplicity, we will use the impact function given by (3.1.4) in this paper and suppose the parameters Θ and K to be random variables.

Considering the market price of the asset to be equal to the fundamental value before the liquidation time τ and to be the impacted asset price after liquidation, we have that the market price is given as

$$S_t^M = 1_{\{0\le t<\tau\wedge T\}}S_t + 1_{\{\tau\wedge T\le t\le T\}}S_t^I(\tau)$$

where S_t and $S_t^I(\tau)$ are given by (3.1.1) and (3.1.3) respectively. Moreover, for any $u \ge 0$, the dynamics of the process $S_t^I(u)$ satisfies the SDE

$$dS_t^I(u) = S_t^I(u)\left(\mu_t^I(u,\Theta,K)dt + \sigma dW_t\right), \quad u \le t \le T$$

where

$$\mu_t^I(u,\Theta,K) = \frac{g'(t-u;\Theta,K)}{g(t-u;\Theta,K)} + \mu.$$

Remark 3.1.2. *The process $(S_t^I(u), t \ge u)$ is adapted with respect to the filtration $\mathbb{F}\vee\sigma(\Theta,K)$ which is the initial enlargement of $\mathbb{F}$ by the random variables (Θ,K). As we suppose $\sigma(\Theta,K)$ is independent of $\mathcal{F}_\infty$, the $(\mathbb{F},\mathbb{P})$-Brownian motion W is also a $(\mathbb{F}\vee\sigma(\Theta,K),\mathbb{P})$-Brownian motion (see e.g. [44, Section 5.9].)*

Thus the market price process of the risky asset, denoted as $S^M = (S_t^M, t \geq 0)$, satisfies the SDE

$$dS_t^M = S_t^M \left(\mu_t^M(\Theta, K)dt + \sigma dW_t\right) \tag{3.1.6}$$

where

$$\mu_t^M(\Theta, K) = 1_{\{0\leq t<\tau\wedge T\}}\mu + 1_{\{\tau\wedge T\leq t\leq T\}}\mu_t^I(\tau, \Theta, K). \tag{3.1.7}$$

We note that the market price admits a regime change at the liquidation time τ, in particular on the drift term. We give an illustrative example as below.

Example 3.1.1. *Suppose that the fundamental value process (3.1.1) is given by the Black-Scholes model with parameters* $S_0^M = 80, \mu = 0.07, \sigma = 0.2, \alpha = 0.9, \Theta = 0.1, K = 0.1$. *Figure 3.3 shows that liquidation triggers a downward jump of the drift term. Afterward the drift term first rises quickly and then declines gradually back to the original drift term. Correspondingly, Figure 3.4 shows the sample market price processes of the asset subject to liquidation impact compared with the fundamental value process.*

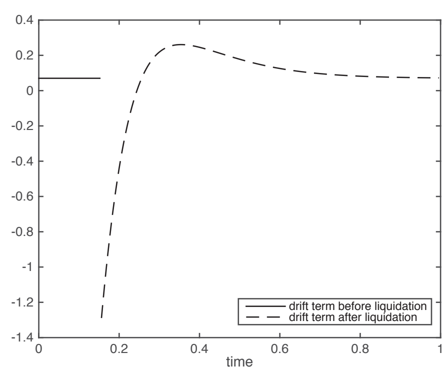

Figure 3.3: Drift μ

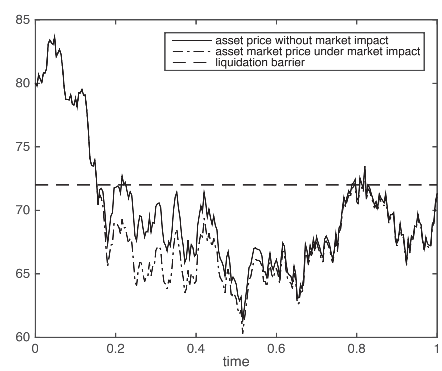

Figure 3.4: Asset price S^M

3.1.2 The optimal investment problem

Our objective is to consider the optimal investment problem from the perspective of investors who trade in the market for the risky asset subject to price impact from disorderly liquidation of the hedge fund's position. For simplicity we assume these agents may trade in the market for the risky asset without transaction costs. We consider fully informed investors, partially informed investors and uninformed investors. We suppose that all investors have access to the market price of the risky asset S^M but their knowledge of the liquidation and price impact are different. We further assume that all the investors know the values of the parameters μ and σ.

Fully informed investors observe the market price and are assumed to have complete knowledge of the mechanism of liquidation and the price impact function. Hence they know, in mathematical terms, the liquidation trigger level α, the impact function g, and the values of the random variables Θ and K when liquidation occurs. Therefore, fully informed investors have complete knowledge of the dynamics of the market price process, together with the information of the price impact.

Partially informed investors are also able to observe the market price and know the liquidation trigger level α, therefore, the liquidation time τ is also observable for them. However, partially informed investors do not have complete information about the price impact function. We suppose the partially informed investors know the functional form of the price impact function g. However, we assume the partially informed investors only know the distributions of Θ and K but not the realized value that is necessary to have full knowledge of the price impact of liquidation.

Uninformed investors are not aware of the liquidation trigger mechanism. They erroneously believe the market price process follows the Black-Scholes dynamics (3.1.1) without price impact. Therefore, they behave under incorrect assumptions, or a misspecification of the market model, which leads them to act like the Merton investor. Considering such uninformed investors allows us to quantify the value of information about the liquidation barrier and price impact, compared to a Merton-type investor.

We denote by $\mathbb{F}^S = (\mathcal{F}_t^S)_{t\geq 0}$ the natural filtration generated by the market price process S^M. Since the market price coincides with the fundamental value process S before liquidation, the liquidation time τ, which is an $\mathbb{F}$-stopping time, is also an $\mathbb{F}^S$-stopping time. We summarize the knowledge of the various investors in the following assumption.

Assumption 3.1.1. *All investors observe the market price of the risky asset and* *know the values of the parameters μ and σ. In addition certain market participants possess additional information:*

(i) The observable information for fully informed investors is modeled by the filtration

$$\mathcal{G}_t^{(2)} = \mathcal{F}_t^S \vee \sigma(\Theta, K) = \mathcal{F}_t \vee \sigma(\Theta, K),$$

they further know the liquidation barrier α, as well as the form of the impact function g.

(ii) The observable information for partially informed investors is modeled by the filtration

$$\mathcal{G}_t^{(1)} = \mathcal{F}_t^S,$$

they further know the liquidation barrier α, the form of the impact function g, and the distribution of (Θ, K).

(iii) To compare with the above two types of insiders, we consider uninformed investors who act as Merton-type investors, erroneously considering Black-Scholes dynamics with constant μ over the entire period $[0, T]$. They have no information about the liquidation mechanism. Further, they do not update their knowledge of the drift process after τ.

Remark 3.1.3. *The common information to three types investors are represented by the "public" filtration $\mathbb{F}^S$ since the market price of the risky asset is observable to all investors. Assumption 3.1.1 implies that partially informed investors know the law of $\mu_t^M(\Theta, K)$. This is similar to the weak information case of [12].*

The essential differences among these three types of investors lie in their knowledge on the drift term $\mu^M(\Theta, K)$ defined in (3.1.7). Fully informed investors are able to completely observe the drift term. Partially informed investors partially observe the drift term, corresponding to the case of partial observations considered by [47]. Partially informed investors may obtain an estimate of the drift term which is adapted to their observation process using filtering theory. Uninformed investors do not have any information about the liquidation mechanism and market impact which causes *them to erroneously specify the drift term as μ. That is, uninformed investors believe that the market prices follow the Black-Scholes dynamics (3.1.1). If the uninformed investor treated the drift of (3.1.1) as an unobservable process he could perhaps apply filtering theory to improve his investment decisions even without knowing anything about the liquidation mechanism or market impact function. However, in this paper we shall only consider the case of Assumption 3.1.1, that is of uninformed investors who estimate the drift at the beginning of the period and do not update it, since from their own view point no liquidation event happened during the period $[0, T]$. The uninformed investors are mainly considered as a benchmark for comparison with the Merton model.*

We shall study the portfolio optimization problem for three types of investors in the remainder of this paper under logarithmic and power utility.

3.2 Fully informed investors

Fully informed investors choose their trading strategy to adjust the portfolio of assets according to their information accessibility. As discussed in Section 3.1 fully informed investors know the realized values of the random variables Θ and K. The investment strategy is characterized by a $\mathbb{G}^{(2)}$-predictable process $\pi^{(2)}$ which represents the proportion of wealth invested in the risky asset. The admissible strategy set $\mathcal{A}^{(2)}$ is a collection of $\pi^{(2)}$ such that, for any $(\theta, k) \in (0, +\infty) \times (0, 1)$,

$$\int_0^T |\pi_t^{(2)} \mu_t^M(\theta, k)| dt + \int_0^T |\pi_t^{(2)} \sigma|^2 dt < \infty. \tag{3.2.1}$$

The risk aversion of the investors is modeled by classic utility functions U defined on $(0, \infty)$ that are strictly increasing, strictly concave, with continuous derivative $U'(x)$ on $(0, \infty)$, and satisfying

$$\lim_{x \to 0^+} U'(x) = +\infty \qquad \lim_{x \to \infty} U'(x) = 0.$$

We define the $\mathbb{G}^{(2)}$-martingale measure $\mathbb{Q}$ by the likelihood process

$$L_t := \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}_t^{(2)}} = \exp\left\{ -\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} dW_v - \int_0^t \frac{\left(\mu_v^M(\Theta, K)\right)^2}{2\sigma^2} dv \right\}. \tag{3.2.2}$$

As mentioned in Remark 3.1.2, W is a $(\mathbb{G}^{(2)}, \mathbb{P})$-Brownian motion. By Girsanov's theorem, the process $W^{\mathbb{Q}}$ defined as

$$W_t^{\mathbb{Q}} = W_t + \int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} dv \tag{3.2.3}$$

is an $(\mathbb{G}^{(2)}, \mathbb{Q})$-Brownian motion and the dynamics of the asset price S^M under $\mathbb{Q}$ may be written as

$$dS_t^M = S_t^M \sigma dW_t^{\mathbb{Q}}.$$

By taking a strategy $\pi^2 \in \mathcal{A}^{(2)}$, the wealth process with initial endowment $X_0 \in \mathcal{G}_0^{(2)}$ evolves as

$$dX_t^{(2)} = X_t^{(2)} \pi_t^{(2)} (\mu_t^M(\Theta, K) dt + \sigma dW_t), \qquad 0 \leq t \leq T \tag{3.2.4}$$

that is

$$X_t^{(2)} = X_0 + \int_0^T \pi_v^{(2)} \sigma S_v^M dW_v^{\mathbb{Q}}. \tag{3.2.5}$$

The fully informed investors' objective is to maximize their expected utility of terminal wealth

$$V_0^{(2)} := \sup_{\pi^{(2)} \in \mathcal{A}^{(2)}} \mathbb{E}\left[U\left(X_T^{(2)}\right)\right] \tag{3.2.6}$$

or

$$V_0^{(2)}(\Theta, K) := \operatorname*{ess\,sup}_{\pi^{(2)}(\Theta, K) \in \mathcal{A}^{(2)}} \mathbb{E}\left[U\left(X_T^{(2)}\right) | \mathcal{G}_0^{(2)}\right] \tag{3.2.7}$$

where $\mathcal{G}_0^{(2)} = \sigma(\Theta, K)$. The link between the optimization problems (3.2.6) and (3.2.7) is given by [8]; if the supremum in (3.2.7) is attained by some strategy in $\mathcal{A}^{(2)}$, then the ω-wise optimum is also a solution to (3.2.6).

As (Θ, K) is independent[1] of $\mathbb{F}$, a martingale representation theorem holds for $(\mathbb{G}^{(2)}, \mathbb{Q})$-local martingale, thus we adopt the standard "martingale approach" (see [48]) to solve the utility optimization problem (3.2.7). We may consider the following static optimization problem

$$\sup_{X_T^{(2)} \in \mathcal{V}} \mathbb{E}\left[U\left(X_T^{(2)}\right) | \mathcal{G}_0^{(2)}\right] \tag{3.2.8}$$

[1] This assumption can be relaxed into a density Jacod hypothesis, using then the result of [6, Proposition 4.6] for a martingale representation theorem.

where

$$\mathcal{V} = \left\{ X_T^{(2)} \middle| X_T^{(2)} = X_0 + \int_0^T \pi_v^{(2)} \sigma S_v^M dW_v^{\mathbb{Q}} \ , \ \pi^{(2)} \in \mathcal{A}^{(2)} \right\}.$$

The optimization problem (3.2.8) can be solved by using the method of Lagrange multipliers (see [8, Proposition 4.5]). The optimal terminal wealth $\hat{X}_T^{(2)}$ is given by

$$\hat{X}_T^{(2)} = I(\Lambda L_T), \tag{3.2.9}$$

where $I = (U')^{-1}$ and the $\mathcal{G}_0^{(2)}$-measurable random variable Λ is determined by

$$\mathbb{E}^{\mathbb{Q}} \left[I(\Lambda L_T) | \, \mathcal{G}_0^{(2)} \right] = X_0. \tag{3.2.10}$$

In order to find the optimal strategy $\hat{\pi}^{(2)}$ one should provide the dynamics of the optimal wealth process

$$\hat{X}_t^{(2)} = \mathbb{E}^{\mathbb{Q}} \left[\hat{X}_T^{(2)} | \mathcal{G}_t^{(2)} \right]. \tag{3.2.11}$$

Since $(X_t^{(2)})_{t\in[0,T]}$ is a $(\mathbb{G}^{(2)}, \mathbb{Q})$-martingale, there exists a $\mathbb{G}^{(2)}$-adapted process J such that

$$\hat{X}_t^{(2)} = \mathbb{E}^{\mathbb{Q}}[\hat{X}_T^{(2)} \big| \, \mathcal{G}_0^{(2)}] + \int_0^t J_v dW_v^{\mathbb{Q}}. \tag{3.2.12}$$

Substituting (3.2.10) into (3.2.12) we have

$$\hat{X}_t^{(2)} = X_0 + \int_0^t J_v dW_v^{\mathbb{Q}}. \tag{3.2.13}$$

Comparing (3.2.5) with (3.2.13), we obtain the optimal strategy

$$\hat{\pi}_t^{(2)} = \frac{J_t}{\sigma \hat{X}_t^{(2)}}.$$

Notice that the optimal strategy $(\hat{\pi}_t^{(2)})_{t\in[0,T]}$ involves the process J which is implicitly determined by the martingale representation as in (3.2.12). To obtain an explicit expression for the optimal strategy, we will consider power and logarithmic utilities in the following subsections.

3.2.1 Power utility

We first consider the power utility

$$U(x) = \frac{x^p}{p}, \quad 0 < p < 1.$$

Using the fact that $I(x) = x^{1/(p-1)}$ and by (3.2.9)-(3.2.10) we obtain the optimal terminal wealth

$$\hat{X}_T^{(2)} = \frac{X_0}{\mathbb{E}\left[(L_T)^{\frac{p}{p-1}} | \mathcal{G}_0^{(2)}\right]} (L_T)^{\frac{1}{p-1}} \tag{3.2.14}$$

where L_T is given by (3.2.2). The following proposition gives then the optimal expected utility as well as the optimal strategy:

Proposition 3.2.1. *For power utility $U(x) = \frac{x^p}{p}$, $0 < p < 1$, the optimal expected utility is*

$$V_0^{(2)}(\Theta, K) = \frac{(X_0)^p}{p} \left(\mathbb{E}\left[(L_T)^{\frac{p}{p-1}} |\mathcal{G}_0^{(2)}\right]\right)^{1-p} \tag{3.2.15}$$

and the optimal strategy is given by

$$\hat{\pi}_t^{(2)} = 1_{\{0 \le t < \tau \wedge T\}} \hat{\pi}_t^{(2,b)} + 1_{\{\tau \wedge T \le t \le T\}} \hat{\pi}_t^{(2,a)}, \quad t \in [0, T] \tag{3.2.16}$$

where

$$\hat{\pi}_t^{(2,b)} = \frac{\mu}{(1-p)\sigma^2} + \frac{Z_t^H}{\sigma H_t}, \quad t \in [\![0, \tau \wedge T[\![, \tag{3.2.17}$$

$$\hat{\pi}_t^{(2,a)} = \frac{\mu_t^I(\tau, \Theta, K)}{(1-p)\sigma^2}, \quad t \in [\![\tau \wedge T, T]\!] \tag{3.2.18}$$

with (H, Z^H) satisfying the following linear BSDE

$$H_t = 1 + \int_t^T \left(\frac{p\left(\mu_v^M(\Theta, K)\right)^2}{2(1-p)^2\sigma^2} H_v + \frac{p\mu_v^M(\Theta, K)}{(1-p)\sigma} Z_v^H \right) dv - \int_t^T Z_v^H dW_v. \tag{3.2.19}$$

Proof. Following [17] we find the explicit expression for the optimal strategy, by computing the dynamics of the optimal wealth process. Applying the abstract Bayes' formula to (3.2.11), we obtain

$$\begin{aligned} \hat{X}_t^{(2)} =& \mathbb{E}^{\mathbb{Q}}\left[\hat{X}_T^{(2)} | \mathcal{G}_t^{(2)}\right] \\ =& \frac{1}{L_t} \mathbb{E}\left[\hat{X}_T^{(2)} L_T | \mathcal{G}_t^{(2)}\right]. \end{aligned} \tag{3.2.20}$$

Substituting (3.2.2) and (3.2.14) into (3.2.20) we have

$$\begin{aligned}
&\hat{X}_t^{(2)} \\
=&\frac{X_0}{\mathbb{E}[(L_T)^{\frac{p}{p-1}}|\mathcal{G}_0^{(2)}]L_t}\mathbb{E}\left[(L_T)^{\frac{p}{p-1}}|\mathcal{G}_t^{(2)}\right] \\
=&\frac{X_0}{\mathbb{E}[(L_T)^{\frac{p}{p-1}}|\mathcal{G}_0^{(2)}]L_t}\mathbb{E}\left[\exp\left\{\int_0^T \frac{p\mu_v^M(\Theta,K)}{(1-p)\sigma}dW_v + \int_0^T \frac{p\left(\mu_v^M(\Theta,K)\right)^2}{2(1-p)\sigma^2}dv\right\}|\mathcal{G}_t^{(2)}\right] \\
=&\frac{X_0(L_t)^{\frac{1}{p-1}}}{\mathbb{E}[(L_T)^{\frac{p}{p-1}}|\mathcal{G}_0^{(2)}]}\mathbb{E}\left[\exp\left\{\int_t^T \frac{p\mu_v^M(\Theta,K)}{(1-p)\sigma}dW_v + \int_t^T \frac{p\left(\mu_v^M(\Theta,K)\right)^2}{2(1-p)\sigma^2}dv\right\}|\mathcal{G}_t^{(2)}\right]
\end{aligned}$$

Defining

$$H_t := \mathbb{E}[\exp\left\{\int_t^T \frac{p\mu_v^M(\Theta,K)}{(1-p)\sigma}dW_v + \int_t^T \frac{p(\mu_v^M(\Theta,K))^2}{2(1-p)\sigma^2}dv\right\}|\mathcal{G}_t^{(2)}]$$

the optimal wealth process writes as

$$\hat{X}_t^{(2)} = \frac{X_0}{H_0}(L_t)^{\frac{1}{p-1}}H_t. \tag{3.2.21}$$

In order to find the dynamics of $(H_t)_{t\in[0,T]}$, we first remark that $(M_t := H_tD_t)_{t\in[0,T]}$ is a $(\mathbb{G}^{(2)},\mathbb{P})$-martingale, where

$$D_t := \exp\left\{\int_0^t \frac{p\mu_v^M(\Theta,K)}{(1-p)\sigma}dW_v + \int_0^t \frac{p(\mu_v^M(\Theta,K))^2}{2(1-p)\sigma^2}dv\right\}. \tag{3.2.22}$$

By the martingale representation theorem there exists a $\mathbb{G}^{(2)}$-adapted process Z^M such that

$$M_t = M_0 + \int_0^t M_vZ_v^MdW_v.$$

From equation (3.2.22)

$$d(\frac{1}{D_t}) = (\frac{1}{D_t})\left\{\left(\frac{p^2(\mu_t^M(\Theta,K))^2}{2(1-p)^2\sigma^2} - \frac{p(\mu_t^M(\Theta,K))^2}{2(1-p)\sigma^2}\right)dt - \frac{p\mu_t^M(\Theta,K)}{(1-p)\sigma}dW_t\right\}$$

which leads to the following dynamics for the process $(H_t)_{t\in[0,T]}$

$$\begin{aligned}
dH_t = H_t\Bigg\{&\left(\frac{p^2(\mu_t^M(\Theta,K))^2}{2(1-p)^2\sigma^2} - \frac{p(\mu_t^M(\Theta,K))^2}{2(1-p)\sigma^2} - \frac{p\mu_t^M(\Theta,K)}{(1-p)\sigma}Z_t^M\right)dt \\
&+ \left(Z_t^M - \frac{p\mu_t^M(\Theta,K)}{(1-p)\sigma}\right)dW_t\Bigg\}.
\end{aligned}$$

Denoting

$$Z_t^H := H_tZ_t^M - \frac{p\mu_t^M(\Theta,K)}{(1-p)\sigma}$$

and using the terminal condition

$$H_T = 1,$$

then $(H_t)_{t\in[0,T]}$ satisfies the following BSDE

$$H_t = 1 + \int_t^T \left(\frac{p(\mu_v^M(\Theta,K))^2}{2(1-p)^2\sigma^2} H_v + \frac{p\mu_v^M(\Theta,K)}{(1-p)\sigma} Z_v^H \right) dv - \int_t^T Z_v^H dW_v. \qquad (3.2.23)$$

Thus the dynamics of the optimal wealth process, using (3.2.21), are

$$d\hat{X}_t^{(2)} = \hat{X}_t^{(2)} \left(\frac{\left(\mu_t^M(\Theta,K)\right)^2}{(1-p)\sigma^2} + \frac{\mu_t^M(\Theta,K) Z_t^H}{\sigma H_t} \right) dt + \hat{X}_t^{(2)} \left(\frac{\mu_t^M(\Theta,K)}{(1-p)\sigma} + \frac{Z_t^H}{H_t} \right) dW_t. \qquad (3.2.24)$$

that leads to the optimal strategy (by comparing (3.2.24) with (3.2.4))

$$\hat{\pi}_t^{(2)} = \frac{\mu_t^M(\Theta,K)}{(1-p)\sigma^2} + \frac{Z_t^H}{\sigma H_t}. \qquad (3.2.25)$$

We decompose the time horizon $[0,T]$ into two random time intervals $[\![0,\tau\wedge T[\![$ and $[\![\tau\wedge T, T]\!]$. On the random interval $[\![\tau\wedge T, T]\!]$, the fully informed investor observes the drift term μ^M thus the BSDE (3.2.23) can be solved explicitly on $[\![\tau\wedge T, T]\!]$:

$$H_t = \exp\left\{ \int_t^T \frac{p(\mu_v^I(\tau,\Theta,K))^2}{2(1-p)^2\sigma^2} dv \right\}, \qquad (3.2.26)$$

$$Z_t^H = 0. \qquad (3.2.27)$$

Recalling (3.1.7) and using (3.2.26)-(3.2.27) we may decompose the optimal strategy in (3.2.25) into two parts:

$$\hat{\pi}_t^{(2)} = 1_{\{0\le t<\tau\wedge T\}} \hat{\pi}_t^{(2,b)} + 1_{\{\tau\wedge T\le t\le T\}} \hat{\pi}_t^{(2,a)} \qquad (3.2.16)$$

where

$$\hat{\pi}_t^{(2,b)} = \frac{\mu}{(1-p)\sigma^2} + \frac{Z_t^H}{\sigma H_t}, \quad t\in[\![0,\tau\wedge T[\![, \qquad (3.2.17)$$

$$\hat{\pi}_t^{(2,a)} = \frac{\mu_t^I(\tau,\Theta,K)}{(1-p)\sigma^2}, \quad t\in[\![\tau\wedge T, T]\!]. \qquad (3.2.18)$$

□

The optimal strategy after liquidation in (3.2.18) is essentially a Merton-type strategy. The part before liquidation in (3.2.17) is the sum of a Merton strategy and an extra component[2] which is determined by the solution of the BSDE (3.2.19). It is hard to obtain a closed-form solution for the BSDE (3.2.19), however, we may solve the BSDE (3.2.19) numerically which will be discussed in Section 3.5.

We next consider the case of logarithmic utility for the fully informed investor.

3.2.2 Logarithmic utility

In this section we consider the logarithmic utility

$$U(x) = \ln(x).$$

Using the fact that $I(x) = \frac{1}{x}$ and by (3.2.9)-(3.2.10) we obtain the optimal terminal wealth

$$\hat{X}_T^{(2)} = \frac{X_0}{L_T} \tag{3.2.28}$$

where L_T is given by (3.2.2). The optimal expected utility is

$$V_0^{(2)}(\Theta, K) = \ln(X_0) - \mathbb{E}\left[\ln(L_T)\right].$$

Applying the abstract Bayes' formula to (3.2.11) and using (3.2.28), we obtain

$$\begin{aligned}\hat{X}_t^{(2)} =& \mathbb{E}^{\mathbb{Q}}\left[\hat{X}_T^{(2)}|\mathcal{G}_t^{(2)}\right]\\ =& \frac{1}{L_t}\mathbb{E}\left[\hat{X}_T^{(2)}L_T|\mathcal{G}_t^{(2)}\right]\\ =& \frac{X_0}{L_t}\end{aligned}$$

whose dynamics is given by (3.2.2) as

$$d\hat{X}_t^{(2)} = \hat{X}_t^{(2)}\Big(\frac{\big(\mu_t^M(\Theta, K)\big)^2}{\sigma^2}dt + \frac{\mu_t^M(\Theta, K)}{\sigma}dW_t\Big) = \hat{X}_t^{(2)}\frac{\mu_t^M(\Theta, K)}{\sigma}dW_t^{\mathbb{Q}}. \tag{3.2.29}$$

Comparing (3.2.29) with (3.2.4) we obtain the optimal strategy

$$\hat{\pi}_t^{(2)} = \frac{\mu_t^M(\Theta, K)}{\sigma^2}.$$

Recalling (3.1.7) we may decompose the optimal strategy into two parts

$$\hat{\pi}_t^{(2)} = 1_{\{0\le t<\tau\wedge T\}}\hat{\pi}_t^{(2,b)} + 1_{\{\tau\wedge T\le t\le T\}}\hat{\pi}_t^{(2,a)}$$

[2] This extra term is called "hedging demand for parameter risk" by Björk et al. [17].

where

$$\hat{\pi}_t^{(2,b)} = \frac{\mu}{\sigma^2}, \quad t \in [\![0, \tau \wedge T[\![, \tag{3.2.30}$$
$$\hat{\pi}_t^{(2,a)} = \frac{\mu_t^I(\tau, \Theta, K)}{\sigma^2}, \quad t \in [\![\tau \wedge T, T]\!].$$

The optimal trading strategy for the fully informed investor is composed of two Merton strategies before and after-liquidation. Accordingly we decompose the optimal wealth process $\hat{X}_t^{(2)}$ as

$$\hat{X}_t^{(2)} = 1_{\{0 \leq t < \tau \wedge T\}} \hat{X}_t^{(2,b)} + 1_{\{\tau \wedge T \leq t \leq T\}} \hat{X}_t^{(2,a)}$$

where $\hat{X}^{(2,b)}$ and $\hat{X}^{(2,a)}$ satisfy the following SDEs

$$d\hat{X}_t^{(2,b)} = \hat{X}_t^{(2,b)} \hat{\pi}_t^{(2,b)} (\mu_t dt + \sigma dW_t), \qquad t \in [\![0, \tau \wedge T[\![, \tag{3.2.31}$$
$$d\hat{X}_t^{(2,a)} = \hat{X}_t^{(2,a)} \hat{\pi}_t^{(2,a)} \left\{ \mu_t^I(\tau, \Theta, K) dt + \sigma dW_t \right\}, \qquad t \in [\![\tau \wedge T, T]\!]. \tag{3.2.32}$$

Then we decompose the expected utility of terminal wealth into two parts depending on if liquidation occurs before or after time T:

$$V_0^{(2)}(\Theta, K) = \mathbb{E}[1_{\{\tau > T\}} \ln(\hat{X}_T^{(2,b)}) | \mathcal{G}_0^{(2)}] + \mathbb{E}[1_{\{\tau \leq T\}} \ln(\hat{X}_T^{(2,a)}) | \mathcal{G}_0^{(2)}]. \tag{3.2.33}$$

The two conditional expectations in (3.2.33) are calculated in Lemma C.1 and C.2 respectively. Combining those lemmas we obtain the following result.

Proposition 3.2.2. *The optimal log expected utility for fully informed investors is*

$$\begin{aligned}
&V_0^{(2)}(\Theta, K) = \\
&\left\{ \mathcal{N}\left(\frac{-\frac{\ln \alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}} \right) - \exp\left(\frac{2\mu}{\sigma^2} - \ln \alpha \right) \mathcal{N}\left(\frac{\frac{\ln \alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}} \right) \right\} \\
&\times \left(\ln(X_0) + \frac{1}{2}(\mu - \frac{\mu^2}{\sigma^2})T \right) \\
&+ \int_{\frac{\ln \alpha}{\sigma}}^{0} \int_y^\infty \frac{2\mu x(x - 2y)}{\sqrt{2\pi T^3}} \exp\left\{ (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)x - \frac{1}{2}(\frac{\mu}{\sigma} - \frac{1}{2}\sigma)^2 T - \frac{1}{2T}(2y - x)^2 \right\} dxdy \\
&- \frac{\ln \alpha}{\sigma} \int_0^T \frac{1}{\sqrt{2\pi t^3}} \exp\left\{ -\frac{1}{2t} \left(\frac{\ln \alpha}{\sigma} - (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t \right)^2 \right\} h^{(2)}(t, \Theta, K) dt
\end{aligned}$$

where

$$h^{(2)}(t; \theta, k) := \ln X_0 + \frac{\mu \ln \alpha}{\sigma^2} + \frac{\mu}{2}t - \frac{\mu^2}{2\sigma^2}t + \int_t^T \frac{\left(\mu_v^I(t, \theta, k)\right)^2}{2\sigma^2} dv.$$

In the next section we consider the optimization problem for the partially informed investors.

3.3 Partially informed investors

The portfolio strategy for partially informed investors is supposed to be $\mathbb{G}^{(1)}$-adapted and denoted by $\pi^{(1)} = (\pi_t^{(1)}, 0 \le t \le T)$. The wealth process evolves as

$$dX_t^{(1)} = X_t^{(1)}\pi_t^{(1)}(\mu_t^M(\Theta, K)dt + \sigma dW_t), \qquad 0 \le t \le T. \tag{3.3.1}$$

Similar to (3.2.1), the admissible strategy set $\mathcal{A}^{(1)}$ is a collection of $\pi^{(1)}$ such that, for any $(\theta, k) \in (0, +\infty) \times (0, 1)$,

$$\int_0^T |\pi_t^{(1)}\mu_t^M(\theta, k)|dt + \int_0^T |\pi_t^{(1)}\sigma|^2 dt < \infty.$$

The portfolio optimization problem for partially informed investors is

$$V_0^{(1)} = \sup_{\pi^{(1)} \in \mathcal{A}^{(1)}} \mathbb{E}\left[U\left(X_T^{(1)}\right)\right]. \tag{3.3.2}$$

Note that the optimization problem (3.3.1)-(3.3.2) is the case of partial observations since the drift term in (3.3.1) is not $\mathbb{G}^{(1)}$-adapted.

Following [47] we first reduce the optimization problem of partial observation to the case of complete observation. Recall that the probability measure $\mathbb{Q}$ is defined as

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{G}_T^{(2)}} = L_T.$$

with the density process L given by (3.2.2) which is a $(\mathbb{G}^{(2)}, \mathbb{P})$-martingale.

We next define the filtered estimate of the drift $\mu_t^M(\Theta, K)$, based on the observation of the market price, by

$$\bar{\mu}_t^M = \mathbb{E}\left[\mu_t^M(\Theta, K)|\mathcal{G}_t^{(1)}\right]. \tag{3.3.3}$$

We define the innovations process $\tilde{W}$ by

$$d\tilde{W}_t = dW_t + \frac{\mu_t^M(\Theta, K) - \bar{\mu}_t^M}{\sigma}dt, \quad 0 \le t \le T. \tag{3.3.4}$$

By [17, Lemma 4.1] we know $\tilde{W}$ is a standard $(\mathbb{G}^{(1)}, \mathbb{P})$-Brownian motion. Then we may rewrite (3.1.6) as

$$dS_t^M = S_t^M \left(\bar{\mu}_t^M dt + \sigma d\tilde{W}_t\right),$$

and the wealth process $X^{(1)}$ as

$$dX_t^{(1)} = X_t^{(1)} \pi_t^{(1)} (\bar{\mu}_t^M dt + \sigma d\tilde{W}_t), \qquad 0 \leq t \leq T$$

with initial wealth $x_0 \in]0, +\infty[$. Now the dynamics of the wealth process $X^{(1)}$ is within the framework of a full observation model since $\bar{\mu}^M$ is $\mathbb{G}^{(1)}$-adapted.

Similar to the case of fully informed investors, the optimization problem (3.3.2) can be solved by the martingale approach.

Proposition 3.3.1. (i) *The optimal terminal wealth of a partially informed investors, with utility function U and $I = (U')^{-1}$ is given by*

$$\hat{X}_T^{(1)} = I(\lambda \bar{L}_T).$$

The Lagrange multiplier λ is determined by the budget constraint

$$\mathbb{E}^{\tilde{\mathbb{Q}}} \left[I(\lambda \bar{L}_T)\right] = x_0$$

and $\bar{L}_T$ is the density of the risk neutral probability measure $\tilde{\mathbb{Q}}$ for the filtration $\mathbb{G}^{(1)}$ defined by

$$\bar{L}_t = \left.\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right|_{\mathcal{G}_t^{(1)}} = \exp\left\{-\int_0^t \frac{\bar{\mu}_v^M}{\sigma} d\tilde{W}_v - \int_0^t \frac{\left(\bar{\mu}_v^M\right)^2}{2\sigma^2} dv\right\} \tag{3.3.5}$$

with $\bar{\mu}_t^M = \mathbb{E}\left[\mu_t^M(\Theta, K) | \mathcal{G}_t^{(1)}\right]$ and the innovation process $\tilde{W}$ given by (3.3.4) is a $(\mathbb{G}^{(1)}, \mathbb{P})$-Brownian motion.

(ii) *The filtered drift estimate $\bar{\mu}^M$ can be computed as*

$$\bar{\mu}_t^M = \begin{cases} \mu, & t \in [\![0, \tau \wedge T[\![\\ \dfrac{\int_0^\infty \int_0^1 \left\{\mu_t^M(\theta,k) \exp\left\{\int_0^t \frac{\mu_v^M(\theta,k)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{\left(\mu_v^M(\theta,k)\right)^2}{2\sigma^2} dv\right\}\right\} \varphi(\theta,k) d\theta dk}{\int_0^\infty \int_0^1 \left\{\exp\left\{\int_0^t \frac{\mu_v^M(\theta,k)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{\left(\mu_v^M(\theta,k)\right)^2}{2\sigma^2} dv\right\}\right\} \varphi(\theta,k) d\theta dk}, & t \in [\![\tau \wedge T, T]\!]. \end{cases}$$

Proof. (i) We define the process $\bar{L}_t = \mathbb{E}[L_t|\mathcal{G}_t^{(1)}]$ and begin by proving that it equals the right-hand side of (3.3.5). By rewriting L_t in (3.2.2) as

$$L_t = \exp\left\{-\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} dW_v^{\mathbb{Q}} + \int_0^t \frac{\left(\mu_v^M(\Theta, K)\right)^2}{2\sigma^2} dv\right\},$$

and noting that the process $1/L_t$ satisfies the equation

$$\frac{1}{L_t} = 1 + \int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} \frac{1}{L_v} dW_v^{\mathbb{Q}}, \tag{3.3.6}$$

we have by taking the conditional expectation of (3.3.6) that

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{L_t}|\mathcal{G}_t^{(1)}\right] = 1 + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} \frac{1}{L_v} dW_v^{\mathbb{Q}}|\mathcal{G}_t^{(1)}\right]. \tag{3.3.7}$$

By [54, Theorem 5.14], we have

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} \frac{1}{L_v} dW_v^{\mathbb{Q}}|\mathcal{G}_t^{(1)}\right] = \int_0^t \mathbb{E}^{\mathbb{Q}}\left[\frac{\mu_v^M(\Theta, K)}{\sigma} \frac{1}{L_v}|\mathcal{G}_t^{(1)}\right] dW_v^{\mathbb{Q}}. \tag{3.3.8}$$

Using the Bayes formula [44, Proposition 1.7.1.5] and (3.2.2), we have respectively

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{L_t}|\mathcal{G}_t^{(1)}\right] = \frac{1}{\mathbb{E}[L_t|\mathcal{G}_t^{(1)}]} \tag{3.3.9}$$

and

$$\int_0^t \mathbb{E}^{\mathbb{Q}}\left[\frac{\mu_v^M(\Theta, K)}{\sigma} \frac{1}{L_v}|\mathcal{G}_t^{(1)}\right] dW_v^{\mathbb{Q}} = \int_0^t \frac{\mathbb{E}[\mu_v^M(\Theta, K)|\mathcal{G}_t^{(1)}]}{\sigma} \frac{1}{\mathbb{E}[L_t|\mathcal{G}_t^{(1)}]} dW_v^{\mathbb{Q}}. \tag{3.3.10}$$

Substituting (3.3.9) and (3.3.10) into (3.3.7) and (3.3.8) we obtain

$$\frac{1}{\mathbb{E}[L_t|\mathcal{G}_t^{(1)}]} = 1 + \int_0^t \frac{\mathbb{E}[\mu_v^M(\Theta, K)|\mathcal{G}_t^{(1)}]}{\sigma} \frac{1}{\mathbb{E}[L_t|\mathcal{G}_t^{(1)}]} dW_v^{\mathbb{Q}}$$

which implies

$$\bar{L}_t = \exp\left\{-\int_0^t \frac{\bar{\mu}_v^M}{\sigma} dW_v^{\mathbb{Q}} + \int_0^t \frac{\left(\bar{\mu}_v^M\right)^2}{2\sigma^2} dv\right\}. \tag{3.3.11}$$

Combining (3.2.3) and (3.3.4) we have

$$dW_t^{\mathbb{Q}} = d\tilde{W}_t + \frac{\bar{\mu}_t^M}{\sigma} dt. \tag{3.3.12}$$

Substituting (3.3.12) into (3.3.11) we find

$$\bar{L}_t = \exp\left\{-\int_0^t \frac{\bar{\mu}_v^M}{\sigma} d\tilde{W}_v - \int_0^t \frac{\left(\bar{\mu}_v^M\right)^2}{2\sigma^2} dv\right\}, \tag{3.3.13}$$

which is a $(\mathbb{G}^{(1)}, \mathbb{P})$-martingale, and we define the risk neutral probability measure $\tilde{\mathbb{Q}}$ by

$$\left.\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right|_{\mathcal{G}_t^{(1)}} = \bar{L}_t.$$

By the fact that $\tilde{W}$ is a $(\mathbb{G}^{(1)}, \mathbb{P})$-Brownian motion and the Girsanov's theorem, the process $W^{\tilde{\mathbb{Q}}}$ defined as

$$W_t^{\tilde{\mathbb{Q}}} = \tilde{W}_t + \int_0^t \frac{\bar{\mu}_v^M}{\sigma} dv, \quad 0 \le t \le T$$

is a $(\mathbb{G}^{(1)}, \tilde{\mathbb{Q}})$-Brownian motion.

Following the same procedure as in Section 3.2 we find the optimal terminal wealth $\hat{X}_T^{(1)}$ given by

$$\hat{X}_T^{(1)} = I(\lambda \bar{L}_T),$$

where $I = (U')^{-1}$ and the Lagrange multiplier λ is determined by

$$\mathbb{E}\left[I(\lambda \bar{L}_T)\bar{L}_T\right] = x_0.$$

(ii) Recall that

$$\frac{1}{L_t} = \left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{G}_t^{(2)}} = \exp\left\{\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{\left(\mu_v^M(\Theta, K)\right)^2}{2\sigma^2} dv\right\}$$

is $(\mathbb{G}^{(2)}, \mathbb{Q})$-martingale. By Bayes' formula, we have

$$\begin{aligned}
\bar{\mu}_t^M =& \mathbb{E}\left[\mu_t^M(\Theta, K)|\mathcal{G}_t^{(1)}\right] \\
=& \frac{\mathbb{E}^{\mathbb{Q}}\left[\mu_t^M(\Theta, K)L_T|\mathcal{G}_t^{(1)}\right]}{\mathbb{E}^{\mathbb{Q}}[L_t|\mathcal{G}_t^{(1)}]} \\
=& \frac{\mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\mu_t^M(\Theta, K)L_T|\mathcal{G}_t^{(2)}\right]\Big|\mathcal{G}_t^{(1)}\right]}{\mathbb{E}^{\mathbb{Q}}[L_t|\mathcal{G}_t^{(1)}]} \\
=& \frac{\mathbb{E}^{\mathbb{Q}}\left[\mu_t^M(\Theta, K)L_t|\mathcal{G}_t^{(1)}\right]}{\mathbb{E}^{\mathbb{Q}}[L_t|\mathcal{G}_t^{(1)}]} \\
=& \frac{\mathbb{E}^{\mathbb{Q}}\left[\mu_t^M(\Theta, K)\exp\left\{\int_0^t \frac{\mu_v^M(\Theta,K)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{\left(\mu_v^M(\Theta,K)\right)^2}{2\sigma^2} dv\right\}|\mathcal{G}_t^{(1)}\right]}{\mathbb{E}^{\mathbb{Q}}[\exp\left\{\int_0^t \frac{\mu_v^M(\Theta,K)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{(\mu_v^M(\Theta,K))^2}{2\sigma^2} dv\right\}|\mathcal{G}_t^{(1)}]}.
\end{aligned}$$

Since the measure $\mathbb{Q}$ coincides with $\mathbb{P}$ on $\mathcal{G}_0^{(2)} = \sigma(\Theta, K)$, the distribution of (Θ, K) under $\mathbb{Q}$ is identical to the one under $\mathbb{P}$. Recall that the Brownian motion $W^{\mathbb{Q}}$ is independent of $\sigma(\Theta, K)$ we have

$$\bar{\mu}_t^M = \frac{\int_0^\infty \int_0^1 \left\{ \mu_t^M(\theta,k) \exp\left\{ \int_0^t \frac{\mu_v^M(\theta,k)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{\left(\mu_v^M(\theta,k)\right)^2}{2\sigma^2} dv \right\} \right\} \varphi(\theta,k) d\theta dk}{\int_0^\infty \int_0^1 \left\{ \exp\left\{ \int_0^t \frac{\mu_v^M(\theta,k)}{\sigma} dW_v^{\mathbb{Q}} - \int_0^t \frac{(\mu_v^M(\theta,k))^2}{2\sigma^2} dv \right\} \right\} \varphi(\theta,k) d\theta dk}. \tag{3.3.14}$$

For $t < \tau \wedge T$ we have $\bar{\mu}_t^M = \mu$ due to the fact that $\mu_t^M = \mu$. □

Remark 3.3.1. *Note that $\bar{\mu}_t^M$ is an unbiased estimator of μ_t^M.*

Following [31] there exits a martingale representation theorem with respect to the $(\mathbb{G}^{(1)}, \tilde{\mathbb{Q}})$-Brownian motion $W^{\tilde{\mathbb{Q}}}$. Similar to the case of fully informed investors, the optimal strategy $\pi^{(1)}$ relies on the martingale representation theorem. For a general utility function, the optimal strategy $\pi^{(1)}$ does not have explicit expression. In the next subsections, we will consider power and logarithmic utilities.

3.3.1 Power utility

We first consider the power utility

$$U(x) = \frac{x^p}{p}, \quad 0 < p < 1.$$

The optimal terminal wealth at T is given by

$$\hat{X}_T^{(1)} = \frac{x_0}{\mathbb{E}\left[\left(\bar{L}_T\right)^{\frac{p}{p-1}}\right]} \left(\bar{L}_T\right)^{\frac{1}{p-1}}$$

where $\bar{L}_T$ is given by equation (3.3.5). The optimal expected utility is

$$V_0^{(1)} = \frac{(x_0)^p}{p} \left(\mathbb{E}\left[\left(\bar{L}_T\right)^{\frac{p}{p-1}}\right]\right)^{1-p}. \tag{3.3.15}$$

Similar to the case of fully informed investors, we may decompose the optimal strategy $\hat{\pi}^{(1)}$ into two parts:

$$\hat{\pi}_t^{(1)} = 1_{\{0 \le t < \tau \wedge T\}} \hat{\pi}_t^{(1,b)} + 1_{\{\tau \wedge T \le t \le T\}} \hat{\pi}_t^{(1,a)}.$$

Following a similar procedure as in Section 3.2.1 we obtain

$$\hat{\pi}_t^{(1,b)} = \frac{\mu}{(1-p)\sigma^2} + \frac{Z_t^{\bar{H}}}{\sigma \bar{H}_t}, \quad t \in [\![0, \tau \wedge T[\![,$$
$$\hat{\pi}_t^{(1,a)} = \frac{\bar{\mu}_t^I}{(1-p)\sigma^2}, \quad t \in [\![\tau \wedge T, T]\!],$$

where $(\bar{H}, Z^{\bar{H}})$ satisfies the linear BSDE

$$\bar{H}_t = 1 + \int_t^T \left(\frac{p \left(\bar{\mu}_v^M\right)^2}{2(1-p)^2\sigma^2} \bar{H}_v + \frac{p\bar{\mu}_v^M}{(1-p)\sigma} Z_v^{\bar{H}} \right) dv - \int_t^T Z_v^{\bar{H}} d\tilde{W}_v. \tag{3.3.16}$$

We will discuss the numerical solution of BSDE (3.3.16) in Section 3.5.

3.3.2 Log utility

In this section we consider the logarithmic utility

$$U(x) = \ln(x).$$

The optimal terminal wealth at T is given by

$$\hat{X}_T^{(1)} = \frac{x_0}{\bar{L}_T}.$$

The optimal expected utility is

$$V_0^{(1)} = \ln(x_0) - \mathbb{E}\left[\ln(\bar{L}_T)\right].$$

The optimal investment process $\hat{\pi}^{(1)}$ is given by

$$\hat{\pi}_t^{(1)} = 1_{\{0 \le t < \tau \wedge T\}} \hat{\pi}_t^{(1,b)} + 1_{\{\tau \wedge T \le t \le T\}} \hat{\pi}_t^{(1,a)}$$

where

$$\hat{\pi}_t^{(1,b)} = \frac{\mu}{\sigma^2}, \quad t \in [\![0, \tau \wedge T[\![, \tag{3.3.17}$$
$$\hat{\pi}_t^{(1,a)} = \frac{\bar{\mu}_t^I}{\sigma^2}, \quad t \in [\![\tau \wedge T, T]\!].$$

We decompose the optimal wealth process into before and after liquidation parts as

$$\hat{X}_t^{(1)} = 1_{\{0 \le t < \tau \wedge T\}} \hat{X}_t^{(1,b)} + 1_{\{\tau \wedge T \le t \le T\}} \hat{X}_t^{(1,a)}$$

where $\hat{X}_t^{(1,b)}$ and $\hat{X}_t^{(1,a)}$ satisfy the following SDEs

$$
\begin{aligned}
d\hat{X}_t^{(1,b)} &= \hat{X}_t^{(1,b)}\hat{\pi}^{(1,b)}(\mu dt + \sigma dW_t), && t \in [\![0, \tau \wedge T[\![,\\
d\hat{X}_t^{(1,a)} &= \hat{X}_t^{(1,a)}\hat{\pi}_t^{(1,a)}\left\{\bar{\mu}_t^a dt + \sigma d\tilde{W}_t\right\}, && t \in [\![\tau \wedge T, T]\!].
\end{aligned}
$$

Then we decompose the expected utility of terminal wealth $V_0^{(1)}$ into two parts depending on if liquidation occurs before or after time T:

$$
V_0^{(1)} = \mathbb{E}\left[1_{\{T<\tau\}} \ln\left(\hat{X}_T^{(1,b)}\right)\right] + \mathbb{E}\left[1_{\{T\geq\tau\}} \ln\left(\hat{X}_T^{(1,a)}\right)\right]. \tag{3.3.18}
$$

Comparing (3.2.30) and (3.3.17), we know partially informed investors holds the same optimal strategy as the fully informed investor before liquidation. The optimal terminal wealth for partially and fully informed investors are identical if no liquidation occurs before T, that is

$$
\mathbb{E}\left[1_{\{T<\tau\}} \ln\left(\hat{X}_T^{(1,b)}\right)\right] = \mathbb{E}\left[1_{\{T<\tau\}} \ln\left(\hat{X}_T^{(2,b)}\right)\right].
$$

Thus the first expectation in (3.3.18) has been calculated in Lemma C.1 and the other expectation is calculated in Lemma C.3. Combining those lemmas we obtain the following result.

Proposition 3.3.2. *The optimal log expected utility for the fully informed investor is*

$$
\begin{aligned}
&V_0^{(1)} = \\
&\left\{\mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}}\right) - \exp\left(\frac{2\mu}{\sigma^2} - \ln\alpha\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}}\right)\right\}\\
&\times\left(\ln(x_0) + \frac{1}{2}(\mu - \frac{\mu^2}{\sigma^2})T\right)\\
&+ \int_{\frac{\ln\alpha}{\sigma}}^{0}\int_{y}^{\infty} \frac{2\mu x(x-2y)}{\sqrt{2\pi T^3}} \exp\left\{(\frac{\mu}{\sigma} - \frac{1}{2}\sigma)x - \frac{1}{2}(\frac{\mu}{\sigma} - \frac{1}{2}\sigma)^2 T - \frac{1}{2T}(2y-x)^2\right\} dxdy\\
&- \frac{\ln\alpha}{\sigma}\int_0^T \frac{1}{\sqrt{2\pi t^3}} \exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma} - (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t\right)^2\right\} h^{(1)}(t)dt
\end{aligned}
$$

where

$$
h^{(1)}(t) := \ln x_0 + \frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}t - \frac{\mu^2}{2\sigma^2}t + \int_t^T \frac{\left(\bar{\mu}_v^M\right)^2}{2\sigma^2} dv.
$$

We will consider the optimization problem for the uninformed investors.

3.4 Uninformed investors

The uninformed investors erroneously believe the market price of the asset follows a Black-Scholes dynamics with constant μ. That is, uninformed investors act as Merton investors. To compare with the fully informed and partially informed investors, we shall consider both the power utility and logarithmic utility in the following sections.

3.4.1 Power Utility

We first consider the power utility, i.e.

$$U(x) = \frac{x^p}{p}, \quad 0 < p < 1.$$

The uninformed investors adopt the Merton strategy

$$\hat{\pi}^{(0)} = \frac{\mu}{(1-p)\sigma^2}. \tag{3.4.1}$$

However, the market price process of the asset is given by (3.1.6). Therefore, corresponding to the sub-optimal strategy given by (3.4.1), the wealth process $\hat{X}_t^{(0)}$ is written as

$$\hat{X}_t^{(0)} = 1_{\{0 \le t < \tau \wedge T\}} \hat{X}_t^{(0,b)} + 1_{\{\tau \wedge T \le t \le T\}} \hat{X}_t^{(0,a)}$$

where $\hat{X}_t^b$ and $\hat{X}_t^{(0,a)}$ are given by

$$\begin{aligned} d\hat{X}_t^{(0,b)} &= \hat{X}_t^{(0,b)} \hat{\pi}^{(0)} \left(\mu dt + \sigma dW_t\right), & t \in [\![0, \tau \wedge T[\![, \\ d\hat{X}_t^{(0,a)} &= \hat{X}_t^{(0,a)} \hat{\pi}^{(0)} \left(\mu_t^I(\tau, \Theta, K) dt + \sigma dW_t\right), & t \in [\![\tau \wedge T, T]\!]. \end{aligned}$$

We next compute the expected utility of final wealth $\mathbb{E}[U(\hat{X}_T^0)]$ using the investment strategy given by (3.4.1). We decompose $\mathbb{E}[U(\hat{X}_T^0)]$ into two parts depending on whether or not liquidation occurs before time T

$$\mathbb{E}\left[U\left(\hat{X}_T^0\right)\right] = \mathbb{E}\left[1_{\{\tau > T\}} U\left(\hat{X}_T^{(0,b)}\right)\right] + \mathbb{E}\left[1_{\{\tau \le T\}} U\left(\hat{X}_T^{(0,a)}\right)\right]. \tag{3.4.2}$$

The two expectations in (3.4.2) are computed in Lemma C.4 and C.5 respectively.

Proposition 3.4.1. *The expected power utility of an uninformed investor who follows the suboptimal strategy (3.4.1) is*

$$\mathbb{E}\left[U\left(\hat{X}_T^0\right)\right] = \frac{x_0^p}{p}\exp\left(\frac{p\mu^2 T}{2(1-p)\sigma^2}\right) \times \left\{\mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma} + (\frac{\mu}{(1-p)\sigma} - \frac{\sigma}{2})T}{\sqrt{T}}\right)\right.$$
$$\left. - \exp\left(\frac{2\mu\ln\alpha}{(1-p)\sigma^2} - \ln\alpha\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma} + (\frac{\mu}{(1-p)\sigma} - \frac{\sigma}{2})T}{\sqrt{T}}\right)\right\}$$
$$- \frac{\ln\alpha}{\sigma}\int_0^1\int_0^\infty\int_0^T \frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma} - (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t\right)^2\right\} l^{(0)}(t,\theta,k)\varphi(\theta)dtd\theta dk$$

where

$$l^{(0)}(t,\theta,k) = \frac{x_0^p}{p}\exp\left\{\frac{\mu\ln\alpha}{(1-p)\sigma^2} + \frac{1}{2}\frac{\mu}{(1-p)}t - \frac{1}{2}\frac{\mu^2}{(1-p)^2\sigma^2}t\right.$$
$$\left. + \int_t^T\left(\frac{p\mu\mu_v^I(t,\theta,k)}{(1-p)\sigma^2}\right)dv\right\}.$$

We next consider the same problem for the uniformed investor under logarithmic utility.

3.4.2 Logarithmic Utility

In case of logarithmic utility, uninformed investors adopt the Merton strategy

$$\hat{\pi}^{(0)} = \frac{\mu}{\sigma^2}. \tag{3.4.3}$$

We denote by $\hat{X}_t^{(0)}$ the wealth process for uninformed investors as holding the suboptimal strategy $\hat{\pi}_t^{(0)}$ given by (3.4.3). Similar to the case of power utility we calculate the expectation $\mathbb{E}[U(\hat{X}_T^{(0)})]$ using the decomposition

$$E[\ln(\hat{X}_T^{(0)})] = \mathbb{E}[1_{\{\tau>T\}}\ln(\hat{X}_T^{(0,b)})] + \mathbb{E}[1_{\{\tau\le T\}}\ln(\hat{X}_T^{(0,a)})]. \tag{3.4.4}$$

Comparing (3.2.30) and (3.4.3), we know uninformed investors hold the same optimal strategy as the fully informed investors before liquidation. The terminal wealth for uninformed and fully informed investors are identical if no liquidation occurs before T, that is

$$\mathbb{E}\left[1_{\{T<\tau\}}\ln\left(\hat{X}_T^{(0,b)}\right)\right] = \mathbb{E}\left[1_{\{T<\tau\}}\ln\left(\hat{X}_T^{(2,b)}\right)\right].$$

Thus the first expectation in (3.4.4) has been calculated in Lemma C.1 and the other expectation is calculated in Lemma C.6.

Proposition 3.4.2. *The expected log utility of an uniformed investor who follows the suboptimal investment strategy (3.4.3) is*

$$
\begin{aligned}
&E[\ln(\hat{X}_T^{(0)})] = \\
&\left\{\mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma}+(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)T}{\sqrt{T}}\right)-\exp\left(\frac{2\mu}{\sigma^2}-\ln\alpha\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma}+(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)T}{\sqrt{T}}\right)\right\} \\
&\times\left(\ln(x_0)+\frac{1}{2}(\mu-\frac{\mu^2}{\sigma^2})T\right) \\
&+\int_{\frac{\ln\alpha}{\sigma}}^{0}\int_{y}^{\infty}\frac{2\mu x(x-2y)}{\sqrt{2\pi T^3}}\exp\left\{(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)x-\frac{1}{2}(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)^2T-\frac{1}{2T}(2y-x)^2\right\}dxdy \\
&-\frac{\ln\alpha}{\sigma}\int_0^1\int_0^\infty\int_0^T\frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{\left(\frac{\ln\alpha}{\sigma}-(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)t\right)^2}{2t}\right\}h^{(0)}(t,\theta,k)\varphi(\theta,k)dtd\theta dk
\end{aligned}
$$

where

$$
h^{(0)}(t,\theta,k) := \ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}t-\frac{\mu^2}{2\sigma^2}t+\int_t^T\left(\frac{2\mu\mu_v^I(t,\theta,k)-\mu^2}{2\sigma^2}\right)dv
$$

We next present some numerical results.

3.5 Numerical results

In this section we illustrate numerical results of the optimization problem for the three types of investors. We set the parameters $\mu = 0.07, \sigma = 0.2$ and the initial value $S_0 = 80$. We let the investment horizon $T = 1$. The liquidation trigger level is chosen as $\alpha = 0.9$. The stochastic processes are discretized using an Euler scheme with $M = 250$ steps and time intervals of length $\Delta t = \frac{1}{250}$. The number of simulations is $N = 10^5$. We suppose the distribution of (Θ, K) is uniform on $[0.05, 0.15]\times[0.02, 0.08]$. The initial wealth is assumed to be $x_0 = 80$. The power utility function is specified as $U(x) = 2x^{\frac{1}{2}}$.

3.5.1 Filtered estimate of the drift

The time horizon $[0,1]$ is discretized equally as $0 = t_0 < t_1 < \cdots < t_M = 1$. For $0 \le m \le M$ we denote by $\mu^M_{t_m}(\Theta, K)$ the discretized approximation of $\mu^M(\Theta, K)$ at time t_m. For $0 \le m \le M-1$, we denote by ΔW_m the increment of the Brownian motion over the time interval $[t_m, t_{m+1}]$. The approximation of the increment of the $(\mathbb{G}^{(2)}, \mathbb{Q})$-Brownian motion is

$$\Delta W^{\mathbb{Q}}_m = \Delta W_m + \frac{\mu^M_{t_m}(\Theta, K)}{\sigma}\Delta t.$$

We approximate the filtered drift estimate in (3.3.14) at time t_m by

$$\hat{\mu}^M_{t_m} = \frac{\int_0^\infty \int_0^1 \left\{\mu^M_{t_m}(\theta, k) G(\theta, k; \Theta, K)\right\} \varphi(\theta, k) d\theta dk}{\int_0^\infty \int_0^1 \left\{G(\theta, k; \Theta, K)\right\} \varphi(\theta, k) d\theta dk} \tag{3.5.1}$$

where $G(\theta, k; \Theta, K)$ is defined as

$$G(\theta, k; \Theta, K) = \\ \exp\left\{ \sum_{0\le i\le m-1} \left(\frac{\mu^M_{t_i}(\theta, k)}{\sigma}(\Delta W_i + \frac{\mu^M_{t_i}(\Theta, K)}{\sigma}\Delta t) - \frac{\left(\mu^M_{t_i}(\theta, k)\right)^2}{2\sigma^2}\Delta t \right) \right\}.$$

We use Monte-Carlo method to estimate the integral in (3.5.1). Suppose the number of simulation is N. For $1 \le n \le N$, we denote by (θ^n, k^n) the realized value of the random variable (Θ, K) in the nth simulation. We estimate $\hat{\mu}^M_{t_m}$ in (3.5.1) by the sample mean

$$\tilde{\mu}^M_{t_m} = \frac{\sum_{1\le n\le N} \left\{\mu^M_{t_m}(\theta^n, k^n) G(\theta^n, k^n; \Theta, K)\right\}}{\sum_{1\le n\le N} \left\{G(\theta^n, k^n; \Theta, K)\right\}}.$$

Remark 3.5.1. *The accuracy of numerical estimation $\tilde{\mu}^M_{t_m}$ can be improved by using robust filtering or non-parametric filtering.*

In Figure 3.5 we illustrate a sample filter estimate $\tilde{\mu}^M$ compared with the drift term $\mu^M(\Theta, K)$ in a specific scenario where the realized value of the liquidation random variables are $(\Theta, K) = (0.1, 0.05)$. From Figure 3.5 we note that the filtered estimate of the drift is very close to the realized drift. This result suggests that knowing the functional form of the market impact is more relevant than the actual realization of (Θ, K).

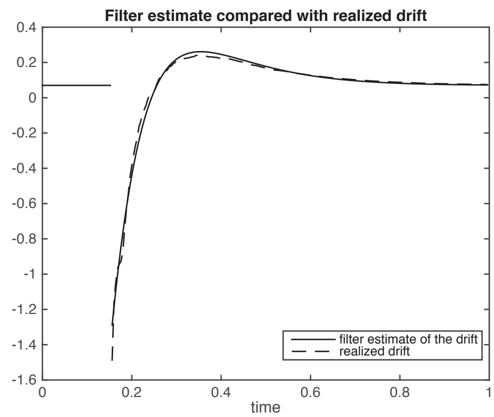

Figure 3.5: Filter estimate of the drift compared with the realized drift

3.5.2 Optimal strategy for power utility

In this section we illustrate the optimal strategies for fully and partially informed investors in case of power utility by solving the related BSDE numerically. We skip the discussion of log utility since the optimal strategies are simply the "myopic" Merton strategy. In case of fully informed investors, we approximate the BSDE (3.2.23) by the following discretized BSDE

$$\begin{aligned} \tilde{H}_{t_{m+1}} =& \tilde{H}_{t_m} - \left(\frac{p(\mu^M_{t_m}(\tilde{\theta},\tilde{k}))^2}{2(1-p)^2\sigma^2} \tilde{H}_{t_m} + \frac{p\mu^M_{t_m}(\tilde{\theta},\tilde{k})}{(1-p)\sigma} \tilde{Z}^H_{t_m} \right) \Delta t + \tilde{Z}^H_{t_m} \Delta W_{t_m}, \quad t_0 \leq t_m < t_M \\ & \qquad (3.5.2) \\ \tilde{H}_{t_M} =& 1. \qquad (3.5.3) \end{aligned}$$

The BSDE (3.5.2)-(3.5.3) can be solved using the following recursive scheme (see [33])

$$\tilde{Z}^H_{t_m} = \frac{1}{\Delta t} \mathbb{E}[\tilde{H}_{t_{m+1}} \Delta W_{t_m} | \mathcal{G}^{(2)}_{t_m}], \qquad (3.5.4)$$

$$\tilde{H}_{t_m} = \frac{\mathbb{E}[\tilde{H}_{t_{m+1}} | \mathcal{G}^{(2)}_{t_m}] + \frac{p\mu^M_{t_m}(\tilde{\theta},\tilde{k})}{(1-p)\sigma} \tilde{Z}^H_{t_m} \Delta t}{1 - \frac{p(\mu^M_{t_m}(\tilde{\theta},\tilde{k}))^2}{2(1-p)^2\sigma^2} \Delta t}. \qquad (3.5.5)$$

We estimate the conditional expectation in (3.5.4) and (3.5.5) by the Monte-Carlo regression approach proposed by [33]. Note that the market price process S^M_t is not Markovian with respect to $(\mathbb{G}^{(2)}, \mathbb{P})$. We define the running minimum process

$\tilde{S}_t^M = \inf\{S_v^M | 0 \le v \le t\}$ and note that the pair $(S_t^M, \tilde{S}_t^M)$ is Markovian with respect to $(\mathbb{G}^{(2)}, \mathbb{P})$. Hence we may choose the regression basis functions: $1, x, x^2, y, y^2$ and xy. By the regression method of [33] the conditional expectations in (3.5.4) and (3.5.5) can be estimated by

$$
\begin{aligned}
& c_1 + c_2(S_t^M - \alpha S_0) + c_3(S_t^M - \alpha S_0)^2 + c_4(\tilde{S}_t^M - \alpha S_0) + c_5(\tilde{S}_t^M - \alpha S_0)^2 \\
& + c_6(\tilde{S}_t^M - \alpha S_0)(\tilde{S}_t^M - \alpha S_0)
\end{aligned}
$$

for some coefficients $c_i, 1 \le i \le 6$.

We approximate the optimal strategy for fully informed investor $\hat{\pi}^{(2)}$ by $\tilde{\pi}^{(2,b)}$ as follows

$$
\tilde{\pi}_{t_m}^{(2)} = \frac{\mu}{(1-p)\sigma^2} + \frac{\tilde{Z}_{t_m}^H}{\sigma \tilde{H}_{t_m}}, \quad 0 \le t_m \le t_M.
$$

Following a similar procedure we may solve the related BSDE for partially informed investors and obtain the approximate optimal strategy.

Figure 3.6 illustrates the approximated optimal strategies for fully and partially investors respectively corresponding to one sample path of the risky asset price where liquidation occurs well before the terminal time T. In particular for the path of the asset price in Figure 3.6 liquidation occurs at time $t = 0.1540$. Before liquidation the two strategies are indistinguishable due to the scale. We plot the optimal strategies before liquidation in Figure 3.7 and note that there is some tracking error before liquidation. This difference may be due to the fact that the before liquidation strategy of both investors contains a component which depends on the solution of a BSDE, which is accomplished backward in time, and in particular depends recursively on the filtered drift estimate for the partially informed investor. Hence, owing to tracking error typical to filtering problems some errors may be propogated to the before liquidation strategy through the numerical solution procedure for the associated BSDE. Table 3.1 presents the approximated optimal strategies for fully and partially investors at times before liquidation corresponding to Figure 3.7. The negligible difference is due to the common information accessible to fully and partially informed investors before liquidation.

Figure 3.8 illustrates the approximated optimal strategies for fully and partially

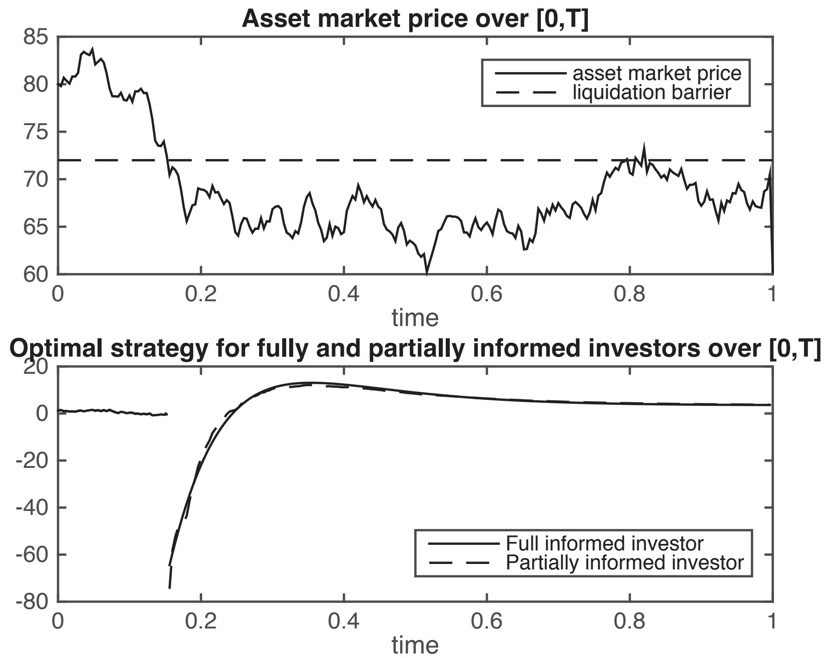

Figure 3.6: Approximated optimal strategy for fully and partially informed investors over $[0, T]$

t_m	0.1200	0.1240	0.1280	0.1320	0.1360	0.1400	0.1440	0.1480	0.1520
$S^M_{t_m}$	79.0600	79.0766	77.9106	76.2818	74.0479	73.5371	73.4940	73.9593	72.4905
$\pi^{(1)}_{t_m}$	-0.1127	-0.3614	-0.0063	-0.5712	-0.2756	-0.1780	0.0699	-0.1559	0.1043
$\pi^{(2)}_{t_m}$	-0.0898	-0.3399	-0.0224	-0.7831	-0.6907	-0.6265	-0.3766	-0.5502	-0.3760

Table 3.1: Approximated optimal strategies before liquidation

investors respectively corresponding to a realized path of the asset price that does not induce liquidation. In particular, the optimal trading strategies of the fully informed and partially informed investors appear almost identical. We also observe a general tendancy for the optimal strategies to decrease the position in the stock as its price moves toward the liquidation barrier and increase the position in the stock as the price moves away from the liquidation barrier. However, as the time to the end of the investment horizon shortens and the probability of liquidation appears less likely the overall trend to increase the position in the stock, toward the level of the Merton strategy, dominates.

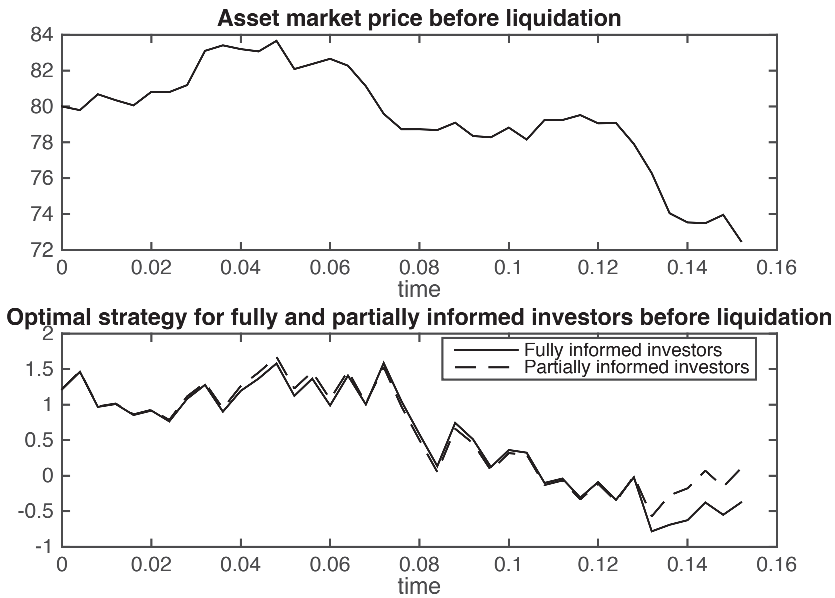

Figure 3.7: Approximated optimal strategy for fully and partially informed investors before liquidation

3.5.3 Optimal expected utility

In this subsection we implement the Monte-Carlo method to find the optimal expected power and log utilities. In case of uninformed investors, since the "optimal" strategy is simply the Merton strategy, we may approximate the wealth process $X^{(0)}$ directly using the Euler scheme. For $0 \leq m \leq M$ and $1 \leq n \leq N$, we denote by $X_{t_m}^{(0),n}$ the realized wealth for uninformed investors at time t_m in the nth simulation. The expected utility $\mathbb{E}[U(X^{(0)})]$ is approximated by the sample mean

$$\bar{V}^{(0)} = \frac{1}{N} \sum_{1 \leq n \leq N} U(X_{t_M}^{(0),n}).$$

The standard error of the sample mean is

$$SE^{(0)} = \sqrt{\frac{1}{(N-1)N} \sum_{1 \leq n \leq N} \Big(U(X_{t_M}^{(0),n}) - \bar{V}^{(0)} \Big)^2}.$$

The relative standard error of the sample mean is

$$RSE^{(0)} = (SE^{(0)})/(|\bar{V}^{(0)}|).$$

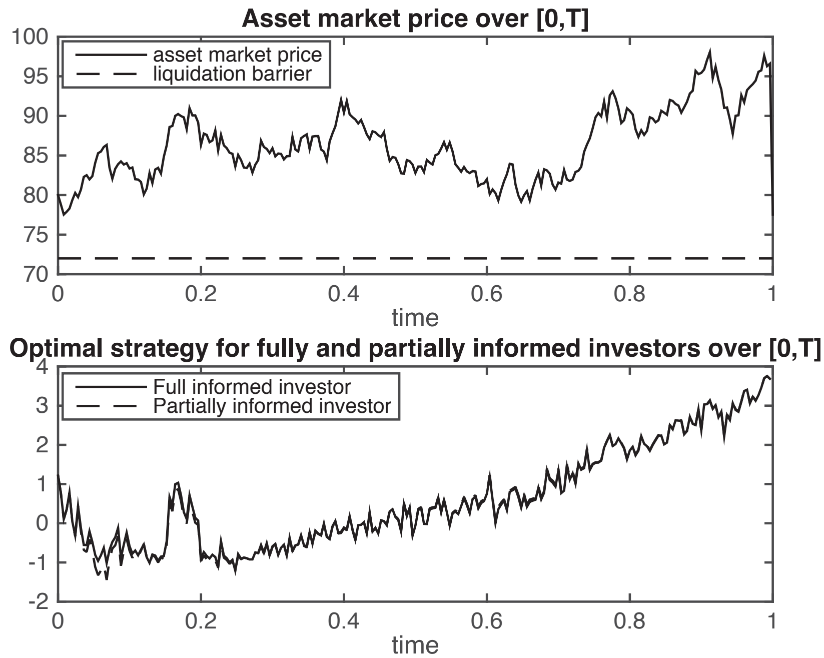

Figure 3.8: Approximated optimal strategy for fully and partially informed investors without liquidation

The 95% confidence interval estimate of the sample mean is

$$[\bar{V}^{(0)} - 1.96 * SE^{(0)}, \bar{V}^{(0)} + 1.96 * SE^{(0)}].$$

This simulation scheme also applies to the log utility for fully and partially informed investors.

However, in case of the power utility for fully and partially informed investors we cannot approximate the wealth process directly since the optimal strategies are not explicitly determined. Although we can first approximate the optimal strategies by solving the related BSDE, this would increase the size of simulation error. Instead we simulate the likelihood process L in (3.2.2) and $\bar{L}$ in (3.3.13) since the optimal expected power utilities are functionals of L_T and $\bar{L}_T$ given by (3.2.15) and (3.3.15) respectively. For instance, in case of power utility for fully informed investors, we denote the discretized realization of L_t in nth simulation by $L^n_{t_m}$ for $0 \leq m \leq M$ and

$1 \leq n \leq N$. The expectation $\mathbb{E}[(L_T)^{\frac{p}{p-1}}]$ is estimated by the sample mean

$$\bar{\xi} = \frac{1}{N} \sum_{1 \leq n \leq N} (L^n_{t_M})^{\frac{p}{p-1}}.$$

The standard error of the sample mean is

$$SE^{(2)} = \sqrt{\frac{1}{(N-1)N} \sum_{1 \leq n \leq N} \left((L^n_{t_M})^{\frac{p}{p-1}} - \bar{\xi} \right)^2}.$$

The relative standard error of the sample mean is

$$RSE^{(2)} = (SE^{(2)})/(|\bar{\xi}|).$$

The 95% confidence interval estimate of the sample mean is

$$[\bar{\xi} - 1.96 * SE^{(2)}, \bar{\xi} + 1.96 * SE^{(2)}].$$

By (3.2.15) the optimal expected utility for fully informed investors is estimated by

$$\bar{V}^{(2)} = \frac{x_0^p}{p} (\bar{\xi})^{1-p}.$$

The 95% confidence interval estimate of optimal expected utility is

$$[\frac{x_0^p}{p} \left(\bar{\xi} - 1.96 * SE^{(2)} \right)^{1-p}, \frac{x_0^p}{p} \left(\bar{\xi} + 1.96 * SE^{(2)} \right)^{1-p}].$$

A similar scheme can be applied to the case of power utility for partially informed investors.

We present the numerical results on the optimal expected utilities for the three types of investors in the Table 3.2 and Table 3.3 for power and log utilities respectively. As should be expected there exists certain gaps among the optimal expected utilities of different types of investors. We may interpret those gaps as the value of information asymmetry. The results are more pronounced in the case of power utility than in the case of logarithmic utility. Nevertheless, in both cases there are statistically significant differences in optimal expected wealth given that the confidence intervals do not overlap. In the power utility case the optimal strategy of the partially informed investor is very close to that of the fully informed investor. However, the inability to fully capture the potential gains from trading against liquidation, owning to the need

Expected utilities	Numerical evaluation		
	Sample mean	Relative standard error	95% estimated confidence interval
Fully informed	48.9602	0.0883	[44.5223, 53.0279]
Partially informed	31.3099	0.0172	[30.7767, 31.8342]
Uninformed	18.9228	0.0012	[18.8796, 18.9661]

Table 3.2: Numerical evaluation of optimal power utilities for three types of investors

Expected utilities	Numerical evaluation		
	Sample Mean	Relative standard error	95% estimated confidence interval
Fully informed	4.8282	0.0073	[4.8219, 4.8346]
Partially informed	4.7579	0.0080	[4.7520, 4.7638]
Uninformed	4.3665	0.0005	[4.3621, 4.3709]

Table 3.3: Numerical evaluation of optimal log utilities for three types of investors

to estimate the drift and the tracking error, leads to a significantly lower optimal expected utility.

In this chapter, we characterize the market impact of liquidation by a function of certain form. We consider the portfolio optimization problem for three types of investors with different levels of information about the liquidation trigger mechanism and the market impact. In case of logarithmic utility, we find the closed-form optimal strategy for all three types of investors. In the case of power utility it is not as straightforward to find the closed-form optimal strategy for the partially informed investors, therefore we use numerical solutions of the BSDEs characterizing the optimal strategies. Finally, we present some numerical results using Monte-Carlo simulation method.

These results indicate that there is significant value, in terms of optimal expected utility, of increased information about the liquidation trigger and market impact in order to trade optimally against an investor who may need to liquidate a large position in a disorderly fashion. The fully and partially informed investors may take advantage of additional information to gain larger portfolio utility than uninformed investors. Fully and partially informed investors tend to short large positions after they observe the occurrence of liquidation since they expect a market price decrease under the temporary market impact. However, fully and partially informed investors also take on very high risk since large short positions might incur large losses in extreme scenarios. In order to maintain and control the risk, we consider the portfolio optimization problem under risk constraints for fully and partially informed investors in next chapter.

Chapter 4

Portfolio optimization with risk constraints

As discussed in the previous chapter, fully and partially informed investors tend to short large positions after they observe the occurrence of liquidation. In order to reduce the risk of large short positions we consider optimal trading problems with risk constraints for fully and partially informed investors in this chapter.

4.1 Risk measures

The uncertainty of the terminal value of a wealth process is characterized by a random variable X on a probability space $(\Omega, \mathcal{F})$. We denote by $\mathcal{X}$ a set of random variables on $(\Omega, \mathcal{F})$. Then a quantitative measure of risk is given by a mapping ρ from $\mathcal{X}$ to $\mathbb{R}$. We first review the definitions for different types of risk measures such as monetary, convex and coherent risk measures (see Föllmer and Schied [30], Artzner et al. [10]).

Definition 4.1.1. *A mapping $\rho(\cdot) : \mathcal{X} \to \mathbb{R}$ is called a monetary risk measure if $\rho(0)$ is finite and if $\rho(\cdot)$ satisfies the following properties for all $X, Y \in \mathcal{X}$.*

- *Monotonicity: $\rho(X) \leq \rho(Y)$ for $X \leq Y$.*
- *Translation invariance: $\rho(X + m) = \rho(X) - m$ for $m \in \mathbb{R}$.*

The interpretation for monotonicity is clear in that the risk of a position increases as the position size increases. The translation invariance property is also called cash invariance. The constant m is interpreted as a capital requirement that makes the position X acceptable from the point view of a supervising agency.

Definition 4.1.2. *A monetary risk measure ρ is called a convex risk measure if it satisfies*

- *Convexity:* $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y), \quad$ *for any* $\lambda \in [0,1]$.

The convexity property shows that a convex risk can benefit from the diversification of positions, i.e., the risk of a diversified position $\lambda X + (1-\lambda)\rho(Y)$ is less or equal to the weighted average of the individual risks.

Definition 4.1.3. *A convex measure of risk ρ is called a coherent risk measure if it satisfies*

- *Positive homogeneity:* $\rho(\lambda X) = \lambda\rho(X)$ *for* $\lambda > 0$.

In the following sections we will focus on convex measures. Föllmer and Schied [30, Theorem 5] provides a representation for convex measures of risk.

Proposition 4.1.1. *Let $\mathcal{X}$ be the set of random variables on $(\Omega, \mathcal{F})$. Then $\rho(\cdot) : \mathcal{X} \to \mathbb{R}$ is a convex measure of risk if and only if there exists a penalty function $\alpha : \mathcal{Q} \to (-\infty, \infty)$ such that*

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{Q}} \{\mathbb{E}[-X] - \alpha(\mathbb{Q})\} \tag{4.1.1}$$

where $\mathcal{Q}$ is the set of probability measures on $(\Omega, \mathcal{F})$. In a particular case, we can set $\alpha(\cdot)$ as the relative entropy.

Remark 4.1.1. *Financial interpretation: X is a random payoff or return, the convex risk measure $\rho(X)$ can be interpreted as the minimal amount of required capital against the financial risk (market risk or credit risk).*

We next introduce a particular convex risk measure, which is called utility-based shortfall risk.

Definition 4.1.4. *(Loss function). A function $L : (-\infty, 0) \to \mathbb{R}$ is called loss function, if it is strictly increasing, strictly convex, continuously differentiable and satisfies $\lim_{x\to 0} L'(x) > -\infty$ and $\lim_{x\to -\infty} L'(x) = 0$.*

We may define a utility-based shortfall risk measure as follows (refer to Föllmer and Schied [30]).

Definition 4.1.5. *(Utility-based shortfall risk) A risk measure ρ is called utility-based shortfall risk, if there exists a loss function L defined according to Definition 4.1.4, such that ρ can be written in the form*

$$\rho(X) = \inf\{m \in \mathbb{R} \ : \ \mathbb{E}[L(-X - m)] \leq \epsilon\}.$$

By taking an exponential loss function, $L(x) = \exp(\gamma x)$, we can construct a typical utility-based risk measure which is called the entropic risk measure, defined as

$$e_\gamma(X) = \frac{1}{\gamma}\left(\ln \mathbb{E}[e^{-\gamma X}]\right) \tag{4.1.2}$$

where $\lambda > 0$ is the parameter of risk aversion. Dai Pra et al. [22] shows that the entropic risk measure has the following dual representation

$$e_\gamma(X) = \sup_{\mathbb{Q}\in\Omega(\mathbb{P})} \{\mathbb{E}^{\mathbb{Q}}[-X] - H(\mathbb{Q}|\mathbb{P})\} \tag{4.1.3}$$

where $\Omega(\mathbb{P})$ is the set of probability measures absolutely continuous with respect to $\mathbb{P}$ and $H(\mathbb{Q}|\mathbb{P})$ is the relative entropy.

4.2 Optimization problem with risk constraints

In this section we consider the optimal trading problems with risk constraint for both fully and partially informed investors. Recall from previous section that the market price process of the risky asset S^M satisfies the SDE

$$dS_t^M = S_t^M \left(\mu_t^M(\Theta, K)dt + \sigma dW_t\right) \tag{4.2.1}$$

where

$$\mu_t^M(\Theta, K) = 1_{\{0\leq t<\tau\wedge T\}}\mu + 1_{\{\tau\wedge T\leq t\leq T\}}\mu_t^I(\tau, \Theta, K). \tag{4.2.2}$$

Both fully and partially informed investors may invest in the risky asset and a riskless asset. Without loss of generality we suppose that the interest rate of the riskless asset is zero. As assumed in Section 3.1.2 the information accessible to fully and partially informed investors is characterized by two different filtrations $\mathcal{G}_t^{(2)}$ and $\mathcal{G}_t^{(1)}$ respectively.

4.2.1 Fully informed investors

Similarly to Section 3.2 we denote the admissible strategy set by $\mathcal{A}^{(2)}$. By taking a strategy $\pi^2 \in \mathcal{A}^{(2)}$, the wealth process with initial endowment $X_0 \in \mathcal{G}_0^{(2)}$ evolves as

$$dX_t^{(2)} = X_t^{(2)}\pi_t^{(2)}(\mu_t^M(\Theta, K)dt + \sigma dW_t), \qquad 0 \leq t \leq T. \tag{4.2.3}$$

Given the terminal wealth for fully informed investor is $X_T^{(2)}$, we define the conditional entropic risk of the terminal wealth for fully informed investors

$$e_\gamma(X_T^{(2)}) = \frac{1}{\gamma}\left(\ln \mathbb{E}[e^{-\gamma X_T^{(2)}}|\mathcal{G}_0^{(2)}])\right). \tag{4.2.4}$$

We consider the following optimal trading problem

$$V_0^{(2)}(\Theta, K) := \operatorname*{ess\,sup}_{\pi^{(2)} \in \mathcal{A}^{(2)}} \mathbb{E}\left[U\left(X_T^{(2)}\right)|\mathcal{G}_0^{(2)}\right] \tag{4.2.5}$$

subject to the risk constraint

$$e_\gamma(X_T^{(2)}) \leq \epsilon^{(2)} \tag{4.2.6}$$

where $\mathcal{G}_0^{(2)} = \sigma(\Theta, K)$. The risk constraint $\epsilon^{(2)}$ plays as a role to restrain the risk level of the terminal wealth.

We next find the proper bounds for $\epsilon^{(2)}$ that guarantee the problem (4.2.5)-(4.2.6) has a solution. We first define the upper bound as

$$\epsilon_{max}^{(2)} := e_\lambda(\hat{X}_T^{(2)}). \tag{4.2.7}$$

where $\hat{X}_T^{(2)}$ is the optimal terminal wealth for problem (3.2.7). Notice that $\hat{X}_T^{(2)}$ is still optimal for the problem (4.2.5)-(4.2.6) if $\epsilon \geq \epsilon_{max}$. In this case, the portfolio optimization with and without the risk constraint have the same optimal solution, where the risk constraint is unbinding.

We next define the lower bound as

$$\epsilon_{min}^{(2)} := \operatorname*{ess\,inf}_{\pi^{(2)}\in\mathcal{A}^{(2)}} \{e_\lambda(X_T^{(2)})\}. \tag{4.2.8}$$

Recall the definition of $e_\lambda(\cdot)$ in (4.1.2) and we can rewrite (4.2.8) as

$$\epsilon_{min}^{(2)} = \operatorname*{ess\,inf}_{\pi^{(2)}\in\mathcal{A}^{(2)}} \{\frac{1}{\gamma}(\ln \mathbb{E}[e^{-\gamma X_T^{(2)}}|\mathcal{G}_0^{(2)}])\}. \tag{4.2.9}$$

$$1 - e^{\gamma\epsilon_{min}^{(2)}} = \operatorname*{ess\,sup}_{\pi^{(2)}\in\mathcal{A}^{(2)}} \{(\mathbb{E}[1 - e^{-\gamma X_T^{(2)}}|\mathcal{G}_0^{(2)}])\}. \tag{4.2.10}$$

To compute $e^{\gamma\epsilon_{min}^{(2)}}$ is equivalent to consider the following portfolio optimization problem

$$\operatorname*{ess\,sup}_{\pi^{(2)}\in\mathcal{A}^{(2)}} \{U(X_T^{(2)})\} \tag{4.2.11}$$

where $U(\cdot)$ is defined as the exponential utility function, i.e.

$$U(x) = 1 - e^{-\gamma x}. \tag{4.2.12}$$

Then it is easy to find

$$I(x) = (U')^{-1}(x) = -\frac{1}{\gamma}\ln(\frac{x}{\gamma}).$$

From (3.2.9)-(3.2.10) we know the optimal terminal wealth $\hat{X}_T^{(2)}$ is given by

$$\hat{X}_T^{(2)} = -\frac{1}{\gamma}\ln(\frac{\Lambda L_T}{\gamma}), \tag{4.2.13}$$

where L_T is defined in (3.2.2) and the $\mathcal{G}_0^{(2)}$-measurable random variable Λ is determined by

$$\mathbb{E}^{\mathbb{Q}}\left[I(\Lambda L_T)|\, \mathcal{G}_0^{(2)}\right] = X_0. \tag{4.2.14}$$

Using (4.2.10) we compute

$$\epsilon_{min}^{(2)} = \frac{1}{\gamma}\ln\mathbb{E}[\frac{\Lambda L_T}{\gamma}|\mathcal{G}_0^{(2)}]. \tag{4.2.15}$$

Clearly if $\epsilon < \epsilon_{min}^{(2)}$ the portfolio optimization problem with risk constraint (4.2.5)-(4.2.6) has no solution. From now on, we assume $\epsilon_{min}^{(2)} < \epsilon < \epsilon_{max}$.

To solve the problem (4.2.5)-(4.2.6) , we first write the risk constraint (4.2.6) as

$$\frac{1}{\gamma}(\ln \mathbb{E}[e^{-\gamma X_T^{(2)}}|\mathcal{G}_0^{(2)}]) \leq \epsilon^{(2)}, \tag{4.2.16}$$

which is equivalent to

$$\mathbb{E}[e^{-\gamma X_T^{(2)}}|\mathcal{G}_0^{(2)}] \leq e^{\gamma\epsilon^{(2)}}. \tag{4.2.17}$$

We use the method of Lagrange multipliers. We only need to solve the following unconstrained optimization problem:

$$\begin{aligned}
&\operatorname*{ess\,sup}_{X_T^{(2)}} \left\{\mathbb{E}[U(X_T^{(2)})|\mathcal{G}_0^{(2)}] - y_1\left(\mathbb{E}[X_T^{(2)}|\mathcal{G}_0^{(2)}] - X_0\right) - y_2\left(\mathbb{E}[e^{-\gamma X_T^{(2)}}|\mathcal{G}_0^{(2)}] - e^{\gamma\epsilon^{(2)}}\right)\right\}\\
=&\operatorname*{ess\,sup}_{X_T^{(2)}} \left\{\mathbb{E}[U(X_T^{(2)}) - y_1 X_T^{(2)} - y_2 e^{-\gamma X_T^{(2)}} + y_1 X_0 + y_2 e^{\gamma\epsilon^{(2)}}|\mathcal{G}_0^{(2)}]\right\}.
\end{aligned}$$

Using similar technique to [8, Proposition 4.5] we define

$$\tilde{U}(x; y_2) := U(x) - y_2 e^{-\gamma x}. \tag{4.2.18}$$

We find the optimal terminal wealth $\hat{X}_T^{(2)}$ is given by

$$\hat{X}_T^{(2)} = \tilde{I}(\hat{y}_1 L_T; \hat{y}_2) \tag{4.2.19}$$

where $\tilde{I}(x; y_2)$ is the inverse function

$$\tilde{I}(x; \hat{y}_2) := (\tilde{U}')^{-1}(x; \hat{y}_2).$$

and $\hat{y}_1, \hat{y}_2$ satisfies

$$\mathbb{E}[L_T \tilde{I}(\hat{y}_1 L_T; \hat{y}_2)|\mathcal{G}_0^{(2)}] = X_0, \tag{4.2.20}$$

$$\mathbb{E}[\exp(-\gamma \tilde{I}(\hat{y}_1 L_T; \hat{y}_2))|\mathcal{G}_0^{(2)}] = e^{\gamma\epsilon^{(2)}}. \tag{4.2.21}$$

In order to find the optimal strategy $\hat{\pi}^{(2)}$ we combine the terminal condition (4.2.19) with the dynamics of the wealth process (4.2.3) to form the following BSDE

$$X_t^{(2)} = \tilde{I}(\hat{y}_1 L_T; \hat{y}_2) - \int_t^T \frac{\mu_v^M(\Theta, K)}{\sigma} Z_v dv - \int_t^T Z_v dW_v. \tag{4.2.22}$$

Then the optimal strategy is given by

$$\tilde{\pi}_t^{(2)} = \frac{Z_t}{X_t^{(2)}\sigma}.$$

It is hard to find explicit solutions for the equation system (4.2.20)-(4.2.21) and the BSDE (4.2.22). We will discuss numerical solutions in Section 4.2.3.

4.2.2 Partially informed investors

The portfolio strategy for partially informed investors is supposed to be $\mathbb{G}^{(1)}$-adapted and denoted by $\pi^{(1)}$. The wealth process evolves as

$$dX_t^{(1)} = X_t^{(1)}\pi_t^{(1)}(\mu_t^M(\Theta, K)dt + \sigma dW_t), \qquad 0 \leq t \leq T. \tag{4.2.23}$$

The set of the admissible strategies set is denoted by $\mathcal{A}^{(1)}$.

Recalling the filtered drift term $\bar{\mu}_t^M$ in (3.3.3) and the innovations process $\tilde{W}$ in (3.3.4), We may write the wealth process $X^{(1)}$ as

$$dX_t^{(1)} = X_t^{(1)}\pi_t^{(1)}(\bar{\mu}_t^M dt + \sigma d\tilde{W}_t), \qquad 0 \leq t \leq T$$

with initial wealth $x_0 \in]0, +\infty[$. Now the dynamics of the wealth process $X^{(1)}$ is within the framework of a full observation model since $\bar{\mu}^M$ is $\mathbb{G}^{(1)}$-adapted.

Given the terminal wealth for partially informed investor is $X_T^{(1)}$, we define the entropic risk of the terminal wealth for partially informed investors

$$e_\gamma(X_T^{(1)}) = \frac{1}{\gamma}\left(\ln \mathbb{E}[e^{-\gamma X_T^{(1)}}]\right). \tag{4.2.24}$$

Add the constraint to Problem 3.3.2 to form the following optimization problem

$$\sup_{\pi^{(1)} \in \mathcal{A}^{(1)}} \mathbb{E}\left[U\left(X_T^{(1)}\right)\right] \tag{4.2.25}$$

subject to the risk constraint

$$e_\gamma(X_T^{(1)}) \leq \epsilon^{(1)}. \tag{4.2.26}$$

Similar to the case of fully informed investors we may determine the upper bound $\epsilon_{max}^{(1)}$ and the lower bound $\epsilon_{min}^{(1)}$ for the risk constraint $\epsilon^{(1)}$. The upper bound is defined as

$$\epsilon_{max}^{(1)} := e_\lambda(\hat{X}_T^{(1)}).$$

where $\hat{X}_T^{(1)}$ is the optimal terminal wealth for problem (3.3.2). The lower bound is defined as

$$\epsilon_{min}^{(1)} = \frac{1}{\gamma}\ln \mathbb{E}[\frac{\lambda \bar{L}_T}{\gamma}].$$

where $\bar{L}$ is defined in (3.3.5) and λ is determined by

$$\mathbb{E}\left[\bar{L}_T I(\lambda \bar{L}_T)\right] = x_0.$$

To guarantee the existence of a solution the problem (4.2.25)-(4.2.26), we choose $\epsilon^{(1)}_{min} < \epsilon^{(1)} < \epsilon^{(1)}_{max}$.

Following a similar procedure for fully informed investor, we find the optimal terminal wealth for partially informed investor $\hat{X}^{(1)}_T$ given by

$$\hat{X}^{(1)}_T = \tilde{I}(\hat{y}_1 \bar{L}_T; \hat{y}_2), \tag{4.2.27}$$

where $\bar{L}_t$ is defined in (3.3.5) and $\hat{y}_1, \hat{y}_2$ are determined by

$$\mathbb{E}[\bar{L}_T \tilde{I}(\hat{y}_1 \bar{L}_T; \hat{y}_2)] = x_0, \tag{4.2.28}$$

$$\mathbb{E}[\exp(-\gamma \tilde{I}(\hat{y}_1 \bar{L}_T; \hat{y}_2))] = e^{\gamma \epsilon^{(1)}}. \tag{4.2.29}$$

The optimal strategy $\hat{\pi}^{(2)}$ is given by

$$\tilde{\pi}^{(2)}_t = \frac{Z_t}{X^{(1)}_t \sigma},$$

where $(X^{(1)}_t, Z_t)$ satisfies the following BSDE

$$X^{(1)}_t = \tilde{I}(\hat{y}_1 \bar{L}_T; \hat{y}_2) - \int_t^T \frac{\bar{\mu}^M_v}{\sigma} Z_v dv - \int_t^T Z_v dW_v. \tag{4.2.30}$$

We will discuss the numerical solutions for the equation system (4.2.28)-(4.2.29) and the BSDE (4.2.30) in Section 4.2.3.

4.2.3 Numerical Results

In this section we illustrate numerical results. We compute the optimal utility and optimal trading strategy numerically for both fully and partially informed investors. In order to compare with the results of optimal trading problems without risk constraints, we use the same assumptions for model parameters as in Section 3.5. We set the parameters $\mu = 0.07, \sigma = 0.2$ and the initial value $S_0 = 80$. We let the investment horizon $T = 1$. The liquidation trigger level is chosen as $\alpha = 0.9$. The stochastic processes are discretized using an Euler scheme with $M = 250$ steps and time intervals of length $\Delta t = \frac{1}{250}$. The number of simulations is $N = 10^5$. We suppose the distribution of (Θ, K) is uniform on $[0.05, 0.15] \times [0.02, 0.08]$. The initial wealth is assumed to be $x_0 = 80$. We assume the risk aversion parameter $\gamma = 1$. We consider the power utility function $U(x) = 2x^{\frac{1}{2}}$.

We first consider the case for fully informed investors. Before we compute the optimal utility we need first determine the lower and upper bound for the risk constraint Recall from (3.2.14) that the optimal terminal wealth for the fully informed investor without risk constraint is given by

$$\hat{X}_T^{(2)} = \frac{X_0}{\mathbb{E}\left[(L_T)^{\frac{p}{p-1}} | \mathcal{G}_0^{(2)}\right]} (L_T)^{\frac{1}{p-1}} \tag{4.2.31}$$

Substituting (4.2.31) into (4.2.7) we have

$$\epsilon_{max}^{(2)} = \frac{1}{\gamma} \ln \left(\mathbb{E} \left[\exp \left\{ \frac{-\gamma X_0 (L_T)^{\frac{1}{p-1}}}{\mathbb{E}\left[(L_T)^{\frac{p}{p-1}} | \mathcal{G}_0^{(2)}\right]} \right\} | \mathcal{G}_0^{(2)} \right]) \right). \tag{4.2.32}$$

We use Monte-Carlo methods based on N realizations of L_T to approximate the expectation involved in (4.2.31). Recall that the process L_t is given by

$$L_t = \exp\left\{ -\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} dW_v - \int_0^t \frac{\left(\mu_v^M(\Theta, K)\right)^2}{2\sigma^2} dv \right\}.$$
$$L_0 = 1.$$

With a time discretization over $[0, T]$ we use the Euler scheme with M time-steps to generate N realizations of L_T. We estimate the upper bound as $\epsilon_{max}^{(2)} = 79.843$.

Recall from (4.2.15) that the lower bound $\epsilon_{min}^{(2)}$ is given by

$$\epsilon_{min}^{(2)} = \frac{1}{\gamma} \ln \mathbb{E}[\frac{\Lambda L_T}{\gamma} | \mathcal{G}_0^{(2)}]. \tag{4.2.33}$$

Similarly we use Monte-Carlo methods based on N realizations of L_T to approximate the expectation in (4.2.33). We estimate the lower bound as $\epsilon_{min}^{(2)} = 17.951$.

We now choose a particular risk constraint $\epsilon_{min}^{(2)} = 50$. To compute the optimal utility, we need to solve the equation system (4.2.20)-(4.2.21). Taking the left hand sides of (4.2.20)-(4.2.21) we define the budget function

$$F(y_1, y_2) = \mathbb{E}[L_T \tilde{I}(y_1 L_T; y_2) | \mathcal{G}_0^{(2)}] \tag{4.2.34}$$

and the risk function

$$G(y_1, y_2) = \mathbb{E}[\exp(-\gamma \tilde{I}(y_1 L_T; y_2)) | \mathcal{G}_0^{(2)}]. \tag{4.2.35}$$

We use Monte-Carlo methods based on N realizations of L_T to approximate the expectations that define the functions $F(y_1, y_2)$ and $G(y_1, y_2)$. Recall that the process L_t is given by

$$L_t = \exp\left\{ -\int_0^t \frac{\mu_v^M(\Theta, K)}{\sigma} dW_v - \int_0^t \frac{\left(\mu_v^M(\Theta, K)\right)^2}{2\sigma^2} dv \right\}.$$
$$L_0 = 1.$$

With a time discretization over $[0, T]$ we use the Euler scheme with M time-steps to generate N realizations of L_T.

We need to solve the equation system

$$F(y_1, y_2) = X_0 \tag{4.2.36}$$
$$G(y_1, y_2) = e^{\gamma \epsilon^{(2)}}. \tag{4.2.37}$$

Gabih et al. [32] studied the property of the functions $F(y_1, y_2)$ and $G(y_1, y_2)$. Particularly the properties of their partial derivatives guarantee the uniqueness and existence of of the solutions $\hat{y}_1$ and $\hat{y}_2$. However, usually the equation system cannot be solved analytically. We need to apply iterative numerical solution procedures. Since the budget and risk functions F and G also involve expectations, we use a Monte Carlo based Newton method. We denote by $y_1^{(k)}$ and $y_2^{(k)}$ the approximation of the solutions in the kth iteration. We design the numerical scheme which is a nested Newton method as follows:

1. Let $y_2^{(0)} = 0$ and substitute it into (4.2.36) to find $F(y_1, 0) = X_0$ which can be solved for $y_1^{(0)}$. Set the initial guess of the iteration as $y_1^{(0)}$ and $y_2^{(0)}$.

2. Substitute the approximation of the kth iteration $y_1^{(k)}$ into (4.2.36) to form the equation:

$$F(y_1^{(k)}, y_2^{(k)}) = X_0 \tag{4.2.38}$$

Use Newton method to solve (4.2.38) for $y_2^{(k)}$.

3. Given the approximation of the kth iteration $y_1^{(k)}$ and $y_2^{(k)}$, we may construct the the approximation of the $k+1$th iteration by using (4.2.37).

4. Repeat step 2 and 3 until a sufficiently accurate value is reached.

Using the algorithm above we solve the equation system (4.2.36)-(4.2.37) numerically for $\hat{y}_1 = 1.348$ and $\hat{y}_2 = 1.651$. Recall from (4.2.19) that the optimal terminal wealth is given by

$$\hat{X}_T^{(2)} = \tilde{I}(\hat{y}_1 L_T; \hat{y}_2)$$

The optimal expected utility is computed as

$$\tilde{V}_0^{(2)} = \mathbb{E}[\frac{1}{2}\left(\hat{X}_T^{(2)}\right)^2].$$

Using Monte-Carlo methods based on N realizations of L_T we approximate

$$\tilde{V}_0^{(2)} \approx 41.07.$$

We next find the optimal strategy by solving the BSDE (4.2.22) numerically. we approximate the BSDE (4.2.22) by the following discretized BSDE

$$X_{t_{m+1}}^{(2)} = X_{t_m} + \left(\frac{\mu_{t_m}^M(\Theta, K)}{\sigma} Z_{t_m}\right)\Delta t + Z_{t_m}\Delta W_{t_m}, \quad t_0 \le t_m < t_M \tag{4.2.39}$$

$$X_{t_M}^{(2)} = \tilde{I}(\hat{y}_1 L_T; \hat{y}_2). \tag{4.2.40}$$

The BSDE (4.2.39)-(4.2.40) can be solved using the following recursive scheme (see [33])

$$Z_{t_m} = \frac{1}{\Delta t}\mathbb{E}[X_{t_{m+1}}^{(2)} \Delta W_{t_m} | \mathcal{G}_{t_m}^{(2)}], \tag{4.2.41}$$

$$X_{t_m}^{(2)} = \mathbb{E}[X_{t_{m+1}}^{(2)} | \mathcal{G}_{t_m}^{(2)}] - \left(\frac{\mu_{t_m}^M(\Theta, K)}{\sigma} Z_{t_m}\right)\Delta t. \tag{4.2.42}$$

We estimate the conditional expectation in (4.2.41) and (4.2.42) by the Monte-Carlo regression approach proposed by [33]. Note that the market price process S_t^M is not Markovian with respect to $(\mathbb{G}^{(2)}, \mathbb{P})$. We define the running minimum process $\tilde{S}_t^M = \inf\{S_v^M | 0 \leq v \leq t\}$ and note that the pair $(S_t^M, \tilde{S}_t^M)$ is Markovian with respect to $(\mathbb{G}^{(2)}, \mathbb{P})$. Hence we may choose the regression basis functions: $1, x, x^2, y, y^2$ and xy. By the regression method of [33] the conditional expectations in (3.5.4) and (3.5.5) can be estimated by

$$
\begin{aligned}
& c_1 + c_2(S_t^M - \alpha S_0) + c_3(S_t^M - \alpha S_0)^2 + c_4(\tilde{S}_t^M - \alpha S_0) + c_5(\tilde{S}_t^M - \alpha S_0)^2 \\
& + c_6(\tilde{S}_t^M - \alpha S_0)(\tilde{S}_t^M - \alpha S_0)
\end{aligned}
$$

for some coefficients $c_i, 1 \leq i \leq 6$.

We approximate the optimal strategy for fully informed investor $\tilde{\pi}^{(2)}$ as follows

$$
\tilde{\pi}_{t_m}^{(2)} = \frac{Z_{t_m}}{\sigma X_{t_m}^{(2)}}, \quad 0 \leq t_m \leq t_M.
$$

Following a similar procedure we may estimate the upper bound as $\epsilon_{max}^{(1)} = 77.631$ and the lower bound as $\epsilon_{min}^{(1)} = 18.175$ for partially informed investors. Then we choose the same risk constraint $\epsilon^{(1)} = 50$. Using Monte-Carlo simulation we approximate the optimal utility with risk constraint as $\tilde{V}_0^{(1)} \approx 25.77$. By solving the related BSDE (4.2.30) numerically we obtain the approximate optimal strategy for partially informed investors.

We illustrate the optimal strategy under risk constraint for both fully and partially informed investors in Figure 4.1, where the corresponding sample path of the risky asset price is same as the one in Figure 3.6 . In particular for the path of the asset price in Figure 4.1 liquidation occurs at time $t = 0.1540$. To make the difference between optimal strategies for these two types of investors more observable, we also present the difference process $\tilde{\pi}^{(1)} - \tilde{\pi}^{(2)}$ in Figure 4.1. Notice that the relatively significant difference occurs after the liquidation time, which means that different levels of information on the liquidation mechanism imposes the main impact on the optimal strategies. Since Figure 4.1 and Figure 3.6 rely on the same path of risky asset price, we may observe the impact of risk constraint on the optimal strategies.

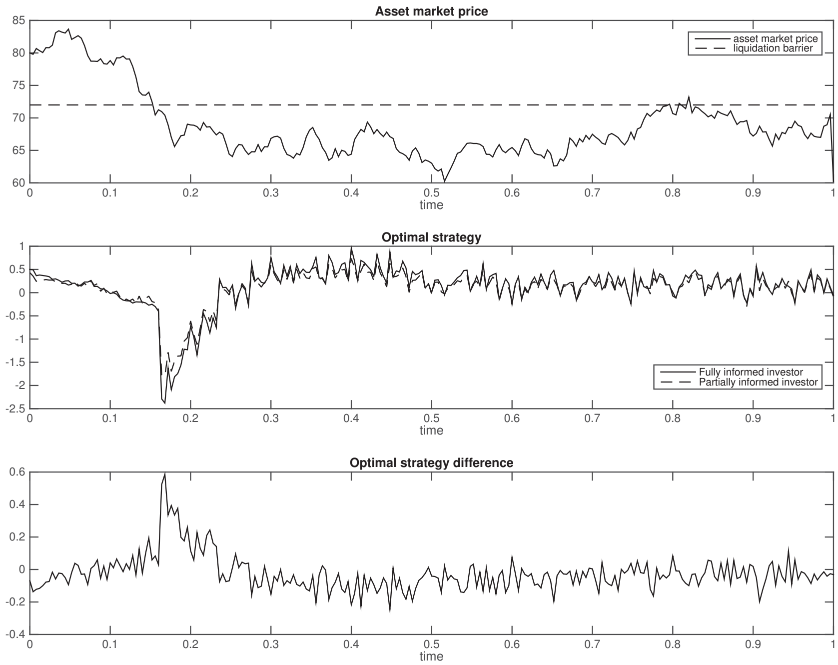

Figure 4.1: Optimal strategy

The optimal strategies under risk constraints take less extreme short positions than that without risk constraint for both fully and partially informed investors. This implies that imposing risk constraints on the portfolio produces essentially the same effect as imposing short selling constraints. Thus the risk constraint may serve as an alternative measure to short selling constraint, which can be imposed by the regulator on financial institutions to manage portfolio risk.

In Table 4.1 we compute the optimal utilities for both fully and partially informed investors by varying the risk constraint. The risk constraint becomes tighter as $\epsilon^{(2)}$ and $\epsilon^{(1)}$ decrease, and the corresponding optimal utilities become smaller as shown in Figure 4.2. In a special case where the risk constraints $\epsilon^{(2)}$ and $\epsilon^{(1)}$ are equal to 80 which exceeds the upper bounds, the optimal utilities coincide with those without risk constraints shown in Table 3.2. As we discussed before, this is the case where is the risk constraint is unbinding.

In this chapter, we imposed risk constraints on the optimal trading problem for

Risk constraint $\epsilon^{(2)}$ or $\epsilon^{(1)}$	Optimal utility	
	Fully informed	Partially informed
20	14.25	12.10
30	28.13	18.79
40	35.87	23.14
50	41.07	25.77
60	45.34	28.51
70	47.03	30.15
80	48.96	31.31

Table 4.1: Optimal utilities under risk constraints

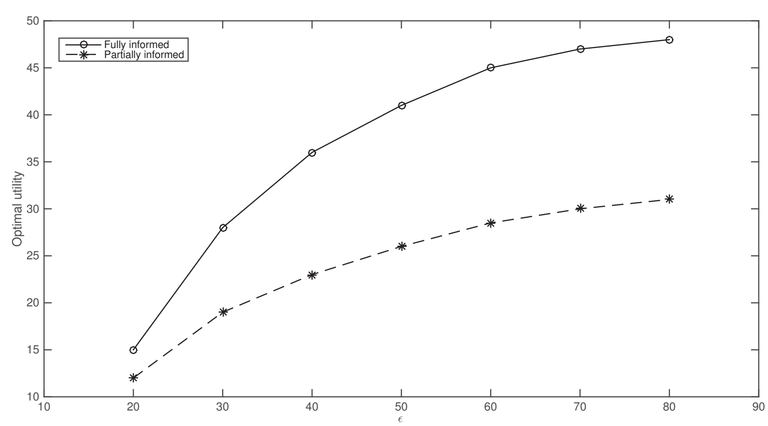

Figure 4.2: Optimal utilities under risk constraints

fully and partially informed investors. We showed that the optimal strategies are less risky than those without risk constraints. There is still an advantage to having additional information about the liquidation trigger and market impact functions. Figure 4.2 shows the utility difference between fully and partially informed investors for variant values of risk constraints. The gap becomes smaller as the risk constraint decreases which means the advantage of additional information is reduced for stricter risk constraints.

Chapter 5

Conclusions and future work

In summary this thesis consists of two projects: the optimal measure transformation problem and the optimal trading problem under asymmetric information. We next summarize our main contributions for both projects and present some possible directions for our future research.

5.1 The OMT problem

We summarize our contributions to the first project as below.

- We developed a new approach for pricing of bonds, futures and forwards based on the solution of an optimal measure transformation (OMT) problem.

- The solution for the OMT problem is characterized by a related FBSDE. We studied the explicit solutions for FBSDEs derived from the OMT problems which extend Hyndman [40] and Hyndman and Zhou [38].

- An equivenlence relationship has been discovered between the optimal stochastic control (OSC) approach in Gombani and Runggaldier [35] and our optimal measure transformation approach. As an advantage over the OSC approach, we extended to models with jumps or even to models for defaultable bonds.

- The OMT problem also provides a financial interpretation of the pricing problem in terms of maximization of returns subject to an entropy penalty term that quantifies financial risk.

In the following subsections we will discuss the future research for the OMT project and show some related methodology.

5.1.1 Zero-recovery defaultable bonds

Let us recall that the price of defautable bonds is given by

$$D(t,T) = E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} \cdot C_T|\mathcal{F}_t], \tag{5.1.1}$$

where C_T represents the random terminal payoff of the defautable bond. We find that the optimal measure turns out to be the martingale measure using the defautable bond as numéraire. The optimal measure is given by

$$\frac{d\mathbb{Q}^{\star}_{|\mathcal{F}_t}}{d\mathbb{P}_{|\mathcal{F}_t}} = \frac{e^{-\int_t^T r(X_s)ds} \cdot C_T}{E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} \cdot C_T|\mathcal{F}_t]}, \tag{5.1.2}$$

where C_T is the random terminal payoff. To avoid the denominator in 5.1.2 to be zero, we excluded the possibility of complete default by assuming that $\mathbb{P}(C_T = 0) = 0$.

Let us review the standard credit risk models for defaultable bonds following Jeanblanc et al. [44]. Default time is modeled by a random viable τ. The random payoff of the defautable bond is $1_{\{\tau > T\}}$. The survival distribution function is defined as

$$G_t = \mathbb{P}(\tau > t)$$

and the hazard function Γ_t is defined by $\Gamma_t = -\ln(G_t)$. Suppose G_t admits a derivative g, then we compute the derivative of Γ as

$$\gamma_t = \frac{-g_t}{G_t} = \lim_{h\to 0} \frac{1}{h\mathbb{P}(\tau > t)} \mathbb{P}(t < \tau \leq t+h) = \mathbb{P}(\tau \in dt|\tau > t)dt$$

Note that γ_t is the conditional density of the default time τ. Then from Jeanblanc et al. [44], the defaultable bond price is given by

$$D(t,T) = \exp\left(-\int_t^T (r_s + \gamma_s)ds\right).$$

Inspired by the above result, we consider a toy model. The complete default time, τ^c, is assumed to be the first jump time of a Cox process N_t with intensity λ_t. The process λ_t is assumed to be independent of the factor process X_t. The actual payoff of the defaultable bond, C_T, is given by

$$C_T = \tilde{C}_T 1_{\{\tau^c > T\}} = \tilde{C}_T 1_{\{N_T = 0\}} \tag{5.1.3}$$

where $\tilde{C}$ is a positive random variable that represents the terminal payoff in cases other than complete default. Let $\mathcal{F}_t^\lambda$ be the natural filtration generated by λ_t and define $\tilde{\mathcal{F}}_t = \mathcal{F}_t \vee \mathcal{F}_t^\lambda$. Then we rewrite the defautable bond price as

$$D(t,T) = E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} \cdot C_T | \tilde{\mathcal{F}}_t]. \tag{5.1.4}$$

Recall that the Cox process has the following property

$$\mathbb{P}(N_T = 0 | \mathcal{F}_t^\lambda) = e^{-\int_t^T \lambda_s ds} 1_{\{N_t = 0\}} \tag{5.1.5}$$

We compute

$$\begin{aligned} D(t,T) =& E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} \cdot C_T | \tilde{\mathcal{F}}_t], \\ =& E_{\mathbb{P}}[E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} \cdot \tilde{C}_T 1_{\{N_T=0\}} | \mathcal{F}_T \vee \mathcal{F}_t^\lambda] | \tilde{\mathcal{F}}_t], \\ =& E_{\mathbb{P}}[e^{-\int_t^T r(X_s)ds} \cdot \tilde{C}_T E_{\mathbb{P}}[1_{\{N_T=0\}} | \mathcal{F}_T \vee \mathcal{F}_t^\lambda] | \tilde{\mathcal{F}}_t], \\ =& 1_{\{\tau > t\}} E_{\mathbb{P}}[e^{-\int_t^T (r(X_s)+\lambda_s)ds} \cdot \tilde{C}_T | \tilde{\mathcal{F}}_t] \end{aligned} \tag{5.1.6}$$

Define $\tilde{r}_t = r(X_t) + \lambda_t$ and we rewrite (5.1.6) as

$$D(t,T) = E_{\mathbb{P}}[e^{-\int_t^T \tilde{r}_s ds} \cdot \tilde{C}_T | \tilde{\mathcal{F}}_t], \tag{5.1.7}$$

where $\tilde{r}_t$ is the complete default risk adjusted interest rate and $\tilde{C}_T$ is almost surely positive. Thus we can incorporate this special model with complete default into our model without complete default. In the future, we aim to incorporate complete default under a more general framework.

5.1.2 Application to other derivatives

Finally we may consider the pricing problem of other derivatives from the perspective of OMT problems. Besides the zero coupon bonds, we have already discussed the pricing problem of futures and forward contract. We may try to apply the OMT approach to other interest rate derivatives such as swaps, forward rate agreement and swaptions.

5.1.3 Wishart process

Richter [60] considered a class of quadratic FBSDEs and gave explicit solutions relying on a system of generalized Riccati equations. The author then applied the results to the problem of maximizing expected utility. This inspires us to explore the connection between our OMT problem with the utility maximization problem.

On the other hand, Richter [60] characterized the forward process X_t by Wishart processes. Recall that the Wishart process X_t, which was first studied in Bru [18], is a $d \times d$-matrix valued process satisfying

$$dX_t = (b + HX_t + X_tH')dt + \sqrt{X_t}dW_t\Sigma + \Sigma'(dW_t)'\sqrt{X_t}, \tag{5.1.8}$$

where b, H, Σ are $d \times d$ matrices and W_t is a $d \times d$-matrix Brownian motion. We may extend our short interest model to consider the Wishart process as the factors process.

5.2 The optimal trading problem

We summarize our contributions to the second project as below.

- We consider a optimal trading problem from a new perspective of small investors trading against a large investor.
- We consider the portfolio optimization problem for three types of investors with different level of information about the liquidation trigger mechanism and the market impact.
- We compute optimal utilities for all three types of investors. The utility differences quantify the values of different levels of information.

- In case of logarithmic utility, we find the closed-form optimal strategies for all three types of investors. In the case of power utility, the optimal strategy is essentially a Merton type strategy corrected by an extra component which is determined by a FBSDE.

- The risky trading strategies taken by fully and partially informed investors can be restrained by imposing risk constraints on the optimal trading problem for fully and partially informed investors.

- We presented some numerical results using Monte-Carlo simulation method.

In the following subsections we will discuss the future research for the optimal trading project and present related methodology possibly applicable.

5.2.1 More general market impact modeling

In our current model liquidation is triggered once the asset price hits the threshold α. The market impact starts to take effect after the liquidation triggered time τ with varying impact magnitude depending on the length of time after τ, which is modeled by the function $g(t; K, \Theta)$ of the form

$$g(t; K, \Theta) = 1 - \frac{Kt}{\Theta} e^{1-\frac{t}{\Theta}} \tag{5.2.1}$$

Notice that the impact function in (5.2.1) only characterizes the negative market impact incurred by liquidation of large long position. It would be more flexible if we also incorporate the positive market impact resulted from the covering a large short position. We may define a positive impact function as below

$$h(t; K', \Theta') = -1 + \frac{K't}{\Theta'} e^{1-\frac{t}{\Theta'}}. \tag{5.2.2}$$

Combine (5.2.1) and (5.2.2) to compose the general impact function

$$\tilde{g}(t; K, K', \Theta, \Theta') = 1_{\{p=1\}} g(t; K, \Theta) + 1_{\{p=-1\}} h(t; K', \Theta') \tag{5.2.3}$$

where p is a state indicator variable with $p = 1$ standing for long position liquidation and $p = -1$ standing for short position liquidation.

The choice of the smooth function $g(t; K, \Theta)$, rather than a instant jump to model the liquidation impact, was a compromise to avoid technical problems. However it is worth looking at the limit situation of the impact function as the realized value of Θ approaches zero. When the realized value of Θ is rather small, the shape of the impact function $g(t; K, \Theta)$ is very steep and hence is a good enough approximation of an instant jump. Besides theoretical analysis of the limit situation, we need to pay attention to the numerical implementation since rather small values of Θ cause bigger approximation errors. Note that the impact function $g(t; K, \Theta)$ defined in (3.1.4) decreases very quickly after zero when Θ is very small. Rather small values of Θ not only generate bigger discretization error for the numerical integration where the integrand involves the function $g(t; K, \Theta)$, but also results in larger variance for the Monte Carlo simulation. One direct solution is to increase the number of discretization points for numerical integration and the number of Monte-Carlo simulations. However, both methods would cost much running time for rather small Θ to guarantee a certain level of accuracy. Actually we can refine the partition especially over the interval closer to zero instead of the whole range of Θ. This discretization technique would promote the accuracy with less cost of running time. On the other hand, we may apply some proper variance reduction technique for Monte-Carlo simulation to improve the implementation efficiency for rather small Θ.

We are mainly concerned with the optimal trading problem from the perspective of market participants whose transactions do not have market impact on the asset price. However from Figure (3.6) we can observe that the optimal strategy for fully informed investors involves short selling of rather large volume, so it is more realistic to incorporate a standard market impact which depends on the volume of transaction. We may refer to Almgren and Chriss [4] for standard market impact model. On the other hand, in real-life markets there exists short selling constraints and transaction fees which would prevent the fully informed investors shorting sell as much as they want. In the future we may study the optimization problem with strategy constraints and transition fees.

5.2.2 Occupation time

We next consider an alternative model in which the liquidation impact is modeled as an instant jump. Following Jiao and Pham [45], Hillairet and Jiao [36] we suppose the asset admits a jump at the liquidation time, i.e.

$$S_\tau = S_{\tau-}(1-\gamma_\tau)$$

where $\gamma < 1$ represents the proportional jump at liquidation. Liquidation is triggered if the asset price S_t crosses down the barrier α and stays below longer than a certain amount of time β. Mathematically, default time τ is defined as

$$\tau = \inf\{t \geq 0 | \Lambda_t \geq \beta\}, \tag{5.2.4}$$

where Λ_t is the occupation time of S_t staying below α over time period $[0,t]$

$$\Lambda_t = \int_0^t 1_{S_u<\alpha} du. \tag{5.2.5}$$

We distinguish three types of investors as below:

- Fully informed investors: both β and α are known values.
- Partially informed investors: α is known, β is a random variable.
- Unknown investors: both β and α are unknown random variables.

Jiao and Pham [45] consider a similar optimization problem for the uninformed investors using the theory of progressive enlargement of filtration. Hillairet and Jiao [36] studied a similar optimization problem for the fully informed investors using the theory of initial enlargement of filtration. We review the following so called density hypothesis for the liquidation time τ which plays a key role.

For any $t \in [0,T]$, the conditional distribution of τ given $\mathcal{F}_t$ admits a density with respect to Lebesgue measure, i.e.

$$\mathbb{P}(\tau \in d\theta | \mathcal{F}_t) = \alpha_t(\theta) d\theta.$$

where $\alpha_t(\theta)$ is positive. In our framework if we suppose $\mathbb{P}(\alpha \in dx|\mathcal{F}_t) = \varphi_t(x)dx$, we can compute the conditional density of τ

$$\mathbb{P}[\tau \in d\theta|\mathcal{F}_t] = 1_{\{S_\theta < \alpha\}}\varphi_t(\Lambda_\theta)d\theta. \tag{5.2.6}$$

Since $1_{\{S_\theta < \alpha\}}\varphi_t(\Lambda_\theta)$ is not always positive, the density hypothesis is not satisfied in our model. Therefore, further research on enlargement of filtrations when the density hypothesis is not satisfied, or a new approach, would be required to resolve this technical hurdle.

Bibliography

[1] Frédéric Abergel, Jean-Philippe Bouchaud, Thierry Foucault, Charles-Albert Lehalle, and Mathieu Rosenbaum. *Market microstructure: confronting many viewpoints*. John Wiley & Sons, 2012.

[2] Dong-Hyun Ahn, Robert F Dittmar, and A Ronald Gallant. Quadratic term structure models: theory and evidence. *The Review of Financial Studies*, 15(1): 243–288, 2002.

[3] Franklin Allen and Douglas Gale. Stock-price manipulation. *The Review of Financial Studies*, 5(3):503–529, 1992.

[4] Robert Almgren and Neil Chriss. Value under liquidation. *Risk*, 12(12):61–63, 1999.

[5] Edward Altman, Andrea Resti, and Andrea Sironi. Default recovery rates in credit risk modelling: a review of the literature and empirical evidence. *Economic Notes*, 33(2):183–208, 2004.

[6] Jürgen Amendinger. Martingale representation theorems for initially enlarged filtrations. *Stochastic Processes and their Applications*, 89(1):101–116, 2000.

[7] Jürgen Amendinger, Peter Imkeller, and Martin Schweizer. Additional logarithmic utility of an insider. *Stochastic Processes and their Applications*, 75(2): 263–286, 1998.

[8] Jürgen Amendinger, Dirk Becherer, and Martin Schweizer. A monetary value

for initial information in portfolio optimization. *Finance and Stochastics*, 7(1): 29–46, 2003.

[9] Stefan Ankirchner, Christophette Blanchet-Scalliet, and Anne Eyraud-Loisel. Optimal portfolio liquidation with additional information. *Mathematics and Financial Economics*, 10:1–14, 2016.

[10] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[11] Kerry Back. Insider trading in continuous time. *The Review of Financial Studies*, 5(3):387–409, 1992.

[12] Fabrice Baudoin. Modeling anticipations on financial markets. *Lecture Notes in Mathematics*, pages 43–94, 2003.

[13] Tomasz R Bielecki and Marek Rutkowski. *Credit risk: modeling, valuation and hedging*. Springer-Verlag, Berlin-Heidelberg-New York, 2002.

[14] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[15] Tomas Björk. *Arbitrage theory in continuous time*. Oxford University Press, Oxford, 2004.

[16] Tomas Björk and Camilla Landén. On the term structure of futures and forward prices. In *Mathematical Finance-Bachelier Congress 2000*, pages 111–149. Springer, 2002.

[17] Tomas Björk, Mark HA Davis, and Camilla Landén. Optimal investment under partial information. *Mathematical Methods of Operations Research*, 71(2):371–399, 2010.

[18] Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4): 725–751, 1991.

[19] Jacques F Carriere. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: Mathematics and Economics*, 19(1):19–30, 1996.

[20] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[21] Qiang Dai and Kenneth J Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.

[22] Paolo Dai Pra, Lorenzo Meneghini, and Wolfgang J Runggaldier. Connections between stochastic control and dynamic games. *Mathematics of Control, Signals and Systems*, 9(4):303–326, 1996.

[23] Łukasz Delong. *Backward stochastic differential equations with jumps and their actuarial and financial applications*. Springer, London-Heidelberg-New York-Dordrecht, 2013.

[24] Darrell Duffie and Rui Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 1996.

[25] Darrell Duffie and Kenneth J Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720, 1999.

[26] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[27] Darrell Duffie, Damir Filipović, and Walter Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003.

[28] Robert J Elliott and Monique Jeanblanc. Incomplete markets with jumps and informed agents. *Mathematical Methods of Operations Research*, 50(3):475–492, 1999.

[29] Robert J. Elliott and John van der Hoek. Stochastic flows and the forward measure. *Finance and Stochastics*, 5:511–525, 2001.

[30] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

[31] Masatoshi Fujisaki, Gopinath Kallianpur, and Hiroshi Kunita. Stochastic differential equations for the non linear filtering problem. *Osaka Journal of Mathematics*, 5(9):19–40, 1972.

[32] Abdelali Gabih, Jörn Sass, and Ralf Wunderlich. Utility maximization under bounded expected loss. *Stochastic Models*, 25(3):375–407, 2009.

[33] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin. A regression-based monte carlo method to solve backward stochastic differential equations. *Annals of Applied Probability*, 15(3):2172–2202, 2005.

[34] Selim Gökay, Alexandre F Roch, and H Mete Soner. *Liquidity models in continuous and discrete time*. Springer, 2011.

[35] Andrea Gombani and Wolfgang J Runggaldier. Arbitrage-free multifactor term structure models: A theory based on stochastic control. *Mathematical Finance*, 23(4):659–686, 2013.

[36] Caroline Hillairet and Ying Jiao. Portfolio optimization with insider's initial information and counterparty risk. *Finance and Stochastics*, 19(1):109–134, 2015.

[37] John C Hull and Alan D White. Numerical procedures for implementing term structure models ii: Two-factor models. *The Journal of Derivatives*, 2(2):37–48, 1994.

[38] Cody Hyndman and Xinghua Zhou. Explicit solutions of quadratic fbsdes arising from quadratic term structure models. *Stochastic Analysis and Applications*, 33 (3):464–492, 2015.

[39] Cody Blaine Hyndman. Forward–backward SDEs and the CIR model. *Statistics & Probability Letters*, 77(17):1676–1682, 2007.

[40] Cody Blaine Hyndman. A forward–backward SDE approach to affine models. *Mathematics and Financial Economics*, 2(2):107–128, 2009.

[41] Kiyosi Itô. On a stochastic integral equation. *Proceedings of the Japan Academy*, 22(1-4):32–35, 1946.

[42] Robert A Jarrow. Market manipulation, bubbles, corners, and short squeezes. *Journal of Financial and Quantitative Analysis*, 27(3):311–336, 1992.

[43] Robert A Jarrow. Derivative security markets, market manipulation, and option pricing theory. *Journal of Financial and Quantitative Analysis*, 29(02):241–261, 1994.

[44] Monique Jeanblanc, Marc Yor, and Marc Chesney. *Mathematical Methods for Financial Markets*. Springer Finance. Springer-Verlag London, Ltd., London, 2009.

[45] Ying Jiao and Huyên Pham. Optimal investment with counterparty risk: a default-density model approach. *Finance and Stochastics*, 15(4):725–753, 2011.

[46] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer-Verlag, New York, 1998.

[47] Ioannis Karatzas and Xiaoliang Zhao. Bayesian adaptive portfolio optimization. *Option pricing, interest rates and risk management*, pages 632–669, 2001.

[48] Ioannis Karatzas, Steven E Shreve, I Karatzas, and Steven E Shreve. *Methods of mathematical finance*, volume 39. Springer, 1998.

[49] Robert Kissell and Morton Glantz. *Optimal trading strategies: quantitative approaches for managing market impact and trading risk*. Amacom, 2003.

[50] Magdalena Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*, 28(2): 558–602, 2000.

[51] Albert S Kyle. Continuous auctions and insider trading. *Econometrica*, 52(6): 1315–1335, 1985.

[52] Wolfgang Lemke. *Term structure modeling and estimation in a state space framework*, volume 565. Springer Berlin Heidelberg, 2006.

[53] Jingya Li, Adam Metzler, and R. Mark Reesor. A contingent capital bond study: Short-selling incentives near conversion to equity. *Working Paper*, 2014.

[54] Robert Liptser and Albert N Shiryaev. *Statistics of random processes: I. general theory*, volume 5. Springer, 2013.

[55] Francis A Longstaff and Eduardo S Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, 47 (4):1259–1282, 1992.

[56] Francis A Longstaff and Eduardo S Schwartz. Valuing american options by simulation: a simple least-squares approach. *The Review of Financial Studies*, 14(1):113–147, 2001.

[57] Jin Ma and Jiongmin Yong. *Forward-backward stochastic differential equations and their applications*. Springer-Verlag, Berlin, 1999.

[58] Robert C Merton. On the pricing of corporate debt: The risk structure of interest rates*. *The Journal of Finance*, 29(2):449–470, 1974.

[59] Etienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[60] Anja Richter. Explicit solutions to quadratic bsdes and applications to utility maximization in multivariate affine stochastic volatility models. *Stochastic Processes and their Applications*, 124(11):3578–3611, 2014.

[61] Elliott Robert. J, Geman Helyette, and Korkie Bob M. Portfolio optimization and contingent claim pricing with differential information. *Stochastics: An International Journal of Probability and Stochastic Processes*, 60(3-4):185–203, 1997.

[62] Steven E Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, Berlin, 2004.

[63] Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

[64] Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1): 155–167, 1971.

Appendix

A Solvability of SDEs and BSDEs

This appendix reviews fundamental results on existence and uniqueness of solutions to stochastic differential equations (SDEs) and backward stochastic differential equations (BSDEs). The numerical solution approach for BSDEs is also covered.

A.1 Existence and uniqueness of strong solutions of SDEs

On a filtered probability space $(\Omega, \mathcal{A}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ we consider a stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad 0 \leq t \leq T, \ X_0 = \xi \tag{A.1}$$

where ξ is an $\mathcal{F}_0$-measurable random variable and W is a $\mathbf{R}^m$ valued Brownian motion w.r.t $(\mathcal{F}_t, \mathbb{P})$. Let $T > 0$ and $\mu(\cdot,\cdot) : [0, T] \times \mathbf{R}^n \to \mathbf{R}^n, \sigma(\cdot,\cdot) : [0, T] \times \mathbf{R}^n \to \mathbf{R}^{n\times m}$ be measurable functions.

Let us first review the definitions of existence and uniqueness of a strong solution of the SDE (A.1) (see Karatzas and Shreve [46]).

Definition A.1. *A strong solution of the SDE* (A.1) *on the given probability space* $(\Omega, \mathbb{F}, \mathbb{P})$ *and with respect to the fixed Brownian motion* W *and initial condition* ξ, *is a process* $X = \{X_t; 0 \leq t < \infty\}$ *with continuous sample paths and with the following properties:*

(i) X *is adapted to the filtration* $\{\mathcal{F}_t\}_{t\geq 0}$,

(ii) $\mathbb{P}(X_0 = \xi) = 1$,

(iii) $\mathbb{P}(\int_0^t \{|\mu_i(s, X_s)| + \sigma_{ij}^2(s, X_s)\} ds < \infty) = 1$ *holds for every* $1 \leq i \leq n$, $1 \leq j \leq m$ *and* $0 \leq t < \infty$, *and*

(iv) the integral version of (A.1)

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad 0 \leq t \leq T,$$

holds almost surely.

Definition A.2. *Suppose X and $\tilde{X}$ are two strong solutions of* (A.1) *relative to W with initial condition ξ, then the strong uniqueness holds if*

$$\mathbb{P}(X_t = \tilde{X}_t; 0 \leq t < \infty) = 1.$$

Itô [41] first proposed the Lipschitz conditions that guarantee the existence and uniqueness of a solution to the SDE (A.1), which is shown in the following theorem.

Theorem A.1. *Suppose that the coefficients $\mu(t, x)$, $\sigma(t, x)$ satisfy the global Lipschitz and linear growth conditions*

$$||\mu(t, x) - \mu(t, y)|| + ||\sigma(t, x) - \sigma(t, y)|| \leq K||x - y||, \quad x, y \in \mathbf{R}^n, t \in [0, T], \quad \text{(A.2)}$$

$$||\mu(t, x)||^2 + ||\sigma(t, x)||^2 \leq K^2(1 + ||x||^2), \quad x \in \mathbf{R}^n, t \in [0, T]. \quad \text{(A.3)}$$

Let ξ be an $\mathbf{R}^n$-valued random vector, independent of the m-dimensional Brownian motion W, and with finite second moment:

$$\mathbb{E}[||\xi||^2] < \infty.$$

Then there exists a continuous, adapted process X which is a strong solution of (A.1) *relative to W, with initial condition ξ.*

In the one-dimensional case, Yamada and Watanabe [64] proposed the Yamada-Watanabe condition which relaxes the Lipschitz condition.

Proposition A.1. *Suppose the coefficients of the one-dimensional equation*

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad 0 \le t \le T, \ X_0 = \xi \tag{A.4}$$

satisfies the conditions

$$|\mu(t, x) - \mu(t, y)| \le K|x - y|,$$
$$|\sigma(t, x) - \sigma(t, y)| \le h(|x - y|),$$

for every $0 \le t < \infty$ *and* $x \in \mathbf{R}$, $y \in \mathbf{R}$, *where* K *is a positive constant and* $h : [0, \infty) \to [0, \infty)$ *is a strictly increasing function with* $h(0) = 0$ *and*

$$\int_{(0,\epsilon)} h^{-2}(v)dv = \infty, \quad \forall \epsilon > 0.$$

Then strong uniqueness holds for the equation (A.4).

The Yamada-Watanabe condition can be used to show the strong uniqueness of solutions of a SDE with certain non-Lipschitz coefficients. For example, we may apply Proposition A.1 to the Cox-Ingersoll-Ross (CIR) model, which describes the stochastic evolution of interest rate r_t by the SDE

$$dr_t = \alpha(\mu - r_t)dt + \sigma\sqrt{r_t}dW_t, \quad t \ge 0, \tag{A.5}$$

with $r_0 > 0$, where α, μ with $\alpha\mu \ge 0$ and σ denote real constants. Notice that the volatility coefficient of (A.5) is Hölder continuous, i.e.,

$$|\sigma(t, x) - \sigma(t, y)| \le C\sqrt{|x - y|}, \quad \forall\ 0 \le t < \infty \text{ and } x, y \in \mathbf{R},$$

where C is a positive constant. Thus the SDE (A.5) satisfies the Yamada-Watanabe condition, which implies that the strong uniqueness of the solution of (A.5) holds.

In affine term structure models (ATSMs) the interest rate is driven by a multi-dimensional affine diffusion process

$$dX_t = (AX_t + \tilde{B})dt + S\text{diag}\left(\sqrt{\alpha_i + \beta_i X_t}\right)dW_t \tag{A.6}$$

where W is an n-dimensional Brownian motion, A is an $(n \times n)$-matrix of scalars, $\tilde{B}$ is an $(n \times 1)$-vector of scalars, for each each $i \in \{1, \ldots, n\}$ the α_i are scalars, for each

$i \in \{1, \dots, n\}$ the $\beta_i = (\beta_{i1}, \dots, \beta_{in})$ are $1 \times n$-vectors taking values in $\mathbf{R}^n$, and S is a non-singular $n \times n$-matrix. Regarding the strong uniqueness of solutions of (A.6), Duffie and Kan [24] proved the following proposition by using Proposition A.1.

Proposition A.2. *Suppose the following conditions:*

(A-I) for all x such that $\alpha_i + \beta_i x = 0, \beta_i(Ax + \tilde{B}) > \beta_i S S' \beta_i'/2$,

(A-II) for all j, if $(\beta_i S)_j \neq 0$, then $\alpha_i + \beta_i x = \alpha_j + \beta_j x$

are satisfied then there exits a unique strong solution X_t to the SDE (A.6) *that takes values in D which is the open domain implied by nonnegative volatilities, i.e.,*

$$D = \{x \in \mathbf{R}^n : \quad \alpha_i + \beta_i x > 0, \quad i \in \{1, \dots, n\}\}.$$

A.2 Existence and uniqueness of solutions of BSDEs

Backward stochastic differential equations (BSDEs) were introduced by Bismut [14] and Pardoux and Peng [59]. Let $(\Omega, \mathcal{A}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. Let W be an n-dimensional standard Brownian motion, and assume $\mathbb{F}$ is the filtration generated by W. BSDEs are of the following form

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t, \quad Y_T = \xi \tag{A.7}$$

or equivalently

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \tag{A.8}$$

where $Y_t \in \mathbf{R}^d$, ξ is $\mathcal{F}_T$-measurable, and f is $\mathcal{P} \otimes \mathcal{B}^d \otimes \mathcal{B}^{d\times n}$-measurable. $\mathcal{P}$ is the predictable σ-algebra, and $\mathcal{B}^d$ is the Borel σ-algebra on $\mathbf{R}^d$. The function f is called the generator of the BSDE. A solution is a pair (Y, Z) such that Y is continuous and adapted, and Z is predictable and satisfies $\int_0^T |Z_s|^2 ds < \infty$. We denote by $L_T^2(\mathbf{R}^d)$ the space of $\mathcal{F}_T$-measurable random variable ξ and by $H_T^2(\mathbf{R}^d)$ the space of predictable processes Y.

Pardoux and Peng [59] proved the following theorem on the existence and uniqueness of solutions to the BSDE (A.7).

Theorem A.2. *Suppose that $\xi \in L^2_T(\mathbf{R}^d)$, $f(\cdot, 0, 0) \in H^2_T(\mathbf{R}^d)$, and f is uniformly Lipschitz; i.e., there exists $C > 0$ such that $d\mathbb{P} \otimes dt$ a.s.*

$$|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \leq C(|y_1 - y_2| + |z_1 - z_2|) \quad \forall (y_1, z_1), \forall (y_2, z_2).$$

Then there exists a unique pair (Y, Z) which solves the BSDE (A.7).

The classical result on the existence of a solution to a quadratic BSDE driven by a Brownian motion is due to Kobylanski [50]. We show a simple quadratic BSDE as an example below:

$$Y_t = \xi + \int_t^T \frac{1}{2}|Z_s|^2 ds - \int_t^T Z_s dW_s.$$

Since the generator is of quadratic growth in Z, which does not satisfy the Lipschitz condition. In order to find conditions for existence and uniqueness of solutions of quadratic BSDEs Kobylanski [50] defined three types of conditions.

(i) The coefficient f is said to satisfy condition (H1) with α_0, β_0, b, c if for all $(t, y, z) \in \mathbf{R}^+ \times \mathbf{R} \times \mathbf{R}^d$,

$$f(t, y, z) = a_0(t, y, z)y + f_0(t, y, z),$$

with

$$\begin{aligned} &\beta_0 \leq a_0(t, y, z) \leq \alpha_0, \quad a.s. \\ &|f_0(t, y, z) \leq b + c(|y|)|z|^2, \quad a.s. \end{aligned} \tag{H1}$$

(ii) The coefficient f is said to satisfy condition $(H2)$ on $[-M, M]$ with l, k and C if for all $t \in (0, \infty), y \in [-M, M], z \in \mathbf{R}^d$,

$$\begin{aligned} &|f(t, y, z)| \leq l(t) + C|z|^2, \quad a.s. \\ &|\frac{\partial f}{\partial z}(t, y, z)| \leq k(t) + C|z| \quad a.s. \end{aligned} \tag{H2}$$

(iii) the coefficient f is said to satisfy condition $(H3)$ with c_ϵ and ϵ if for all $t \in (0, \infty), y \in \mathbf{R}, z \in \mathbf{R}$,

$$\frac{\partial f}{\partial y}(t, y, z) \leq l_\epsilon(t) + \epsilon|z|^2 \quad a.s. \tag{H3}$$

Theorem A.3. *If the generator f of* (A.7) *satisfies (H1) with $\alpha_0, \beta_0, b \in \mathbf{R}$ and $c : (0, \infty) \to (0, \infty)$ be a continuous increasing function, then the BSDE* (A.7) *has at least one solution. If for all $\epsilon, M > 0$ there exists $l, l_\epsilon, C \in \mathbf{R}$ such that f satisfies condition $(H2)$ on $[-M, M]$ with l, k, C and satisfies condition $(H3)$ on $[-M, M]$ with l_ϵ and ϵ, then the BSDE* (A.7) *has a unique solution.*

Before we review the results on solutions of BSDEs with jumps we need introduce a random measure (refer to Delong [23, Chapter 2]).

Definition A.3. *A function N defined on $\Omega \times [0, T] \times \mathbf{R}$ is called a random measure if*

(i) for any $\omega \in \Omega, N(\omega, \cdot)$ is a σ-finite measure on $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbf{R})$,

(ii) for any $A \in \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbf{R})$, $N(\cdot, A)$ is a random variable on $(\Omega, \mathbb{F}, \mathbb{P})$.

Next we introduce a predictable compensator of a random measure.

Definition A.4. *For a random measure N we define*

$$\mathbb{E}_N(A) = \mathbb{E}\left[\int_{[0,T]\times\mathbf{R}} 1_A(\omega, t, z) N(\omega, dt, dz)\right], \quad A \in \mathbb{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbf{R}). \quad \text{(A.9)}$$

If there exists an $\mathbb{F}$-predictable random measure v such that

(i) $\mathbb{E}_v$ is a σ-finite measure on $\mathcal{P} \otimes \mathcal{B}(\mathbf{R})$,

(ii) the measures $\mathbb{E}_N$ and $\mathbb{E}_v$ are identical on $\mathcal{P} \otimes \mathcal{B}(\mathbf{R})$,

then we say that the random measure N has a compensator v.

Given the compensator v of a random measure N, we can define the compensated random measure

$$\tilde{N}(\omega, dt, dz) = N(\omega, dt, dz) - v(\omega, dt, dz).$$

Random measures are usually related to jumps of discontinuous processes. We assume the random measure N is an integer-valued random measure with the compensator

$$v(dt, dz) = \mathcal{Q}(t, dz)\eta(t)dt,$$

where $\eta : \Omega \times [0,T] \to [0,\infty)$ is a predictable process, and $\mathcal{Q}$ is a kernel from $(\Omega \times [0,T], \mathcal{P})$ into $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ satisfying

$$\int_0^T \int_{\mathbf{R}} z^2 \mathcal{Q}(t,dz)\eta(t)dt < \infty.$$

Delong [23] investigated BSDEs driven by a Brownian motion and a compensated random measure

$$Y_t = \xi + \int_t^T f(s, Y_{s-}, X_s, U_s(\cdot))ds - \int_t^T Z_s dW_s - \int_t^T \int_{\mathbf{R}} U_s(z)\tilde{N}(ds,dz). \quad \text{(A.10)}$$

The following theorem presents the existence and uniqueness of solutions for BSDEs with jumps (see Delong [23]).

Theorem A.4. *Assume the terminal value ξ and the generator f satisfy the following conditions:*

(i) the terminal value $\xi \in L^2(\mathbf{R})$,

(ii) the generator $f : \Omega \times [0,T] \times \mathbf{R} \times \mathbf{R} \times L^2_{\mathcal{Q}}(\mathbf{R}) \to \mathbf{R}$ is predictable and Lipschitz continuous in the sense that

$$\begin{aligned} &|f(\omega,t,y,z,u) - f(\omega,t,y',z',u')|^2 \\ &\qquad \le K\left(|y-y'|^2 + |z-z'|^2 + \int_{\mathbf{R}} |u(x)-u'(x)|^2 \mathcal{Q}(t,dx)\eta(t)\right), \end{aligned}$$

(iii) $\mathbb{E}[\int_0^T |f(t,0,0,0)|^2 dt] < \infty$.

Then the BSDE (A.10) *has a unique solution.*

A.3 Numerical methods for BSDEs

BSDEs usually do not have closed form solutions in which cases numerical solution methods can be applied. We review the regression based on Monte-Carlo approach proposed by Gobet et al. [33]. Consider a BSDE

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t, \quad Y_T = \xi. \quad \text{(A.11)}$$

can be discretized using the algorithm. Given a partition $\pi : 0 = t_0 < \cdots < t_n = T$ of the interval $[0, T]$, we write the Euler discretization of the BSDE (A.11) as

$$Y^\pi_{t_i} - Y^\pi_{t_{i-1}} = -f(t_{i-1}, Y^\pi_{t_{i-1}}, Z^\pi_{t_{i-1}}) + Z^\pi_{t_{i-1}} \cdot (W_{t_i} - W_{t_{i-1}}) \tag{A.12}$$

together with the terminal condition $Y^\pi_{t_n} = g(W_{t_n})$. A backward induction scheme is obtained by taking conditional expectations as follows:

- $Y^\pi_{t_n} = g(W_{t_n})$,
- $Z^\pi_{t_{i-1}} = (t_i - t_{i-1})^{-1}\mathbb{E}[Y^\pi_{t_i}(W_{t_i} - W_{t_{i-1}})|\mathcal{F}_{t_{i-1}}]$,
- $Y^\pi_{t_{i-1}} = \mathbb{E}[Y^\pi_{t_i}|\mathcal{F}_{t_{i-1}}] + f(t_{i-1}, Y^\pi_{t_{i-1}}, Z^\pi_{t_{i-1}})(t_i - t_{i-1})$

for all $i = 1, \cdots, n$. Then the conditional expectations involved in the above scheme reduce to the regression of $Y^\pi_{t_i}$ and $Y^\pi_{t_i}(W_{t_i} - W_{t_{i-1}})$ on the random variable $W_{t_{i-1}}$. For instance one can use the classical kernel regression estimation (see Carriere [19]) or the basis projection method studied by Longstaff and Schwartz [56].

B Optimal measure transformation

This appendix provides technical proofs for the optimal measure transformation part.

B.1 Riccati equations

We prove the existence and uniqueness of solutions to the Riccati equations based on Gombani and Runggaldier [35, Theorem B.1].

Proposition B.1. *The following decoupled Riccati equations admit a pair of unique explicit solutions.*

$$\dot{q}_s + q_s A + A' q_s + \frac{1}{2}(q'_s + q_s)\Sigma\Sigma'(q'_s + q_s) - Q = 0_{n\times n}, \quad s \in [0, T] \tag{B.1}$$

$$\dot{u}_s + u_s A + B'(q'_s + q_s) + u_s \Sigma\Sigma'(q'_s + q_s) - R' = 0_{1\times n}, \quad s \in [0, T] \tag{B.2}$$

$$q_T = 0_{n\times n}, \quad u_T = 0_{1\times n}. \tag{B.3}$$

Proof. We first prove equation (B.1) admits a unique explicit solution. By taking the transpose of both sides of equation (B.1) we find

$$\dot{q}_s' + A'q_s' + q_s'A + \frac{1}{2}(q_s' + q_s)\Sigma\Sigma'(q_s' + q_s) - Q = 0_{n\times n}, \tag{B.4}$$

Adding equation (B.4) to equation (B.1) gives

$$(\dot{q}_s + \dot{q}_s') + A'(q_s + q_s') + (q_s + q_s')A + (q_s' + q_s)\Sigma\Sigma'(q_s' + q_s) - 2Q = 0_{n\times n}, \tag{B.5}$$

and subtracting equation (B.4) from equation (B.1) to find

$$(\dot{q}_s - \dot{q}_s') + A'(q_s - q_s') + (q_s - q_s')A = 0. \tag{B.6}$$

Define

$$U_s = \frac{q_s' + q_s}{2}, \quad V_s = \frac{q_s - q_s'}{2},$$

and by the terminal condition (B.3) we have

$$U_T = 0_{n\times n}, \quad V_T = 0.$$

Hence U_s and V_s satisfy the following equations

$$\dot{U}_s + A'U_s + U_sA + U_s\Sigma\Sigma'U_s - Q = 0_{n\times n}, \tag{B.7}$$

$$\dot{V}_s + A'V_s + V_sA = 0_{n\times n}, \tag{B.8}$$

$$U_T = 0_{n\times n}, \quad V_T = 0_{n\times n}. \tag{B.9}$$

By Gombani and Runggaldier [35, Theorem B.1] there exists a pair of unique (U_s, V_s) satisfying equations (B.7)-(B.9). Moreover, we actually have $V_s = 0_{n\times n}$ which means $q_s = q_s'$, so q_s is symmetric, and $q_s = U_s$.

After we obtain the solution q_s, equation (B.2) is simplified as an ODE for u_s, which can be solved explicitly as in Gombani and Runggaldier [35, Corollary B.3].

□

B.2 Solutions of quadratic FBSDEs

We provide proof for Theorem 2.4.2.

Proof. The solvability of FBSDE (2.4.20)-(2.4.21) is guaranteed by Kobylanski [50]. Apply Itô's formula to the function $f(t,x) = -(x'q_t x + u_t x + p_t)$, q_t and u_t are solutions to (2.4.22)-(2.4.23) and p_t satisfies (2.4.27). Let $Y_t = f(t, X_t)$, then we have

$$
\begin{aligned}
dY_t = -\Big\{ & X_t'\dot{q}_t X_t + \dot{u}_t X_t + \big(k - u_t B - \frac{1}{2}\mathrm{tr}(q_t + q_t')\Sigma\Sigma' - \frac{1}{2}u_t\Sigma\Sigma' u_t' + \frac{1}{2}z_t z_t'\big) \\
& + \big(X_t'A' + B' + X'(q_t + q_t')\Sigma\Sigma' + u_t\Sigma\Sigma' + z_t\Sigma'\big)q_t X_t \\
& + X_t' q_t\big(AX_t + B + \Sigma\Sigma'(q_t + q_t')X_t + \Sigma\Sigma' u_t' + \Sigma z_t'\big) \\
& + u_t\big[AX_t + B + \Sigma\Sigma'(q_t + q_t')X_t + \Sigma\Sigma' u_t' + \Sigma z_t'\big] + \frac{1}{2}\mathrm{tr}(q_t + q_t')\Sigma\Sigma'\Big\} dt \\
& - \big\{X_t'(q_t + q_t')\Sigma + u_t\Sigma + z_t\big\} dW_t^{\mathbb{P}} \\
= -\Big\{ & \big[X_t'\dot{q}_t X_t + X_t'(q_t A + A' q_t)X_t + \frac{1}{2}X_t'(q_t' + q_t)\Sigma\Sigma'(q_t' + q_t)X_t - X_t' Q X_t\big] \\
& + \big[\dot{u}_t X_t + u_t A X_t + B'(q_t' + q_t)X_t + u_t\Sigma\Sigma'(q_t' + q_t)X_t - R'X_t\big] \\
& + \big[(X_t' Q X_t + R' X_t + k) + (\frac{1}{2}X_t'(q_t' + q_t)\Sigma\Sigma'(q_t' + q_t)X_t \\
& + X_t'(q_t + q_t')\Sigma\Sigma' u_t' + X_t'(q_t + q_t')\Sigma z_t' + \frac{1}{2}u_t\Sigma\Sigma' u_t' + u_t\Sigma z_t' + \frac{1}{2}z_t z_t')\big]\Big\} dt \\
& - \big\{X_t'(q_t + q_t')\Sigma + u_t\Sigma + z_t\big\} dW_t^{\mathbb{P}}
\end{aligned} \tag{B.10}
$$

Substituting (2.4.22)-(2.4.23) and (2.4.26) into (B.10) we have

$$
dY_t = -(X_t' Q X_t + R' X_t + k - \frac{1}{2}Z_t Z_t')dt - Z_t dW_t^{\mathbb{P}}
$$

Thus (Y_t, Z_t) defined by (2.4.25)-(2.4.26) satisfies

$$
Y_t = Y_T + \int_t^T (X_s' Q X_s + R' X_s + k - \frac{1}{2}Z_s Z_s')ds + \int_t^T Z_s dW_s^{\mathbb{P}}
$$

By the boundary conditions of (2.4.24) and we have

$$
Y_t = -\ln C_T + \int_t^T (X_s' Q X_s + R' X_s + k + \frac{1}{2}Z_s Z_s')ds + \int_t^T Z_s dW_s^{\mathbb{P}}.
$$

□

C Optimal trading problem

This appendix provides technical proofs for the optimal trading project.

C.1 Explicit expression of optimal utilities

We prove several lemmas which are used for our main results Propositions 3.2.2, 3.3.2, 3.4.1 and 3.4.2.

Lemma C.1.

$$
\begin{aligned}
&\mathbb{E}[1_{\{\tau>T\}}\ln(\hat{X}_T^{(2,b)})|\mathcal{G}_0^{(2)}] = \left(\ln(X_0) + \frac{1}{2}(\mu - \frac{\mu^2}{\sigma^2})T\right) \\
&\times \left\{\mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}}\right) - \exp\left(\frac{2\mu\ln\alpha}{\sigma^2} - \ln\alpha\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}}\right)\right\} \\
&+ \int_{\frac{\ln\alpha}{\sigma}}^{0}\int_{y}^{\infty} \frac{2\mu x(x-2y)}{\sigma\sqrt{2\pi T^3}}\exp\left\{(\frac{\mu}{\sigma} - \frac{1}{2}\sigma)x - \frac{1}{2}(\frac{\mu}{\sigma} - \frac{1}{2}\sigma)^2 T - \frac{1}{2T}(2y-x)^2\right\}dxdy
\end{aligned}
$$

where

$$
\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{u^2}{2}}du
$$

is the cumulative distribution function of a standard normal random variable.

Proof. By (3.1.1) we have

$$
S_t = S_0\exp\left\{\sigma\left((\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t + W_t\right)\right\}. \tag{C.1}
$$

Define

$$
B_t = (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t + W_t
$$

and

$$
\tilde{B}_t = \inf\{B_v|0 \le v \le t\}.
$$

Recalling the definition of τ in (3.1.2) we find

$$
1_{\{\tau>T\}} = 1_{\{\tilde{B}_T > \frac{\ln\alpha}{\sigma}\}}. \tag{C.2}
$$

Let $\kappa = \frac{\mu}{\sigma} - \frac{1}{2}\sigma$. From [44] we know

$$
\begin{aligned}
&\mathbb{P}(B_T \in dx, \tilde{B}_T \in dy) \\
&= 1_{\{x>y\}}1_{\{y<0\}}\frac{2(x-2y)}{\sqrt{2\pi T^3}}\exp\{\kappa x - \frac{1}{2}\kappa^2 T - \frac{1}{2T}(2y-x)^2\}dxdy.
\end{aligned} \tag{C.3}
$$

and

$$\mathbb{P}(\tau > T)$$
$$= \mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}}\right) - \exp\left(\frac{2\mu\ln(\alpha)}{\sigma^2} - \ln(\alpha)\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma} + (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T}{\sqrt{T}}\right). \tag{C.4}$$

On the other hand, by (3.2.31) we know

$$\begin{aligned}
\hat{X}_T^{(2,b)} &= X_0 \exp\left\{\left(\hat{\pi}^{(2,b)}\mu - \frac{1}{2}(\hat{\pi}^{(2,b)}\sigma)^2\right)T + \hat{\pi}^{(2,b)}\sigma W_T\right\} \\
&= X_0 \exp\left\{\frac{\mu}{\sigma}\left((\frac{\mu}{\sigma} - \frac{1}{2}\sigma)T + W_T\right) + \frac{1}{2}\mu(1 - \frac{\mu}{\sigma^2})T\right\} \\
&= X_0 \exp\left\{\frac{\mu}{\sigma}B_T + \frac{1}{2}\mu(1 - \frac{\mu}{\sigma^2})T\right\}.
\end{aligned} \tag{C.5}$$

Using (C.2) and (C.5) we compute

$$\begin{aligned}
&\mathbb{E}\left[1_{\{\tau>T\}} \ln\left(\hat{X}_T^{(2,b)}\right) |\mathcal{G}_0^{(2)}\right] \\
=&\mathbb{E}\left[1_{\{\tau>T\}}\left\{\ln(X_0) + \frac{\mu}{\sigma}B_T + \frac{1}{2}\mu(1 - \frac{\mu}{\sigma^2})T\right\} |\mathcal{G}_0^{(2)}\right] \\
=&\mathbb{P}(\tau > T)\left\{\ln(X_0) + \frac{1}{2}(\mu - \frac{\mu^2}{\sigma^2})T\right\} + \mathbb{E}\left[1_{\{\hat{B}_T > \frac{\ln\alpha}{\sigma}\}}\frac{\mu}{\sigma}B_T\right]
\end{aligned} \tag{C.6}$$

since (Θ, K) is independent of $\mathbb{F}$ and X_0 is $\mathcal{G}_0^{(2)}$-measurable. Finally we apply (C.3) and (C.4) to (C.6) to obtain the result. □

Lemma C.2.

$$\mathbb{E}[1_{T\geq\tau}\ln(\hat{X}_T^{(2,a)})|\mathcal{G}_0^{(2)}]$$
$$= -\frac{\ln\alpha}{\sigma}\int_0^T \frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma} - (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t\right)^2\right\} h^{(2)}(t,\Theta,K)dt$$

where

$$h^{(2)}(t,\theta,k) := \ln X_0 + \frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}t - \frac{\mu^2}{2\sigma^2}t + \int_t^T \frac{\left(\mu_v^I(t,\theta,k)\right)^2}{2\sigma^2}dv. \tag{C.7}$$

Proof. Let $t = \tau$ in (C.1) we have

$$S_\tau = S_0\exp\{(\mu - \frac{1}{2}\sigma^2)\tau + \sigma W_\tau\}.$$

Using the fact $S_\tau = \alpha S_0$ we find

$$W_\tau = \frac{1}{\sigma}\left\{\ln\alpha - (\mu - \frac{1}{2}\sigma^2)\tau\right\}. \tag{C.8}$$

By (3.2.31) and (C.8) we compute

$$\begin{aligned}
\hat{X}_\tau^{(2,b)} &= X_0 \exp\left\{(\hat{\pi}^{(2,b)} - \frac{1}{2}(\hat{\pi}^{(2,b)})^2\sigma^2)\tau + \hat{\pi}^b\sigma W_\tau\right\} \\
&= X_0 \exp\left\{\frac{\mu^2}{2\sigma^2}t + \frac{\mu}{\sigma}W_\tau\right\} \\
&= X_0 \exp\left\{\frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}\tau - \frac{\mu^2}{2\sigma^2}\tau\right\}.
\end{aligned} \tag{C.9}$$

Solving (3.2.32) we obtain

$$\begin{aligned}
&\hat{X}_T^{(2,a)} \\
=&\hat{X}_\tau^{(2,b)} \exp\{\int_\tau^T (\hat{\pi}_v^{2,a}(\Theta, K)\mu_v^I(\tau,\Theta,K) - \frac{1}{2}(\hat{\pi}_v^{2,a}(\Theta,K))^2\sigma^2)dv \\
&+ \int_\tau^T \hat{\pi}_v^{2,a}(\Theta,K)\sigma dW_v\} \\
=&\hat{X}_\tau^{(2,b)} \exp\{\int_\tau^T \frac{(\mu_v^I(\tau,\Theta,K))^2}{2\sigma^2}dv + \int_\tau^T \frac{\mu_v^I(\tau,\Theta,K)}{\sigma}dW_v\}.
\end{aligned} \tag{C.10}$$

Using (C.9) and (C.10) we compute

$$\begin{aligned}
&\mathbb{E}\left[1_{T\geq\tau}\ln\left(\hat{X}_T^{(2,a)}\right)|\mathcal{G}_0^{(2)}\right] \\
=&\mathbb{E}\left[1_{T\geq\tau}\left(\ln\left(\hat{X}_\tau^{(2,b)}\right) + \int_\tau^T \frac{\left(\mu_v^I(\tau,\Theta,K)\right)^2}{2\sigma^2}dv + \int_\tau^T \frac{\mu_v^I(\tau,\Theta,K)}{\sigma}dW_v\right)|\mathcal{G}_0^{(2)}\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[1_{T\geq\tau}\left(\ln(\hat{X}_\tau^{(2,b)}) + \int_\tau^T \frac{\left(\mu_v^I(\tau,\Theta,K)\right)^2}{2\sigma^2}dv\right.\right.\right. \\
&\left.\left.\left.+\int_\tau^T \frac{\mu_v^I(\tau,\Theta,K)}{\sigma}dW_v\right)|\sigma(\tau),\mathcal{G}_0^{(2)}\right]|\mathcal{G}_0^{(2)}\right] \\
=&\mathbb{E}\left[1_{T\geq\tau}\left(\ln X_0 + \frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}\tau - \frac{\mu^2}{2\sigma^2}\tau + \int_\tau^T \frac{\left(\mu_v^I(t,\Theta,K)\right)^2}{2\sigma^2}dv\right)|\mathcal{G}_0^{(2)}\right].
\end{aligned}$$

Recall that (Θ, K) is independent to $\mathbb{F}$ and that from [44, Sect. 3.3.1] that the density of τ is

$$\mathbb{P}(\tau\in dt) = -\frac{\ln\alpha}{\sigma}\frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma} - (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t\right)^2\right\}dt. \tag{C.11}$$

Using (C.11), and the definition of the function $h^{(2)}(t,\theta,k)$ in (C.7), we obtain the result. $\square$

Lemma C.3.

$$\mathbb{E}[1_{T\geq\tau}\ln(\hat{X}_T^{1,a})] = -\frac{\ln\alpha}{\sigma}\int_0^T \frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma}-(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)t\right)^2\right\}h^{(1)}(t)dt$$

where

$$h^{(1)}(t) := \ln x_0 + \frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}t - \frac{\mu^2}{2\sigma^2}t + \int_t^T \frac{\left(\bar{\mu}_v^M\right)^2}{2\sigma^2}dv. \tag{C.12}$$

Proof. Similar to the proof of Lemma C.2, we find the terminal wealth $\hat{X}_T^{(1,a)}$ if liquidation occurs before T

$$\hat{X}_T^{(1,a)} = x_0\exp\left\{\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\frac{(\bar{\mu}_v^M)^2}{2\sigma^2}dv+\int_\tau^T\frac{\bar{\mu}_v^M}{\sigma}d\tilde{W}_v\right\}.$$

We compute

$$\begin{aligned}
&\mathbb{E}\left[1_{T\geq\tau}\ln\left(\hat{X}_T^{(1,a)}\right)\right]\\
=&\mathbb{E}\left[1_{T\geq\tau}\left\{\ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\frac{(\bar{\mu}_v^M)^2}{2\sigma^2}dv+\int_\tau^T\frac{\bar{\mu}_v^M}{\sigma}d\tilde{W}_v\right\}\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[1_{T\geq\tau}\left\{\ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\frac{(\bar{\mu}_v^M)^2}{2\sigma^2}dv+\int_\tau^T\frac{\bar{\mu}_v^M}{\sigma}d\tilde{W}_v\right\}\middle|\sigma(\tau)\right]\right]\\
=&\mathbb{E}\left[1_{T\geq\tau}\left\{\ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\frac{(\bar{\mu}_v^M)^2}{2\sigma^2}dv\right\}\right]
\end{aligned}$$

Using the density of τ given in (C.11) and the definition of the function $h^{(1)}(t)$ in (C.12), we obtain the result. $\square$

Lemma C.4.

$$\begin{aligned}
&\frac{1}{p}\mathbb{E}[1_{\{\tau>T\}}(\hat{X}_T^{(0,b)})^p] =\\
&\quad\frac{x_0^p}{p}\exp\left(\frac{p\mu^2T}{2(1-p)\sigma^2}\right)\times\left\{\mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma}+(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})T}{\sqrt{T}}\right)\right.\\
&\quad\left.-\exp\left(\frac{2\mu\ln\alpha}{(1-p)\sigma^2}-\ln\alpha\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma}+(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})T}{\sqrt{T}}\right)\right\}.
\end{aligned}$$

Proof. The proof is basically the same as that of Lemma C.1 except using the power utility function instead of log utility function. We compute

$$
\begin{aligned}
&\frac{1}{p}\mathbb{E}[1_{\{\tau>T\}}(\hat{X}_T^{(0,b)})^p] \\
=&\frac{x_0^p}{p}\exp\left(\frac{p\mu^2 T}{2(1-p)\sigma^2}\right)\times\int_{\frac{\ln\alpha}{\sigma}}^{0}\int_{y}^{\infty}\frac{2(x-2y)}{\sqrt{2\pi T^3}} \\
&\quad\times\exp\left\{(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})x-\frac{1}{2}(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})^2T-\frac{1}{2T}(2y-x)^2\right\}dxdy.
\end{aligned}
\tag{C.13}
$$

Define

$$
C_t=(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})t+W_t
$$

and

$$
\tilde{C}_t=\inf\{C_s|0\le s\le t\}.
$$

Note that the integral in (C.13) is equal to $\mathbb{P}(\tilde{C}_T>\frac{\ln\alpha}{\sigma})$. By [44, Sect. 3.2.2] we know

$$
\begin{aligned}
&\mathbb{P}(\tilde{C}_T>\frac{\ln\alpha}{\sigma}) \\
=&\mathcal{N}\left(\frac{-\frac{\ln\alpha}{\sigma}+(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})T}{\sqrt{T}}\right)-\exp\left(\frac{2\mu\ln\alpha}{(1-p)\sigma^2}-\ln\alpha\right)\mathcal{N}\left(\frac{\frac{\ln\alpha}{\sigma}+(\frac{\mu}{(1-p)\sigma}-\frac{\sigma}{2})T}{\sqrt{T}}\right).
\end{aligned}
\tag{C.14}
$$

Substituting (C.14) into (C.13) we obtain the result. □

Lemma C.5.

$$
\begin{aligned}
&\frac{1}{p}\mathbb{E}[1_{T\ge\tau}(\hat{X}_T^{(0,a)})^p]=-\frac{\ln\alpha}{\sigma} \\
&\times\int_0^1\int_0^\infty\int_0^T\frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma}-(\frac{\mu}{\sigma}-\frac{1}{2}\sigma)t\right)^2\right\}l^{(0)}(t,\theta,k)\varphi(\theta,k)dtd\theta dk
\end{aligned}
$$

where

$$
\begin{aligned}
&l^{(0)}(t,\theta,k)= \\
&\quad\frac{x_0^p}{p}\exp\left\{\frac{\mu\ln\alpha}{(1-p)\sigma^2}+\frac{1}{2}\frac{\mu}{(1-p)}t-\frac{1}{2}\frac{\mu^2}{(1-p)^2\sigma^2}t+\int_t^T\left(\frac{p\mu\mu_v^I(t,\theta,k)}{(1-p)\sigma^2}\right)dv\right\}.
\end{aligned}
\tag{C.15}
$$

Proof. Similar to the proof of Lemma C.2 we find the wealth value at liquidation time τ as follows

$$\hat{X}_{\tau}^{(0,b)} = x_0 \exp\left\{\frac{\mu \ln \alpha}{(1-p)\sigma^2} + \frac{1}{2}\frac{\mu}{(1-p)}\tau - \frac{1}{2}\frac{\mu^2}{(1-p)^2\sigma^2}\tau\right\}.$$

Then the terminal wealth $\hat{X}_T^{(0,a)}$ if liquidation occurs before T is

$$\hat{X}_T^{(0,a)} = \hat{X}_{\tau}^{(0,b)} \exp\left\{\int_{\tau}^{T}\left(\frac{\mu\mu_v^I(\tau,\Theta,K)}{(1-p)\sigma^2} - \frac{\mu^2}{2(1-p)^2\sigma^2}\right)dv + \int_{\tau}^{T}\frac{\mu}{(1-p)\sigma}dW_v\right\}.$$

We compute

$$\begin{aligned}
&\mathbb{E}[1_{T\geq\tau}U(\hat{X}_T^{(0,a)})] \\
=&\frac{1}{p}\mathbb{E}\left[1_{\{T\geq\tau\}}(\hat{X}_{\tau}^{b})^p \exp\left\{\int_{\tau}^{T}\left(\frac{p\mu\mu_v^I(\tau,\Theta,K)}{(1-p)\sigma^2} - \frac{p\mu^2}{2(1-p)^2\sigma^2}\right)dv\right.\right. \\
&\quad\left.\left.+\int_{\tau}^{T}\frac{p\mu}{(1-p)\sigma}dW_v\right\}\right] \\
=&\frac{1}{p}\mathbb{E}\left[\mathbb{E}\left[1_{\{T\geq\tau\}}(\hat{X}_{\tau}^{b})^p \exp\left\{\int_{\tau}^{T}\left(\frac{p\mu\mu_v^I(\tau,\Theta,K)}{(1-p)\sigma^2} - \frac{p\mu^2}{2(1-p)^2\sigma^2}\right)dv\right.\right.\right. \\
&\quad\left.\left.\left.+\int_{\tau}^{T}\frac{p\mu}{(1-p)\sigma}dW_v\right\}|\sigma(\tau)\right]\right] \\
=&\frac{1}{p}\mathbb{E}\left[1_{\{T\geq\tau\}}(\hat{X}_{\tau}^{b})^p \exp\left\{\int_{\tau}^{T}\left(\frac{p\mu\mu_v^I(\tau,\Theta,K)}{(1-p)\sigma^2} - \frac{p\mu^2}{2(1-p)\sigma^2}\right)dv\right\}\right]
\end{aligned}$$

Using the density of τ given in (C.11) and the definition of the function $l^{(0)}(t,\theta,k)$ in (C.15), we obtain the result. □

Lemma C.6.

$$\begin{aligned}
&\mathbb{E}[1_{T\geq\tau}\ln(\hat{X}_T^{(0,a)})] = -\frac{\ln\alpha}{\sigma} \\
&\int_0^{\infty}\int_0^{\infty}\int_0^{T}\frac{1}{\sqrt{2\pi t^3}}\exp\left\{-\frac{1}{2t}\left(\frac{\ln\alpha}{\sigma} - (\frac{\mu}{\sigma} - \frac{1}{2}\sigma)t\right)^2\right\}h^{(0)}(t,\theta,k)\varphi(\theta,k)dtd\theta dk
\end{aligned}$$

where

$$h^{(0)}(t,\theta,k) := \ln x_0 + \frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}t - \frac{\mu^2}{2\sigma^2}t + \int_t^T\left(\frac{2\mu\mu_v^I(t,\theta,k) - \mu^2}{2\sigma^2}\right)dv. \quad \text{(C.16)}$$

Proof. Similar to the proof of Lemma C.2, we find the terminal wealth $\hat{X}_T^{(0,a)}$ if liquidation occurs before T

$$\begin{aligned}
&\hat{X}_T^{(0,a)} = \\
&\quad x_0\exp\left\{\frac{\mu\ln\alpha}{\sigma^2} + \frac{\mu}{2}\tau - \frac{\mu^2}{2\sigma^2}\tau + \int_{\tau}^{T}\left(\frac{\mu\mu_v^I(\tau,\Theta,K)}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right)dt + \int_{\tau}^{T}\frac{\mu}{\sigma}dW_t\right\}.
\end{aligned}$$

We compute

$$
\begin{aligned}
&\mathbb{E}\left[1_{T\geq\tau}\ln(\hat{X}_T^{(0,a)})\right]\\
=&\mathbb{E}\left[1_{T\geq\tau}\left\{\ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\left(\frac{\mu\mu_v^I(\tau,\Theta,K)}{\sigma^2}-\frac{\mu^2}{2\sigma^2}\right)dt\right.\right.\\
&\left.\left.+\int_\tau^T\frac{\mu}{\sigma}dW_t\right\}\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[1_{T\geq\tau}\left\{\ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\left(\frac{\mu\mu_v^I(\tau,\Theta,K)}{\sigma^2}-\frac{\mu^2}{2\sigma^2}\right)dt\right.\right.\right.\\
&\left.\left.\left.+\int_\tau^T\frac{\mu}{\sigma}dW_t\right\}|\sigma(\tau)\right]\right]\\
=&\mathbb{E}\left[1_{T\geq\tau}\left\{\ln x_0+\frac{\mu\ln\alpha}{\sigma^2}+\frac{\mu}{2}\tau-\frac{\mu^2}{2\sigma^2}\tau+\int_\tau^T\left(\frac{\mu\mu_v^I(\tau,\Theta,K)}{\sigma^2}-\frac{\mu^2}{2\sigma^2}\right)dt\right\}\right]
\end{aligned}
$$

Using the density of τ given in (C.11) and the definition of the function $h^{(0)}(t,\theta)$ in (C.16), we obtain the result. □