Mean-variance portfolio selection under partial information

Jie Xiong [a,b,*] and Xun Yu Zhou [c,†]

[a] Department of Mathematics, University of Tennessee,
Knoxville, TN 37996-1300, USA

[b] Department of Mathematics, Hebei Normal University,
Shijiazhuang 050016, PRC

[c] Department of Systems Engineering and Engineering Management,
The Chinese University of Hong Kong, Shatin, Hong Kong

October 4, 2006

Abstract

This paper is concerned with a continuous-time mean–variance portfolio selection problem in a (possibly incomplete) market with multiple stocks and a bond. Only the past price movements of the stocks and the bond are the information available to the investors. A separation principle is shown to hold in this setting. Efficient strategies based on the aforementioned partial information are derived, which involve the optimal filter of the stock appreciation rate processes. The main methodological contribution of the paper is to employ the particle system representation to develop analytical and numerical approaches in obtaining the filter as well as solving the related backward stochastic differential equation.

KEY WORDS: Mean–variance portfolio selection, continuous time, partial information, nonlinear filtering, backward stochastic differential equation, particle system representation.

MSC 2000 SUBJECT CLASSIFICATIONS: 91B28, 93C41, 93E11

1 Introduction

In the Nobel-Prize-winning work [19], Markowitz proposed the mean–variance portfolio selection model for a single investment period, where an agent seeks to minimize the

*Research supported partially by NSA Grant H98230-05-1-0043.

†Research supported by RGC Grant CUHK4175/03E and the Croucher Senior Research Fellowship.

risk of his investment, measured by the variance of his return, subject to a given mean return. The dynamic extension of the Markowitz model, especially in continuous time, has been studied extensively in recent years; see, e.g., Li and Ng [15], Zhao and Ziemba [30], Zhou and Li [31], Lim [16], Bielecki *et al* [2], and Xia [25] (In particular, refer to Steinbach [23] and Bielecki *et al* [2] for elaborative discussions on the history of the mean–variance model.) In many of these works, explicit, analytic forms of efficient portfolios have been obtained. However, in all these works it is assumed that the driving Brownian motions are completely observable by an investor, which in reality is more an exception than a rule. Practically, the investor can only observe the stock prices (including past and present) based on which he will have to make his investment decisions. This leads to the so-called partially observed portfolio selection problem, and this paper aims to solve the problem in the realm of mean–variance. An important finding in this paper is that the separation principle (for separating filtering and optimization) turns out to hold in the mean–variance setting, which in turn greatly simplify the problem. Another main contribution of the paper is to employ the particle system representation, which has been developed quite recently for solving stochastic PDEs, to develop analytical and numerical approaches in obtaining the filter as well as solving the related backward stochastic differential equation.

Asset allocation and asset pricing based on partial information under various setups have been studied extensively in the financial economics literature; see, to name a few, Lakner [14], Brennan and Xia [3], Xia [26], Rogers [22], Nagai and Peng [20], Yang and Xiong [27]. Detemple [4], Dothan and Feldman [5] and Gennotte [7] established a separation principle. However, all these works are predominantly done within the expected utility framework. (Refer to [2, 23, 30] for discussions on crucial differences between the utility and mean–variance models.) Pham [21] considered a mean–variance hedging problem for a general semimartingale model, and proved a separation principle for a diffusion model (though it is not a multi-dimensional geometric Brownian motion as in the present paper). Although in theory a mean–variance problem *a là* Markowitz can be formulated as a mean–variance hedging problem, there are subtleties such as the feasibility and the determination of the Lagrange multiplier (see [2, 16]). Moreover, the analysis in [21] is rather involved due to the martingale method employed, whereas here we will give a very direct, clean and short proof for a separation principle.

The rest of this article is organized as follows: In Section 2, we formulate the mean–variance portfolio selection model under partial information. In Section 3, we derive the innovation process associated with the filtering problem, which leads to the separation principle. Section 4 studies the optimal filter in details for two cases. Section 5 is devoted to the optimization part as well as the final solution to the partially observed mean–variance problem. A numerical solution to a related backward SDE is presented, which is of independent interest.

2 The Model

We consider a market consisting of d stocks and a bond whose prices are stochastic processes $S_i(t)$, $i = 0, 1, \cdots, d$, governed by the following stochastic differential equations (SDEs):

$$\begin{cases} dS_i(t) = S_i(t)\left(\mu_i(t)dt + \sum_{j=1}^m \tilde{\sigma}_{ij}(t)d\tilde{W}_j(t)\right), & i = 1, 2, \cdots, d, \\ dS_0(t) = S_0(t)\mu_0(t)dt, \quad t \geq 0, \end{cases}$$

where $\tilde{W} := (\tilde{W}_1, \cdots, \tilde{W}_m)^*$ is a standard Brownian motion defined on a filtered complete probability space $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{t\geq 0})$, $\mu_i(t)$, $i = 1, 2, \cdots, d$, are the appreciation rate processes of the stocks, $\mu_0(t)$ is the interest rate process, and the $d \times m$ matrix valued process $\tilde{\Sigma}(t) := (\tilde{\sigma}_{ij}(t))$ is the volatility process. Here and throughout the paper A^* denotes the transpose of a matrix A.

Let

$$\mathcal{G}_t := \sigma(S_i(s) : \ s \leq t, \ i = 0, 1, 2, \cdots, d), \quad t \geq 0.$$

In our model $\mathcal{G}_t$, rather than $\mathcal{F}_t^{\tilde{W}}$ (the filtration generated by $\tilde{W}$), is the only information available to the investors at time t.

By Itô's formula, we have

$$d \log S_i(t) = \left(\mu_i(t) - \frac{1}{2}a_{ii}(t)\right) dt + \sum_{j=1}^m \tilde{\sigma}_{ij}(t)d\tilde{W}_j(t), \quad i = 1, 2, \cdots, d, \tag{2.1}$$

where

$$a_{ij}(t) := \sum_{k=1}^m \tilde{\sigma}_{ik}(t)\tilde{\sigma}_{jk}(t), \quad i, j = 1, 2, \cdots, d.$$

The following assumptions on the market coefficients will be in force throughout this paper.

Assumption (ND): For any $t \geq 0$, the $d \times d$ matrix $A(t) := (a_{ij}(t))$ is of full rank almost surely.

Assumption (BC): There exists a finite constant C such that $\forall\, t \geq 0,\ \ \forall i, j,$

$$|\tilde{\sigma}_{ij}(t)| \leq C, \qquad \text{a.s.}$$

Assumption (IC): $\mathbb{E}\int_0^T (\mu_0(t)^2 + |\mu(t)|^2)dt < \infty$.

Remark 2.1. In this article, we allow $d < m$ as long as the condition (ND) is satisfied; in other words, the market itself is allowed to be incomplete. It is interesting to note that, unlike the full information case, the incompleteness of the market does not impose essential difficulty in the partial information case. This can be explained as follows. In the classical model of incomplete market (with full information), there are vast amount of information available. Namely, one has to seek optimal portfolios in the class of all $\mathcal{F}_t^{\tilde{W}}$-adapted portfolios. When $m > d$, the number of available stocks is less than that of the (independent) random factors and hence, some of the market risks can not be completely eliminated by composing an appropriate stock portfolio. As a result, portfolio selection problems become harder than the case with a complete market because some contingent claims cannot be replicated. In our current setup, the available information comes *only* from the stocks themselves, and any other information is not observable anyway. Therefore, the model is essentially "complete", as also will be evident in what follows, although the market is indeed incomplete in the conventional sense.

It is easy to show that the quadratic covariation process between $\log S_i(t)$ and $\log S_j(t)$ is given by $\int_0^t a_{ij}(s)ds$. Therefore, the matrix valued process $(a_{ij}(t))$ is $\mathcal{G}_t$-adapted. Let $\Sigma(t) \equiv (\sigma_{ij}(t))$ be the square root of $A(t)$. Then, $\sigma_{ij}(t)$ is $\mathcal{G}_t$-adapted, i.e., it is completely observable. As we shall see in (3.5) below the stock price $S_i(t)$ satisfies an equivalent SDE which depends on $\sigma_{ij}(t)$ instead of $\tilde{\sigma}_{ij}(t)$. Moreover, $\mu_0(t) = \frac{d}{dt}\log S_0(t)$ is also $\mathcal{G}_t$-adapted. Therefore, we do not need to consider the filtering problem for the stochastic interest rate and volatility processes.

However, the stochastic process $\mu(t) := (\mu_1(t), \cdots, \mu_d(t))^*$ is not necessarily $\mathcal{G}_t$-adapted and hence, its value is not available to the investors. Note that $\mu(t)$, being a very general process, does not need to be even $\mathcal{F}_t^{\tilde{W}}$-adapted.

Denote by $L^2_{\mathcal{G}}(0, T; \mathbb{R}^n)$ the set of $\mathbb{R}^n$-valued, $\mathcal{G}_t$-adapted processes $f(t)$ with $\mathbb{E}\int_0^T |f(t)|^2 dt < \infty$. (Similar notation $L^2_{\mathcal{H}}(0, T; \mathbb{R}^n)$ can be defined for any filtration

$\mathcal{H}_t$.) $L^2_{\mathcal{G}}(0,T;\mathbb{R}^n)$ becomes a Hilbert space endowed with the norm $\| f \|_{L^2_{\mathcal{G}}(0,T;\mathbb{R}^n)} := \left(\mathbb{E}\int_0^T |f(t)|^2 dt\right)^{\frac{1}{2}}$.

We now define the class of admissible portfolios (investment strategies).

Definition 2.2. A d-dimensional process $u(t) \equiv (u_1(t), \cdots, u_d(t))^*$ is an *admissible portfolio* if $u(t) \in L^2_{\mathcal{G}}(0,T;\mathbb{R}^d)$.

In the preceding definition, $u_i(t)$ represents the worth of an agent's wealth (dollar amount) in the ith stock, $i = 1, 2, \cdots, d$. It is well known that under a so-called self-financed portfolio, the wealth process of an agent, starting with an initial wealth x_0, satisfies the following *wealth equation* (see, e.g., [11]):

$$\begin{cases} dx(t) = \left(\mu_0(t)x(t) + \sum_{i=1}^d (\mu_i(t) - \mu_0(t))u_i(t)\right) dt + \sum_{i=1}^d \sum_{j=1}^m \tilde{\sigma}_{ij}(t)u_i(t)d\tilde{W}_j(t), \quad t \geq 0, \\ x(0) = x_0. \end{cases} \tag{2.2}$$

The partially observed mean–variance portfolio selection model is formulated as the following optimization model:

$$\begin{array}{ll} \text{Minimize} & \mathrm{Var}(x(T)) = \mathbb{E}(x(T) - \mathbb{E}x(T))^2, \\ \text{subject to} & \begin{cases} u(t) \text{ is self-financed and admissible,} \\ (x(t), u(t)) \text{ satisfies equation (2.2) with initial wealth } x_0, \\ \mathbb{E}x(T) = z, \end{cases} \end{array} \tag{2.3}$$

where $x_0, z \in \mathbb{R}$ are given constants.

3 Separation Principle

In this section, we consider the filtering problem associated with our model (2.3) and establish a separation principle. Specifically, we define the innovation process for the filtering problem. Based on this process, we will derive a $\mathcal{G}_t$-adapted representation for the wealth process corresponding to any self-financed admissible portfolio.

Theorem 3.1. *Under any self-financed admissible portfolio $u(t)$, the corresponding wealth process $x(t)$ satisfies the following SDE:*

$$\begin{cases} dx(t) = \left(x(t)\mu_0(t) + \sum_{i=1}^d (\bar{\mu}_i(t) - \mu_0(t))u_i(t)\right) dt + \sum_{i,j=1}^d \sigma_{ij}(t)u_i(t)d\nu_j(t), \ t \geq 0, \\ x(0) = x_0, \end{cases} \tag{3.1}$$

where $\bar{\mu}_i(t) := \mathbb{E}(\mu_i(t)|\mathcal{G}_t)$ *is the optimal filter of* $\mu_i(t)$*, the* innovation process $\nu(t) \equiv (\nu_1(t), \cdots, \nu_d(t))^*$ *is a d-dimensional Brownian motion given by*

$$d\nu(t) := \Sigma(t)^{-1} d \log S(t) - \Sigma(t)^{-1} \left(\bar{\mu}(t) - \frac{1}{2}\tilde{A}(t) \right) dt, \qquad (3.2)$$

$S(t) := (S_1(t), \cdots, S_d(t))^*$, $\log S(t) := (\log S_1(t), \cdots, \log S_d(t))^*$, $\bar{\mu}(t) := (\bar{\mu}_1(t), \cdots, \bar{\mu}_d(t))^*$, *and* $\tilde{A}(t) := (a_{11}(t), \cdots, a_{dd}(t))^*$.

Proof: From (2.1), we see that

$$\log S_i(t) - \log S_i(0) - \int_0^t \left(\mu_i(s) - \frac{1}{2} a_{ii}(s) \right) ds = \sum_{j=1}^m \int_0^t \tilde{\sigma}_{ij}(s) d\tilde{W}_j(s), \qquad i = 1, 2, \cdots, d$$

are martingales with quadratic covariation process $\int_0^t A(s)ds = \int_0^t \Sigma(s)^2 ds$. By the martingale representation theorem, there exists a standard Brownian motion $W \equiv (W_1, \cdots, W_d)$ on $(\Omega, \mathcal{F}, P)$ such that

$$\sum_{j=1}^m \tilde{\sigma}_{ij}(t) d\tilde{W}_j(t) = \sum_{j=1}^d \sigma_{ij}(t) dW_j(t), \qquad i = 1, \cdots, d. \qquad (3.3)$$

Thus,

$$d \log S_i(t) = \left(\mu_i(t) - \frac{1}{2} a_{ii}(t) \right) dt + \sum_{j=1}^d \sigma_{ij}(t) dW_j(t), \qquad i = 1, \cdots, d. \qquad (3.4)$$

Equivalently, the stock prices satisfy the following modified SDE:

$$dS_i(t) = S_i(t) \left(\mu_i(t) dt + \sum_{j=1}^d \sigma_{ij}(t) dW_j(t) \right), \qquad i = 1, \cdots, d. \qquad (3.5)$$

Note that $\Sigma(t)$ is invertible. Let $\tilde{S}(t)$ be defined by

$$d\tilde{S}(t) := \Sigma(t)^{-1} d \log S(t).$$

We can write the observation equation (3.4) into the classical form (cf. (8.1.1) in [10]):

$$\tilde{S}(t) = \tilde{S}(0) + \int_0^t \Sigma(s)^{-1} \left(\mu(s) - \frac{1}{2}\tilde{A}(s) \right) ds + W(t) \qquad (3.6)$$

with the observation σ-field $\mathcal{G}_t$. By Theorem 8.1.3 and Remark 8.1.1 in Kallianpur [10], $(\nu(t), \mathcal{G}_t)$ is a d-dimensional Brownian motion such that $\sigma(\nu(u) - \nu(s) : u \geq s \geq t)$ is independent of $\mathcal{G}_t$.

By (3.6) and (3.2), we get

$$\Sigma(t)dW(t) = \Sigma(t)d\nu(t) + (\bar{\mu}(t) - \mu(t))dt. \tag{3.7}$$

The desired form of wealth equation (3.1) then follows from (2.2), (3.3) and (3.7). □

Remark 3.2. A notorious difficulty in tackling general stochastic optimization problems with partial information is that one usually cannot separate the filtering and optimization, except for some very rare situations. The significance of Theorem 3.1 is that for the specific mean–variance portfolio selection problem, the separation principle happens to hold: one can simply replace the appreciation rate by its filter in the wealth equation, and then solve the resulting optimization problem as in the complete information case.

4 Filtering

In this section, we study the filtering problem (for the appreciation rate process) by considering two cases associated with the volatility processes. The aim is to study the optimal filter $U(t)$ given by

$$\langle U(t), f\rangle := \mathbb{E}(f(\mu(t))|\mathcal{G}_t), \qquad \forall\, f \in C_b(\mathbb{R}^d).$$

4.1 Case 1: Non-random $\tilde{\Sigma}$

In this subsection we consider the case when the original volatility process $\tilde{\Sigma}(t)$ is a deterministic matrix-valued function of t and $\mu(t)$ a d-dimensional Markov process (with a generator L) independent of $\tilde{W}$. By the definitions of $\Sigma(t)$ and W, where W is defined via (3.3), it is clear that $\Sigma(t)$ is also non-random and $\mu(t)$ is independent of W. Then the filtering problem becomes a classical one with the signal $\mu(t)$ and observation $\tilde{S}(t)$ given by (3.6). In this case, $\mathcal{G}_t = \mathcal{F}_t^{\tilde{S}}$.

Remark 4.1. By Theorem 8.3.1 in [10], every square integrable $\mathcal{F}_t^{\tilde{S}}$-martingale (and hence, every $\mathcal{G}_t$-martingale in the present case) Y_t can be represented as

$$Y_t = \mathbb{E}(Y_0) + \int_0^t \Phi(s)^* d\nu(s) \tag{4.1}$$

where $\Phi(s) \in L^2_{\mathcal{F}^{\tilde{S}}}(0, T; \mathbb{R}^d)$. This fact will be useful in Proposition 5.5 below in establishing an optimal portfolio since $\mathbb{H} = AC(\mathcal{G})$ in this case (cf. Definition 5.1). Namely, the market is complete in the sense of Definition 5.2.

In Kurtz and Xiong ([12], [13]), a large class of stochastic partial differential equations, with the filtering equations as a special case, and their numerical solutions are studied based on a technique called the particle system representation. In this subsection, we demonstrate that the particle system representation itself can be used to derive the filtering equations directly. Note that the generator L of the appreciation rate process $\mu(t)$ is not required to be a second-order differential operator. In fact, even the continuity of $\mu(t)$ in t needs not to be assumed. It is also worth mentioning that the results in this and the next subsections are not covered by those of [12], [13].

Introduce the following integrability condition:

$$\int_0^T \left| \left(\mu(s)^* - \frac{1}{2}\tilde{A}(s)^* \right) \Sigma(s)^{-1} \right|^2 ds < \infty, \qquad \text{a.s.} \tag{4.2}$$

Applying Girsanov's formula to (3.6) and noting that μ and W are independent, under the probability $\tilde{P}$ defined below, we get that $\tilde{S}(t)$ is a Brownian motion independent of $\mu(t)$, where $dP = M(T)d\tilde{P}$ with

$$\begin{aligned} M(t) \quad := \quad \exp \Bigg(& \int_0^t \left(\mu(s)^* - \frac{1}{2}\tilde{A}(s)^* \right) \Sigma(s)^{-1} d\tilde{S}(s) \\ & -\frac{1}{2} \int_0^t \left| \left(\mu(s)^* - \frac{1}{2}\tilde{A}(s)^* \right) \Sigma(s)^{-1} \right|^2 ds \Bigg). \end{aligned}$$

By Kallianpur-Striebel's formula, we can represent the optimal filter $U(t)$ as

$$\langle U(t), f \rangle = \frac{\langle V(t), f \rangle}{\langle V(t), 1 \rangle} \tag{4.3}$$

where for $f \in C_b(\mathbb{R}^d)$,

$$\langle V(t), f \rangle := \tilde{\mathbb{E}} \left(M(t) f(\mu(t)) | \mathcal{G}_t \right)$$

is the unnormalized filter. Here $\tilde{\mathbb{E}}$ refers to the expectation under the new probability measure $\tilde{P}$.

Lemma 4.2. *Suppose that $\mu(t)$ is a d-dimensional Markov process (with generator L) independent of W satisfying (4.2). Then, $V(t)$ is represented as*

$$\langle V(t), f \rangle = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n M^i(t) f(\mu^i(t)) \tag{4.4}$$

where $\mu^1(t),\ \mu^2(t),\ \cdots$ *are independent copies of* $\mu(t)$ *and*

$$\begin{aligned} M^i(t) \ &:= \ \exp\left(\int_0^t \left(\mu^i(s)^* - \frac{1}{2}\tilde{A}(s)^*\right)\Sigma(s)^{-1}d\tilde{S}(s) \right.\\ &\left. -\frac{1}{2}\int_0^t \left|\left(\mu^i(s)^* - \frac{1}{2}\tilde{A}(s)^*\right)\Sigma(s)^{-1}\right|^2 ds\right). \end{aligned} \tag{4.5}$$

Proof: Note that

$$dM^i(t) = M^i(t)\left(\mu^i(t)^* - \frac{1}{2}\tilde{A}(t)^*\right)\Sigma(t)^{-1}d\tilde{S}(t). \tag{4.6}$$

Denote by $\mathcal{X}$ the collection of all those processes $\mu(t)$ satisfying (4.2) and by $\mathcal{Y}$ the collection of all measurable $\mathbb{R}^{2d}$-valued random vectors. Then $\mathcal{X}$ and $\mathcal{Y}$ are measurable spaces. For $\mu^i \in \mathcal{X}$, the SDE (4.6) has a unique strong solution. Therefore, for each fixed $t \in [0,T]$ there is a measurable functional $F_t :\ C([0,T],\mathbb{R}^d) \times \mathcal{X} \to \mathcal{Y}$ such that $(\mu^i(t), M^i(t)) = F_t(\tilde{S}, \mu^i)$. As a consequence, under the conditional probability $\tilde{P}(\cdot|\mathcal{F}_t^{\tilde{S}})$ $(\mu^i(t), M^i(t))$ is completely determined by μ^i, $i = 1, 2, \cdots$. Since $\tilde{S},\ \mu^1,\ \mu^2,\ \cdots$ are independent, the strong law of large numbers yields

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n M^i(t)f(\mu^i(t)) = \tilde{\mathbb{E}}\left(M(t)f(\mu(t))|\mathcal{F}_t^{\tilde{S}}\right). \tag{4.7}$$

Since $\mathcal{F}_t^{\tilde{S}} = \mathcal{G}_t$, we obtain (4.4). □

Now we derive via Itô's formula and (4.4)-(4.5) a stochastic partial differential equation (SPDE) for the unnormalized filter $V(t)$. Let $\mu^i(t)$, $i = 1, 2, \cdots$, be as in Lemma 4.2, which are independent Markov processes with generator L. It is well-known (see the standard textbooks Ethier and Kurtz [6] or Stroock and Varadhan [24]) that there are independent martingales $N_f^i(t)$ such that

$$df(\mu^i(t)) = Lf(\mu^i(t))dt + dN_f^i(t), \quad f \in D(L), \tag{4.8}$$

where $D(L)$ is the domain of L.

Applying Itô's formula to (4.6) and (4.8), we get

$$\begin{aligned} d(M^i(t)f(\mu^i(t))) \ &= \ M^i(t)\left(Lf(\mu^i(t))dt + dN_f^i(t)\right) \\ &+M^i(t)f(\mu^i(t))\left(\left(\mu^i(t)^* - \frac{1}{2}\tilde{A}(t)^*\right)\Sigma(t)^{-1}d\tilde{S}(t)\right). \end{aligned}$$

Taking average for $i = 1, 2, \cdots, k$, letting $k \to \infty$, and applying Lemma 4.2, we see that $V(t)$ satisfies the following Zakai equation:

$$d\langle V(t), f\rangle = \langle V(t), Lf\rangle\, dt + \langle V(t), G_t f\rangle\, d\tilde{S}(t) \tag{4.9}$$

where

$$G_t f(\mu) := f(\mu)\left(\mu^* - \frac{1}{2}\tilde{A}(t)^*\right)\Sigma(t)^{-1}.$$

Making use of (4.3), by Itô's formula, we have

$$\begin{aligned} d\langle U(t), f\rangle &= \langle U(t), Lf\rangle\, dt \\ &\quad + \left(\langle U(t), G_t f\rangle - \langle U(t), G_t 1\rangle \langle U(t), f\rangle\right)\left(d\tilde{S}(t) - \langle U(t), G_t 1\rangle\, dt\right). \end{aligned}$$

Note that

$$\begin{aligned} \langle U(t), G_t 1\rangle &= \mathbb{E}\left(\left(\mu(t)^* - \frac{1}{2}\tilde{A}(t)^*\right)\Sigma(t)^{-1}\middle|\mathcal{G}_t\right) \\ &= \left(\bar{\mu}(t)^* - \frac{1}{2}\tilde{A}(t)^*\right)\Sigma(t)^{-1}. \end{aligned}$$

Hence (3.2) can be rewritten as

$$d\tilde{S}(t) - \langle U(t), G_t 1\rangle\, dt = d\nu(t).$$

Therefore, $U(t)$ satisfies the following Fujisaki-Kallianpur-Kunita (FKK) equation:

$$d\langle U(t), f\rangle = \langle U(t), Lf\rangle\, dt + \left(\langle U_t, G_t f\rangle - \langle U(t), G_t 1\rangle \langle U(t), f\rangle\right) d\nu(t). \tag{4.10}$$

Remark 4.3. Since L is not necessarily a second-order differential operator which was used heavily in [12] in proving that SPDE of the form (4.10) has a unique strong solution, we cannot establish such a property for a general L by the same argument in [12]. However, by Theorem 9.1 in Bhatt *et al* [1], under suitable conditions, the solution to (4.10) is indeed unique.

Next, let us discuss the numerical implementation of the preceding filter. Recall that $\mu^1(t)$, $\mu^2(t)$, $\cdots$ are independent copies of $\mu(t)$. For $\delta > 0$, let $\eta_\delta(t) := j\delta$ for $j\delta \le t < (j+1)\delta$, $j = 0, 1, \cdots$. We approximate $M^i(t)$ by the Euler scheme:

$$\begin{aligned} M^{\delta,i}(t) &:= \exp\left(\int_0^t \left(\mu^i(\eta_\delta(s))^* - \frac{1}{2}\tilde{A}(\eta_\delta(s))^*\right)\Sigma(s)^{-1} d\tilde{S}(s)\right. \\ &\quad \left. -\frac{1}{2}\int_0^t \left|\left(\mu^i(\eta_\delta(s))^* - \frac{1}{2}\tilde{A}(\eta_\delta(s))^*\right)\Sigma(s)^{-1}\right|^2 ds\right). \end{aligned}$$

Define

$$\left\langle V^{n,\delta}(t), f\right\rangle := \frac{1}{n}\sum_{i=1}^n M^{\delta,i}(t) f(\mu^i(t)) \text{ and } \langle V^n(t), f\rangle := \frac{1}{n}\sum_{i=1}^n M^i(t) f(\mu^i(t)).$$

We then combine both approximations by using $\bar{V}^n := V^{n,1/n^{1/2\alpha}}$ to approximate the unnormalized filter, where $\alpha > 0$ is given in (4.11) below.

Although we did not assume the sample path continuity of $\mu(t)$, we need some kind of continuity in the sense of moments. We assume that $|\Sigma(t)^{-1}| \le K$ and with some $\alpha > 0$,

$$\mathbb{E}\left(|\mu(t)-\mu(s)|^2\right) \le |t-s|^\alpha. \tag{4.11}$$

For example, if $\mu(t)$ is a compound Poisson process, then (4.11) holds with $\alpha = 1$.

Let $\bar{\mathbb{R}}^d$ be the one-point compactification of $\mathbb{R}^d$. Then $C_b(\bar{\mathbb{R}}^d)$ is a separable Banach space. Let $\mathcal{M}_F(\bar{\mathbb{R}}^d)$ be the space of finite Borel measures on $\bar{\mathbb{R}}^d$, and d be a distance defined on $\mathcal{M}_F(\bar{\mathbb{R}}^d)$ whose topology coincides with the weak convergence topology. More precisely, let $\{f_k\} \subset C_b^1(\bar{\mathbb{R}}^d)$ be a dense subset of $C_b(\bar{\mathbb{R}}^d)$. We define

$$d(\nu_1,\nu_2) := \sum_{k=1}^\infty \frac{|\langle \nu_1-\nu_2, f_k\rangle|}{2^k \|f_k\|_L}, \quad \nu_1,\nu_2 \in \mathcal{M}_F(\bar{\mathbb{R}}^d),$$

where

$$\|f\|_L := \sup_{x\in\bar{\mathbb{R}}^d} |f(x)| + \sup_{x,y\in\bar{\mathbb{R}}^d} \frac{|f(x)-f(y)|}{|x-y|}, \quad f \in C_b^1(\bar{\mathbb{R}}^d).$$

Theorem 4.4. *Suppose that $|\Sigma(t)^{-1}| \le K$ and (4.11) holds. Then, for each fixed t, there exists $M > 0$, such that for all n,*

$$\tilde{\mathbb{E}}\left(d(\bar{V}^n(t), V(t))\right) \le \frac{M}{\sqrt{n}}. \tag{4.12}$$

Proof: By the conditional independence, it is easy to show that $\forall\ f \in C_b(\bar{\mathbb{R}}^d)$,

$$\tilde{\mathbb{E}}\left(\langle V^n(t)-V(t), f\rangle^2\right) \le \frac{c_1(T)^2\|f\|_\infty^2}{n}.$$

Next we note that

$$\begin{aligned}
&\tilde{\mathbb{E}}\left\{\left|\log M^{\delta,i}(t) - \log M^i(t)\right|^2\right\} \\
\le\ & 2\int_0^T \mathbb{E}\left|\left(\mu^i(\eta_\delta(s))^* - \frac{1}{2}\tilde{A}(\eta_\delta(s))^*\right)\Sigma(s)^{-1} - \left(\mu^i(s)^* - \frac{1}{2}\tilde{A}(s)^*\right)\Sigma(s)^{-1}\right|^2 ds \\
& + T\int_0^T \mathbb{E}\left|\left|\left(\mu^i(\eta_\delta(s))^* - \frac{1}{2}\tilde{A}(\eta_\delta(s))^*\right)\Sigma(s)^{-1}\right|^2 - \left|\left(\mu^i(s)^* - \frac{1}{2}\tilde{A}(s)^*\right)\Sigma(s)^{-1}\right|^2\right|^2 ds \\
\le\ & c_2(T)\delta^\alpha.
\end{aligned}$$

Let $\tilde{d}$ be the Waserstein metric on $\mathcal{M}_F(\bar{\mathbb{R}}^d)$, namely,

$$\tilde{d}(v_1, v_2) := \inf \{ |\langle v_1 - v_2, f\rangle| : \quad |f(x)| \le 1 \; \forall x, \quad |f(x) - f(y)| \le |x - y| \; \forall x, y \}.$$

Then

$$\tilde{d}(V^{n,\delta}(t), V^n(t)) \le \frac{1}{n} \sum_{i=1}^n \left(M^{\delta,i}(t) \vee M^i(t) \right) \left(|\mu^{\delta,i}(t) - \mu^i(t)| + |\log M^{\delta,i}(t) - \log M^i(t)| \right).$$

Thus

$$\tilde{\mathbb{E}} \tilde{d}(V^{n,\delta}(t), V^n(t)) \le c_3(T) \delta^\alpha.$$

It is clear that $d \le \tilde{d}$. So

$$\begin{aligned} \tilde{\mathbb{E}} \left(d(\bar{V}^n(t), V(t)) \right) &\le \tilde{\mathbb{E}} \tilde{d}(V^{n,1/n^{1/2\alpha}}, V^n(t)) + \tilde{\mathbb{E}} d(V^n(t), V(t)) \\ &\le c_2(T) (n^{-1/2\alpha})^\alpha + \frac{c_1(T)}{\sqrt{n}} = \frac{c_3(T)}{\sqrt{n}}. \end{aligned}$$

□

4.2 Case 2: Random $\tilde{\Sigma}$

In this subsection we discuss a case when the volatility is a random process with the following structure: $\sigma(t)$ is a function of $\mu(t)$ plus a white noise, namely,

$$d\sigma_{ij}(t) = h^{ij}(\mu(t)) dt + dB^{ij}(t), \qquad 1 \le i \le j \le d, \tag{4.13}$$

where $B^{ij}(t)$, $1 \le i \le j \le d$, are independent Brownian motions. Note that the independence assumption is imposed for the ease of presentation only. In fact, the arguments below remain valid if we replace the $\frac{d(d+1)}{2}$-dimensional Brownian motion $(B^{ij}(t),\ 1 \le i \le j \le d)$ by a linear transformation of this process.

In this case, we have a classical filtering problem with observations $\tilde{S}(t)$ and $\Sigma(t) \equiv (\sigma_{ij}(t))$ given by (3.6) and (4.13) respectively. Then

$$\tilde{B}_{ij}(t) := \sigma_{ij}(t) - \int_0^t \langle U(s), h^{ij} \rangle \, ds$$

are Brownian motions adapted to $\mathcal{G}_t$ and are independent of $\nu(t)$. Namely, $\tilde{\nu}(t) := (\tilde{B}_{ij}(t),\ \nu(t);\ 1 \le i \le j \le d)$ forms the $\left(\frac{d(d+1)}{2} + d \right)$-dimensional innovation process for the filtering problem. The FKK equation can be derived similar to the arguments as in Subsection 4.1 leading to (4.10). Namely, we only need to replace $\nu(t)$ and

$\Big(\mu(t) - \frac{1}{2}\tilde{A}(t)\Big)^* \Sigma(t)^{-1}$ by $\tilde{\nu}(t)$ and $\Big(h_{ij}(\mu(t)),\ \Big(\mu(t) - \frac{1}{2}\tilde{A}(t)\Big)^* \Sigma(t)^{-1},\ 1 \le i \le j \le d\Big)$ respectively. The numerical scheme can also be given by employing a similar method in Subsection 4.1. We leave the details to the interested reader.

Remark 4.5. The condition (4.1) does not hold for the present model. In fact, $\tilde{B}_{ij}(t)$ is a $\mathcal{G}_t$-martingale independent of $\nu(t)$; hence it cannot be represented as the stochastic integrals with respect to $\nu(t)$.

5 Optimization

In this section, we derive the optimal strategy of the partially observed mean–variance problem (2.3) in three steps. First, we derive from (3.1) a constraint on the terminal wealth $x(T)$. Then, we solve a static optimization problem under this constraint to find the best terminal wealth, $x^*(T)$. After that, we show that there is a portfolio such that $x^*(T)$ is its terminal wealth. Finally, we give a numerical scheme in solving the backward SDE involved in deriving the optimal portfolio and prove the convergence of the proposed scheme.

5.1 The optimal value of $x(T)$

Let

$$\begin{aligned}\rho(t) \quad := \quad & \exp\Bigg(- \int_0^t \mu_0(s)ds - \sum_{i,j=1}^{d} \int_0^t \sigma_{ij}^{-1}(s)(\bar{\mu}_i(s) - \mu_0(s))d\nu_j(s) \\ & -\frac{1}{2}\sum_{j=1}^{d} \int_0^t \left| \sum_{i=1}^{d} \sigma_{ij}^{-1}(s)(\bar{\mu}_i(s) - \mu_0(s)) \right|^2 ds\Bigg),\end{aligned}$$

where, with an abuse of notation, $\sigma_{ij}^{-1}(s)$ denotes the ijth element of $\Sigma(s)^{-1}$. Denote

$$\theta_j(t) := \sum_{i=1}^{d} \sigma_{ij}^{-1}(t)(\bar{\mu}_i(t) - \mu_0(t)). \tag{5.1}$$

By Itô's formula, we get

$$d\rho(t) = -\rho(t)\mu_0(t)dt - \sum_{j=1}^{d} \rho(t)\theta_j(t)d\nu_j(t), \quad \rho(0) = 1. \tag{5.2}$$

Applying Itô's formula to (5.2) and (3.1), we have

$$d(x(t)\rho(t)) = \rho(t)\sum_{i,j=1}^{d}\sigma_{ij}(t)u_i(t)d\nu_j(t) - x(t)\sum_{j=1}^{d}\rho(t)\theta_j(t)d\nu_j(t). \tag{5.3}$$

Therefore, $x(t)\rho(t)$ is a $\mathcal{G}_t$-martingale and hence,

$$x(t) = \rho(t)^{-1}\mathbb{E}(\rho(T)x(T)|\mathcal{G}_t).$$

In particular, taking $t = 0$ we have

$$\mathbb{E}(\rho(T)x(T)) = x(0) = x_0. \tag{5.4}$$

Definition 5.1. A contingent claim $v \in \mathbb{H} := L^2(\Omega, \mathcal{G}_T, P)$ is called *attainable* if there is $\Phi(s) \in L^2_{\mathcal{G}}(0, T; \mathbb{R}^d)$ such that

$$v\rho(T) = \mathbb{E}(v\rho(T)) + \int_0^T \Phi(s)^* d\nu(s).$$

Denote the collection of all attainable contingent claims by $AC(\mathcal{G})$. It is easy to see that $AC(\mathcal{G})$ is a subspace of $\mathbb{H}$. Denote by $\mathbb{H}_0$ the closure of $AC(\mathcal{G})$.

Definition 5.2. The market is complete if $AC(\mathcal{G}) = \mathbb{H}$.

Remark 5.3. If $\Sigma(t)$ is non-random as in Subsection 4.1, then the market is complete.

Now we seek

$$\min_{v\in\mathbb{H}_0}\mathbb{E}(v - z)^2 \tag{5.5}$$

subject to constraints

$$\mathbb{E}v = z \quad \text{and} \quad \mathbb{E}(\rho(T)v) = x_0. \tag{5.6}$$

Theorem 5.4. *Let α and β be the orthogonal projections on $\mathbb{H}_0$ of 1 and $\rho(T)$ respectively. Then the optimal solution to the optimization problem (5.5) and (5.6) is given by*

$$v = \frac{\left(z\langle\beta,\beta\rangle_{\mathbb{H}} - x_0\langle\alpha,\beta\rangle_{\mathbb{H}}\right)\alpha + \left(-z\langle\alpha,\beta\rangle_{\mathbb{H}} + x_0\langle\alpha,\alpha\rangle_{\mathbb{H}}\right)\beta}{\langle\alpha,\alpha\rangle_{\mathbb{H}}\langle\beta,\beta\rangle_{\mathbb{H}} - \langle\alpha,\beta\rangle_{\mathbb{H}}^2}. \tag{5.7}$$

Proof: Note that

$$\mathbb{E}(v-z)^2 = \mathbb{E}(v-z\alpha)^2 + z^2\mathbb{E}(1-\alpha)^2.$$

So, the optimization problem becomes

$$\min_{v\in\mathbb{H}_0} \|v-z\alpha\|_{\mathbb{H}}^2$$

subject to constraints

$$\langle v,\alpha\rangle_{\mathbb{H}} = z \quad \text{and} \quad \langle v,\beta\rangle_{\mathbb{H}} = x_0. \tag{5.8}$$

Using Lagrange multipliers, we define

$$f(v,\lambda_1,\lambda_2) := \|v-z\alpha\|_{\mathbb{H}}^2 - 2\lambda_1(\langle v,\alpha\rangle_{\mathbb{H}} - z) - 2\lambda_2(\langle v,\beta\rangle_{\mathbb{H}} - x_0),\ (v,\lambda_1,\lambda_2)\in\mathbb{H}\times\mathbb{R}^2.$$

Taking Fréchet derivative and setting it to be zero, we have

$$2(v-z\alpha) - 2\lambda_1\alpha - 2\lambda_2\beta = 0.$$

This implies

$$v = z\alpha + \lambda_1\alpha + \lambda_2\beta.$$

Plugging the above into the constraints (5.8), we obtain the values of λ_1 and λ_2, which lead to (5.7). □

5.2 Replicate v

In this subsection, we seek the wealth process $x(t)$ which satisfies (3.1) and $x(T)=v$, where $v\in\mathbb{H}_0$ is given by (5.7). Namely, we seek a solution to the following backward SDE:

$$\begin{cases} dx(t) = \Big(x(t)\mu_0(t) + \sum_{j=1}^d(\bar{\mu}_j(t)-\mu_0(t))u_j(t)\Big)\,dt + \sum_{i,j=1}^d \sigma_{ij}(t)u_i(t)d\nu_j(t),\ 0\le t\le T, \\ x(T) = v. \end{cases} \tag{5.9}$$

Let

$$Z_j(t) := \sum_{i=1}^d \sigma_{ij}(t)u_i(t).$$

Then

$$u_i(t) = \sum_{j=1}^{d} \sigma_{ij}^{-1}(t) Z_j(t) \tag{5.10}$$

and (5.9) becomes

$$\begin{cases} dx(t) = \Big(x(t)\mu_0(t) + \sum_{j=1}^{d} \theta_j(t) Z_j(t)\Big)\, dt + \sum_{j=1}^{d} Z_j(t) d\nu_j(t), \ 0 \le t \le T, \\ x(T) = v. \end{cases} \tag{5.11}$$

If $\mathcal{F}_t^{\nu} = \mathcal{G}_t$, then (5.11) is the usual backward SDE (BSDE) whose solution exists, assuming that $\mu_0(t)$ and $\theta_j(t)$ are essentially bounded (cf. [28]). However, it is well-known that, in general, $\mathcal{F}_t^{\nu} \neq \mathcal{G}_t$.

Now we give a proof to the existence of a unique square integrable solution to the BSDE (5.11), under the following additional condition:

Assumption (UB): For some $\delta > 0$, $A(t) \ge \delta I$ a.s., a.e.$t \ge 0$, and $\mu_0(t)$ and $\mu(t)$ are essentially bounded.

Under this assumption, $\theta_j(t)$ is also essentially bounded.

Proposition 5.5. *If $v \in \mathbb{H}_0$, then (5.11) has a unique $\mathcal{G}_t$-adapted, square integrable solution $(x(t), Z_j(t),\ j = 1, 2, \cdots, d)$.*

Proof: If $(x(t), Z_j(t),\ j = 1, 2, \cdots, d)$ is a $\mathcal{G}_t$-adapted, square integrable solution to (5.11), as in (5.3), we have

$$x(t)\rho(t) = x_0 + \sum_{j=1}^{d} \int_0^t \rho(s) \left(\sum_{i=1}^{d} \sigma_{ij}(s) u_i(s) - x(s)\theta_j(s) \right) d\nu_j(s)$$

is a $\mathcal{G}_t$-local martingale. Hence there is an increasing sequence of $\mathcal{G}_t$-stopping times $\{\tau_n\}$ with $\tau_n \to T$ as $n \to \infty$ such that for each n,

$$x(t \wedge \tau_n)\rho(t \wedge \tau_n) = \mathbb{E}(x(T \wedge \tau_n)\rho(T \wedge \tau_n)|\mathcal{G}_t).$$

For any fixed $t \in [0, T]$,

$$x(t \wedge \tau_n)\rho(t \wedge \tau_n) \le \sup_{0 \le s \le T} x(s) \sup_{0 \le s \le T} \rho(s),$$

whereas the right hand side of the above is a square integrable random variable by virtue of the Cauchy–Schwartz inequality and the standard L^2 estimation on the super norm of the solutions to SDEs. Hence we obtain by dominated convergence theorem that

$$x(t)\rho(t) = \mathbb{E}(x(T)\rho(T)|\mathcal{G}_t).$$

Namely,

$$x(t) = \rho(t)^{-1}\mathbb{E}(v\rho(T)|\mathcal{G}_t). \tag{5.12}$$

This implies the uniqueness of the solution.

To prove the existence we first assume that $v \in AC(\mathcal{G})$. We show that $x(t)$ given by (5.12) is a solution to (5.11). As $v \in AC(\mathcal{G})$, $v\rho(T)$ is square integrable in view of Definition 5.1, and we have the representation

$$\mathbb{E}(v\rho(T)|\mathcal{G}_t) = \mathbb{E}(v\rho(T)) + \sum_{j=1}^{d}\int_0^t \Phi^j(s)d\nu_j(s), \tag{5.13}$$

where each Φ^j is square integrable. Define

$$Z_j(t) := x(t)\theta_j(t) + \rho(t)^{-1}\Phi^j(t), \quad 0 \le t \le T. \tag{5.14}$$

By Itô's formula, it is easy to show that $(x(t), Z(t)) \equiv (x(t), Z_j(t),\ j = 1, 2, \cdots, d)$ satisfies (5.11). Moreover, a stopping time argument exactly as in [28, pp. 352–353] establishes the square integrability of $(x(t), Z(t))$.

Next, let $v \in \mathbb{H}_0$. Then there is a sequence $\{v_n\} \subset AC(\mathcal{G})$ such that $v_n \to v$ in $\mathbb{H}$. By the above proof there is a unique square integrable solution $(x_n(t), Z_n(t))$ to (5.11) with $x_n(T) = v_n$. Moreover,

$$\sup_n \mathbb{E}\int_0^T (|x_n(t)|^2 + |Z_n(t)|^2)dt \le K \sup_n \mathbb{E}|v_n|^2 < +\infty;$$

see p. 349, Theorem 2.2 in [28]. This implies that $(x_n(t), Z_n(t))$ is a bounded sequence in $L^2_{\mathcal{G}}(0, T; \mathbb{R}^{d+1})$. Hence there is a subsequence (still denoted as $(x_n(t), Z_n(t))$), along with $(x(t), Z(t)) \in L^2_{\mathcal{G}}(0, T; \mathbb{R}^{d+1})$, so that

$$(x_n(t), Z_n(t)) \to (x(t), Z(t)) \quad \text{weakly in } L^2_{\mathcal{G}}(0, T; \mathbb{R}^{d+1}).$$

By Mazur's theorem there is a sequence $(\tilde{x}_n(t), \tilde{Z}_n(t))$, each element of which is a convex combination of those in $\{(x_n(t), Z_n(t))\}$, so that

$$(\tilde{x}_n(t), \tilde{Z}_n(t)) \to (x(t), Z(t)) \quad \text{(strongly) in } L^2_{\mathcal{G}}(0, T; \mathbb{R}^{d+1}).$$

Since (5.11) is a linear equation, by a standard technique we conclude that $(x(t), Z(t))$ is a square integrable solution to (5.11). $\square$

Summarizing, we get

Theorem 5.6. *For every z, there exists an optimal portfolio to the mean–variance problem (2.3), which is a self-financed, admissible portfolio replicating v given by (5.7). Moreover, this optimal portfolio $u(t)$ is given by (5.10) and the corresponding optimal wealth process is $x(t)$, where $(x(t), Z_j(t),\ j = 1, 2, \cdots, d)$ is the square integrable solution to (5.11).*

So solving our partially observed mean–variance problem boils down to solving the BSDE (5.11). Numerical solutions to some classes of nonlinear BSDEs have been developed lately [29, 8]. However, in those works the drift coefficients of the BSDEs are assumed to be deterministic functions. In our model, the coefficients $\mu_0(t)$ and $\theta_j(t)$ are random in general. It is to the best of our knowledge that how to solve such a BSDE numerically remains open. In the next subsection, we shall give a numerical solution to this BSDE based upon the constructive proof of the last theorem.

5.3 Numerical solution

In this subsection, we assume that the market is complete (see Definition 5.2) and seek a numerical solution to (5.11). By virtue of the constructive proof of Proposition 5.5 the solution is given by (5.12) and (5.14). We now propose a numerical scheme to approximate (5.12) and (5.14).

To start with, note that as the market is complete, $\alpha = 1$ and $\beta = \rho(T)$. By (5.7), we get

$$v = z + \frac{x_0 - z\mathbb{E}\rho(T)}{\mathrm{Var}(\rho(T))}(\rho(T) - \mathbb{E}\rho(T)). \tag{5.15}$$

As in the proof of Proposition 5.5, the key to solving (5.11) is the martingale representation of the $\mathcal{G}_t$-martingale $\mathbb{E}(\rho(T)v|\mathcal{G}_t)$. We will establish a particle representation for this martingale.

Let (θ^1, ν^1), (θ^2, ν^2), $\cdots$ be independent copies of (θ, ν) which appeared in (5.2). Now we define $\rho^i(t, t')$, $t, t' \geq 0$, in two steps. First, for $t \leq t'$, let $\rho^i(t, t') := \rho(t)$ which is given by (5.2). Second, for $t > t'$, let $\rho^i(t, t')$ be given by (5.2) with (b, ν) replaced by (b^i, ν^i):

$$d\rho^i(t,t') = -\rho^i(t,t')\mu_0(t)dt - \sum_{j=1}^{d} \rho^i(t,t')\theta^i_j(t)d\nu^i_j(t), \quad \rho^i(t',t') = \rho(t'). \tag{5.16}$$

Let $v^i(T,t)$ be given by (5.15) with $\rho(T)$ replaced by $\rho^i(T,t)$:

$$v^i(T,t) = z + \frac{x_0 - z\mathbb{E}\rho(T)}{\mathrm{Var}(\rho(T))}(\rho^i(T,t) - \mathbb{E}\rho(T)). \tag{5.17}$$

Theorem 5.7. *Let v be given by (5.15). Then*

$$\mathbb{E}(\rho(T)v|\mathcal{G}_t) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \rho^i(T,t)v^i(T,t). \tag{5.18}$$

Proof: Note that the SDE (5.16) has a unique strong solution. Therefore, there exists a measurable functional $F_{t,T}$ such that

$$(\rho^i(T,t), v^i(T,t)) = F_{t,T}(\theta|_{[0,t]}, \nu|_{[0,t]}, \theta^i|_{[t,T]}, \nu^i|_{[t,T]}).$$

Note that $(\bar{\mu}(t), \nu(t))$ are $\mathcal{G}_t$-measurable. Further, as discussed in Section 2 $\sigma^{-1}_{ij}(t)$ and $\mu_0(t)$ are also $\mathcal{G}_t$-measurable. Thus, $\theta(t)$ defined by (5.1) is $\mathcal{G}_t$-measurable and we have,

$$(\rho^i(T,t), v^i) = G_{t,T}(S|_{[0,t]}, \theta^i|_{[t,T]}, \nu^i|_{[t,T]})$$

for a measurable functional $G_{t,T}$. By the independence of $S|_{[0,t]}$, $(\theta^i|_{[t,T]}, \nu^i|_{[t,T]})$, $i = 1, 2, \cdots$, it follows from the strong law of law numbers under the conditional probability (given $\mathcal{G}_t$) that (5.18) holds. □

For notational simplicity, we now assume $d = T = 1$. Denote the process in (5.18) by $N(t)$. By (5.13), we have

$$\langle N, \nu\rangle_t = \int_0^t \Phi(s)ds$$

which can be approximated by (cf. Jacod and Shiryaev [9], Theorem 1.4.47)

$$\sum_{k=1}^{m} (N(t_k) - N(t_{k-1}))(\nu(t_k) - \nu(t_{k-1}).$$

Based on the above, we now approximate Φ by piecewise constants, i.e., approximate Φ on $(\frac{k}{n}, \frac{k+1}{n})$ by

$$\begin{aligned}\Phi^n\left(\frac{k}{n}\right) \quad &:= \quad n\sum_{j=1}^{m}\left(N\left(\frac{k-1}{n}+\frac{j}{mn}\right) - N\left(\frac{k-1}{n}+\frac{j-1}{mn}\right)\right) \\ &\qquad \times\left(\nu\left(\frac{k-1}{n}+\frac{j}{mn}\right) - \nu\left(\frac{k-1}{n}+\frac{j-1}{mn}\right)\right) \\ &\qquad n = 1, 2, \cdots\end{aligned} \tag{5.19}$$

where $m = m_n$ is to be chosen late.

Lemma 5.8. *For $1 < p < 2$, we have*

$$
\begin{aligned}
& \mathbb{E}\left|\Phi^n\left(\frac{k}{n}\right) - n\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(s)ds\right|^p \qquad (5.20) \\
\leq\ & c\left(m^{\epsilon-p} + n^{p/2}m^{-p/2}\right)\left(\mathbb{E}\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(t)^2 dt\right)^{\frac{p}{2}},
\end{aligned}
$$

for any $\epsilon > 0$, where c depends only on p.

Proof: Let

$$
\pi(t) := \frac{k-1}{n} + \frac{j-1}{mn}, \qquad \text{for } \frac{k-1}{n} + \frac{j-1}{mn} \leq t < \frac{k-1}{n} + \frac{j}{mn}.
$$

Apply Itô's formula, we have

$$
\begin{aligned}
& \Phi^n\left(\frac{k}{n}\right) - n\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(s)ds \\
=\ & n\int_{\frac{k-1}{n}}^{\frac{k}{n}} \left\{\Phi(t)(\nu(t) - \nu(\pi(t))) + (N(t) - N(\pi(t)))\right\} d\nu(t).
\end{aligned}
$$

For $p \in (1,2)$, by Doob's inequality, we have

$$
\begin{aligned}
& \mathbb{E}\left(\left|\Phi^n\left(\frac{k}{n}\right) - n\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(s)ds\right|^p\right) \\
\leq\ & c_p n^p \mathbb{E}\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \left\{\Phi(t)^2(\nu(t) - \nu(\pi(t)))^2 + (N(t) - N(\pi(t)))^2\right\} dt\right)^{\frac{p}{2}} \\
\leq\ & c_p n^p \mathbb{E}\left(\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(t)^2 dt\right)^{\frac{p}{2}} \sup_{\frac{k-1}{n} \leq t \leq \frac{k}{n}} (\nu(t) - \nu(\pi(t)))^p\right) \\
& + c_p n^p \mathbb{E}\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} (N(t) - N(\pi(t)))^2 dt\right)^{\frac{p}{2}} \\
\leq\ & c_p n^p \left(\mathbb{E}\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(t)^2 dt\right)^{\frac{p}{2}} \left(\mathbb{E} \sup_{\frac{k-1}{n} \leq t \leq \frac{k}{n}} (\nu(t) - \nu(\pi(t)))^{\frac{2p}{2-p}}\right)^{\frac{2-p}{2}} \\
& + c_p n^p \left(\mathbb{E}\int_{\frac{k-1}{n}}^{\frac{k}{n}} (N(t) - N(\pi(t)))^2 dt\right)^{\frac{p}{2}}.
\end{aligned}
$$

Note that for i.i.d. normal random variables ξ_i with mean 0 and variance σ^2, we have

$$\mathbb{E} \sup_{1\le i\le m} \xi_i^p \le \left(\mathbb{E} \sup_{1\le i\le m} \xi_i^{p/\epsilon}\right)^{\epsilon} \le \left(\mathbb{E}\sum_{i=1}^m \xi_i^{p/\epsilon}\right)^{\epsilon} \le cm^{\epsilon}\sigma^p$$

for any $\epsilon > 0$. Thus

$$\left(\mathbb{E} \sup_{\frac{k-1}{n}\le t\le \frac{k}{n}} (\nu(t)-\nu(\pi(t)))^{\frac{2p}{2-p}}\right)^{\frac{2-p}{2}} \le \frac{cm^{\epsilon}}{(nm)^p}.$$

On the other hand,

$$\begin{aligned}\mathbb{E}\int_{\frac{k-1}{n}}^{\frac{k}{n}} (N(t)-N(\pi(t)))^2 dt &= \mathbb{E}\int_{\frac{k-1}{n}}^{\frac{k}{n}}\int_{\pi(t)}^{t} \Phi(s)^2 ds dt \\ &= \mathbb{E}\sum_{j=1}^m \int_{\frac{k-1}{n}+\frac{j-1}{mn}}^{\frac{k-1}{n}+\frac{j}{mn}} \int_{\frac{k-1}{n}+\frac{j-1}{mn}}^{t} \Phi(s)^2 ds dt \\ &\le \frac{1}{nm}\mathbb{E}\int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi(t)^2 dt.\end{aligned}$$

(5.20) then follows easily. □

Next we need to approximate $N(k\delta)$, $k = 0, 1, \cdots, mn$, where $\delta = \frac{1}{mn}$. To this end, we need to approximate $\rho^i(t,t')$ by time-discretization. Recall that $\rho^i(t,t')$ is given by (5.2) for $t \le t'$ and (5.16) for $t > t'$.

For $j \le k$, let

$$\begin{aligned}\rho^{i,\delta}(j\delta,k\delta) &:= \rho^i((j-1)\delta,k\delta) - \rho^i((j-1)\delta,k\delta)\mu_0((j-1)\delta)\delta \\ &\quad -\rho^i((j-1)\delta,k\delta)\theta((j-1)\delta)(\nu(j\delta)-\nu((j-1)\delta)),\end{aligned}$$

with $\rho^{i,\delta}(0,k\delta) := 1$. For $j > k$, let

$$\begin{aligned}\rho^{i,\delta}(j\delta,k\delta) &:= \rho^i((j-1)\delta,k\delta) - \rho^i((j-1)\delta,k\delta)\mu_0((j-1)\delta)\delta \\ &\quad -\rho^i((j-1)\delta,k\delta)\theta((j-1)\delta)(\nu^i(j\delta)-\nu^i((j-1)\delta)).\end{aligned}$$

Let

$$v^{i,\delta}(T,k\delta) := z + \frac{x_0 - z\mathbb{E}\rho(T)}{\mathrm{Var}(\rho(T))}(\rho^{i,\delta}(T,k\delta) - \mathbb{E}\rho(T)),$$

and

$$N^{n,\delta}(k\delta) := \frac{1}{n}\sum_{i=1}^n \rho^{i,\delta}(T,k\delta)v^{i,\delta}(T,k\delta). \tag{5.21}$$

Now, in view of (5.14), we define

$$Z^n(t) := \rho^n\left(\frac{[nt]}{n}\right)^{-1}\left(\theta^n\left(\frac{[nt]}{n}\right)N^{n,\delta}\left(\frac{[nt]}{n}\right)+\Phi^{n,\delta}\left(\frac{[nt]}{n}\right)\right) \tag{5.22}$$

where $\delta = \frac{1}{m_n n}$ (with m_n suitably chosen), $\Phi^{n,\delta}$ is defined as in (5.19) for Φ^n with N replaced by $N^{n,\delta}$, the approximate θ^n and ρ^n of θ and ρ can be defined by the same method as in Section 4.1 such that $\forall\ \beta > 1$,

$$\mathbb{E}\sup_{0\le t\le T}\left|\rho^n\left(\frac{[nt]}{n}\right)^{-1}\theta^n\left(\frac{[nt]}{n}\right)N^{n,\delta}\left(\frac{[nt]}{n}\right)-\rho(t)^{-1}\theta(t)N(t)\right|^\beta \to 0 \tag{5.23}$$

and

$$\mathbb{E}\sup_{0\le t\le T}\left|\rho^n\left(\frac{[nt]}{n}\right)^{-1}-\rho(t)^{-1}\right|^\beta \to 0. \tag{5.24}$$

Finally, we define the *approximate portfolio* by

$$u^n(t) := \Sigma(t)^{-1}Z^n(t).$$

Since we do not know the continuity of $\Phi(s)$, we cannot obtain $\Phi^n(s) \to \Phi(s)$ (cf. (5.23)). Therefore, it is not clear whether $u^n(t) \to u(t)$. However, we have the convergence of their corresponding terminal wealths.

Theorem 5.9. *Suppose that* $\mathbb{E}(|v\rho(T)|^2) < \infty$. *Then*

$$\mathbb{E}|x^n(T) - x(T)| \to 0 \quad \textit{as } n \to \infty. \tag{5.25}$$

Proof: For simplicity of notation, we assume $\mu_0(t) \equiv 0$. Then by (5.11), we have

$$\begin{aligned}\mathbb{E}|x^n(T)-x(T)| \;&\le\; c\mathbb{E}\int_0^T|\theta^n(t)Z^n(t)-\theta(t)Z(t)|^2dt\\ &\quad +c\mathbb{E}\left(\int_0^T|Z^n(t)-Z(t)|^2dt\right)^{1/2}.\end{aligned}$$

We only estimate the second term (the first can be evaluated similarly). Note that by (5.22) and (5.19),

$$\begin{aligned}|Z^n(t)-Z(t)| \;&\le\; \left|\rho^n\left(\frac{[nt]}{n}\right)^{-1}\theta^n\left(\frac{[nt]}{n}\right)N^{n,\delta}\left(\frac{[nt]}{n}\right)-\rho(t)^{-1}\theta(t)N(t)\right|\\ &\quad +\left|\rho^n\left(\frac{[nt]}{n}\right)^{-1}\Phi^{n,\delta}\left(\frac{[nt]}{n}\right)-\rho(t)^{-1}\Phi(t)\right|.\end{aligned}$$

By (5.23), the first term on the right hand side of the above inequality tends to 0. Note that

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^T \left|\rho^n\left(\frac{[nt]}{n}\right)^{-1}\Phi^{n,\delta}\left(\frac{[nt]}{n}\right) - \rho(t)^{-1}\Phi(t)\right|^2 dt\right)^{1/2}\right] \\
\le\ &\mathbb{E}\left[\left(\int_0^T \rho^n\left(\frac{[nt]}{n}\right)^{-2}\left|\Phi^{n,\delta}\left(\frac{[nt]}{n}\right) - \Phi^n\left(\frac{[nt]}{n}\right)\right|^2 dt\right)^{1/2}\right] \\
&+\mathbb{E}\left[\left(\int_0^T \left|\rho^n\left(\frac{[nt]}{n}\right)^{-2} - \rho(t)^{-2}\right|\Phi^n\left(\frac{[nt]}{n}\right)^2 dt\right)^{1/2}\right] \\
&+\mathbb{E}\left[\left(\frac{1}{n}\sum_{k=1}^n \left|\Phi^n\left(\frac{k}{n}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}}\Phi(s)ds\right|^2 \sup_{0\le t\le T}\rho(t)^{-2}\right)^{1/2}\right] \\
&+\mathbb{E}\left[\left(\sum_{k=1}^n \int_{\frac{k-1}{n}}^{\frac{k}{n}}\left|\Phi(t) - n\int_{\frac{k-1}{n}}^{\frac{k}{n}}\Phi(s)ds\right|^2 \sup_{0\le t\le T}\rho(t)^{-2}\right)^{1/2}\right].
\end{aligned}$$

The convergence of the first term follows from the similar arguments as in Subsection 4.1, that of the second term from (5.24), and that of the fourth term from the same arguments as in the proof of Lebesgue's continuity theorem, namely, first approximate Φ by uniformly continuous functions and then prove the conclusion for such functions. Finally, the third term is dominated by

$$\begin{aligned}
&\mathbb{E}\max_{1\le k\le n}\left|\Phi^n\left(\frac{k}{n}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}}\Phi(s)ds\right| \sup_{0\le t\le T}\rho(t)^{-1} \\
\le\ &\left(\mathbb{E}\sum_{k=1}^n \left|\Phi^n\left(\frac{k}{n}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}}\Phi(s)ds\right|^p\right)^{1/p}\left(\mathbb{E}\sup_{0\le t\le T}\rho(t)^{-\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \\
\le\ &cn\left(m^{\epsilon-p} + n^{p/2}m^{-p}\right)
\end{aligned}$$

which converges to 0 if we take $m = n^\beta$ with $\beta > \max\left(\frac{1+\epsilon}{p}, \frac{1}{2}\right)$. □

Remark 5.10. It follows from (5.25) that $\mathbb{E}x^n(T) \to \mathbb{E}x(T)$. However, it is not clear whether $\text{Var}(x^n(T)) \to \text{Var}(x(T))$. To achieve this, we need a higher moment (in p) in (5.20) for which we require $\int_0^T \Phi(s)^2 ds$ to have a higher moment. To be precise, if we assume that $v\rho(T)$ has a moment of order $p' > 2$, then

$$\mathbb{E}\left[\left(\int_0^T \Phi(s)^2 ds\right)^{p'/2}\right] < \infty.$$

The proof of (5.20) can be adapted for $2 < p < p'$ and, hence, (5.25) can be strengthen to

$$\mathbb{E}\left(|x^n(T) - x(T)|^2\right) \to 0.$$

In this case we get the convergence of both $\mathbb{E}x^n(T)$ and $\mathrm{Var}(x^n(T))$.

Remark 5.11. If the market is not complete, the numerical approximation of this section remains valid if we replace 1 and $\rho^i(T,t)$ by their projections α and β^i on $\mathbb{H}_0$ (i.e., use the formula (5.7) instead of (5.15)). However, it remains *open* how to calculate α and β^i numerically.

Acknowledgment: Most of this work was done when the first author was visiting the Chinese University of Hong Kong, the hospitality and financial support of which are gratefully acknowledged. The authors also thank the associate editor and two anonymous referees for constructive comments that have led to an improved version.

References

[1] A.G. Bhatt, G. Kallianpur and R.L. Karandikar (1995). Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering. *Ann. Probab.* **23**, 1895-1938.

[2] T.R. Bielecki, H. Jin, S.R. Pliska and X.Y. Zhou (2005). Continuous-time mean–variance portfolio selection with bankruptcy prohibition. *Math. Fin.,* **15**, 213-244.

[3] M.J. Brennan and Y. Xia (2001). Assessing asset pricing anomalies. *Rev. Finan. Stud.*, **14**, 905-942.

[4] J.B. Detemple (1986). Asset pricing in a production economy with incomplete information, *J. Finan.*, **41**, 383-391.

[5] M.U. Dothan and D. Feldman (1986). Equilibrium interest rates and multiperiod bonds in a partially observable economy, *J. Finan.*, **41**, 369-382.

[6] S. N. Ethier and T. G. Kurtz (1985). *Markov Processes : Characterization and Convergence.* Wiley, New York.

[7] G. Gennotte (1986). Optimal portfolio choice under incomplete information, *J. Finan.*, **41**, 733-746.

[8] E. Gobet, J.-P. Lemor and X. Warin (2004). A regression-based Monte Carlo method to solve backward stochastic differential equations, *Ann. Appl. Probab.*, **15**, 2172-2202.

[9] J. Jacod and A. N. Shiryaev (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag, New York.

[10] G. Kallianpur (1980). *Stochastic Filtering Theory.* Springer-Verlag, New York.

[11] Karatzas, I. and S.E. Shreve (1991). *Methods of Mathematical Finance.* Springer-Verlag.

[12] T. Kurtz and J. Xiong (1999). Particle representations for a class of nonlinear SPDEs. *Stoch. Proc. & Their Appl.,* **83**, 103-126.

[13] T. Kurtz and J. Xiong (2000). Numerical solutions for a class of SPDEs with application to filtering. *Stochastics in Finite and Infinite Dimension: In Honor of Gopinath Kallianpur.* Edited by T. Hida, R. Karandikar, H. Kunita, B. Rajput, S. Watanabe and J. Xiong. *Trends in Mathematics.* Birkhauser, Boston.

[14] P. Lakner (1998). Optimal trading strategy for an investor: the case of partial information. *Stoch. Proc. & Their Appl.*, **76**, 77-97.

[15] D. Li and W.L. Ng (2000). Optimal dynamic portfolio selection: multi-period mean-variance formulation. *Math. Fin.*, **10**, 387-406.

[16] A.E.B. Lim (2004). Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Math. Oper. Res.,* **29**, 132-161.

[17] R.S. Liptser and A.N. Shiryaev (2001). *Statistics of Random Processes, I, General Theory.* Second Edition. Springer, New York.

[18] J. Ma, P. Protter and J. Yong (1994). Solving forward-backward stochastic differential equations explicitly– a four step scheme. *Prob. Th. & Rel. Fields,* **98**, 339-359.

[19] H.M. Markowitz (1952). Portfolio selection. *J. of Finan.* **7**, 77-91.

[20] H. Nagai and S. Peng (2002). Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon. *Ann. Appl. Probab.*, **12**, 173-195.

[21] H. Pham (2001). Mean–variance hedging for partially observed drift processes. *Int. J. Theor. & Appl. Finan.*, **4**, 263-284.

[22] L.C.G. Rogers (2001). The relaxed investor and parameter uncertainty. *Finan. & Stoch.*, **5**, 131-154.

[23] M.C. Steinbach (2001). Markowitz revisited: mean–variance models in financial portfolio analysis. *SIAM Rev.*, **43**, 31-85.

[24] D.W. Stroock and S.R.S. Varadhan (1979) *Multidimensional Diffusion Processes.* Springer-Verlag, Berlin.

[25] J. Xia (2005). Mean–variance portfolio choice: quadratic partial hedging. *Math. Fin.*, **15**, 533-538.

[26] Y. Xia (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *J. of Finan.*, **LVI**, 205-246.

[27] Z.J. Yang and J. Xiong. Maximizing the expected utility from terminal wealth with partial information and the valuation of information. Submitted.

[28] J. Yong and X.Y. Zhou (1999). *Stochastic Control: Hamiltonian Systems and HJB Equations.* Springer, New York.

[29] J. Zhang (2004). A numerical scheme for BSDEs. *Ann. Appl. Probab.*, **14**, 459-488.

[30] Y.G. Zhao and W.T. Ziemba (2000). Mean-variance versus expected utility in dynamic investment analysis. Working paper, Faculty of Commerce and Business Administration, University of British Columbia.

[31] X.Y. Zhou and D. Li (2000). Continuous-time mean-variance portfolio selection: a stochastic LQ framework. *Appl. Math. & Optim.*, **42**, 19-33.