

# Optimal investment and hedging under partial information

MICHAEL MONOYIOS

Mathematical Institute, University of Oxford

[www.maths.ox.ac.uk/~monoyios](http://www.maths.ox.ac.uk/~monoyios)

Tutorial Lectures for the session on

*Inverse and Partial Information Problems*

at the

Special Semester on Stochastics with Emphasis on Finance

Johan Radon Institute for Computational and Applied Mathematics

Linz

Austria

September 2008

September 2, 2008

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Filtering theory</b>	<b>4</b>
2.1	Observation model . . . . .	4
2.2	Innovations process . . . . .	4
2.2.1	The Innovations Conjecture . . . . .	5
2.3	Signal process model . . . . .	7
2.4	Fundamental filtering equation . . . . .	7
2.4.1	Linear observations . . . . .	9
2.4.2	Linear observations and linear signal . . . . .	9
2.5	Multi-dimensional Kalman-Bucy filter . . . . .	11
<b>3</b>	<b>Merton problem with uncertain drift</b>	<b>12</b>
3.1	Full information case . . . . .	13
3.1.1	Portfolio optimisation via convex duality . . . . .	13
3.2	Partial information case . . . . .	15
<b>4</b>	<b>Optimal hedging of basis risk with partial information</b>	<b>18</b>
4.1	Basis risk model: full information case . . . . .	18
4.1.1	Perfect correlation case . . . . .	19
4.1.2	Incomplete case . . . . .	20
4.2	Partial information case . . . . .	22
4.2.1	Choice of prior . . . . .	22
4.2.2	Two-dimensional Kalman-Bucy filter . . . . .	23
4.2.3	Optimal hedging with random drifts . . . . .	25
4.2.4	The primal problem . . . . .	26
4.2.5	Dual problem and optimal hedge . . . . .	27
4.2.6	Stochastic control representation of the indifference price . . . . .	30
4.2.7	Analytic approximation for the indifference price . . . . .	30

<b>5</b>	<b>Investment with inside information and drift uncertainty</b>	<b>31</b>
5.1	Linear filtering on an expanded filtration . . . . .	32
5.2	Computing the information drift . . . . .	34
5.3	Optimal investment for an insider with drift parameter uncertainty . . . . .	36
5.3.1	Anticipative Brownian information . . . . .	37

### Abstract

We first give an exposition of filtering theory, then consider the Merton optimal investment problem when the agent does not know the drift parameter of the underlying stock. This is taken to be a random variable with a Gaussian prior distribution, which is updated via a Kalman filter. The resulting problem of optimal investment with a random drift can be treated as a full information problem, and an explicit solution is possible. We then treat an incomplete market hedging problem. A claim on a non-traded asset is hedged using a correlated traded asset, and the hedger is once again uncertain of the true values of the drifts of each asset. After filtering, the resulting problem with random drifts is solved in the case that each asset's prior distribution has the same variance. Analytic approximations for the optimal hedging strategy are obtained. Finally, we examine an optimal investment problem with inside information, in which the insider does not know the true drift of the stock. Explicit solutions are possible, after first enlarging the filtration to accommodate the insider's additional knowledge, then filtering the asset price drift.

## 1 Introduction

These lectures examine some problems of optimal investment, and of optimal hedging of a contingent claim in an incomplete market, when the agent's information set is restricted to stock price observations, possibly augmented by some additional information related to the terminal value of a stock price.

In classical models of financial mathematics, one usually specifies a probability space  $(\Omega, \mathcal{F}, P)$  equipped with a background filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ , and then writes down some stochastic process  $S = (S_t)_{0 \leq t \leq T}$  for an asset price, such that  $S$  is adapted to the filtration  $\mathbb{F}$ . A typical example would be the Black-Scholes (henceforth, BS) model of a stock price, following the geometric Brownian motion

$$dS_t = \sigma S_t(\lambda dt + dB_t), \quad (1)$$

where  $B$  is an  $\mathbb{F}$ -Brownian motion and the volatility  $\sigma > 0$  and the Sharpe ratio  $\lambda$  are assumed to be known constants. Of course, this is a strong assumption that the agent is assumed to be able to observe the Brownian motion (henceforth, BM) process  $B$ , as well as the stock price process  $S$ . We refer to this as a *full information* scenario. In this case, an agent would use  $\mathbb{F}$ -adapted trading strategies in  $S$ .

We shall relax the full information assumption. We shall assume that the agent can only observe the stock price process, and not the Brownian motion  $B$ , and that the constants  $\sigma, \lambda$  are not (heroically) assumed to be known. The agent's trading strategies must also be adapted to the *observation filtration*  $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}$  generated by  $S$ . We refer to this as a *partial information* scenario.

In this case, the parameter  $\lambda$  would be regarded as an unknown constant whose value needs to be determined from price data. In principle, one would also have to apply this philosophy to the volatility  $\sigma$ , but we shall make the approximation that price observations are continuous, so that  $\sigma$  can be computed from the quadratic variation  $[S]_t$  of the stock price, since we have

$$\frac{d[S]_t}{dt} = \sigma^2 S_t^2. \quad (2)$$

One way to model the uncertainty in our knowledge of  $\lambda$  is to consider it as an  $\mathbb{F}$ -adapted process, or as a random variable (measurable with respect to  $\mathcal{F}_0$ ) with a given initial distribution (the *prior* distribution), which is updated in the face of new price information, that is, as the observation filtration  $\hat{\mathbb{F}}$  evolves. This is an example of a *filtering* problem, which is to compute the best estimate of a random variable given observations up to time  $t \in [0, T]$ , and hence given

the sigma algebra  $\hat{\mathcal{F}}_t, t \in [0, T]$  In the case of the BS model (1), where we model  $\lambda$  as a random variable, we might be interested in computing the conditional expectation

$$\hat{\lambda}_t := E[\lambda | \hat{\mathcal{F}}_t].$$

We shall see that the effect of filtering is that the model (1) may be replaced by a model specified on the filtered probability space  $(\Omega, \hat{\mathcal{F}}_T, \hat{\mathbb{F}}, P)$  and written as

$$dS_t = \sigma S_t(\hat{\lambda}_t dt + d\hat{B}_t), \quad (3)$$

where  $\hat{B}$  is an  $\hat{\mathbb{F}}$ -BM. This model may now be treated as a full information model, since both  $\hat{B}$  and  $\hat{\lambda}$  are  $\hat{\mathbb{F}}$ -adapted processes. The price we have paid for restoring a full information scenario is that the constant parameter  $\lambda$  has been replaced by a random process  $\hat{\lambda}$ .

The process of replacing a partial information model with an effective full information model is usually only achievable in special circumstances, such as Gaussian prior distributions and certain linearity properties in the relation between the observable and unobservable processes, as we shall see in the next section.

In the rest of these lectures, we first give an exposition of filtering theory, culminating in the linear filtering case (the Kalman-Bucy filter). We then apply the results to the Merton problem of optimal investment, which seeks a trading strategy to maximise expected utility of terminal wealth. The problem is then to choose a trading strategy  $\pi = (\pi_t)_{0 \leq t \leq T}$ , the wealth placed in stock at each time  $t \in [0, T]$ , to maximise the functional

$$J(x; \pi) := E \left[ x + \int_0^T \frac{\pi_t}{S_t} dS_t \right],$$

where  $x$  is the initial capital at time 0. The *value function* is the maximum expected utility expressed as a function of  $x$ :

$$u(x) := \sup_{\pi \in \mathcal{A}} J(x; \pi), \quad (4)$$

where  $\mathcal{A}$  denotes a set of admissible trading strategies, which will depend on whether we have a full information model, a partial information model, or one with some other relevant information structure. The main distinction between a full information model and a partial information model is that:

- in a full information setting,  $\pi$  is  $\mathbb{F}$ -adapted, where  $\mathbb{F}$  is (typically) the Brownian filtration generated by  $B$ , and we take  $S$  to follow the stochastic differential equation (1), with  $\sigma, \lambda$  known constants;
- in a partial information setting,  $\pi$  is  $\hat{\mathbb{F}}$ -adapted, where  $\hat{\mathbb{F}}$  is the observation filtration generated by  $S$ , and we take  $S$  to follow the stochastic differential equation (3), with  $\sigma$  a constant given by (2) and  $\hat{\lambda}$  an  $\hat{\mathbb{F}}$ -adapted process.

We shall see how to characterise  $\hat{\lambda}$  using filtering theory.

We shall then move on to the hedging of a claim in an incomplete market setting under partial information. Specifically, we shall consider a *basis risk* model involving the optimal hedging of a contingent claim on a non-tradeable asset  $Y$  using a traded stock  $S$ , correlated with  $Y$ , when the hedger is restricted to trading strategies in  $S$  that are adapted to the observation filtration  $\hat{\mathbb{F}}$  generated by the asset prices. In the full information case the asset prices are correlated log-Brownian motions given by

$$dS_t = \sigma S_t(\lambda dt + dB_t), \quad dY_t = \beta Y_t(\theta dt + dW_t),$$

where the Brownian motions  $B, W$  are correlated with correlation  $\rho \in [-1, 1]$ . The parameters  $\sigma > 0$ ,  $\lambda, \beta > 0$  and  $\theta$  are assumed constant.

This market is complete when the correlation is perfect, but incomplete otherwise. A number of studies have used exponential indifference valuation methods to hedge the claim in an optimal

manner in a full information scenario, and we outline these results before moving on to the partial information case, where we assume the hedger does not know with certainty the drifts of  $S$  and  $Y$ .

Finally we return to the Merton optimal investment problem (4) when the agent is assumed to have some additional information in the form of knowledge of the value of a random variable  $I$ , whose value is measurable with respect to time- $T$  information.

## 2 Filtering theory

A good reference for this section is Chapter VI.8 of Rogers & Williams [21].

The setting is a probability space  $(\Omega, \mathcal{F}, P)$  equipped with a filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ . All our processes are assumed to be  $\mathbb{F}$ -adapted. Note that  $\mathbb{F}$  is *not* the observation filtration. Let us call  $\mathbb{F}$  the *background filtration*. We consider two processes, both taken to be one-dimensional (for simplicity):

- a *signal process*  $U = (U_t)_{0 \leq t \leq T}$  which is not directly observable (so unobservable);
- an *observation process*  $O = (O_t)_{0 \leq t \leq T}$ , which is observable and somehow correlated with  $U$ , so that by observing  $O$  we can say something about the distribution of  $U$ .

Let  $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}$  denote the observation filtration generated by  $O$ . That is,

$$\hat{\mathcal{F}}_t := \sigma(O_s; 0 \leq s \leq t), \quad 0 \leq t \leq T.$$

The *filtering problem* is to *compute the conditional distribution* of the signal  $U_t, t \in [0, T]$ , given observations up to that time. Or, equivalently, to compute the conditional expectation

$$E[f(U_t) | \hat{\mathcal{F}}_t], \quad 0 \leq t \leq T,$$

where  $f: \mathbb{R} \rightarrow \mathbb{R}$  is some test function.

To proceed further, we need to specify some particular model for the signal and observation processes.

### 2.1 Observation model

Let  $B = (B_t)_{0 \leq t \leq T}$  be an  $\mathbb{F}$ -Brownian motion, let  $H = (H_t)_{0 \leq t \leq T}$  be an  $\mathbb{F}$ -adapted process satisfying

$$E \int_0^T H_t^2 dt < \infty,$$

and we shall assume the observation process  $O$  is of the form

$$O_t = \int_0^t H_u du + B_t, \quad 0 \leq t \leq T. \quad (5)$$

The typical situation will be where  $H_t = h(t, U_t)$ , a deterministic function  $h: [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  of time and the current signal value. The yet more specialised situation (that we shall mainly focus on) will be the *linear* case when  $h(t, x) = G(t)x$ , with  $G(\cdot)$  a deterministic function. Then  $H_t = G(t)U_t$  and the observation process stochastic differential equation (SDE) is

$$dO_t = G(t)U_t dt + dB_t, \quad (\text{linear observation model}).$$

### 2.2 Innovations process

Introduce the notation

$$\hat{\phi}_t := E[\phi_t | \hat{\mathcal{F}}_t],$$

for any process  $\phi$ . Define the  $\hat{\mathbb{F}}$ -adapted *innovations process*

$$N_t := O_t - \int_0^t \hat{H}_u du, \quad 0 \leq t \leq T. \quad (6)$$

**Proposition 1** *The innovations process  $N$  is an  $\hat{\mathbb{F}}$ -Brownian motion.*

**Proof** From (5) and (6) we have

$$N_t = \int_0^t (H_u - \hat{H}_u) du + B_t.$$

With  $s \leq t$ , we have

$$\begin{aligned} E[N_t | \hat{\mathcal{F}}_s] - N_s &= E \left[ \int_s^t (H_u - \hat{H}_u) du + B_t \middle| \hat{\mathcal{F}}_s \right] \\ &= E \left[ \int_s^t \left( E[H_u | \hat{\mathcal{F}}_u] - \hat{H}_u \right) du \middle| \hat{\mathcal{F}}_s \right] + E \left[ E(B_t - B_s | \mathcal{F}_s) \middle| \hat{\mathcal{F}}_s \right] = 0, \end{aligned}$$

using the Tower property of conditional expectation. So  $N$  is a continuous  $\hat{\mathbb{F}}$ -martingale with quadratic variation  $[N]_t = [B]_t = t$ , so  $N$  is an  $\hat{\mathbb{F}}$ -Brownian motion. □

### 2.2.1 The Innovations Conjecture

Denote by  $\mathbb{F}^N := (\mathcal{F}_t^N)_{0 \leq t \leq T}$  the filtration generated by  $N$ , so that  $\mathcal{F}_t^N := \sigma(N_s; s \leq t)$ . Since  $N$  is  $\hat{\mathbb{F}}$ -adapted, we have  $(\mathcal{F}_t^N) \subseteq (\hat{\mathcal{F}}_t)$ . For *linear* systems, we shall see that we also have

$$\mathcal{F}_t^N = \hat{\mathcal{F}}_t, \quad 0 \leq t \leq T, \quad (7)$$

so that in this case the observations and the innovations represent the same information (because there is an invertible map that derives one from the other). The “innovations conjecture” (of Kailath) is that the identity (7) holds in general, but we now know that this is not the case. However, the following positive and very useful result *is* true.

**Theorem 1** *Every local  $\hat{\mathbb{F}}$ -martingale  $M$  admits a representation of the form*

$$M_t = M_0 + \int_0^t \Phi_s dN_s, \quad 0 \leq t \leq T,$$

where  $\Phi$  is  $\hat{\mathbb{F}}$ -adapted and  $\int_0^T \Phi_t^2 dt < \infty$  a.s. If  $M$  happens to be a square-integrable martingale, then  $\Phi$  can be chosen so that  $E \int_0^T \Phi_t^2 dt < \infty$ .

To prove this result, recall the following well-known result on representation of local martingales with respect to a Brownian filtration. See Theorems 3.4.2 and 3.4.15 in Karatzas and Shreve [11] for more details on this and other classical results in stochastic calculus.

**Theorem 2 (Local martingale representation)** *Let  $W$  be a Brownian motion and let  $\mathbb{F}^W$  denote its natural filtration. Every local martingale  $M$  with respect to  $\mathbb{F}^W$  admits a representation of the form*

$$M_t = M_0 + \int_0^t b_s dW_s, \quad 0 \leq t < \infty,$$

for an  $\mathbb{F}^W$ -adapted process satisfying  $\int_0^t b_s^2 ds < \infty$  almost surely for every  $0 < t < \infty$ . In particular, every such  $M$  has continuous sample paths. If  $M$  happens to be a square-integrable martingale ( $EM_t^2 < \infty, \forall t \geq 0$ ), then  $b$  can be chosen so that  $E \int_0^t b_s^2 ds < \infty$  for every  $0 < t < \infty$ .

Note that, if only the innovations conjecture (7) were true in general, then Theorem 1 would follow directly from Theorem 2. As the innovations conjecture is not true in general, we shall prove the theorem by performing a measure change that turns  $O$  into a Brownian motion, then apply Theorem 2, then invert the change of measure to revert back to the innovations process  $N$ .

**Proof of Theorem 1** We carry this out only in the case of bounded  $H$ , and for the case when  $M$  is a martingale, to present the ideas with the minimum of technical fuss.

If  $H$  is bounded, then so is  $\hat{H}$ , so the process

$$Z_t := \mathcal{E}(-\hat{H} \cdot N)_t, \quad 0 \leq t \leq T,$$

is an  $\hat{\mathbb{F}}$ -martingale. By the Girsanov Theorem, since  $N$  is an  $\hat{\mathbb{F}}$ -BM under  $P$ , then the process

$$N_t + \int_0^t \hat{H}_s ds = O_t,$$

that is, the observation process, is a  $(\tilde{P}, \hat{\mathbb{F}})$ -BM under the probability measure  $\tilde{P}$  defined on  $(\Omega, \hat{\mathbb{F}})$  by

$$\left. \frac{d\tilde{P}}{dP} \right|_{\hat{\mathcal{F}}_t} = Z_t, \quad 0 \leq t \leq T.$$

Notice that the inverse likelihood ratio is

$$\begin{aligned} \left. \frac{dP}{d\tilde{P}} \right|_{\hat{\mathcal{F}}_t} &= \Lambda_t := \frac{1}{Z_t} = \exp \left( \int_0^t \hat{H}_s dN_s + \frac{1}{2} \hat{H}_s^2 ds \right) \\ &= \exp \left( \int_0^t \hat{H}_s dO_s - \frac{1}{2} \hat{H}_s^2 ds \right) \\ &= \mathcal{E}(\hat{H} \cdot O)_t. \end{aligned}$$

The stochastic differential equations for  $Z, \Lambda$  are therefore

$$Z_t = 1 - \int_0^t Z_s \hat{H}_s dN_s, \quad \Lambda_t = 1 + \int_0^t \Lambda_s \hat{H}_s dO_s. \quad (8)$$

Using the so-called *Bayes rule*, for  $s < t$  and an  $\hat{\mathcal{F}}_t$ -measurable random variable  $X$ :

$$\tilde{E}[X | \hat{\mathcal{F}}_s] = \frac{1}{Z_s} E[Z_t X | \hat{\mathcal{F}}_s],$$

we find that, since  $M$  is a  $P$ -martingale, then  $\Lambda M$  is a  $\tilde{P}$ -martingale: for  $s < t$ , we have

$$\tilde{E}[\Lambda_t M_t | \hat{\mathcal{F}}_s] = \frac{1}{Z_s} E[Z_t \Lambda_t M_t | \hat{\mathcal{F}}_s] = \Lambda_s M_s.$$

An application of Theorem 2 gives a representation of the form

$$\Lambda_t M_t = \Lambda_0 M_0 + \int_0^t \Psi_s dO_s = \Lambda_0 M_0 + \int_0^t \Psi_s (dN_s + \hat{H}_s ds). \quad (9)$$

Now from (9), (8) and the integration by parts formula,<sup>1</sup> we obtain

$$\begin{aligned} M_t &= (\Lambda_t M_t) Z_t \\ &= (\Lambda_t M_t) Z_t + \int_0^t (\Lambda_s M_s) dZ_s + \int_0^t Z_s d(\Lambda_s M_s) + [\Lambda M, Z]_t \\ &= M_0 + \int_0^t \Lambda_s M_s (-Z_s \hat{H}_s dN_s) + \int_0^t Z_s \Psi_s (dN_s + \hat{H}_s ds) - \int_0^t Z_s \hat{H}_s \Psi_s d[N]_s \\ &= M_0 + \int_0^t (Z_s \Psi_s - M_s \hat{H}_s) dN_s \\ &= M_0 + \int_0^t \Phi_s dN_s, \end{aligned}$$

for  $\Phi = Z\Psi - M\hat{H}$ .

<sup>1</sup>For any two processes  $X, Y$ ,

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X, Y]_t.$$

□

To proceed further, we now need yet more structure, this time on the signal process  $U$ .

### 2.3 Signal process model

We take the signal process to be of the form

$$U_t = U_0 + \int_0^t b(s, U_s) ds + \int_0^t \sigma(s, U_s) dW_s, \quad (10)$$

where  $W$  is a Brownian motion correlated with  $B$  in the observation model (5)

$$\frac{d}{dt}[W, B]_t = \rho, \quad \rho \in [-1, 1],$$

and the functions  $b : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  and  $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  satisfy the Lipschitz and linear growth conditions

$$\begin{aligned} |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| &\leq K|x - y|, \quad \forall x, y \in \mathbb{R}, \\ |b(t, x)| + |\sigma(t, x)| &\leq K(1 + |x|), \quad \forall x \in \mathbb{R}, \end{aligned}$$

for some real  $K > 0$ . Then there exists a unique process  $U$  that satisfies (10).

Let  $f \in C_0^2(\mathbb{R})$  be a twice continuously differentiable function with compact support<sup>2</sup>. The generator of  $U$  is  $\mathcal{A}_t$  given by

$$\mathcal{A}_t f(x) = b(t, x)f'(x) + \frac{1}{2}\sigma^2(t, x)f''(x).$$

For brevity we use the notation

$$f_t \equiv f(U_t), \quad (\mathcal{A}f)_t \equiv \mathcal{A}_t f(U_t), \quad 0 \leq t \leq T.$$

By Itô's formula we have that

$$M_t^f := f_t - f_0 - \int_0^t (\mathcal{A}f)_s ds = \int_0^t \sigma(s, U_s) f'(U_s) dW_s \quad (11)$$

is an  $\mathbb{F}$ -local martingale, and in fact a square-integrable martingale if  $f$  is of compact support. We assume this is the case. The cross-variation of  $M^f$  with  $B$  is

$$[M^f, B]_t = \int_0^t \sigma(s, U_s) f'(U_s) d[W, B]_s = \int_0^t \rho \sigma(s, U_s) f'(U_s) ds =: \int_0^t \alpha_s^f ds, \quad (12)$$

where we have defined the process  $\alpha^f$  by

$$\alpha_t^f := \rho \sigma(t, U_t) f'(U_t), \quad 0 \leq t \leq T.$$

### 2.4 Fundamental filtering equation

The fundamental filtering theorem is the following.

**Theorem 3** *For the observation and signal process models of (5) and (10) we have, for every  $f \in C_0^2(\mathbb{R})$ , and with  $f_t \equiv f(U_t)$ , using the notation  $\hat{\phi}_t := E[\phi_t | \hat{\mathcal{F}}_t]$  for any process  $\phi$ , the fundamental filtering equation*

$$\hat{f}_t = \hat{f}_0 + \int_0^t (\widehat{\mathcal{A}f})_s ds + \int_0^t \left( \widehat{f_s H_s} - \hat{f}_s \hat{H}_s + \hat{\alpha}_s^f \right) dN_s, \quad 0 \leq t \leq T. \quad (13)$$

The proof will require the following lemma.

**Lemma 1** *Consider two  $\mathbb{F}$ -adapted processes  $V, C$  with  $E|V_t| < \infty, \forall t \in [0, T]$  and  $E \int_0^T |C_t| dt < \infty$ . If  $J_t := V_t - \int_0^t C_s ds$  is an  $\mathbb{F}$ -martingale, then*

$$\hat{J}_t := \hat{V}_t - \int_0^t \hat{C}_s ds \quad \text{is an } \hat{\mathbb{F}}\text{-martingale.}$$

<sup>2</sup>So  $f$  is zero outside of a compact set.

**Proof** For  $s \leq t$ , writing  $\int_0^t C_u du = \int_0^s C_u du + \int_s^t C_u du$  and using the fact that  $J$  is an  $\mathbb{F}$ -martingale, we have

$$\begin{aligned} E \left[ V_t - \int_0^s C_u du - \int_s^t C_u du \middle| \mathcal{F}_s \right] &= V_s - \int_0^s C_u du \\ \Rightarrow E \left[ V_t - \int_s^t C_u du \middle| \mathcal{F}_s \right] &= V_s, \end{aligned} \quad (14)$$

which we shall use shortly.

Now consider

$$\begin{aligned} E[\hat{J}_t | \hat{\mathcal{F}}_s] &= E \left[ \hat{V}_t - \int_0^t \hat{C}_u du \middle| \hat{\mathcal{F}}_s \right] \\ &= E \left[ \hat{V}_t - \int_0^s \hat{C}_u du - \int_s^t \hat{C}_u du \middle| \hat{\mathcal{F}}_s \right] \\ &= E \left[ \hat{V}_t - \int_s^t \hat{C}_u du \middle| \hat{\mathcal{F}}_s \right] - \int_0^s \hat{C}_u du \\ &= E \left[ E[V_t | \hat{\mathcal{F}}_t] - \int_s^t E[C_u | \hat{\mathcal{F}}_u] du \middle| \hat{\mathcal{F}}_s \right] - \int_0^s \hat{C}_u du \\ &= E[V_t | \hat{\mathcal{F}}_s] - \int_s^t E[C_u | \hat{\mathcal{F}}_s] du - \int_0^s \hat{C}_u du \\ &= E[E[V_t | \mathcal{F}_s] | \hat{\mathcal{F}}_s] - \int_s^t E[E[C_u | \mathcal{F}_s] | \hat{\mathcal{F}}_s] du - \int_0^s \hat{C}_u du \\ &= E \left[ E \left[ V_t - \int_s^t C_u du \middle| \mathcal{F}_s \right] \middle| \hat{\mathcal{F}}_s \right] - \int_0^s \hat{C}_u du \\ &= E[V_s | \hat{\mathcal{F}}_s] - \int_0^s \hat{C}_u du \quad (\text{using (14)}) \\ &= \hat{V}_s - \int_s^t \hat{C}_u du \\ &= \hat{J}_s. \end{aligned}$$

□

**Proof of Theorem 3** Recall from (11) that

$$f_t = f_0 + \int_0^t (\mathcal{A}f)_s ds + M_t^f, \quad (15)$$

where  $M_t^f = \int_0^t \sigma(U_s) f'(U_s) dW_s$  is an  $\mathbb{F}$ -martingale. So, using Lemma 1 and Theorem 1 we have

$$\hat{M}_t^f := \hat{f}_t - \hat{f}_0 - \int_0^t (\widehat{\mathcal{A}f})_s ds = \hat{\mathbb{F}}\text{-martingale} =: \int_0^t \Phi_s dN_s, \quad (16)$$

for a suitable  $\hat{\mathbb{F}}$ -adapted process  $\Phi$  such that  $E \int_0^T \Phi_t^2 dt < \infty$ . We want to compute  $\Phi$ , to establish that

$$\Phi_t = \widehat{f_t H_t} - \hat{f}_t \hat{H}_t + \hat{\alpha}_t^f, \quad 0 \leq t \leq T. \quad (17)$$

This will be accomplished by computing  $E[f_t O_t | \hat{\mathcal{F}}_t] = \hat{f}_t O_t$  in two ways and comparing the results.



On the one hand, using (15), (5), (12) and the integration by parts formula,

$$\begin{aligned} f_t O_t &= f_0 O_0 + \int_0^t f_s dO_s + \int_0^t O_s df_s + [f, O]_t \\ &= \int_0^t f_s (H_s ds + dB_s) + \int_0^t O_s ((\mathcal{A}f)_s ds + dM_s^f) + [M^f, B]_t \\ &= \int_0^t (f_s H_s + O_s (\mathcal{A}f)_s + \alpha_s^f) ds + \mathbb{F}\text{-martingale.} \end{aligned}$$

So by Lemma 1,

$$\widehat{f}_t O_t = \hat{f}_t O_t = \int_0^t \left( \widehat{f}_s H_s + O_s (\widehat{\mathcal{A}f})_s + \hat{\alpha}_s^f \right) ds + \widehat{\mathbb{F}}\text{-martingale.} \quad (18)$$

On the other hand, from (16), (6) and the integration by parts formula, we obtain

$$\begin{aligned} \hat{f}_t O_t &= \hat{f}_0 O_0 + \int_0^t \hat{f}_s dO_s + \int_0^t O_s d\hat{f}_s + [\hat{f}, O]_t \\ &= \int_0^t \hat{f}_s (\hat{H}_s ds + dN_s) + \int_0^t O_s \left( (\widehat{\mathcal{A}f})_s ds + \Phi_s dN_s \right) + \int_0^t \Phi_s ds \\ &= \int_0^t \left( \hat{f}_s \hat{H}_s + O_s (\widehat{\mathcal{A}f})_s + \Phi_s \right) ds + \widehat{\mathbb{F}}\text{-martingale.} \end{aligned} \quad (19)$$

Comparing (18) and (19), the difference between the bounded variation parts is a continuous martingale of bounded variation, so is constant, and is null at zero, so is identically zero, and therefore (17) holds. □

### 2.4.1 Linear observations

Take  $H_t = h(t, U_t) = G(t)U_t$  and  $f(x) = x^k, k = 1, 2, \dots$ . Then we obtain from (13):

$$\hat{U}_t = \hat{U}_0 + \int_0^t \widehat{b(s, U_s)} ds + \int_0^t \left[ G(s) \left( \widehat{U_s^2} - (\hat{U}_s)^2 \right) + \rho \widehat{\sigma(s, U_s)} \right] dN_s, \quad (k=1), \quad (20)$$

$$\begin{aligned} \widehat{U}_t^k &= \widehat{U}_0^k + k \int_0^t \left( \widehat{b(s, U_s) U_s^{k-1}} + \frac{1}{2} (k-1) \sigma^2(s, U_s) \widehat{U_s^{k-2}} \right) ds \\ &+ \int_0^t \left[ G(s) \left( \widehat{U_s^{k+1}} - \hat{U}_s \widehat{U_s^k} \right) + k \rho \sigma(s, U_s) \widehat{U_s^{k-1}} \right] dN_s, \quad k = 2, 3, \dots \end{aligned} \quad (21)$$

Equations (20) and (21) convey the complexity of non-linear filtering. To solve the equation for the  $k^{\text{th}}$  conditional moment, one needs to know the  $(k+1)^{\text{th}}$  conditional moment as well as  $E[g(s, U_s) | \widehat{\mathcal{F}}_s]$  for  $g(s, x) = b(s, x)x^{k-1}$ ,  $g(s, x) = \sigma^2(s, x)x^{k-2}$ ,  $g(s, x) = \sigma(s, x)x^{k-1}$ . This means the computation of conditional moments cannot be achieved by induction on  $k$  and the problem is inherently infinite dimensional except in the linear case.

### 2.4.2 Linear observations and linear signal

Now take  $h(t, x) = G(t)x$  as before, and  $b(t, x) = A(t)x$ ,  $\sigma(t, x) = C(t)$ , for deterministic functions  $A(\cdot), C(\cdot)$ , and suppose that the signal process has a Gaussian initial distribution, so that the signal and observation processes follow

$$\begin{aligned} dU_t &= A(t)U_t dt + C(t)dW_t, & U_0 &\sim N(\mu, \mathbf{v}), \\ dO_t &= G(t)U_t dt + dB_t, & O_0 &= 0, \end{aligned}$$

with  $U_0$  independent of  $B$  and of  $W$ . The two-dimensional process  $(X, Y)$  is then Gaussian, so the conditional distribution of  $U_t$  given  $\hat{\mathcal{F}}_t$  is *normal* (so, in particular, is completely characterised by its mean and variance), with mean

$$\hat{U}_t := E[U_t | \hat{\mathcal{F}}_t]$$

and variance

$$\text{var}[U_t | \hat{\mathcal{F}}_t] =: V_t = E[(U_t - \hat{U}_t)^2 | \hat{\mathcal{F}}_t] = \widehat{U}_t^2 - (\hat{U}_t)^2.$$

Notice that the initial values are

$$\hat{U}_0 = E[U_0 | \hat{\mathcal{F}}_0] = EU_0 = \mu,$$

and

$$V_0 = E[(U_0 - \hat{U}_0)^2 | \hat{\mathcal{F}}_0] = E[(U_0 - \mu)^2] = \text{var}(U_0) = v.$$

The problem then boils down to finding an algorithm for computing the sufficient statistics  $\hat{U}_t, V_t$  from their initial values  $\hat{U}_0 = \mu, V_0 = v$ .

Now, from (20) we obtain, along with the initial condition  $\hat{U}_0 = \mu$ , the SDE

$$d\hat{U}_t = A(t)\hat{U}_t dt + (G(t)V_t + \rho C(t)) dN_t, \quad \hat{U}_0 = \mu. \quad (22)$$

From (21) with  $k = 2$  we obtain

$$d\widehat{U}_t^2 = \left( C^2(t) + 2A(t)\widehat{U}_t^2 \right) dt + \left[ G(t) \left( \widehat{U}_t^3 - \hat{U}_t \widehat{U}_t^2 \right) + 2\rho C(t)\hat{U}_t \right] dN_t, \quad \widehat{U}_0^2 - \mu^2 = v. \quad (23)$$

But for a normal random variable  $X \sim N(m, \Sigma^2)$ , we have

$$E[X^3] = m(m^2 + 3\Sigma^2),$$

whence

$$\widehat{U}_t^3 = E[U_t^3 | \hat{\mathcal{F}}_t] = E[U_t | \hat{\mathcal{F}}_t] \left( \left( E[U_t | \hat{\mathcal{F}}_t] \right)^2 + 3\text{var}[U_t | \hat{\mathcal{F}}_t] \right) = \hat{U}_t \left[ \left( \hat{U}_t \right)^2 + 3V_t \right],$$

and therefore

$$\widehat{U}_t^3 - \hat{U}_t \widehat{U}_t^2 = \hat{U}_t \left[ \left( \hat{U}_t \right)^2 + 3V_t - \widehat{U}_t^2 \right] = 2V_t \hat{U}_t.$$

Using this, (22), (23) and the Itô formula, we obtain

$$\begin{aligned} dV_t &= d \left[ \widehat{U}_t^2 - \left( \hat{U}_t \right)^2 \right] \\ &= \left( C^2(t) + 2A(t)\widehat{U}_t^2 \right) dt + \left[ G(t) \left( 2V_t \hat{U}_t \right) + 2\rho C(t)\hat{U}_t \right] dN_t - 2\hat{U}_t d\hat{U}_t - d[\hat{U}]_t, \end{aligned}$$

which simplifies to the *non-stochastic* Riccati equation

$$\frac{dV_t}{dt} = (1 - \rho^2)C^2 + 2(A(t) - \rho C(t)G(t))V_t = G^2(t)V_t^2, \quad V_0 = v. \quad (24)$$

In other words, the conditional variance  $V_t$  is a *deterministic* function of  $t$ , and given by the solution of (24). Thus, there is in fact only one sufficient statistic, the conditional mean  $\hat{U}_t$  which satisfies the linear SDE (22), which is the celebrated *Kalman-Bucy filter*. We summarise all this below.

**Theorem 4 (One-dimensional Kalman-Bucy filter)** *On a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, P)$ , with  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ , let  $U$  be an  $\mathbb{F}$ -adapted signal process satisfying*

$$dU_t = A(t)U_t dt + C(t)dW_t, \quad 0 \leq t \leq T,$$

and let  $O$  be an  $\mathbb{F}$ -adapted observation process satisfying

$$dO_t = G(t)U_t dt + dB_t, \quad 0 \leq t \leq T,$$

where  $W, B$  are  $\mathbb{F}$ -Brownian motions with correlation  $\rho$ , and the coefficients  $A(\cdot), C(\cdot), G(\cdot)$  are deterministic functions satisfying

$$\int_0^T (|A(t)| + C^2(t) + G^2(t)) dt < \infty.$$

Define the observation filtration  $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}$  by

$$\hat{\mathcal{F}}_t := \sigma(O_s; 0 \leq s \leq t).$$

Suppose  $U_0$  is an  $\mathcal{F}_0$ -measurable random variable, and that the distribution of  $U_0$  is Gaussian with mean  $\mu$  and variance  $v$ , independent of  $W$  and  $B$ . Then the conditional expectation  $\hat{U}_t := E[U_t | \hat{\mathcal{F}}_t]$  for  $0 \leq t \leq T$  satisfies

$$d\hat{U}_t = A(t)\hat{U}_t dt + [G(t)V_t + \rho C(t)] dN_t, \quad \hat{U}_0 = \mu,$$

where  $N$  is the innovations process, an  $\hat{\mathbb{F}}$ -Brownian motion satisfying

$$dN_t = dO_t - G(t)\hat{U}_t dt,$$

and  $V_t = \text{var}[U_t | \hat{\mathcal{F}}_t]$  is the conditional variance, which is independent of  $\hat{\mathcal{F}}_t$  and satisfies the deterministic Riccati equation

$$\frac{dV_t}{dt} = (1 - \rho^2)C^2(t) + 2[A(t) - \rho C(t)G(t)]V_t - G^2(t)V_t^2, \quad V_0 = v.$$

**Remark 1 (Validity of innovations conjecture for linear systems)** It is now straightforward to confirm the validity of the innovations conjecture  $\mathcal{F}_t^N = \hat{\mathcal{F}}_t, 0 \leq t \leq T$ , for linear systems. The solution  $\hat{U}$  of the SDE (22) is adapted to the filtration  $\mathbb{F}^N$  of the driving Brownian motion  $N$ , so  $(\mathcal{F}_t^{\hat{U}}) \subseteq (\mathcal{F}_t^N)$ , where  $\mathcal{F}_t^{\hat{U}} = \sigma(\hat{U}_s; 0 \leq s \leq t)$ . So from the relation  $O_t = N_t + \int_0^t G(s)\hat{U}_s ds$  we see that  $O$  is  $\mathbb{F}^N$ -adapted, i.e.  $(\hat{\mathcal{F}}_t) \subseteq (\mathcal{F}_t^N)$ . Because the reverse inclusion  $(\mathcal{F}_t^N) \subseteq (\hat{\mathcal{F}}_t)$  always holds, the two filtrations are the same.

## 2.5 Multi-dimensional Kalman-Bucy filter

A multi-dimensional version of the Kalman-Bucy filter can be derived along similar lines to the one-dimensional case. We state the result below.

**Theorem 5** Consider a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, P)$ , with  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ , and two  $\mathbb{F}$ -adapted processes  $U, O$  as given below.

Let  $U = (U_t)_{0 \leq t \leq T}$  be an  $n$ -dimensional signal process satisfying

$$dU_t = A(t)U_t dt + C(t)dW_t, \quad U_0 \sim N(\mu, v), \quad (\text{linear signal}), \quad (25)$$

where  $U_0 \sim N(\mu, v)$  denotes an  $n$ -dimensional  $\mathcal{F}_0$ -measurable Gaussian vector with mean  $\mu \in \mathbb{R}^n$  and covariance matrix  $v \in \mathbb{R}^n \times \mathbb{R}^n$ , independent of the  $d$ -dimensional Brownian motion  $W$ , and where  $A(t) \in \mathbb{R}^n \times \mathbb{R}^n, C(t) \in \mathbb{R}^n \times \mathbb{R}^d$ .

Let  $O = (O_t)_{0 \leq t \leq T}$  be an  $m$ -dimensional observation process satisfying

$$dO_t = G(t)U_t dt + D(t)dB_t, \quad Z_0 = 0, \quad (\text{linear observations}),$$

where  $G(t) \in \mathbb{R}^m \times \mathbb{R}^n, C(t) \in \mathbb{R}^m \times \mathbb{R}^k$ , and  $B$  is a  $k$ -dimensional Brownian motion independent of  $W$  and  $U_0$ .

We assume that  $A, C, G, D$  are bounded on bounded intervals, that  $DD^T$  is non-singular, and that  $(D(t)D(t)^T)^{-1}$  is bounded on every bounded  $t$ -interval.

Let  $\hat{\mathbb{F}} = (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}$  denote the observation filtration generated by  $O$ , so that  $\hat{\mathcal{F}}_t = \sigma(O_s; 0 \leq s \leq t)$ .

The conditional expectation vector  $\hat{U}_t := E[U_t | \hat{\mathcal{F}}_t], 0 \leq t \leq T$ , satisfies the SDE

$$\begin{aligned} d\hat{U}_t &= A(t)\hat{U}_t dt + V_t G^T(t) (D(t)D^T(t))^{-1} (dO_t - G(t)\hat{U}_t dt), \quad \hat{U}_0 = \mu, \\ &= A(t)\hat{U}_t dt + V_t G^T(t) (D(t)D^T(t))^{-1} dN_t, \quad \hat{U}_0 = \mu, \end{aligned} \quad (26)$$

where  $N$  is the innovations process, defined by

$$N_t := O_t - \int_0^t G(s)\hat{U}_s ds, \quad 0 \leq t \leq T,$$

and satisfying

$$N_t = \int_0^t D(s)d\hat{B}_s, \quad (27)$$

where  $\hat{B}$  is a standard  $k$ -dimensional  $\hat{\mathbb{F}}$ -Brownian motion.

The error  $U_t - \hat{U}_t$  is independent of  $\hat{\mathcal{F}}_t$  and the error covariance

$$V_t := E[(U_t - \hat{U}_t)(U_t - \hat{U}_t)^T | \hat{\mathcal{F}}_t] = \text{var}[U_t | \hat{\mathcal{F}}_t],$$

satisfies the deterministic matrix Riccati equation

$$\frac{dV_t}{dt} = A(t)V_t + V_t A^T(t) - V_t G^T(t) (D(t)D^T(t))^{-1} G(t)V_t + C(t)C^T(t), \quad V_0 = v.$$

We make two remarks.

- Notice that by (27) we can rewrite (26) as

$$d\hat{U}_t = A(t)\hat{U}_t dt + V_t G^T(t) (D(t)D^T(t))^{-1} D(t)d\hat{B}_t, \quad \hat{U}_0 = \mu,$$

which is a linear SDE of the same type as (25).

- Since  $U, \hat{U}$  satisfy (25), (26) and  $U_0$  is Gaussian, then  $U_t, \hat{U}_t$  are Gaussian vectors for each  $t$ , and the error  $U_t - \hat{U}_t$  is also Gaussian:  $U_t - \hat{U}_t$  has mean 0 and covariance  $V_t$ , and  $\text{Law}(U_t | \hat{\mathcal{F}}_t) = \mathcal{N}(\hat{U}_t, V_t)$ .

### 3 Merton problem with uncertain drift

A stock price process  $S = (S_t)_{0 \leq t \leq T}$  follows

$$dS_t = \sigma S_t (\lambda dt + dB_t), \quad (28)$$

on a complete probability space  $(\Omega, \mathcal{F}, P)$  equipped with a filtration  $\mathbb{F} := (\mathcal{F}_t)_{0 \leq t \leq T}$ , and  $B = (B_t)_{0 \leq t \leq T}$  is an  $\mathbb{F}$ -Brownian motion. In a full information model,  $\sigma > 0$  and  $\lambda$  are assumed to be known constants.

In a partial information model with continuous observations, we no longer assume that we know the values of the parameters in the stochastic differential equation (SDE) (28), and we restrict trading strategies to be  $\hat{\mathbb{F}}$ -adapted, where  $\hat{\mathbb{F}}$  is the filtration generated by the stock price. Then  $\sigma$  is still known from the quadratic variation of  $S$  but the value of  $\lambda$  is not. In the partial information case, we shall model the uncertainty in  $\lambda$  by treating it as a random variable with a Gaussian prior distribution,  $\lambda \sim \mathcal{N}(\lambda_0, v_0)$  (the normal probability law of mean  $\lambda_0$  and variance  $v_0$ ), independent of  $B$ .

We consider a Merton optimal investment problem where an agent with a power utility function

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

may invest a portion of his wealth in shares and the remaining wealth in a cash account with zero interest rate (for simplicity).

### 3.1 Full information case

In this case, with  $\sigma, \lambda$  in (28) treated as known constants, the agent may use an  $\mathbb{F}$ -adapted strategy  $\theta = (\theta_t)_{0 \leq t \leq T}$ , where  $\theta_t$  is the proportion of wealth invested in shares at time  $t$ , an  $\mathbb{F}$ -adapted process satisfying  $\int_0^T \theta_t^2 dt < \infty$  almost surely. The  $\mathbb{F}$ -adapted wealth process  $X = (X_t)_{0 \leq t \leq T}$  then follows

$$dX_t = \sigma \theta_t X_t (\lambda dt + dB_t). \quad (29)$$

We define the set of admissible strategies  $\mathcal{A}$  as those strategies  $\theta$  such that  $X_t \geq 0$  almost surely for all  $t \in [0, T]$ .

The objective is to maximise expected utility of terminal wealth

$$E[U(X_T) | X_0 = x],$$

over all admissible strategies.

Let  $Q$  denote the unique martingale measure for this market. The change of measure martingale  $Z := (Z_t)_{0 \leq t \leq T}$  is given by

$$Z_t := \frac{dQ}{dP} \Big|_{\mathcal{F}_t} = \mathcal{E}(-\lambda \cdot B)_t = \exp\left(-\lambda B_t - \frac{1}{2} \lambda^2 t\right),$$

and satisfies the SDE

$$dZ_t = -\lambda Z_t dB_t, \quad 0 \leq t \leq T.$$

Under  $Q$ , the process  $B^Q$  defined by

$$B_t^Q := B_t + \lambda t,$$

is a Brownian motion. Then we can write (29) as

$$dX_t = \sigma \theta_t X_t dB_t^Q, \quad (30)$$

so  $X$  is a local  $Q$ -martingale. Using the Itô formula we write the solution to (30) given  $X_0 = x$  as

$$X_t = x \mathcal{E}(\sigma \theta \cdot B^Q)_t, \quad 0 \leq t \leq T,$$

so by the Novikov condition  $X$  will be a  $Q$ -martingale provided that  $\theta$  satisfies the integrability condition

$$E \exp \frac{1}{2} \sigma^2 \int_0^T \theta_t^2 dt < \infty.$$

Let us assume this is the case from now on, and we shall see that the optimal strategy does indeed satisfy this integrability condition.

#### 3.1.1 Portfolio optimisation via convex duality

Given a continuous, increasing, concave utility function  $U(\cdot)$ , define its *convex conjugate*  $V : \mathbb{R}^+ \rightarrow \mathbb{R}$  by

$$V(\eta) := \sup_{x \in \text{dom}(U)} [U(x) - x\eta], \quad \eta > 0. \quad (31)$$

Note that (31) is equivalent to the bidual relation

$$U(x) = \inf_{y \in \mathbb{R}^+} [V(y) + xy], \quad x \in \text{dom}(U). \quad (32)$$

We shall employ the classical dual approach to portfolio optimisation, which is summarised by the following theorem. See Karatzas [9] for a proof (though we shall justify the crucial results).

**Theorem 6** *In a complete market model with continuous wealth process  $X$  and with unique martingale measure  $Q$ , define the primal value function by*

$$u(x) := \sup_{\theta \in \mathcal{A}} EU(X_T), \quad x \in \text{dom}(U), \quad (33)$$

and define the dual value function by

$$v(\eta) := EV \left( \eta \frac{dQ}{dP} \right), \quad \eta > 0.$$

Then we have

1.  $u(x)$  and  $v(\eta)$  are conjugate:

$$v(\eta) = \sup_{x \in \text{dom}(U)} [u(x) - x\eta], \quad u(x) = \inf_{\eta > 0} [v(\eta) + x\eta],$$

so that  $u'(x) = \eta$  (equivalently,  $v'(\eta) = -x$ );

2. The optimal terminal wealth in (33) is  $X_T^*$  satisfying

$$U'(X_T^*) = \eta \frac{dQ}{dP}, \quad \text{equivalently,} \quad X_T^* = I \left( \eta \frac{dQ}{dP} \right);$$

3. The following properties for  $u'(x)$  and  $v'(\eta)$  hold true:

$$u'(x) = EU'(X_T^*), \quad v'(\eta) = E^Q V' \left( \eta \frac{dQ}{dP} \right). \quad (34)$$

To get an idea of how this theorem is obtained, we consider the maximisation of the objective functional  $EU(X_T)$  subject to the constraint  $E^Q X_T = E[Z_T X_T] = x$  via the Lagrangian

$$\begin{aligned} L(X_T, \eta) &:= EU(X_T) + \eta(x - E[Z_T X_T]) \\ &= \eta x + E[U(X_T) - \eta Z_T X_T] \\ &\leq \eta x + EV(\eta Z_T) \\ &= \eta x + v(\eta). \end{aligned}$$

where  $\eta > 0$  is a Lagrange multiplier whose role is to enforce the constraint. The first order condition with respect to the terminal wealth gives that the optimal terminal wealth  $X_T^*$  satisfies

$$U'(X_T^*) = \eta Z_T,$$

or

$$X_T^* = I(\eta Z_T),$$

where  $I$  is the inverse of  $U'$ . Substituting this into the constraint fixes the multiplier  $\eta$  as satisfying

$$E[Z_T I(\eta Z_T)] = x.$$

For power utility,  $U(x) = x^\gamma / \gamma$ , we have

$$\begin{aligned} V(\eta) &= -\frac{\eta^q}{q}, \quad q = -\frac{\gamma}{1-\gamma}, \\ I(\eta) &= \eta^{-(1-q)}. \end{aligned}$$

Applying these results we obtain the optimal terminal wealth as

$$X_T^* = (\eta Z_T)^{-(1-q)}.$$

The constraint gives the relation between the Lagrange multiplier and the initial wealth  $x$  as

$$\eta^{-(1-q)} E[Z_T^q] = x.$$

For  $t \in [0, T]$  an easy computation using the explicit expression for  $Z_T$  gives

$$E[Z_T^q | \mathcal{F}_t] = Z_t^q \exp\left(-\frac{1}{2}q(1-q)\lambda^2(T-t)\right), \quad (35)$$

so that  $\eta$  is given explicitly by

$$\eta^{-(1-q)} = x \exp\left(\frac{1}{2}q(1-q)\lambda^2 T\right).$$

The optimal wealth process  $X^* = (X_t^*)_{0 \leq t \leq T}$  is a  $Q$ -martingale, so for  $t \leq T$  we have

$$\begin{aligned} X_t^* &= E^Q[X_T^* | \mathcal{F}_t] \\ &= \frac{1}{Z_t} E[Z_T X_T^* | \mathcal{F}_t] \\ &= x \exp\left(\frac{1}{2}q(1-q)\lambda^2 t\right) Z_t^{-(1-q)}, \end{aligned} \quad (36)$$

on using the explicit forms of  $\eta$  and  $X_T^*$  along with (35).

The optimal trading strategy  $\theta^*$  is given by applying Itô's formula to compute  $dX_t^*$  and noting that the coefficient of  $dB_t$  is equal to  $\sigma\theta_t^* X_t^*$ . This gives that the optimal strategy is to keep a constant proportion of wealth in the stock:

$$\theta_t^* = \frac{\lambda}{\sigma(1-\gamma)}. \quad (37)$$

The constant proportion on the right-hand side of (37) is called the Merton proportion.

### 3.2 Partial information case

Recall the stock price process

$$dS_t = \sigma S_t(\lambda dt + dB_t) = \sigma S_t d\xi_t,$$

In the partial information case the agent is restricted to using  $\hat{\mathbb{F}}$ -adapted trading strategies, where  $\hat{\mathbb{F}}$  is the observation filtration:  $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}$ , with

$$\hat{\mathcal{F}}_t := \sigma\{\xi_u, 0 \leq u \leq t\} = \sigma\{S_u, 0 \leq u \leq t\},$$

and where the *observation process* is  $\xi = (\xi_t)_{0 \leq t \leq T}$ , defined by

$$\xi_t := \frac{1}{\sigma} \int_0^t \frac{dS_u}{S_u} = \lambda t + B_t, \quad (38)$$

corresponding to noisy observations of  $\lambda$ , with  $B$  representing the noise. The agent is restricted to using  $\hat{\mathbb{F}}$ -adapted strategies, so with continuous observations  $\sigma$  is a known constant and  $\lambda$  is an unknown constant, and hence modelled as a random variable. Let us assume that the distribution of  $\lambda$  is Gaussian,  $\lambda \sim (\lambda_0, v_0)$ , independent of  $B$ .

We are faced with a Kalman-Bucy type filtering problem whose unobservable *signal process* is the market price of risk  $\lambda$ . The signal process SDE is

$$d\lambda = 0, \quad (39)$$

and the observation process SDE is (38).

We apply Theorem 4 to the signal process  $\lambda$  in (39) and observation process  $\xi$  in (38), so that  $A(t) = C(t) = 0$  and  $G(t) = 1$  for all  $t \in [0, T]$ . Then the optimal filter  $\hat{\lambda}_t := E[\lambda | \hat{\mathcal{F}}_t]$ ,  $0 \leq t \leq T$ , satisfies

$$d\hat{\lambda}_t = v_t d\hat{B}_t, \quad \hat{\lambda}_0 = \lambda_0, \quad (40)$$

where  $v_t := E[(\lambda - \hat{\lambda}_t)^2 | \hat{\mathcal{F}}_t]$ ,  $0 \leq t \leq T$ , the conditional variance, satisfies the Riccati equation

$$\frac{dv_t}{dt} = -v_t^2, \quad (41)$$

with initial value  $v_0$ , so that

$$v_t = \frac{v_0}{1 + v_0 t}. \quad (42)$$

The process  $\hat{B}$  is an  $\hat{\mathbb{F}}$ -Brownian motion, the innovations process, satisfying

$$d\hat{B}_t = d\xi_t - \hat{\lambda}_t dt, \quad 0 \leq t \leq T. \quad (43)$$

Using this in (40), the optimal filter can also be written in terms of the observable  $\xi$  as

$$\hat{\lambda}_t = \frac{\lambda_0 + v_0 \xi_t}{1 + v_0 t}. \quad (44)$$

The effect of the filtering is that the agent is now investing in a stock with dynamics given by  $dS_t = \sigma S_t d\xi_t$  which, using (43), becomes

$$dS_t = \sigma S_t (\hat{\lambda}_t dt + d\hat{B}_t), \quad 0 \leq t \leq T. \quad (45)$$

The ( $\hat{\mathbb{F}}$ -adapted) wealth process  $X^0$  then follows

$$dX_t^0 = \sigma \theta_t^0 X_t^0 (\hat{\lambda}_t dt + d\hat{B}_t), \quad 0 \leq t \leq T, \quad (46)$$

where  $\theta_t^0$  is the proportion of wealth invested in shares at time  $t$ , an  $\hat{\mathbb{F}}$ -adapted process satisfying  $\int_0^T (\theta_t^0)^2 dt < \infty$  almost surely, and such that  $X_t^0 \geq 0$  almost surely for all  $t \in [0, T]$ . Denote by  $\mathcal{A}_0$  the set of such admissible strategies.

The objective is to maximise expected utility of terminal wealth

$$E[U(X_T^0) | X_0^0 = x],$$

over all admissible strategies. This may now be treated as a full information problem, which we solve via duality techniques.

Let  $Q^0$  denote the unique martingale measure for this market. The change of measure martingale  $Z^0 := (Z_t^0)_{0 \leq t \leq T}$  is given by

$$Z_t^0 := \frac{dQ^0}{dP} \Big|_{\hat{\mathcal{F}}_t} = \mathcal{E}(-\hat{\lambda} \cdot \hat{B})_t = \exp \left( - \int_0^t \hat{\lambda}_s d\hat{B}_s - \frac{1}{2} \int_0^t \hat{\lambda}_s^2 ds \right), \quad (47)$$

and satisfies the SDE

$$dZ_t^0 = -\hat{\lambda}_t Z_t^0 d\hat{B}_t, \quad 0 \leq t \leq T. \quad (48)$$

We may write  $Z_t^0 = f(t, \hat{\lambda}_t)$  where  $f(t, x)$  is a smooth function  $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^+$ , and apply Itô's formula along with the SDE (40) for  $\hat{\lambda}_t$  to give

$$dZ_t^0 = \left[ f_t(t, \hat{\lambda}_t) + \frac{1}{2} v_t^2 f_{xx}(t, \hat{\lambda}_t) \right] dt + v_t f_x(t, \hat{\lambda}_t) d\hat{B}_t, \quad 0 \leq t \leq T, \quad (49)$$

with subscripts of  $f$  denoting partial derivatives. Equating (48) and (49) yields the partial differential equations (PDEs)

$$\begin{aligned} v_t f_x(t, x) &= -x f(t, x), \\ f_t(t, x) + \frac{1}{2} v_t^2 f_{xx}(t, x) &= 0, \end{aligned}$$



with  $f(0, \cdot) = Z_0^0 = 1$ . The solution to these PDEs gives  $Z_t^0$  in the form

$$Z_t^0 = \left( \frac{v_0}{v_t} \right)^{\frac{1}{2}} \exp \left[ -\frac{1}{2} \left( \frac{\hat{\lambda}_t^2}{v_t} - \frac{\lambda_0^2}{v_0} \right) \right], \quad 0 \leq t \leq T. \quad (50)$$

By convex duality, the primal value function

$$u_0(x) := \sup_{\theta^0 \in \mathcal{A}_0} E[U(X_T^0) | X_0^0 = x], \quad x > 0, \quad (51)$$

and dual value function

$$v_0(y) := E[\tilde{U}(yZ_T^0)], \quad y > 0, \quad (52)$$

are convex conjugates:

$$v_0(y) = \sup_{x > 0} (u_0(x) - xy), \quad (53)$$

where  $V(y) := \sup_{x > 0} [U(x) - xy]$  is the conjugate of  $U$ . For power utility,  $V$  is given by

$$V(y) = -\frac{y^q}{q}, \quad q = -\frac{\gamma}{1-\gamma}. \quad (54)$$

Hence the dual value function is given by

$$v_0(y) = -\frac{y^q}{q} H_0,$$

where

$$H_0 := E \left[ (Z_T^0)^q \right],$$

and the primal value function is given by

$$u_0(x) = \frac{x^\gamma}{\gamma} H_0^{1-\gamma}. \quad (55)$$

The optimal terminal wealth  $X_T^{0,*}$ , attained by adopting the strategy that achieves the supremum in (51), is given by  $U'(X_T^{0,*}) = yZ_T^0$ , or, with  $I \equiv -V'$  denoting the inverse of  $U'$ ,

$$X_T^{0,*} = -V'(yZ_T^0), \quad y = u'_0(x).$$

Hence

$$X_T^{0,*} = \frac{x}{H_0} (Z_T^0)^{-(1-q)}. \quad (56)$$

The optimal wealth process  $X^{0,*}$  is a  $(Q^0, \hat{\mathbb{F}})$ -martingale, so

$$X_t^{0,*} = E^{Q^0} [X_T^{0,*} | \hat{\mathcal{F}}_t] = \frac{1}{Z_t^0} E[Z_T^0 X_T^{0,*} | \hat{\mathcal{F}}_t], \quad (57)$$

where  $E^{Q^0}$  denotes expectation under  $Q^0$ .

From (40) we deduce that for  $t \leq T$ , conditional on  $\hat{\mathcal{F}}_t$ ,  $\hat{\lambda}_T$  is normally distributed,  $\hat{\lambda}_T \sim N(\hat{\lambda}_t, v_t - v_T)$ . For a normally distributed random variable  $Y \sim N(m, s^2)$ , we have

$$E \exp(cY^2) = \frac{1}{\sqrt{1-2cs^2}} \exp \left( \frac{cm^2}{1-2cs^2} \right), \quad (58)$$

so that, given the explicit expression (50) for  $Z_t^0$ , the constant  $H_0$  and the right-hand-side of (57) can be computed in closed form. Then,  $H_0$  is given by

$$H_0 = \left[ \frac{(1+v_0T)^q}{1+qv_0T} \right]^{1/2} \exp \left[ -\frac{1}{2} \frac{q(1-q)\lambda_0^2 T}{1+qv_0T} \right], \quad (59)$$

so that (55) gives the value function explicitly, as

$$u_0(x) = \frac{x^\gamma}{\gamma} \left[ \frac{1 - \gamma - \gamma v_0 T}{(1 - \gamma)(1 + v_0 T)} \right]^{\gamma/2} \left( \frac{1 - \gamma}{1 - \gamma - \gamma v_0 T} \right)^{1/2} \exp \left( \frac{1}{2} \frac{\gamma \lambda_0^2 T}{1 - \gamma - \gamma v_0 T} \right). \quad (60)$$

Note that for a well-posed problem we require the risk aversion coefficient  $\gamma$  to satisfy  $1 - \gamma - \gamma v_0 T > 0$  or

$$\gamma < \frac{1}{1 + v_0 T}.$$

For the optimal wealth process, straightforward computations give the formula

$$X_t^{0,*} = x \left( \frac{\Psi_t^0}{\Psi_0^0} \right)^{\frac{1}{2}} \exp \left[ \frac{1}{2} (1 - q) (\Lambda_t^0 - \Lambda_0^0) \right], \quad (61)$$

where, for  $t \in [0, T]$ ,

$$\Psi_t^0 := \frac{v_t}{1 + qv_t(T - t)}, \quad \Lambda_t^0 := \frac{\hat{\lambda}_t^2}{v_t(1 + qv_t(T - t))}. \quad (62)$$

To compute the optimal trading strategy  $\theta^{0,*}$ , we apply the Itô formula to (61) and compare the coefficient of  $d\hat{B}_t$  with that in (46) for the case of the optimal wealth process. This gives

$$\theta_t^{0,*} = \frac{\hat{\lambda}_t}{\sigma(1 - \gamma)} \left( \frac{1}{1 + qv_t(T - t)} \right). \quad (63)$$

The classical Merton formula is thus altered in two ways: the constant  $\lambda$  is replaced by its filtered estimate  $\hat{\lambda}_t$ , and the risky asset proportion is decreased by the factor  $(1 + v_t(T - t))^{-1}$ .

As noted by Rogers [20], the more risk averse the investor, the less likely he is to invest in shares, and as  $t \rightarrow T$ , the optimal strategy gets closer and closer to the Merton rule.

## 4 Optimal hedging of basis risk with partial information

Now we analyse the hedging of a contingent claim in an example of an incomplete market, a basis risk model, first under a full information assumption, and then under a partial information scenario. See [14, 15, 16] for more details on these models.

### 4.1 Basis risk model: full information case

In a full information model, the setting is a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ , where the filtration  $\mathbb{F}$  is the  $P$ -augmentation of that generated by a two-dimensional Brownian motion  $(B, B^\perp)$ . A traded stock price  $S := (S_t)_{0 \leq t \leq T}$  follows a log-Brownian process given by

$$dS_t = \sigma S_t (\lambda dt + dB_t) =: \sigma S_t d\xi_t, \quad (64)$$

where  $\sigma > 0$  and  $\lambda$  are known constants. For simplicity, the interest rate is taken to be 0. The process  $\xi$  in (64) defined by  $d\xi_t := \lambda dt + dB_t$  will subsequently play a role as one component of an observation process in a partial information model, when  $\lambda$  will be treated as a random variable rather than as a known constant.

A non-traded asset price  $Y := (Y_t)_{0 \leq t \leq T}$  follows the correlated log-Brownian motion

$$dY_t = \beta Y_t (\theta dt + dW_t) =: \beta Y_t d\zeta_t, \quad (65)$$

with  $\beta > 0$  and  $\theta$  known constants. The Brownian motion  $W$  is correlated with  $B$  according to

$$d[B, W]_t = \rho dt, \quad W = \rho B + \sqrt{1 - \rho^2} B^\perp, \quad \rho \in [-1, 1],$$

and the process  $\zeta$ , given by  $d\zeta_t := \theta dt + dW_t$ , will act as the second component of an observation process in a partial information model, when  $\theta$  will be considered a random variable. We shall

henceforth refer to the Sharpe ratios  $\lambda$  (respectively,  $\theta$ ) as the drift of  $S$  (respectively,  $Y$ ), for brevity.

A European contingent claim pays the non-negative random variable  $h(Y_T)$  at time  $T$ . In what follows we shall consider utility maximisation problems with the additional random terminal endowment  $nh(Y_T)$ , for  $n \in \mathbb{R}$ , and we make the following assumption on the random endowment.

**Assumption 1** The random endowment  $nh(Y_T)$  is continuous and bounded below, with finite expectation under any martingale measure.

An agent may trade the stock in a self-financing fashion, leading to the portfolio wealth process  $X = (X_t)_{0 \leq t \leq T}$  satisfying

$$dX_t = \sigma \pi_t (\lambda dt + dB_t),$$

where  $\pi := (\pi_t)_{0 \leq t \leq T}$  is the wealth in the stock, representing the agent's trading strategy, satisfying  $\int_0^T \pi_t^2 dt < \infty$  almost surely.

#### 4.1.1 Perfect correlation case

This market is incomplete for  $|\rho| \neq 1$ . If the correlation is perfect, however, the market becomes complete and perfect hedging is possible as shown below (see Monoyios [15] for more details),

The minimal martingale measure  $Q^M$  has density process with respect to  $P$  given by

$$\left. \frac{dQ^M}{dP} \right|_{\mathcal{F}_t} = \mathcal{E}(-\lambda \cdot B)_t, \quad 0 \leq t \leq T.$$

Under  $Q^M$ ,  $(S, Y)$  follow

$$\begin{aligned} dS_t &= \sigma S_t dB_t^{Q^M}, \\ dY_t &= \beta(\theta - \rho\lambda) Y_t dt + \beta Y_t dW_t^{Q^M}, \end{aligned}$$

where  $B^{Q^M}, W^{Q^M}$  are correlated Brownian motions under  $Q^M$ . The stock price  $S$  is a local  $Q^M$ -martingale, but this is not the case for the non-traded asset, unless we have the perfect correlation case,  $\rho = 1$ . In this case  $Y$  is effectively a traded asset (as  $Y_t$  is then a function of  $S_t$ ), so the  $Q^M$ -drift of  $Y$  vanishes. Therefore, given  $\sigma, \beta$ , in the  $\rho = 1$  case the drifts are related by

$$\theta = \lambda.$$

In this case the market becomes complete, and perfect hedging is possible. It is easy to show that with  $\rho = 1$ , so that  $W = B$ , we have

$$\begin{aligned} Y_t &= Y_0 \left( \frac{S_t}{S_0} \right)^{\beta/\sigma} e^{ct} \\ c &= \frac{1}{2} \sigma \beta \left( 1 - \frac{\beta}{\sigma} \right) \end{aligned}$$

Let the claim price process be  $v(t, Y_t)$ ,  $0 \leq t \leq T$ , where  $v : [0, T] \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$  is smooth enough to apply the Itô formula, so that

$$dv(t, Y_t) = [v_t(t, Y_t) + \mathcal{A}^Y v(t, Y_t)] dt + \beta Y_t v_y(t, Y_t) dW_t,$$

where  $\mathcal{A}^Y$  is the generator of the process  $Y$  in (65). The replication conditions are

$$X_t = v(t, Y_t), \quad 0 \leq t \leq T, \quad dX_t = dv(t, Y_t).$$

Standard arguments then show that to perfectly hedge the claim one must hold  $\Delta_t$  shares of  $S$  at  $t \in [0, T]$ , given by

$$\Delta_t = \frac{\beta Y_t}{\sigma S_t} \frac{\partial v}{\partial y}(t, Y_t), \tag{66}$$

and the claim pricing function  $v(t, y)$  satisfies

$$\begin{aligned} v_t(t, y) + \beta(\theta - \lambda)yv_y(t, y) + \frac{1}{2}\beta^2y^2v_{yy}(t, y) &= 0, \\ v(T, y) &= h(y). \end{aligned}$$

But with  $\rho = 1$ ,  $\theta = \lambda$ , so we get the BS partial differential equation (PDE), and

$$v(t, Y_t) = \text{BS}(t, Y_t),$$

where  $\text{BS}(t, y)$  denotes the BS option pricing formula at time  $t$ , with underlying asset price  $y$ .

Therefore, a position in  $n$  claims is hedged by  $\Delta_t^{(\text{BS})}$  units of  $S$  at  $t \in [0, T]$ , where

$$\Delta_t^{(\text{BS})} = -n \frac{\beta Y_t}{\sigma S_t} \frac{\partial}{\partial y} \text{BS}(t, Y_t; \beta), \quad (67)$$

and where  $\text{BS}(t, y; \beta)$  denotes the BS formula at time  $t$  for underlying asset price  $y$  and volatility  $\beta$ . From our perspective, the salient feature of (67) is that the perfect hedge does not require knowledge of the values of the drifts  $\lambda, \theta$ .

#### 4.1.2 Incomplete case

Now suppose the correlation is not perfect, so that the market is incomplete. We embed the problem in a utility maximisation framework in a manner that is by now classical. Let the agent have risk preferences expressed via the exponential utility function

$$U(x) = -\exp(-\alpha x), \quad x \in \mathbb{R}, \quad \alpha > 0.$$

The agent maximises expected utility of terminal wealth at time  $T$ , with a random endowment of  $n$  units of claim payoff:

$$J(t, x, y; \pi) = E[U(X_T + nh(Y_T)) | X_t = x, Y_t = y].$$

The value function is  $u^{(n)}(t, x, y) \equiv u(t, x, y)$ , defined by

$$u(t, x, y) := \sup_{\pi \in A} J(t, x, y; \pi), \quad (68)$$

$$u(T, x, y) = U(x + nh(y)). \quad (69)$$

Denote the optimal trading strategy that achieves the supremum in (68) by  $\pi^* \equiv \pi^{*,n}$ , and denote the optimal wealth process by  $X^* \equiv X^{*,n}$ .

The following definitions of utility-based price and hedging strategy are now standard (see [14, 17] for instance).

**Definition 1 (Indifference price)** *The indifference price per claim at  $t \in [0, T]$ , given  $X_t = x, Y_t = y$ ,  $p(t, x, y) \equiv p^{(n)}(t, x, y)$ , is defined by*

$$u^{(n)}(t, x - np^{(n)}(t, x, y), y) = u^{(0)}(t, x, y)$$

We allow for possible dependence on  $t, x, y$  of  $p^{(n)}$  in the above definition, but with exponential preferences it turns out that there is no dependence on  $x$ .

**Definition 2 (Optimal hedging strategy)** *The optimal hedging strategy for  $n$  units of the claim is  $\pi^H := (\pi_t^H)_{0 \leq t \leq T}$  given by*

$$\pi_t^H := \pi_t^{*,n} - \pi_t^{*,0}, \quad 0 \leq t \leq T.$$

The solution to the optimisation problem (68) is well-known, using the so-called distortion technique [23]. See [14] for more details.

The HJB equation for the value function  $u$  is

$$u_t + \mathcal{A}^Y u - \frac{(\lambda u_x + \rho \beta y u_{xy})^2}{2u_{xx}} = 0.$$

The optimal trading strategy  $\pi^*$  is given by  $\pi_t^* = \Pi^*(t, X_t^*, Y_t)$ , where the function  $\Pi^* : [0, T] \times \mathbb{R} \times \mathbb{R}^+$  is given by

$$\Pi^*(t, x, y) := - \left( \frac{\lambda u_x + \rho \beta y u_{xy}}{\sigma u_{xx}} \right). \quad (70)$$

We have the following representation for the value function and indifference price.

**Proposition 2** [6, 14, 17] *The value function  $u \equiv u^{(n)}$  and indifference price  $p \equiv p^{(n)}$ , given  $X_t = x, Y_t = y$  for  $t \in [0, T]$ , are given by*

$$\begin{aligned} u^{(n)}(t, x, y) &= -e^{-\alpha x - \frac{1}{2}\lambda^2(T-t)} [F(t, Y)]^{1/(1-\rho^2)}, \\ F(t, y) &= E^{Q^M} [\exp(-\alpha(1-\rho^2)nh(Y_T)) | Y_t = y], \\ p^{(n)}(t, y) &= -\frac{1}{\alpha(1-\rho^2)n} \log F(t, y). \end{aligned} \quad (71)$$

The function  $F(t, y)$  satisfies a linear PDE by virtue of the stochastic representation (71) and the Feynman-Kac theorem. The indifference pricing function  $p(t, y) \equiv p^{(n)}(t, y)$  then satisfies

$$p_t + \beta(\theta - \rho\lambda)yp_y + \frac{1}{2}\beta^2y^2p_{yy} - \frac{1}{2}\beta^2y^2\alpha n(1-\rho^2)(p_y)^2 = 0.$$

Given the above results, it is easy to show that the expression (70) for the optimal control loses its dependence on  $x$  and simplifies to

$$\Pi^*(t, y) := \frac{1}{\alpha\sigma b(t, T)} \left( \lambda + \frac{\rho\beta y}{1-\rho^2} \frac{F_y}{F} \right).$$

Then, applying Definition 2 gives the optimal hedging strategy for a position in  $n$  claims (see [14] for further details of this derivation).

**Proposition 3** *The optimal hedging strategy for a position in  $n$  claims is to hold  $\Delta_t^I$  shares at  $t \in [0, T]$ , given by*

$$\begin{aligned} \Delta_t^I &= -n\rho \frac{\beta Y_t}{\sigma S_t} \frac{\partial p^{(n)}}{\partial y}(t, Y_t), \\ p(t, y) &= -\frac{1}{\alpha n(1-\rho^2)} \log E^{Q^M} [\exp(-\alpha(1-\rho^2)nh(Y_T)) | Y_t = y], \end{aligned} \quad (72)$$

We note that if  $n = 1$  and  $\rho = 1$ , we recover the perfect delta hedge (66), and that the claim price then satisfies the BS PDE. The measure  $Q^M$  is the minimal martingale measure for the model, under which  $(S_t)_{0 \leq t \leq T}$  is a local martingale and under which  $Y$  follows

$$dY_t = \beta Y_t \left[ (\theta - \rho\lambda)dt + dW_t^{Q^M} \right], \quad (73)$$

for a  $Q^M$ -Brownian motion  $W^{Q^M}$ .

In [14, 15] the hedging strategy in (72) is shown to be superior to the BS-style hedge (67), in terms of the terminal hedging error distribution produced by selling the claim at the appropriate price (the indifference price or the BS price) and investing the proceeds in the corresponding hedging portfolio. But from (73) we see that the exponential hedge requires knowledge of  $\lambda, \theta$ , which are impossible to estimate accurately (see Rogers [20] or Monoyios [15]). This can ruin the effectiveness of indifference hedging, as shown in [15]. It is therefore dubious to draw any meaningful conclusions on the effectiveness of utility-based hedging in this model without relaxing the assumption that the agent knows the true values of the drifts.

## 4.2 Partial information case

Now we assume the hedger does not know the values of the return parameters  $\lambda, \theta$ , so these are considered to be random variables. Equivalently, the agent cannot observe the Brownian motions  $B, W$  driving the asset prices, so is required to use strategies adapted to the observation filtration  $\hat{\mathbb{F}}$  generated by asset returns.

### 4.2.1 Choice of prior

We take the the two-dimensional random variable

$$U := \begin{pmatrix} \lambda \\ \theta \end{pmatrix}$$

to have a Gaussian distribution which will be updated as the agent attempts to filter the values of the drifts from asset observations during the hedging interval  $[0, T]$ .

The choice of Gaussian prior is motivated by the idea that the agent has some past observations of  $S, Y$  before time 0, uses these to obtain classical point estimates of the drifts, and the joint distribution of the estimators is used as the prior in a Bayesian framework. Ultimately, in order to obtain explicit solutions, we shall assume that the agent uses observations before time 0 of equal length for both assets. In setting the prior this way, we make the approximation that the asset price observations are continuous, so that  $\sigma, \beta, \rho$  are known from the quadratic variation and co-variation of  $S, Y$ . This is because our goal here is to focus on the severest problem of drift parameter uncertainty.

So, consider, for the moment, an observer with data for  $S$  over a time interval of length  $t_S$ , and for  $Y$  over a window of length  $t_Y$ , who considers  $\lambda$  and  $\theta$  as *constants*, and records the returns  $dS_t/S_t$  and  $dY_t/Y_t$  in order to estimate the values of the drifts. The best estimator of  $\lambda$  is  $\bar{\lambda}(t_S)$  given by

$$\begin{aligned} \bar{\lambda}(t_S) &= \frac{1}{t_S} \int_{t_0}^{t_0+t_S} \frac{dS_u}{\sigma S_u} \\ &= \lambda + \frac{B_{t_0+t_S}}{t_S} \\ &\sim \text{N}\left(\lambda, \frac{1}{t_S}\right), \end{aligned}$$

where  $\text{N}(\mu, \Sigma)$  denotes the normal probability law with mean  $\mu$  and variance  $\Sigma$ . The estimator of  $\lambda$  is normally distributed, with a similar computation for the estimator of  $\theta$ . The estimator,  $(\bar{\lambda}, \bar{\theta})$ , of the (supposed constant) vector  $(\lambda, \theta)$  is bivariate normal. Defining  $v_0 := 1/t_S$  and  $w_0 := 1/t_Y$  it is easily checked that

$$\begin{pmatrix} \bar{\lambda} \\ \bar{\theta} \end{pmatrix} \sim \text{N}(M, C_0),$$

where the mean vector  $M$  and covariance matrix  $C_0$  are given by

$$M = \begin{pmatrix} \lambda \\ \theta \end{pmatrix}, \quad C_0 = \begin{pmatrix} v_0 & \rho \min(v_0, w_0) \\ \rho \min(v_0, w_0) & w_0 \end{pmatrix}. \quad (74)$$

With this in mind, we shall suppose that  $(\lambda, \theta)$ , now considered as a *random variable*, is bivariate normal according to

$$\lambda \sim \text{N}(\lambda_0, v_0), \quad \theta \sim \text{N}(\theta_0, w_0), \quad \text{cov}(\lambda, \theta) = c_0 := \rho \min(v_0, w_0),$$

for some chosen values  $\lambda_0, \theta_0$ , typically obtained from past data prior to time 0. This distribution will be updated via subsequent observations of

$$\xi_t := \frac{1}{\sigma} \int_0^t \frac{dS_u}{S_u} = \lambda t + B_t, \quad \zeta_t := \frac{1}{\beta} \int_0^t \frac{dY_u}{Y_u} = \theta t + W_t,$$

over the hedging interval  $[0, T]$ .

### 4.2.2 Two-dimensional Kalman-Bucy filter

We are firmly within the realm of a two-dimensional Kalman filtering problem, which we treat as follows. Define the observation filtration by

$$\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}, \quad \hat{\mathcal{F}}_t = \sigma(\xi_s, \zeta_s; 0 \leq s \leq t).$$

The *observation process*,  $O$ , and unobservable *signal process*,  $U$ , are defined by

$$O := \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}_{0 \leq t \leq T}, \quad U := \begin{pmatrix} \lambda \\ \theta \end{pmatrix},$$

satisfying the stochastic differential equations

$$dO_t = U dt + D d\mathbf{B}_t, \quad dU = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

where

$$D = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} B_t \\ B_t^\perp \end{pmatrix}.$$

The optimal filter is  $\hat{U}_t := E[U|\hat{\mathcal{F}}_t]$ ,  $0 \leq t \leq T$ , a two-dimensional process defining the best estimates of  $\lambda$  and  $\theta$  given observations up to time  $t \in [0, T]$ :

$$\hat{U}_t \equiv \begin{pmatrix} \hat{\lambda}_t \\ \hat{\theta}_t \end{pmatrix} := \begin{pmatrix} E[\lambda|\hat{\mathcal{F}}_t] \\ E[\theta|\hat{\mathcal{F}}_t] \end{pmatrix}, \quad \begin{pmatrix} \hat{\lambda}_0 \\ \hat{\theta}_0 \end{pmatrix} = \begin{pmatrix} \lambda_0 \\ \theta_0 \end{pmatrix}. \quad (75)$$

The solution to this filtering problem converts the partial information model to a full information model with random drifts, given in the following proposition. To avoid a proliferation of symbols, we abuse notation and write  $\hat{\lambda}_t \equiv \hat{\lambda}(t, S_t)$  and  $\hat{\theta}_t \equiv \hat{\theta}(t, Y_t)$  for processes  $\hat{\lambda}, \hat{\theta}$  that will turn out to be functions of time and current asset price.

**Proposition 4** *The partial information model is equivalent to a full information model in which the asset price dynamics in the observation filtration  $\hat{\mathbb{F}}$  are*

$$dS_t = \sigma S_t (\hat{\lambda}_t dt + d\hat{B}_t), \quad (76)$$

$$dY_t = \beta Y_t (\hat{\theta}_t dt + d\hat{W}_t), \quad (77)$$

where  $\hat{B}, \hat{W}$  are  $\hat{\mathbb{F}}$ -Brownian motions with correlation  $\rho$ , and the random drifts  $\hat{\lambda}, \hat{\theta}$  are  $\hat{\mathbb{F}}$ -adapted processes.

If  $\lambda$  and  $\theta$  have common initial variance  $v_0$ , then  $\hat{\lambda}, \hat{\theta}$  are given by

$$\begin{pmatrix} \hat{\lambda}_t \\ \hat{\theta}_t \end{pmatrix} = \begin{pmatrix} \lambda_0 \\ \theta_0 \end{pmatrix} + \int_0^t v_u \begin{pmatrix} d\hat{B}_u \\ d\hat{W}_u \end{pmatrix}, \quad 0 \leq t \leq T, \quad (78)$$

where  $(v_t)_{0 \leq t \leq T}$  is the deterministic function

$$v_t := \frac{v_0}{1 + v_0 t}.$$

Equivalently,  $\hat{\lambda}, \hat{\theta}$  are given as functions of time and current asset price by

$$\hat{\lambda}_t = \hat{\lambda}(t, S_t) = \frac{\lambda_0 + v_0 \xi_t}{1 + v_0 t}, \quad \hat{\theta}_t = \hat{\theta}(t, Y_t) = \frac{\theta_0 + v_0 \zeta_t}{1 + v_0 t}, \quad (79)$$

with

$$\xi_t = \frac{1}{\sigma} \log \left( \frac{S_t}{S_0} \right) + \frac{1}{2} \sigma t, \quad \zeta_t = \frac{1}{\beta} \log \left( \frac{Y_t}{Y_0} \right) + \frac{1}{2} \beta t. \quad (80)$$

**Proof** By the Kalman-Bucy filter, Theorem 5,  $\hat{U}$  satisfies the stochastic differential equation

$$d\hat{U}_t = C_t (DD^T)^{-1} (dO_t - \hat{U}_t dt) =: C_t (DD^T)^{-1} dN_t, \quad (81)$$

where  $(N_t)_{0 \leq t \leq T}$  is the innovations process, defined by

$$\begin{aligned} N_t &:= O_t - \int_0^t \hat{U}_s ds \\ &= \begin{pmatrix} \xi_t - \int_0^t \hat{\lambda}_s ds \\ \zeta_t - \int_0^t \hat{\theta}_s ds \end{pmatrix} \\ &=: \begin{pmatrix} \hat{B}_t \\ \hat{W}_t \end{pmatrix}, \end{aligned} \quad (82)$$

and  $\hat{B}, \hat{W}$  are  $\hat{\mathbb{F}}$ -Brownian motions with correlation  $\rho$ . The deterministic matrix function  $C_t$  is the conditional variance-covariance matrix defined by

$$C_t := E \left[ (U - \hat{U}_t)(U - \hat{U}_t)^T \middle| \hat{\mathcal{F}}_t \right] = E \left[ (U - \hat{U}_t)(U - \hat{U}_t)^T \right],$$

( $T$  denoting transpose) where the last equality follows because the error  $U - \hat{U}_t$  is independent of  $\hat{\mathcal{F}}_t$ .

Using (82), and writing  $dS_t$  in terms of  $d\xi_t$ , as in (64), gives the dynamics (76) of  $S$  in the observation filtration; (77) is established similarly.

The matrix  $C = (C_t)_{0 \leq t \leq T}$  satisfies the Riccati equation

$$\frac{dC_t}{dt} = -C_t (DD^T)^{-1} C_t,$$

with  $C_0$  given in (74). Then  $R_t := C_t^{-1}$  satisfies the Lyapunov equation

$$\frac{dR_t}{dt} = (DD^T)^{-1}.$$

Define the elements of the conditional covariance matrix by

$$C_t =: \begin{pmatrix} v_t & c_t \\ c_t & w_t \end{pmatrix}.$$

Then the filtering equation (81) is a pair of coupled stochastic differential equations:

$$\begin{aligned} \begin{pmatrix} d\hat{\lambda}_t \\ d\hat{\theta}_t \end{pmatrix} &= \frac{1}{1 - \rho^2} \begin{pmatrix} v_t - \rho c_t & c_t - \rho w_t \\ c_t - \rho w_t & w_t - \rho c_t \end{pmatrix} \begin{pmatrix} d\xi_t - \hat{\lambda}_t dt \\ d\zeta_t - \hat{\theta}_t dt \end{pmatrix} \\ &= \frac{1}{1 - \rho^2} \begin{pmatrix} v_t - \rho c_t & c_t - \rho w_t \\ c_t - \rho w_t & w_t - \rho c_t \end{pmatrix} \begin{pmatrix} d\hat{B}_t \\ d\hat{W}_t \end{pmatrix}. \end{aligned}$$

Solving the Lyapunov equation yields 3 equations for  $v_t, w_t, c_t$ :

$$\begin{aligned} \frac{v_t}{v_t w_t - c_t^2} - \frac{v_0}{v_0 w_0 - c_0^2} &= \frac{t}{1 - \rho^2}, \\ \frac{w_t}{v_t w_t - c_t^2} - \frac{w_0}{v_0 w_0 - c_0^2} &= \frac{t}{1 - \rho^2}, \\ \frac{c_t}{v_t w_t - c_t^2} - \frac{c_0}{v_0 w_0 - c_0^2} &= \frac{\rho t}{1 - \rho^2}, \end{aligned} \quad (83)$$

where we have written  $c_0 \equiv \rho \min(v_0, w_0)$  for brevity.

Now make the simplification  $w_0 = v_0$ . From the discussion in Section 4.2.1, we see that this corresponds to using past observations over the same length of time,  $t_S = t_Y$ , for both  $S$  and  $Y$



in fixing the prior. Then  $c_0 = \rho v_0$ , and the solution to the system of equations (83) gives the entries of the matrix  $C_t$  as

$$v_t = \frac{v_0}{1 + v_0 t}, \quad w_t = v_t, \quad c_t = \rho v_t.$$

With this simplification, the equation for the optimal filter simplifies to

$$\begin{pmatrix} d\hat{\lambda}_t \\ d\hat{\theta}_t \end{pmatrix} = v_t \begin{pmatrix} d\xi_t - \hat{\lambda}_t dt \\ d\zeta_t - \hat{\theta}_t dt \end{pmatrix} = v_t \begin{pmatrix} d\hat{B}_t \\ d\hat{W}_t \end{pmatrix}, \quad (84)$$

which, along with the initial condition in (75), yields (78) and (79).

Finally, the expressions in (80) for  $\xi_t, \zeta_t$  follow directly from the solutions of (64) and (65) for  $S$  and  $Y$ .

□

Armed with Proposition 4 we may now treat the model as a full information model with random drift parameters  $(\hat{\lambda}_t, \hat{\theta}_t)$ , and this is done in the next section. We make the following observations:

- The formulae  $\hat{\lambda}(t, S_t), \hat{\theta}(t, Y_t)$  for the random drifts, in terms of current asset price, allow the model to be expressed in Markovian form with only one extra state variable (the stock price  $S$ ) in the stochastic control problems for the optimal price and hedge compared with the full information case with constant drifts, as we shall see.
- The simplification  $w_0 = v_0$  is made to allow for a simple analytic solution to the marginal optimal hedging problem. One can proceed without this assumption, but the formulae become more complicated.

#### 4.2.3 Optimal hedging with random drifts

On the stochastic basis  $(\Omega, \hat{\mathcal{F}}, \hat{\mathbb{F}}, P)$ , the wealth process associated with trading strategy  $\pi := (\pi_t)_{0 \leq t \leq T}$ , an  $\hat{\mathbb{F}}$ -adapted process satisfying the integrability condition  $\int_0^T \pi_t^2 dt < \infty$  a.s., is  $X = (X_t)_{0 \leq t \leq T}$ , satisfying

$$dX_t = \sigma \pi_t (\hat{\lambda}_t dt + d\hat{B}_t). \quad (85)$$

The class  $\mathcal{M}$  of local martingale measures for this model consists of measures  $Q$  with density processes defined by

$$Z_t := \frac{dQ}{dP} \Big|_{\hat{\mathcal{F}}_t} = \mathcal{E}(-\hat{\lambda} \cdot \hat{B} - \psi \cdot \hat{B}^\perp)_t, \quad 0 \leq t \leq T, \quad (86)$$

for integrands  $\psi$  satisfying  $\int_0^t \psi_u^2 du < \infty$  a.s., for all  $t \in [0, T]$  (it is not hard to show that  $\int_0^t \hat{\lambda}_u^2 dt < \infty, 0 \leq t \leq T$ ). For  $\psi = 0$  we obtain the minimal martingale measure  $Q^M$ .

Under  $Q \in \mathcal{M}$ ,  $(\hat{B}^Q, \hat{B}^{\perp, Q})$  is two-dimensional Brownian motion, where

$$d\hat{B}_t^Q := d\hat{B}_t^Q + \hat{\lambda}_t dt, \quad d\hat{B}_t^{\perp, Q} := d\hat{B}_t^\perp + \psi_t dt,$$

and the asset prices and random drifts satisfy

$$\begin{aligned} dS_t &= \sigma S_t d\hat{B}_t^Q, \\ dY_t &= \beta Y_t [(\hat{\theta}_t - \rho \hat{\lambda}_t - \sqrt{1 - \rho^2} \psi_t) dt + d\hat{W}_t^Q], \\ d\hat{\lambda}_t &= v_t [-\hat{\lambda}_t dt + d\hat{B}_t^Q], \\ d\hat{\theta}_t &= v_t [-(\rho \hat{\lambda}_t + \sqrt{1 - \rho^2} \psi_t) dt + d\hat{W}_t^Q], \end{aligned}$$

where  $\hat{W}^Q = \rho \hat{B}^Q + \sqrt{1 - \rho^2} \hat{B}^{\perp, Q}$ .

The relative entropy between  $Q \in \mathcal{M}$  and  $P$  is defined by

$$\begin{aligned} \mathcal{H}(Q, P) &:= E \left[ \frac{dQ}{dP} \log \frac{dQ}{dP} \right] \\ &= E^Q \left[ - \int_0^T \hat{\lambda}_t d\hat{B}_t^Q - \int_0^T \psi_t d\hat{B}_t^{\perp, Q} + \frac{1}{2} \int_0^T (\hat{\lambda}_t^2 + \psi_t^2) dt \right]. \end{aligned}$$

Using the  $Q$ -dynamics of  $\hat{\lambda}_t$  it is straightforward to establish that  $E^Q \int_0^t \hat{\lambda}_u^2 du < \infty$  for all  $t \in [0, T]$ . If, in addition, we have the integrability condition

$$E^Q \int_0^t \psi_u^2 du < \infty, \quad 0 \leq t \leq T, \quad (87)$$

then

$$\mathcal{H}(Q, P) = E^Q \left[ \frac{1}{2} \int_0^T (\hat{\lambda}_t^2 + \psi_t^2) dt \right] < \infty. \quad (88)$$

In this case we write  $Q \in \mathcal{M}_f$ , where  $\mathcal{M}_f$  denotes the set of martingale measures  $Q$  with finite relative entropy with respect to  $P$ , and we define  $\mathcal{H}(Q, P) := \infty$  otherwise. From (88) we note that the minimal entropy measure  $Q^E$  is given by

$$\mathcal{H}(Q^E, P) = E^Q \left[ \frac{1}{2} \int_0^T \hat{\lambda}_t^2 dt \right],$$

corresponding to  $\psi \equiv 0$  in (88). This means that the minimal martingale measure and the minimal entropy measure in this model coincide:  $Q^E = Q^M$ .

For an initial time  $t \in [0, T]$ , we define the conditional entropy between  $Q \in \mathcal{M}$  and  $P$  by

$$H_t(Q, P) := E \left[ \frac{Z_T}{Z_t} \log \left( \frac{Z_T}{Z_t} \right) \middle| \hat{\mathcal{F}}_t \right], \quad 0 \leq t \leq T, \quad (89)$$

satisfying  $H_0(Q, P) \equiv \mathcal{H}(Q, P)$ . Provided the integrability condition (87) is satisfied, then

$$H_t(Q, P) = E^Q \left[ \frac{1}{2} \int_t^T (\hat{\lambda}_u^2 + \psi_u^2) du \middle| \hat{\mathcal{F}}_t \right],$$

and we define  $H_t(Q, P) := \infty$  otherwise. In particular, therefore, recalling that  $\hat{\lambda}_t \equiv \hat{\lambda}(t, S_t)$  is a smooth and Lipschitz function of time and current stock price, and that the  $Q$ -dynamics of  $\hat{\lambda}_t$  do not depend on  $\psi_t$  for any  $Q \in \mathcal{M}$ , the minimal conditional entropy  $(H_t(Q^E, P))_{0 \leq t \leq T}$  will be a deterministic function of time and stock price, given by  $H_t(Q^E, P) \equiv H^E(t, S_t)$  for a  $C^{1,2}([0, T] \times \mathbb{R}^+)$  function  $H^E$  defined by

$$H^E(t, s) := E^{Q^E} \left[ \frac{1}{2} \int_t^T \hat{\lambda}^2(u, S_u) du \middle| S_t = s \right]. \quad (90)$$

#### 4.2.4 The primal problem

We use an exponential utility function,  $U(x) = -\exp(-\alpha x)$ ,  $x \in \mathbb{R}$ ,  $\alpha > 0$ . The primal value function  $u \equiv u^{(n)}$  is defined as the maximum expected utility of wealth at  $T$  from trading  $S$  and receiving  $n$  units of the claim on  $Y$ , when starting at time  $t \in [0, T]$ :

$$u^{(n)}(t, x, s, y) := \sup_{\pi \in \mathcal{A}} E[U(X_T + nh(Y_T)) | X_t = x, S_t = s, Y_t = y], \quad (91)$$

where  $\mathcal{A}$  denotes the set of admissible trading strategies. The dynamics of the state variables  $X, S, Y$  are given by (85) and (76,77). For starting time 0 we write  $u^{(n)}(x) \equiv u^{(n)}(0, x, \cdot, \cdot)$ .

The set of admissible strategies is defined as follows. Denote by  $\Delta := \pi/S$  be the adapted  $S$ -integrable process for the number of shares held. We follow Becherer [2] and take the space of permitted strategies as

$$\mathcal{A} = \{\Delta : (\Delta \cdot S) \text{ is a } (Q, \hat{\mathbb{F}})\text{-martingale for all } Q \in \mathcal{M}_f\},$$

where  $(\Delta \cdot S)_t = \int_0^t \Delta_u dS_u$  is the gain from trading over  $[0, t], t \in [0, T]$ . Other choices for  $\mathcal{A}$  are possible, see for example Schachermayer [22] or Delbaen et al [5]. However, these all lead to the same solution for the dual problem (see [5, 8, 22]), and hence to the same primal solution for the utility maximisation problem.

Note that we could also write the value function in (91) as a function of wealth, the non-traded asset price, and the current values of the random drifts:  $u(t, x, y; \hat{\lambda}_t, \hat{\theta}_t)$ , with the dynamics of the drifts given by (84). We do not pursue this as the dimension of the resulting HJB equation is initially higher than in our formulation, and could be reduced by changing variables using the formulae for the random drifts in (79).

Denote the optimal trading strategy by  $\pi^* \equiv \pi^{*,n}$ , and the optimal wealth process by  $X^* \equiv X^{*,n}$ . The utility-based price and hedge for a position in  $n$  claims are defined in the now classical manner. The indifference price per claim at  $t \in [0, T]$ , given  $X_t = x, S_t = s, Y_t = y$ , is  $p^{(n)}$  given by

$$u^{(n)}(t, x - np^{(n)}(t, x, s, y), s, y) = u^{(0)}(t, x, s, y).$$

The optimal hedging strategy is to hold  $(\Delta_t^H)_{0 \leq t \leq T}$  shares of stock at time  $t$ , where  $\Delta_t^H S_t =: \pi_t^H S_t$ , and  $\pi^H := (\pi_t^H)_{0 \leq t \leq T}$ , is defined by

$$\pi_t^H := \pi_t^{*,n} - \pi_t^{*,0}, \quad 0 \leq t \leq T. \quad (92)$$

It is well known that with exponential utility the indifference price is independent of the initial cash wealth  $x$ , so we shall write  $p^{(n)}(t, x, s, y) \equiv p^{(n)}(t, s, y)$  from now on.

For small positions in the claim (or, equivalently, for small risk aversion), we shall later approximate the indifference price by the marginal utility-based price introduced by Davis [4]. This is the indifference price for infinitesimal diversions of funds into the purchase or sale of claims, and is equivalent (as is well-known) to the limit of the indifference price as  $n \rightarrow 0$ .

**Definition 3 (Marginal price)** The marginal utility-based price of the claim at  $t \in [0, T]$  is  $\hat{p}(t, s, y)$  defined by

$$\hat{p}(t, s, y) := \lim_{n \rightarrow 0} p^{(n)}(t, s, y).$$

It is well known that with exponential utility the marginal price is also equivalent to the limit of the indifference price as risk aversion goes to zero. Under appropriate conditions (satisfied in this model) it is given by the expectation of the payoff under the optimal measure of the dual problem without the claim. For exponential utility this measure is the minimal entropy measure  $Q^E$  and, as we have already seen, in our model  $Q^E = Q^M$ , giving the representation  $\hat{p}(t, s, y) = E^{Q^M}[h(Y_T)|S_t = s, Y_t = y]$ , as we shall see in the next section.

#### 4.2.5 Dual problem and optimal hedge

We attack the primal utility maximisation problem (91) using well-known duality results that are now a classical tool for incomplete market optimisation problems (see the seminal papers by Kramkov and Schachermayer [12] and Karatzas et al [10]). For a problem with the random terminal endowment of a European claim, and with exponential utility, as in this paper, Delbaen et al [5] establish the required duality relations between the primal and dual problems in a semimartingale setting. We shall use these results below to establish a simple algebraic relation (Lemma 2) between the primal value function and the indifference price, which we shall then exploit to derive the representation for the optimal hedging strategy.

The dual problem with starting time 0 has value function defined by

$$v^{(n)}(\eta) := \inf_{Q \in \mathcal{M}} E[V(\eta Z_T) + \eta Z_T n h(Y_T)],$$

where  $Z$  is the density process in (86) and  $V$  is the convex conjugate of the utility function:

$$V(\eta) := \sup_{x \in \mathbb{R}} [U(x) - x\eta].$$

For exponential utility  $V$  is given by

$$V(\eta) = \frac{\eta}{\alpha} \left[ \log \left( \frac{\eta}{\alpha} \right) - 1 \right].$$

Hence the dual value function has the well-known entropic representation

$$v^{(n)}(\eta) = V(\eta) + \frac{\eta}{\alpha} \inf_{Q \in \mathcal{M}} [\mathcal{H}(Q, P) + \alpha n E^Q h(Y_T)].$$

Denoting the dual minimiser that attains the above infimum by  $Q^{*,n}$ , we observe that  $Q^{*,n} \in \mathcal{M}_f$ .

For a starting time  $t \in [0, T]$  the dual value function is defined by

$$v^{(n)}(t, \eta, s, y) := \inf_{Q \in \mathcal{M}} E \left[ V \left( \eta \frac{Z_T}{Z_t} \right) + \eta \frac{Z_T}{Z_t} n h(Y_T) \middle| S_t = s, Y_t = y \right], \quad (93)$$

and we write  $v^{(n)}(\eta) \equiv v^{(n)}(0, \eta, \cdot, \cdot)$ .

**Lemma 2** *The primal value function and indifference price are related by*

$$u^{(n)}(t, x, s, y) = u^{(0)}(t, x, s) \exp \left( -\alpha n p^{(n)}(t, s, y) \right), \quad (94)$$

where the value function without the claim is given by

$$u^{(0)}(t, x, s) = -\exp \left( -\alpha x - H^E(t, s) \right), \quad (95)$$

and  $H^E(t, s)$  is the conditional minimal entropy function defined in (90).

**Proof** For brevity, we give the proof for  $t = 0$ . The proof for a general starting time follows similar lines, and we make some comments on how to adapt the following argument for that case at the end of the proof.

The fundamental duality linking the primal and dual problems in Delbaen et al [5] implies that the value functions  $u^{(n)}(x)$  and  $v^{(n)}(\eta)$  are conjugate:

$$v^{(n)}(\eta) = \sup_{x \in \mathbb{R}} [u^{(n)}(x) - x\eta], \quad u^{(n)}(x) = \inf_{\eta > 0} [v^{(n)}(\eta) + x\eta].$$

The value of  $\eta$  attaining the above infimum is  $\eta^*$ , given by  $v_{\eta}^{(n)}(\eta^*) = -x$ , so that

$$u^{(n)}(x) = v^{(n)}(\eta^*) + x\eta^*,$$

which translates to

$$u^{(n)}(x) = -\exp \left( -\alpha x - \inf_{Q \in \mathcal{M}} [\mathcal{H}(Q, P) + \alpha n E^Q h(Y_T)] \right). \quad (96)$$

So, in particular,

$$u^{(0)}(x) = -\exp \left[ -\alpha x - \mathcal{H}(Q^E, P) \right], \quad (97)$$

where  $Q^E$  is the minimal entropy measure:  $Q^E = Q^{*,0}$

Combining the dual representations (96) and (97) for the primal problems with and without the claim, with the definition of the indifference price, gives the dual representation for the utility-based price in the form

$$p^{(n)} = \frac{1}{\alpha n} \left[ \inf_{Q \in \mathcal{M}} [\mathcal{H}(Q, P) + \alpha n E^Q h(Y_T)] - \mathcal{H}(Q^E, P) \right], \quad (98)$$

which is the representation found in Delbaen et al [5], modified slightly as we have a random endowment of  $n$  claims ([5] considered the case  $n = -1$ ).

In particular, for  $n \rightarrow 0$  or  $\alpha \rightarrow 0$ , we obtain the marginal price of Davis [4]:

$$\hat{p} := \lim_{n \rightarrow 0} p^{(n)} = E^{Q^E} h(Y_T) = E^{Q^M} h(Y_T), \quad (99)$$

the last inequality following from the equality of  $Q^M$  and  $Q^E$ , as implied by (88).

From (96)–(98), the relation between the primal value functions and indifference price then follows immediately, as

$$\begin{aligned} u^{(n)}(x) &= -\exp\left(-\alpha x - \mathcal{H}(Q^E, P) - \alpha n p^{(n)}\right) \\ &= u^{(0)}(x) \exp\left(-\alpha n p^{(n)}\right). \end{aligned}$$

Similarly, a corresponding relation for a starting time  $t \in [0, T]$  may also be derived. This is achieved using the definition (93) of the dual value function for an initial time  $t \in [0, T]$ , the conjugacy of  $u^{(n)}(t, x, s, y)$  and  $v^{(n)}(t, \eta, s, y)$  and the definitions (89) and (90) of the conditional entropy and conditional minimal entropy.

□

Using Lemma 2 we obtain the following representation for the optimal hedging strategy associated with the indifference price. In what follows we assume that the indifference price is a suitably smooth function of  $(t, s, y)$ , so that (given Lemma 2) we may assume the primal value function is smooth enough to be a classical solution of the associated Hamilton-Jacobi-Bellman (HJB) equation. This smoothness property is confirmed in [16].

**Theorem 7** *The optimal hedge for a position in  $n$  claims is to hold  $\Delta_t^H$  units of  $S$  at  $t \in [0, T]$ , where*

$$\Delta_t^H = -n \left( p_s^{(n)}(t, S_t, Y_t) + \rho \frac{\beta Y_t}{\sigma S_t} p_y^{(n)}(t, S_t, Y_t) \right).$$

**Remark 2** We note the extra term in the hedging formula compared with the corresponding full information result (72). The drift parameter uncertainty results in additional risk, manifested as dependence of the indifference price on the stock price, and hence the derivative with respect to the stock price appears in the theorem.

**Proof** The HJB equation associated with the primal the value function is

$$u_t^{(n)} + \max_{\pi} \mathcal{A}_{X,S,Y} u^{(n)} = 0,$$

where  $\mathcal{A}_{X,S,Y}$  is the generator of  $(X, S, Y)$  under  $P$ . Performing the maximisation over  $\pi$  yields the optimal Markov control as  $\pi_t^{*,n} = \pi^{*,n}(t, X_t^{*,n}, S_t, Y_t)$ , where

$$\pi^{*,n}(t, x, s, y) = - \left( \frac{\hat{\lambda} u_x^{(n)} + \sigma s u_{xs}^{(n)} + \rho \beta y u_{xy}^{(n)}}{\sigma u_{xx}^{(n)}} \right),$$

and where the arguments of the functions on the right-hand-side are omitted for brevity. For the case  $n = 0$  there is no dependence on  $y$  in the value function  $u^{(0)}$ , and we have  $\pi_t^{*,0} = \pi^{*,0}(t, X_t^{*,0}, S_t)$ , where

$$\pi^{*,0}(t, x, s) = - \left( \frac{\hat{\lambda} u_x^{(0)} + \sigma s u_{xs}^{(0)}}{\sigma u_{xx}^{(0)}} \right).$$

Applying the definition (92) of the optimal hedging strategy along with the representations (94) and (95) from Lemma 2 for the value functions, gives the result.

□

#### 4.2.6 Stochastic control representation of the indifference price

The dual representation (98) of  $p^{(n)}$  gives the price of the claim at time 0 as the value function of a control problem:

$$p^{(n)} = \inf_{\psi} E^Q \left[ \frac{1}{2\alpha n} \int_0^T \psi_t^2 dt + h(Y_T) \right],$$

to be minimised over control processes  $(\psi_t)_{0 \leq t \leq T}$ , such that  $Q \in \mathcal{M}_f$ , and with dynamics for  $S, Y$  given by

$$\begin{aligned} dS_t &= \sigma S_t d\hat{B}_t^Q, \\ dY_t &= \beta Y_t [(\hat{\theta}(t, Y_t) - \rho \hat{\lambda}(t, S_t) - \sqrt{1 - \rho^2} \psi_t) dt + d\hat{W}_t^Q]. \end{aligned}$$

For a starting time  $t \in [0, T]$  we have

$$p^{(n)}(t, s, y) = \inf_{\psi} E^Q \left[ \frac{1}{2\alpha n} \int_t^T \psi_u^2 du + h(Y_T) \middle| S_t = s, Y_t = y \right].$$

The HJB dynamic programming PDE associated with  $p^{(n)}(t, s, y)$  is

$$\begin{aligned} p_t^{(n)} + \mathcal{A}_{S,Y}^{Q^M} p^{(n)} + \inf_{\psi} \left[ \frac{1}{2\alpha n} \psi^2 - \beta \sqrt{1 - \rho^2} \psi y p_y^{(n)} \right] &= 0, \\ p(T, s, y) &= h(y), \end{aligned}$$

where  $\mathcal{A}_{S,Y}^{Q^M}$  is generator of  $(S, Y)$  under minimal measure:

$$\mathcal{A}_{S,Y}^{Q^M} f(t, s, y) = \beta (\hat{\theta}(t, y) - \rho \hat{\lambda}(t, s)) y f_y + \frac{1}{2} s^2 f_{ss} + \frac{1}{2} \beta^2 y^2 f_{yy} + \rho \sigma \beta s y f_{sy}.$$

The optimal Markov control is  $\psi_t^{*,n} \equiv \psi^{*,n}(t, S_t, Y_t)$ , where

$$\psi^{*,n}(t, s, y) = \alpha n \sqrt{1 - \rho^2} \beta y p_y^{(n)}(t, s, y),$$

and note that  $\psi^{*,0} = 0$ . Substituting back into the HJB equation, we find that  $p^{(n)}$  is expected to solve the semi-linear PDE

$$\begin{aligned} p_t^{(n)} + \mathcal{A}_{S,Y}^{Q^M} p^{(n)} - \frac{1}{2} \alpha n (1 - \rho^2) \beta^2 y^2 \left( p_y^{(n)} \right)^2 &= 0, \\ p^{(n)}(T, s, y) &= h(y). \end{aligned} \tag{100}$$

We note that for  $n = 0$  this becomes a linear PDE for the marginal price  $\hat{p}$ , so that by the Feynman-Kac Theorem we have

$$\hat{p}(t, s, y) = E_{t,s,y}^{Q^M} h(Y_T), \tag{101}$$

consistent with the general result (99). We shall see that in this case the marginal price is given by a BS-type formula.

#### 4.2.7 Analytic approximation for the indifference price

To obtain analytic results we approximate the indifference price by the marginal price in (101). The marginal price (and hence the associated trading strategy) can be computed in analytic form since, under  $Q^M$ ,  $\log Y_T$  is Gaussian. We have the following result.

**Proposition 5** *Under  $Q^M$ , conditional on  $S_t = s, Y_t = y$ ,  $\log Y_T \sim N(m, \Sigma^2)$ , where  $m \equiv m(t, s, y)$  and  $\Sigma^2 \equiv \Sigma^2(t)$  are given by*

$$\begin{aligned} m(t, s, y) &= \log y + \beta \left( \hat{\theta}(t, y) - \rho \hat{\lambda}(t, s) - \frac{1}{2} \beta \right) (T - t) \\ \Sigma^2(t) &= [1 + (1 - \rho^2) v_t (T - t)] \beta^2 (T - t) \end{aligned}$$

**Proof** This is established by computing the SDEs for  $Y$  and for  $\hat{\theta}_t - \rho\hat{\lambda}_t$  under  $Q^M$ . Indeed, applying the Itô formula to  $\log Y_t$  under  $Q^M$ , we obtain, for  $t < T$ ,

$$\log Y_T = \log Y_t + \beta \int_t^T (\hat{\theta}_u - \rho\hat{\lambda}_u) du - \frac{1}{2}\beta^2(T-t) + \beta \int_t^T d\hat{W}_u^{Q^M}, \quad (102)$$

where  $\hat{W}^{Q^M}$  is a Brownian motion under  $Q^M$ . The dynamics of  $\hat{\theta}_t - \rho\hat{\lambda}_t$  under  $Q^M$  are

$$d(\hat{\theta}_t - \rho\hat{\lambda}_t) = \sqrt{1 - \rho^2} v_t d\hat{B}_t^{\perp, Q^M},$$

where  $\hat{B}^{\perp, Q^M}$  is a  $Q^M$ -Brownian motion perpendicular to that driving the stock, related to  $\hat{W}^{Q^M}$  by  $\hat{W}^{Q^M} = \rho\hat{B}^{Q^M} + \sqrt{1 - \rho^2}\hat{B}^{\perp, Q^M}$ , and where  $\hat{B}^{Q^M}$  is the Brownian motion driving  $S$ . Hence, for  $u > t$ , after changing the order of integration in a double integral, we obtain

$$\int_t^T (\hat{\theta}_u - \rho\hat{\lambda}_u) du = (\hat{\theta}_t - \rho\hat{\lambda}_t)(T-t) + \sqrt{1 - \rho^2} \int_t^T v_u(T-u) d\hat{B}_u^{\perp, Q^M}.$$

This can be inserted into (102) to yield the desired result. □

We are thus able to obtain BS-style formulae for the price and hedge. For a put option of strike  $K$  we easily obtain the following explicit formulae for the marginal price and the associated optimal hedging strategy, where  $\Phi$  denotes the standard cumulative normal distribution function.

**Corollary 1** *With  $m$  and  $\Sigma$  as in Proposition 5, define  $b \equiv b(t, s, y)$  by*

$$m = \log y + b - \frac{1}{2}\Sigma^2.$$

*Then the marginal price at time  $t \in [0, T]$  of a put option with payoff  $(K - Y_T)^+$  is  $\hat{p}(t, S_t, Y_t)$ , where*

$$\begin{aligned} \hat{p}(t, s, y) &= K\Phi(-d_1 + \Sigma) - ye^b\Phi(-d_1), \\ d_1 &= \frac{1}{\Sigma} \left[ \log\left(\frac{y}{K}\right) + b + \frac{1}{2}\Sigma^2 \right]. \end{aligned}$$

*The optimal hedging strategy given by Theorem 7 with  $\hat{p}$  as an approximation to the indifference price is  $\hat{\Delta}_t \equiv \hat{\Delta}(t, S_t, Y_t)$ , where*

$$\hat{\Delta}(t, s, y) = n\rho \frac{\beta y}{\sigma s} e^b \Phi(-d_1). \quad (103)$$

In [16] these results are used to conduct a simulation study of the effectiveness of the optimal hedge under partial information (that is, with Bayesian learning about the drift parameters of the assets), compared with the BS-style hedge and the optimal hedge without learning. The results show that optimal hedging combined with a filtering algorithm to deal with drift parameter uncertainty can indeed give improved hedging performance over methods which take  $S$  as a perfect proxy for  $Y$ , and over methods which do not incorporate learning via filtering.

## 5 Investment with inside information and drift uncertainty

In the last topic of these lectures we consider the Merton optimal investment problem of Section 3, but with the added feature that the agent has some additional information at time 0, in the form of an  $\mathcal{F}_T$ -measurable random variable  $I$ , where  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$  is the background filtration. This work is based on [3].

We first expand the filtration  $\mathbb{F}$  by  $\sigma(I)$ , denoting the expanded filtration by  $\mathbb{F}^I$ . We then write the stock price dynamics under this expanded filtration which allows us to apply a filtering

algorithm that estimates the unknown drift parameter based on stock price observations as well as the anticipative knowledge possessed by the insider, represented by the random variable  $I$ . The filtering algorithm is an extension to the Kalman-Bucy filter presented earlier. It results in filtering equations whose *dynamics* are the same as the usual Kalman-Bucy equations. The main difference is the background filtration under which the filter is applied and this leads to several modifications to the initial conditions of the linear system. The goal is to restrict the insider's admissible trading strategies to be adapted to  $\hat{\mathbb{F}}^I$ , the filtration generated by stock returns, but expanded by the anticipative information.

In expanding a filtration, a crucial step is to identify the information drift,  $\nu$ , a drift added to an  $\mathbb{F}$ -Brownian motion that creates an  $\mathbb{F}^I$ -Brownian motion. We give a lemma that allows us to compute  $\nu$  explicitly.

The rest of this section is organised as follows. We first set up a linear system under an expanded filtration and present a (linear) filtering theorem that allows us to estimate an unobservable signal process based on an observation process plus a piece of anticipative information known at the beginning of the time horizon. The development of the theorem is along the lines of standard linear filters.

We then derive a lemma using a theorem in Mansuy and Yor [13] (the result stems originally from Theorem 2.1 in Jacod [7]) that gives an explicit expression for  $\nu$  in terms of the regular conditional density of  $I$ . We give an example where  $I$  is (noisy) knowledge of the terminal value  $B_T$  of a Brownian motion, and compute  $\nu$  explicitly. We then consider the investment problem to maximise expected utility of terminal wealth where  $I$  is the terminal value of an  $\mathbb{F}$ -Brownian motion distorted by noise and the agent's trading strategy is required to be  $\hat{\mathbb{F}}^I$ -adapted.

For an extension of these results to the case where  $I$  represents noisy knowledge of the terminal stock price  $S_T$ , see [3].

## 5.1 Linear filtering on an expanded filtration

We present a filtering algorithm applied to a linear system where estimation of the signal process is based on current stock price observations as well as a piece of anticipative information. The development of this filter is very much along the lines of classical linear filters, with the exception that the background filtration contains a signal process, and observation process (as in the classical case) but in addition incorporates knowledge of an  $\mathcal{F}_T$ -measurable random variable  $I$ .

A stock price process  $S = (S_t)_{0 \leq t \leq T}$  is assumed to follow

$$dS_t = \sigma S_t(\lambda dt + dB_t) =: \sigma S_t d\xi_t, \quad 0 \leq t \leq T, \quad (104)$$

where  $B$  is a Brownian motion on a complete probability space  $(\Omega, \mathcal{F}, P)$  equipped with a filtration  $\mathbb{F} := (\mathcal{F}_t)_{0 \leq t \leq T}$ . Here,  $B, S$  are  $\mathbb{F}$ -adapted, with  $\lambda$  an  $\mathcal{F}_0$ -measurable random variable. We may take, for instance,

$$\mathcal{F}_t := \sigma(\lambda, B_s, \xi_s; 0 \leq s \leq t),$$

or

$$\mathcal{F}_t := \sigma(\lambda, \xi_s; 0 \leq s \leq t),$$

since, when we observe  $\lambda$  and  $\xi$ , we also observe  $B$ .

It turns out that it is more convenient to consider the return process  $\xi$  as our observation instead of the stock price process  $S$  itself. Indeed from (104), it is clear that observations of one is equivalent to the other. Here  $B$  is an  $\mathbb{F}$ -Brownian motion.

In a full information setting,  $\lambda$  is a known constant. The aim of the partial information model of Section 3 was to remove this assumption and infer the conditional distribution of  $\lambda$  given a smaller observation filtration  $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \leq t \leq T}$  with  $\hat{\mathcal{F}}_t := \sigma(\xi_s; 0 \leq s \leq t)$ , so that  $\lambda$  is estimated using stock price observations.

Here, we first expand the background filtration  $\mathbb{F}$  with an  $\mathcal{F}_T$ -measurable r.v.  $I$ . By considering the partial information problem on an expanded filtration, we are incorporating extra information to assist in the estimation of the unknown market price of risk  $\lambda$ . To do this, we reformulate the linear system in terms of quantities adapted to the expanded filtration.



Since  $\mathbb{F}$  incorporates a Brownian filtration, we have various initial expansion results at our disposal (see [1],[7],[13],[19]). Denote the expanded filtration by  $\mathbb{F}^I := (\mathcal{F}_t^I)_{0 \leq t \leq T}$  with  $\mathcal{F}_t^I := \sigma(I, \lambda, B_s, \xi_s; 0 \leq s \leq t)$ . For our purposes it is sufficient to note that, assuming  $I$  is such that  $\mathbb{F}$ -semimartingales remain  $\mathbb{F}^I$ -semimartingales, then there exists an  $\mathbb{F}^I$ -adapted process  $\nu$  such that

$$B_t := B_t^I + \int_0^t \nu_s ds, \quad 0 \leq t \leq T, \quad (105)$$

where  $B^I$  is an  $\mathbb{F}^I$ -Brownian motion. The process  $\nu$  is called the information drift. The assumption that  $\mathbb{F}$ -semimartingales remain  $\mathbb{F}^I$ -semimartingales relies on *Jacod's criterion* [7] (see also Protter [19], Theorem VI.10), and will be satisfied in examples where  $I$  is noisy information on  $B_T$  or  $S_T$ . For cases where noise is absent, the expanded filtration  $\mathbb{F}^I$  is defined only up to time  $T-$  (Jacod's criterion is violated at time  $T$ ).

Using (105), the stock price SDE (104) can be written in terms of  $\mathbb{F}^I$ -adapted processes as

$$dS_t = \sigma S_t (\lambda_t^I dt + dB_t^I), \quad 0 \leq t \leq T, \quad (106)$$

where  $\lambda_t^I := \lambda + \nu_t$ . We interpret (106) as the insider's stock price process (in the case where  $\lambda$  is a known constant, which is precisely the condition we shall relax later in the paper), with a non-constant market price of risk, on the probability space  $(\Omega, \mathcal{F}_T^I, P)$  equipped with the filtration  $\mathbb{F}^I$ .

With the stock price process appropriately written under the expanded filtration  $\mathbb{F}^I$ , we now remove the assumption that  $\lambda$  is a known constant. We model  $\lambda$  as a Gaussian random variable with prior distribution  $N(\lambda_0, v_0)$  (the normal probability law of mean  $\lambda_0$  and variance  $v_0$ ), independent of  $B$ . The process  $\lambda^I$  is then an unobservable signal process satisfying some linear SDE under  $\mathbb{F}^I$ . We then acknowledge that estimation of the signal  $\lambda^I$  is based on  $\hat{\mathbb{F}}^I := (\hat{\mathcal{F}}_t^I)_{0 \leq t \leq T}$  defined by

$$\hat{\mathcal{F}}_t^I := \sigma(I, \xi_s; 0 \leq s \leq t)$$

with the observation process  $\xi$  satisfying

$$d\xi_t = \lambda_t^I dt + dB_t^I, \quad 0 \leq t \leq T.$$

Our filtering algorithm will convert the partial information model (106) to a full information model of the form

$$dS_t = \sigma S_t (\hat{\lambda}_t^I dt + d\hat{B}_t^I), \quad 0 \leq t \leq T \quad (107)$$

where  $\hat{B}^I$  is an  $\hat{\mathbb{F}}^I$ -Brownian motion and  $\hat{\lambda}_t^I$  is the best estimate of  $\lambda_t^I$  given  $\hat{\mathcal{F}}_t^I$

$$\hat{\lambda}_t^I := E[\lambda_t^I | \hat{\mathcal{F}}_t^I].$$

It gives us the SDE satisfied by  $\hat{\lambda}_t^I$  and, as it turns out, a Riccati equation satisfied by the error function

$$R(t) := E[(\lambda_t^I - \hat{\lambda}_t^I)^2 | \hat{\mathcal{F}}_t^I].$$

We solve for both of these quantities explicitly.

Once we have the model (107), we can treat it like a standard full information model and compute the maximum utility via duality.

Our filtering algorithm deviates from standard situations in which one seeks the best estimate of the signal given observations  $\xi$ . Here, we have these observations plus the additional knowledge of the random variable  $I$ . The following theorem gives the desired filtering equations.

**Theorem 8** *On a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ , let  $\zeta, \xi$  be  $\mathbb{F}$ -adapted processes and let  $I$  be an  $\mathcal{F}_T$ -measurable random variable. Denote by  $\mathbb{F}^I = (\mathcal{F}_t^I)_{0 \leq t \leq T}$  the filtration  $\mathbb{F}$  expanded by  $\sigma(I)$ .*

*Let  $\zeta$  be a signal process satisfying*

$$d\zeta_t = A(t)\zeta_t dt + C(t)dW_t^I, \quad 0 \leq t \leq T$$

and let  $\xi$  be an observation process satisfying

$$d\xi_t = G(t)\zeta_t dt + dB_t^I, \quad 0 \leq t \leq T$$

where  $W^I, B^I$  are  $\mathbb{F}^I$ -Brownian motions with correlation  $\rho$ , and the coefficients  $A(\cdot), C(\cdot), G(\cdot)$  are deterministic functions satisfying

$$\int_0^T \left( |A(t)| + C^2(t) + G^2(t) \right) dt < \infty.$$

Define  $\hat{\mathbb{F}}^I := (\hat{\mathcal{F}}_t^I)_{0 \leq t \leq T}$  by

$$\hat{\mathcal{F}}_t^I := \sigma(I, \xi_s; 0 \leq s \leq t).$$

Suppose the distribution of  $\zeta_0$  given  $\hat{\mathcal{F}}_0^I$  is Gaussian with mean  $m$  and variance  $\Sigma$ , independent of  $W^I$  and  $B^I$ . Then the conditional expectation  $\hat{\zeta}_t := E[\zeta_t | \hat{\mathcal{F}}_t^I]$  for  $0 \leq t \leq T$  satisfies

$$d\hat{\zeta}_t = A(t)\hat{\zeta}_t dt + \left( G(t)R(t) + \rho C(t) \right) dN_t, \quad \hat{\zeta}_0 = m \quad (108)$$

where  $N$  is the innovations process and an  $\hat{\mathbb{F}}^I$ -Brownian motion satisfying

$$dN_t = d\xi_t - G(t)\hat{\zeta}_t dt$$

and  $R(t)$  is the error function, which is independent of  $\hat{\mathcal{F}}_t^I$  and satisfies the Riccati equation

$$\frac{dR(t)}{dt} = (1 - \rho^2)C^2(t) + 2 \left( A(t) - \rho C(t)G(t) \right) R(t) - G^2(t)R^2(t), \quad R(0) = \Sigma. \quad (109)$$

The outline of the proof is as follows. We first show that the innovations process  $N$  is an  $\hat{\mathbb{F}}^I$ -Brownian motion. Since our system is linear, the filtration  $\hat{\mathbb{F}}^I$  is equivalent to the filtration  $\mathbb{F}^{N,I} := (\mathcal{F}_t^{N,I})_{0 \leq t \leq T}$  with  $\mathcal{F}_t^{N,I} := \sigma(I, N_s; 0 \leq s \leq t)$ . This is the same as standard linear filters. This allows us to express  $\hat{\mathbb{F}}^I$ -local martingales as stochastic integrals in terms of the innovations  $N$ . The remainder of the proof follows that of classical filtering theory. For brevity, we present the proof that  $N$  is an  $\hat{\mathbb{F}}^I$ -Brownian motion. The rest of the proof of Theorem 8 follows along the same lines as our earlier proofs.

**Lemma 3** *The innovations process  $N$  is an  $\hat{\mathbb{F}}^I$ -Brownian motion.*

**Proof of Lemma 3** For  $s \leq t$

$$\begin{aligned} E[N_t | \hat{\mathcal{F}}_s^I] - N_s &= E[N_t - N_s | \hat{\mathcal{F}}_s^I] \\ &= E \left[ \int_s^t G(u)(\zeta_u - \hat{\zeta}_u) ds \middle| \hat{\mathcal{F}}_s^I \right] + E \left[ E[B_t^I - B_s^I | \mathcal{F}_s^I] \middle| \hat{\mathcal{F}}_s^I \right] \\ &= 0. \end{aligned}$$

Since  $\langle N \rangle_t = \langle W \rangle_t = t$ ,  $N$  is an  $\hat{\mathbb{F}}^I$ -Brownian motion. □

## 5.2 Computing the information drift

We now present expansion of filtration results that allow us to compute the information drift in (105) explicitly. We begin with the background filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ . To expand  $\mathbb{F}$  with a  $\mathcal{F}_T$ -measurable random variable  $I$ , we first consider the process  $(\pi_t(f))_{0 \leq t \leq T}$  defined for any bounded Borel function  $f : \mathbb{R} \rightarrow \mathbb{R}$  as the continuous version of the martingale  $(E[f(I) | \mathcal{F}_t])_{0 \leq t \leq T}$ :

$$\pi_t(f) := E[f(I) | \mathcal{F}_t], \quad 0 \leq t \leq T.$$

There then exists a predictable family of measures  $(\mu_t(dx))_{0 \leq t \leq T}$  such that

$$\pi_t(f) = \int_{\mathbb{R}} f(x) \mu_t(dx).$$

For fixed  $t \in [0, T]$ , the measure  $\mu_t(dx)$  is the conditional distribution of  $I$  given  $\mathcal{F}_t$ . We assume the existence of a density function  $g(t, x, y)$  for each  $t \in [0, T]$  such that

$$\pi_t(f) = \int_{\mathbb{R}} f(x) \mu_t(dx) = \int_{\mathbb{R}} f(x) g(t, x, B_t) dx. \quad (110)$$

To state the enlargement decomposition formula, introduce the  $\mathbb{F}$ -predictable process  $(\dot{\pi}_t(f))_{0 \leq t \leq T}$  such that

$$\pi_t(f) = Ef(I) + \int_0^t \dot{\pi}_s(f) dB_s,$$

which exists by the representation property of Brownian martingales as stochastic integrals with respect to  $B$ .

We assume the existence of a predictable family of measures  $(\dot{\mu}_t(dx))_{0 \leq t \leq T}$  such that

$$\dot{\pi}_t(f) = \int_{\mathbb{R}} f(x) \dot{\mu}_t(dx).$$

We suppose that for each  $t \in [0, T]$  the measure  $\dot{\mu}_t(dx)$  is absolutely continuous with respect to  $\mu_t(dx)$  and define  $\alpha(t, x)$  by

$$\dot{\mu}_t(dx) = \alpha(t, x) \mu_t(dx).$$

Now suppose we have a continuous  $\mathbb{F}$ -martingale  $M$  given by

$$M_t = \int_0^t m_s dB_s, \quad 0 \leq t \leq T.$$

Then Theorem 1.6 in Mansuy and Yor [13] states that there exists an  $\mathbb{F}^I$ -local martingale  $M^I$  such that

$$M_t = M_t^I + \int_0^t \alpha(s, I) d\langle M, B \rangle_s,$$

provided that, almost surely,

$$\int_0^t |\alpha(s, I)| d\langle M, B \rangle_s < \infty.$$

In particular, if  $\int_0^t |\alpha(s, I)| ds < \infty$  a.s. then  $B$  decomposes as

$$B_t = B_t^I + \int_0^t \alpha(s, I) ds,$$

with  $B^I$  an  $\mathbb{F}^I$ -Brownian motion.

**Lemma 4** *The information drift is  $\nu_t = \alpha(t, I)$ , where  $\alpha(t, x)$  is given in terms of the conditional density  $g(t, x, B_t)$  by*

$$\alpha(t, x) = \frac{g_y(t, x, B_t)}{g(t, x, B_t)}, \quad 0 \leq t \leq T.$$

**Proof** From the definition of  $\alpha(t, x)$  we have

$$\begin{aligned} \dot{\pi}_t(f) &= \int_{\mathbb{R}} f(x) \alpha(t, x) \mu_t(dx) \\ &= \int_{\mathbb{R}} f(x) \alpha(t, x) g(t, x, B_t) dx. \end{aligned}$$

Hence,

$$\begin{aligned} d\pi_t(f) &= \dot{\pi}_t(f)dB_t \\ &= \left( \int_{\mathbb{R}} f(x)\alpha(t,x)g(t,x,B_t)dx \right) dB_t, \end{aligned}$$

so that

$$d\langle \pi(f), M \rangle_t = \left( \int_{\mathbb{R}} f(x)\alpha(t,x)g(t,x,B_t)dx \right) d\langle B, M \rangle_t. \quad (111)$$

But from the defining representation (110), the right-hand side of which is a smooth function of  $B_t$ , the Itô formula gives

$$d\langle \pi(f), M \rangle_t = \left( \int_{\mathbb{R}} f(x)g_y(t,x,B_t)dx \right) d\langle B, M \rangle_t, \quad (112)$$

and comparing (111) with (112) yields the result.  $\square$

Hence, given  $I$ , the information drift can be computed explicitly from the conditional density. We give an example that we shall use shortly.

**Example 1 (Noisy information on  $B_T$ )** The random variable  $I$  takes the form

$$I = aB_T + (1-a)\epsilon$$

where  $0 < a \leq 1$  and  $\epsilon \sim N(0,1)$  independent of  $B$ . The conditional distribution of  $I$  given  $\mathcal{F}_t$  is

$$N(aB_t, a^2(T-t) + (1-a)^2) = N(aB_t, a^2(T_a - t)),$$

where

$$T_a := T + [(1-a)/a]^2, \quad (113)$$

and so the conditional density is

$$g(t,x,B_t) = \frac{1}{a\sqrt{2\pi(T_a-t)}} \exp \left[ -\frac{1}{2} \frac{(x-aB_t)^2}{a^2(T_a-t)} \right]. \quad (114)$$

So by Lemma 4,

$$\alpha(t,x) = \frac{x-aB_t}{a(T_a-t)},$$

and the information drift is

$$\nu_t = \alpha(t,I) = \frac{I-aB_t}{a(T_a-t)} = \frac{a(B_T-B_t) + (1-a)\epsilon}{a(T_a-t)}. \quad (115)$$

For  $a=1$ , we get

$$\nu_t = \frac{B_T - B_t}{T-t},$$

which is a well-known result.

### 5.3 Optimal investment for an insider with drift parameter uncertainty

We now attack the main problem of this section, the optimal investment decisions of an insider who possesses some noisy information on the final value of the stock price or its driving Brownian motion. We do not assume the insider observes the underlying Brownian motion driving the stock price. The insider will therefore be restricted to using  $\hat{\mathbb{F}}^I$ -adapted strategies, where  $\hat{\mathbb{F}}^I$  is the observation filtration associated with stock returns, but expanded with the information carried by the r.v.  $I$ .

### 5.3.1 Anticipative Brownian information

We consider an example in which the insider has knowledge at time 0 of the random variable

$$I = aB_T + (1 - a)\epsilon$$

considered in Example 1 where  $0 < a \leq 1$  and  $\epsilon \sim N(0, 1)$  independent of  $B$ . This problem is artificial in the sense that we assume the insider has knowledge of  $aB_T + (1 - a)\epsilon$ , but we shall then restrict him to using  $\hat{\mathbb{F}}^B$ -adapted strategies where  $\hat{\mathbb{F}}^B$  denotes the stock return filtration expanded by  $aB_T + (1 - a)\epsilon$ . In other words, we allow the insider to have advance Brownian information, but then assume that he does not observe the Brownian filtration before time  $T$ .

We write the stock price SDE (104) in terms of  $\mathbb{F}^B$ -adapted processes, with the background filtration  $\mathbb{F}^B = (\mathcal{F}_t^B)_{0 \leq t \leq T}$  defined by

$$\mathcal{F}_t^B := \sigma(\lambda, aB_T + (1 - a)\epsilon, \xi_s; 0 \leq s \leq t).$$

Using the information drift  $\nu_t$  of Example 1, we have

$$dS_t = \sigma S_t(\lambda_t^B dt + dB_t^B), \quad 0 \leq t \leq T,$$

where  $B^B$  is an  $\mathbb{F}^B$ -Brownian motion and  $\lambda^B$  is an  $\mathbb{F}^B$ -adapted process given by

$$\lambda_t^B := \lambda + \nu_t = \lambda + \frac{I - aB_t}{a(T_a - t)} =: h(t, B_t)$$

with

$$h(t, x) = \lambda + \frac{I - ax}{a(T_a - t)}.$$

Applying Itô's formula to  $\lambda^B$ , we obtain

$$d\lambda_t^B = -\frac{1}{T_a - t} dB_t^B. \quad (116)$$

With  $\xi$  being the returns process in (38), we also have

$$d\xi_t = \lambda_t^B dt + dB_t^B. \quad (117)$$

We now regard  $\lambda$  as an unknown constant, with prior distribution  $\lambda \sim N(\lambda_0, v_0)$ . Then,  $(\lambda_t^B)_{0 \leq t \leq T}$  is an unobservable signal process following (116), and  $\xi$  is an observation process following (117). In such a filtering framework, define  $\hat{\mathbb{F}}^B := (\hat{\mathcal{F}}_t^B)_{0 \leq t \leq T}$  by

$$\hat{\mathcal{F}}_t^B := \sigma(aB_T + (1 - a)\epsilon, \xi_s; 0 \leq s \leq t).$$

Then, the best estimate of  $\lambda_t^B$  given  $\hat{\mathcal{F}}_t^B$  will be obtained from observations of the process  $\xi$  following (117) and the r.v.  $aB_T + (1 - a)\epsilon$  whose value is known at time 0. At time 0, since  $\lambda$  is assumed to be independent of  $B$ , we have

$$\text{Law}(\hat{\lambda}_0^B | \hat{\mathcal{F}}_0^B) = N\left(\lambda_0 + \frac{I}{aT_a}, v_0\right).$$

This defines the prior distribution of the signal process  $\lambda^B$ . Note that since  $I$  is  $\hat{\mathcal{F}}_t^B$ -measurable, it does not contribute to the initial variance.

Theorem 8 yields that the optimal filter

$$\hat{\lambda}_t^B := E[\lambda_t^B | \hat{\mathcal{F}}_t^B], \quad 0 \leq t \leq T,$$

satisfies the SDE

$$d\hat{\lambda}_t^B = \left(R(t) - \frac{1}{T_a - t}\right) d\hat{B}_t^B, \quad \hat{\lambda}_0^B = \lambda_0 + \frac{I}{aT_a}, \quad (118)$$

where  $\hat{B}_t^B$  is the innovations process, an  $\hat{\mathbb{F}}^B$ -Brownian motion related to the observation process by

$$d\hat{B}_t^B := d\xi_t - \hat{\lambda}_t^B dt, \quad (119)$$

and  $R(t)$  is the conditional variance of  $\lambda_t^B$ :

$$R(t) := E \left[ \left( \lambda_t^B - \hat{\lambda}_t^B \right)^2 \middle| \hat{\mathcal{F}}_t^B \right],$$

which satisfies

$$\frac{dR(t)}{dt} = \frac{2}{T_a - t} R(t) - R^2(t), \quad R(0) = v_0. \quad (120)$$

The Riccati-type ODE (120) is solved by converting it to a linear ODE for  $1/R(t)$ , yielding

$$R(t) = \left( \frac{v_0}{1 + (v_0 - (1/T_a))t} \right) \left( \frac{T_a}{T_a - t} \right).$$

More pertinently, (118) becomes

$$d\hat{\lambda}_t^B = v_t^B d\hat{B}_t^B, \quad \hat{\lambda}_0^B = \lambda_0 + \frac{I}{aT_a}, \quad (121)$$

where we have defined

$$v_t^B := R(t) - \frac{1}{T_a - t} = \frac{v_0 - (1/T_a)}{1 + (v_0 - (1/T_a))t}, \quad 0 \leq t \leq T. \quad (122)$$

Note that (121) is of exactly the same form as (40) with  $v_t$  replaced by  $v_t^B$  and with  $\hat{B}_t$  replaced by  $\hat{B}_t^B$ . Indeed,  $v_t^B$  plays the role of an ‘effective variance’, and satisfies the same Riccati equation as  $v_t$ , but with a modified initial condition:

$$\frac{dv_t^B}{dt} = - (v_t^B)^2, \quad v_0^B = v_0 - (1/T_a).$$

Using (119) in the SDE for  $\hat{\lambda}^B$ , the optimal filter is given explicitly in terms of the observable  $\xi_t$  by

$$\hat{\lambda}_t^B = \frac{\lambda_0 + (I/(aT_a)) + (v_0 - (1/T_a))\xi_t}{1 + (v_0 - (1/T_a))t} \quad (123)$$

$$= \frac{\hat{\lambda}_0^B + (v_0 - (1/T_a))\xi_t}{1 + (v_0 - (1/T_a))t} \quad (124)$$

$$= \frac{\hat{\lambda}_0^B + v_0^B \xi_t}{1 + v_0^B t}, \quad 0 \leq t. \quad (125)$$

This is of the same form as (44), with  $\lambda_0$  replaced by  $\lambda_0 + I/(aT_a)$  and  $v_0$  replaced by  $v_0^B = v_0 - (1/T_a)$ .

To solve the insider’s utility maximisation problem, we proceed in the same way as Section 3. Instead of (45), the insider is now investing in a stock with dynamics

$$dS_t = \sigma S_t \left( \hat{\lambda}_t^B dt + d\hat{B}_t^B \right), \quad 0 \leq t < T.$$

The insider’s ( $\hat{\mathbb{F}}^B$ -adapted) wealth process is  $X^B$ , following

$$dX_t^B = \sigma \theta_t^B X_t^B \left( \hat{\lambda}_t^B dt + d\hat{B}_t^B \right), \quad 0 \leq t < T,$$

where  $\theta^B$  is now a  $\hat{\mathbb{F}}^B$ -adapted strategy satisfying  $\int_0^T (\theta_t^B)^2 dt < \infty$  almost surely, and such that  $X_t^B \geq 0$  almost surely for all  $t \in [0, T)$ . Denote by  $\mathcal{A}_B$  the set of such admissible strategies.

We introduce the change of measure martingale

$$Z_t^B := \mathcal{E} \left( -\hat{\lambda}^B \cdot \hat{B}^B \right)_t.$$

By the same methods as in Section 3 we can establish a similar formula to (50) for  $Z_t^B$ . Ultimately, this simply amounts to making the replacements

$$\begin{aligned} v_t &\rightarrow v_t^B, & v_0 &\rightarrow v_0^B = v_0 - 1/T_a, \\ \hat{\lambda}_t &\rightarrow \hat{\lambda}_t^B, & \lambda_0 &\rightarrow \hat{\lambda}_0^B = \lambda_0 + I/(aT_a). \end{aligned} \quad (126)$$

Hence  $Z_t^B$  is given by

$$Z_t^B = \left( \frac{v_0^B}{v_t^B} \right)^{\frac{1}{2}} \exp \left[ -\frac{1}{2} \left( \frac{(\hat{\lambda}_t^B)^2}{v_t^B} - \frac{(\hat{\lambda}_0^B)^2}{v_0^B} \right) \right].$$

The insider's value function is

$$u_B(x) := \sup_{\theta \in \mathcal{A}_B} E[U(X_T^B) | X_0^B = x], \quad (127)$$

and we note that this is implicitly conditioned on  $\hat{\mathcal{F}}_0^B$ . The maximal expected utility is then given by a similar formula to (55):

$$u_B(x) = \frac{x^\gamma}{\gamma} H_B^{1-\gamma},$$

where

$$H_B := E \left[ (Z_T^B)^q \right],$$

and where  $H_B$  is given by a similar formula to (59), with the replacements (126). This results in

$$H_B = \left[ \frac{(1 + v_0^B T)^q}{1 + qv_0^B T} \right]^{1/2} \exp \left[ -\frac{1}{2} \frac{q(1-q)(\hat{\lambda}_0^B)^2 T}{1 + qv_0^B T} \right],$$

and the explicit expression

$$u_B(x) = \frac{x^\gamma}{\gamma} \left[ \frac{1 - \gamma - \gamma v_0^B T}{(1 - \gamma)(1 + v_0^B T)} \right]^{\gamma/2} \left[ \frac{1 - \gamma}{1 - \gamma - \gamma v_0^B T} \right]^{1/2} \exp \left\{ \frac{1}{2} \frac{\gamma \lambda_0^2 T}{1 - \gamma - \gamma v_0^B T} \right\}. \quad (128)$$

Of course, this depends on the value of  $aB_T + (1-a)\epsilon$ , since the expectation in (127) is implicitly conditioned on  $\hat{\mathcal{F}}_0^B$ , and the strategies are  $\hat{\mathbb{F}}^B$ -adapted.

The optimal wealth process  $X^{B,*}$  is given by a similar formula to (61):

$$X_t^{B,*} = x \left( \frac{\Psi_t^B}{\Psi_0^B} \right)^{\frac{1}{2}} \exp \left[ \frac{1}{2} (1-q)(\Lambda_t^B - \Lambda_0^B) \right],$$

where, for  $t \in [0, T]$ ,

$$\Psi_t^B := \frac{v_t^B}{1 + qv_t^B(T-t)}, \quad \Lambda_t^B := \frac{(\hat{\lambda}_t^B)^2}{v_t^B(1 + qv_t^B(T-t))}.$$

The optimal strategy  $\theta^{B,*}$  is given by a similar formula to (63):

$$\theta_t^{B,*} = \frac{\hat{\lambda}_t^B}{\sigma(1-\gamma)} \left( \frac{1}{1 + qv_t^B(T-t)} \right).$$

It is clear that the solution of the utility maximisation problem for the insider who has knowledge of  $aB_T + (1-a)\epsilon$  and who must optimally filter the resulting stock price drift is obtained from the solution of the partial information problem of Section 3 by making the replacements in (126).

## References

- [1] Amendinger J, Imkeller P and Schweizer M 1998 Additional logarithmic utility of an insider *Stochastic Processes and their Applications* **75** 263–286
- [2] Becherer D 2004 Utility-indifference hedging and valuation via reaction-diffusion systems *Proc. R. Soc. Lond. A* **460** 27–51
- [3] Danilova A, Monoyios M and Ng A 2008 Optimal investment with inside information and parameter uncertainty, *preprint*
- [4] Davis M H A 1997 Option pricing in incomplete markets In: *Mathematics of Derivative Securities* (eds: Dempster M A H and Pliska S R), pp 216–226 (Cambridge University Press: Cambridge, UK)
- [5] Delbaen F, Grandits P, Rheinländer T, Samperi D, Schweizer M and Stricker C 2002 Exponential hedging and entropic penalties *Mathematical Finance* **12** 99–123
- [6] Henderson V 2002 Valuation of claims on nontraded assets using utility maximization *Mathematical Finance* **12** 351–373
- [7] Jacod J 1985 Grossissement initial, hypothèse (H') et théorème de Girsanov In: *Grossissements de filtrations: exemples et applications, Lecture Notes in Math.* (eds: Jeulin T, Yor M), **1118** 15–35 (Springer: Berlin)
- [8] Kabanov Y and Stricker C 2002 On the optimal portfolio for the exponential utility maximization: remarks to the six-author paper *Mathematical Finance* **12** 125–134
- [9] Karatzas I 1996 Lectures on the Mathematics of Finance, *CRM Monographs*, **8**, American Mathematical Society
- [10] Karatzas I, Lehoczky J P, Shreve S E and Xu G-L 1991 Martingale and duality methods for utility maximization in an incomplete market *SIAM Journal on Control and Optimization* **29** 702–730
- [11] Karatzas I and Shreve S E 1991 *Brownian motion and stochastic calculus*, Second Edition, Springer-Verlag, New York
- [12] Kramkov D and Schachermayer W 1999 The asymptotic elasticity of utility functions and optimal investment in incomplete markets *Annals of Applied Probability* **9** 904–950
- [13] Mansuy R and Yor M 2006 Random times and enlargements of filtrations in a Brownian setting *Lecture Notes in Mathematics* **1873**, Springer
- [14] Monoyios M 2004 Performance of utility-based strategies for hedging basis risk *Quantitative Finance* **4** 245–255
- [15] Monoyios M 2007 Optimal hedging and parameter uncertainty *IMA Journal of Management Mathematics* **18** 331–351
- [16] Monoyios M 2008 Marginal utility-based hedging of claims on non-traded assets with partial information, *preprint*
- [17] Musiela M and Zariphopoulou T 2004 An example of indifference prices under exponential preferences *Finance & Stochastics* **8** 229–239
- [18] Pikovsky I & Karatzas I 1996 Anticipative portfolio optimization *Adv. in App. Prob.* **28** 1095–1122
- [19] Protter P 2004 *Stochastic integration and differential equations*, Second Edition, Springer, Berlin
- [20] Rogers L C G 2001 The relaxed investor and parameter uncertainty *Finance & Stochastics* **5** 131–154
- [21] Rogers L C G & Williams D *Diffusions, Markov Processes, and Martingales*, Vol. 2: Itô Calculus, Wiley, New York
- [22] Schachermayer W 2001 Optimal investment in incomplete markets when wealth may become negative *Annals of Applied Probability* **11** 694–734
- [23] Zariphopoulou T 2001 A solution approach to valuation with unhedgeable risks *Finance & Stochastics* **5** 61–82