

JIANHUI HUANG (Edmonton)

## A RECURSIVE ROBUST BAYESIAN ESTIMATION IN PARTIALLY OBSERVED FINANCIAL MARKET

*Abstract.* I propose a nonlinear Bayesian methodology to estimate the latent states which are partially observed in financial market. The distinguishable character of my methodology is that the recursive Bayesian estimation can be represented by some deterministic partial differential equation (PDE) (or evolution equation in the general case) parameterized by the underlying observation path. Unlike the traditional stochastic filtering equation, this dynamical representation is continuously dependent on the underlying observation path and thus it is robust to the modeling errors. Moreover, its advantages in financial econometrics are also discussed.

**1. Introduction.** The cutting-edge works of Merton (1969, 1971) and Black and Scholes (1973) open some frontiers of stochastic finance which models the financial states by stochastic processes such as Markov processes, or more specially, the Itô diffusion processes. Among them, the most canonical example is the Black–Scholes formula where the stock price is assumed to be some geometric Brownian motion (GBM). Through judicious choices of its parameters, this model provides enough flexibility to accommodate a wide range of dynamics of financial variables we are interested in. The motivation of my work stems from the real-world phenomenon that in many applications, the financial states of interest are often unobservable directly from the market and corrupted with some noise. Such states bear the name of “latent states”. For this reason, the parametric specification of the underlying model is just the preliminary step while in the next step, it is of

---

2000 *Mathematics Subject Classification*: 60G35, 60J60, 62M05, 62P05, 91B28, 91B70.

*Key words and phrases*: asset value, gauge transform, incomplete information, latent state, partial differential equation, partial observation, portfolio allocation, stochastic volatility.

I am grateful for the useful comments and review efforts from the anonymous referee and copy editor.

more importance to provide a reliable estimation of these latent variables, which is the main objective of filtering theory.

The filtering for discrete-time observed diffusion processes has been extensively studied in a large number of papers, including Ionides (2004) for the general case, and Nielsen and Vestergaard (2000), Durham and Gallant (2002) and Pitt (2002) for financial applications in stochastic volatility (SV).

I consider an important problem which has not been discussed before: the nonlinear robust filtering of a latent variable with continuous-time observations in financial market. In this paper, I develop an efficient PDE/evolution-based continuously dependent estimator for the latent states which bears the name of robust filter in filtering theory. My methodology offers a large number of theoretical and computational advantages which will be discussed later in detail.

The academic contributions of my work are two-fold, in the fields of econometrics and stochastic analysis: (1) The linear robust filter is already widely applied in electronic engineering even though the nonlinear robust filter is still under development. On the other hand, to the best of my knowledge, neither the linear nor nonlinear robust filter has been applied to finance research before. Therefore, this paper is the first one to discuss the robust filtering methodology in finance research. For this reason, this paper has more contributions to financial econometrics. (2) Besides its econometric content, this paper also contributes to stochastic filtering theory. A path-dependent probability change approach is introduced which explicitly characterizes the distribution change of the underlying signal process and can be applied in the more general martingale problem (MP) framework.

The rest of this paper is organized as follows: In Section 2, I provide the basic setup and introduce the estimation problem with incomplete information. Section 3 is devoted to deriving the continuously dependent estimator with the help of a gauge transform and path-dependent probability change. Some implications of my approach are summarized in Section 4. Section 5 contains the concluding remarks and potential extensions.

**2. The basic setup.** Without loss of generality, we consider a finite time interval  $[0, T]$  for given  $T > 0$ . Let  $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0})$  be a filtered complete probability space on which all stochastic processes are defined and adapted to  $\{\mathcal{F}_t\}_{t \geq 0}$ . Let  $\mathcal{N}$  denote the collection of null sets in  $(\Omega, \mathcal{F}, \mathbb{P})$ . For any stochastic process  $K$ , define its augmented natural filtration as

$$\mathcal{F}_t^K \triangleq \sigma\{K_u : 0 \leq u \leq t\} \vee \mathcal{N}.$$

**2.1. General incomplete information model.** In stochastic finance, filtration is used to represent information progressively released to the public market. We use the term “incomplete information” to represent the fact that

the information available to public market is a sub-filtration of  $\{\mathcal{F}_t\}_{t \geq 0}$ . Following Duffie and Lando (2005), the financial signal we consider is an  $n$ -dimensional latent variable  $V = (V_1, \dots, V_n)$  which is characterized by

$$(2.1) \quad dV_t = \alpha(t, V_t)dt + \sigma(t, V_t)dW_t,$$

$$(2.2) \quad \alpha = \alpha(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

$$(2.3) \quad \sigma = \sigma(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}.$$

Here,  $W_t$  is a standard,  $d$ -dimensional Brownian motion.

REMARK 2.1. The deterministic trend of  $V$  is modeled by its drift function  $\alpha$  while its randomness variance is captured by the diffusion function  $\sigma$ .

For sake of brevity and simplicity, throughout this paper, we only focus on the following one-dimensional and time-invariant case:

$$(2.4) \quad dV_t = \alpha(V_t)dt + \sigma(V_t)dW_t.$$

All results can be generalized to the general multi-dimensional and time-varying case (2.1) with parallel arguments. Moreover, to ensure the existence and uniqueness of its strong solution, we invoke the following regularity conditions.

**A1.**  $\alpha$  and  $\sigma$  satisfy the Lipschitz and linear growth conditions, that is, there exists some fixed constant  $K > 0$  such that, for all  $x, y \in \mathbb{R}$ ,

$$(2.5) \quad \alpha^2(x) + \sigma^2(x) \leq K(1 + x^2),$$

$$(2.6) \quad |\sigma(x_1) - \sigma(x_2)| + |\alpha(x_1) - \alpha(x_2)| \leq K|x_1 - x_2|.$$

It is well known that  $V$  in (2.4) is the unique solution of the martingale problem proposed by Stroock and Varadhan (1979):

DEFINITION 2.1. Let  $\mathcal{D}(A) = C_0^2(\mathbb{R})$ , the space of twice continuously differentiable functions with compact support, and

$$(2.7) \quad Af(x) = \frac{1}{2} \sigma^2(x) \frac{d^2 f(x)}{dx^2} + \alpha(x) \frac{df(x)}{dx}.$$

Then for  $f \in \mathcal{D}(A)$ ,

$$M^f(t) = f(V_t) - \int_0^t Af(V_s) ds$$

is an  $\mathcal{F}_t^V$ -martingale.

The latent variable  $V$  cannot be observed directly but is corrupted with a Gaussian white noise. Therefore, the cumulative observation up to  $t$  is

$$(2.8) \quad Y_t = \int_0^t h(V_s) ds + B_t.$$

Here, the observation noise  $B$  is a standard Brownian motion independent of  $V$  and the sensor function  $h = h(x)$  is some known Borel measurable function. To derive the continuously dependent estimator, we need further assume

**A2.**  $h \in \mathcal{D}(A)$ .

REMARK 2.2. Note that A2 implies the finite energy condition in stochastic filtering theory:

$$\mathbb{E} \int_0^T |h(V_t)|^2 dt < \infty.$$

The combined relations (2.4), (2.8) form a general incomplete information model. Similar to the work of Duffie and Lando (2005), I do not consider the insider-trading problem and the related asymmetric information in this paper.

REMARK 2.3. Consider the incomplete information model with unknown parameter vector  $\theta$ ,

$$(2.9) \quad dV_t = \alpha(V_t, \theta)dt + \sigma(V_t, \theta)dW_t.$$

It can be shown easily that our approach still works by incorporating the parameter  $\theta$  into the augmented state space  $(\theta, V)$ . The common example is the SV model in which case  $\theta$  itself can also be described by some diffusion process.

For ease of notation, we define the Lie bracket for the operator  $A$ :

DEFINITION 2.2. For  $f_1, f_2 \in \mathcal{D}(A)$ , the *Lie bracket for  $A$*  is defined as

$$(2.10) \quad [f_1, f_2] \triangleq Af_1f_2 - f_1Af_2 - f_2Af_1.$$

For  $f_1, f_2 \in \mathcal{D}(A)$ , the cross variation  $\langle M^{f_1}, M^{f_2} \rangle_t$  is well defined and we have

PROPOSITION 2.1. For  $f_1, f_2 \in \mathcal{D}(A)$ ,

$$(2.11) \quad \langle M^{f_1}, M^{f_2} \rangle_t = \int_0^t [f_1, f_2](V_s) ds.$$

In particular, if A2 holds true, then

$$(2.12) \quad \langle M^h, M^h \rangle_t = \int_0^t [h, h](V_s) ds.$$

**2.2. Incomplete information in credit spreads.** The first example of incomplete information again comes from Duffie and Lando (2005) where the impact of accounting observation noise on the term structure of bond's credit spreads is investigated. In their model, the public market's investors are not

fully informed of the status of the firm with bond issuance. In other words, the investor cannot observe the firm's asset value  $V$  directly. Instead, the cumulative observation  $Y_t$  is

$$(2.13) \quad Y_t = \int_0^t \ln V_s ds + B_t.$$

That is just the case

$$h(x) = \ln x \quad \text{for } x > 0$$

of the incomplete information model (2.4), (2.8). The observation model discussed by Nielsen and Vestergaard (2000) also fits into our framework (2.4), (2.8) when setting the sensor function to be

$$h(x) = x.$$

Generally, when  $h(x) = Cx$  for some constant  $C$ , our problem reduces to the well-known Kalman–Bucy filter with the state equation

$$(2.14) \quad dV_t = \alpha V_t dt + dW_t$$

and the Gaussian observation

$$(2.15) \quad dY_t = CV_t dt + dB_t.$$

**2.3. Incomplete information in portfolio allocation.** As mentioned in the introduction, besides the asset value, our specification (2.4) is still general enough to cover a large set of other financial quantities such as the exchange rate, equity return rate or interest rate. Now we turn to another motivating example of an incomplete information model (2.4), (2.8): estimation of the equity return rate in an optimal portfolio allocation problem.

There are many research papers on optimal portfolio choice and market equilibrium with incomplete information, including Detemple (1986), Dothan and Feldman (1986) and Back (2003). In those works, the asset price process satisfies

$$(2.16) \quad \frac{dS}{S} = \mu_t dt + \sigma dW_t$$

and the instantaneous equity return rate  $\mu_t$  satisfies

$$(2.17) \quad d\mu_t = \phi_1(\theta - \mu_t)dt + \phi_2 dB_t,$$

where  $W$  and  $B$  are independent Brownian motions. In contrast to the general setting of stochastic volatility, here we suppose that the investor can observe  $S$  and the volatility  $\sigma$  is some known constant. Instead, he cannot observe  $\mu$  directly, which acts as the latent variable here. Based on this, the filtration available to the investor is  $\mathcal{F}_t^S$  and his primary target is to extract some estimation of the return rate from the observation  $S$ . Now

introduce

$$Y_t = \frac{1}{\sigma} \ln S_t.$$

Then from the Itô formula, we have

$$(2.18) \quad dY_t = \left[ \frac{\mu_t}{\sigma} - \frac{\sigma}{2} \right] dt + dW_t.$$

Hence, the model (2.17), (2.18) is a special case of our general model (2.4), (2.8) if we set

$$h(x) = \frac{x}{\sigma} - \frac{\sigma}{2}.$$

**3. Continuous-dependence estimation with incomplete information.** We now consider the problem of estimating the latent states in the incomplete information model (2.4), (2.8). This problem belongs to filtering theory. In this case, the information reaching the market is  $\{\mathcal{F}_t^Y\}_{t \geq 0}$  and the minimal square variance estimation of  $V$  is equivalent to the conditional distribution

$$(3.1) \quad \pi_t(\cdot) \triangleq \mathbb{P}[V_t \in \cdot \mid \mathcal{F}_t^Y],$$

or the conditional expectation

$$(3.2) \quad \pi_t(f) \triangleq \mathbb{E}[f(V_t) \mid \mathcal{F}_t^Y]$$

for a rich enough class of test functions  $f$ .

**3.1. Stochastic filtering equation.** One of the main results of stochastic filtering theory, due to Fujisaki, Kallianpur and Kunita (1972), is the following Fujisaki–Kallianpur–Kunita (FKK) equation: for  $f \in \mathcal{D}(A)$ ,  $\pi_t(f)$  satisfies

$$(3.3) \quad d\pi_t(f) = \pi_t(Af)dt + [\pi_t(hf) - \pi_t(h)\pi_t(f)][dY_t - \pi_t(h)dt].$$

The FKK equation links the optimal filter  $\pi_t$  with some measure-valued stochastic partial differential equation driven by the observation process. However, the structure of the FKK equation is too complicated to apply, both for the purposes of numerical computation and theoretical analysis. One of the undesirable features is its nonlinearity with respect to  $f$ .

On the other hand, an equivalent but much more tractable equation is developed which bears the name of Duncan–Mortensen–Zakai (DMZ) equation (see Zakai (1969)). The DMZ equation is a linear measure-valued stochastic filtering equation and characterizes the dynamics of the unnormalized conditional distribution  $\sigma_t$  which links to  $\pi_t$  through the Kallianpur–Striebel–Bayes formula. From A2 and (2.8), we know that

$$(3.4) \quad \Lambda_t \triangleq \exp\left(-\int_0^t h(V_s) dY_s + \frac{1}{2} \int_0^t |h(V_s)|^2 ds\right)$$

is an  $\{\mathcal{F}_t\}_{t \geq 0}$  martingale with  $\mathcal{F}_t \triangleq \sigma\{V_s, Y_s : 0 \leq s \leq t\} \vee \mathcal{N}$ . Therefore,

$$\frac{d\mathbb{P}}{d\overline{\mathbb{P}}}\bigg|_{\mathcal{F}_t} = (\Lambda_t)^{-1}$$

defines a probability measure  $\overline{\mathbb{P}}$  which is mutually absolutely continuous with respect to  $\mathbb{P}$  and as a standard result in filtering theory, we get

PROPOSITION 3.1. *Suppose A2 and (2.8) are true. Then under  $\overline{\mathbb{P}}$ ,  $Y$  is a Brownian motion independent of  $V$  and the law of  $V$  keeps unchanged.*

Let  $B(\mathbb{R})$  denote the set of all bounded and measurable functions. Then for  $f \in B(\mathbb{R})$ , the Kallianpur–Striebel–Bayes formula links  $\pi_t$  to the unnormalized filter  $\sigma_t$  by

$$(3.5) \quad \pi_t(f) = \frac{\sigma_t(f)}{\sigma_t(1)},$$

where

$$(3.6) \quad \sigma_t(f) \triangleq \overline{\mathbb{E}}(f(X_t)\Lambda_t^{-1} | \mathcal{F}_t^Y).$$

Note that by (3.3), (3.5) and

$$(3.7) \quad \sigma_t(1) = \sigma_0(1) + \int_0^t \sigma_s(h) dY_s,$$

from the Itô formula, it follows that the unnormalized filter is characterized by the following DMZ equation: for  $f \in \mathcal{D}(A)$ ,

$$(3.8) \quad \sigma_t(f) = \sigma_0(f) + \int_0^t \sigma_s(Af) ds + \int_0^t \sigma_s(hf) dY_s.$$

Moreover, if we assume that the unnormalized conditional distribution  $\sigma_t$  admits a density process  $q_t$ , then it turns out to satisfy the following Kushner equation (Kushner (1964, 1967)):

$$(3.9) \quad dq_t(x) = A^*q_t(x)dt + h(x)q_t(x)dY_t,$$

where

$$A^*f(x) \triangleq \frac{1}{2} \frac{\partial}{\partial x} \left( \sigma^2(x) \frac{\partial f(x)}{\partial x} \right) - \frac{\partial}{\partial x} (\alpha(x)f(x))$$

is the Fokker–Planck operator associated with  $A$  in (2.7).

Before making the filter computation, it is important to make sure the FKK and DMZ equations admit a unique solution. The interested readers may refer to Rozovskii (1991) for the results on uniqueness.

It is remarkable that the Zakai and Kushner equations are superior to the FKK equation in that they are linear with respect to  $f$ . However, they still require some stochastic integrations to be performed. As explained by Clark (1978, 2005), these stochastic filtering equations are actually impractical to

implement in real life. This is because the stochastic integrations involved are not consistent with modeling errors which seem to be unavoidable and ubiquitous. Meanwhile, the above filtering equations are not continuously dependent on the underlying observation process. Empirical results conducted in engineering show that the continuously dependent filter does indeed perform well in real data problems. This is especially the case when the assumption of Brownian observation noise is unrealistic. We will present a detailed discussion of this point in a later section.

As explained above, to estimate the latent variables with incomplete information, we need to derive some recursive representation of the filter  $\pi_t$  which should be continuously dependent on the underlying observation and avoid stochastic integration. We solved this problem in two different but related directions: (1) The PDE filtering equation corresponding to the Kushner equation. (2) The evolution filtering equation corresponding to the DMZ equation. The filters derived along these two directions are essentially equivalent while the latter is more general as it holds true for the general martingale problem setup.

**3.2. Filtering equation based on PDE.** Let us hereafter denote by  $y = \{y_t, 0 \leq t \leq T\}$  an arbitrary but fixed observation trajectory. In other words,  $y_t = Y(t, \omega)$  for all  $t \geq 0$  and some  $\omega \in \Omega$ . As the first direction, we will derive some PDE parameterized by the observation path using the elegant tools originally due to Krylov and Rozovskii (1981) as well as Pardoux (1979) based on the following “gauge transform” to density process.

DEFINITION 3.1. The gauge transform of  $q_t(x)$  in (3.9) is defined as

$$(3.10) \quad \hat{q}_t(x) \triangleq q_t(x) \cdot \exp(\kappa_t(x)),$$

where

$$\kappa_t(x) \triangleq e^{-h(x)Y_t}.$$

REMARK 3.1.  $\hat{q}_t(x)$  and  $q_t(x)$  are equivalent as we can recover  $q_t(x)$  once  $\hat{q}_t(x)$  is given. The converse also holds true.

The main result of this part is

THEOREM 3.1. Assume A1, A2 and suppose  $\alpha, \sigma$  are continuously differentiable. Then the gauge transform  $\hat{q}_t(x)$  satisfies the following linear second-order partial differential equation:

$$(3.11) \quad \frac{\partial \hat{q}_t(x)}{\partial t} = \frac{1}{2} \sigma^2(x) \frac{\partial^2}{\partial x^2} \hat{q}_t(x) + \Phi_1(t, x, Y_t) \frac{\partial}{\partial x} \hat{q}_t(x) + \Phi_2(t, x, Y_t) \hat{q}_t(x),$$

where

$$(3.12) \quad \Phi_1(t, x, y) \triangleq y\sigma^2(x)h'(x) + \sigma(x)\sigma'(x) - \alpha(x),$$

$$(3.13) \quad \Phi_2(t, x, y) \triangleq \frac{1}{2} \sigma^2(x)(h''(x)y + (h'(x)y)^2) + yh'(x)(\sigma(x)\sigma'(x) - \alpha(x)) - \left( \alpha'(x) + \frac{1}{2} h^2(x) \right).$$

Moreover,  $\widehat{q}_t(x)$  is continuously dependent on  $Y_t$ .

*Proof.* First,  $\kappa_t$  satisfies the equation

$$d\kappa_t(x) = \kappa_t(x) \left( -h(x)dY_t + \frac{1}{2} h^2(x)dt \right).$$

Integration by parts yields

$$(3.14) \quad \begin{aligned} d\widehat{q}_t(x) &= q_t(x)d\kappa_t(x) + \kappa_t(x)dq_t(x) + d\langle q(x), \kappa(x) \rangle_t \\ &= q_t(x)\kappa_t(x) \left( -h(x)dY_t + \frac{1}{2} h^2(x)dt \right) \\ &\quad + \kappa_t(x)(h(x)q_t(x)dY_t + A^*q_t(x)dt) \\ &\quad + \langle h(x)q_t(x)dY_t, \kappa_t(-h(x))dY_t \rangle \\ &= \frac{1}{2} q_t(x)\kappa_t(x)h^2(x)dt + \kappa_t(x)A^*q_t(x)dt - q_t(x)\kappa_t(x)h^2(x)dt \\ &= \kappa_t(x)A^*q_t(x)dt - \frac{1}{2} q_t(x)\kappa_t(x)h^2(x)dt \\ &= \kappa_t(x)A^* \left( \frac{\widehat{q}_t(x)}{\kappa_t(x)} \right) dt - \frac{1}{2} \widehat{q}_t(x)h^2(x)dt. \end{aligned}$$

That is,

$$(3.15) \quad \frac{\partial \widehat{q}_t(x)}{\partial t} = \kappa_t(x)A^* \left( \frac{\widehat{q}_t(x)}{\kappa_t(x)} \right) - \frac{1}{2} \widehat{q}_t(x)h^2(x).$$

Rearranging the resulting terms, we get the representation of  $\Phi_1$  and  $\Phi_2$  immediately. Since  $\alpha, \sigma$  are continuously differentiable, the solution of (3.11)–(3.13) is continuously dependent on its coefficients. This implies the continuous dependence on the observations. This completes the proof. ■

An important example will be the case when  $h(x) = x$  and  $\alpha(x) = \alpha x$ ,  $\sigma(x) = \sigma x$  as in Nielsen and Vestergaard (2000). In this case, we have

$$(3.16) \quad \Phi_1(t, x, y) \triangleq y\sigma^2x^2 + \sigma^2x - \alpha x,$$

$$(3.17) \quad \Phi_2(t, x, y) \triangleq \frac{1}{2} \sigma^2x^2y^2 + y[\sigma^2x - \alpha x] - \left( \alpha + \frac{1}{2} x^2 \right).$$

**3.3. Filtering equation based on an evolution equation.** As another direction, we now derive some evolution equation parameterized by the observation path  $y_t$ . This evolution equation is another continuously dependent representation of the optimal filter with respect to the latent variables. My proof relies on the path-dependent probability change and the following

“gauge transform” for  $\sigma_t(f)$ . The papers of Davis (1980, 1981) discuss the semigroup approach to nonlinear robust filter.

DEFINITION 3.2. For  $f \in B(\mathbb{R})$ , the *gauge transform* of  $V$  is defined as

$$(3.18) \quad \nu_t(f) \triangleq \overline{\mathbb{E}}[f(V_t)A_t^{-1} \exp(-Y_t h(V_t)) | \mathcal{F}_t^Y]$$

and it is equivalent to  $\sigma_t$  because

$$(3.19) \quad \nu_t(f) = \sigma_t(f \exp(-Y_t h)) \quad \text{and} \quad \sigma_t(f) = \nu_t(f \exp(Y_t h)).$$

The main result of this part is

THEOREM 3.2. Assume A1 and A2 hold true. Then  $\nu_t(f)$  satisfies the evolution equation

$$(3.20) \quad \frac{d}{dt} \nu_t(f) = \nu_t(\widehat{A}^y f) + \nu_t\left(f \left\{ -Y_t A h - \frac{1}{2} h^2 + \frac{Y_t^2}{2} [h, h] \right\}\right)$$

for all  $f \in \mathcal{D}(\widehat{A}_t^y) = \mathcal{D}(A)$  where  $\widehat{A}_t^y f = Af - Y_t[h, f]$  and  $A$  is defined in (2.7).

To prove it, we need the following preliminary results.

PROPOSITION 3.2. For a given observation path  $y_t$ ,

$$\nu_t(f) = \overline{\mathbb{E}}[f(V_t)A_t^y \Xi_t],$$

where

$$(3.21) \quad A_t^y \triangleq \exp\left(-\int_0^t y_s dM_s^h - \int_0^t \frac{y_s^2}{2} [h, h](V_s) ds\right),$$

$$(3.22) \quad \Xi_t \triangleq \exp\left(\int_0^t \left\{ -y_s A h(V_s) - \frac{1}{2} h^2(V_s) + \frac{y_s^2}{2} [h, h](V_s) \right\} ds\right).$$

*Proof.* From the independence of  $B, V$  and integration by parts, we have

$$(3.23) \quad \int_0^t h(V_s) dY_s = Y_t h(V_t) - \int_0^t Y_s dh(V_s).$$

It follows that

$$\begin{aligned} A_t^{-1} &= \exp\left(\int_0^t h(V_s) dY_s - \frac{1}{2} \int_0^t h^2(V_s) ds\right) \\ &= \exp\left(Y_t h(V_t) - \int_0^t Y_s dh(V_s) - \frac{1}{2} \int_0^t h^2(V_s) ds\right). \end{aligned}$$

Consequently,

$$(3.24) \quad \nu_t(f) = \overline{\mathbb{E}}\left[f(V_t) \exp\left(-\int_0^t Y_s dh(V_s) - \frac{1}{2} \int_0^t h^2(V_s) ds\right) \middle| \mathcal{F}_t^Y\right].$$

Therefore, for a given observation path  $y_t$ ,

$$(3.25) \quad \nu_t(f) = \mathbb{E} \left[ f(V_t) \exp \left( - \int_0^t y_s A h(V_s) ds - \frac{1}{2} \int_0^t h^2(V_s) ds - \int_0^t y_s dM_s^h \right) \right].$$

Hence the result. ■

LEMMA 3.1.  $A_t^y$  is a continuous martingale.

*Proof.* It is easy to see that  $A_t^y$  is a continuous process and

$$(3.26) \quad dA_t^y = -A_t^y y_t dM_t^h,$$

which implies  $A_t^y$  is a local martingale. Moreover, from A2,

$$\mathbb{E} \exp \left( \frac{1}{2} \int_0^t [h, h](V_s) ds \right) < \infty, \quad \forall t \in [0, T]. \quad (1)$$

Then from the Novikov criterion,  $A_t^y$  is a martingale. ■

Now, we are ready to introduce the path-dependent probability transform.

DEFINITION 3.3. For given  $y_t$ , introduce the path-dependent probability measure  $\widehat{\mathbb{P}}^y$  by

$$\left. \frac{d\widehat{\mathbb{P}}^y}{d\mathbb{P}} \right|_{\mathcal{F}_t} = A_t^y.$$

After the first measure change to  $\overline{\mathbb{P}}$ , the law of the signal process  $V$  remains unchanged. In contrast, with the second path-dependent measure change, the law of  $V$  will be changed and characterized by some observation-dependent martingale problem. The gauge transform  $\nu(\cdot)$  then takes the form of the Feynman–Kac multiplicative functional

$$\nu_t(f) = \widehat{\mathbb{E}}^y[f(V_t)\Xi_t],$$

where

$$(3.27) \quad \Xi_t = \exp \left( \int_0^t \left( -y_s A h(V_s) - \frac{1}{2} h^2(V_s) + \frac{y_s^2}{2} [h, h](V_s) \right) ds \right).$$

LEMMA 3.2. Assume A1–A2. Then under  $\widehat{\mathbb{P}}^y$ ,  $V$  is the unique solution to the  $\widehat{A}_t^y$  martingale problem

$$(3.28) \quad df(V_t) = \widehat{A}_t^y f(V_t) dt + d\widehat{M}_t^f,$$

where  $\mathcal{D}(\widehat{A}_t^y) = \mathcal{D}(A)$  and for  $f \in \mathcal{D}(\widehat{A}_t^y)$ ,

$$(3.29) \quad \widehat{A}_t^y f \triangleq Af - y_t[h, f],$$

$$(3.30) \quad d\widehat{M}_t^f \triangleq dM_t^f + y_t[h, f](X_t) dt.$$

---

(1) Equivalently,  $\mathbb{E} \exp(\frac{1}{2} \int_0^T [h, h](V_s) ds) < \infty$ .

*Proof.* For  $f \in \mathcal{D}(A)$ , it is obvious that

$$(3.31) \quad df(V_t) = \widehat{A}_t^y f(V_t)dt + d\widehat{M}_t^f$$

and we only need to show  $\widehat{M}_t^f$  is a martingale. By the continuity of  $A^y$ , we have

$$\langle M^f, A^y \rangle_t = \int_0^t -A_s^y y_s [f, h] ds$$

and

$$(3.32) \quad d\widehat{M}_t^f = dM_t^f - \frac{1}{A_t^y} d[M^f, A^y]_t.$$

Thus, from the Girsanov–Meyer theorem,  $\widehat{M}_t^f$  is a local martingale under  $\widehat{\mathbb{P}}^y$ . Moreover, from A2,

$$\widehat{\mathbb{E}}^y[\widehat{M}^f, \widehat{M}^f]_t = \overline{\mathbb{E}}([M^f, M^f]A^y)_t < \infty.$$

It follows that  $\widehat{M}_t^f$  is a martingale under  $\widehat{\mathbb{P}}^y$ . This completes the proof. ■

*Proof of Theorem 3.2.* From Proposition 3.2, we have

$$\nu_t(f) = \widehat{\mathbb{E}}^y(f(V_t)\Xi_t).$$

From Lemma 3.2 and integration by parts,

$$(3.33) \quad d(f(V_t)\Xi_t) = \widehat{A}_t^y f(V_t)\Xi_t dt + f(V_t)\Xi_t \left\{ -y_t Ah(V_t) - \frac{1}{2}h^2(V_t) + \frac{y_t^2}{2} [h, h](V_t) \right\} dt + \Xi_t d\widehat{M}_t^f.$$

Note that  $\int_0^t \Xi_s d\widehat{M}_s^f$  is a  $\widehat{\mathbb{P}}^y$ -martingale under A1 and A2. Hence taking the expectation under  $\widehat{\mathbb{P}}^y$ , we get

$$\nu_t(f) = \nu_0(f) + \int_0^t \nu_s(\widehat{A}_s^y f) ds + \int_0^t \nu_s \left( f \left\{ -y_s Ah - \frac{1}{2}h^2 + \frac{y_s^2}{2} [h, h] \right\} \right) ds.$$

Replace the fixed observation path  $y_t$  with the observation process  $Y_t$  to get the result. ■

**4. Implications of the continuous-dependence estimator.** Our continuously dependent estimation approach, summarized by the PDE (3.11) and evolution equation (3.20), distinguishes itself from other extant estimation methods in two theoretical features:

REMARK 4.1. The equations (3.11), (3.20) are well defined for all observation trajectories  $y_t \in C[0, T]$  whereas the FKK, DMZ or Kushner equations are only defined on a subset of  $C[0, T]$  with Wiener measure 1. Here,  $C[0, T]$  is the set of all continuous functions on  $[0, T]$ .

REMARK 4.2. The equation (3.11) is a parabolic second-order partial differential equation which is continuously dependent on the observation trajectory  $y_t$ ; the evolution equation (3.20) is also continuously dependent on the observation trajectory  $y_t$ .

Our approach also provides some computational and practical advantages which are demonstrated as follows:

**4.1. Particle filter and PDE.** One advantage of our continuously dependent estimation (3.11), (3.20) is its computational efficiency which becomes the first concern in real-time online computation. It is worthwhile to point out that the numerical methods for computing the nonlinear filter in stochastic filtering equations are poorly developed. The widely employed approach to solving the stochastic filtering equation is to make some Monte-Carlo simulation based on the transition distribution which naturally leads to the development of a particle filter algorithm.

However, the real-time computation outcomes demonstrate that the convergence rate of the particle filter is not always satisfactory and reliable, except for a few cases such as the finite-dimensional linear, Gaussian state space model (the Kalman–Bucy filter). On the other hand, it is remarkable that our continuously dependent estimations (3.11), (3.20) have no stochastic integration involved. Thus these equations can be solved without the introduction of Monte Carlo simulation and the associated particle filter. As a result, the nonlinear estimation of latent variables with incomplete information can be solved using only the finite difference or finite element methods, which turns out to be more computationally effective and reliable. Some numerical computations are in progress and will be included in companion papers.

**4.2. Discrete and continuous-time observation.** Another advantage of our continuously dependent estimation is as follows. Stemming from the background of practical finance, there has been a dichotomy between models with discrete-time observations and those with continuous-time observations. Note that while the former models gain popularity due to their favorable statistical properties (see Gordon, Salmond and Smith (1993)), the latter are more elegant and powerful in theoretical analysis (see Lipster and Shiryaev (2001)).

Although our continuously dependent estimator is derived on the basis of continuous-time observation, it can be shown that our method still provides some sustainable and consistent estimation of the underlying latent variable under some mild conditions, whether the observations are discrete-time or continuous-time.

We consider the filtering problem in which the observation sampling intervals converge to zero. An illustrating example is the high frequency data.

Of course, we can relax this assumption using the technique of “filling-in-the-missing-data” by Johannes, Polson and Stroud (2000). Suppose we are given a discrete-time observation model. That means the available observation stream can only take some discrete-time form, for example, the discrete time instances

$$0 = t_0 < t_1 < \dots < t_n = t$$

instead of continuous-time observation. Correspondingly, the observation time series will be  $\{Y_{t_i}\}_{i=1}^n$ . However, we can get some continuous-time observation path  $\widehat{Y}_{[0,t]}$  from the linear interpolation of  $\{Y_{t_i}\}_{i=1}^n$ . The derived path  $\widehat{Y}_{[0,t]}$  carries the same information on the latent state variable. It is also close to the observation path  $Y_{[0,t]}$  which is implied in our observation model (2.8) in the sup-norm defined on  $C[0, T]$ . Consequently, due to the continuity of the observation path  $Y_{[0,t]}$ , the continuously dependent estimator provides a sensible approximation to the true conditional distribution and it can still be adapted to handle the discrete-time observation case. In particular, this is always the case for high frequency sampling.

**4.3. Color noise added.** When using the filtering method, a problem naturally arising is how to quantify or calibrate the estimation error caused by the modeling inaccuracy. The most common modeling inaccuracy is caused by the popular use of Brownian motion as the idealized observation noise in filtering framework. In fact, the additive white noise observation model (2.8) is an idealization of the realistic observation mechanism. Empirical research demonstrates it is more appropriate to impose the colored noise (e.g. Ornstein–Uhlenbeck process) instead of white noise (Brownian motion) on the observation. Unlike the traditional filtering algorithm which is more susceptible to modeling errors, our continuously dependent estimator is “immune” to such errors (see Clark (2005) for details).

**5. Concluding remarks.** The proposed continuously dependent estimators, which are the solutions of (3.11), (3.20), also allow some stimulating generalizations such as the Markov state process or the more general martingale problem setup of Ethier and Kurtz (1986). This paper also motivates some innovative viewpoints on the existing financial research topics (e.g. the SV problem) which are studied in our parallel work. Some simulation works providing computational evidence for our proposed approach will be presented in some companion papers with J. Li.

## References

- K. Back (2003), *Partial and Asymmetric Information*, Lecture Notes in CIME-EMS, 2003.  
 F. Black and M. Scholes (1973), *The pricing of options and corporate liabilities*, J. Political Economy 81, 637–654.

- J. M. C. Clark (1978), *The design of robust approximations to the stochastic differential equations of nonlinear filtering*, in: Communication Systems and Random Process Theory (Darlington, 1977), NATO Adv. Study Inst. Ser., Ser. E Appl. Sci. 25, Sijthoff & Noordhoff, Alphen aan den Rijn, 1978, 721–734.
- J. M. C. Clark and D. Crisan (2005), *On a robust version of the integral representation formula of nonlinear filtering*, Probab. Theory Related Fields 133, 43–56.
- M. H. A. Davis (1980), *On a multiplicative functional transformation arising in nonlinear filtering theory*, Z. Wahrsch. Verw. Gebiete 54, 125–139.
- M. H. A. Davis (1981), *Factorization of a multiplicative functional of nonlinear filtering theory*, System Control Lett. 25, 260–278.
- J. Detemple (1986), *Asset pricing in a production economy with incomplete information*, J. Finance 41, 383–391.
- M. Dothan and D. Feldman (1986), *Equilibrium interest rates and multi-period bonds in a partially observable economy*, *ibid.* 41, 369–382.
- D. Duffie and D. Lando (2001), *Term structures of credit spreads with incomplete accounting information*, Econometrica 69, 633–664.
- G. B. Durham and A. R. Gallant (2002), *Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes*, J. Business Econom. Statist. 20, 297–316.
- S. N. Ethier and T. G. Kurtz (1986), *Markov Processes: Characterization and Convergence*, Wiley, New York.
- M. Fujisaki, G. Kallianpur and H. Kunita (1972), *Stochastic differential equations for the nonlinear filtering problem*, Osaka J. Math. 9, 19–40.
- N. Gordon, D. Salmond and A. Smith (1993), *Novel approach to non-linear/nongaussian Bayesian state estimation*, IEEE Proc.-F. 140, 107–113.
- E. L. Ionides (2004), *Inference and filtering for partially observed diffusion processes via sequential Monte Carlo*, Univ. of Michigan Statistics Department, Technical Report 405.
- M. Johannes, N. Polson and J. Stroud (2003), *Nonlinear filtering of stochastic differential equations with jumps*, J. Econometrics 9, 19–40.
- N. V. Krylov and B. L. Rozovskii (1981), *Stochastic evolution equations*, J. Soviet Math. 16, 1233–1276.
- H. J. Kushner (1964), *On the differential equations satisfied by conditional probability densities of Markov processes with applications*, SIAM J. Control 2, 106–119.
- H. J. Kushner (1967), *Dynamical equations for optimal nonlinear filtering*, J. Differential Equation 3, 179–190.
- R. S. Lipster and A. N. Shiryaev (2001), *The Statistics of Random Processes*, Springer, Berlin.
- R. C. Merton (1969), *Lifetime portfolio selection under uncertainty: the continuous-time case*, Rev. Econom. Statist. 51, 247–257.
- R. C. Merton (1971), *Optimum consumption and portfolio rules in a continuous time model*, J. Econom. Theory 3, 373–413.
- J. Nielsen and M. Vestergaard (2000), *Estimation in continuous-time stochastic volatility models using nonlinear filters*, Int. J. Theor. Appl. Finance 3, 279–308.
- E. Pardoux (1979), *Stochastic partial differential equations and filtering of diffusion processes*, Stochastics 3, 127–167.
- M. Pitt (2002), *Smooth particle filters for likelihood evaluation and maximization*, working paper, Univ. of Warwick.
- B. L. Rozovskii (1991), *A simple proof of uniqueness for Kushner and Zakai equations*, in: Stochastic Analysis, E. Mayer-Wolf *et al.* (eds.), Academic Press, New York, 449–458.

D. W. Stroock and S. R. S. Varadhan (1979), *Multidimensional Diffusion Processes*, Springer, Berlin.

M. Zakai (1969), *On the optimal filtering of diffusion processes*, Z. Wahrsch. Verw. Gebiete 11, 230–243.

Department of Statistics  
University of Alberta  
Edmonton, AB, Canada T6G 2G1  
E-mail: jhuang@stat.ualberta.ca

*Received on 19.1.2007;*  
*revised version on 17.5.2007*

(1856)