A FILTERING APPROACH TO TRACKING VOLATILITY FROM PRICES OBSERVED AT RANDOM TIMES

BY JAKŠA CVITANIĆ,[1] ROBERT LIPTSER AND BORIS ROZOVSKII[2]

Caltech, Tel Aviv University and University of Southern California

This paper is concerned with nonlinear filtering of the coefficients in asset price models with stochastic volatility. More specifically, we assume that the asset price process $S = (S_t)_{t \geq 0}$ is given by

$$dS_t = m(\theta_t)S_t \, dt + v(\theta_t)S_t \, dB_t,$$

where $B = (B_t)_{t \geq 0}$ is a Brownian motion, v is a positive function and $\theta = (\theta_t)_{t \geq 0}$ is a cádlág strong Markov process. The random process θ is unobservable. We assume also that the asset price S_t is observed only at random times $0 < \tau_1 < \tau_2 < \cdots$. This is an appropriate assumption when modeling high frequency financial data (e.g., tick-by-tick stock prices).

In the above setting the problem of estimation of θ can be approached as a special nonlinear filtering problem with measurements generated by a multivariate point process $(\tau_k, \log S_{\tau_k})$. While quite natural, this problem does not fit into the "standard" diffusion or simple point process filtering frameworks and requires more technical tools. We derive a closed form optimal recursive Bayesian filter for θ_t, based on the observations of $(\tau_k, \log S_{\tau_k})_{k \geq 1}$. It turns out that the filter is given by a recursive system that involves only deterministic Kolmogorov-type equations, which should make the numerical implementation relatively easy.

1. Introduction. In the classical Black–Scholes model for financial markets, the stock price S_t is modeled as a geometric Brownian motion, that is, with diffusion coefficient equal to σS_t, where "volatility" σ is assumed to be constant. The volatility parameter is of great importance in applications of the model, for example, for option pricing. Consequently, many researchers have generalized the constant volatility model to so-called stochastic volatility models, where σ_t is itself random and time dependent. There are two basic classes of models: complete and incomplete. In complete models, the volatility is assumed to be a functional of the stock price; in incomplete models, it is driven by some other source of noise that is possibly correlated with the original Brownian motion. In this paper we study a particular incomplete model in which the volatility process is independent

of the driving Brownian motion process. This has the economic interpretation of the volatility being influenced by market, political, financial and other factors that are independent of the "systematic risk" (the Brownian motion process) associated with the particular stock price under study. Option traders, investment banks, economic analysts and others depend on modeling future volatility for their trading, economic forecasts, risk management and so on.

Estimating volatility from observed stock prices is not a trivial task in either complete or incomplete models, in part because the prices are observed at discrete, possibly random time points. Since volatility itself is not observed, it is natural to apply filtering methods to estimate the volatility process from historical stock price observations. Nevertheless, this has only recently been investigated in continuous-time models, in particular, by Frey and Runggaldier [5]. See [22] for an up-to-date survey. See also [2] for a discrete-time approach with equally spaced observations, [6] for an approximating algorithm in continuous time, [17] for a nonparametric approach, as well as [3, 10, 19] for still other approaches. There is also a rich econometrics, time-series literature on ARCH–GARCH models of stochastic volatility, that presents an alternative way to model and estimate volatility; see [7] for a survey.

Our paper was prompted by Frey and Runggaldier [5]. Like that paper, we assume that the asset price process $S = (S_t)_{t \geq 0}$ is given by

$$dS_t = m(\theta_t) S_t \, dt + v(\theta_t) S_t \, dB_t,$$

where $B = (B_t)_{t \geq 0}$ is a Brownian motion, v is a positive function, and $\theta = (\theta_t)_{t \geq 0}$ is a cádlág strong Markov process. The "volatility" process θ is unobservable, while the asset price S_t is observed only at random times $0 < \tau_1 < \tau_2 < \cdots$. This assumption is designed to reflect the discrete nature of high frequency financial data such as tick-by-tick stock prices. The random time moments τ_k can be interpreted as "instances at which a large trade occurs or at which a market maker updates his quotes in reaction to new information" (see [4]). Hence, it is natural to assume that $\{\tau_k\}_{k \geq 1}$ might also be correlated with θ.

In the above setting the problem of volatility estimation can be regarded as a special nonlinear filtering problem.

Frey and Runggaldier [5] derive a Kallianpur–Striebel type formula (see, e.g., [9]) for the optimal mean-square filter for θ_t based on the observations of $S_{\tau_1}, S_{\tau_2}, \ldots$ for all $\tau_k \leq t$ and investigate Markov chain approximations for this formula. We extend this result in that we derive the exact filtering equations for θ_t that allow us to compute the conditional distribution of θ_t given $S_{\tau_1 \wedge t}, S_{\tau_2 \wedge t}, \ldots$. Moreover, our framework includes general random times of observations, not just doubly stochastic Poisson processes.

We remark that, while being natural, the Frey and Runggaldier model adopted in this paper does not quite fit into the "standard" diffusion or simple point process filtering frameworks (cf. [12, 15, 20]) and requires more technical tools. In particular, the general filtering theory for diffusion processes requires that the diffusion

coefficient of the observation process does not depend on the state process, while in our case the presence of θ_t in the diffusion coefficient is crucial. The "standard" filtering theory for point processes is also not applicable in the present setting since the observation process $(\tau_i, S_{\tau_i})_{i \geq 1}$ is a multivariate process (see also Remark 2).

It turns out that the resulting filtering equations are simpler than their counterparts in the case of continuous observations. In the latter case, the nonlinear filters are described by infinite-dimensional stochastic differential equations. For example, if θ_t is a diffusion process, the filtering equations (e.g., Kushner filter or Zakai filter) are given by stochastic partial differential equations (see, e.g., [20]). In contrast, in our setting, the filtering equation can be reduced to a recursive system of linked *deterministic* equations of Kolmogorov type. Therefore, the numerical implementation of the filter is much simpler (see the follow up paper [1]).

We describe the model in Section 2, state the main results and examples in Section 3, provide the proofs in Section 4, and present more detailed examples in Section 5.

2. Mathematical model.

2.1. *Risky asset and observation times.* Let us fix a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ equipped with a filtration $\mathbf{F} = (\mathcal{F}_t)_{t \geq 0}$ that satisfies the "usual" conditions (see, e.g., [16]). All random processes considered in the paper are assumed to be defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and adapted to $\mathbf{F}$.

It is assumed that there is a risky asset with the price process $S = (S_t)_{t \geq 0}$ given by the Itô equation

$$(2.1) \qquad dS_t = m(\theta_t) S_t \, dt + v(\theta_t) S_t \, dB_t,$$

where $B = (B_t)_{t \geq 0}$ is a standard Brownian motion and $\theta = (\theta_t)_{t \geq 0}$ is a cádlág Markov jump-diffusion process in $\mathbb{R}$ with the generator $\mathcal{L}$. To simplify the discussion, it is assumed that $m(x)$ and $v(x)$ are measurable bounded functions on $\mathbb{R}$, the initial condition S_0 is constant, and $v(x)$ and S_0 are positive.

The process $(\theta_t)_{t \geq 0}$ is referred to as the *volatility process*. It is unobservable, and the only observable quantities are the values of the log-price process $X_t = \log S_t$ taken at stopping times $(\tau_k)_{k \geq 0}$, so that $\tau_0 = 0$, $\tau_k < \tau_{k+1}$ if $\tau_k < \infty$, and $\tau_k \uparrow \infty$ as $k \uparrow \infty$.

In accordance with (2.1), the log-price process is given by

$$X_t = \int_0^t \left(m(\theta_s) - \tfrac{1}{2} v^2(\theta_s) \right) ds + \int_0^t v(\theta_s) \, dB_s.$$

For notational convenience, set $X_k := X_{\tau_k}$. Thus, the observations are given by the sequence $(\tau_k, X_k)_{k \geq 0}$.

REMARK 1 (*Note on the reading sequence*). The reader interested primarily in applying our results to real data can focus her attention on Example 3.1,

which appears to be the most practical model to work with. That example provides self-contained formulas for estimating the conditional (filtering) distribution of the volatility process. We report on the numerical results related to this example in the follow-up paper [1].

Clearly, the observation process $(\tau_k, X_k)_{k \geq 0}$ is a multivariate (marked) point process (see, e.g., [8, 13]) with the counting measure

$$\mu(dt, dy) = \sum_{k \geq 1} \mathbf{I}_{\{\tau_k < \infty\}} \delta_{\{\tau_k, X_k\}}(t, y) \, dt \, dy,$$

where $\delta_{\{\tau_k, X_k\}}$ is the Dirac delta-function on $\mathbb{R}_+ \times \mathbb{R}$.

We introduce two filtrations related to $(\tau_k, X_k)_{k \geq 0}$: $(\mathcal{G}(n))_{n \geq 0}$ and $(\mathcal{G}_t)_{t \geq 0}$, where

$$\mathcal{G}(n) := \sigma\{(\tau_k, X_k)_{k \leq n}\},$$

$$\mathcal{G}_t := \sigma\big(\mu([0, r] \times \Gamma) : r \leq t, \Gamma \in \mathcal{B}(\mathbb{R})\big),$$

where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$.

It is a standard fact (see Theorem 31 in Chapter III, Section 3 in [8]) that

$$(2.2) \qquad\qquad \mathcal{G}_{\tau_k} = \mathcal{G}(k), \qquad k = 0, 1, \dots,$$

and $\{\tau_k\}$ is a system of stopping times with respect to $(\mathcal{G}_t)_{t \geq 0}$.

REMARK 2. Although $\mathcal{G}_{\tau_k}$ contains all the relevant information carried by the observations obtained up to time τ_k, the filtration $(\mathcal{G}_t)_{t \geq 0}$ provides additional information between the observation times. To elucidate this point on a more intuitive level, we note that the length of the time elapsed between τ_k and τ_{k+1} carries additional information about the state of θ_t after τ_k. Specifically, if the frequency of observations is proportional to the stock's volatility $v(\theta_t)$, $t \in \,]\!]\tau_k, \tau_{k+1}[\![$, the larger values of $t - \tau_k$ might indicate lower values of $v(\theta_t)$.

2.2. *Volatility process.* A more precise description of the volatility process is in order now. Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $(\mathbb{R}_+ \times \mathbb{R}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}))$ be measurable spaces with Borel σ-algebras. The volatility process $\theta = (\theta_t)_{t \geq 0}$ is defined by the Itô equation

$$(2.3) \quad d\theta_t = b(t, \theta_t) \, dt + \sigma(t, \theta_t) \, dW_t + \int_{\mathbb{R}} u(\theta_{t-}, x)(\mu^\theta - \nu^\theta)(dt, dx),$$

where W_t is a standard Wiener process and $\mu^\theta = \mu^\theta(dt, dx)$ is a Poisson measure on $(\mathbb{R}_+ \times \mathbb{R}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}))$ with the compensator $\nu^\theta(dt, dx) = K(dx) \, dt$, where $K(dx)$ is a σ-finite nonnegative measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We assume that $E\theta_0^2 < \infty$, the functions $b(t, z), \sigma(t, z)$ and $u(z, x)$ are Lipschitz continuous in z uniformly with respect to other variables, and

$$|b(t, z)|^2 + |\sigma(t, z)|^2 + \int_{\mathbb{R}} |u(z, x)|^2 K(dx) \leq C(1 + |z|^2).$$

It is well known that under these assumptions (2.3) possesses a unique strong solution adapted to $\mathbf{F}$, and $E\theta_t^2 < \infty$ for any $t \geq 0$.

The generator $\mathcal{L}$ of the volatility process is given by

$$\mathcal{L}f(x) := b(t, x)f'(x) + \tfrac{1}{2}\sigma^2(t, x)f''(x)$$
$$+ \int_{\mathbb{R}} \big(f(x + u(x, y)) - f(x) - f'(x)u(x, y)\big)K(dy).$$

Before proceeding with the assumptions and main results, we shall introduce additional notation. Set

$$(2.4) \qquad a(s, t) = \int_s^t \big(m(\theta_u) - \tfrac{1}{2}v^2(\theta_u)\big)\,du$$

and

$$(2.5) \qquad \sigma^2(s, t) = \int_s^t v^2(\theta_u)\,du.$$

For simplicity, it is assumed that $v^2(s, t)$ is bounded away from zero. Let us denote by $\rho_{s,t}(y)$ the density function of the normal distribution with mean $a(s, t)$ and the variance $\sigma^2(s, t)$:

$$(2.6) \qquad \rho_{s,t}(y) := \frac{1}{\sqrt{2\pi}\sigma(s, t)}e^{-(y-a(s,t))^2/(2\sigma^2(s,t))}.$$

Clearly, ρ is the conditional density of the stock's log-increments $X_t - X_s$ given θ.

Let $\mathcal{F}_\infty^\theta = (\mathcal{F}_t^\theta)_{t \geq 0}$ be the right-continuous filtration generated by $(\theta_t)_{t \geq 0}$ and augmented by P-zero sets from $\mathcal{F}$. Denote by G_k^θ the conditional distribution of τ_{k+1} with respect to $\mathcal{F}_\infty^\theta \vee \mathcal{G}(k)$ (here and below $\mathcal{F}^1 \vee \mathcal{F}^2$ stands for the σ-algebra generated by the σ-algebras $\mathcal{F}^1$ and $\mathcal{F}^2$). That is, G_k^θ is the distribution of the time of the next observation, given previous history, and given θ,

$$(2.7) \qquad G_k^\theta(dt) = \mathsf{P}\big(\tau_{k+1} \in dt | \mathcal{F}_\infty^\theta \vee \mathcal{G}(k)\big).$$

Without loss of generality, we can and will assume that $G_k^\theta(dt)$ is the regular version of the RHS of (2.7).

Let $N = (N_t)_{t \geq 0}$ be the counting process with interarrival times: $\tau_0 = 0$, $(\tau_k - \tau_{k-1})_{k \geq 1}$, that is,

$$(2.8) \qquad N_t = \sum_{k \geq 1} I(\tau_k \leq t).$$

2.3. *Assumptions.* The following assumptions will be in force throughout the paper:

ASSUMPTION A.0. For every $\mathcal{G}$-predictable and a.s. finite stopping time S,

$$\mathsf{P}(N_S - N_{S-} \neq 0 | \mathcal{G}_{S-}) = 0 \quad \text{or} \quad 1.$$

ASSUMPTION A.1.   The Brownian motion B is independent of (θ, N).

ASSUMPTION A.2.   For every k, there exists a $\mathcal{G}(k)$-measurable integrable random measure Φ_k on $\mathcal{B}(\mathbb{R}_+)$ so that for almost all $\omega \in \Omega$, $\Phi_k([0, \tau_k(\omega)]) = 0$ and G_k^θ is absolutely continuous with respect to Φ_k.

Denote by $\phi(\tau_k, t) = \phi(\theta, \tau_k, t)$ the Radon–Nikodym derivative of $G_k^\theta(dt)$ with respect to $\Phi_k(dt)$, that is, for almost every ω,

$$(2.9) \qquad \phi(\tau_k, t) := \frac{dG_k^\theta((\tau_k, t])}{d\Phi_k((\tau_k, t])}.$$

Assumption A.0 is not essential for the derivation of the filter. However, under this assumption, the structure of the optimal filter is simpler, and in the practical examples important for this paper, this assumption holds anyway. In particular, Assumption A.0 is verified if the conditional distribution $G_k^\theta = \mathsf{P}(\tau_{k+1} \leq t|\mathcal{F}_\infty^\theta \vee \mathcal{G}(k))$ is absolutely continuous with respect to the Lebesgue measure or if the arrival times τ_k are nonrandom (more generally, it holds if the compensator of the counting process N_t is a continuous process).

The following two simple but important examples illustrate Assumption A.2.

EXAMPLE 2.1.   Let $(\tau_k)_{k \geq 0}$ be the jump times of a doubly stochastic Poisson process (Cox process) with the intensity $n(\theta_t)$. In this case,

$$\mathsf{P}(\tau_{k+1} \leq t|\mathcal{F}_\infty^\theta \vee \mathcal{G}(k)) = \begin{cases} 1 - e^{-\int_{\tau_k}^t n(\theta_s)\,ds}, & t \geq \tau_k, \\ 0, & \text{otherwise.} \end{cases}$$

Then, one can take $\Phi_k(ds) = ds$ and $\phi(\tau_k, s) = n(\theta_t)\exp(-\int_{\tau_k}^s n(\theta_u)\,du)$. If $n(\theta_t) = n$ is a constant, one could also choose

$$\Phi_k(ds) = n\exp\{n(\tau_k - s)\}\,ds \quad \text{and} \quad \phi(\tau_k, s) = 1.$$

EXAMPLE 2.2.   If the filtering is based on nonrandom observation times τ_k (e.g., $\tau_k = kh$ where h is a fixed time step), then a natural choice would be $\Phi_k(ds) = \delta_{\{\tau_{k+1}\}}(s)\,ds$ and $\phi(\tau_k, s) = 1$.

For practical purposes, $\Phi_k(ds)$ must be known or easily computable as soon as the observations $(\tau_i, X_i)_{i \leq k}$ become available. In contrast, the Radon–Nikodym density $\phi(\tau_k)$ is, in general, a function of the volatility process and is subject to estimation.

We note that Assumption A.2 could be weakened slightly by replacing G_k^θ by a regular version of the conditional distribution of τ_{k+1} with respect to $\mathcal{F}_{\tau_{k+1}-}^\theta \vee \mathcal{G}(k)$. The latter assumption would make the proof a little bit more involved and we leave it to the interested reader.

3. Main results and introductory examples.

3.1. *Main result.* For a measurable function f on $\mathbb{R}$ with $E|f(\theta_t)| < \infty$, define the conditional expectation estimator $\pi_t(f)$ by

$$(3.1) \qquad \pi_t(f) := E\big(f(\theta_t)|\mathcal{G}_t\big) = \int_{\mathbb{R}} f(z)\pi_t(dz),$$

where $\pi_t(dz) := d\mathsf{P}(\theta_t \leq z|\mathcal{G}_t)$ is the filtering distribution. [Note that we omit the argument θ_t of f in the estimator $\pi_t(f)$.] In the spirit of the Bayesian approach, it is assumed that the a priori distribution

$$\pi_0(dx) = \mathsf{P}(\theta_0 \in dx)$$

is given.

Let $\sigma\{\theta_{\tau_k}\}$ be the σ-algebra generated by θ_{τ_k}. For $t > \tau_k$, let us define the following *structure functions*:

$$(3.2) \qquad \psi_k\big(f; t, y, \theta_{\tau_k}\big) := E\big(f(\theta_t)\rho_{\tau_k, t}(y - X_k)\phi(\tau_k, t)|\sigma\{\theta_{\tau_k}\} \vee \mathcal{G}(k)\big),$$

and its integral with respect to y,

$$
\begin{aligned}
(3.3) \qquad \bar{\psi}_k\big(f; t, \theta_{\tau_k}\big) &:= \int_{\mathbb{R}} \psi_k\big(f; t, y, \theta_{\tau_k}\big) \, dy \\
&= E\big(f(\theta_t)\phi(\tau_k, t)|\sigma\{\theta_{\tau_k}\} \vee \mathcal{G}(k)\big),
\end{aligned}
$$

where ρ and ϕ are given by (2.6) and (2.9), respectively.

If $f \equiv 1$, the argument f in ψ and $\bar{\psi}$ is replaced by 1.

Write

$$\Phi_k(\{\tau_{k+1}\}) := \int_0^\infty I(t = \tau_{k+1})\Phi_k(dt),$$

that is, $\Phi_k(\{\tau_{k+1}\})$ is the jump of $\Phi_k(dt)$ at τ_{k+1}.

Finally, for $t \geq \tau_k$ and a bounded function f, define

$$\mathcal{M}_k(f; t, \pi_t) := \frac{\pi_{\tau_k}(\bar{\psi}_k(f; t)) - \pi_{t-}(f)\pi_{\tau_k}(\bar{\psi}_k(1; t))}{\int_t^\infty \pi_{\tau_k}(\bar{\psi}_k(1; s))\Phi_k(ds)}$$

whenever the numerator is not zero. If the numerator is zero, set $\mathcal{M}_k(f; t, \pi_t)$ to be equal to zero.

The main result of this paper is as follows:

THEOREM 3.1. *Let Assumptions* A.0–A.2 *hold. Then for every measurable bounded function f in the domain of the generator $\mathcal{L}$ such that $\int_0^t E|\mathcal{L}f(\theta_s)| \, ds < \infty$ for any $t \geq 0$, the following system of equations holds*:

(1) *For every $k = 0, 1, \ldots$,*

$$\pi_{\tau_{k+1}}(f) = \left. \frac{\pi_{\tau_k}(\psi_k(f; t, y))}{\pi_{\tau_k}(\psi_k(1; t, y))} \right|_{\left\{ \substack{t = \tau_{k+1} \\ y = X_{k+1}} \right\}}$$

(3.4)

$$- \mathcal{M}_k(f; t, \pi_t)_{|\{t = \tau_{k+1}\}} \Phi(\{\tau_{k+1}\}).$$

(2) *For every $k = 0, 1, \ldots$ and $t \in \,]\!]\tau_k, \tau_{k+1}[\![$,*

(3.5) $$d\pi_t(f) = \pi_t(\mathcal{L}f) \, dt - \mathcal{M}_k(f; t, \pi_t) \Phi_k(dt).$$

3.2. *Remarks.*

1. Equations (3.4) and (3.5) form a closed system of equations for the filter $\pi_t(f)$. It is often convenient and customary (see, e.g., [20, 21] and the references therein) to write a differential equation for a measure-valued process $H_t(dx)$ in its variational form, that is, as the related system of equations for $H_t(f)$ for all f from a sufficiently rich class of test functions belonging to the domain of the operator $\mathcal{L}$. In our setting, such a reduction to the variational form is a necessity, since in some cases the filtering measure $\pi_s(dx) = \mathsf{P}(\theta_s \in dx | \mathcal{G}_s)$ may not belong to the domain of $\mathcal{L}$. However, in the important examples discussed below, there is no need to resort to the variational form. The interested reader who is unaccustomed to the variational approach might benefit from looking first into the examples at the end of this section and in Section 5, where the filtering equations are written as equations for posterior distributions.

2. The system (3.4) simplifies considerably if

(3.6) $$\mathcal{M}_k(f; t, \pi_t)_{|\{t = \tau_{k+1}\}} \Phi(\{\tau_{k+1}\}) = 0 \qquad \text{for all } k.$$

Obviously, (3.6) holds if, for all k, $\Phi_k(dt)$ is continuous at $t = \tau_{k+1}$, as in the case when N_t is a Cox process. In fact, (3.6) holds true in many other interesting cases, even when $\Phi_k(dt)$ has jumps at all τ_{k+1}, as in the case of fixed observation intervals (see Example 5.3 below). We note then that the following *separation principle* holds.

COROLLARY 1.   *Assume* (3.6). *Then the filtering at the observation times* $\{\tau_k\}_{k \geq 1}$ *does not require filtering between them; it is done by the Bayes type recursion:*

(3.7) $$\pi_{\tau_{k+1}}(f) = \left. \frac{\pi_{\tau_k}(\psi_k(f; t, y))}{\pi_{\tau_k}(\psi_k(1; t, y))} \right|_{\left\{ \substack{t = \tau_{k+1} \\ y = X_{k+1}} \right\}}.$$

3. Note that for high-frequency observations, even if condition (3.6) is not met, for all practical purposes, it may suffice to compute the volatility estimates only at the observation times. In that case, one would only use the relatively simple recursion formula (3.4), and disregard equation (3.5).

4. Clearly, the "structure functions" ψ and $\bar{\psi}$ are of paramount importance for computing the posterior distribution of the volatility process. We would like to stress that these do not involve the observations and could be pre-computed "off-line" using just the a priori distribution. Then, "on-line," when the observations become available, one needs only to plug in the obtained measurements (τ_k, X_k), and to compute $\pi_t(f)$ by recursion. This feature is important for developing efficient numerical algorithms.

5. Note also that, for almost every $\omega \in \Omega$, filtering equation (3.5) is a *linear deterministic* equation of Kolmogorov's type, rather than a *nonlinear stochastic* partial differential equation. The latter is typical of the nonlinear filtering of diffusion processes. The well-posedness and the regularity properties of equation (3.5) are well researched in the literature on second-order parabolic deterministic integro-differential equations (see, e.g., [11, 14, 18] and the references therein).

EXAMPLE 3.1 (*Volatility as a Markov chain*). Let us now assume that the counting process is a Cox process with intensity $n(\theta_t)$, and take $\phi(\tau_k, s) = n(\theta_t)e^{-\int_{\tau_k}^{s} n(\theta_u)\,du}$ and $\Phi_k(ds) = ds$. Also assume $\theta = (\theta_t)_{t \leq T}$ is a homogeneous Markov jump process taking values in the finite alphabet $\{a_1, \ldots, a_M\}$ with the intensity matrix $\Lambda = \|\lambda(a_i, a_j)\|$ and the initial distribution $p_q = \mathsf{P}(\theta_0 = a_q)$, $q = 1, \ldots, M$. (This is one of the two models of the state process discussed in [5].) In this case,

$$\mathcal{L}f(\theta_s) = \sum_j \lambda(\theta_s, a_j) f(a_j).$$

Denote by θ_t^j the process θ_t starting from a_j, and

$$p_{ji}(t) := \mathsf{P}(\theta_t = a_i | \theta_0 = a_j), \qquad \pi_j(t) = \mathsf{P}(\theta_t = a_j | \mathcal{G}_t),$$

$$r_{ji}(t, z) := E\big(e^{-\int_0^t n(\theta_u^j)\,du} \rho_{0,t}^j(z) | \theta_t^j = a_i\big),$$

where $\rho_{0,t}^j(z)$ is obtained by substituting θ_s^j for θ_s in $\rho_{0,t}(z)$. It follows from Theorem 3.1 (for details, see Example 5.1), with $f(\theta_t) := I_{\{\theta_t = a_i\}}$, that

$$(3.8) \quad \pi_i(\tau_k) = \frac{n(a_i)\sum_j r_{ji}(\tau_k - \tau_{k-1}, X_k - X_{k-1})p_{ji}(\tau_k - \tau_{k-1})\pi_j(\tau_{k-1})}{\sum_{i,j} n(a_i)r_{ji}(\tau_k - \tau_{k-1}, X_k - X_{k-1})p_{ji}(\tau_k - \tau_{k-1})\pi_j(\tau_{k-1})}.$$

This recursion can be easily computed, once one computes ("off-line") the values r_{ij}. This example is also treated in more detail in Section 5.

4. Proofs. In the proof of the main result we want to show that

$$d\pi_t(f) = \pi_t(\mathcal{L}f)\,dt + dM_t,$$

where M_t is a martingale, and then we find a (integral) martingale representation of M_t with respect to the measure $\mu - \nu$, where ν is a compensator of μ. We first find the compensator.

4.1. $(\mathcal{G}_t)$-*compensator of* μ. Denote by $\mathcal{P}(\mathcal{G})$ the predictable σ-algebra on $\Omega \times [0, \infty)$ with respect to $\mathcal{G}$, and set

$$\widetilde{\mathcal{P}}(\mathcal{G}) = \mathcal{P}(\mathcal{G}) \otimes \mathcal{B}(\mathbb{R}).$$

A nonnegative random measure $\nu(dt, dy)$ on $\widetilde{\mathcal{P}}(\mathcal{G})$ is called a $\widetilde{\mathcal{P}}(\mathcal{G})$-compensator of μ if, for any $\widetilde{\mathcal{P}}(\mathcal{G})$-measurable, nonnegative function $\varphi(t, y) = \varphi(\omega, t, y)$:

(4.1)

$$\text{(i)} \quad \int_0^t \int_{\mathbb{R}} \varphi(s, y) \nu(ds, dy) \text{ is } \mathcal{P}(\mathcal{G})\text{-measurable,}$$

$$\text{(ii)} \quad E \int_0^\infty \int_{\mathbb{R}} \varphi(t, y) \mu(dt, dy) = E \int_0^\infty \int_{\mathbb{R}} \varphi(t, y) \nu(dt, dy).$$

Let $G_k(ds, dx) = G_k(\omega, ds, dx)$ be a regular version of the conditional distribution of (τ_{k+1}, X_{k+1}) given $\mathcal{G}(k)$ (it is assumed that $G_k([0, \tau_k], dx) = 0$):

$$\text{(4.2)} \qquad \mathsf{G}_k(dt, dy) = d\mathsf{P}\big(\tau_{k+1} \le t, X_{k+1} \le y | \mathcal{G}(k)\big).$$

Denote $G_k(dt) = G_k(dt, \mathbb{R})$, that is, $G_k(t) = \mathsf{P}(\tau_{k+1} \le t | \mathcal{G}(k))$ (with probability one).

By Theorem III.1.33 in [8] (see also Proposition 3.4.1 in [16]),

$$\text{(4.3)} \qquad \nu(dt, dy) = \sum_{k \ge 0} I_{]\!]\tau_k, \tau_{k+1}]\!]}(t) \frac{G_k(dt, dy)}{G_k([t, \infty), \mathbb{R})}.$$

We now derive a representation, suitable for the filtering purposes, of the $\widetilde{\mathcal{P}}(\mathcal{G})$-compensator ν in terms of the structure functions (3.2), (3.3) and the posterior distribution of θ.

LEMMA 4.1. *The* $\widetilde{\mathcal{P}}(\mathcal{G})$-*compensator* ν *admits the following version*:

$$\text{(4.4)} \qquad \nu(dt, dy) = \sum_{k \ge 0} I_{]\!]\tau_k, \tau_{k+1}]\!]}(t) \frac{\pi_{\tau_k}(\psi_k(1; t, y))}{\int_t^\infty \pi_{\tau_k}(\bar{\psi}_k(1; s)) \Phi_k(ds)} \Phi_k(dt) \, dy.$$

PROOF. By Assumption A.1, for $t > \tau_k$, with probability 1,

$$\mathsf{P}\big(\tau_{k+1} \le t, X_{k+1} \le y | \mathcal{F}_\infty^\theta \vee \mathcal{G}(k)\big)$$

$$= E\big(\mathsf{P}\big(\tau_{k+1} \le t, X_{k+1} \le y | \mathcal{F}^\theta \vee \mathcal{G}(k) \vee \sigma(\tau_{k+1})\big) | \mathcal{F}_\infty^\theta \vee \mathcal{G}(k)\big)$$

$$\text{(4.5)} \qquad = E\big(I_{(\tau_{k+1} \le t)} \mathsf{P}\big(X_{k+1} \le y | \mathcal{F}_\infty^\theta \vee \mathcal{G}(k) \vee \sigma(\tau_{k+1})\big) | \mathcal{F}_\infty^\theta \vee \mathcal{G}(k)\big)$$

$$= E\bigg(I_{(\tau_{k+1} \le t)} \int_{-\infty}^y \rho_{\tau_k, \tau_{k+1}}(z - X_k) \, dz \Big| \mathcal{F}_\infty^\theta \vee \mathcal{G}(k)\bigg)$$

$$= \int_{\tau_k}^t \int_{-\infty}^y \rho_{\tau_k, s}(z - X_k) \, dz \, G_k^\theta(ds),$$

where we recall that G_k^θ is a regular version of the conditional distribution of τ_{k+1} with respect to $\mathcal{F}_\infty^\theta \vee \mathcal{G}(k)$. Thus, by Assumption A.2, for $t > \tau_k$, with probability 1,

(4.6)
$$\mathsf{P}\big(\tau_{k+1} \le t, X_{k+1} \le y | \mathcal{F}_\infty^\theta \vee \mathcal{G}(k)\big)$$
$$= \int_{\tau_k}^t \int_{-\infty}^y \rho_{\tau_k, s}(z - X_k) \phi(\tau_k, s) \, dz \, \Phi_k(ds).$$

By (3.2), using notation (3.1), we see that

$$E\big(E\big[\phi(\tau_k, s) \rho_{\tau_k, s}(z - X_k) | \sigma\{\theta_{\tau_k}\} \vee \mathcal{G}(k)\big] | \mathcal{G}(k)\big) = \pi_{\tau_k}\big(\psi_k(1; s, z)\big).$$

This, together with (4.6), yields, recalling definition (4.2),

(4.7)
$$G_k(ds, dz) = \pi_{\tau_k}\big(\psi_k(1; s, z)\big) \Phi_k(ds) \, dz.$$

In the same way, for $t > \tau_k$, with probability 1,

(4.8)
$$G_k([t, \infty], \mathbb{R}) = \int_t^\infty \pi_{\tau_k}\big(\bar{\psi}_k(1; s)\big) \Phi_k(ds).$$

This completes the proof. □

REMARK 3. If the right-hand side of (4.8) is zero, then

$$\mathsf{P}\big(\tau_{k+1} \ge t | \mathcal{G}(k)\big) = 0.$$

Hence, $I_{\rrbracket \tau_k, \tau_{k+1} \llbracket}(t) = 0$ with probability 1 and, by the $0/0 = 0$ convention, the corresponding term in (4.4) is zero.

4.2. *Semimartingale representation of the optimal filter.* In this section we will prove the following result.

THEOREM 4.1. *For any bounded function f from the domain of the operator $\mathcal{L}$ such that $\int_0^t E|\mathcal{L} f(\theta_s)| \, ds < \infty$ for all $t < \infty$, the Itô differential of the optimal filter $\pi_s(f)$ is given by equation*

$$d\pi_s(f) = \pi_s(\mathcal{L} f) \, ds$$

(4.9)
$$+ \int_{\mathbb{R}} \left(\sum_{k \ge 0} I_{\rrbracket \tau_k, \tau_{k+1} \llbracket}(s) \frac{\pi_{\tau_k}\big(\psi_k(f; s, y)\big)}{\pi_{\tau_k}\big(\psi_k(1; s, y)\big)} - \pi_{s-}(f) \right)$$

$$\times (\mu - \nu)(ds, dy).$$

PROOF. It suffices to verify the statement for twice continuously differentiable functions f with f, f' f'' bounded. By Itô's formula,

$$f(\theta_t) = f(\theta_0) + \int_0^t \mathcal{L}f(\theta_s)\,ds + \int_0^t f'(\theta_s)\sigma(\theta_s)\,dW_s$$

$$+ \int_0^t \int_R (f(\theta_{s-} + u(\theta_{s-}, x)) - f(\theta_{s-}))(\mu^\theta - \nu^\theta)(ds, dx).$$

Denote

$$L_t = \int_0^t f'(\theta_s)\sigma(\theta_s)\,dW_s$$

$$+ \int_0^t \int_{\mathbb{R}} (f(\theta_{s-} + u(\theta_{s-}, x)) - f(\theta_{s-}))(\mu^\theta - \nu^\theta)(ds, dx).$$

Then, we have

$$\pi_t(f) = E\big(f(\theta_0)|\mathcal{G}_t\big) + E\bigg(\int_0^t \mathcal{L}f(\theta_s)\,ds\Big|\mathcal{G}_t\bigg) + E(L_t|\mathcal{G}_t).$$

Set

$$M_t = \big\{E\big(f(\theta_0)|\mathcal{G}_t\big) - \pi_0(f)\big\}$$

$$+ \bigg\{E\bigg(\int_0^t \mathcal{L}f(\theta_s)\,ds\Big|\mathcal{G}_t\bigg) - \int_0^t \pi_s(\mathcal{L}f)\,ds\bigg\} + E(L_t|\mathcal{G}_t).$$

Obviously, the process $E\big(f(\theta_0)|\mathcal{G}_t\big) - \pi_0(f)$ is a $\mathcal{G}_t$-martingale. Process L_t is a $\mathcal{F}_t$-martingale. Since $\mathcal{G}_t \subseteq \mathcal{F}_t$, for $t > t'$,

$$E\big(E(L_t|\mathcal{G}_t)|\mathcal{G}_{t'}\big) = E\big(E(L_t|\mathcal{F}_{t'})|\mathcal{G}_{t'}\big) = E(L_{t'}|\mathcal{G}_{t'}).$$

Consequently, $E(L_t|\mathcal{G}_t)$ is a martingale too. Finally, $E(\int_0^t \mathcal{L}f(\theta_s)\,ds|\mathcal{G}_t) - \int_0^t \pi_s((\mathcal{L}f))\,ds$ is also a $\mathcal{G}_t$-martingale. Indeed, for $t > s > t'$, we have $E(\pi_s(\mathcal{L}f)|\mathcal{G}_{t'}) = E(\mathcal{L}f(\theta_s)|\mathcal{G}_{t'})$, which yields

$$E\bigg[E\bigg(\int_0^t \mathcal{L}f(\theta_s)\,ds\Big|\mathcal{G}_t\bigg) - \int_0^t \pi_s(\mathcal{L}f)\,ds\Big|\mathcal{G}_{t'}\bigg]$$

$$= E\bigg(\int_0^{t'} \mathcal{L}f(\theta_s)\,ds\Big|\mathcal{G}_{t'}\bigg) - \int_0^{t'} \pi_s(\mathcal{L}f)\,ds.$$

Thus, M_t is a $\mathcal{G}_t$-martingale. In particular, this means that $\pi_t(f)$ is a $\mathcal{G}$-semimartingale with paths in the Skorokhod space $\mathbb{D}_{[0,\infty)}(\mathbb{R})$, so that $\pi_t(f)$ is a right continuous process with limits from the left. By the martingale representation theorem (see, e.g., Theorem 1 and Problem 1.c in Chapter 4, Section 8 in [16]),

$$M_t = \int_0^t \int_{\mathbb{R}} H(s, y)(\mu - \nu)(ds, dy).$$

It is a standard fact that $P(N_S - N_{S-} \neq 0 | \mathcal{G}_{S-}) = \nu(\{S\}, \mathbb{R}_+)$. Hence, due to Assumption A.0, by Theorem 4.10.1 from [16] [see formulae (10.6) and (10.15)],

$$(4.10) \qquad H(t, y) = \mathsf{M}^{\mathsf{P}}_\mu(\triangle M | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y),$$

where $\triangle M_t = M_t - M_{t-}$ and the conditional expectation $\mathsf{M}^{\mathsf{P}}_\mu(g | \widetilde{\mathcal{P}}(\mathcal{G}))$ is defined by the following relation (see, e.g., [16], Chapter 2, Section 2 and Chapter 10, Section 1): for any $\widetilde{\mathcal{P}}(\mathcal{G})$-measurable bounded and compactly supported function $\varphi(t, y)$,

$$E \int_0^\infty \int_{\mathbb{R}} \varphi(t, y) g_t \mu(dt, dy)$$
$$= E \int_0^\infty \int_{\mathbb{R}} \varphi(t, y) \mathsf{M}^{\mathsf{P}}_\mu(g | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y) \nu(dt, dy).$$

By Lemma 4.10.2, [16],

$$(4.11) \qquad \mathsf{M}^P_\mu(\pi_t(f) | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y) = \mathsf{M}^P_\mu(f | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y).$$

Since $\pi_{t-}(f)$ is $\widetilde{\mathcal{P}}(\mathcal{G})$-measurable [which implies $\mathsf{M}^{\mathsf{P}}_\mu(\pi_-(f) | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y) = \pi_{t-}(f)$], by (4.11),

$$\mathsf{M}^{\mathsf{P}}_\mu(\triangle M | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y)$$
$$(4.12) \qquad = \mathsf{M}^{\mathsf{P}}_\mu(\pi_t(f) - \pi_{t-}(f) | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y)$$
$$= \mathsf{M}^{\mathsf{P}}_\mu(f | \widetilde{\mathcal{P}}(\mathcal{G}))(t, y) - \pi_{t-}(f).$$

To complete the proof, one needs to show that

$$(4.13) \qquad \mathsf{M}^{\mathsf{P}}_\mu(f(\theta_\cdot) | \widetilde{\mathcal{P}}(\mathcal{G}))(s, y) = \sum_{k \geq 0} I_{]\!]\tau_k, \tau_{k+1}]\!]}(s) \frac{\pi_{\tau_k}(\psi_k(f; s, y))}{\pi_{\tau_k}(\psi_k(1; s, y))}.$$

To prove (4.13), it suffices to demonstrate that, for any $\widetilde{\mathcal{P}}(\mathcal{G})$-measurable bounded and compactly supported function $\varphi(t, y)$,

$$E \sum_{k \geq 0} \int_{(\tau_k, \tau_{k+1}] \cap (\tau_k, \infty)} \int_{\mathbb{R}} \varphi(t, y) \frac{\pi_{\tau_k}(\psi_k(f; t, y))}{\pi_{\tau_k}(\psi_k(1; t, y))} \nu(dt, dy)$$
$$(4.14)$$
$$= E \int_0^\infty \int_{\mathbb{R}} \varphi(t, y) f(\theta_t) \mu(dt, dy).$$

By monotone class arguments, we can assume that $\varphi(t, x) = v(t) g(x)$, where $v(t)$ is a $\mathcal{P}(\mathcal{G})$-measurable process and $g(x)$ is a continuous function on $\mathbb{R}$. By Lemma III.1.39 in [8], since $v(t)$ is $\mathcal{P}(\mathcal{G})$-measurable, it must be of the form

$$(4.15) \qquad v(t) = v_0 + \sum_{k \geq 1}^\infty v_k(t) I_{]\!]\tau_k, \tau_{k+1}]\!]}(t),$$

where v_0 is a constant and $v_k(t)$ are $\mathcal{G}(k) \otimes \mathcal{B}(\mathbb{R}_+)$-measurable functions.

Owing to (4.15) and Lemma 4.1, in order to prove (4.14), it suffices to verify the equality

(4.16)
$$
E\left[\int_{(\tau_k, \tau_{k+1}] \cap (\tau_k, \infty)} \int_{\mathbb{R}} g(y) v_k(t) \frac{\pi_{\tau_k}(\psi_k(f; t, y))}{\pi_{\tau_k}(\psi_k(1; t, y))} \Phi_k(dt) \, dy\right]
$$
$$
= E\left[v_k(\tau_{k+1}) g(X_{k+1}) f(\theta_{\tau_{k+1}}) \mathbf{1}_{\{\tau_{k+1} < \infty\}}\right].
$$

The next step follows the ideas of Theorem III.1.33 in [8]. We have

$$
E\left[v_k(\tau_{k+1}) g(X_{k+1}) f(\theta_{\tau_{k+1}}) \mathbf{1}_{\{\tau_{k+1} < \infty\}}\right]
$$
$$
= E\left[E\left(v_k(\tau_{k+1}) g(X_{k+1}) f(\theta_{\tau_{k+1}}) \mathbf{1}_{\{\tau_{k+1} < \infty\}} | \mathcal{G}(k) \vee \mathcal{F}_\infty^\theta\right)\right]
$$
$$
= E\left(\int_{(\tau_k, \infty)} \int_{\mathbb{R}} v_k(s) g(y) E[f(\theta_s) G_k^\theta(ds, dy) | \mathcal{G}(k)]\right),
$$

where, as before, $G_k^\theta(ds, dy)$ is the regular version of the conditional distribution of (τ_{k+1}, X_{k+1}) with respect to $\mathcal{F}_\infty^\theta \vee \mathcal{G}(k)$.

By Fubini's theorem, and recalling notation (4.2),

(4.17)
$$
E\left(\int_{(\tau_k, \infty)} \int_{\mathbb{R}} v_k(s) g(y) E[f(\theta_s) G_k^\theta(ds, dy) | \mathcal{G}(k)]\right)
$$
$$
= E\left(\int_{(\tau_k, \infty)} \int_{\mathbb{R}} v_k(s) g(y) \frac{E[f(\theta_s) G_k^\theta(ds, dy) | \mathcal{G}(k)]}{G_k([s, \infty]; \mathbb{R})}\right.
$$
$$
\left. \times \int_{[s, \infty]} G_k(du, \mathbb{R})\right)
$$
$$
= E\left(\int_{\tau_k}^{\tau_{k+1}} \int_{\mathbb{R}} v_k(s) g(y) \frac{E[f(\theta_s) G_k^\theta(ds, dy) | \mathcal{G}(k)]}{G_k([s, \infty]; \mathbb{R})}\right).
$$

By (4.6),

(4.18)
$$
G_k^\theta(ds, dy) = \rho_{\tau_k, s}(z - X_k) \phi(\tau_k, s) \Phi_k(ds) \, dy.
$$

Hence, for $s > \tau_k$,

$$
E[f(\theta_s) G_k^\theta(ds, dy) | \mathcal{G}(k)]
$$
$$
= E\left(E\left(f(\theta_s) \rho_{\tau_k, s}(y - X_k) \phi(\tau_k, s) | \sigma\{\theta_{\tau_k}\} \vee \mathcal{G}(k)\right) | \mathcal{G}(k)\right) \Phi_k(ds) \, dy.
$$
$$
= \pi_{\tau_k}(\psi_k(f; s, y)) \, dy \, \Phi_k(ds).
$$

This, together with (4.8), yields

$$
E\left(\int_{\tau_k}^{\tau_{k+1}} \int_{\mathbb{R}} v_k(s) g(y) \frac{E[f(\theta_s) G_k^\theta(ds, dy) | \mathcal{G}(k)]}{G_k([s, \infty]; \mathbb{R})}\right)
$$
$$
= E\left(\int_{\tau_k}^{\tau_{k+1}} \int_{\mathbb{R}} v_k(s) g(y) \frac{\pi_{\tau_k}(\psi_k(f; s, y)) \, dy}{\int_s^\infty \pi_{\tau_k}(\bar{\psi}(1; t)) \Phi_k(dt)} \Phi_k(ds)\right),
$$

so that (4.16) is satisfied, and the proof follows. $\quad\square$

4.3. *Proof of Theorem* 3.1. In this section we show that Theorem 3.1 follows from Lemma 4.1 and Theorem 4.1.

PROOF. First, we note that the stochastic integral in the RHS of (4.9) can be written as the difference of the integrals with respect to μ and ν. Indeed, since f is bounded, this follows from [8], Proposition II.1.28.

By applying Lemma 4.1 and integrating over y, one gets that, for $t \in \rrbracket\tau_k, \tau_{k+1}\rrbracket$,

$$
\int_{\mathbb{R} \times (\tau_k, t]} \left(\frac{\pi_{\tau_k}(\psi_k(f; s, y))}{\pi_{\tau_k}(\psi_k(1; s, y))} - \pi_{s-}(f) \right) \nu(ds, dy)
$$
$$
= \int_{(\tau_k, t]} \frac{\pi_{\tau_k}(\bar{\psi}_k(f; s)) - \pi_{s-}(f)\pi_{\tau_k}(\bar{\psi}_k(1; s))}{\int_s^\infty \pi_{\tau_k}(\bar{\psi}_k(1; u))\Phi_k(du)} \Phi_k(ds).
$$

This equation verifies that (3.5) follows from the semimartingale representation (4.9), for t between the consecutive observation times.

For the jump part (3.4), we note that

$$
\int_0^t \int_{\mathbb{R}} \pi_{s-}(f)\mu(ds, dy) = \sum_{\tau_{k+1} \leq t} \pi_{(\tau_{k+1})-}(f)
$$

and

$$
\int_0^t \int_{\mathbb{R}} \frac{\pi_{\tau_k}(\psi_k(f; s, y))}{\pi_{\tau_k}(\psi_k(1; s, y))} \mu(ds, dy) = \sum_{\tau_{k+1} \leq t} \frac{\pi_{\tau_k}(\psi_k(f; s, y))}{\pi_{\tau_k}(\psi_k(1; s, y))} \bigg|_{\left\{ \begin{smallmatrix} s=\tau_{k+1} \\ y=X_{k+1} \end{smallmatrix} \right\}}.
$$

Now, (4.9) can be rewritten as follows:

$$
\begin{aligned}
\pi_t(f) &= \pi_0(f) + \int_0^t \pi_s(\mathcal{L}f)\,ds \\
&\quad + \sum_{\tau_{k+1} \leq t} \left(\frac{\pi_{\tau_k}(\psi_k(f; s, y))}{\pi_{\tau_k}(\psi_k(1; s, y))} \bigg|_{\left\{ \begin{smallmatrix} s=\tau_{k+1} \\ y=X_{k+1} \end{smallmatrix} \right\}} - \pi_{(\tau_{k+1})-}(f) \right) \\
&\quad - \sum_{k \geq 0} \int_{(\tau_k, t \wedge \tau_{k+1}]} \mathcal{M}_k(f; s, \pi_s)\Phi_k(ds).
\end{aligned}
$$

(4.19)

Suppose $t \in \rrbracket\tau_k, \tau_{k+1}\llbracket$. Then,

$$
\pi_t(f) = \pi_{\tau_k}(f) + \int_{\tau_k}^t \pi_s(\mathcal{L}f)\,ds - \int_{\tau_k}^t \mathcal{M}_k(f; s, \pi_s)\Phi_k(ds).
$$

It follows that

$$
\begin{aligned}
&\pi_{(\tau_{k+1})-}(f) \\
&= \pi_{\tau_k}(f) + \int_{\tau_k}^{\tau_{k+1}} \pi_s(\mathcal{L}f)\,ds - \int_{\tau_k}^{(\tau_{k+1})-} \mathcal{M}_k(f; s, \pi_s)\Phi_k(ds).
\end{aligned}
$$

Therefore, from (4.19),

$$\pi_{\tau_{k+1}}(f) = \frac{\pi_{\tau_k}(\psi_k(f; s, y))}{\pi_{\tau_k}(\psi_k(1; s, y))}\bigg|_{\left\{\substack{s=\tau_{k+1} \\ y=X_{k+1}}\right\}} - \mathcal{M}_k(f; t, \pi_t)_{|\{t=\tau_{k+1}\}}\Phi(\{\tau_{k+1}\}).$$

This completes the proof. □

5. Examples. In this section we consider some important special cases of Theorem 3.1.

EXAMPLE 5.1 (*Markov chain volatility and Cox process arrivals*). Recall the setting of Example 3.1 and its notation r_{ij}, $\pi_j(t)$ and θ^j. It follows from Example 2.1 that in this case $\Phi_k(\{\tau_{k+1}\}) = 0$ for all k. Hence, the second term in the RHS of (3.4) is zero. By (3.2), for $f(\theta_t) = \mathbf{1}_{\{\theta_t=a_i\}}$ and $t > \tau_k$,

$$\psi_k(f; t, y, \theta_{\tau_k}) = n(a_i)\big[E\big(I_{\{\theta_t=a_i\}}e^{-\int_s^t n(\theta_u)\,du}\rho_{s,t}(y-x)|\theta_s\big)\big]_{\left\{\substack{s=\tau_k \\ x=X_k}\right\}}.$$

Thus, owing to the homogeneity of θ_t, for $t > \tau_k$,

$$\begin{aligned}
&\pi_{\tau_k}\big(\psi_k(f; t, y)\big) \\
&\quad = \sum_j n(a_i)E\big(I_{\{\theta_t=a_i\}}e^{-\int_s^t n(\theta_u)\,du}\rho_{s,t}(y-x)|\theta_s=a_j\big)_{\left\{\substack{s=\tau_k \\ x=X_k}\right\}}\pi_j(\tau_k) \\
&\quad = \sum_j n(a_i)E\big(I_{\{\theta_{t-s}^j=a_i\}}e^{-\int_0^{t-s} n(\theta_u)\,du}\rho_{0,t-s}^j(y-x)\big)_{\left\{\substack{s=\tau_k \\ x=X_k}\right\}}\pi_j(\tau_k) \\
&\quad = \sum_j n(a_i)E\big[I_{\{\theta_{t-s}^j=a_i\}}E\big(e^{-\int_0^{t-s} n(\theta_u)\,du}\rho_{0,t-s}^j(y-x)|\theta_{t-s}^j\big)\big]_{\left\{\substack{s=\tau_k \\ x=X_k}\right\}}\pi_j(\tau_k) \\
&\quad = \sum_j n(a_i)r_{ji}(t-\tau_k, y-X_k)p_{ji}(t-\tau_k)\pi_j(\tau_k).
\end{aligned}$$

Similar formula holds for the denominator of the first term of the RHS of the equation. Now equation (3.8) follows from (3.4).

Mimicking the previous calculations and using the notation

$$\bar{r}_{ji}(t) := E\big(e^{-\int_0^t n(\theta_u^j)\,du}|\theta_t^j = a_i\big),$$

it is readily checked that, for $t > \tau_k$,

$$\pi_{\tau_k}\big(\bar{\psi}_k(\mathbf{1}_{\{\theta_t=a_i\}}; t)\big) = n(a_i)\sum_j \pi_j(\tau_k)\bar{r}_{ji}(t-\tau_k)p_{ji}(t-\tau_k)$$

and

$$\pi_{\tau_k}\big(\bar{\psi}_k(1, t)\big) = \sum_{i,j}\pi_j(\tau_k)n(a_i)\bar{r}_{ji}(t-\tau_k)p_{ji}(t-\tau_k),$$

which are needed in computing (3.5). It is easily verified that, in the setting of this example, equation (3.5) reduces to the following:

(5.1)
$$d\pi_i(t) = \sum_j \lambda(a_j, a_i)\pi_j(t)\, dt$$
$$+ \bar{D}(\tau_k, t)\pi_i(t)\, dt + D_i(\tau_k, t)\, dt,$$

where

$$D_i(\tau_k, t) = -\frac{n(a_i)\sum_j \bar{r}_{ji}(t - \tau_k)p_{ji}(t - \tau_k)\pi_j(\tau_k)}{\int_t^\infty \sum_{i,j} n(a_i)\bar{r}_{ji}(s - \tau_k)p_{ji}(s - \tau_k)\pi_j(\tau_k)\, ds},$$

$$\bar{D}(\tau_k, t) = \frac{\sum_{l,j} n(a_l)\bar{r}_{jl}(t - \tau_k)p_{jl}(t - \tau_k)\pi_j(\tau_k)}{\int_t^\infty \sum_{i,j} n(a_i)\bar{r}_{ji}(s - \tau_k)p_{ji}(s - \tau_k)\pi_j(\tau_k)\, ds}.$$

Note that equation (5.1) is considered for a fixed ω and $t > \tau_k(\omega)$. Therefore, τ_k and $\pi.(\tau_k)$ should be viewed as known quantities.

EXAMPLE 5.2 (*Poisson arrivals*).   Let θ be still the same as in Example 5.1. Suppose that the interarrival times between the observations are exponential with constant intensity $n(\theta) \equiv \lambda$. In other words, N_t is a Poisson process with constant parameter λ. In this case, the volatility process θ is independent of N_t. Then, on the interval $\tau_k < t < \tau_{k+1}$, equation (5.1) reduces to

(5.2)
$$d\pi_i(t) = \sum_j \lambda(a_j, a_i)\pi_j(t)\, dt$$
$$- \lambda\left(\sum_j p_{ji}(t - \tau_k)\pi_j(\tau_k) - \pi_i(t)\right) dt.$$

On the other hand, owing to the independence of N and θ, it is readily checked that on the interval $\tau_k < t < \tau_{k+1}$,

$$\pi_i(t) = \sum_j p_{ji}(t - \tau_k)\pi_j(\tau_k).$$

Therefore, the filtering equation (5.2) is simply the forward Kolmogorov equation for θ.

A similar effect appears also in the following example.

EXAMPLE 5.3 (*Fixed observation intervals*).   Assume for simplicity that the Markov process θ_t is homogeneous. Also assume that $\tau_k = kh$, where h is a fixed

time step. Notice that

$$(5.3) \qquad \mathcal{G}_t = \mathcal{G}(k) \qquad \text{for any } t \in [\![\tau_k, \tau_{k+1}[\![.$$

Denote by $P(t, x, dy)$ the transition probability kernel of the process θ_t, given that $\theta_0 = x$, and let T_t denote the associated transition operator.

In accordance with Example 2.2, one can take

$$\phi(\tau_k, t) \equiv 1 \quad \text{and} \quad \Phi_k(dt) = \delta_{\{\tau_{k+1}\}}(t)\, dt.$$

Thus, we get

$$(5.4) \qquad \psi_k(f; t, y, \theta_{\tau_k}) = E\big[f(\theta_t)\rho_{\tau_k, t}(y - X_k)|\sigma\{\theta_{\tau_k}\} \vee \mathcal{G}(k)\big],$$

$$(5.5) \qquad \bar{\psi}_k(f; t, \theta_{\tau_k}) = T_{t-\tau_k} f(\theta_{\tau_k}) := \int f(y) \mathsf{P}(t - \tau_k, \theta_{\tau_k}, dy).$$

Since $\Phi_k(dt) = 0$ on $[\![\tau_k, \tau_{k+1}[\![$, (3.5) is reduced to the forward Kolmogorov equation

$$\frac{\partial_t}{\partial t} \pi_t(f) = \pi_t(\mathcal{L}f),$$

subject to the initial condition $\pi_{\tau_k}(f)$. The unique solution of this equation is given by $\pi_t(f) = \pi_{\tau_k}(T_{t-\tau_k} f)$, $t < \tau_{k+1}$. Hence,

$$(5.6) \qquad \pi_{\tau_{k+1}-}(f) = \pi_{\tau_k}(T_h f).$$

Since $\phi(\tau_k, t) \equiv 1$, the denominator of $\mathcal{M}_k$ is equal to 1 when $t = \tau_{k+1}$. This together with the formula $\Phi(\{\tau_{k+1}\}) = 1$ yields

$$(5.7) \qquad \mathcal{M}_k(f; t, \pi_t)|_{t=\tau_{k+1}} \Phi(\{\tau_{k+1}\}) = \pi_{\tau_k}(T_h f) - \pi_{\tau_{k+1}-}(f).$$

Owing to (5.7), we get $\mathcal{M}_k(f; t, \pi_t)|_{t=\tau_{k+1}} \Phi(\{\tau_{k+1}\}) = 0$.

This yields the following recursion formula:

$$\pi_{\tau_{k+1}}(f) = \frac{\pi_{\tau_k}(\psi_k(f; t, y))}{\pi_{\tau_k}(\psi_k(1; t, y))}\bigg|_{t=\tau_{k+1}, y=X_{\tau_{k+1}}}$$

$$= \frac{\int_{\mathbb{R}} E(f(\theta_{t-\tau_k})\rho_{0, t-\tau_k}(y - z)|\theta_0 = z)\pi_{\tau_k}(dz)}{\int_{\mathbb{R}} E(\rho_{0, t-\tau_k}(y - z)|\theta_0 = z)\pi_{\tau_k}(dz)}\bigg|_{t=\tau_{k+1}, y=X_{\tau_{k+1}}}.$$

REFERENCES

[1] CVITANIĆ, J., ROZOVSKII, B. and ZALIAPIN, IL. (2006). Numerical estimation of volatility values from discretely observed diffusion data. *Journal of Computational Finance*. To appear.

[2] ELLIOTT, R. J., HUNTER, W. C. and JAMIESON, B. M. (1998). Drift and volatility estimation in discrete time. *J. Econom. Dynam. Control* **22** 209–218. MR1488361

[3] FOUQUE, J.-P., PAPANICOLAOU, G. and SIRCAR, R. (2000). *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge Univ. Press. MR1768877

[4] FREY, R. (1997). Derivative asset analysis in models with level-dependent and stochastic volatility. *CWI Quarterly* **10** 1–34. MR1472800

[5] FREY, R. and RUNGGALDIER, W. (2001). A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *Internat. J. Theoret. Appl. Finance* **4** 199–210. MR1831267

[6] GALLANT, A. R. and TAUCHEN, G. (1998). Reprojecting partially observed systems with application to interest rate diffusions. *J. Amer. Statist. Assoc.* **93** 10–24.

[7] GOURIEROUX, C. (1997). *ARCH Models and Financial Applications*. Springer, New York. MR1439744

[8] JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin. MR0959133

[9] KALLIANPUR, G. and STRIEBEL, C. (1969). Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes. *Teor. Veroyatnost. i Primenen.* **14** 597–622. MR0264780

[10] KALLIANPUR, G. and XIONG, J. (2001). Asset pricing with stochastic volatility. *Appl. Math. Optim.* **43** 47–62. MR1804394

[11] KREIN, S. G. (1982). *Linear Equations in Banach Spaces*. Birkhäuser, Boston. MR0684836

[12] KRYLOV, N. V. and ZATEZALO, A. (2000). Filtering of finite-state time-non homogeneous Markov processes, a direct approach. *Appl. Math. Optim.* **42** 229–258. MR1795610

[13] LAST, G. and BRANDT, A. (1995). *Marked Point Processes on the Real Line*: *A Dynamic Approach*. Springer, New York. MR1353912

[14] LIONS, J.-L. and MAGENES, E. (1968). *Problèmes aux Limites Non Homogènes et Applications*. Dunod, Paris.

[15] LIPTSER, R. S. and SHIRYAEV, A. N. (2000). *Statistics of Random Processes II. Applications*. Springer, New York. MR1800858

[16] LIPTSER, R. S. and SHIRYAYEV, A. N. (1989). *Theory of Martingales*. Kluwer Acad. Publ., Dordrecht. MR1022664

[17] MALLIAVIN, P. and MANCINO, M. E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics* **6** 49–62. MR1885583

[18] MIKULEVICIUS, R. and PRAGARAUSKAS, H. (1992). On the Cauchy problem for certain integro-differential operators in Sobolev and Hölder spaces. *Lithuanian Math. J.* **32** 238–263. MR1246036

[19] ROGERS, L. C. G. and ZANE, O. (1998). Designing and estimating models of high-frequency data. Preprint.

[20] ROZOVSKII, B. L. (1990). *Stochastic Evolution Systems. Linear Theory and Applications to Non-Linear Filtering*. Kluwer Acad. Publ., Dordrecht. MR1135324

[21] ROZOVSKII, B.L. (1991). A simple proof of uniqueness for Kushner and Zakai equations. In *Stochastic Analysis* (E. Mayer-Wolf et al., eds.) 449–458. Academic Press, Boston. MR1119843

[22] RUNGGALDIER, W. J. (2004). Estimation via stochastic filtering in financial market models. In *Mathematics of Finance* (G. Yin and Q. Zhang, eds.) 309–318. Amer. Math. Soc., Providence, RI. MR2076550

J. CVITANIĆ
CALTECH, M/C 228-77
1200 E. CALIFORNIA BLVD.
PASADENA, CALIFORNIA 91125
USA
E-MAIL: cvitanic@hss.caltech.edu

R. LIPTSER
DEPARTMENT OF ELECTRICAL ENGINEERING-SYSTEMS
TEL AVIV UNIVERSITY
69978 TEL AVIV
ISRAEL
E-MAIL: liptser@eng.tau.ac.il

B. ROZOVSKII
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA 90089-1113
USA
E-MAIL: rozovski@math.usc.edu