Statistics and Probability Letters 78 (2008) 3195–3202

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On extending classical filtering equations

Michael A. Kouritzin [a], Hongwei Long [b,*]

[a] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada T6G 2G1

[b] Department of Mathematical Sciences, Florida Atlantic University, Boca Raton, FL 33431, USA

ARTICLE INFO

Article history:
Received 12 March 2007
Received in revised form 19 February 2008
Accepted 2 June 2008
Available online 17 June 2008

MSC:
primary 60G35
secondary 60H15

ABSTRACT

In this paper, we give a direct derivation of the Duncan–Mortensen–Zakai filtering equation, without assuming right continuity of the signal, nor its filtration, and without the usual finite energy condition. As a consequence, the Fujisaki–Kallianpur–Kunita equation is also derived. Our results can be applied to filtering problems in which the signal process has α-stable ($\alpha > 1$) components, and the sensor function is linear.

1. Introduction

Classical continuous-time filtering theory requires that the signal process satisfies the finite energy condition given below in (1). However, some signals like the α-stable processes, or stochastic processes satisfying a large class of stochastic differential equations (SDEs) driven by α-stable processes, will no longer satisfy the moment condition (e.g. the second moment of an α-stable random variable does not exist when $1 < \alpha < 2$). In this paper, we generalize the classical filtering equations, by weakening the finite energy condition. Our new results can be applied to parameter estimation, and filtering for the geophysical models discussed in e.g. Ditlevsen (1999a,b). In these models, climate change is related to an SDE driven by an α-stable process (with $\alpha \approx 1.75$), and inference is done through massive ice core samples corrupted by such things as dating error, ice shifting and melting and refreezing.

Let $\{X_t, t \geq 0\}$ be a measurable Markov process, taking values in a complete separable metric space S, and living on a complete probability space $(\Omega, \mathcal{F}, P)$. The classical filtering problem is to describe the conditional distribution of signal X_t, given the collection $\{Y_s, 0 \leq s \leq t\}$ of distorted, corrupted, partial observations

$$Y_t = \int_0^t h(X_s)\mathrm{d}s + W_t,$$

where $h : S \to \mathbb{R}^d$ is measurable and W is a standard Brownian motion independent of X. A common goal is to derive a stochastic differential equation (SDE) for the conditional distribution. Under the finite energy condition

$$\int_0^T E[|h(X_t)|^2]\mathrm{d}t < \infty, \tag{1}$$

with $|\cdot|$ denoting the Euclidean distance, Fujisaki et al. (1972) obtained a SDE for the conditional expectations $E[f(X_t)|Y_s, 0 \leq s \leq t]$ with f belonging to the domain of the generator of X. We shall refer to this equation as the FKK equation, but it is equally well known as the Kushner-Stratonovich equation. An equivalent, yet simpler, equation is the

* Corresponding author.
E-mail addresses: mkouritz@math.ualberta.ca (M.A. Kouritzin), hlong@fau.edu (H. Long).

0167-7152/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2008.06.005

Duncan–Mortensen–Zakai (DMZ) equation for the unnormalized conditional expectation (see Zakai (1969)). There are two methods to derive the DMZ equation. The first one is to use the Kallianpur–Striebel formula, the FKK equation and Ito's rule (e.g., Szpirglas (1978) and Davis and Marcus (1980)). The second method is to derive the DMZ equation directly under (1) as done by Ocone (1984) (see also references therein).

The finite energy condition (1) has previously been imposed in the derivation of the FKK and DMZ equations. Yet, these equations (see (6) and (7) in Section 2) are well defined only under the condition $E\int_0^T |h(X_s)|\mathrm{d}s < \infty$. Indeed, it follows from Section 3, that a martingale formulation of the filtering problem holds under the condition $\int_0^T |h(X_s)|^2\mathrm{d}s < \infty$ a.s. Therefore, it is natural to ponder (see Remark 3 of Kurtz and Ocone (1988)), whether solutions to these filtering equations exist under conditions weaker than (1). In this paper, we answer this question in the affirmative.

Herein, we derive the DMZ equation without assuming right continuity or the standard finite energy condition. In Section 2, we provide our notation and main results. In Section 3, we consider a martingale problem related to the unnormalized filter and prove Theorem 1. In Section 4, we derive the DMZ equation and FKK equation (Theorem 2). The Appendix contains some technical results used in previous Sections.

2. Notation and main results

We let X and Y be as in the introduction, and define $\mathcal{F}_t^X = \sigma\{X_s, 0 \le s \le t\}$, $\overline{\mathcal{F}}_t^X = \sigma\{\mathcal{F}_t^X, \mathcal{N}\}$, $\overline{\mathcal{F}}_{t+}^X = \cap_{\varepsilon>0}\overline{\mathcal{F}}_{t+\varepsilon}^X$, $\mathcal{F}_t^W = \sigma\{W_s, 0 \le s \le t\}$, $\mathcal{F}_t^Y = \sigma\{\sigma\{Y_s, s \le t\}, \mathcal{N}\}$, where $\mathcal{N}$ is the collection of P-null sets. Let $B(S)$ be the class of $\mathbb{R}$-valued bounded measurable function on S and $P(t, x, \Gamma)$ $(t \ge 0,\ x \in S, \Gamma \in \mathcal{B}(S))$ be the transition function for X. Its transition semigroup T_t, defined by $T_t f(x) = \int_S f(y)P(t, x, \mathrm{d}y)$ for $f \in B(S)$, is generally only measurable. We let $J_0 = \{f \in B(S) : bp - \lim_{t\searrow 0} T_t f = f\}$, where bp-lim stands for bounded pointwise limit, and assume that $(T_t f)(x)$ is a jointly measurable function of (t, x) for all $f \in J_0$. Then, J_0 would contain the continuous bounded functions if X were right continuous, which we do not assume. We define

$$\mathcal{D} = \left\{f \in J_0 :\ \text{there exists } g^f \in J_0 \text{ such that } (T_t f)(x) = f(x) + \int_0^t (T_s g^f)(x)\mathrm{d}s,\ \forall x \in S\right\}$$

g^f is uniquely determined so we let $Lf = g^f$ for $f \in \mathcal{D}$. L is called the *weak generator* of $\{X_t, t \ge 0\}$ and $\mathcal{D}$ is its domain (e.g. Dynkin (1965)). It follows from Kallianpur and Karandikar (1985) that $\mathcal{D}$ is measure-determining class if J_0 is bp-dense in $B(S)$.

To calculate the conditional expectations $\pi_t(f) = E[f(X_t)|\mathcal{F}_t^Y]$, $f \in B(S)$, we fix $T > 0$ and use the conditions on h

$$\int_0^T |h(X_t)|^2\mathrm{d}t < \infty \quad \text{a.s. and} \tag{2}$$

$$\int_0^T E|h(X_t)|\mathrm{d}t < \infty. \tag{3}$$

Compared with the finite energy condition (1), our new conditions (2) and (3) are more general, allowing for instance $h(x) = x$ and X_t to be an α-stable process with $\alpha > 1$. We set $\mathcal{F}_t^{X,Y} \doteq \sigma\{\mathcal{F}_t^X, \mathcal{F}_t^Y\}$, $\mathcal{F}_{t+}^{X,Y} = \cap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}^{X,Y}$, $\mathcal{Y}^t = \bigcap_{\varepsilon>0}\sigma\{\mathcal{F}_{t+\varepsilon}^Y, \mathcal{F}_\infty^X\}$, and

$$M_t(f) = f(X_t) - f(X_0) - \int_0^t Lf(X_s)\mathrm{d}s, \quad \forall f \in \mathcal{D}. \tag{4}$$

It follows by Lemma 6 in the Appendix that $\int_0^t Lf(X_s)\mathrm{d}s \in \overline{\mathcal{F}}_t^X$ and $\{M_t(f), t \ge 0\}$ is an $\{\overline{\mathcal{F}}_t^X\}$-martingale under P. We define the innovation process $\nu_t = Y_t - \int_0^t \pi_s(h)\mathrm{d}s$. Then, by Lemma 2.2 and Remark 2.1 of Fujisaki et al. (1972), we know that $(\nu_t, \mathcal{F}_t^Y, P)$ is a d-dimensional martingale with a continuous version under the condition (3). Now, we define

$$A_t = \exp\left\{\int_0^t \langle h(X_s), \mathrm{d}Y_s\rangle - \frac{1}{2}\int_0^t |h(X_s)|^2\mathrm{d}s\right\} \tag{5}$$

and find that $\{A_t^{-1}, t \in [0, T]\}$ is a $\left\{\mathcal{Y}^t\right\}$-martingale and $\left\{(W_t, \mathcal{Y}^t),\ t \ge 0\right\}$ is a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, P)$. These results are standard in filtering theory, and can be easily derived by using the independence of W and X.

Now, we define a new probability measure via Girsanov's theorem

$$\frac{\mathrm{d}\overline{P}|_{\mathcal{Y}^t}}{\mathrm{d}P|_{\mathcal{Y}^t}} = A_t^{-1} = E\left[A_T^{-1}|\mathcal{Y}^t\right].$$

Then, $\overline{P}\circ(X, Y)^{-1} = P\circ(X, W)^{-1}$, $\{(A_t, \mathcal{Y}^t),\ t \in [0, T]\}$ is a martingale under $\overline{P}$, and $\{Y_t\}$ is a standard Brownian motion independent of $\{X_t\}$ under $\overline{P}$. Moreover, one has the Kallianpur–Striebel formula

$$\pi_t(f) = E[f(X_t)|\mathcal{F}_t^Y] = \frac{\overline{E}[f(X_t)A_t|\mathcal{F}_t^Y]}{\overline{E}[A_t|\mathcal{F}_t^Y]} = \frac{\bar{p}_t^Y(f)}{\bar{p}_t^Y(1)},$$

where $\bar{p}_t^Y(f) \doteq \bar{E}[f(X_t)A_t|\mathcal{F}_t^Y]$ for all $f \in \mathcal{D}$. Hence, the conditional statistics of X_t given $\mathcal{F}_t^Y$, in terms of P, can be calculated from those for the new measure $\overline{P}$. Now, we state our main results.

Theorem 1. *Suppose that h satisfies (2). Then, $\bar{p}_t^Y(f) - \bar{p}_0^Y(f) - \int_0^t \bar{p}_s^Y(Lf)\mathrm{d}s$ is an $\{\mathcal{F}_t^Y\}$-martingale under $\overline{P}$ for each $f \in \mathcal{D}$.*

Theorem 2. *Suppose h satisfies (2) and (3). Then, the DMZ and FKK equations*

$$\bar{p}_t^Y(f) = \bar{p}_0^Y(f) + \int_0^t \bar{p}_s^Y(Lf)\mathrm{d}s + \int_0^t \langle \bar{p}_s^Y(hf), \mathrm{d}Y_s\rangle \quad a.s. \tag{6}$$

$$\pi_t(f) = \pi_0(f) + \int_0^t \pi_s(Lf)\mathrm{d}s + \int_0^t \langle \pi_s(hf) - \pi_s(h)\pi_s(f), d\nu_s\rangle \quad a.s. \tag{7}$$

hold for all $t \in [0, T]$, $f \in \mathcal{D}$.

Remark 1. In Eqs. (6) and (7) as well as their proofs we take $\{\bar{p}_t^Y(hf),\ t \geq 0\}$ and $\{\pi_t(hf),\ t \geq 0\}$ to be the optional projections of $\{\bar{E}[f(X_t)h(X_t)A_t|\mathcal{F}_t^Y],\ t \geq 0\}$ and $\{E[f(X_t)h(X_t)|\mathcal{F}_t^Y],\ t \geq 0\}$, respectively.

Remark 2. When Kurtz and Ocone (1988) considered the uniqueness of solutions to (6) and (7) (in the case S is locally compact), they used the condition (3) and $\int_0^T |\pi_s(h)|^2\mathrm{d}s < \infty$ a.s. which is similar to our condition (2). In Theorem 2, we provide existence of solutions to (6) and (7) under conditions (2) and (3), which complements the results in Kurtz and Ocone (1988).

3. Proof of Theorem 1

We give some preliminary results before proving Theorem 1. Note that the augmented filtration $\{\mathcal{F}_t^{X,Y}\}_{t\geq 0}$ need not be right continuous. By Lemma 7 in the Appendix, $\{A_t, t \geq 0\}$ is indistinguishable from an almost surely continuous $\{\mathcal{F}_t^{X,Y}\}$-progressively measurable process. In the sequel, $\{A_t, t \geq 0\}$ is taken to be this progressively measurable process.

Note that $M_t(f) = f(X_t) - f(X_0) - \int_0^t Lf(X_s)\mathrm{d}s$ is an $\{\overline{\mathcal{F}}_t^X\}$-martingale but not necessarily to be cadlag, and $\{\overline{\mathcal{F}}_t^X\}$ is not necessarily right continuous. We define

$$M_t^+(f) = \lim_{s\searrow t} M_s(f),$$

which makes sense by upcrossing inequality. Then, $M_t^+(f)$ is actually a cadlag $\{\overline{\mathcal{F}}_{t+}^X\}$-martingale (see Meyer (1966)). Let $Z_t(f) = f(X_0) + \int_0^t Lf(X_s)\mathrm{d}s + M_t^+(f)$. We first prove the following result:

Lemma 3. $\bar{p}_t^Y(f) = \overline{E}[f(X_t)A_t|\mathcal{F}_t^Y] = \overline{E}[Z_t(f)A_t|\mathcal{F}_t^Y]$.

Proof. We need only show $\overline{E}[M_t^+(f)A_t|\mathcal{F}_t^Y] = \overline{E}[M_t(f)A_t|\mathcal{F}_t^Y]$. However, one has that

$$\begin{aligned}\overline{E}[(M_{t+s}(f) - M_t(f))A_t|\mathcal{F}_t^Y] &= \overline{E}[\overline{E}[(M_{t+s}(f) - M_t(f))A_t|\mathcal{F}_t^{X,Y}]|\mathcal{F}_t^Y] \\ &= \overline{E}[A_t\overline{E}[(M_{t+s}(f) - M_t(f))|\mathcal{F}_t^{X,Y}]|\mathcal{F}_t^Y] \\ &= 0.\end{aligned}$$

Using dominated convergence theorem and letting $s \to 0$, one has $\overline{E}[M_t^+(f)A_t|\mathcal{F}_t^Y] = \overline{E}[M_t(f)A_t|\mathcal{F}_t^Y]$. □

Lemma 4. *$t \to A_tM_t^+(f)$ is an $\left\{\mathcal{F}_{t+}^{X,Y}\right\}$-martingale under $\overline{P}$.*

Proof. For $\tau < t$, we have that

$$\begin{aligned}\overline{E}\left[A_tM_t^+(f) \mid \mathcal{F}_{\tau+}^{X,Y}\right] &= \overline{E}\left[\overline{E}\left[A_t \mid \mathcal{Y}^\tau\right]M_t^+(f)|\mathcal{F}_{\tau+}^{X,Y}\right] \\ &= A_\tau\overline{E}\left[M_t^+(f) \mid \mathcal{F}_{\tau+}^{X,Y}\right] \\ &= A_\tau M_\tau^+(f) \quad \text{a.s.} \quad \square\end{aligned}$$

We define

$$\mathbb{M}_t(f) \doteq Z_t(f)A_t - f(X_0) - \int_0^t A_sLf(X_s)\mathrm{d}s, \quad \forall f \in \mathcal{D}. \tag{8}$$

Lemma 5. *$\{\mathbb{M}_t(f), t \geq 0\}$ is an $\{\mathcal{F}_{t+}^{X,Y}\}$-martingale under $\overline{P}$.*

Proof. By integration by parts, we have

$$\begin{aligned} Z_t(f)A_t &= M_t^+(f)A_t + f(X_0)A_t + \int_0^t Lf(X_s)\mathrm{d}s \cdot A_t \\ &= M_t^+(f)A_t + f(X_0)A_t + \int_0^t A_s Lf(X_s)\mathrm{d}s + \int_0^t \int_0^u Lf(X_s)\mathrm{d}s\mathrm{d}A_u. \end{aligned}$$

This implies that $\mathbb{M}_t(f)$ is an $\{\mathcal{F}_{t+}^{X,Y}\}$-local martingale. Moreover, by (8), we find

$$\begin{aligned} |\mathbb{M}_t(f)| &\le \left| \left(f(X_0) + \int_0^t Lf(X_s)\mathrm{d}s + M_t^+(f) \right) A_t \right| + |f(X_0)| + \left| \int_0^t A_s Lf(X_s)\mathrm{d}s \right| \\ &\le \left(\sup_{x\in S} |f(x)| + t \sup_{x\in S} |Lf(x)| \right) A_t + |M_t^+(f)A_t| + \sup_{x\in S} |f(x)| + \sup_{x\in S} |Lf(x)| \int_0^t A_s\mathrm{d}s. \end{aligned}$$

So, $\mathbb{M}_t(f)$ is of class DL, hence an $\{\mathcal{F}_{t+}^{X,Y}\}$-martingale, by the fact that A_t and $M_t^+(f)A_t$ are both $\{\mathcal{F}_{t+}^{X,Y}\}$-martingales, hence of class DL. □

Proof of Theorem 1. We take conditional expectations on (8), and use Fubini's theorem as well as independent increments to find

$$\overline{E}[Z_t(f)A_t \mid \mathcal{F}_t^Y] = \overline{E}[f(X_0)] + \int_0^t \overline{E}[A_s Lf(X_s) \mid \mathcal{F}_s^Y]\mathrm{d}s + \mathcal{M}_t(f) \quad \text{a.s.},$$

where $\mathcal{M}_t(f) \doteq \overline{E}\left[\mathbb{M}_t(f) \mid \mathcal{F}_t^Y\right]$ is an $\left\{\mathcal{F}_t^Y\right\}$-martingale by Lemma 5. Hence, by Lemma 3, $\bar{p}_t^Y(f) - \bar{p}_0^Y(f) - \int_0^t \bar{p}_s^Y(Lf)\mathrm{d}s$ is a zero mean $\{\mathcal{F}_t^Y\}$-martingale. □

4. Proof of Theorem 2

To derive the DMZ equation, we identify $\mathcal{M}_t(f)$ as the desired stochastic integral. By martingale representation (e.g., Problem 3.4.16 of Karatzas and Shreve (1988)), we know that $\mathcal{M}_t(f)$ is continuous, and there exists $\mathbb{R}^d$-valued $\{\mathcal{F}_t^Y\}$-progressively measurable process $\{\alpha_t^f, t \ge 0\}$ such that $\mathcal{M}_t(f) = \int_0^t \langle \alpha_s^f, \mathrm{d}Y_s\rangle$ and $\int_0^T |\alpha_t^f|^2\mathrm{d}t < \infty$ a.s. Also, $\mathcal{M}_t(f)$ is the unique $\mathcal{F}_t^Y$-measurable random variable with $\overline{E}[\mathcal{M}_t(f)\xi_t] = \overline{E}[\mathbb{M}_t(f)\xi_t]$ for all bounded $\xi_t \in \mathcal{F}_t^Y$. Without loss of generality, we can take $\xi_t = \overline{E}\left[\xi_T \mid \mathcal{F}_t^Y\right]$ with $\xi_T \in \mathcal{F}_T^Y$. Since $\left\{\mathcal{F}_t^Y\right\}_{t\ge 0}$ is continuous, we have that $t \to \overline{E}\left[\xi_T \mid \mathcal{F}_t^Y\right]$ has a continuous modification (see II.2.9 of Revuz and Yor (1991)), and by almost sure monotonicity of conditional expectation, we can make this modification bounded. Thus, $\mathcal{M}_t(f) = \int_0^t \langle \alpha_s^f, \mathrm{d}Y_s\rangle$ is the $\mathcal{F}_t^Y$-measurable random variable with $\overline{E}[\mathcal{M}_t(f)\xi_t] = \overline{E}[\mathbb{M}_t(f)\xi_t]$ for all continuous $\{\mathcal{F}_t^Y\}$-martingales $\xi_t = \int_0^t \langle \Phi_s, \mathrm{d}Y_s\rangle$ with Φ progressively measurable and

$$c_\xi \doteq \sup_{\omega,t} |\xi_t(\omega)| < \infty.$$

We can also take Φ_s bounded by Lemma 8 in the Appendix.

Now, in order to calculate $\overline{E}[\mathbb{M}_t(f)\xi_t]$, we define the stopping times

$$\sigma_N \doteq \inf\left\{ t > 0 : \left| \int_0^t \langle \alpha_s^f, \mathrm{d}Y_s\rangle \right| > N \right\}.$$

Then, $\int_0^t \overline{E}\left[|\alpha_s^f|^2 1_{s\le\sigma_N}\right]\mathrm{d}s < \infty$ so that $|\alpha_s^f| 1_{s\le\sigma_N} \in \mathcal{L}^2(\overline{P})$ almost everywhere.

Proof of Theorem 2. Since the FKK equation can be easily derived by using Ito's formula, integration by parts and the DMZ equation, we just derive the DMZ equation here. From the proof of Theorem 1, we know that

$$\bar{p}_t^Y(f) - \bar{p}_0^Y(f) - \int_0^t \bar{p}_s^Y(Lf)\mathrm{d}s = \mathcal{M}_t(f), \tag{9}$$

where $\mathcal{M}_t(f) \doteq \overline{E}\left[\mathbb{M}_t(f) \mid \mathcal{F}_t^Y\right] = \int_0^t \langle \alpha_s^f, \mathrm{d}Y_s\rangle$ is an $\left\{\mathcal{F}_t^Y\right\}$-martingale. To derive the DMZ equation, it suffices to prove that $\alpha_t^f = \bar{p}_t^Y(hf)$ a.s.. We set $\xi_t = \int_0^t \langle \Phi_s, \mathrm{d}Y_s\rangle$, let $\{\tau_0^m, \tau_1^m, \ldots, \tau_m^m\}$ be a refining partition of $[0, T]$ and define operator $\Delta_i^m A_t = A_{t\wedge\tau_i^m} - A_{t\wedge\tau_{i-1}^m}$. By Doob's optional sampling theorem, we have

$$\mathcal{M}_{t\wedge\sigma_N}(f) = \overline{E}\left[\overline{E}\left[\mathbb{M}_t(f)|\mathcal{F}_t^Y\right]|\mathcal{F}_{t\wedge\sigma_N}^Y\right] = \overline{E}\left[\mathbb{M}_t(f)|\mathcal{F}_{t\wedge\sigma_N}^Y\right] \quad \text{a.s.}$$

It follows that

$$\overline{E}[\mathcal{M}_{t\wedge\sigma_N}(f)\xi_{t\wedge\sigma_N}] = \overline{E}\left[\mathbb{M}_t(f)\xi_{t\wedge\sigma_N}\right]. \tag{10}$$

By martingale difference technique, the $\{\mathcal{Y}^t\}$-martingale property of $\xi_{t\wedge\sigma_N}$ and Lemma 5, we find that

$$\begin{aligned}
\overline{E}\left[\mathbb{M}_t(f)\xi_{t\wedge\sigma_N}\right] &= \sum_{i=1}^m \overline{E}\left[\Delta_i^m(\mathbb{M}_t(f)\xi_{t\wedge\sigma_N})\right] \\
&= \sum_{i=1}^m \overline{E}[\Delta_i^m\mathbb{M}_t(f)\cdot\xi_{t\wedge\sigma_N\wedge\tau_{i-1}^m}] + \sum_{i=1}^m \overline{E}[\mathbb{M}_{t\wedge\tau_{i-1}^m}(f)\cdot\Delta_i^m\xi_{t\wedge\sigma_N}] + \sum_{i=1}^m \overline{E}[\Delta_i^m\mathbb{M}_t(f)\Delta_i^m\xi_{t\wedge\sigma_N}] \\
&= \sum_{i=1}^m \overline{E}[\Delta_i^m\mathbb{M}_t(f)\Delta_i^m\xi_{t\wedge\sigma_N}].
\end{aligned} \tag{11}$$

Then, using (8), we have for some constant $C_1 > 0$

$$\begin{aligned}
\left|\sum_{i=1}^m \overline{E}\left[\Delta_i^m\mathbb{M}_t(f)\Delta_i^m\xi_{t\wedge\sigma_N}\right] - \sum_{i=1}^m \overline{E}\left[\Delta_i^m(Z_t(f)A_t)\Delta_i^m\xi_{t\wedge\sigma_N}\right]\right| &= \left|\sum_{i=1}^m \overline{E}\int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m} A_sLf(X_s)\Delta_i^m\xi_{t\wedge\sigma_N}\mathrm{d}s\right| \\
&\le C_1\overline{E}\left[\int_0^t A_s\mathrm{d}s\cdot\max_{j\le m}|\Delta_j^m\xi_{t\wedge\sigma_N}|\right],
\end{aligned} \tag{12}$$

which tends to zero when $m\to\infty$ by dominated convergence theorem. Therefore, $\overline{E}\left[\mathbb{M}_t(f)\xi_{t\wedge\sigma_N}\right]$ is the limit of

$$\sum_{i=1}^m \overline{E}[\Delta_i^m A_t\Delta_i^m\xi_{t\wedge\sigma_N}Z_{t\wedge\tau_i^m}(f)] + \sum_{i=1}^m \overline{E}[A_{t\wedge\tau_{i-1}^m}\Delta_i^m\xi_{t\wedge\sigma_N}\Delta_i^m Z_t(f)] \tag{13}$$

as $m\to\infty$. By the $\left\{\mathcal{Y}^t\right\}$-martingale property of $\xi_{t\wedge\sigma_N}$, it follows that $\overline{E}[H\Delta_i^m\xi_{t\wedge\sigma_N}] = 0, \forall H\in\mathcal{L}^1\left(\Omega,\mathcal{Y}^{t\wedge\tau_{i-1}^m},\overline{P}\right)$ so the second term in (13) is zero. For the first term in (13), one uses integration by parts to find that

$$\Delta_i^m A_t\Delta_i^m\xi_{t\wedge\sigma_N} - \int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m} 1_{s\le\sigma_N}\langle h(X_s)A_s,\Phi_s\rangle\mathrm{d}s = \int_{t\wedge\tau_{i-1}^m\wedge\sigma_N}^{t\wedge\tau_i^m\wedge\sigma_N}\langle A_s\Phi_s,\mathrm{d}Y_s\rangle + \int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m}\langle h(X_s)A_s\xi_{s\wedge\sigma_N},\mathrm{d}Y_s\rangle \tag{14}$$

and notes that the right hand side of (14) is a $\{\mathcal{Y}^t\}$-local martingale difference by Proposition 3.2.24 of Karatzas and Shreve (1988) ($h(X_\cdot)$ is $\mathcal{Y}^t$-progressive). Moreover, there is some constant $C_2 > 0$ such that

$$\left|\Delta_i^m A_t\Delta_i^m\xi_{t\wedge\sigma_N} - \int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m} 1_{s\le\sigma_N}\langle h(X_s)A_s,\Phi_s\rangle\mathrm{d}s\right| \le C_2\left(|\Delta_i^m A_t| + \int_0^t A_s|h(X_s)|\mathrm{d}s\right).$$

Thus, if we define

$$\rho_k \doteq \inf\left\{t>0 : \left|\int_0^t\langle A_s\Phi_s,\mathrm{d}Y_s\rangle\right| \vee \left|\int_0^t\langle h(X_s)A_s\xi_{s\wedge\sigma_N},\mathrm{d}Y_s\rangle\right| > k\right\}\wedge T,$$

then $\left\{\Delta_i^m A_{t\wedge\rho_k}\Delta_i^m\xi_{t\wedge\rho_k\wedge\sigma_N} - \int_{t\wedge\rho_k\wedge\tau_{i-1}^m}^{t\wedge\rho_k\wedge\tau_i^m} 1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle\mathrm{d}s\right\}_{k=1}^\infty$ is uniformly integrable by condition (3) and the fact that $\{A_t,\ t\ge 0\}$ is a $\{\mathcal{Y}^t\}$-martingale under $\overline{P}$, hence of class DL. Therefore, by (14), Lemma 9 in the Appendix and the fact that $Z_{t\wedge\tau_i^m}(f)$ is bounded and $\mathcal{Y}^{t\wedge\tau_{i-1}^m}$-measurable, we have

$$\begin{aligned}
&\overline{E}[\Delta_i^m A_t\Delta_i^m\xi_{t\wedge\sigma_N}Z_{t\wedge\tau_i^m}(f)] - \int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m}\overline{E}[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle Z_{t\wedge\tau_i^m}(f)]\mathrm{d}s \\
&= \lim_{k\to\infty}\overline{E}\left[\Delta_i^m A_{t\wedge\rho_k}\Delta_i^m\xi_{t\wedge\rho_k\wedge\sigma_N}Z_{t\wedge\tau_i^m}(f) - \int_{t\wedge\rho_k\wedge\tau_{i-1}^m}^{t\wedge\rho_k\wedge\tau_i^m} 1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle Z_{t\wedge\tau_i^m}(f)\mathrm{d}s\right] \\
&= \lim_{k\to\infty}\overline{E}\left[Z_{t\wedge\tau_i^m}(f)\left(\int_{t\wedge\tau_{i-1}^m\wedge\sigma_N\wedge\rho_k}^{t\wedge\tau_i^m\wedge\sigma_N\wedge\rho_k}\langle A_s\Phi_s,\mathrm{d}Y_s\rangle + \int_{t\wedge\tau_{i-1}^m\wedge\rho_k}^{t\wedge\tau_i^m\wedge\rho_k}\langle h(X_s)A_s\xi_{s\wedge\sigma_N},\mathrm{d}Y_s\rangle\right)\right] \\
&= \lim_{k\to\infty}\overline{E}\left[Z_{t\wedge\tau_i^m}(f)\overline{E}\left[\left(\int_{t\wedge\tau_{i-1}^m\wedge\sigma_N\wedge\rho_k}^{t\wedge\tau_i^m\wedge\sigma_N\wedge\rho_k}\langle A_s\Phi_s,\mathrm{d}Y_s\rangle + \int_{t\wedge\tau_{i-1}^m\wedge\rho_k}^{t\wedge\tau_i^m\wedge\rho_k}\langle h(X_s)A_s\xi_{s\wedge\sigma_N},\mathrm{d}Y_s\rangle\right)\middle|\mathcal{Y}^{t\wedge\tau_{i-1}^m}\right]\right] \\
&= 0.
\end{aligned} \tag{15}$$

By condition (3) and the fact that $1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle\in\sigma\{\overline{\mathcal{F}}_s^X,\mathcal{F}_\infty^Y\}$, we find that

$$\begin{aligned}\overline{E}\left[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle(M^+_{t\wedge\tau_i^m}(f)-M_s(f))\right]&=\overline{E}\left[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle(M^+_{t\wedge\tau_i^m}(f)-M^+_s(f))\right]\\&\quad+\overline{E}\left[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle(M^+_s(f)-M_s(f))\right]\\&=0\quad\text{a.e.}\end{aligned}\tag{16}$$

For any $s\in[t\wedge\tau_{i-1}^m,t\wedge\tau_i^m]$, we have by (4)

$$\begin{aligned}Z_{t\wedge\tau_i^m}(f)&=f(X_0)+\int_0^s Lf(X_u)du+\int_s^{t\wedge\tau_i^m}Lf(X_u)du+M^+_{t\wedge\tau_i^m}(f)\\&=f(X_s)+M^+_{t\wedge\tau_i^m}(f)-M_s(f)+\int_s^{t\wedge\tau_i^m}Lf(X_u)\mathrm{d}u.\end{aligned}$$

Consequently, by (16), the boundedness of Lf, Φ as well as condition (3), we have

$$\begin{aligned}&\lim_{m\to\infty}\sum_{i=1}^m\int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m}\overline{E}[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle Z_{t\wedge\tau_i^m}(f)]\mathrm{d}s=\lim_{m\to\infty}\sum_{i=1}^m\int_{t\wedge\tau_{i-1}^m}^{t\wedge\tau_i^m}\Bigg\{\overline{E}\left[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle f(X_s)\right]\\&\quad+\overline{E}\left[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle(M^+_{t\wedge\tau_i^m}(f)-M_s(f))\right]+\int_s^{t\wedge\tau_i^m}\overline{E}[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle Lf(X_u)]\mathrm{d}u\Bigg\}\,\mathrm{d}s\\&=\int_0^t\overline{E}[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle f(X_s)]\mathrm{d}s.\end{aligned}\tag{17}$$

Therefore, by (10)–(17), it follows that

$$\overline{E}\left[\mathcal{M}_{t\wedge\sigma_N}(f)\xi_{t\wedge\sigma_N}\right]=\int_0^t\overline{E}[1_{s\le\sigma_N}A_s\langle h(X_s),\Phi_s\rangle f(X_s)]\mathrm{d}s.\tag{18}$$

Now, by Ito's isometry property, we note that

$$\overline{E}\left[\mathcal{M}_{t\wedge\sigma_N}(f)\xi_{t\wedge\sigma_N}\right]=\int_0^t\overline{E}[1_{s\le\sigma_N}\langle\alpha_s^f,\Phi_s\rangle]\mathrm{d}s.\tag{19}$$

Combining (18) and (19), we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{E}\left[\mathcal{M}_{t\wedge\sigma_N}(f)\xi_{t\wedge\sigma_N}\right]=\overline{E}\left[1_{t\le\sigma_N}\langle\alpha_t^f,\Phi_t\rangle\right]=\overline{E}\left[1_{t\le\sigma_N}\langle h(X_t)f(X_t)A_t,\Phi_t\rangle\right].$$

By (3), it follows that $\bar{E}[h(X_t)f(X_t)A_t|\mathcal{F}_t^Y]$ exists for a.e. $t\in(0,T]$ and

$$\begin{aligned}\alpha_t^f=\lim_{N\to\infty}1_{t\le\sigma_N}\alpha_t^f&=\lim_{N\to\infty}\overline{E}\left[1_{t\le\sigma_N}h(X_t)f(X_t)A_t|\mathcal{F}_t^Y\right]\\&=\lim_{N\to\infty}1_{t\le\sigma_N}\overline{E}\left[h(X_t)f(X_t)A_t|\mathcal{F}_t^Y\right]\\&=\overline{E}\left[h(X_t)f(X_t)A_t|\mathcal{F}_t^Y\right],\end{aligned}\tag{20}$$

a.s. for a.e. $t\in(0,T]$ since $1_{t\le\sigma_N}$ is $\mathcal{F}_t^Y$-measurable. Thus, we find that $\alpha_t^f=\bar{p}_t^Y(hf)$ a.s. for a.e. $t\in(0,T]$. Hence, by Fubini's theorem, we have that

$$\int_0^T\left|\alpha_t^f-\bar{p}_t^Y(hf)\right|^2\mathrm{d}t=0$$

a.s. for all $T>0$ and

$$\bar{p}_t^Y(f)=\bar{p}_0^Y(f)+\int_0^t\bar{p}_s^Y(Lf)\mathrm{d}s+\int_0^t\langle\bar{p}_s^Y(hf),\mathrm{d}Y_s\rangle\quad\text{a.s.}$$

This completes the proof. □

Acknowledgements

The authors are grateful to the referee for very helpful comments. The first author was supported by an NSERC Discovery Grant. The second author was supported by FAU Start-up funding at the C. E. Schmidt College of Science.

Appendix

In this Appendix, we provide some technical results which have been used previously.

Lemma 6. *$\int_0^t Lf(X_s)\mathrm{d}s \in \overline{\mathcal{F}}_t^X$ and $\{M_t(f), t \geq 0\}$ is an $\{\overline{\mathcal{F}}_t^X\}$-martingale.*

Proof. It follows as in Remark 3.2.11 of Karatzas and Shreve (1988) that $\int_0^\cdot Lf(X_s)\mathrm{d}s$ is indistinguishable from an $\{\overline{\mathcal{F}}_t^X\}$-progressively measurable process and consequently we can redefine $\int_0^t Lf(X_s)\mathrm{d}s$ such that $\int_0^t Lf(X_s)\mathrm{d}s \in \overline{\mathcal{F}}_t^X$ for all $t \geq 0$. Then, it follows from Ethier and Kurtz (1986) p. 162 that

$$E\left[(M_\tau(f) - M_t(f))|\mathcal{F}_t^X\right] = 0 \quad \text{a.s.}, \ \tau \geq t.$$

The lemma follows by noting $\mathcal{G} \doteq \left\{G \in \overline{\mathcal{F}}_t^X : E\left[(M_\tau(f) - M_t(f))1_G\right] = 0\right\}$ is a monotone class. □

Lemma 7. *$\{A_t, t \geq 0\}$ is indistinguishable from an almost surely continuous $\{\mathcal{F}_t^{X,Y}\}$-progressively measurable process.*

Proof. There exist $\mathbb{R}^d$-valued $\{\overline{\mathcal{F}}_t^X\}$-simple processes $\{U^n\}$ with non-random times $\{t_j^n\}$ satisfying $0 = t_0^n < t_1^n < \cdots, \lim_{j\to\infty} t_j^n = \infty$ such that

$$\sup_{t\leq T}\left|\sum_{j:t_j^n\leq t}\langle U_{t_{j-1}^n}^n, Y_{t_j^n} - Y_{t_{j-1}^n}\rangle - \int_0^t \langle h(X_s), \mathrm{d}Y_s\rangle\right| \longrightarrow 0$$

in probability (e.g. Proposition 3.2.26, Remark 3.2.11, the development on p. 146, and the development of Proposition 3.2.6 of Karatzas and Shreve (1988)). Now, we define the $\{\mathcal{Y}^t\}$-cadlag martingales

$$A_t^n = \exp\left\{\sum_{j:t_j^n\leq t}\left[\langle U_{t_{j-1}^n}^n, Y_{t_j^n} - Y_{t_{j-1}^n}\rangle - \frac{1}{2}|U_{t_{j-1}^n}^n|^2(t_j^n - t_{j-1}^n)\right]\right\}.$$

By Doob's inequality, we have for any given $\varepsilon > 0$

$$\overline{P}\left(\sup_{t\leq T}|A_t^n - A_t| \geq \varepsilon\right) \leq \overline{E}\left|A_T^n - A_T\right|/\varepsilon.$$

Now, by Lemma 9 (to follow), $A_T^n \to A_T$ in probability and $\overline{E}|A_T^n| = \overline{E}|A_T| = 1$ imply that $\overline{E}|A_T^n - A_T| \to 0$ as $n \to \infty$. Therefore,

$$\lim_{n\to\infty}\overline{P}\left(\sup_{t\leq T}|A_t^n - A_t| \geq \varepsilon\right) = 0, \quad \forall \varepsilon > 0.$$

Thus, there is a subsequence n_m such that $\sup_{t\leq T}\left|A_t^{n_m} - A_t\right| \longrightarrow 0$ a.s. Moreover, $\{A_t^{n_m}, t \geq 0\}$ is $\{\mathcal{F}_t^{X,Y}\}$-progressively measurable. Now, we redefine $A_t = \limsup_{m\to\infty} A_t^{n_m}$, $\{A_t, t \geq 0\}$ is continuous a.s. and $\{\mathcal{F}_t^{X,Y}\}$-progressively measurable. □

Let $\mathcal{L}_M^\infty = \{$ continuous $\{\mathcal{F}_t^Y\}$ -martingales M on $[0, T]$ with $\sup_{t,\omega}|M_t(\omega)| < \infty\}$ and

$$\bar{\mathcal{L}}_M^\infty = \left\{\int_0^\cdot \langle \beta_s, \mathrm{d}Y_s\rangle : \beta \text{is } \{\mathcal{F}_t^Y\} \text{ -progressively measurable and } \sup_{t\leq T,\omega}\left(|\beta_t(\omega)| \vee \left|\int_0^t \langle \beta_s, \mathrm{d}Y_s\rangle(\omega)\right|\right) < \infty\right\}.$$

Lemma 8. *$\bar{\mathcal{L}}_M^\infty$ is $\mathcal{L}^2$-dense in $\mathcal{L}_M^\infty$.*

Proof. We let $M_t = \int_0^t \langle \beta_s, \mathrm{d}Y_s\rangle$ be in $\mathcal{L}_M^\infty$ and define $\beta_s^n \doteq \beta_s 1_{|\beta_s|\leq n}$. By dominated convergence, we find

$$\overline{E}\left|\int_0^T \langle \beta_s - \beta_s^n, \mathrm{d}Y_s\rangle\right|^2 = \overline{E}\int_0^T |\beta_s - \beta_s^n|^2 \mathrm{d}s \to 0.$$

Now, we define $\lambda^{n,m} \doteq \inf\left\{t > 0 : \left|\int_0^t \langle \beta_s^n, \mathrm{d}Y_s\rangle\right| > m\right\} \wedge T$ and find that $\lambda^{n,m} \nearrow T$ as $m \to \infty$ by continuity. Then,

$$M_t^{n,m} \doteq \int_0^{t\wedge\lambda^{n,m}} \langle \beta_s^n, \mathrm{d}Y_s\rangle \in \bar{\mathcal{L}}_M^\infty$$

and there is a sequence M_T^{n,m_n} that converges to M_T in $\mathcal{L}^2$. □

For the reader's convenience, we state the following basic result (see Theorem 4.5.4 of Chung (1974)) that we have relied upon heavily.

Lemma 9. *Let* $0 < r < \infty$, $V_n \in L^r$ *and* $V_n \to V$ *in probability. Then the following three properties are equivalent:*

(i) $\{|V_n|^r\}$ *is uniformly integrable;*
(ii) $V_n \to V$ *in* L^r*;*
(iii) $E(|V_n|^r) \to E(|V|^r)$.

References

Chung, K.L., 1974. A Course in Probability Theory, 2nd ed. Academic Press, New York.
Davis, M.H.A., Marcus, S.I., 1980. An Introduction to nonlinear filtering. In: Hazewinkel, M., Willems, J.C. (Eds.), Stochastic System: The Mathematics of Filtering and Identification and Application. D. Reidel Publishing Company, pp. 53–75.
Ditlevsen, P.D., 1999a. Observation of α-stable noise induced millennial climate changes from an ice-core record. Geophys. Res. Lett. 26, 1441–1444.
Ditlevsen, P.D., 1999b. Anomalous jumping in a double-well potential. Physical Review E 60, 172–179.
Dynkin, E.B., 1965. Markov Processes, vol. 1. Springer-Verlag, Berlin.
Ethier, S.N., Kurtz, T.G., 1986. Markov Processes, Characterization and Convergence. Wiley, New York.
Fujisaki, M., Kallianpur, G., Kunita, H, 1972. Stochastic differential equations for the nonlinear filtering problem. Osaka J. Math. 9, 19–40.
Kallianpur, G., Karandikar, R.L., 1985. White noise calculus and nonlinear filtering theory. Ann. Probab. 13, 1033–1107.
Karatzas, I., Shreve, S.E., 1988. Brownian Motion and Stochastic Calculus. Springer-Verlag, Berlin.
Kurtz, T.G., Ocone, D.L., 1988. Unique characterization of conditional distributions in nonlinear filtering. Ann. Probab. 16, 80–107.
Meyer, P.A., 1966. Probability and Potentials. Blaisdell, Waltham, MA.
Ocone, D., 1984. Remarks on the finite energy condition in additive white noise filtering. Systems Control Lett. 5, 197–203.
Revuz, D., Yor, M., 1991. Continuous Martingales and Brownian Motion. Springer-Verlag, Berlin.
Szpirglas, J., 1978. Sur l'équivalence d'équations différentielle stochastiques à valeurs mesures intervenant dans le filtrage Markovien non linéaire. Ann. Inst. H. Poincaré Sec. B 14, 33–59.
Zakai, M., 1969. On the optimal filtering of diffusion processes. Z. Wahrsch. Verw. Gebiete 11, 230–243.