The Annals of Probability
1995, Vol. 23, No. 4, 1895–1938

UNIQUENESS AND ROBUSTNESS OF SOLUTION OF MEASURE-VALUED EQUATIONS OF NONLINEAR FILTERING[1]

BY ABHAY G. BHATT, G. KALLIANPUR AND RAJEEVA L. KARANDIKAR

University of North Carolina at Chapel Hill,
University of North Carolina at Chapel Hill,
and Indian Statistical Institute

We consider the Zakai equation for the unnormalized conditional distribution σ when the signal process X takes values in a complete separable metric space E and when h is a continuous, possibly unbounded function on E. It is assumed that X is a Markov process which is characterized via a martingale problem for an operator A_0. Uniqueness of solution for the measure-valued Zakai and Fujisaki–Kallianpur–Kunita equations is proved when the test functions belong to the domain of A_0. It is also shown that the conditional distributions are robust.

1. Introduction. The nonlinear filtering problem can be described as follows: X is the system or signal process which is unobservable. We get information about X by observing the process Y which is a function of X corrupted by noise. The usual model for Y is

$$Y_t = \int_0^t h(X_s)\,ds + W_t, \tag{1.1}$$

where W is assumed to be an $\mathbb{R}^k$-valued Brownian motion and h is a measurable function. The observation σ-field $\mathscr{F}_t^Y = \sigma\{Y_s: 0 \le s \le t\}$ contains all the available information about X_t. The primary aim of filtering theory is to get an estimate of X_t based on the information $\mathscr{F}_t^Y$. This is given by the conditional distribution π_t or, equivalently, the conditional expectations $E[f(X_t)|\mathscr{F}_t^Y]$ for a rich enough class of functions f. This estimate also minimizes the squared error loss and hence π is called the optimal filter.

In this paper the signal process will be assumed to take values in a complete, separable metric space. We will restrict our attention to the case when the signal process X is a Markov process which is characterized via the martingale problem for a certain operator A_0. The operator A_0 can be considered as a restriction to a suitable domain of the infinitesimal generator of X.

Received April 1994; revised February 1995.
[1]Research partially supported by the NSF, by Air Force Office of Scientific Research Grant F49620-92-J-0154 and by Army Research Office Grant DAAL-03-92-G-0008.
AMS 1991 *subject classifications*. Primary 60G35, 62M20, 93E11; secondary 60G44, 60G57, 60H15, 60J35.
Key words and phrases. Nonlinear filtering, Zakai equation, martingale problem, robustness.

It is desirable to get a formula for π_t which is recursive in t. In other words, it would be nice if π were to satisfy a (stochastic) differential equation. It does satisfy an infinite-dimensional stochastic differential equation (SDE) (or when X is a finite-dimensional diffusion, a stochastic partial differential equation). This is widely known as the Kushner or the Fujisaki–Kallianpur–Kunita (FKK) equation. See Kushner (1967), Fujisaki, Kallianpur and Kunita (1972) and Kallianpur (1980). Zakai (1969) obtained an equivalent stochastic partial differential equation for a measure-valued process σ_t which is easier to handle because it is linear in σ_t; σ_t has the property that $\pi_t = \sigma_t / \langle \sigma_t, 1 \rangle$. The process σ_t is thus called the unnormalized conditional distribution of X_t given $\mathscr{F}_t^Y$.

The filtering problem can be said to be completely solved if it can be shown that π_t (or σ_t) is the unique solution of the FKK (respectively, Zakai) equation. This has been done by various authors using essentially two types of techniques. Szpirglas (1978), Pardoux (1982), Baras, Blankenship and Hopkins (1983), Sheu (1983), Chaleyat-Maurel, Michel and Pardoux (1990) and Rozovskii (1991) considered the problem of uniqueness when the signal X is finite dimensional and when A_0 is the full generator of X. This is done essentially using operator techniques.

The other approach is via martingale problems. Hijab (1989) proved that the optimal filter π is the unique solution of a martingale problem. Kurtz and Ocone (1988) proved uniqueness of solution for the FKK and Zakai equations when X is the unique solution corresponding to the martingale problem for A_0. The state space for the signal process is assumed to be a locally compact, separable metric space. They considered unbounded h, but they still required that hf be bounded for every f in the domain of A_0.

In many cases, the signal can be modelled as a solution of a stochastic partial differential equation which in turn can be considered as an infinite-dimensional stochastic differential equation. In such a case the state space is a nonlocally compact, complete separable metric space. In this article we will consider the measure-valued Zakai equation when the state space E is a complete, separable metric space and when h is a continuous, possibly unbounded function and prove uniqueness of solution of the Zakai equation. Equivalent results for the FKK equation will be deduced. We will assume throughout the article that h satisfies

$$E\int_0^T |h(X_s)|^2 \, ds < \infty. \tag{1.2}$$

This guarantees the existence of a solution of the Zakai and FKK equations. For uniqueness we assume an additional integrability condition (3.6). When the state space is a locally compact separable metric space, these results are still an improvement on those of Kurtz and Ocone (1988). For more precise comparison with existing work, see Remark 3.1 at the end of Section 3.

The paper is organized as follows. In Section 2 we give some preliminary results which are used in the course of the article. Section 3 is the main section of this article. We prove pathwise uniqueness of solution of the Zakai equation under a set of conditions on the operator A which corresponds to the process (X, Y) in the sense of martingale problems. We consider the correlated signal and noise case. The important special case of independent signal and noise is summarized in Section 4. In this case the above-mentioned set of conditions can be deduced from similar conditions on the operator A_0.

Although, with practical applications in mind, the noise can be taken to be finite dimensional, the analysis of Sections 3 and 4 can be easily extended to infinite dimensions. This is done in Section 5. One possible application of the result can be when the signal and observation are both given by a stochastic partial differential equation.

In Section 6, we apply the results of the previous sections to signals taking values in a Hilbert space. An example of a pollution process is also considered. In Section 7 we apply our results to a signal which is modelled by a semilinear stochastic differential equation on a Hilbert space. This example was recently considered by Zabczyk (1994). Here we show that our results directly imply the uniqueness of the density-valued solutions of the Zakai equation.

In Section 8, we consider the problem of statistical robustness of the filter in the case of independent signal and noise. We show that if signal processes X^n converge to X in law, then the corresponding unnormalized conditional distributions σ^n converge to σ in law. Further, we show that σ^n and σ can be expressed as functionals on Wiener space and there the convergence is in probability. The proofs of these results depend on the essential use of the Kallianpur–Striebel formula, which enables the required conditional expectation to be evaluated by an explicit integration. Finally, in Section 9, we state the uniqueness and robustness results for the normalized conditional distribution π.

In the remainder of this section we explain the model and the terminology. Let (E, d) be a complete, separable metric space and fix $T > 0$. Let A_0 be an operator with domain $\mathscr{D}(A_0) \subseteq C_b(E)$, the space of bounded continuous functions on E and such that the range of the operator is a subset of $C(E)$. We will assume that the E-valued signal process X is a solution of the martingale problem for A_0. That is, the process

$$M_t^f = f(X_t) - f(X_0) - \int_0^t A_0 f(X_s)\, ds$$

is a martingale for every $f \in \mathscr{D}(A_0)$. Also throughout this article h will be a continuous function satisfying (1.2). The filtering model is as in (1.1). We will assume that the past of X and W is independent of the future increments of W, that is, for every $t > 0$,

$$\sigma\{X_s, W_s\colon 0 \le s \le t\} \text{ and } \sigma\{W_v - W_u\colon t < u < v \le T\} \text{ are independent.} \tag{1.3}$$

We will show later that if (X, Y) is a Markov process arising as a solution to a martingale problem, then under some conditions on the domain of the operator A there exist linear operators $D^i\colon \mathscr{D}(A_0) \to C_b(E \times \mathbb{R}^k)$ such that the cross quadratic variation process $\langle M^f, W^i \rangle$ is given by

$$\langle M^f, W^i \rangle_t = \int_0^t D^i f(X_s, Y_s)\, ds, \qquad 1 \le i \le k. \tag{1.4}$$

The conditions we assume also imply that the process (X_t, Y_t) is a Hunt process and that M^f and W are martingales of class (D). Hence this result also follows from a deep result of Motoo and Watanabe (1965). See also Fujisaki, Kallianpur and Kunita (1972). We will give an independent proof of this fact.

Let $\mathscr{P}(E)$ and $\mathscr{M}_+(E)$ denote the spaces of probability measures and positive finite measures on E, respectively, equipped with the topology of weak convergence; $\mathscr{P}(E)$ is a complete separable metric space under the Prohorov metric. A natural extension of the Prohorov metric to $\mathscr{M}_+(E)$ [see, e.g., Problem 6, Chapter IX of Ethier and Kurtz (1986)] also makes it a complete separable metric space. For $\mu \in \mathscr{M}_+(E)$, $f \in C_b(E)$, we will write

$$\langle \mu, f \rangle = \int_E f(x)\, d\mu(x).$$

For a process (Z_t), we will denote by $\mathscr{F}_t^Z$ the smallest σ-field generated by the null sets in the underlying probability space and the random variables Z_s, $0 \le s \le t$.

For $0 \le t \le T$, define $\pi_t \in \mathscr{P}(E)$ by

$$\langle \pi_t, f \rangle = E\big[f(X_t) | \mathscr{F}_t^Y \big], \qquad \forall f \in C_b(E). \tag{1.5}$$

If $X \in D([0,T], E)$, then there exists a version of π which belongs to $D([0,T], \mathscr{P}(E))$ [see Yor (1977)]. We will use this version without further comment. The process π satisfies the FKK equation

$$\begin{aligned} \langle \pi_t, f \rangle = \langle \pi_0, f \rangle &+ \int_0^t \langle \pi_s, A_0 f \rangle\, ds \\ &+ \int_0^t \sum_{i=1}^k \big(\langle \pi_s, h^i f + D^i f(\cdot, Y_s) \rangle - \langle \pi_s, h^i \rangle \langle \pi_s, f \rangle \big)\, dI_s^i, \\ &\qquad \forall f \in \mathscr{D}(A_0), \end{aligned} \tag{1.6}$$

where I_t defined by

$$I_t^i = Y_t^i - \int_0^t \langle \pi_s, h^i \rangle\, ds \tag{1.7}$$

is the innovation process. Define $\sigma_t \in \mathscr{M}_+(E)$ by

$$\langle \sigma_t, f \rangle = \langle \pi_t, f \rangle \exp\left\{ \int_0^t \sum_{i=1}^k \langle \pi_s, h^i \rangle\, dY_s^i - \tfrac{1}{2} \sum_{i=1}^k \int_0^t \big| \langle \pi_s, h^i \rangle \big|^2\, ds \right\}, \qquad \forall f \in C_b(E). \tag{1.8}$$

Applying Itô's formula to (1.8) and using (1.6) and the fact that I is a Wiener martingale with respect to $(\mathscr{F}_t^Y)$ [see Fujisaki, Kallianpur and Kunita (1972)], we get that $\{\sigma_t\}$ satisfies the Zakai equation

$$\begin{aligned}\langle \sigma_t, f\rangle &= \langle \pi_0, f\rangle + \int_0^t \langle \sigma_s, A_0 f\rangle\, ds \\ &\quad + \int_0^t \sum_{i=1}^k \big\langle \sigma_s, h^i f(\cdot) + D^i f(\cdot, Y_s)\big\rangle\, dY_s^i, \qquad \forall f \in \mathscr{D}(A_0).\end{aligned} \tag{1.9}$$

2. Preliminaries. Suppose L is the generator of (the semigroup corresponding to) a Markov process (U_t). Then the law of U_t is the unique solution to the measure-valued evolution equation

$$\langle \mu_t, f\rangle = \langle \mu_0, f\rangle + \int_0^t \langle \mu_s, Lf\rangle\, ds, \qquad \forall f \in \mathscr{D}(L), \tag{2.1}$$

where μ_0 is the law of U_0. Now an interesting question is, does uniqueness still hold if the *test functions* f in (2.1) are taken from a subdomain $\mathscr{D}_0$ of $\mathscr{D}(L)$? In this section we will show that the answer is affirmative if the martingale problem for B is well posed, where B is the restriction of L to $\mathscr{D}_0$, and if B satisfies some additional conditions. This result is deduced from an analogous result in Bhatt and Karandikar (1993b) and is given here in a form that is suitable for our purpose, namely, in deducing uniqueness for the Zakai equation.

We begin with some notation and definitions. Let S be a complete, separable metric space, and let B be an operator on $C(S)$ with domain $\mathscr{D}(B) \subset C_b(S)$. Let the bp-closure of a set V be the smallest set containing V which is closed under bounded pointwise (bp) convergence of sequences. We will denote this set by bp-closure(V). Suppose that B satisfies the following conditions.

C1. There exists $\Theta \in C(S)$, satisfying

$$|Bf(x)| \le C_f \Theta(x), \qquad \forall f \in \mathscr{D}(B),\ x \in S.$$

C2. There exists a countable subset $\{f_n\} \subset \mathscr{D}(B)$ such that

$$\text{bp-closure}\big(\{(f_n, \Theta^{-1}Bf_n) \colon n \ge 1\}\big) \supset \{(f, \Theta^{-1}Bf) \colon f \in \mathscr{D}(B)\}.$$

C3. $\mathscr{D}(B)$ is an algebra that separates points in S and contains the constant functions.

DEFINITION 2.1. A (S-valued) process $(Z_t)_{0 \le t \le T}$ defined on some probability space $(\Omega, \mathscr{F}, P)$ is said to be a solution to the martingale problem for (B, μ) if:

(i) $P \circ Z_0^{-1} = \mu$;
(ii) $\int_0^t E\Theta(Z_s)\, ds < \infty$, for every $t \le T$;
(iii) for all $f \in \mathscr{D}(B)$, $f(Z_t) - \int_0^t Bf(Z_s)\, ds$ is a martingale.

When $T = \infty$, in (ii) above, $t \leq T$ is to be replaced by $t < \infty$. We will say that the martingale problem for (B, μ) is well posed if there exists a solution Z to the martingale problem for (B, μ), and if Z_1 and Z_2 are any two solutions, then $\mathscr{L}(Z_1) = \mathscr{L}(Z_2)$. Here and in what follows, $\mathscr{L}(Z)$ denotes the law of Z, where Z could be a process or a random variable.

We will say that the $D([0,T], S)$ martingale problem for (B, μ) is well posed if there exists an r.c.l.l. solution $(Z_t)_{0 \leq t \leq T}$ to the martingale problem and for any two solutions with r.c.l.l. paths, their laws are the same. We also assume the following conditions.

C4. The $D([0,T], S)$ martingale problem for (B, δ_z) is well posed for every $z \in S$.

C5. For all $\mu \in \mathscr{P}(S)$, any progressively measurable solution to the martingale problem for (B, μ) admits a cadlag modification.

Note that conditions C4 and C5 together imply that the martingale problem for (B, δ_z) is well posed in the class of progressively measurable solutions for every $z \in S$. Also conditions C2 and C4 together imply that the solution Z is a strong Markov process [see Theorems IV.4.2 and IV.4.6 of Ethier and Kurtz (1986) and Remark 2.1 in Horowitz and Karandikar (1990)].

Let B satisfy conditions C1–C4 and, for $z \in S$, let $P_z \in \mathscr{P}(D([0,T], S))$ be the law of the solution to the $D([0,T], S)$ martingale problem for (B, δ_z). Let ζ_t denote the coordinate random variables on $D([0,T], S)$ and let

$$(2.2)\quad \mathscr{M}_+(B, \Theta) = \left\{ \mu \in \mathscr{P}(S) : \int_S \left(\int_{D([0,T],S)} \int_0^T \Theta(\zeta_t)\, dt\, dP_z \right) d\mu < \infty \right\}.$$

It can be proved that the $D([0,T], S)$ martingale problem for (B, μ) is well posed if and only if $\mu \in \mathscr{M}_+(B, \Theta)$.

The following version of a result in Bhatt and Karandikar (1993b) on uniqueness of solution to a measure-valued evolution equation will be required later on.

THEOREM 2.1. *Suppose B satisfies conditions* C1–C5. *If* $\{\mu_t^i\} \subset \mathscr{P}(S)$, $i = 1, 2$, *satisfy*

$$(2.3)\quad \text{for every Borel set } U \subset S,\ t \mapsto \mu_t^i(U) \text{ is measurable}$$

and

$$(2.4)\quad \int_0^T \langle \Theta, \mu_s^i \rangle\, ds < \infty$$

and if, for every $0 \leq t \leq T$, $\{\mu_t^i\}$ *is a solution to the evolution equation*

$$(2.5)\quad \langle \mu_t^i, f \rangle = \langle \mu_0, f \rangle + \int_0^t \langle \mu_s^i, Bf \rangle\, ds, \qquad \forall f \in \mathscr{D}(B),$$

then $\mu_t^1 = \mu_t^2$, *for all* $0 \leq t \leq T$.

PROOF. Define the operator $\tilde{B}$ on $C(S_0)$, where $S_0 = [0,T] \times S$, by

$$\mathscr{D}(\tilde{B}) = \{f \otimes g: f \in \mathscr{D}(B), g \in C^1([0,T])\},$$
$$\tilde{B}(f \otimes g) = \left(\frac{\partial}{\partial t} + B\right)(f \otimes g).$$

Then

$$|\tilde{B}F(t,z)| \le C_F(1 + \Theta(z)), \qquad \forall F \in \mathscr{D}(\tilde{B}), (t,z) \in S_0.$$

It is well known that $\tilde{Z}_t = (t, Z_t)$ is a solution to the martingale problem for $(\tilde{B}, \delta_0 \otimes \mu_0)$. We make the following (deterministic) time change. Let

$$\Gamma(t) = T\big(1 - \exp\big(-\{\sqrt{t+1} - 1\}\big)\big), \qquad t \ge 0.$$

Then $\Gamma'(t) = T/(2\sqrt{t+1})\exp(-\{\sqrt{t+1} - 1\})$. Hence Γ' is bounded. Note that Γ is a strictly increasing function of t.

Let $Y_t = (\Gamma(t), Z_{\Gamma(t)})$ and $DF(t,z) = \Gamma'(t)\tilde{B}F(t,z)$, for all $F \in \mathscr{D}(\tilde{B})$. Then $(Y_t: 0 \le t < \infty)$ is a solution to the martingale problem for $(D, \delta_0 \otimes \mu_0)$. It is also easy to see that D satisfies condition C1 with $\Theta'(t,z) = 1 + \Theta(z)$ and conditions C2, C4 and C5.

Let $\tilde{\mu}_t^i := \delta_t \otimes \mu_t^i$ and let $\nu_t^i = \tilde{\mu}_{\Gamma(t)}^i$. Then ν_t^i satisfies (2.3) and, for $0 \le t < \infty$,

$$\langle F, \nu_t^i\rangle = \langle F, \nu_0^i\rangle + \int_0^t \langle DF, \nu_s^i\rangle\, ds, \qquad \forall F \in \mathscr{D}(D).$$

Further, for $\beta > 0$,

$$\int_0^\infty e^{-\beta t}\langle 1 + \Theta, \nu_t^i\rangle\, dt = \int_0^\infty e^{-\beta t}\frac{1}{\Gamma'(t)}\langle 1 + \Theta, \tilde{\mu}_{\Gamma(t)}^i\rangle \Gamma'(t)\, dt.$$

Note that $e^{-\beta t}/(\Gamma'(t)) = (2/T)\sqrt{t+1}\exp(-\{\beta t + \sqrt{t+1} - 1\}) \to 0$ as $t \to \infty$. Hence $e^{-\beta t}/(\Gamma'(t))$ is bounded. Therefore,

$$\begin{aligned}\int_0^\infty e^{-\beta t}\langle 1 + \Theta, \nu_t^i\rangle\, dt &\le C\int_0^\infty \langle 1 + \Theta, \tilde{\mu}_{\Gamma(t)}^i\rangle \Gamma'(t)\, dt \\ &= C\int_0^T \langle 1 + \Theta, \tilde{\mu}_t^i\rangle\, dt < \infty\end{aligned}$$

by (2.4). Thus D and ν^i satisfy all the conditions of Theorem 3.1 of Bhatt and Karandikar (1993b) and as a consequence of that result, $\nu_t^1 = \nu_t^2$. This in turn implies that $\mu_t^1 = \mu_t^2$, for all $0 \le t \le T$. □

We will now prove a result on Markov processes arising as solutions to a martingale problem. This identifies the cross quadratic variation of some processes related to the Markov process.

Now suppose $S = S_0 \times \mathbb{R}^k$. For functions $f \in C(S_0)$ and $g \in C(\mathbb{R}^k)$ we will denote by $f \otimes g$ the function on S defined by

$$f \otimes g(x, z) = f(x)g(z), \qquad x \in S_0,\ z \in \mathbb{R}. \tag{2.6}$$

Let us suppose that

$$1 \otimes g \in \mathscr{D}(B), \qquad \forall g \in C_b^2(\mathbb{R}^k) \tag{2.7}$$

and that the following condition holds.

C6. If $Z_t = (\xi_t, \eta_t)$ is any r.c.l.l. solution to the martingale problem for B, then η has continuous paths.

THEOREM 2.2. *Suppose B satisfies* C1–C4 *and* C6. *Let $\mathscr{D}_0 \subset C_b(S_0)$ be such that*

$$f \otimes g \in \mathscr{D}(B), \qquad \forall f \in \mathscr{D}_0,\ g \in C_b^2(\mathbb{R}^k). \tag{2.8}$$

Then for $1 \le i \le k$, there exist mappings C^i: $\mathscr{D}_0 \mapsto C(S)$ such that if $Z_t = (\xi_t, \eta_t)$ is any r.c.l.l. solution to the martingale problem for B, then η_t is a semimartingale and

$$\left[f(\xi_.), \eta_.^i \right]_t = \int_0^t C^i f(\xi_s, \eta_s)\, ds. \tag{2.9}$$

Here $[f(\xi_.), \eta_.^i]$ denotes the cross quadratic variation between the semimartingales $f(\xi_.)$ and $\eta_.^i$.

REMARK 2.1. As we noted earlier, conditions C2 and C4 together imply that the process (ξ, η) is a strong Markov process. Now conditions C1 and C3 imply that (ξ, η) is quasi-left-continuous. See Theorem IV.3.12 of Ethier and Kurtz (1986) for a proof of this fact. Hence we get that under conditions C1–C4 the process (ξ, η) is a Hunt process. Further, the semimartingales $f(\xi_.)$ and $\eta_.$ are of class (DL). Hence the conclusion of the theorem will follow from some deep results on additive functionals of Markov processes proved in Motoo and Watanabe (1965). In Fujisaki, Kallianpur and Kunita (1972) the FKK equation (1.6) for the case of Markov processes was derived under the assumption that the signal and observation processes are Hunt processes and that Meyer's hypothesis (DL) is satisfied.

We are going to use (2.9) in the next section in the context of filtering theory where we have the special form of the operator as in (2.8) above and where we also know that condition C6 is satisfied. In the presence of these two extra conditions the existence of the operators C^i as in (2.9) can be proved independently as follows. The proof is also much simpler than that of Motoo and Watanabe (1965).

PROOF OF THEOREM 2.2. For $f \in \mathscr{D}_0$, $g \in C_b^2(\mathbb{R}^k)$, let $\Lambda(f, g)$ be defined by

$$\Lambda(f,g) = B(f \otimes g) - (f \otimes 1)(B(1 \otimes g)) - (B(f \otimes 1))(1 \otimes g).$$

Let us note that if (ξ_t, η_t) is any r.c.l.l. solution to the martingale problem for B and if $M_t^f = f(\xi_t) - \int_0^t B(f \otimes 1)(\xi_s, \eta_s)\, ds$ and $N_t^g = g(\eta_t) - \int_0^t B(1 \otimes g)(\xi_s, \eta_s)\, ds$, then M^f, N^g are martingales and

$$M_t^f N_t^g - \int_0^t \Lambda(f,g)(\xi_s, \eta_s)\, ds$$

is a martingale and hence $\langle M^f, N^g \rangle_t = \int_0^t \Lambda(f, g)(\xi_s, \eta_s)\, ds$. Since N^g is a continuous martingale, it follows that $[M^f, N^g]_t = \langle M^f, N^g \rangle_t$ and hence that

$$[f(\xi.), g(\eta.)]_t = [M^f, N^g]_t = \int_0^t \Lambda(f,g)(\xi_s, \eta_s)\, ds. \tag{2.10}$$

Fix $1 \le i \le k$. For $m \ge 1$, let $g^m \in C_b^2(\mathbb{R}^k)$ be such that $g^m(u) = u^i$ for $u = (u^1, u^2, \dots, u^m) \in \mathbb{R}^k$, $|u^i| \le (m+1)$.

We will prove that

$$\Lambda(f, g^m)(x,u) = \Lambda(f, g^{m+1})(x,u), \qquad \forall\, x \in S_0,\ -m \le u^i \le m, \tag{2.11}$$

and the required functional C could then be defined by

$$C^i f(x,u) = \Lambda(f, g^m)(x,u), \qquad \forall\, x \in S_0,\ -m \le u^i \le m. \tag{2.12}$$

In view of (2.11), it would follow that Cf is well defined and is continuous. Fix $x \in S_0$, $u \in \mathbb{R}^k$ and let $|u^i| \le m$. Let (ξ_t, η_t) be a r.c.l.l. solution to the martingale problem for $(B, \delta_{(x,u)})$.

Let $\tau = \inf\{t \ge 0: |\eta_t^i| \ge (m+1)\}$. The continuity of η implies that $\tau > 0$. Now clearly

$$[f(\xi.), g^m(\eta.)]_{(t \wedge \tau)} = [f(\xi.), g^{m+1}(\eta.)]_{(t \wedge \tau)}$$

and hence

$$\int_0^{t \wedge \tau} \Lambda(f, g^m)(\xi_s, \eta_s)\, ds = \int_0^{t \wedge \tau} \Lambda(f, g^{m+1})(\xi_s, \eta_s)\, ds.$$

Now dividing by $t \wedge \tau$, taking the limit as $t \to 0$ and using the fact that paths of (ξ, η) are right continuous with $\xi_0 = x$, $\eta_0 = u$, it follows that

$$\Lambda(f, g^m)(x,u) = \Lambda(f, g^{m+1})(x,u)$$

and thus, as noted above, Cf is well defined. Now $g^m(\eta)$ is a semimartingale for all m and hence it follows that η is a semimartingale. In addition, if (ξ, η) is a solution to the martingale problem for B and if $\tau^m = \inf\{t \ge 0: |\eta_t| \ge (m+1)\}$, then arguing as above it follows that

$$\begin{aligned}
[f(\xi.), \eta.]_{t \wedge \tau^m} &= [f(\xi.), g^m(\eta.)]_{t \wedge \tau^m} \\
&= \int_0^{t \wedge \tau^m} \Lambda(f, g^m)(\xi_s, \eta_s)\, ds \\
&= \int_0^{t \wedge \tau^m} C^i f(\xi_s, \eta_s)\, ds.
\end{aligned}$$

Since this holds for every $m \ge 1$, this completes the proof. □

3. Uniqueness of solution of the Zakai equation. Let $(\Omega, \mathscr{F}, P)$ be a given probability space on which are defined processes X and W satisfying (1.3) and where W is a k-dimensional standard Brownian motion. The observation model is given by (1.1); h is assumed to be a continuous function satisfying (1.2), and the signal process X is assumed to be a Markov process which is characterized as the unique solution to the martingale problem for A_0. Suppose that the operator A_0 satisfies conditions C1–C5. (See Section 6 for examples of A_0 satisfying these conditions.) Suppose that the $(E \times \mathbb{R}^k)$-valued process (X, Y) is a Markov process arising as the unique solution to the martingale problem for an operator A with domain $\mathscr{D}(A)$ consisting of finite linear combinations of functions of the form $f \otimes g$, where $f \in \mathscr{D}(A_0)$ and $g \in C_b^\infty(\mathbb{R}^k)$. Suppose that the martingale problem for A is well posed. Then it follows that A satisfies the conditions of Theorem 2.2 and as a consequence there exist mappings D^i: $\mathscr{D}(A_0) \to C(E \times \mathbb{R}^k)$ such that the cross quadratic variation process $\langle M^f, W^i \rangle$ is given by

$$\langle M^f, W^i \rangle_t = \int_0^t D^i f(X_s, Y_s)\, ds, \qquad 1 \le i \le k, \tag{3.1}$$

where

$$M_t^f = f(X_t) - f(X_0) - \int_0^t A_0 f(X_s)\, ds.$$

See also Remark 2.1. It can now be deduced that the operator A is given by

$$A(f \otimes g) = A_0 f \otimes g + \tfrac{1}{2} f \otimes \Delta g + \sum_{i=1}^{k} (D^i f)(1 \otimes g^i) + h^i f \otimes g^i, \tag{3.2}$$

where g^i denotes the partial derivative of g w.r.t. the ith component. Let us assume that

$$|D^i f(x, y)| \le C_f \Theta(x), \quad \forall\, f \in \mathscr{D}(A_0), 1 \le i \le k, \forall\, x \in E,\, y \in \mathbb{R}^k. \tag{3.3}$$

Then A satisfies condition C1 with Θ' in place of Θ, where

$$\Theta'(x, y) = 1 + \Theta(x) + \max_{1 \le i \le k} |h^i(x)|. \tag{3.4}$$

Further, if π_0 is the law of X_0, then assuming

$$\pi_0 \otimes \delta_0 \in \mathscr{M}_+(A, \Theta') \tag{3.5}$$

is equivalent to assuming that a solution X to the $D([0, T], E)$ martingale problem for (A_0, π_0) exists. Hence note that (1.2) and (3.5) imply that a solution to the Zakai equation exists, namely, σ_t. The next result is on uniqueness.

THEOREM 3.1. *Suppose that the operator A defined by* (3.2) *satisfies conditions* C2–C5 *and that it satisfies* C1 *with* Θ' *in place of* Θ. *Suppose* π_0 *satisfies* (1.2) *and* (3.5). *If* $\{\rho_t\}$ *is an* $\mathscr{F}_t^Y$*-adapted* $\mathscr{M}_+(E)$*-valued cadlag process which is a solution to the Zakai equation*

$$\langle \rho_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho_s, A_0 f\rangle\, ds + \int_0^t \sum_{i=1}^k \big\langle \rho_s, h^i f(\cdot) + D^i f(\cdot, Y_s)\big\rangle\, dY_s^i, \quad \forall f \in \mathscr{D}(A_0),$$

and which satisfies

$$E\int_0^T \langle \rho_t, \Theta'\rangle \exp\left\{ -\sum_{i=1}^k \int_0^t \langle \pi_s, h^i\rangle\, dI_s^i - \frac{1}{2}\sum_{i=1}^k \int_0^t |\langle \pi_s, h^i\rangle|^2\, ds\right\} dt < \infty, \tag{3.6}$$

then $\rho_t = \sigma_t$, *for all* $t \le T$ *a.s., where* σ_t *is defined by* (1.8).

PROOF. Let $R_t = \exp\{\sum_{i=1}^k \int_0^t \langle \pi_s, h^i\rangle\, dY_s^i - \frac{1}{2}\sum_{i=1}^k \int_0^t |\langle \pi, h^i\rangle|^2\, ds\}$. Define $\{\mu_t\} \in \mathscr{M}_+(E)$ by

$$\langle \mu_t, f\rangle = \frac{\langle \rho_t, f\rangle}{R_t}, \qquad \forall f \in C_b(E).$$

Note that $\langle \sigma_t, f\rangle / R_t = \langle \pi_t, f\rangle$ and hence σ satisfies (3.6).

Now using (1.7) and Itô's formula we get

$$\langle \mu_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \mu_s, A_0 f\rangle\, ds + \sum_{i=1}^k \int_0^t \Big(\big\langle \mu_s, h^i f(\cdot) + D^i f(\cdot, Y_s)\big\rangle - \langle \pi_s, h^i\rangle \langle \mu_s, f\rangle\Big)\, dI_s^i.$$

We will prove that $\mu_t = \pi_t$ for $0 \le t \le T$ a.s. This will imply that $\rho_t = \sigma_t$, for all $0 \le t \le T$ a.s. Now for $f \otimes g \in \mathscr{D}(A)$ an application of Itô's formula shows that

$$\begin{aligned} \langle \mu_t, f\rangle g(Y_t) = {}& \langle \pi_0, f\rangle g(0) + \int_0^t \big\langle \mu_s, A(f\otimes g)(\cdot, Y_s)\big\rangle\, ds \\ & + \sum_{i=1}^k \int_0^t \Big(\langle \mu_s, f\rangle g^i(Y_s) + \Big(\big\langle \mu_s, h^i f(\cdot) + D^i f(\cdot, Y_s)\big\rangle \\ & - \langle \pi_s, h^i\rangle \langle \mu_s, f\rangle\Big) g(Y_s)\Big)\, dI_s^i. \end{aligned} \tag{3.7}$$

Note that

$$(3.8)\quad \mathscr{F}_t^Y = \sigma\left\{Y_t, \int_0^t g(Y_s, s)\, ds: g \in C_b(\mathbb{R}^k \times [0,\infty)),\, g \geq 0\right\} \vee \mathscr{N},$$

where $\mathscr{N}$ is the collection of all P null sets. Since μ_t and π_t are $\mathscr{F}_t^Y$-measurable, we claim that to prove the theorem it suffices to show

$$(3.9)\quad E\left[\langle \mu_t, f\rangle G(Y_t, V_t^0, \dots, V_t^m)\right] = E\left[\langle \pi_t, f\rangle G(Y_t, V_t^0, \dots, V_t^m)\right],$$

for all $f \in C_b(E)$, $G \in C_b(\mathbb{R}^k \times [0,\infty)^{m+1})$ and all choices of $m \geq 1$, $g_1, \dots, g_m \in C_b(\mathbb{R}^k \times [0,\infty))$, $g_i \geq 0$, where the process V^i is defined by

$$(3.10)\quad V_t^0 = t, \qquad V_t^i = \int_0^t g_i(Y_s, s)\, ds,$$

for $1 \leq i \leq m$. For, it follows from (3.8)–(3.10) that, for every $f \in C_b(E)$, $\langle \mu_t, f\rangle = \langle \pi_t, f\rangle$ a.s. Since $C_b(E)$ is separable and μ_t and π_t are right continuous, we conclude that $\mu_t = \pi_t$, for all $0 \leq t \leq T$ a.s.

We prove (3.9) by using Theorem 2.1 for a suitably defined operator B. Let $E_0 = E \times \mathbb{R}^k \times [0,\infty)^{m+1}$. Choose and fix g_i as above for $i = 1, 2, \dots, m$. Let $g_0 \equiv 1$. Let $\mathscr{D}(B)$ be the algebra generated by functions of the form

$$\left\{g \in C_b(E_0): g(x, y, v_0, \dots, v_m) = f(x, y) \prod_{i=0}^m f_i(v_i),\right.$$
$$\left. f \in \mathscr{D}(A),\, f_i \in C_b^1([0,\infty))\right\}$$

and define

$$(3.11)\quad \begin{aligned} &Bg(x, y, v_0, \dots, v_m)\\ &\quad = Af(x, y) \prod_{i=0}^m f_i(v_i) + f(x, y) \sum_{i=0}^m g_i(y, v_0) f_i'(v_i) \prod_{j \neq i} f_j(v_j). \end{aligned}$$

Here f_i' denotes the derivative of f_i; $\mathscr{D}(B)$ separates points in E_0 since $\mathscr{D}(A)$ separates points in $E \times \mathbb{R}^k$ and $C_b^1([0,\infty))$ does so in $[0,\infty)$. Further, since $C_b^1([0,\infty))$ is separable under the metric given by the norm $|g| = \|g\| + \|g'\|$ and since $\mathscr{D}(A)$ satisfies the separability condition C2, so does B. Note that

$$|Bg(x, y, v_0, \dots, v_m)| \leq C_g(1 + \Theta'(x, y)), \qquad \forall\, (x, y, v_0, \dots, v_m) \in E_0.$$

Arguing as in the proof of Theorem 3.3 of Kurtz and Ocone (1988), we can show that $(X, Y, V^0, \dots, V^m)$ is the unique solution to the $D([0,T], E_0)$ martingale problem for B. Condition C5 can be verified from the fact that it is satisfied for A and that the $[0,\infty)^{m+1}$ component has a modification with right-continuous paths having left limits [see Theorem 4.3.8 of Ethier and Kurtz (1986)]. Hence $(B, \mathscr{D}(B))$ satisfy the conditions of Theorem 2.1.

Now using Itô's formula, (3.7), (3.10) and recalling (3.11) we get that, for $f \in \mathscr{D}(A_0)$, $g \in C_0^\infty(\mathbb{R}^k)$, $f_i \in C_b^1([0,\infty))$,

$$\begin{aligned}
\langle \mu_t, f\rangle g(Y_t) \prod_{i=0}^m f_i(V_t^i) \\
&= \langle \pi_0, f\rangle g(0) \prod_{i=0}^m f_i(V_0^i) \\
&\quad + \int_0^t \left\langle \mu_s, B\left(f \otimes g \bigotimes_{i=0}^m f_i\right)(\cdot, Y_s, V_s^0, \dots, V_s^m)\right\rangle ds \\
&\quad + \sum_{l=1}^k \int_0^t \big(\langle \mu_s, h^l f + D^l f(\cdot, Y_s)\rangle g(Y_s) \\
&\qquad - \langle \mu_s, f\rangle\langle \pi_s, h^l\rangle g(Y_s) + \langle \mu_s, f\rangle g^l(Y_s)\big) \prod_{i=0}^m f_i(V_s^i)\, dI_s^l.
\end{aligned}$$

Since $\{\rho_t\}$ satisfies (3.6),

$$\begin{aligned}
&\langle \mu_t, f\rangle g(Y_t) \prod_{i=0}^m f_i(V_t^i) - \langle \pi_0, f\rangle g(0) \prod_{i=0}^m f_i(V_0^i) \\
&\quad - \int_0^t \left\langle \mu_s, B\left(f \otimes g \bigotimes_{i=0}^m f_i\right)(\cdot, Y_s, V_s^0, \dots, V_s^m)\right\rangle ds
\end{aligned} \tag{3.12}$$

is a martingale. Now define $\nu_t^1 \in \mathscr{P}(E_0)$ by

$$\langle \nu_t^1, G\rangle = E\big[\langle \mu_t, G(\cdot, Y_t, V_t^0, \dots, V_t^m)\rangle\big], \qquad G \in C_b(E_0),$$

and similarly define ν^2 for $(\pi_t, Y_t, V_t^0, \dots, V_t^m)$. Then, taking expectations in (3.12), we get

$$\langle \nu_t^1, G\rangle = \langle \nu_0^1, G\rangle + \int_0^t \langle \nu_s^1, BG\rangle\, ds, \qquad \forall\, G \in \mathscr{D}(B). \tag{3.13}$$

A similar argument shows that ν^2 satisfies (3.13). Also

$$\begin{aligned}
\int_0^T &\langle \nu_t^1, \Theta'\rangle\, dt \\
&= \int_0^T E\langle \mu_t, \Theta'\rangle\, dt \\
&= \int_0^T E\langle \rho_t, \Theta'\rangle R_t^{-1}\, dt \\
&= E\int_0^T \langle \rho_t, \Theta'\rangle \exp\left\{-\sum_{i=1}^k \int_0^t \langle \pi_s, h^i\rangle\, dI_s^i - \tfrac{1}{2}\sum_{i=1}^k \int_0^t \left|\langle \pi_s, h^i\rangle\right|^2 ds\right\} dt \\
&< \infty
\end{aligned}$$

by (3.6). Similarly $\int_0^T \langle \nu_t^2, \Theta'\rangle\, dt < \infty$. We now apply Theorem 2.1 to conclude that $\nu_t^1 = \nu_t^2$ for all $t \le T$. In particular, (3.9) holds and, as remarked earlier, the proof is complete. □

Thus we have proved *pathwise uniqueness* of the solution to the Zakai equation. The filter π_t can be represented as a functional of the path Y as

$$\pi_t = H_t(Y), \tag{3.14}$$

and the functional σ_t can be represented as

$$\sigma_t = F_t(Y). \tag{3.15}$$

Here, for each t, H_t and F_t are functions from $D([0,T], E)$ into $\mathscr{P}(E)$ and $\mathscr{M}_+(E)$, respectively.

Using the well known Yamada–Watanabe argument, we can deduce the following result on uniqueness in law for the Zakai equation [see, e.g., Ikeda and Watanabe (1981)].

THEOREM 3.2. *Suppose (Y') is an r.c.l.l. process with values in $\mathbb{R}^k$ defined on a probability space $(\Omega', \mathscr{F}', P')$ such that*

$$\mathscr{L}(Y') = \mathscr{L}(Y),$$

and let $\pi'_t = H_t(Y')$. If $\{\rho'_t\}$ is an $\mathscr{F}_t^{Y'}$-adapted $\mathscr{M}_+(E)$-valued cadlag process satisfying the Zakai equation

$$\langle \rho'_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho'_s, A_0 f\rangle\, ds + \int_0^t \sum_{i=1}^k \left(\left\langle \rho'_s, h^i f + D^i f(\cdot, Y'_s)\right\rangle dY_s'^{i}\right),$$

$$\forall f \in \mathscr{D}(A_0),$$

and the integrability condition

$$E_{P'}\int_0^T \langle \rho'_t, \Theta'\rangle \exp\left\{-\int_0^t \sum_{i=1}^k \langle \pi'_s, h^i\rangle\, dI_s^i - \frac{1}{2}\sum_{i=1}^k \int_0^t \left|\langle \pi'_s, h^i\rangle\right|^2 ds\right\} dt < \infty, \tag{3.16}$$

then the law of the process $(Y'_\cdot, \rho'_\cdot)$ is the same as the law of the process $(Y_\cdot, \sigma_\cdot)$ and

$$\rho'_t = F_t(Y') \quad a.s.$$

Suppose that the measure P_0 [on $(\Omega, \mathscr{F})$] defined by

$$\frac{dP_0}{dP} = \exp\left\{-\int_0^T \sum_{i=1}^k h^i(X_s)\, dW_s^i - \frac{1}{2}\sum_{i=1}^k \int_0^T \left|h^i(X_s)\right|^2 ds\right\} \tag{3.17}$$

is a probability measure. This is so, for example, if X and W are independent or if h is a bounded function [see Kallianpur (1980)]. It is well known that Y is a k-dimensional standard Brownian motion on $(\Omega, \mathscr{F}, P_0)$. Thus while considering the Zakai equation, it is convenient to work on the probability space $(\Omega, \mathscr{F}, P_0)$. The probability measure P_0 is called a reference probability. Also $\langle \sigma_t, f\rangle$ can now be expressed as

$$\langle \sigma_t, f\rangle = E_{P_0}\left[f(X_t)\frac{dP}{dP_0}\middle|\mathscr{F}_t^Y\right], \qquad \forall f \in C_b(E) \tag{3.18}$$

[see Theorem 18.21 of Elliott (1982), page 291]. This is an analogue of the Kallianpur–Striebel Bayes formula.

Condition (3.6) can then be rewritten as

$$E_{P_0}\left[\int_0^T \langle \rho_t, \Theta' \rangle \, dt\right] < \infty. \tag{3.19}$$

We now give a result on solutions to the Zakai equation which may not be adapted to the observations process. This result can be proved using arguments similar to the one given in Theorem 3.1 and hence we will only give a sketch below.

THEOREM 3.3. *Suppose that the conditions of Theorem* 3.1 *are satisfied and that* P_0 *defined by* (3.17) *satisfies* $P_0(\Omega) = 1$. *Suppose* (Y^*) *is an* $\mathbb{R}^k$*-valued* $(\mathscr{G}_t)$*-Wiener process on a probability space* $(\Omega^*, \mathscr{F}^*, P^*)$. *If* $\{\rho_t^*\}$ *is a* $(\mathscr{G}_t)$*-adapted* $\mathscr{M}_+(E)$*-valued cadlag process satisfying the Zakai equation*

$$\langle \rho_t^*, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho_s^*, A_0 f\rangle \, ds + \int_0^t \sum_{i=1}^k \left(\left\langle \rho_s^*, h^i f + D^i f(\cdot, Y_s^*)\right\rangle\right) dY_s^{*i},$$

$$\forall f \in \mathscr{D}(A_0),$$

and the integrability condition

$$E_{P^*}\int_0^T \langle \rho_t^*, \Theta^* \rangle \, dt < \infty, \tag{3.20}$$

then one has

$$E_{P^*}\left[\langle \rho_t^*, f\rangle | \mathscr{F}_t^{Y^*}\right] = \langle F_t(Y^*), f\rangle. \tag{3.21}$$

In particular, if (ρ^*) *is* $(\mathscr{F}_t^{Y^*})$*-adapted, then*

$$\rho_t^* = F_t(Y^*) \tag{3.22}$$

and

$$P^* \circ (Y_\cdot^*, \rho_\cdot^*)^{-1} = P_0 \circ (Y_\cdot, \sigma_\cdot)^{-1}. \tag{3.23}$$

PROOF. Let $g_1, g_2, \dots, g_m$, E_0 and B be as in the proof of Theorem 3.1. Define V_t^i as in (3.10) with Y^* in place of Y. Recalling that Y^* is a Brownian motion and using Itô's formula, it follows that

$$\begin{aligned} &\langle \rho_t^*, f\rangle g(Y_t^*) \prod_{i=0}^m f_i(V_t^i) - \langle \pi_0, f\rangle g(0) \prod_{i=0}^m f_i(V_0^i) \\ &\quad - \int_0^t \left\langle \rho_s^*, B\left(f \otimes g \bigotimes_{i=0}^m f_i\right)(\cdot, Y_s^*, V_s^0, \dots, V_s^m)\right\rangle ds \end{aligned} \tag{3.24}$$

is a $(\mathscr{G}_t)$ martingale. Hence defining ν_t^1 by

$$\langle \nu_t^1, G\rangle = E_{P^*}\left[\left\langle \rho_t^*, G(\cdot, Y_t^*, V_t^0, \dots, V_t^m)\right\rangle\right], \qquad G \in C_b(E_0), \tag{3.25}$$

and taking expectation in (3.24) we get

$$\langle \nu_t^1, G\rangle = \langle \nu_0^1, G\rangle + \int_0^t \langle \nu_s^1, BG\rangle \, ds, \qquad \forall G \in \mathscr{D}(B). \tag{3.26}$$

Noting that $F_t(Y^*)$ is a solution to the Zakai equation, it follows that ν_t^2 defined by the rhs in (3.25) with ρ_t^* replaced by $F_t(Y^*)$ also satisfies (3.26). As in Theorem 3.1, we get $\nu_t^1 = \nu_t^2$ for all t, that is,

$$(3.27)\quad E\big[\langle \rho_t^*, f\rangle G(Y_t, V_t^0, \dots, V_t^m)\big] = E\big[\langle F_t(Y^*), f\rangle G(Y_t, V_t^0, \dots, V_t^m)\big],$$

for all choices of f, G. This yields (3.21). The other part follows easily from this. □

This result leads us to the following notion of a weak solution to the Zakai equation.

DEFINITION 3.1. An $\mathscr{M}_+(E)$-valued process ρ defined on some probability space $(\Omega^*, \mathscr{F}^*, P^*)$ is a weak solution of the Zakai equation if there exists an $\mathbb{R}^k$-valued $(\mathscr{G}_t)$ Wiener process Y^* (defined possibly on an extended probability space) and such that ρ is $(\mathscr{G}_t)$-adapted and (ρ, Y) satisfies the Zakai equation

$$\langle \rho_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho_s, A_0 f\rangle\, ds + \int_0^t \sum_{i=1}^k \big(\langle \rho_s, h^i f + D^i f(\cdot, Y_s^*)\rangle\big)\, dY_s^{*i},$$
$$\forall f \in \mathscr{D}(A_0).$$

Suppose that the D^i's appearing in (3.1) depend only on the first coordinate and further suppose, for simplicity of writing, that $D^i f$ is a bounded function for every i and every f. That is, the operator A defined by (3.2) has the form

$$(3.28)\quad A(f\otimes g) = A_0 f\otimes g + \tfrac{1}{2} f\otimes \Delta g + \sum_{i=1}^k \big(D^i f\otimes g^i + h^i f\otimes g^i\big),$$

where now $D^i\colon \mathscr{D}(A_0) \to C_b(E)$ for every i. It can be shown that assuming that A has this form is equivalent to assuming that if (X', Y') is a solution to the martingale problem for A, then defining W' by $W_t' = Y_t' - \int_0^t h(X_s')\, ds$ one has that (X', W') is a solution to a martingale problem for an operator A^* with domain $\mathscr{D}(A)$.

Further, the operator A^* is given as follows. For $f\otimes g \in \mathscr{D}(A^*)$,

$$A^*(f\otimes g) = A_0 f\otimes g + \tfrac{1}{2} f\otimes \Delta g + \sum_{i=1}^k D^i f\otimes g^i,$$
$$A^*(f\otimes 1) = A_0 f,\qquad A^*(1\otimes g) = \tfrac{1}{2}\Delta g,\qquad A^*(1\otimes 1) = 0.$$

In this case, we can also show that the martingale problem for A is well posed if and only if the martingale problem for A^* is well posed [see Lemma 4.4 of Kurtz and Ocone (1988)], and then that A satisfies C1–C5 if and only if A^* satisfies C1–C5.

3.1. *Martingale problem for* σ. We will now consider the martingale problem corresponding to σ. Let

$$\begin{aligned}\mathscr{D}(\mathscr{A}) = \{F \in C_b(\mathscr{M}_+(E))\colon F(\mu) &= g(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle), \\ &\text{for some } n \geq 1,\ f_1, \dots, f_n \in \mathscr{D}(A_0),\ g \in C_0^2(\mathbb{R}^n)\}.\end{aligned} \tag{3.29}$$

For such an $F \in \mathscr{D}(\mathscr{A})$ define

$$\begin{aligned}\mathscr{A}F(\mu) = {} & \sum_{l=1}^{n} g^l(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle)\langle \mu, A_0 f_l\rangle \\ & + \tfrac{1}{2} \sum_{l,j=1}^{n} \sum_{i=1}^{k} \big(g^{lj}(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle) \\ & \qquad \times \langle \mu, h^i f_l + D^i f_l\rangle \langle \mu, h^i f_j + D^i f_j\rangle\big).\end{aligned} \tag{3.30}$$

Note that

$$|\mathscr{A}F(\mu)| \leq C_F \Xi(\mu),$$

where $\Xi(\mu) = \langle \mu, \Theta\rangle + |\langle \mu, 1\rangle|^2 + \sum_{i=1}^{k} \langle \mu, |h^i|\rangle^2$. Suppose σ satisfies

$$E_{P_0} \int_0^T \Xi(\sigma_s)\, ds < \infty. \tag{3.31}$$

Now (1.9) and an application of Itô's formula to $F(\sigma_t)$ for $F \in \mathscr{D}(\mathscr{A})$ implies that $\{\sigma_t\}$ is a solution to the martingale problem for $\mathscr{A}$ on $(\Omega, \mathscr{F}, P_0)$.

THEOREM 3.4. *Suppose that the operator* A *is given by* (3.28) *and that the conditions of Theorem* 3.1 *are satisfied by* A. *Let* π_0 *satisfy* (3.5). *Further suppose that if* (X, Y) *is the solution to the martingale problem for* $(A, \pi_0 \otimes \delta_0)$, *then* P_0 *defined by* (3.17) *is a probability measure and that* (1.2) *and* (3.31) *hold.*

Then any solution $\{\rho_t\}$ *of the* $D([0,T], \mathscr{M}_+(E))$ *martingale problem for* $(\mathscr{A}, \delta_{\pi_0})$ *and satisfying*

$$E \int_0^T \Xi(\rho_s)\, ds < \infty$$

is a weak solution of the Zakai equation.

PROOF. The existence of a solution follows from the remarks just preceding this theorem. Let $\{f_n\}$ be the countable set appearing in C2. Without loss of generality, suppose that $\{f_n\}$ is an algebra. Then a process X is a solution to the martingale problem for A_0 if and only if it is a solution to the martingale problem for $A_0|_{\{f_n\}}$ [see Proposition IV.3.1. of Ethier and Kurtz (1986)]. Let $\overline{A_0} = A_0|_{\{f_n\}}$. Let $\bar{A}$ be defined correspondingly by (3.2). It follows from the hypothesis on A and the above-mentioned fact that $\bar{A}$ satisfies the conditions of Theorem 3.1.

Let ρ be a solution to the $D([0,T], \mathscr{M}_+(E))$ martingale problem for $\mathscr{A}$. A standard argument shows that

$$M_t^l = \langle \rho_t, f_l \rangle - \langle \rho_0, f_l \rangle - \int_0^t \langle \rho_s, \overline{A_0} f_l \rangle \, ds \tag{3.32}$$

is a martingale for every $l \geq 1$ and that the cross quadratic variation process $\langle M^l, M^j \rangle$ is given by

$$\langle M^l, M^j \rangle_t = \sum_{i=1}^k \int_0^t \langle \rho_s, h^i f_l + D^i f_l \rangle \langle \rho_s, h^i f_j + D^i f_j \rangle \, ds. \tag{3.33}$$

Let $a_l = 2^{-l}(\|f_l\| + \|Df_l\|)^{-1}$. Then note that if we write $\overline{h} = \sum_{i=1}^k |h^i|$,

$$\begin{aligned} E \sup_{0 \leq t \leq T} \left(a_l M_t^l \right)^2 &\leq 4 a_l^2 E \int_0^T \left| \langle \rho_s, hf_i + Df_i \rangle \right|^2 ds \\ &\leq \frac{8}{2^{-2l}} E \int_0^T \left(\left| \langle \rho_s, \overline{h} \rangle \right|^2 + \left| \langle \rho_s, 1 \rangle \right|^2 \right) ds \\ &< \infty. \end{aligned} \tag{3.34}$$

The finiteness of the last expression follows since ρ is a solution to the martingale problem for $\mathscr{A}$. Let $e_l = (\delta_{lj})$ be the standard basis in ℓ^2. Let $Z_t^l = a_l M_t^l$. Then (3.34) implies that $Z_t = \sum_{l=1}^\infty Z_t^l e_l$ is an ℓ^2-valued martingale. In fact, $Z \in D([0,T], \ell^2)$. Let $\Sigma_s(\omega) \in \mathscr{L}(\ell^2)$ be defined by

$$\left(\Sigma_s(\omega) e_l, e_j \right) = \begin{cases} a_j \left\langle \rho_s(\omega), h^l f_j + D^l f_j \right\rangle, & 1 \leq l \leq k, \\ 0, & l \geq k, \end{cases}$$

for every $j \geq 1$. Here $(\cdot, \cdot)$ denotes the inner product in ℓ^2. Note

$$\begin{aligned} E \int_0^T \|\Sigma_s\|_{\mathrm{HS}}^2 \, ds &= E \int_0^T \sum_{l,j=1}^\infty \left(\Sigma_s e_l, e_j \right)^2 ds \\ &= E \int_0^T \sum_{j=1}^\infty \sum_{l=1}^k a_j^2 \langle \rho_s, h^l f_j + D^l f_j \rangle^2 \, ds \\ &\leq 8 E \int_0^T \left(\left| \langle \rho_s, \overline{h} \rangle \right|^2 + \left| \langle \rho_s, 1 \rangle \right|^2 \right) ds \\ &< \infty. \end{aligned}$$

Hence arguing as in Yor (1974) we get the existence of a $(\mathscr{G}_t)$ cylindrical Brownian motion β, possibly on an enlarged probability space, where ρ is $(\mathscr{G}_t)$-adapted and such that

$$Z_t = \int_0^t \Sigma_s \, d\beta_s. \tag{3.35}$$

Let $\beta_s^l = \beta_s(e_l)$. Then by (3.35) we get

$$(Z_t, e_l) = \int_0^t \langle \Sigma_s^* e_l, d\beta_s \rangle,$$

where Σ_s^* is the adjoint of the operator Σ_s, or

$$\begin{aligned} a_l M_t^l &= \int_0^t \sum_{j=1}^{\infty} (\Sigma_s^* e_l, e_j)\, d\beta_s^j \\ &= \int_0^t \sum_{j=1}^{k} a_l \langle \rho_s, h^j f_l + D^j f_l \rangle\, d\beta_s^j. \end{aligned}$$

Let Y be the k-dimensional Brownian motion defined by $Y_t^l = \beta_t^l$, $1 \le l \le k$. Then we have

$$M_t^l = \int_0^t \langle \rho_s, hf_l + Df_l \rangle\, dY_s.$$

Hence ρ is a solution of the Zakai equation (1.9), where now the test functions f belong to $\mathscr{D}(\overline{A_0})$. □

REMARK 3.1. In all of the earlier work cited in the Introduction the authors have considered the case when the signal process X takes values in $\mathbb{R}^d$ and when A_0 is the generator of X. Szpirglas (1978) has shown the uniqueness of solutions to these equations when X is a general $\mathbb{R}^d$-valued Markov process, for a bounded h and when the signal and noise are independent. Uniqueness is also proved for special cases of the signal process and special classes of unbounded h in Pardoux (1982), Baras, Blankenship and Hopkins (1983) and Sheu (1983). In a similar setup with h bounded, uniqueness of solution in the class of finite signed measures for the measure-valued Zakai and Kushner or FKK equations is proved in a recent paper by Rozovskii (1991) which also contains references to other work in this area.

When the Markov process X_t takes values in $\mathbb{R}^d$ and admits a density with respect to the Lebesgue measure, the corresponding Zakai equation for the unnormalized conditional density can be considered. Uniqueness in this case has been studied by Pardoux (1979), Krylov and Rozovskii (1981) and Chaleyat-Maurel, Michel and Pardoux (1990).

The two papers which can be directly compared with the results in this section are those of Kurtz and Ocone (1988) and Hijab (1989). As mentioned earlier in the Introduction, Theorem 3.1 is an improvement of the results of Kurtz and Ocone (1988) where they have proved uniqueness of solution for the two equations when the state space is assumed to be a locally compact separable metric space. They have considered unbounded h, although they still require that hf be bounded for every f in the domain of A_0. They prove the result under two sets of conditions on the operator A_0. When h is continuous they require that the domain of A_0 be sufficiently rich. When h is discontinuous they assume that the range of $(\lambda - A_0)$ is sufficiently dense for every $\lambda > 0$. This condition essentially implies that the closure of A_0 is the generator of the Markov process X.

Hijab (1989) also considered the case when the state space is a complete separable metric space. He showed that π is the unique solution of a martingale problem for a suitably defined operator. He also required that h be bounded and belong to the domain of A_0.

Earlier, Kallianpur and Karandikar (1984) obtained results for uniqueness of solution for the analogues of the Zakai and FKK equations for a general Polish space-valued Markov process and for an unbounded h via their finitely additive approach to filtering theory, when A_0 is the generator of the Markov process X. Bhatt and Karandikar (1993b) have extended these results to the case when A_0 may not be the generator, but the martingale problem for A_0 is well posed. The advantage of this is that the test functions are taken from the domain of A_0 as opposed to the full domain of the generator in Kallianpur and Karandikar (1984), which is usually intractable.

4. Independent signal and noise. An important special case of the previous results occurs when X and W are independent. This is the case which usually arises in practical applications. In this case uniqueness of the Zakai equation can be proved under similar conditions on the operator A_0 instead of those on A in the correlated case. Note that in this case $Df \equiv 0$ for all $f \in \mathscr{D}(A_0)$. Recall from (3.17) that

$$\frac{dP_0}{dP} = \exp\left\{ -\int_0^T \sum_{i=1}^k h^i(X_s)\, dW_s^i - \frac{1}{2} \sum_{i=1}^k \int_0^T |h^i(X_s)|^2\, ds \right\}$$

defines a probability measure on $(\Omega, \mathscr{F})$. Further, under P_0, Y is a Brownian motion independent of X and the law of X under P_0 is the same as the law of X under P. Thus the conditional expectation appearing in (3.18) can be evaluated by integrating with respect to the law of X and thus we get an explicit expression for $\langle \sigma_t, f \rangle$ and hence for $\langle \pi_t, f \rangle$.

Let $\Omega^0 = C([0,T], \mathbb{R}^k)$, let $\mathscr{F}^0$ be the Borel σ-field on Ω^0 and let Q be the Wiener measure. Let $\tilde{Y}$ be the coordinate process on Ω^0. Let $\tilde{X}$ be a process on $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$, where the law of $\tilde{X}$ is the same as the law of X. Let $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}) = (\hat{\Omega}, \hat{\mathscr{F}}, \hat{P}) \otimes (\Omega^0, \mathscr{F}^0, Q)$. Define F by

$$\langle F_t(\omega^0), f \rangle = \int f\big(\tilde{X}_t(\hat{\omega})\big) q_t(\hat{\omega}, \omega^0)\, d\hat{P}(\hat{\omega}), \qquad \forall f \in C_b(E), \tag{4.1}$$

where

$$q_t(\hat{\omega}, \omega^0) = \exp\left\{ \sum_{i=1}^k \int_0^t h^i\big(\tilde{X}_s(\hat{\omega})\big)\, d\tilde{Y}_s^i(\omega^0) - \frac{1}{2} \sum_{i=1}^k \int_0^t \Big(h^i\big(\tilde{X}_s(\hat{\omega})\big)\Big)^2 ds \right\}. \tag{4.2}$$

Then σ_t can be defined by $\langle \sigma_t(\omega), f \rangle = \langle F_t(Y(\omega)), f \rangle$. Now we know that $\langle \pi_t, f \rangle = \langle \sigma_t, f \rangle / \langle \sigma_t, 1 \rangle$. This is the Kallianpur–Striebel Bayes formula. We will use this later in Section 8 when proving the robustness results.

The following result gives sufficient conditions for the uniqueness of solution of the Zakai equation when the signal and noise are independent. This can be deduced from the results in the previous section.

THEOREM 4.1. *Suppose that the signal process X and the noise W are independent. Further suppose that X is the unique solution of the martingale problem for* (A_0, π_0) *and that* A_0 *satisfies conditions* C1–C5. *Also let* (1.2) *hold.*

If $\{\rho_t\}$ *is an* $\mathscr{F}_t^Y$*-adapted* $\mathscr{M}_+(E)$*-valued cadlag process satisfying*

$$\langle \rho_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho_s, A_0 f\rangle\, ds + \sum_{i=1}^k \int_0^t \langle \rho_s, h^i f\rangle\, dY_s^i$$

and

$$E_{P_0}\int_0^T \langle \rho_t, \Theta\rangle\, dt < \infty, \tag{4.3}$$

then $\rho_t = \sigma_t$, *for all* $0 \le t \le T$ *a.s., where* σ_t *is defined by* (1.8).

Uniqueness in law holds for the Zakai equation. Further if $\mathscr{A}$ *is defined by* (3.29) *and* (3.30) *with* $D \equiv 0$ *and if* (3.31) *holds, then any solution to the martingale problem for* $(\mathscr{A}, \delta_{\pi_0})$ *is a weak solution of the Zakai equation.*

PROOF. We first note that $\{\sigma_t\}$ satisfies (4.3). It follows from (4.3) that

$$\langle \rho_t, f\rangle - \int_0^t \langle \rho_s, A_0 f\rangle\, ds$$

is a martingale under P_0. Let $\nu_t \in \mathscr{P}(E)$ be defined by

$$\langle \nu_t, f\rangle = E_{P_0}\langle \rho_t, f\rangle.$$

Then $\{\nu_t\}$ satisfy

$$\langle \nu_t, f\rangle = \langle \nu_0, f\rangle + \int_0^t \langle \nu_s, f\rangle\, ds, \qquad \forall f \in \mathscr{D}(A_0).$$

By Theorem 2.1, ν_t is the unique such solution. As in Theorem 3.1 we can deduce that $E_{P_0}\langle \rho_t, g\rangle = E_{P_0}\langle \sigma_t, g\rangle$, for all $g \in C(E)$. It is easy to see using (1.2) and (3.18) that $\{\sigma_t\}$ satisfies (3.19). Hence, so does $\{\rho_t\}$.

Now define A by (3.2), with $Df \equiv 0$. An easy verification shows that $(A, \mathscr{D}(A))$ satisfy conditions C1, C2 and C3. The well posedness of the $D([0,T], E \times \mathbb{R})$ martingale problem for $(A, \delta_{(x,y)})$ is proved as in Lemma 4.4 of Kurtz and Ocone (1988). To prove condition C5 once again note that if (X, Y) is a solution to the martingale problem for A, then X is a solution to the martingale problem for A_0 and hence admits a cadlag modification. Existence of a cadlag modification for Y follows as in Theorem IV.3.8 of Ethier and Kurtz (1986). Hence condition C5 is satisfied by A. Now Theorem 3.1 is applicable and we get the requisite pathwise uniqueness.

The remaining assertions in the theorem follow from Theorems 3.2, 3.3 and 3.4. □

5. Infinite-dimensional model. A close look at the proofs of the results in Sections 3 and 4 suggest that the analysis will go through even when the noise is assumed to be infinite dimensional. Suppose H_1 is a real separable

Hilbert space. Let W be an H_1-valued cylindrical Brownian motion. Suppose $h\colon E \to H_1$ is a continuous function satisfying

$$E\int_0^T \|h(X_s)\|^2\,ds < \infty. \tag{5.1}$$

The observation model is

$$Y_t = \int_0^t h(X_s)\,ds + W_t. \tag{5.2}$$

For the sake of simplicity let us assume that X and W are independent. Define the *reference* probability measure P_0 as in (3.17). It is not difficult to see that Y is a cylindrical Brownian motion on the transformed probability space $(\Omega, \mathscr{F}, P_0)$.

Define $\sigma_t \in \mathscr{M}_+(E)$ by

$$\langle \sigma_t, f\rangle = E_{P_0}\left[f(X_t)\frac{dP}{dP_0}\middle| \mathscr{F}_t^Y\right], \qquad \forall\, f \in C_b(E).$$

Then it can be verified that σ_t satisfies the Zakai equation

$$\langle \sigma_t, f\rangle = \langle \sigma_0, f\rangle + \int_0^t \langle \sigma_s, A_0 f\rangle\,ds + \int_0^t \langle \sigma_s, hf\rangle\,dY_s, \qquad \forall\, f \in \mathscr{D}(A_0). \tag{5.3}$$

Here A_0 is as in the previous section, namely, X is the unique solution of the martingale problem for A_0. Let $\{e_i\colon i \geq 1\}$ be a complete orthonormal system (CONS) in H_1 and let $\bar{h} = \sum_{i=1}^{\infty} |h^i| e_i$, where $h^i = (h, e_i)$. Suppose σ satisfies

$$E_{P_0}\int_0^T \left(\langle \sigma_s, \Theta\rangle + \|\langle \sigma_s, \bar{h}\rangle\|^2\right) ds < \infty. \tag{5.4}$$

A weak solution of the Zakai equation for the infinite-dimensional observation model can be defined as in Definition 3.1.

THEOREM 5.1. (a) *Let $h\colon E \to H_1$ be a continuous function satisfying* (5.1). *Suppose A_0 satisfies conditions* C1–C5. *Further suppose that the $D([0,T], E)$ martingale problem for (A_0, δ_x) is well posed for every $x \in E$. If $\{\rho_t\}$ is an $\mathscr{F}_t^Y$-adapted $\mathscr{M}_+(E)$-valued cadlag process satisfying*

$$E_{P_0}\left[\int_0^T \langle \rho_s, \Theta\rangle\,ds\right] < \infty$$

and which is a solution of the equation (5.3), *then $\rho_t = \sigma_t$, for every t a.s.*

(b) *Let $\mathscr{D}(\mathscr{A})$ be defined by* (3.29) *and, for $F \in \mathscr{D}(\mathscr{A})$, let*

$$\begin{aligned}\mathscr{A}F(\mu) = {} & \sum_{l=1}^{n} g^l(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle)\langle \mu, A_0 f_l\rangle \\ & + \tfrac{1}{2}\sum_{l,j=1}^{n}\sum_{i=1}^{\infty}\big(g^{lj}(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle) \\ & \qquad \times \langle \mu, h^i f_l + D^i f_l\rangle\langle \mu, h^i f_j + D^i f_j\rangle\big).\end{aligned}$$

Then any solution of the $D([0,T],\mathscr{M}_+(E))$ *martingale problem for* $(\mathscr{A},\delta_{\pi_0})$ *satisfying*

$$E\int_0^T\Big(\langle\rho_s,\Theta\rangle+\big\|\langle\rho_s,\bar{h}\rangle\big\|^2\Big)\,ds$$

is a weak solution of the Zakai equation.

PROOF. (a) Let $\{e_i\}$ be a CONS in H_1. Let P_n be the orthogonal projection onto $[e_1,\dots,e_n]$. Let $\hat{\mathscr{D}}$ be the algebra generated by functions of the form $f\otimes\hat{g}\in C_b(E\times H_1)$, where $f\in\mathscr{D}(A_0)\cup\{1\}$ and $\hat{g}(y)=g(P_n(y))$, $g\in C_b^\infty(\mathbb{R}^n)$, for some $n\ge 1$ or $g\equiv 1$. For $f\otimes\hat{g}\in\hat{\mathscr{D}}$, $g\in C_b^\infty(\mathbb{R}^n)$, define

$$A^*(f\otimes\hat{g})=A_0f\otimes g+\tfrac{1}{2}f\otimes\Delta g$$

and

$$A(f\otimes\hat{g})=A_0f\otimes g+\tfrac{1}{2}f\otimes\Delta g+\sum_{i=1}^n h^if\otimes g^i,$$

where Δ is the Laplacian operator and g^i is the ith partial derivative of g.

Then (X,W) is the unique solution of the martingale problem for A^* and (X,Y) is a solution of the martingale problem for A. We want to prove that the martingale problem for A is well posed. Hence the argument of Lemma 4.4 of Kurtz and Ocone (1988) goes through for every $f\otimes\hat{g}\in\hat{\mathscr{D}}$. Thus we get that if (X_t',Y_t') is a solution of the martingale problem for A, then $[X_t',W_t'=Y_t'-\int_0^t h(X_s')\,ds]$ is a solution of the martingale problem for A^*. This proves that the martingale problem for A is well posed. Conditions C1–C5 can be verified as in the finite-dimensional case. The rest of the argument is exactly as in the proof of Theorem 3.1.

(b) For (b) we can argue as in the finite-dimensional case that since (5.4) is satisfied, σ is a solution of the martingale problem for $\mathscr{A}$. To prove the final assertion we proceed as in the proof of Theorem 3.4. Note that now Σ_s will be defined by

$$\big(\Sigma_s(\omega)e_i,e_j\big)=a_j\big\langle\rho_s(\omega),h_if_j\big\rangle,\qquad i,j\ge 1.$$

Then

$$\begin{aligned}E\int_0^T\|\Sigma_s\|_{\mathrm{HS}}^2\,ds&=E\int_0^T\sum_{i,j=1}^\infty\big(\Sigma_se_i,e_j\big)^2\,ds\\&\le E\int_0^T\sum_{i=1}^\infty\big\langle\rho_s,|h^i|\big\rangle^2\,ds\\&<\infty.\end{aligned}$$

The rest of the proof follows. □

6. Applications to Φ′-valued SDE's. We will apply the results of Section 4 to models of environmental pollution. We begin with some facts about stochastic differential equations in duals of nuclear spaces.

Let Φ be a countably Hilbertian nuclear space. That is, Φ is a real linear space whose topology is given by an increasing sequence $\|\cdot\|_n$: $n \in \mathbb{Z}^+$ of Hilbertian seminorms. Let H_n denote the completion of Φ with respect to the norm $\|\cdot\|_n$, and let Φ' and H_{-n} denote the duals of Φ and H_n, respectively. Let $\{\phi_j\} \subset \Phi$ be a CONS in H_0 and a complete orthogonal system in H_p, $p \in \mathbb{Z}$. Let $\phi_j^p = \|\phi_j\|_p^{-1}\phi_j$. Then ϕ_j^p is a CONS in H_p. For $v \in H_{-p}$, $\phi \in H_p$, the action of v on ϕ will be denoted by $v[\phi]$ and is given by

$$v[\phi] = \sum_{j=1}^{\infty} \langle v, \phi_j^{-p} \rangle_{-p} \langle \phi, \phi_j^p \rangle_p.$$

Consider the following Φ'-valued SDE:

$$X_t = X_0 + \int_0^t b(s, X_s)\,ds + \int_0^t \int_U G(s, X_{s-}, u)\tilde{N}(du\,ds) \tag{6.1}$$

under Assumption S; $\tilde{N}$ is a compensated Poisson random measure with intensity μ.

ASSUMPTION S. For (b, G, μ): $\forall\ T > 0$, $\exists\ p_0 = p_0(T) \in N^+$ such that $\forall\ p \geq p_0$, $\exists\ q \geq p$ and a constant $k = k(q, p, T)$ such that the following conditions hold:

(S1) *Continuity*. For all $t \in [0, T]$, $b(t, \cdot)$: $H_{-p} \to H_{-q}$ is continuous, $\forall$ $t \in [0, T]$ and $v \in H_{-p}$, $G(t, v, \cdot) \in L^2(U, \mu, H_{-p})$ and, for t fixed, the map $v \mapsto G(t, v, \cdot)$ is continuous from H_{-p} to $L^2(U, \mu, H_{-p})$.

(S2) *Coercivity*. For all $t \in [0, T]$ and $\phi \in \Phi$,

$$2b(t, \phi)\big[\theta_p(\phi)\big] \leq k\big(1 + \|\phi\|_{-p}^2\big).$$

(S3) *Growth*. For all $t \in [0, T]$ and $v \in H_{-p}$,

$$\|b(t, v)\|_{-q}^2 \leq k\big(1 + \|v\|_{-p}^2\big) \quad \text{and} \quad \int_U \|G(t, v, u)\|_{-p}^2 \mu(du) \leq k\big(1 + \|v\|_{-p}^2\big).$$

(S4) *Monotonicity*. For all $t \in [0, T]$ and $v_1, v_2 \in H_{-p}$,

$$\big\langle b(t, v_1) - b(t, v_2), v_1 - v_2 \big\rangle_{-q} + \int_U \|G(t, v_1, \cdot) - G(t, v_2, \cdot)\|_{-q}^2\, d\mu$$
$$\leq k\|v_1 - v_2\|_{-q}^2,$$

where θ_p is the isometric linear map from H_{-p} to H_p given by

$$\theta_p\left(\sum_{i=1}^{\infty} \alpha_i \phi_i^{-p}\right) = \sum_{i=1}^{\infty} \alpha_i \phi_i^p.$$

Let

$$\mathscr{D}_0^\infty(\Phi') := \big\{F: \Phi' \to \mathbb{R} \mid \exists\, f \in C_0^\infty(\mathbb{R}) \text{ and } \phi \in \Phi \text{ such that } F(v) = f(v[\phi])\big\}$$

and, for $F \in \mathscr{D}_0^\infty(\Phi')$, define $A'_sF\colon \Phi' \to \mathbb{R}$ by

$$A'_sF(v) = b(s,v)[\phi]f'(v[\phi]) + \int_U \{f(v[\phi] + G(s,v,u)[\phi]) - f(v[\phi]) - G(s,v,u)[\phi]f'(v[\phi])\}\mu(du).$$

Here f' denotes the derivative of f. The following is a summarization of results proved in Hardy, Kallianpur, Ramsubramanian and Xiong (1994).

THEOREM 6.1. (a) *Under Assumption* S *the SDE* (6.1) *has a unique* Φ'-*valued solution if there exists* r_0 *s.t.* $E\|X_0\|_{-r_0}^2 < \infty$. *Furthermore, let* $p(T) = \max(r_0, p_0(T))$ *and* $p_1(T) \geq p(T)$ *be such that the canonical injection from* $H_{p_1(T)}$ *to* $H_{p(T)}$ *is Hilbert–Schmidt. Then* $X|_{[0,T]} \in D([0,T], H_{-p_1(T)})$ *and*

$$E \sup_{0 \leq t \leq T} \|X_t\|_{-p_1(T)}^2 = \tilde{K} < \infty, \tag{6.2}$$

where $\tilde{K}$ *depends only on* k *and* $E\|X_0\|_{-p_1(T)}^2$.

(b) *The process* X *is the unique solution of the* $D([0,\infty), \Phi')$ *martingale problem for* (A'_s).

We want to consider the filtering problem when the signal process is a solution to the martingale problem for (A'_s) and the observation model is as in (1.1), where W is independent of X and h is a continuous function satisfying (1.2). Since we are concerned with a finite time interval $[0,T]$, we fix $T > 0$ and write p_0 and p_1 for $p_0(T)$ and $p_1(T)$, respectively. Let $p_2 \geq p_1$ be such that the canonical injection from H_{p_2} to H_{p_1} is Hilbert–Schmidt. Recall that $H_{-p_1} \subset H_{-p_2}$. We will consider the martingale problem on H_{-p_2}.

To apply Theorem 4.1 we have to get a suitable operator which satisfies the conditions of that theorem. We assume the following extra condition.

(S1′) For all $\phi \in \Phi$, $b(s,v)[\phi]$ and $\int_U G(s,v,u)[\phi]\,\mu(du)$ are continuous in s and v.

Let $\mathscr{D} = \{F\colon H_{-p_2} \to \mathbb{R} \mid \exists\ n \geq 1,\ f \in C_0^\infty(\mathbb{R}^n)$ and $\phi_1, \dots, \phi_n \in \Phi$ such that $F(v) = f(v[\phi_1], \dots, v[\phi_n])\}$. We extend A'_s to A_s on the algebra $\mathscr{D}$ as follows:

$$A_sF(v) = \sum_{i=1}^n f_i(v[\phi])b(s,v)[\phi_i] + \int_U \left\{ f(v[\phi] + G(s,v,u)[\phi]) - f(v[\phi]) - \sum_{i=1}^n G(s,v,u)[\phi_i]f_i(v[\phi]) \right\}\mu(du), \tag{6.3}$$

where $f_i = (\partial/\partial x_i)f$, $\phi = (\phi_1, \dots, \phi_n)$ and $v[\phi] = (v[\phi_1], \dots, v[\phi_n])$.

Note that in view of (S1′), the operator $A^0 := d/ds + A_s$ is an operator on $C_0^1[0, T] \otimes C(H_{-p_2})$.

Define P_0 by (3.17) and let σ_t be defined as in (3.18).

THEOREM 6.2. *If* $\{\rho_t\}$ *is an* $\mathscr{F}_t^Y$*-adapted* $\mathscr{M}_+(H_{-p_2})$*-valued cadlag process satisfying* (3.6) *and*

$$\langle \rho_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho_s, A_s f\rangle\, ds + \sum_{i=1}^k \int_0^t \langle \rho_s, h^i f\rangle\, dY_s^i, \qquad \forall f \in \mathscr{D}, \tag{6.4}$$

then $\rho_t = \sigma_t$ *for all* $0 \le t \le T$ *a.s.*

PROOF. Let $\rho_t^0 = \delta_t \otimes \rho_t$ and let A^0 be as defined above. Then (6.4) is equivalent to

$$\langle \rho_t^0, f\rangle = \langle \rho_0^0, f\rangle + \int_0^t \langle \rho_s^0, A^0 f\rangle\, ds + \sum_{i=1}^k \int_0^t \langle \rho_s^0, h^i f\rangle\, dY_s^i, \qquad \forall f \in \mathscr{D}(A^0).$$

We can apply Theorem 4.1 if we show that $(\mathscr{D}(A^0), A^0)$ satisfy the conditions of that theorem.

Since $C_0^\infty(\mathbb{R}^n)$ separates points in $\mathbb{R}^n$ and $\mathscr{D}$ separates points in H_{-p_2}, we get that $\mathscr{D}(A^0)$ separates points in $[0, T] \times H_{-p_2}$. Also, $(A^0, \mathscr{D}(A^0))$ satisfy the *separability condition*, namely, $\exists\ \{F_n\}_{n \ge 1} \subset C_0^1([0, T]) \otimes \mathscr{D}$ such that bp-closure$\{(F_n, A^0 F_n)\colon\ n \ge 1\} \supseteq \{(F, A^0 F)\colon\ F \in C^1([0, T]) \otimes \mathscr{D}\}$. This condition follows since $C_0^\infty(\mathbb{R}^n)$ is separable under the metric given by the norm $|||f||| = \|f\| + \|\nabla f\|$ and since Φ is separable. Condition C1, with $\Theta(v) = 1 + \|v\|_{-p_2}^2$, follows from condition (S3).

Since any solution of the $D([0, T], H_{-p_2})$ martingale problem for (A_s), considered as a Φ'-valued process, is a solution to the martingale problem for (A'_s), Theorem 6.1 tells us that such a solution has to be necessarily unique. The existence of a solution once again follows from Theorem 6.1 and Itô's formula. Further it is well known that X is a solution to the martingale problem for (A_s) if and only if (s, X_s) is a solution to the martingale problem for A^0.

Hence to prove the theorem we only need to prove condition C3. First note that, for any $\phi \in \Phi$,

$$\begin{aligned} f(X_t[\phi]) &- \int_0^t f'(X_s[\phi]) b(s, X_s)[\phi]\, ds \\ &+ \int_U \{ f(X_s[\phi] + G(s, X_s, u)[\phi]) \\ &\qquad - f(X_s[\phi]) - G(s, X_s, u)[\phi] f'(X_s[\phi])\} \mu(du) \end{aligned}$$

is a martingale for every $f \in C_0^\infty(\mathbb{R})$. It follows that there exists a cadlag modification, say Y_t^ϕ, of $X_t[\phi]$ [see Theorem 4.3.8 of Ethier and Kurtz (1986), page 179]. Now as in Hardy, Kallianpur, Ramsubramanian and Xiong (1994) we can show that

$$M_t^{\phi} = Y_t^{\phi} - Y_0^{\phi} - \int_0^t b(s, X_s)[\phi]\, ds$$

is a martingale with the quadratic variation process given by

$$\langle M^{\phi} \rangle_t = \int_0^t \int_U (G(s, X_s, u)[\phi])^2 \mu(du)\, ds.$$

Let $M_t^j = M_t^{\phi_j^{p_2}}$. Then we have

$$\begin{aligned}
E \sup_{t \le T} \left\| \sum_{j=m}^{n} M_t^j \phi_j^{-p_2} \right\|_{-p_2}^2 &= E \sup_{t \le T} \sum_{j=m}^{n} (M_t^j)^2 \\
&\le 4E \sum_{j=m}^{n} \langle M^j \rangle_T \\
&= 4 \sum_{j=m}^{n} E \int_0^T \int_U \left(G(s, X_s, u)\left[\phi_j^{p_2}\right]\right)^2 \mu(du)\, ds \qquad (6.5)\\
&\le 4T \sum_{j=m}^{n} E \sup_{s \le T} \int_U \|G(s, X_s, u)\|_{-p_1}^2 \|\phi_j^{p_2}\|_{p_1}^2 \mu(du) \\
&\le 4kTE \sup_{s \le T} \left(1 + \|X_s\|_{-p_1}^2\right) \sum_{j=m}^{n} \|\phi_j^{p_2}\|_{p_1}^2 \\
&\to 0 \quad \text{as } m, n \to \infty,
\end{aligned}$$

where k is as in Assumption S. Hence $\sum_{j=1}^{n} M_t^j \phi_j^{-p_2}$ converges uniformly in $[0, T]$ a.s. to, say, M_t in H_{-p_2}. Clearly M has cadlag paths. It can be checked that, for every $\phi \in \Phi$,

$$M_t[\phi] = M_t^{\phi} = X_t[\phi] - X_0[\phi] - \int_0^t b(s, X_s)[\phi]\, ds \quad \text{a.s.,}$$

for every t. Define

$$\tilde{X}_t = M_t + X_0 + \int_0^t b(s, X_s)\, ds.$$

Then $\tilde{X} \in D([0, T], H_{-p_2})$ and is a modification of X. The last assertion follows from Theorem 4.1. □

We now consider Φ'-valued diffusions. Later on in this section we will consider an example where a sequence of Poisson driven SDE's converges to a diffusion process. Let W be a Φ'-valued process such that, for every $\phi \in \Phi$, $W.[\phi]$ is a real-valued Wiener process with $E|W_t[\phi]|^2 = tQ(\phi, \phi)$, where Q is a continuous bilinear form on $\Phi \times \Phi$; W is then called a Φ'-valued Wiener process with covariance Q. Consider

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \Sigma(s, X_s)\, dW_s, \qquad (6.6)$$

where b: $\mathbb{R}_+ \times \Phi' \to \Phi'$ and Σ: $\mathbb{R}_+ \times \Phi' \to \mathscr{L}(\Phi', \Phi')$ are two measurable mappings.

The coefficients b and Σ are assumed to satisfy the following conditions (Assumption D), which are taken from Kallianpur, Mitoma and Wolpert (1990). Let $\Sigma_s^*(v)$: $\Phi \to \Phi$ be the adjoint of the operator $\Sigma_s(v)$ and

$$|Q_{\Sigma_s(v)}|_{-p,-p} = \sum_j Q\big(\Sigma_s^*(v)\phi_j^p, \Sigma_s^*(v)\phi_j^p\big).$$

ASSUMPTION D. For any $T > 0$, $\exists\ p_0 = p_0(T)$ such that, for all $p \geq p_0$, $\exists\ q \geq p$ and a constant $K = K(p, q, T)$ satisfying the following conditions:

(D1) *Continuity*. For all $t \in [0, T]$, $v, v_1, v_2 \in H_{-p}$, $b(t, v) \in H_{-q}$ and $\Sigma(t, v_1)(v_2) \in H_{-p}$. Furthermore, for t fixed, $b(t, v)$ and $|Q_{\Sigma_t(v_1)-\Sigma_t(v_2)}|_{-p,-p}$ are continuous in v, v_1 and v_2.

(D2) *Coercivity*. For all $t \in [0, T]$ and $\phi \in \Phi$,

$$2b(t,\phi)\big[\theta_p(\phi)\big] \leq K\big(1 + \|\phi\|_{-p}^2\big).$$

(D3) *Growth*. For all $t \in [0, T]$ and $v \in H_{-p}$,

$$\|b(t,v)\|_{-q}^2 \leq K\big(1 + \|v\|_{-p}^2\big) \quad \text{and} \quad |Q_{\Sigma_t(v)}|_{-p,-p} \leq K\big(1 + \|v\|_{-p}^2\big).$$

(D4) *Monotonicity*. For all $t \in [0, T]$ and $v_1, v_2 \in H_{-p}$,

$$2\big\langle b(t,v_1) - b(t,v_2), v_1 - v_2\big\rangle_{-q} + |Q_{\Sigma_t(v_1)-\Sigma_t(v_2)}|_{-q,-q} \leq K\|v_1 - v_2\|_{-q}^2.$$

(D5) *Initial*. There exists an index r_0 such that $E\|X_0\|_{-r_0}^2 < \infty$.

We further assume that $b(s, v)[\phi]$ and $Q(\Sigma_s^*(v)[\phi_1], \Sigma_s^*(v)[\phi_2])$ are continuous in s and v for every fixed ϕ, ϕ_1, $\phi_2 \in \Phi$. Recall that

$$\mathscr{D} = \{F\colon H_{-p_2} \to \mathbb{R} \mid \exists\ n \geq 1, f \in C_0^\infty(\mathbb{R}^n) \text{ and } \phi_1, \dots, \phi_n \in \Phi \text{ such that } F(v) = f(v[\phi_1], \dots, v[\phi_n])\}.$$

For $F \in \mathscr{D}$ with $F(v) = f(v[\phi_1], \dots, v[\phi_n])$ define

$$\begin{aligned} B_s F(v) = {} & \sum_{i=1}^n f_i(v[\phi_1], \dots, v[\phi_n]) b(s,v)[\phi_i] \\ & + \tfrac{1}{2} \sum_{i,j=1}^n f_{ij}(v[\phi_1], \dots, v[\phi_n]) Q\big(\Sigma_s^*(v)[\phi_i], \Sigma_s^*(v)[\phi_j]\big). \end{aligned} \tag{6.7}$$

Here f_i and f_{ij} denote the partial derivatives of f. Then, arguing as for the martingale problem for (A_s) and using Lemma 2.1 of Kallianpur and Xiong (1995), we get that the $D([0, T], H_{-p_2})$ martingale problem for (B_s) is well posed. We choose p_2 such that the canonical injection of H_{-p_1} to H_{-p_2} is Hilbert–Schmidt. Using the growth conditions (D3) and proceeding as in (6.5) we can show that in fact the martingale problem for (B_s) is well posed in the class of progressively measurable solutions; that is, C5 holds [see also Bhatt and Karandikar (1993a)]. Hence we have the following theorem, which is similar to Theorem 6.2.

THEOREM 6.3. *Let X defined on $(\Omega, \mathscr{F}, P)$ be the solution of the $D([0,T], H_{-p_2})$ martingale problem for (B_s), and let h and Y be as in* (1.1) *and* (1.2). *If $\{\rho_t\}$ is an $\mathscr{F}_t^Y$-adapted $\mathscr{M}_+(H_{-p_2})$-valued process satisfying* (3.6) *and*

$$\langle \rho_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \rho_s, B_s f\rangle\, ds + \sum_{i=1}^k \int_0^t \langle \rho_s, h^i f\rangle\, dY_s^i, \qquad \forall f \in \mathscr{D}, \tag{6.8}$$

then $\rho_t = \sigma_t$ for all $0 \le t \le T$ a.s., where σ_t is defined by (1.5) *and* (1.8).

We now consider a model for environmental pollution which is studied in Kallianpur and Xiong (1994). The linear problem for filtering is studied there. Here we continue to assume the nonlinear model (1.1), where h satisfies (1.2).

6.1. *Water pollution problem with a tolerance level.* Suppose that undesired chemicals are deposited in a river in terms of Poisson streams. Let $\mathscr{X} = [0, l]$ denote the river. The chemicals are deposited at random times $\tau_1(\omega) < \tau_2(\omega) < \cdots$ with random magnitudes $V_1(\omega), V_2(\omega), \dots$. For the sake of simplicity we will assume that the changes in chemical concentration do not depend on the locations where the chemicals are deposited. For $C \subset \mathbb{R}_+$, let

$$N([0,t] \times C) = \sum_{j:\, \tau_j \le t} I_C(V_j(\omega)).$$

Under the assumption that V_j, $j = 1, 2, \dots,$ are i.i.d. random variables, N is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity measure μ on $\mathbb{R}_+$. Let $D > 0$, V and α be constants. The chemical concentration is denoted by X. We also suppose that there is a mechanism to clean up the river when the chemical density passes a fixed level $\xi(x)$.

We regard $X(t, x)$ as an infinite-dimensional process X_t determined by its action on "smooth" functions ϕ in the sense

$$X_t[\phi] = \int_{\mathscr{X}} X(t, x) \phi(x) \rho(x)\, dx,$$

where $\rho(x) = e^{-2cx}$, $c = V/2D$. Define the operator L on $H = L^2(\mathscr{X}, \rho(x)\, dx)$ with Neumann boundary conditions by

$$L = D \frac{\partial^2}{\partial x^2} - V \frac{\partial}{\partial x}.$$

Then L is a positive definite and self-adjoint operator.

Let λ_j and ϕ_j be defined by

$$\lambda_0 = 0, \qquad \lambda_j = D\left(c^2 + \left(\frac{j\pi}{l}\right)^2\right),$$

$$\phi_0(x) = \sqrt{\frac{2c}{1 - e^{-2cl}}}, \qquad \phi_j(x) = \sqrt{\frac{2}{l}}\, e^{cx} \sin\left(\frac{j\pi}{l}x + \alpha_j\right), \tag{6.9}$$

$$\alpha_j = \tan^{-1}\left(-\frac{j\pi}{lc}\right), \qquad j = 1, 2, \dots .$$

The λ_j and ϕ_j are, respectively, the eigenvalues and eigenfunctions of L with Neumann boundary conditions on H. For $\phi \in H$ and $r \in \mathbb{R}$, let

$$\|\phi\|_r^2 = \sum_j \langle \phi, \phi_j \rangle^2 (1 + \lambda_j)^{2r},$$

where $\langle \cdot, \cdot \rangle$ is the inner product on H. Define $\Phi = \{\phi \in H\colon \|\phi\|_r < \infty \ \forall\ r \in \mathbb{R}\}$ and let H_r be the completion of Φ w.r.t. the norm $\|\cdot\|_r$. The canonical injection from H_{r+r_1} to H_r is Hilbert–Schmidt if $r_1 > d/4$. Hence Φ is a countably Hilbertian nuclear space. We consider the following stochastic model for X:

$$\begin{aligned} X_t[\phi] &= X_0[\phi] + \int_0^t (X_s[-L\phi] - \alpha X_s[\phi])\, ds \\ &\quad + \int_0^t \int_0^\infty a(\xi[\phi] - X_{s-}[\phi]) N(ds, da). \end{aligned} \tag{6.10}$$

We use Theorem 6.1 to get existence and uniqueness of Φ'-valued solutions of (6.10). We assume that μ has finite first and second moments. The solution X belongs to $D([0,T], H_{-p})$ if $\xi \in H_{-p}$ [see Kallianpur and Xiong (1994)]. Then X is also a solution to the following martingale problem. For $F(v) = f(v[\phi_1], \dots, v[\phi_n]) \in \mathscr{D}$,

$$\begin{aligned} AF(v) &= \sum_{i=1}^n \left\{ f_i(v[\phi]) \left\{ -v[L\phi_i] - \alpha v[\phi_i] + \int_0^\infty a(\xi[\phi_i] - v[\phi_i]) \mu(da) \right\} \right\} \\ &\quad + \int_0^\infty \left\{ f(v[\phi] + a(\xi[\phi] - v[\phi])) - f(v[\phi]) \right. \\ &\qquad \left. - \sum_{i=1}^n a(\xi[\phi_i] - v[\phi_i]) f_i(v[\phi]) \right\} \mu(da), \end{aligned} \tag{6.11}$$

where $\phi = (\phi_1, \dots, \phi_n)$ and $v[\phi] = (v[\phi_1], \dots, v[\phi_n])$. Here $b(t, v)[\phi] = v[-L\phi] - \alpha v[\phi] + \int_0^\infty a(\xi[\phi] - v[\phi]) \mu(da)$ and $G(t, v, a)[\phi] = a(\xi[\phi] - v[\phi])$. It can be checked that (b, G, μ) satisfy the conditions in Assumption S.

Suppose there are k stations which are monitoring the pollution levels in the river. Once again the observations are assumed to be noisy functions of the signal X. Let

$$Y_t^i = \int_0^t h^i(X_s)\,ds + W_t^i, \qquad 1 \le i \le k. \tag{6.12}$$

The linear filtering problem was considered in Kallianpur and Xiong (1994). Theorem 6.2 is applicable and we get uniqueness of the Zakai equation

$$\langle \sigma_t, F\rangle = \langle \sigma_0, F\rangle + \int_0^t \langle \sigma_s, AF\rangle\,ds + \sum_{i=1}^k \int_0^t \langle \sigma_s, h^i F\rangle\,dY_s^i, \qquad \forall\, F \in \mathscr{D},$$

when the signal process is a solution of the SDE (6.10).

We now consider the following diffusion approximation model studied by Kallianpur and Xiong (1994) which is appropriate in a situation when pollution is emitted by factories located along a river, densely in some sense. Consider a sequence of SDE's for such pollution processes of the form

$$X_t^n = X_0^n - \int_0^t (LX_s^n + \alpha^n X_s^n)\,ds + \int_0^t \int_0^\infty a(\xi - X_{s-}^n) N^n(ds, da),$$

where N^n is a sequence of Poisson random measures on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity measure μ^n on $\mathbb{R}_+$. We impose the following assumptions:

(E1) $\alpha^n + a^n \to \alpha$ and $b^n \to \beta^2$ as $n \to \infty$, where $a^n = \int_0^\infty a\mu^n(da)$ and $b^n = \int_0^\infty a^2 \mu^n(da)$.
(E2) For any $\varepsilon > 0$, $\mu^n\{a\colon a > \varepsilon\} \to 0$ as $n \to \infty$.
(E3) There exists a sequence c^n such that $c^n \alpha^n \to \gamma$ as $n \to \infty$.
(E4) $\sup_n \int_M^\infty a^2 \mu^n(da) \to 0$ as $M \to \infty$.

For any $\phi, \psi \in \Phi$, let $Q(\phi, \psi) = \langle \phi, \psi\rangle_0$. Let $b\colon \Phi' \to \Phi'$ and $\Sigma\colon \Phi' \to \mathscr{L}(\Phi', \Phi')$ be given by

$$b(v) = -Lv - \alpha v + \gamma\xi \quad\text{and}\quad \Sigma^*(v)\phi = \beta v[\phi]\phi_0.$$

Let $V_t^n = c^n X_t^n$.

Suppose that there exists r_0 such that $\sup_n E\|V_0^n\|_{-r_0}^2 < \infty$ and $\{V_0^n\}$ converges weakly to a Φ'-valued random variable V_0. Then it was shown in Kallianpur and Xiong (1994) that V^n converges in distribution to the solution V_t of the diffusion equation

$$V_t = V_0 + \int_0^t b(V_s)\,ds + \int_0^t \Sigma(V_s)\,dW_s,$$

where W is a Φ'-valued Wiener process with covariance Q. Further, if $\xi \in H_0$ and V_0 is an H_0-valued random variable with $E\|V_0\|_0^2 < \infty$, then $V \in C([0,T], H_0)$ with

$$E \sup_{0 \le t \le T} V_t[\phi]^2 < \infty, \qquad \forall\, \phi \in \Phi. \tag{6.13}$$

It is easy to see that the coefficients b and Σ satisfy the conditions in Assumption D. We continue to assume the model (6.12). We apply Theorem 6.3 to conclude uniqueness of solution to the Zakai equation (6.8).

7. Semilinear SDE's on Hilbert spaces. Recently Ahmed and Zabczyk (1994) and Zabczyk (1994) considered the nonlinear filtering equation when the signal process X is driven by a semilinear stochastic differential equation on a Hilbert space H. They have studied the existence of a solution to the Zakai equation for the unnormalized conditional distribution of X given the observation process Y, where, as in (1.1),

$$Y_t = \int_0^t h(X_s)\,ds + W_t, \qquad 0 \le t \le T, \tag{7.1}$$

and W_t is a one-dimensional standard Brownian motion. Let W^1 be a cylindrical Brownian motion on H with covariance operator I and which is independent of W. Let $b \in H$, $F\colon H \to H$ be bounded Lipschitz, and let R and L be operators on H such that R is a bounded nonnegative operator and L is the infinitesimal generator of a strongly continuous contraction semigroup $\{S_t,\, t \ge 0\}$ on H. The signal process is given by

$$dX_t = (LX_t + F(X_t))\,dt + R^{1/2}\,dW_t^1 + b\,dW_t, \qquad X_0 = \xi. \tag{7.2}$$

Let Q denote the covariance operator of the Wiener process $B_t = W_t^1 + bW_t$. Assume that the following condition is satisfied.

(SL1) The operators

$$Q_t = \int_0^t S_r Q S_r^*\,dr, \qquad t \ge 0, \tag{7.3}$$

are trace class.

Then the equation (7.2) has a unique mild solution. [For definition of a mild solution and the proof of the above-stated fact, see DaPrato and Zabczyk (1992).]

In Ahmed and Zabczyk (1994) the uncorrelated case, namely, $b \equiv 0$ was considered. Existence and uniqueness of density-valued solutions of the Zakai equation was investigated there. Here we state the correlated case from Zabczyk (1994).

(SL2) Let

$$Q_\infty = \int_0^\infty S_r Q S_r^*\,dr \tag{7.4}$$

be a trace class operator.

In addition, let μ be the Gaussian measure with mean zero and covariance operator Q_∞. Further let

$$\begin{aligned}\mathscr{D} = \Big\{ f \in C_b^2(H)\colon &\sup_{x\in H} \|f_{xx}\|_{L^1(H)} < \infty,\ \exists\, f' \in C_b^2(H) \text{ such that} \\ &\sup_{x\in H} \|f'_{xx}\|_{L^1(H)} < \infty \text{ and } f(x) = f'(L^{-1}x),\ x \in H \Big\},\end{aligned} \tag{7.5}$$

where $C_b^2(H)$ denotes the class of functions having bounded and uniformly continuous first and second Fréchet derivatives, the two being, respectively, denoted by f_x and f_{xx}. Define the operator A_0 with domain $\mathscr{D}$ by

$$A_0 f(x) = \tfrac{1}{2}\mathrm{Tr}[Qf_{xx}(x)] + (F(x), f_x(x)) + (x, L^* f_x(x)), \tag{7.6}$$

where $(\cdot,\cdot)$ denotes the inner product in H. Let $\mathscr{V} = W_Q^{1,2}(H)$ be the completion of the space $\mathscr{D}_0$ with respect to the norm

$$\|f\|_{\mathscr{V}} = \left(\int_H \left(|f(x)|^2 + (Qf_x(x), f_x(x))\right)\mu(dx)\right)^{1/2}. \tag{7.7}$$

(SL3) Suppose that $\mathscr{V}$ can be identified with a subset of $\mathscr{H} = L^2(H, \mu)$.

The following theorem is from Zabczyk (1994).

THEOREM 7.1. *In addition to the above setup, assume that the initial measure π_0 is absolutely continuous with respect to μ having a density $p_0 \in \mathscr{H}$ and that*:

(i) Image $Q_\infty \subset \mathscr{D}(L)$ *and* $\exists\, c > 0$ *such that*

$$|(x, LQ_\infty)| \le c|Q^{1/2}|\,|Q^{1/2}y|, \qquad x, y \in H;$$

(ii) $\sup_x |Q^{-1/2}F(x)| < \infty$;
(iii) $b \in (\text{Image } Q_\infty^{1/2}) \cap (\text{Image } R^{1/2})$.

Then the density-valued Zakai equation

$$\begin{aligned} d\langle f, p_t\rangle &= \langle A_0 f, p_t\rangle\, dt \\ &\quad + \left(\langle hf, p_t\rangle + \langle (b, f_x), p_t\rangle\right) dY_t, \qquad \forall\, f \in \mathscr{D}(A_0), \end{aligned} \tag{7.8}$$

has a solution q such that $q \in C([0,T], \mathscr{H}) \cap L^2([0,T], \mathscr{V})$ a.s.

Now fix a CONS $\{\phi_i\}$ in H. Let P_n denote the orthogonal projection onto the linear span of $[\phi_1, \dots, \phi_n]$. Let

$$\begin{aligned} \mathscr{D}_0 = \{ f \in \mathscr{D} \mid \exists\, g \in C_0^2(\mathbb{R}^n) &\text{ for some } n \ge 1 \\ &\text{such that } f(x) = g((x, \phi_1), \dots, (x, \phi_n))\} \end{aligned}$$

and consider the martingale problem for the operator A_0 restricted to $\mathscr{D}_0$. Using standard arguments as in Yor (1974) and as illustrated in the proof of Theorem 3.4 above, it can be shown that any solution of the martingale problem for this restriction of A_0 is a mild solution of (7.2) and hence that the martingale problem for A_0 with domain $\mathscr{D}_0$ is well posed. Furthermore we can define the operator A as in Section 3 and show that (X, Y) is the unique solution of the martingale problem for A and that the conditions of Theorem 3.1 are also satisfied. [For a similar result, see Theorem 2.10 of Bhatt, Kallianpur, Karandikar and Xiong (1995).] Hence Theorem 3.1 can be applied to prove uniqueness of solutions for the equation

$$(7.9)\quad \langle \rho_t, f\rangle = \langle \rho_0, f\rangle + \int_0^t \langle \rho_s, A_0 f\rangle\, ds + \int_0^t \langle \rho_s, hf\rangle\, dY_s, \qquad \forall f \in \mathscr{D}_0.$$

Clearly if q_t is a solution of (7.8), then

$$\rho_t(dx) = q_t(x)\,\mu(dx)$$

is a solution of (7.9). Hence uniqueness of solution of the density-valued Zakai equation (7.8) follows. Further any solution (7.8) is automatically identified as the unnormalized conditional density of X_t given $\mathscr{F}_t^Y$.

8. Robustness of σ. We will now show the continuous dependence of σ on the operators corresponding to the signal process X in the special case when signal and noise are independent. We need to restrict to this case because the Kallianpur–Striebel Bayes formula yields an exact expression of $\langle \sigma_t, f\rangle$ as an expectation (as opposed to conditional expectation). Suppose A_0^n and A_0 are operators with common domain $\mathscr{D}$ and satisfying the conditions of Theorem 4.1, with

$$(8.1)\qquad \|A_0 f(x)\| \le C_f \Theta(x),$$

$$(8.2)\qquad \|A_0^n f(x)\| \le C_f \Theta^n(x), \qquad \forall\, n.$$

We will assume that

$$(8.3)\qquad \sup_n E_{P^n}\left(\int_0^T (\Theta^n(X_r^n))^a\, dr\right) = C_1 < \infty,$$

for some $a > 1$. Let X^n and X, defined, respectively, on $(\Omega^n, \mathscr{F}^n, P^n)$ and $(\Omega, \mathscr{F}, P)$, be solutions of the martingale problems for A_0^n and A_0. We assume that the martingale problem for A_0 is well posed. Further let $X^n \Rightarrow X$. We take the observation models to be

$$Y_t^n = \int_0^t h^n(X_s^n)\, ds + W_t^n$$

and

$$Y_t = \int_0^t h(X_s)\, ds + W_t,$$

where, for every n, X^n and W^n are independent processes defined on $(\Omega^n, \mathscr{F}^n, P^n)$. Also X and W defined on $(\Omega, \mathscr{F}, P)$ are independent; W^n and W are $\mathbb{R}^k$-valued Brownian motions and h^n, h are continuous functions from E into $\mathbb{R}^k$. We will assume also that

$$(8.4)\qquad h^n \to h \quad \text{uniformly on compact subsets of } E.$$

Let P_0 be defined by (3.17). Define P_0^n similarly with X^n in place of X, Y^n in place of Y and h^n in place of h. Suppose that

$$\sup_n E_{P^n} \exp\left\{ b\int_0^T |h^n(X_s^n)|^2 \, ds \right\} = C_2 < \infty, \tag{8.5}$$

for some $b > 1$, and

$$\sup_n \sup_{0 \le t \le T} E_{P^n} |h^n(X_t^n)|^2 = C_3 < \infty. \tag{8.6}$$

Let us note that (8.5) implies that, for some $p' > 1$,

$$\sup_n E_{P^n} \left(\frac{dP^n}{dP_0^n} \right)^{p'} \le C_2 < \infty. \tag{8.7}$$

Let σ^n and σ be the corresponding unnormalized conditional distributions [see (1.8)]. Recall that Y and Y^n are Brownian motions under P_0 and P_0^n, respectively, and that σ^n and σ can be written as

$$\sigma_t = F_t(Y), \qquad \sigma_t^n = F_t^n(Y^n)$$

[see (3.15)], where F_t and F_t^n are Wiener functionals. Further F_t and F_t^n can be defined as in (4.1). Using Skorokhod's representation theorem, get $\tilde{X}^n, \tilde{X}$ on $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$ such that $\tilde{X}^n \to \tilde{X}$ a.s. and $\mathscr{L}(\tilde{X}^n) = \mathscr{L}(X^n)$, $\mathscr{L}(\tilde{X}) = \mathscr{L}(X)$. Recall from Section 4 the definitions of Ω^0, $\hat{\Omega}$ and $\tilde{Y}$. Again let F be defined by (4.1) and, similarly, F^n by

$$\langle F_t^n(\omega^0), f \rangle = \int f(\tilde{X}_t^n(\hat{\omega})) q_t^n(\hat{\omega}, \omega^0) \, d\hat{P}(\hat{\omega}), \qquad \forall f \in C_b(E),$$

where q_t is as in (4.2) and

$$q_t^n(\hat{\omega}, \omega^0) = \exp\left\{ \sum_{i=1}^k \int_0^t h^{n,i}(\tilde{X}_s^n(\hat{\omega})) \, d\tilde{Y}_s^i(\omega^0) - \tfrac{1}{2} \sum_{i=1}^k \int_0^t \left(h^{n,i}(\tilde{X}_s^n(\hat{\omega})) \right)^2 ds \right\}.$$

Let us note that if $\tilde{P}^n$ is defined by $d\tilde{P}^n = q_T^n \, d\tilde{P}$, then defining $\tilde{X}^n$ on $\tilde{\Omega}$ in the obvious way, we get that the law of $\tilde{X}^n$ under $\tilde{P}^n$ is the same as the law of X^n under P^n. The following is our main result on robustness.

THEOREM 8.1. *We have $F^n \to F$ in Q-probability as $D([0,T], \mathscr{M}_+(E))$-valued variables.*

PROOF. The proof is divided into two steps.

Step 1. Fix $t > 0$. We will show that, for every t,

$$F_t^n \to F_t \quad \text{in } Q\text{-probability}. \tag{8.8}$$

Since $\tilde{X}^n \to \tilde{X}$ a.s., it follows from (8.4) that $h^{n,i}(\tilde{X}_s^n) \to h^i(\tilde{X}_s)$ a.s. for every $s > 0$. From (8.5) it is now clear that

$$\lim_{n \to \infty} E_{\tilde{P}} \int_0^t \left| h^{n,i}(\tilde{X}_s^n) - h^i(\tilde{X}_s) \right|^2 ds = 0 \tag{8.9}$$

and hence that

$$\int_0^t h^{n,i}(\tilde{X}_s^n)\,d\tilde{Y}_s^i \to \int_0^t h^i(\tilde{X}_s)\,d\tilde{Y}_s^i \tag{8.10}$$

in $\tilde{P}$-probability. Thus we get $q_t^n \to q_t$ in $\tilde{P}$-probability. Clearly for $f \in C_b(E)$, $f(\tilde{X}_t^n) \to f(\tilde{X}_t)$ in probability. We will now verify that q_t^n is uniformly integrable. A standard computation (using independence of $\tilde{X}^n, \tilde{Y}$) shows that, for $p > 1$,

$$\begin{aligned} E_{\tilde{P}}(q_t^n)^p &= \int\left[\int (q_t^n)^p(\hat{\omega}, \omega^0)\,dQ(\omega^0)\right]d\hat{P}(\hat{\omega}) \\ &= \int \exp\left\{\left(\frac{p(p-1)}{2}\right)\sum_{i=1}^k \int_0^t \left|h^{n,i}(\tilde{X}_s^n(\hat{\omega}))\right|^2 ds\right\} d\hat{P}(\hat{\omega}). \end{aligned} \tag{8.11}$$

Thus choosing $p > 1$ such that $p(p-1)/2 < b$ and using (8.5), it follows that q_t^n is uniformly integrable. Thus

$$\lim_{n\to\infty} E_Q|\langle F_t^n, f\rangle - \langle F_t, f\rangle| = 0. \tag{8.12}$$

Since (8.12) holds for all $f \in C_b(E)$, (8.8) follows.

Step 2. *Now we will show that F^n is a tight sequence of $D([0,T], \mathscr{M}_+(E))$-valued random variables on $(\Omega^0, \mathscr{F}^0, Q)$.*

Since X^n converges weakly to X, X^n satisfies the following *compact containment condition* [see Ethier and Kurtz (1986)]. For every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset E$ satisfying

$$\inf_n P^n(X_t^n \in K_\varepsilon : 0 \le t \le T) \ge 1 - \varepsilon^{2q}, \tag{8.13}$$

where $1/q = 1 - 1/p$ for p as in (8.11). Let

$$\mathscr{K}_\varepsilon = \{\mu \in \mathscr{M}_+(E) : \mu(K^c_{\varepsilon 2^{-m}}) < \varepsilon 2^{-m}, \forall\, m \ge 1\}.$$

Then $\mathscr{K}_\varepsilon$ is compact in $\mathscr{M}_+(E)$, and recalling that X^n has the same law under P^n and under P_0^n and using (8.13) and Hölder's inequality for p and q as above, we get

$$\begin{aligned} &Q\{F_t^n \notin \mathscr{K}_\varepsilon \text{ for some } 0 \le t \le T\} \\ &\quad = Q\{F_t^n(K^c_{\varepsilon 2^{-m}}) > \varepsilon 2^{-m} \text{ for some } m \ge 1 \text{ and for some } 0 \le t \le T\} \\ &\quad \le \sum_{m=1}^\infty Q\left\{\sup_{0\le t\le T} F_t^n(K^c_{\varepsilon 2^{-m}}) > \varepsilon 2^{-m}\right\} \\ &\quad \le \sum_{m=1}^\infty \frac{1}{\varepsilon 2^{-m}} E_Q\left[\sup_{0\le t\le T} F_t^n(K^c_{\varepsilon 2^{-m}})\right] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{m=1}^{\infty} \frac{1}{\varepsilon 2^{-m}} E_Q\left[\sup_{0\le t\le T} \int I_{K^c_{\varepsilon 2^{-m}}}(\tilde{X}^n_t) q^n_t \, d\hat{P}\right] \\
(8.14)\quad &\le \sum_{m=1}^{\infty} \frac{1}{\varepsilon 2^{-m}} E_Q\left[\sup_{0\le t\le T} \left(\int I_{K^c_{\varepsilon 2^{-m}}}(\tilde{X}^n_t)\, d\hat{P}\right)^{1/q} \left(\int (q^n_t)^p \, d\hat{P}\right)^{1/p}\right] \\
&\le \sum_{m=1}^{\infty} \frac{1}{\varepsilon 2^{-m}} \left(E_{\hat{P}}\left[\sup_{0\le t\le T} I_{K^c_{\varepsilon 2^{-m}}}(\tilde{X}^n_t)\right]\right)^{1/q} E_{\tilde{P}}\left[1 + \sup_{0\le t\le T} (q^n_t)^p\right] \\
&\le \sum_{m=1}^{\infty} \frac{1}{\varepsilon 2^{-m}} \left(E_{P^n}\left[\sup_{0\le t\le T} I_{K^c_{\varepsilon 2^{-m}}}(X^n_t)\right]\right)^{1/q} E_{\tilde{P}}\left[1 + C(p)(q^n_T)^p\right] \\
&< (1 + C_2 C(p)) \sum_{m=1}^{\infty} \varepsilon 2^{-m} = (1 + C_2 C(p))\varepsilon.
\end{aligned}
$$

Here we have also used Doob's maximal inequality for the submartingale $(q^n_t)^p$ and $C(p)$ above denotes a constant depending on p. Now we will show that, for every $f \in \mathscr{D}$, $\langle F^n, f\rangle$ is tight in $D([0,T], \mathbb{R})$. This will imply [see Jakubowski (1986)] that F is relatively compact. Fix $f \in \mathscr{D}$ and let

$$
(8.15)\qquad J^n_t = \int_0^t \langle F^n_s, A^n_0 f\rangle \, ds, \qquad M^n_t = \sum_{i=1}^{k} \int_0^t \langle F^n_s, h^{n,i} f\rangle \, dY^i_s.
$$

Then

$$
(8.16)\qquad \langle F^n_t, f\rangle = \langle F^n_0, f\rangle + J^n_t + M^n_t.
$$

Let $p = (1 + p')/2$, where p' is as in (8.7). Now applying Cauchy–Schwarz and Hölder inequalities and using (8.6) and (8.7) we get that, for $\varepsilon > 0$, $\eta > 0$, we can choose $\delta > 0$ such that

$$
\begin{aligned}
&\sup_n Q\Big(\sup_{0\le t-s\le\delta} |\langle M^n\rangle_t - \langle M^n\rangle_s| > \varepsilon\Big) \\
&\quad \le \frac{1}{\varepsilon} \sup_n E_Q\left[\sup_{0\le t-s\le\delta} \int_s^t |\langle F^n_r, h^n f\rangle|^2 \, dr\right] \\
&\quad \le \frac{\|f\|^2}{\varepsilon} \sup_n E_Q\left[\sup_{0\le t-s\le\delta} \int_s^t \langle F^n_r, |h^n|\rangle^2 \, dr\right] \\
&\quad \le \frac{\|f\|^2}{\varepsilon} \sup_n E_Q\left[\sup_{0\le t-s\le\delta} \int_s^t \left(\int \left|h^n(\tilde{X}^n_r(\hat{\omega}))\right| q^n_r(\hat{\omega}, \cdot)\, d\hat{P}(\hat{\omega})\right)^2 dr\right] \\
&\quad \le \frac{\|f\|^2}{\varepsilon} \sup_n E_Q\left[\sup_{0\le t-s\le\delta} \int_s^t \left\{\left(\int \left|h^n(\tilde{X}^n_r(\hat{\omega}))\right|^2 d\hat{P}(\hat{\omega})\right)\right.\right. \\
&\qquad\qquad \left.\left. \times \left(\int (q^n_r(\hat{\omega}, \cdot))^2 \, d\hat{P}(\hat{\omega})\right)\right\} dr\right]
\end{aligned}
$$

$$\leq \frac{\|f\|^2}{\varepsilon}\delta^{(p-1)/p}\sup_n E_Q\left[\int_0^T\left\{\left(\int\left|h^n(\tilde{X}_r^n(\hat{\omega}))\right|^2 d\hat{P}(\hat{\omega})\right)^p \times\left(\int(q_r^n(\hat{\omega},\cdot))^2\,d\hat{P}(\hat{\omega})\right)^p\right\}dr\right]^{1/p}$$

$$\leq \frac{C_3^p\|f\|^2}{\varepsilon}\delta^{(p-1)/p}\left(1+\sup_n E_{\tilde{P}}\int_0^T(q_r^n)^{2p}\,dr\right)$$

$$\leq \frac{C_3^pT\|f\|^2}{\varepsilon}\delta^{(p-1)/p}\left(1+\sup_n E_{\tilde{P}}(q_T^n)^{2p}\right)$$

$$\leq \frac{C_3^pT\|f\|^2}{\varepsilon}\delta^{(p-1)/p}\left(1+\sup_n E_{P^n}\left(\frac{dP^n}{dP_0^n}\right)^{2p-1}\right)$$

$$\leq \frac{(1+C_2)C_3^pT\|f\|^2}{\varepsilon}\delta^{(p-1)/p}$$

$$< \eta.$$

Hence $\langle M^n\rangle$ is C-tight in $D([0,T],\mathbb{R})$ [see Theorem VI.4.13 of Jacod and Shiryaev (1987)]. Similarly for $\varepsilon > 0$, $\eta > 0$, using (8.3) and (8.7), for $1 < p < 2$ we can choose $\delta > 0$ such that

$$\sup_n Q\left(\sup_{0\leq t-s\leq\delta}|J_t^n - J_s^n| > \varepsilon\right)$$

$$= \sup_n Q\left(\sup_{0\leq t-s\leq\delta}\left|\int_s^t\langle F_r^n, A_0^n f\rangle\,dr\right| > \varepsilon\right)$$

$$\leq \frac{1}{\varepsilon}\sup_n E_Q\left[\sup_{0\leq t-s\leq\delta}\int_s^t\langle F_r^n, |A_0^n f|\rangle\,dr\right]$$

$$\leq \frac{C_f}{\varepsilon}\sup_n E_Q\left[\sup_{0\leq t-s\leq\delta}\int_s^t\langle F_r^n, \Theta^n\rangle\,dr\right]$$

$$= \frac{C_f}{\varepsilon}\sup_n E_Q\left[\sup_{0\leq t-s\leq\delta}\int_s^t\int\Theta^n(\tilde{X}_r^n(\hat{\omega}))q_r^n(\hat{\omega},\cdot)\,d\hat{P}(\hat{\omega})\,dr\right]$$

$$\leq \frac{C_f}{\varepsilon}\sup_n E_Q E_{\hat{P}}\left[\sup_{0\leq t-s\leq\delta}\int_s^t\Theta^n(\tilde{X}_r^n)q_r^n\,dr\right]$$

$$\leq \frac{C_f}{\varepsilon}\delta^{(p-1)/p}\sup_n E_{\tilde{P}}\left[\int_0^T\left(\Theta^n(\tilde{X}_r^n)q_r^n\right)^p dr\right]^{1/p}$$

$$\leq \frac{C_f}{\varepsilon}\delta^{(p-1)/p}\left(1+\sup_n E_{\tilde{P}}\left[\int_0^T\left(\Theta^n(\tilde{X}_r^n)q_r^n\right)^p dr\right]\right)$$

$$\leq \frac{C_f}{\varepsilon}\delta^{(p-1)/p}\left(1+\sup_n E_{\tilde{P}}\int_0^T\left(\Theta^n(\tilde{X}_r^n)\right)^p(q_T^n)^p\,dr\right)$$

$$= \frac{C_f}{\varepsilon}\delta^{(p-1)/p}\left(1 + \sup_n E_{P^n}\int_0^T (\Theta^n(X_r^n))^p \left(\frac{dP^n}{dP_0^n}\right)^{p-1} dr\right)$$

$$\le \frac{C_f T}{\varepsilon}\delta^{(p-1)/p}$$

$$\times\left(1 + \sup_n \left[E_{P^n}\int_0^T (\Theta^n(X_r^n))^{p/2-p}\, dr\right]^{2-p}\left[E_{P^n}\frac{dP^n}{dP_0^n}\right]^{p-1}\right)$$

$$\le C\delta^{(p-1)/p} < \eta,$$

where the constant C above depends on a, C_1, C_2, f and ε. Hence by Proposition VI.3.26 of Jacod and Shiryaev (1987), J^n is C-tight in $D([0,T],\mathbb{R})$. Further using Corollary VI.3.33 of the same reference we conclude that $\langle F^n, f\rangle$ is C-tight in $D([0,T],\mathbb{R})$. Thus as remarked earlier we get that F^n is relatively compact.

Now convergence of F^n to F in Q-probability follows from Steps 1 and 2 and the following lemma.

LEMMA 8.2. *Let ξ^n, ξ be $D([0,T],E_1)$-valued random variables on $(\Omega_1, \mathscr{F}_1, P_1)$, where E_1 is a complete, separable metric space. Suppose*

$$\xi_t^n \to \xi_t \quad \text{in } P_1\text{-probability for each } t$$

and ξ^n is tight. Then $\xi^n \to \xi$ in P_1-probability.

PROOF. Let d_0 be a complete metric on $D([0,T],E_1)$ for the Skorokhod topology. For processes η_1, η_2 let $d(\eta_1,\eta_2) = E_{P_1}(d_0(\eta_1,\eta_2)\wedge 1)$. Then it is well known that $d(\xi^n,\xi)\to 0$ if and only if $\xi^n \to \xi$ in P_1-probability.

Consider $\eta_t^n = (\xi_t^n, \xi_t)$ as an $E_1\times E_1$-valued process. Tightness of ξ^n implies that η^n is tight. Clearly for $t_1,\dots,t_k \ge 0$, $(\eta_{t_1}^n,\dots,\eta_{t_k}^n) \to_{P_1} (\eta_{t_1},\dots,\eta_{t_k})$, where $\eta_t = (\xi_t,\xi_t)$. Hence η^n converges in distribution to η. Clearly for any $\varepsilon > 0$, the set

$$G_\varepsilon = \{(\alpha_1,\alpha_2)\in D([0,T],E_1)\times D([0,T],E_1) \colon d_0(\alpha_1,\alpha_2) > \varepsilon\}$$

is a $P_1\circ\eta^{-1}$ continuity set and, hence,

$$P_1(\eta^n \in G_\varepsilon) \to P_1(\eta\in G_\varepsilon),$$

that is,

$$P_1(d_0(\xi^n,\xi) > \varepsilon) \to P_1(d_0(\xi,\xi) > \varepsilon) = 0;$$

hence, $\xi^n \to \xi$ in P_1-probability as $n\to\infty$. □

9. The normalized conditional distribution. All the results presented until now have been with regard to the unnormalized conditional distribution σ [cf. (1.8)]. Similar results regarding π [cf. (1.5)] can be proved. For the sake of completeness, we state the results in this section.

THEOREM 9.1. *Suppose (X,Y), defined on $(\Omega,\mathscr{F},P)$ is related by the model* (1.1). *Suppose that h is a continuous function satisfying* (1.2). *Let A_0*

and D be as before and suppose that A, defined by (3.2), *satisfies the conditions* C1–C5. *Further suppose that $\mathscr{D}(A)$ separates points in $E \times \mathbb{R}^k$.*

Let $\{\mu_t\} \subset \mathscr{P}(E)$ be an $\mathscr{F}_t^Y$-adapted cadlag process which is a solution of the FKK equation

$$\begin{aligned}\langle \mu_t, f\rangle = \langle \mu_0, f\rangle + \int_0^t \langle \mu_s, A_0 f\rangle\, ds \\ + \sum_{i=1}^k \int_0^t \big(\langle \mu_s, h^i f + D^i f(\cdot, Y_s)\rangle - \langle \mu_s, h^i\rangle\langle \mu_s, f\rangle\big)\, dI_s^{\mu^i}, \\ \forall f \in \mathscr{D}(A_0),\end{aligned} \tag{9.1}$$

where

$$I_t^{\mu^i} = Y_t^i - \int_0^t \langle \mu_s, h^i\rangle\, ds.$$

Then, if $\{\rho_t\} \subset \mathscr{M}_+(E)$ defined by

$$\begin{aligned}\langle \rho_t, f\rangle = \langle \mu_t, f\rangle \exp\left\{\sum_{i=1}^k \int_0^t \langle \mu_s, h^i\rangle\, dY_s^i - \tfrac{1}{2}\sum_{i=1}^k \int_0^t |\langle \mu_s, h^i\rangle|^2\, ds\right\}, \\ \forall f \in C_b(E),\end{aligned} \tag{9.2}$$

satisfies (3.6), *then $\mu_t = \pi_t$ for all $t \le T$ a.s.*

PROOF. An application of the Itô formula shows that ρ defined by (9.2) satisfies the Zakai equation (1.9). Theorem 3.1 is applicable and it implies that $\rho_t = \sigma_t$, for all $t \le T$ a.s. Since $\langle \mu_t, f\rangle = \langle \rho_t, f\rangle / \langle \rho_t, 1\rangle$ and $\langle \pi_t, f\rangle = \langle \sigma_t, f\rangle / \langle \sigma_t, 1\rangle$, the result is proved. □

The next result is an analogue of Theorem 3.4 and can be proved similarly. However, first we need to define a "martingale operator" corresponding to π. Let $\mathscr{D}$ be defined by (3.29). Then for $F \in \mathscr{D}$ define

$$\begin{aligned}\mathscr{A}' F(\mu) = \sum_{i=1}^n g_i(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle)\langle \mu, A_0 f_i\rangle \\ + \tfrac{1}{2} \sum_{i,j=1}^n \sum_{l=1}^k g_{ij}(\langle \mu, f_1\rangle, \dots, \langle \mu, f_n\rangle) \\ \times \big[\langle \mu, h^l f_i + D^l f_i\rangle - \langle \mu, h^l\rangle\langle \mu, f_i\rangle\big] \\ \times \big[\langle \mu, h^l f_j + D^l f_j\rangle - \langle \mu, h^l\rangle\langle \mu, f_j\rangle\big],\end{aligned} \tag{9.3}$$

where D is given by (1.4). We will once again assume that D is an operator on $C_b(E)$. It is easy to see that π is a solution to the martingale problem for $\mathscr{A}'$. We also define a weak solution of the FKK equation analogous to Definition 3.1.

DEFINITION 9.1. A $\mathscr{P}(E)$-valued process μ defined on some probability space $(\Omega^*, \mathscr{F}^*, P^*)$ is a weak solution of the FKK equation if there exists an

$\mathbb{R}^k$-valued $(\mathscr{G}_t)$-Wiener process I^μ (defined possibly on an extended probability space), μ is $(\mathscr{G}_t)$-adapted and (μ, I^μ) satisfies (9.1).

THEOREM 9.2. *Suppose that the conditions of Theorem* 3.1 *are satisfied by A. Let* π_0 *satisfy* (3.5). *Suppose that* (3.6) *holds.*

Then any solution $\{\rho_t\}$ *of the* $D([0,T], \mathscr{P}(E))$ *martingale problem for* $(\mathscr{A}, \delta_{\pi_0})$ *and satisfying*

$$E\int_0^T \Xi(\rho_s)\,ds < \infty$$

is a weak solution of the FKK equation.

We now present a related concept that was introduced in Kurtz and Ocone (1988).

DEFINITION 9.2. $(\mu, U) \in D([0,T], \mathscr{P}(E) \times \mathbb{R}^k)$ is a solution to the filtered martingale problem for A if:

(i) μ is $\mathscr{F}^U_{t+}$-adapted;

(ii) $\int_0^t E\langle \mu_s, \Theta'(\cdot, U_s)\rangle\,ds < \infty$ for all t;

(iii) for all $f \in \mathscr{D}(A)$,

$$\langle \mu_t, f(\cdot, U_t)\rangle - \int_0^t \langle \mu_s, Af(\cdot, U_s)\rangle\,ds$$

is an $(\mathscr{F}^U_{t+})$ martingale.

Then we have the following theorem, which is an improvement of Theorem 3.3 in Kurtz and Ocone (1988) and can be proved using Theorem 2.1. The main idea of this proof was also used in the proof of Theorem 3.1.

THEOREM 9.3. *Let* $\mathscr{D}(A)$ *be an algebra that separates points in* $E \times \mathbb{R}^k$ *and contains constant functions. Suppose A satisfies conditions* C1–C5.

Let $\{\mu_t\} \subset \mathscr{P}(S)$ *satisfy*

$$E\int_0^T \langle \Theta, \mu_s\rangle\,ds < \infty.$$

If (μ_t, U_t) *is a solution to the filtered martingale problem for A with* $\mu_0 = \pi_0$, *then* (μ, U) *has the same distribution as* (π, Y).

Finally we state the result for robustness of the optimal filter. Let A_0^n, A_0, h^n, h and (X^n, Y^n), (X, Y) be as in Section 8 and let (8.1)–(8.5) be satisfied. We will continue to assume that the martingale problem for (A_0) is well posed and that $X^n \Rightarrow X$. Let π^n and π be the respective normalized conditional distributions. Recall that these can be expressed as functionals of the observation processes; that is,

$$\pi_t^n = H_t^n(Y^n), \qquad \pi_t = H_t(Y).$$

Then clearly

$$H_t^n = \frac{F_t^n}{\langle F_t^n, 1 \rangle}, \qquad H_t = \frac{F_t}{\langle F_t, 1 \rangle}. \tag{9.4}$$

Note that, in view of (1.8),

$$\inf_t \langle F_t^n(Y^n), 1 \rangle > 0, \qquad \inf_t \langle F^t(Y), 1 \rangle > 0 \quad \text{a.s.}$$

Let $(\Omega^0, \mathscr{F}^0, Q)$ be the canonical Wiener space as in the previous section.

THEOREM 9.4. (a) *$H^n \to H$ in Q-probability as $D([0,T], \mathscr{P}(E))$-valued random variables.*
(b) $P^n \circ (\pi^n)^{-1} \Rightarrow P * (\pi)^{-1}$.

PROOF. The first part follows immediately from Theorem 8.1 and (9.4). For (b) note that, for any $G \in C_b(D([0,T], \mathscr{P}(E)))$,

$$\begin{aligned} E_{P^n}[G(\pi^n)] &= E_{P^n}[G(H^n(Y^n))] \\ &= E_{\tilde{P}}[G(H^n) q_T^n] \\ &\to E_{\tilde{P}}[G(H) q_T] \\ &= E_P[G(H(Y))] \\ &= E_P[G(\pi)]. \end{aligned}$$

Here q_T^n and q_T are as in the previous section and the convergence above follows from part (a) and the fact that q_T^n converge to q_T in $L^1(\tilde{P})$. This now completes the proof. □

While this paper was being written, we received a copy of a paper by Professor E. M. Goggin, which she kindly sent to us. In this paper [Goggin (1994)] the question that is addressed is the convergence of $E[U^n|V^n]$ to $E[U|V]$ when $(U^n, V^n) \to_d (U, V)$. As an application of these results to filtering, the author considers the setup of Di Masi and Runggaldier (1982), where the signal process is a one-dimensional diffusion and the function h is bounded (and the noise appearing in the observation is a mixture of Gaussian and Poisson noise). In this case, the author deduces that π_t^n converges in law to π_t for each t. It is also mentioned that convergence in law of the process π^n to π can be deduced, but no indication of proof is given.

It may be noted that here we have proved convergence of H^n to H in probability on Wiener space, and our signal is a general E-valued Markov process; the functions h^n may depend on n and need not be bounded.

REFERENCES

AHMED, N. U. and ZABCZYK, J. (1994). Nonlinear filtering of semilinear stochastic differential equations on Hilbert spaces. Preprint 522, Institute of Mathematics, Polish Academy of Sciences.

BARAS, J. S., BLANKENSHIP, G. L. and HOPKINS, W. E. (1983). Existence, uniqueness and asymptotic behaviour of solutions to a class of Zakai equations with unbounded coefficients. *IEEE Trans. Automat. Control* **28** 203–214.

BHATT, A. G., KALLIANPUR, G., KARANDIKAR, R. L. and XIONG, J. (1995). On Hilbert space valued diffusions. *Appl. Math. Optim.* To appear.

BHATT, A. G. and KARANDIKAR, R. L. (1993a). Invariant measures and evolution equations for Markov processes characterized via martingale problems. *Ann. Probab.* **21** 2246–2268.

BHATT, A. G. and KARANDIKAR, R. L. (1993b). Evolution equations for Markov processes: applications to the white noise theory of filtering. *Appl. Math. Optim.* **31** 327–348.

CHALEYAT-MAUREL, M., MICHEL, D. and PARDOUX, E. (1990). Un théorème d'unicité pour l'équation de Zakai. *Stochastics Stochastics Rep.* **29** 1–12.

DAPRATO, G. and ZABCZYK, J. (1992). Stochastic equations in infinite dimensions. *Encyclopedia of Mathematics and Its Applications* **45**. Cambridge Univ. Press.

DELLACHERIE, C. and MEYER, P. A. (1978). *Probabilities and Potential.* North-Holland, Amsterdam.

DI MASI, G. B. and RUNGGALDIER, W. J. (1982). On approximation methods for nonlinear filtering. *Nonlinear Filtering and Stochastic Control. Lecture Notes in Math.* **972** 249–259. Springer, Berlin.

ELLIOTT, R. J. (1982). *Stochastic Calculus and Applications.* Springer, New York.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

FUJISAKI, M., KALLIANPUR, G. and KUNITA, H. (1972). Stochastic differential equations for the nonlinear filtering problem. *Osaka J. Math.* **9** 19–40.

GOGGIN, E. M. (1994). Convergence in distribution of conditional expectations. *Ann. Probab.* **22** 1097–1114.

HARDY, G., KALLIANPUR, G., RAMSUBRAMANIAN, S. and XIONG, J. (1994). The existence and uniqueness of solutions of nuclear space valued SDE's driven by Poisson random measures. *Stochastics Stochastics Rep.* **50** 85–122.

HIJAB, O. (1989). On partially observed control of Markov processes. *Stochastics Stochastics Rep.* **28** 123–144.

HOROWITZ, J. and KARANDIKAR, R. L. (1990). Martingale problems associated with the Boltzmann equation. In *Seminar on Stochastic Processes, 1989* (E. Çinlar et al., eds.) 75–122. Birkhäuser, Boston.

IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

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

JAKUBOWSKI, A. (1986). On the Skorokhod topology. *Ann. Inst. H. Poincaré Sec. B* **22** 263–285.

KALLIANPUR, G. (1980). *Stochastic Filtering Theory.* Springer, New York.

KALLIANPUR, G. and KARANDIKAR, R. L. (1984). The nonlinear filtering problem for the unbounded case. *Stochastic Process. Appl.* **18** 57–66.

KALLIANPUR, G. and KARANDIKAR, R. L. (1988). *White Noise Theory of Prediction, Filtering and Smoothing.* Gordon and Breach, New York.

KALLIANPUR, G., MITOMA, I. and WOLPERT, R. L. (1990). Diffusion equations in duals of nuclear spaces. *Stochastics Stochastics Rep.* **29** 1–45.

KALLIANPUR, G. and XIONG, J. (1994). Stochastic models of environmental pollution. *Adv. in Appl. Probab.* **26** 377–403.

KALLIANPUR, G. and XIONG, J. (1995). Diffusion approximation of nuclear space valued SDE driven by Poisson random measure. *Ann. Appl. Probab.* **5** 493–517.

KRYLOV, N. V. and ROZOVSKII, B. L. (1981). Stochastic evolution equations. *J. Sov. Math.* **16** 1233–1276.

KURTZ, T. G. and OCONE, D. L. (1988). Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.* **16** 80–107.

KUSHNER, H. (1967). Dynamical equations for optimal nonlinear filtering. *J. Differential Equations* **3** 179–190.

MOTOO, M. and WATANABE, S. (1965). On a class of additive functionals of Markov processes. *J. Math Kyoto Univ.* **4** 429–469.

PARDOUX, E. (1979). Stochastic partial differential equations and filtering of diffusion processes. *Stochastics* **3** 127–168.

PARDOUX, E. (1982). Equations du filtrage non linéaire de la prédiction et du lissage. *Stochastics* **6** 193–231.

ROZOVSKII, B. L. (1991). A simple proof of uniqueness for Kushner and Zakai equations. In *Stochastic Analysis* (E. Mayer-Wolf, E. Merzbach and A. Shwartz, eds.) 449–458. Academic Press, New York.

SHEU, S. J. (1983). Solution of certain parabolic equations with unbounded coefficients and its application to nonlinear filtering. *Stochastics* **10** 31–46.

SZPIRGLAS, J. (1978). Sur l'équivalence d'équations différentielles stochastiques à valeurs mesures intervenant dans le filtrage markovien non linéaire. *Ann. Inst. H. Poincaré Sec. B* **14** 33–59.

YOR, M. (1974). Existence et unicité de diffusions à valeurs dans un espace de Hilbert. *Ann. Inst. H. Poincaré Sec. B* **10** 55–88.

YOR, M. (1977). Sur les théories du filtrage et de la prédiction. *Séminaire de Probabilités XI. Lecture Notes in Math.* **581** 257–297. Springer, Berlin.

ZABCZYK, J. (1994). Filtering equation for infinite dimensional systems. Preprint. (Based on a paper presented at the Workshop on Stochastic Partial Differential Equations held at the University of North Carolina at Chapel Hill, June 1994.)

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

A. G. BHATT
INDIAN STATISTICAL INSTITUTE
7, S.J.S. SANSANWAL MARG
NEW DELHI, 110016
INDIA

G. KALLIANPUR
CENTER FOR STOCHASTIC PROCESSES
DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL
CB 3260, PHILLIPS HALL
CHAPEL HILL, NORTH CAROLINA 27599-3260

R. L. KARANDIKAR
INDIAN STATISTICAL INSTITUTE
7, S.J.S. SANSANWAL MARG
NEW DELHI, 110016
INDIA