isid/ms/2009/13
December 9, 2009
http://www.isid.ac.in/~statmath/eprints

Uniqueness of solution to the Kolmogorov's forward equation: Applications to White Noise Theory of Filtering

Abhay G. Bhatt
Rajeeva L. Karandikar

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Uniqueness of solution to the Kolmogorov's forward equation: Applications to White Noise Theory of Filtering

by

Abhay G. Bhatt and Rajeeva L. Karandikar

Indian Statistical Institute and *Cranes Software International Limited*

abhay@isid.ac.in and *rkarandikar@gmail.com*

Abstract

We consider a signal process X taking values in a complete, separable metric space E. X is assumed to be a Markov process charachterized via the martingale problem for an operator A. In the context of the finitely additive white noise theory of filtering, we show that the optimal filter $\Gamma_t(y)$ is the unique solution of the analogue of the Zakai equation for every y, not necessarily continuous. This is done by first proving uniqueness of solution to a (perturbed) measure valued evolution equation associated with A. An additional assumption of uniqueness of the local martingale problem for A is imposed.

AMS 2000 *subject classification:* Primary 60G35 Secondary 60J35, 60G44

Key words and phrases: Zakai equation, Markov process, Martingale problem, Evolution equation

1 Introduction

The white noise approach to filtering theory was developed extensively by Kallianpur and Karandikar during the 1980s. A comprehensive account of the theory can be found in their book [10]. (See also [9].) In their set-up, the signal process X was assumed to be a Markov process, defined on some (countably additive) probability space $(\Omega, \mathcal{F}, P)$, and taking values in a complete separable metric space E while the additive noise was modelled as a white noise which exists only on a finitely additive probability space.

Uniqueness of solution of the analogue of the Zakai equation in this context was proved in [8] and [11]. The unique solution is indeed the (unnormalized) optimal filter. However, in this equation the class of test functions was $D(L)$, the domain of the strong generator of the Markov process X.

In [2] and [3], the question of uniqueness was proved via a different approach. Following on the results of [6], a sufficient condition for invariant measures for the Markov process X was proved in terms of an operator A which charachterizes X via martingale problems. To be precise, it was assumed that X is the unique solution of the martingale problem for A. This result in turn was used to prove uniqueness of solution for the following (probability) measure valued evolution equation for A.

$$\int f d\mu_t = \int f d\mu_0 + \int_0^t \left(\int Af d\mu_s \right) ds \qquad \forall f \in \mathcal{D}(A). \tag{1.1}$$

The above is the *weak* version of the Kolmogorov's forward equation for A. Later a perturbed evolution equation was considered. For a non-negative function λ on E uniqueness of solution to the (positive) measure valued equation

$$\int f d\rho_t = \int f d\rho_0 + \int_0^t \left(\int (Af - \lambda f) d\rho_s \right) ds \qquad \forall f \in \mathcal{D}(A) \tag{1.2}$$

was proved. The class of test functions in equations (1.1), (1.2) is $\mathcal{D}(A)$ -the domain of the operator A for which the martingale problem is assumed to be well posed. $\mathcal{D}(A)$ can be much smaller than the domain of the strong generator of the Markov Process X. In the white noise theory of filtering, the Zakai equation is similar to (1.2) where the perturbation λ appears in terms of the observation function h.

In the above mentioned approach it was assumed throughout that $\mathcal{D}(A) \subset C_b(E)$, the space of real valued bounded continuous functions on E and that for every $f \in \mathcal{D}(A)$, Af is continuous. The function λ was also assumed to be continuous. In [2] the results were proved under the additional assumption that Af and λ are bounded functions and were extended to the case of unbounded functions in [3]. The corresponding results on uniqueness of solution to the Zakai equation were proved under the assumptions, respectively, of boundedness and continuity of the observation function h and later extended to a general continuous h satisfying the energy condition.

In both these articles the characterization of the optimal filter as unique solution of the Zakai equation was proved for all *continuous* observation paths y.

Subsequently, the results on invariant measures and evolution equations in [2] and [3] were extended to allow discontinuous, and unbounded, Af in [4], [5]. These in turn were used to prove uniqueness results for the Zakai equation in the *classical* non-linear filtering theory in a fairly general set-up ([4], [1]). However, in this set-up, the results on uniqueness of solution to the unperturbed equation (1.1) sufficed as they were applied to the operator B which characterized the signal-observation pair (X, Y) in terms of martingale problems.

In this article, with a view to applications in the white noise theory of filtering we consider the perturbed evolution equation (1.2) when λ is not only unbounded but also discontinuous. We can no longer use conditioning arguments as in [2] or [3] to prove uniqueness (on a bigger state space) of the martingale problem for the perturbed operator $A - \lambda$. We circumvent this problem by using a suitable change of measure. However, this necessitates an extra assumption on the operator A, *viz.*, that the *local martingale problem* for A is well-posed.

The relevant terminology of martingale problems and the main definitions are given in the next section and the uniqueness of the perturbed evolution equation is proved there. In the last section, this is applied to get uniqueness of solution to the Zakai equation in the context of the white noise theory of filtering.

2 Perturbed Evolution Equation

Throughout this article (E, d) denotes a complete, separable metric space, $C_b(E)$, the space of real valued bounded continuous functions on E, $C(E)$, the space of real valued continuous functions on E, $M(E)$, the class of all real valued Borel measurable functions on E, $\mathcal{B}(E)$ the Borel σ-field on E, $\mathcal{P}(E)$ the space of probability measures on E and $\mathcal{M}_+(E)$ the space of positive finite measures on E.

A denotes an operator with domain $\mathcal{D}(A) \subset C_b(E)$ and with range contained in $M(E)$. $\mathbf{1}$ will denote the constant function taking value 1 while $\mathbf{1}_F$ will denote the indicator function of the set F. For $C \subset M(E)$, we define the *bp*-closure of C to be the smallest subset of $M(E)$ containing C which is closed under bounded pointwise convergence of sequences of functions.

Recall that an operator A is said to satisfy the maximum principle if for $f \in \mathcal{D}(A)$, $x_0 \in E$ is such that $f(x_0) = \sup_{y \in E} f(y)$ then $Af(x_0) \leq 0$.

Let us impose the following conditions on an operator A.

Hypothesis 2.1 *$\mathcal{D}(A) \subset C_b(E)$ is an algebra, $\mathbf{1} \in \mathcal{D}(A)$, $A\mathbf{1} = 0$ and $\mathcal{D}(A)$ separates points in E.*

Hypothesis 2.2 *$A : \mathcal{D}(A) \to M(E)$ is an operator satisfying the maximum principle.*

Hypothesis 2.3 *There exists a complete separable metric space U, an operator $\hat{A} : \mathcal{D}(A) \to C(E \times U)$ and a transition function η from $(E, \mathcal{B}(E))$ into $(U, \mathcal{B}(U))$ such that*

$$(Af)(x) = \int_U \hat{A}f(x, u)\eta(x, du). \tag{2.3}$$

Hypothesis 2.4 *There exists $\hat{\Phi} \in C(E \times U)$ such that for all $f \in \mathcal{D}(A)$, there exists $C_f < \infty$ satisfying*

$$|\hat{A}f(x,u)| \le C_f \hat{\Phi}(x,u) \quad \forall (x,u) \in E \times U, \tag{2.4}$$

$$\Phi(x) = \int_U \hat{\Phi}(x,u)\eta(x,du) < \infty \quad \forall x \in E. \tag{2.5}$$

Note that the above hypotheses imply that

$$|Af(x)| \le C_f \Phi(x) \quad \forall x \in E. \tag{2.6}$$

Hypothesis 2.5 *There exists a countable set $\{f_n : n \ge 1\} \subset \mathcal{D}(A)$ such that*

$$\textit{bp-closure}(\{(f_n, \Phi^{-1}Af_n)\} : n \ge 1) \supset \{(f, \Phi^{-1}Af : f \in \mathcal{D}(A)\}.$$

To emphasize the role of Φ, we will say that (A, Φ) satisfy Hypotheses 2.1 - 2.5.

Definition 2.1 : An E valued process $(X_t)_{0 \le t < T}$ defined on some probability space $(\Omega, \mathcal{F}, P)$ is said to be a solution to the martingale problem for (A, μ) with respect to a filtration $\{\mathcal{F}_t : 0 \le t < T\}$ if

(i) X is $\{\mathcal{F}_t\}$ - progressively measurable,

(ii) $\mathcal{L}(X_0) = \mu$,

(iii) $\mathbb{E}\int_0^T |Af(X_s)|ds < \infty \ : \ \forall f \in \mathcal{D}(A)$

and

(iv) for every $f \in \mathcal{D}(A)$, $M_t^f = f(X_t) - \int_0^t Af(X_s)ds$ is a $\{\mathcal{F}_t\}$ - martingale.

Here and in the sequel, $\mathcal{L}(Z)$ denotes the law of a random variable Z.

We state a result on uniqueness of solution to evolution equation from [4]. The following additional assumption on A, the function Φ and $\mu \in \mathcal{P}(E)$ is needed here.

Hypothesis 2.6 *If $(X_t)_{0 \le t < T}$ and $(Y_t)_{0 \le t < T}$ are solutions to the martingale problem for (A, μ) (defined possibly on different probability spaces) such that $\mathbb{E}\int_0^T \Phi(X_s)ds < \infty$ and $\mathbb{E}\int_0^T \Phi(Y_s)ds < \infty$, then the finite dimensional distributions of the two processes are the same.*

We say that $\{\nu_t : 0 \le t \le T\} \subset \mathcal{M}_+(E)$ is a measurable family if for all Borel sets B in E, $t \mapsto \nu_t(B)$ is Borel measurable.

Theorem **2.1** *Suppose (A, Φ) satisfies Hypotheses 2.1 – 2.5. Suppose $\{\mu_t : 0 \le t \le T\} \subset \mathcal{P}(E)$ is a measurable family satisfying*

$$\int_0^T \int_E \Phi(x)\mu_t(dx)dt < \infty \tag{2.7}$$

and

$$\langle f, \mu_t \rangle = \langle f, \mu_0 \rangle + \int_0^t \langle Af, \mu_s \rangle ds, \quad t \le T, \ f \in \mathcal{D}(A). \tag{2.8}$$

Then there exists a progressively measurable solution $(X_t)_{t<T}$ to the martingale problem for (A, μ_0), with

$$\mathcal{L}(X_t) = \mu_t \quad \forall t < T.$$

In particular, if (A, Φ, μ_0) also satisfies Hypothesis 2.6 then (2.8) admits a unique solution.

In order to deal with the unbounded operators and their pertubations, we introduce the notion of *local martingale problem.*

Definition 2.2 : An E valued process $(X_t)_{0\leq t<T}$ defined on some probability space $(\Omega, \mathcal{F}, P)$ is said to be a solution to the local martingale problem for (A, μ) with respect to a filtration $\{\mathcal{F}_t : 0 \leq t < T\}$ if

(i) X is $\{\mathcal{F}_t\}$ - progressively measurable,

(ii) $\mathcal{L}(X_0) = \mu$,

(iii) $\int_0^t |Af(X_s)|ds < \infty$: *a.s.* $\forall f \in \mathcal{D}(A)$ $\forall t < T$

and

(iv) for all $f \in \mathcal{D}(A)$, $M_t^f = f(X_t) - \int_0^t Af(X_s)ds$ is a $\{\mathcal{F}_t\}$ local martingale.

We introduce another assumption on (A, Φ, μ_0)

Hypothesis 2.7 *If $(X_t)_{0\leq t<T}$ and $(Y_t)_{0\leq t<T}$ are solutions to the local martingale problem for (A, μ_0) (defined possibly on different probability spaces) such that $\int_0^T \Phi(X_s)ds < \infty$ a.s. and $\int_0^T \Phi(Y_s)ds < \infty$ a.s. , then the finite dimensional distributions of the two solutions are the same.*

The following is the main result of this section and gives sufficient conditions for uniqueness of solution to the perturbed evolution equation (2.10) to hold.

Theorem **2.2** *Suppose (A, Φ) satisfies Hypotheses 2.1 – 2.5 and 2.7. Let $\mu_0 \in \mathcal{P}(E)$ and suppose that the martingale problem for (A, μ_0) admits a progressively measurable solution $(X_t^*)_{t<T}$ with*

$$\mathbb{E}\Big[\int_0^T \Phi(X_t^*)dt\Big] < \infty. \tag{2.9}$$

Let $\lambda : E \longrightarrow [0, \infty)$ be a measurable function. Then the equation

$$\langle f, \rho_t\rangle = \langle f, \mu_0\rangle + \int_0^t \langle Af - \lambda f, \rho_s\rangle ds \qquad t \leq T, : \ f \in \mathcal{D}(A) \tag{2.10}$$

admits a unique solution in the class of measurable families $\{\rho_t : 0 \leq t \leq T\} \subset \mathcal{M}_+(E)$ such that

$$\int_0^T \int_E \Phi(x)\rho_t(dx)dt < \infty. \tag{2.11}$$

Proof: We first show that there exists a solution to (2.10). Indeed, it is easy to see that $\{\rho_t : 0 \leq t \leq T\}$ defined by

$$\rho_t(B) = \mathbb{E}\big[\mathbf{1}_B(X_t^*) \exp\{-\int_0^t \lambda(X_s^*)ds\}\big] \qquad B \in \mathcal{B}(E) \tag{2.12}$$

is a measurable family, satisfies (2.11) and is a solution to (2.10).

The proof of the uniqueness is divided in several steps.

Step 1: *To convert the pertubed evolution equation into evolution equation for a suitable operator:*

Let $U, \eta, \hat{\Phi}, \Phi, \{f_n : \ n \geq 1\}$ be such that Hypotheses 2.3 – 2.5 are satisfied. We need to add a point Δ that is not in E. So take $\Delta \notin E$ and let $E^{\Delta} = E \cup \{\Delta\}$. Define a metric d' on E^{Δ} by

$$d'(\Delta, \Delta) = 0,$$
$$d'(x, \Delta) = d'(\Delta, x) = 1 \quad \forall\, x \in E$$
$$d'(x, y) = d(x, y) \wedge 1 \quad \forall\, x, y \in E.$$

Extend the functions $\{f_n, n \geq 1\}$ and λ to E^{Δ} by defining

$$f_n(\Delta) = 0,\ n \geq 1, \text{ and } \lambda(\Delta) = 0.$$

Define operators A^{Δ} and B^{Δ} as follows. Let

$$\mathcal{D}(A^{\Delta}) = \{f \in C_b(E^{\Delta}) : f|_E \in \mathcal{D}(A)\}$$

and for $f \in \mathcal{D}(A^{\Delta})$

$$A^{\Delta} f(x) = Af(x) \quad \forall x \in E$$
$$A^{\Delta} f(\Delta) = 0.$$

Let $\mathcal{D}(B^{\Delta}) = \mathcal{D}(A^{\Delta})$ and for $f \in \mathcal{D}(B^{\Delta})$ and $x \in E^{\Delta}$

$$B^{\Delta} f(x) = A^{\Delta} f(x) - \lambda(x)(f(x) - f(\Delta)).$$

It is easy to see that A^{Δ} (and Φ) satisfies Hypotheses 2.1 – 2.4 with $\Phi, \hat{\Phi}$ extended to E^{Δ} by setting $\Phi(\Delta) = 1$ and $\hat{\Phi}(\Delta, u) = 1$ and $\eta(\Delta, F) = \mathbf{1}_F(\Delta)$. Taking $f_0 = \mathbf{1}$, we can verify that Hypothesis 2.5 is also satisfied with $\{f_n : n \geq 0\}$.

As for B^{Δ}, Hypothesis 2.1 can be verified directly from the definition of B^{Δ}. Hypothesis 2.2 can be verified easily since $\lambda(x) \geq 0\ \ \forall x$. For Hypotheses 2.3 and 2.4, we consider the auxillary space $U_1 = U \times [0, \infty)$. For $x \in E$, $f \in \mathcal{D}(A^{\Delta})$ and $u_1 = (u, t) \in U_1$ define

$$\hat{A}_1 f(x, u_1) = \hat{A} f(x, u) - t(f(x) - f(\Delta)),$$
$$\hat{\Phi}_1(x, u_1) = \hat{\Phi}(x, u) + 2t, \qquad \Phi_1(x) = \Phi(x) + 2\lambda(x),$$
$$\hat{A}_1 f(\Delta, u_1) = 0, \text{ and } \hat{\Phi}_1 f(\Delta, u_1) = 1.$$

Further let $\eta_1(x, F \times G) = \eta(x, F)\mathbf{1}_G(\lambda(x))$ where F, G are Borel subsets of U and $[0, \infty)$ respectively and $x \in E^{\Delta}$. Now we can see that for $f \in \mathcal{D}(B^{\Delta})$

$$B^{\Delta} f(x) = \int_{U_1} \hat{A}_1 f(x, u_1)\eta_1(x, du_1).$$

It follows that with $U_1, \eta_1, \Phi_1, \hat{\Phi}_1, \{f_n : n \geq 0\}$ as given above (B^{Δ}, Φ_1) satisfy Hypothesess 2.3 – 2.5.

Let $\{\rho_t\}$ be a solution to (2.10) satisfying the integrability condition (2.11). Taking $f = \mathbf{1}$, and since $A\mathbf{1} = 0$ we see that

$$\rho_t(E) = \mu_0(E) - \int_0^t \langle \lambda, \rho_s \rangle ds. \tag{2.13}$$

Since $\mu_0(E) = 1$, ρ_t is a sub-probability measure. Hence it follows that μ_t defined by

$$\mu_t(F) = \rho_t(F \cap E) + \mathbf{1}_F(\Delta)(1 - \rho_t(E)) \text{ for } F \in \mathcal{B}(E^\Delta) \tag{2.14}$$

belongs to $\mathcal{P}(E^\Delta)$. Note that for $g \in C_b(E^\Delta)$ such that $g(\Delta) = 0$ we have

$$\int_0^t \langle g, \mu_s \rangle ds = \int_0^t \langle g, \rho_s \rangle ds.$$

Now

$$\langle g, \mu_t \rangle = \langle g, \mu_0 \rangle + \int_0^t \langle B^\Delta g, \mu_s \rangle ds \qquad \forall\, g \in \mathcal{D}(B^\Delta).$$

We can verify this from (2.10) first for g such that $g(\Delta) = 0$ and then use (2.13), (2.14), the fact that $\mu_t \in \mathcal{P}(E^\Delta)$ and that $B^\Delta \mathbf{1} = 0$. Further, we also have

$$\int_0^T \int_E \Phi_1(x)\mu_t(dx)dt = \int_0^T \int_E (\Phi(x) + \lambda(x))\, \rho_t(dx)dt < \infty$$

by (2.11) and (2.13). This completes the first step.

Now invoking Theorem 2.1 it follows that the martingale problem for (B^Δ, μ_0) admits a progressively measurable solution $(X_t)_{t<T}$ defined on some probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ and which satisfies $\mathcal{L}(X_t) = \mu_t$ for $t < T$. Note that this implies $E_{\widetilde{P}}\left[\int_0^T \Phi_1(X_s)ds\right] < \infty$.

The idea behind the rest of the proof is as follows. We can verify that

$$U_t = \mathbf{1}_E(X_t) \exp\left\{\int_0^t \lambda(X_s)ds\right\}$$

is a local martingale and integration by parts can be used to verify that for $f \in \mathcal{D}(A)$

$$\left(f(X_t) - \int_0^t (Af)(X_s)ds\right) U_t$$

is also a local martingale. Thus *if we can construct a probability measure* $\widetilde{Q}$ such that

$$\frac{d\widetilde{Q}}{d\widetilde{P}} = U_{\tau_n} \quad \text{on } \mathcal{F}_{\tau_n}$$

for a suitable sequence of stopping times τ_n increasing to T, we can conclude that under $\widetilde{Q}$, (X_t) is a solution of the local martingale problem for A. Then law of (X_t) is uniquely determined in view of Hypothesis 2.7. From here we can conclude that law of (X_t) under $\widetilde{P}$ is also uniquely determined. To achieve this, we will first construct a copy Z of the process X (*i.e.* Z and X having the same finite dimensional distributions) on a suitable probability space on which we can use a variant of Kolmogorov consistency theorem and thus construct $\widetilde{Q}$ as outlined above.

Step 2: *Construction of Z on a suitable space:*

Let $\{f_k : k \geq 0\} \subset \mathcal{D}(A^\Delta)$ be the countable subset constructed in step 1 above such that hypothesis 2.5 is valid for A^Δ. Without loss of generality, assume that $f_1(x) = \mathbf{1}_E(x)$. Let $a_k = \|g_k\|$. and let $\mathcal{J} : E^\Delta \to \hat{E} := \prod_{k=0}^\infty [-a_k, a_k]$ be defined by

$$\mathcal{J}(x) = (f_0(x), f_1(x), ..., f_k(x), ...)$$

A generic element of $\hat{E}$ will be denoted by ζ and $(\zeta_0, \zeta_1, \ldots, \zeta_k, \ldots)$ will denote its components. Since $\{f_k : k \geq 0\}$ separate points in E^Δ it follows that $\mathcal{J}$ is one-to-one. Hence $\mathcal{J}^{-1}$ is well-defined on $\mathcal{J}(E^\Delta)$. We extend $\mathcal{J}^{-1}$ to $\hat{E}$, (and with an abuse of notation we continue to call the extension by $\mathcal{J}^{-1}$) by setting $\mathcal{J}^{-1} = \Delta$ on the complement of $\mathcal{J}(E^\Delta)$.

Further, let

$$\hat{f}_k = f_k \circ \mathcal{J}^{-1}, \quad \hat{\lambda} = \lambda \circ \mathcal{J}^{-1}, \quad \hat{\Phi}_1 = \Phi_1 \circ \mathcal{J}^{-1},$$
$$g_k = A^\Delta f_k, \quad h_k = B^\Delta f_k = g_k - \lambda,$$
$$\hat{g}_k = g_k \circ \mathcal{J}^{-1}, \quad \hat{h}_k = h_k \circ \mathcal{J}^{-1} = \hat{g}_k - \hat{\lambda}.$$

In view of of (2.6) we have for $k \geq 0$

$$|\hat{g}_k(\zeta)| \leq C_k \hat{\Phi}_1(\zeta), \quad |\hat{h}_k(\zeta)| \leq C_k \hat{\Phi}_1(\zeta) \ \ \forall \zeta \in \hat{E} \tag{2.15}$$

for suitable constants C_k.

Let

$$\mathcal{X}_t(\widetilde{\omega}) = \mathcal{J}(X_t(\widetilde{\omega})) \text{ for } \widetilde{\omega} \in \widetilde{\Omega}.$$

Since

$$f_k(X_t) - \int_0^t B^\Delta f_k(X_s) ds$$

is a martingale, it is easy to see that each component of $\mathcal{X}_t$ admits an r.c.l.l. modification and thus $\mathcal{X}$ itself admits an r.c.l.l. modification, denoted by $\tilde{\mathcal{X}}$. Let $\widetilde{\mathcal{F}}_t = \sigma\{\tilde{\mathcal{X}}_s : s \leq t\})$. Before proceeding, let us also note that

$$M_t^k = \hat{f}_k(\tilde{\mathcal{X}}_t) - \int_0^t \hat{h}_k(\tilde{\mathcal{X}}_s) ds \tag{2.16}$$

is a $(\widetilde{\mathcal{F}}_t)$-martingale ON $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$.

Let $\tilde{X}$ be defined by

$$\tilde{X}_s(\widetilde{\omega}) = \mathcal{J}^{-1}(\tilde{\mathcal{X}}_s(\widetilde{\omega})), \text{ for } \widetilde{\omega} \in \widetilde{\Omega}.$$

It follows that $\tilde{X}$ is a version of X and so for each t, $\mathcal{L}(\tilde{X}_t) = \mu_t$.

For $0 \leq s < T$ and $\zeta \in \hat{E}$, let

$$K(s, \zeta) = \frac{T}{T-s}\left(1 + \hat{\Phi}_1(\zeta)\right).$$

For $0 \leq t < T$, $0 \leq u < \infty$, define processes (α_t), (τ_u) on $\widetilde{\Omega}$ by

$$\alpha_t(\widetilde{\omega}) = \int_0^t K(s, \tilde{\mathcal{X}}_s(\widetilde{\omega})) ds,$$
$$\tau_u(\widetilde{\omega}) = \inf\{t : \alpha_t(\widetilde{\omega}) \geq u\}.$$

Then it can be seen that (α_t) is a $(\widetilde{\mathcal{F}}_t)$ adapted, strictly increasing process and limit of $\alpha_t(\widetilde{\omega})$ as t tends to T is ∞. Further, for each $u < \infty$, τ_u is a $(\widetilde{\mathcal{F}}_t)$ stop time with $\tau_u < T$. Also

$$\tau_{\alpha_t} = t, \ \text{ and } \ \alpha_{\tau_u} = u \ \text{ for } 0 \leq t < T, \ 0 \leq u < \infty.$$

Considering (τ_u) as a random time change, define $\mathcal{Y}_u = \tilde{\mathcal{X}}_{\tau_u}$ and $\widetilde{\mathcal{G}}_u = \widetilde{\mathcal{F}}_{\tau_u}$. It then follows that $(\mathcal{Y}_u)$, (τ_u) are $(\widetilde{\mathcal{G}}_u)$ adapted, and that

$$\tau_u(\widetilde{\omega}) = \int_0^u \frac{1}{K(\tau_v(\widetilde{\omega}), \mathcal{Y}_v(\widetilde{\omega}))} dv, \tag{2.17}$$

$$\alpha_t(\widetilde{\omega}) = \inf\{u \,:\, \tau_u(\widetilde{\omega}) \geq t\} \tag{2.18}$$

and hence α_t is a $(\widetilde{\mathcal{G}}_v)$ stop time. Also, $\tilde{\mathcal{X}}_t = \mathcal{Y}_{\alpha_t}$.

The optional sampling theorem and (2.16) now give us that

$$\begin{aligned} N_u^k &= M_{\tau_u}^k \\ &= \hat{f}_k(\tilde{\mathcal{X}}_{\tau_u}) - \int_0^{\tau_u} \hat{h}_k(\tilde{\mathcal{X}}_s) ds \\ &= \hat{f}_k(\mathcal{Y}_u) - \int_0^u \frac{1}{K(\tau_v, \mathcal{Y}_v)} \hat{h}_k(\mathcal{Y}_v) dv \end{aligned} \tag{2.19}$$

is a local martingale. Using that $\frac{1}{K(s,\zeta)}\hat{h}_k \leq C_k$ (see (2.15)), it follows that

$$(N_u^k, \widetilde{\mathcal{G}}_u) \text{ is a } \widetilde{P} \text{ martingale.} \tag{2.20}$$

Let $\mathbb{D} = D([0,\infty), [0,T) \times \hat{E})$ equipped with the Skorokhod topology (see [7]). We will denote a generic element of $\mathbb{D}$ by (γ, θ) with γ denoting the $[0,T)$ valued function and θ denoting the $\hat{E}$ valued function. Let ψ_u and $\mathbb{Y}_u$ be defined by

$$\psi_u(\gamma,\theta) = \gamma(u) \text{ and } \mathbb{Y}_u(\gamma,\theta) = \theta(u).$$

Consider the mapping $\Lambda : \widetilde{\Omega} \to \mathbb{D}$ given by

$$\Lambda(\widetilde{\omega})(u) = (\tau_u(\widetilde{\omega}), \mathcal{Y}_u(\widetilde{\omega})).$$

For $u \geq 0$, let $\mathcal{H}_u$ be the σ field on $\mathbb{D}$ generated by $\{\gamma(r), \mathbb{Y}(r) \,:\, r \leq u\}$ and $\mathbb{P} = \widetilde{P} \circ [\Lambda]^{-1}$. Then by definition

$$\mathbb{P} \circ (\psi, \mathbb{Y})^{-1} = \widetilde{P} \circ (\tau, \mathcal{Y})^{-1}.$$

Thus in view of (2.17) we have

$$\psi_u(\gamma,\theta) = \int_0^u \frac{1}{K(\psi_v(\gamma,\theta), \mathbb{Y}_v(\gamma,\theta))} dv \quad \mathbb{P} - a.s.$$

For $0 \leq t < T$ let

$$\beta_t(\gamma,\theta) = \inf\{u :\, \psi_u(\gamma,\theta) \geq t\}$$

and let

$$\begin{aligned} \mathbb{Z}_t(\gamma,\theta) &= \mathbb{Y}_{\beta_t(\gamma,\theta)}(\gamma,\theta) \\ Z_t(\gamma,\theta) &= \mathcal{J}^{-1} \circ \mathbb{Z}_t(\gamma,\theta). \end{aligned}$$

It then follows that β_t is a $(\mathcal{H}_u)$ stop time and the joint distribution of

$$\Big\{\psi_u, \mathbb{Y}_u, \beta_t, \mathbb{Z}_t, Z_t \,:\, 0 \leq u < \infty,\, 0 \leq t < T\Big\} \text{ under } \mathbb{P}$$

is the same as the point distribution of

$$\left\{\tau_u, \mathcal{Y}_u, \alpha_t, \tilde{\mathcal{X}}_t, \tilde{X}_t \;:\; 0 \le u < \infty,\, 0 \le t < T\right\} \text{ under } \widetilde{P}.$$

In particular, (Z_t) is a solution of the (B^{Δ}, μ_0) martingale problem with $\mathcal{L}(Z_t) = \mu_t$ for all t,

$$\beta_t(\gamma, \theta) = \int_0^t K(s, \mathbb{Z}_t(\gamma, \theta))ds \text{ a.s. } \mathbb{P}$$

and

$$\psi_u(\gamma, \theta) = \inf\{t \ge 0 :\; \beta_t(\gamma, \theta) \ge u\} \;\text{ a.s. } \mathbb{P}.$$

Also, using (2.19) – (2.20) it follows that

$$R_u^k = \hat{f}_k(\mathbb{Y}_u) - \int_0^u \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{h}_k(\mathbb{Y}_v)dv$$

is a $\mathbb{P}$ martingale for every $k \ge 0$.

Step 3: *Construction of the Probability measure* $\mathbb{Q}$.

Recall that $f_1(x) = \mathbf{1}_E(x)$ and $h_1(x) = -\lambda(x)$ and that $\lambda(\Delta) = 0$. Writing $F = \mathcal{J}(E)$, we have $\hat{f}_1 = \mathbf{1}_F$ and $\hat{h}_1 = -\hat{\lambda} = \hat{h}_1\hat{f}_1$. Thus we have

$$\mathbf{1}_F(\mathbb{Y}_u) + \int_0^u \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{\lambda}(\mathbb{Y}_u)dv \tag{2.21}$$

is a martingale. Using integration by parts, it follows that

$$L_u = \mathbf{1}_F(\mathbb{Y}_u) \exp\left\{\int_0^u \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{\lambda}(\mathbb{Y}_u)dv\right\} \tag{2.22}$$

is a martingale.

Now we can construct a probability measure $\mathbb{Q}$ on $\mathbb{D}$ such that for any set $\mathbb{B} \in \mathcal{H}_u$

$$\mathbb{Q}(\mathbb{B}) = \int_{\mathbb{B}} \mathbf{1}_F(\mathbb{Y}_u) \exp\left\{\int_0^u \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{\lambda}(\mathbb{Y}_u)dv\right\} d\mathbb{P}. \tag{2.23}$$

Indeed, equation (2.23) can be used to define $\mathbb{Q}_m$ on $\mathcal{H}_m$ for every integer $m \ge 1$ and then we can use a version of Kolmogorov's consistency theorem (see [12, Theorem V.4.1]) to construct $\mathbb{Q}$.

For any u, t noting that $u \wedge \beta_t$ is bounded stop time (w.r.t. $(\mathcal{H}_u)$), it follows from (2.23) that for any set $\mathbb{B} \in \mathcal{H}_{u\wedge\beta_t}$

$$\mathbb{Q}(\mathbb{B}) = \int_{\mathbb{B}} \mathbf{1}_F(\mathbb{Y}_{u\wedge\beta_t}) \exp\left\{\int_0^{u\wedge\beta_t} \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{\lambda}(\mathbb{Y}_v)\right\} d\mathbb{P}. \tag{2.24}$$

As a consequence, we get

$$\begin{aligned} &\mathbb{P}(\mathbb{B} \cap \{\mathbb{Y}_{u\wedge\beta_t} \in F\}) \\ &\quad = \int_{\mathbb{B}} \mathbf{1}_F(\mathbb{Y}_{u\wedge\beta_t}) \exp\left\{-\int_0^{u\wedge\beta_t} \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{\lambda}(\mathbb{Y}_v)dv\right\} d\mathbb{Q}. \end{aligned} \tag{2.25}$$

This completes the construction of $\mathbb{Q}$.

Step 4: *Uniqueness of solution of the evolution equation.*

Let $\mathcal{K}_t = \mathcal{H}_{\beta_t}$. Using the fact that that R^k_u and L_u are $(\mathcal{H}_u)$ martingales under $\mathbb{P}$, $\hat{g}_k = \hat{h}_k + \hat{\lambda}$ and the fact that $\hat{h}_k = \hat{h}_k \mathbf{1}_F$, one can check using integration by parts formula that $(S^k_u L_u, \mathcal{H}_u)$ is a $\mathbb{P}$ martingale where

$$S^k_u = \hat{f}_k(\mathbb{Y}_u) - \int_0^u \frac{1}{K(\psi_v, \mathbb{Y}_v)} \hat{g}_k(\mathbb{Y}_v) dv \tag{2.26}$$

and hence it follows that $(S^k_u, \mathcal{H}_u)$ is a $\mathbb{Q}$ martingale. Recalling that $\mathbb{Z}_t = \mathbb{Y}_{\beta_t}$, $\mathbb{Y}_{u\wedge\beta_t} = \mathbb{Z}_{\tau_u\wedge t}$, it follows that

$$W^k_t := S^k_{\beta_t} = \hat{f}_k(\mathbb{Z}_t) - \int_0^t \hat{g}_k(\mathbb{Z}_s) ds \tag{2.27}$$

is a $(\mathcal{K}_t)$ local martingale under $\mathbb{Q}$. Recalling the definition of $\hat{f}_k$, $\hat{g}_k$ and that of Z_t, it follows that

$$W^k_t = f_k(Z_t) - \int_0^t (Af_k)(Z_s) ds.$$

Since (W^k_t) is a local martingale for every k, in view of Hypothesis 2.5, it follows that under $\mathbb{Q}$, (Z_t) is a solution to the local martingale problem for (A, μ_0). It is easy to verify that $\int_0^T \Phi(Z_s) ds < \infty \quad a.s.\,\mathbb{Q}$. Thus the finite dimensional distributions of the process (Z_t) under $\mathbb{Q}$ are the same as those of the process (X_t), in view of hypothesis 2.7 and the assumptions in the Theorem.

Also, $\mathcal{K}_{t\wedge\tau_u} = \mathcal{H}_{u\wedge\beta_t}$. Thus (2.25) can be recast as follows. For any set $\mathbb{B} \in \mathcal{K}_{t\wedge\tau_u}$

$$\mathbb{P}\left(\mathbb{B} \cap \{\mathbb{Z}_{\tau_u\wedge t} \in F\}\right) = \int_{\mathbb{B}} \mathbf{1}_F(\mathbb{Z}_{\tau_u\wedge t}) \exp\left\{-\int_0^{\tau_u\wedge t} \hat{\lambda}(\mathbb{Z}_s) ds\right\} d\mathbb{Q}. \tag{2.28}$$

Hence, for $B \subseteq E$, $B \in \mathcal{B}(E)$, taking $\mathbb{B} = \{\mathbb{Z}_{\tau_u\wedge t} \in \mathcal{J}(B)\}$ it follows that

$$\mathbb{P}\left(\{\mathbb{Z}_{\tau_u\wedge t} \in \mathcal{J}(B)\}\right) = \int \mathbf{1}_{\mathcal{J}(B)}(\mathbb{Z}_{\tau_u\wedge t}) \exp\left\{-\int_0^{\tau_u\wedge t} \hat{\lambda}(\mathbb{Z}_s) ds\right\} d\mathbb{Q}. \tag{2.29}$$

Taking limit as $u \to \infty$ (via the sequence of positive integers) it follows that

$$\mathbb{P}\left(\{\mathbb{Z}_t \in \mathcal{J}(B)\}\right) = \int \mathbf{1}_{\mathcal{J}(B)}(\mathbb{Z}_t) \exp\left\{-\int_0^t \hat{\lambda}(\mathbb{Z}_s) ds\right\} d\mathbb{Q}. \tag{2.30}$$

From the definition of Z, it now follows that

$$\mathbb{P}\left(Z_t \in B\right) = \int \mathbf{1}_B(Z_t) \exp\left\{-\int_0^t \lambda(Z_s) ds\right\} d\mathbb{Q}. \tag{2.31}$$

We have seen that $\mu_t(B) = \mathbb{P}(Z_t \in B)$ and also that the finite dimensional distributions of the process (Z_t) under $\mathbb{Q}$ are the same as those of the process X^* - the solution to the (A, μ_0) martingale problem. Hence it follows that any solution $\{\mu_t\}$ to (2.10) satisfying (2.11) is given by (2.12). This completes the proof.

3 Zakai equation in White noise theory of filtering

In filtering theory, the process of interest or the signal process X is unobservable. In the following, we will assume that X is a (possibly time inhomogeneous) Markov process charachterized via a martingale problem for (A_t) where $A_t, 0 \leq t \leq T$ are operators with common domain $\mathcal{D}$. This is equivalent to saying that the state-time process (t, X_t) is a (time homogeneous) Markov process charachterized via the martingale problem for A where $\mathcal{D}(A)$ consists of finite linear combinations of functions of the form $f(x)\xi(t)$ for $f \in \mathcal{D}, \xi \in C_b^1([0,\infty))$. Then for $g(t,x) = \sum_{i=1}^k f_i(x)\xi_i(t) \in \mathcal{D}(A)$, Ag is defined by

$$Ag(t,x) = \sum_{i=1}^k \left[f_i(x)\frac{\partial}{\partial t}\xi_i(t) + \xi_i(t)A_t f_i(x) \right]. \tag{3.32}$$

We assume that A (and $\mathcal{D}(A)$) satisfy the conditions of Theorem 2.2 with a suitable function Φ and that the signal process satisfies

$$\mathbb{E}\int_0^T \Phi(s, X_s)ds < \infty. \tag{3.33}$$

We will work with the white noise model of filtering proposed by Kallianpur and Karandikar and which we describe below. We just introduce the notations and terminology relevant for our purpose. For a more complete overview see [9] and [10].

Let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let (n_t) be a $\mathcal{H}$ valued white noise process. Such a process does not exist on a countably additive probability space but can be constructed on a finitely additive probability space. We assume that (n_t) is independent of the signal process X.

Let $T < \infty$ and $H = L^2([0,T], \mathcal{H})$, the space of $\mathcal{H}$ valued square integrable functions on $[0,T]$, *i.e.*

$$H = \left\{ f : [0,T] \to \mathcal{H} \quad | \quad \int_0^T \|f_s\|^2 ds < \infty \right\}.$$

Then H is also a Hilbert space.

The observation process (y_t) is modelled as

$$y_t = h_t(X_t) + n_t, \qquad 0 \leq t \leq T \tag{3.34}$$

where the observation function $h : [0,T] \times E \to \mathcal{H}$ satisfies the finite energy condition

$$\mathbb{E}\int_0^T \|h_s(X_s)\|^2 ds < \infty. \tag{3.35}$$

Note that the white noise process (n_t) belongs to H and hence so does (y_t). The main aim of filtering theory is to estimate X_t based on observations $\{y_s : 0 \leq s \leq t\}$.

The conditional distribution $F_t(y)$ of X_t given $\{y_s : 0 \leq s \leq t\}$ defined by

$$F_t(y)(B) = \mathbb{E}\left[I_B(X_t) | y_s : 0 \leq s \leq t\right] \text{ for all Borel sets } B \subset E$$

is the optimal filter. The following result gives an alternative expression for F_t and is from [10]. Also see [9].

Let

$$\lambda^y(s,x) = \tfrac{1}{2}\|h_s(x)\|^2 - \langle h_s(x), y_s\rangle,$$
$$\Gamma_t(y)(B) = \mathbb{E}\left[I_B(X_t)\exp\left\{-\int_0^t \lambda^y(s,X_s)ds\right\}\right].$$

Then

$$F_t(y)(B) = \frac{\Gamma_t(y)(B)}{\Gamma_t(y)(E)} \text{ for all Borel sets } B \subset E.$$

$\Gamma_t(y)$ is called the unnormalised conditional distribution of X_t given the observations upto time t.

It can be shown (see [10]) that $\Gamma_t(y)$ satisfies the following equation.

$$\langle f,\mu_t\rangle = \langle f,\mu_0\rangle + \int_0^t \langle A_s f - \lambda^y(s,\cdot)f, \mu_s\rangle ds \quad 0\le t\le T, \quad f\in\mathcal{D}. \tag{3.36}$$

Equation (3.36) is the analogue of the Zakai equation in the white noise theory of filtering. Since

$$-\int_0^t \lambda^y(s,X_s)ds \le \tfrac{1}{2}\int_0^t \|y_s\|^2 ds$$

using (3.33) it follows that

$$\int_0^T \langle\Phi(t,\cdot),\Gamma_t(y)\rangle dt < \infty. \tag{3.37}$$

For $y\in H$ fixed, let us define a measure $\tilde{\Gamma}_t$ on $[0,T]\times E$ by (for Borel sets $C\subseteq[0,T]$, $D\subseteq E$)

$$\tilde{\Gamma}_t(C\times D) = \mathbf{1}_C(t)\exp\left\{-\tfrac{1}{2}\int_0^t \|y_s\|^2 ds\right\}\Gamma_t(y)(D).$$

Also, let $\tilde{\lambda}^y(s,x) = \frac{1}{2}\|y_s\|^2 + \lambda^y(s,x)$. Note that

$$\tilde{\lambda}^y(s,x)\ge 0 \ \ \forall (s,x)\in[0,T]\times E.$$

It can be easily seen that $\{\tilde{\Gamma}_t\}$ is a solution of

$$\langle g,\rho_t\rangle = \langle g,\rho_0\rangle + \int_0^t \langle Ag - \tilde{\lambda}^y g, \rho_s\rangle ds \quad 0\le t\le T, \quad g\in\mathcal{D}(A). \tag{3.38}$$

and satisfies

$$\int_0^T \langle\Phi(t,\cdot),\tilde{\Gamma}_t(y)\rangle dt < \infty. \tag{3.39}$$

Indeed, if $\{\mu_t:\ 0\le t\le T\}$ is a solution to (3.36) satisfying

$$\int_0^T \langle\Phi(t,\cdot),\mu_t\rangle dt < \infty. \tag{3.40}$$

then it can be seen that $\{\tilde{\mu}_t:\ 0\le t\le T\}$ defined by

$$\tilde{\mu}_t(C\times D) = \mathbf{1}_C(t)\exp\left\{-\tfrac{1}{2}\int_0^t \|y_s\|^2 ds\right\}\mu_t(D)$$

for Borel sets $C \subseteq [0,T]$, $D \subseteq E$ is a solution to (3.38) and satisfies

$$\int_0^T \langle \Phi(t,\cdot), \tilde{\mu}_t \rangle dt < \infty. \tag{3.41}$$

This observation and Theorem 2.2 yield the following characterization of $\Gamma_t(y)$.

THEOREM **3.3** *Suppose that the signal process X is the unique solution of the martingale problem for (A_t) and that the operator A defined by* (3.32) *satisfies the conditions of Theorem 2.2. Suppose $h : [0,T] \times E \to \mathcal{H}$ satisfies the finite energy condition* (3.35)*. Then for all $y \in H$ the unnormalised conditional distribution $\Gamma_t(y)$ is the unique solution of the Zakai equation* (3.36) *in the class of solutions $\{\mu_t\}$ satisfying* (3.40)*.*

References

[1] A. G. Bhatt, G. Kallianpur, and R. L. Karandikar. Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering. *Ann. Probab.*, 23(4):1895–1938, 1995.

[2] A. G. Bhatt and R. L. Karandikar. Invariant measures and evolution equations for markov processes charachterized via martingale problems. *Ann. Probab.*, 21(4):2246–2268, 1993.

[3] A. G. Bhatt and R. L. Karandikar. Evolution equations for markov processes: application to the white-noise theory of filtering. *Appl. Math. Optim.*, 31(3):327–348, 1995.

[4] A. G. Bhatt and R. L. Karandikar. Characterization of the optimal filter: the non-markov case. *Stochastics Stochastics Rep.*, 66(3-4):177–204, 1999.

[5] A. G. Bhatt, R. L. Karandikar, and B. V. Rao. On characterisation of markov processes via martingale problems. *Proc. Indian Acad. Sci. Math. Sci.*, 116(1):83–96, 2006.

[6] P. Echeverra. A criterion for invariant measures of markov processes. *Z. Wahrsch. Verw. Gebiete*, 61(1):1–16, 1982.

[7] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, New York, first edition, 1986.

[8] G. Kallianpur and R. L. Karandikar. Measure-valued equations for the optimum filter in finitely additive nonlinear filtering theory. *Z. Wahrsch. Verw. Gebiete*, 66(1):1–17, 1984.

[9] G. Kallianpur and R. L. Karandikar. White noise calculus and nonlinear filtering theory. *Ann. Probab.*, 13(4):1033–1107, 1985.

[10] G. Kallianpur and R. L. Karandikar. *White noise theory of Prediction, Filtering and Smoothing*, volume 3 of *Stochastics Monographs.* Gordon and Breach, New York, first edition, 1988.

[11] R. L. Karandikar. On the feynman kac formula and its applications to filtering theory. *Appl. Math. Optim.*, 16(3):263–276, 1987.

[12] K. R. Parthasarathy. *Probability measures on metric Spaces*, volume 3 of *Probability and Mathematical Statistics, A series of Monographs and textbooks.* Academic Press, New York and London, first edition, 1967.