Report 99-028
A Bayes Formula for Gaussian Noise Processes and Its Applications.
Pranab K. Mandal
V. Mandrekar
ISSN: 1389-2355

A BAYES FORMULA FOR GAUSSIAN NOISE PROCESSES AND ITS APPLICATIONS

PRANAB K. MANDAL[1] AND V. MANDREKAR[2]

Abstract. An elementary approach is used to derive a Bayes type formula, extending the Kallianpur-Striebel formula, for nonlinear filters associated with Gaussian noise processes. In particular cases of certain Gaussian processes, recent results of Kunita and of Le Breton on fractional Brownian motion are derived. We also use the classical approximation of the Brownian motion by the Ornstein-Uhlenbeck dispersion process to solve the "instrumentability" problem of Balakrishnan. We give precise conditions for the convergence of the filter based on the Ornstein-Uhlenbeck dispersion process to the filter based on the Brownian motion. It is also shown that the solution of the Zakai equation can be approximated by that of a (deterministic) partial differential equation.

Key words. Filtering, Gaussian noise process, Bayes formula, Ornstein-Uhlenbeck dispersion process, Zakai equation, Fractional Brownian motion.

AMS subject classifications. 60G35, 60G15, 62M20, 93E11

1 Introduction

The general filtering problem can be described as follows. The *signal* or *system process* $(X_t, 0 \leq t \leq T)$ is unobservable. Information about (X_t) is obtained by observing another process Y which is a function of X corrupted by noise, i. e.,

$$Y_t = \beta_t + N_t, \quad 0 \leq t \leq T, \tag{1.1}$$

where β_t is measurable with respect to $\mathcal{F}_t^X$, the σ-field generated by the signal $\{X_u,\ 0 \leq u \leq t\}$ (augmented by the inclusion of zero probability sets) and (N_t) is some noise process. The observation σ-field $\mathcal{F}_t^Y = \sigma\{Y_u,\ 0 \leq u \leq t\}$ contains

[1]EURANDOM, P.O. Box 513, 5600 MB Eindhoven, Netherlands

[2]Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824

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 $\mathcal{F}_t^Y$. This is given by the conditional distribution ν_t of X_t given $\mathcal{F}_t^Y$, or equivalently, the conditional expectation $E(f(X_t)|\mathcal{F}_t^Y)$ for a rich enough class of functions f. Since this estimate minimizes the squared error loss, ν is called the *optimal filter*.

In the classical case one considers the observation model

$$dY_t = h(t, X_t)\, dt + dW_t, \tag{1.2}$$

where W is the Wiener process independent of X and h satisfies the conditions for the Girsanov theorem (for details see [9]). Kallianpur and Striebel ([11]) derived a Bayes type formula for the conditional distribution ν_t of the form $\nu_t = \frac{\sigma_t}{<\sigma_t, 1>}$, where σ_t is the so called unnormalized conditional distribution. In the case when the signal process X_t is a Markov process, satisfying the SDE

$$dX_t = A(t, X_t)\, dt + B(t, X_t)\, d\tilde{W}_t,$$

where $\tilde{W}$ is another Wiener process independent of W, Zakai ([20]) showed that σ_t is the unique solution of a measure valued stochastic differential equation.

That the noise process (N_t) is a Wiener process plays an important part in deriving all of the above equations and formulas. However, in the real physical system, the noise process (N_t) may not be exactly a Wiener process. In this case no effective way of computing the filter is known. In a recent paper Kunita ([12]) considered the filtering problem with the observation process

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

where N_t is a particular Gaussian process connected to W_t by a kernel. He derived a Bayes type formula extending the one by Kallianpur and Striebel. We generalize this result to any Gaussian noise process N_t with β in the model (1.1) belonging a.s. to the RKHS of the covariance of (N_t). It should be noted that this result with a modified Kallianpur-Striebel proof was first obtained by one of the authors ([18]).

However, the proof presented here is entirely new and is based on an extension of a one-dimensional result which makes (Y_t), under a change of measure, Gaussian with the same distribution as that of (N_t) and independent of (X_t). As an immediate consequence we get the result of Kunita, and Kallianpur and Striebel with a simple proof.

Recently, stochastic models appropriate for long-range dependent phenomena have been given a great deal of interest and numerous theoretical resets and successful applications have been already reported (see, e.g., Beran [3] and references therein). In this view we consider the filtering problem with the fractional Brownian motion noise process. We obtain a general form of the filter in this case. In particular, if $X_t = \eta$ for all t, then we obtain all the results in [5], under his assumptions.

We also discuss the issue raised by Balakrishnan ([2]) regarding "instrumenting" the filtering problem. An approach to this problem using finitely additive measures was given by Kallianpur and Karandikar in their well-known monograph ([10]). They work on the Cameron-Martin space with a finitely additive measure and approximate the filter through an extension. Our method is to follow the classical approach of Physics; namely, to approximate the Wiener noise process by the Ornstein-Uhlenbeck dispersion process (see, e.g., Nelson [15]). Using our Bayes formula we show that the usual filtering theory with the Wiener process can be obtained as a limit. The latter uses the ideas of Kunita ([12]) on stability. We give here the precise conditions for the validity of stability. It should be observed that the theory with the Ornstein-Uhlenbeck dispersion process can be instrumented. We approximate the dispersion process by neglecting a term of order σ^{-1} for σ large (cf. (6.15)) and for this process we obtain a Zakai equation which can be approximated by an ordinary partial differential equation.

The article is organized as follows. In section 2, we give a brief overview of RKHS and its connection with stochastic processes. The extension of the Kallianpur

Striebel formula is obtained in section 3. We discuss Kunita's result in section 4. Section 5 deals with the filtering problem with the fractional Brownian motion as the noise process. Finally, in section 6, the filtering problem corresponding to the Ornstein-Uhlenbeck dispersion noise process is considered along with its limit.

2 Reproducing Kernel Hilbert Space and Stochastic Processes

Definition : A Hilbert space H consisting of real valued functions on some set $\mathbf{T}$ is said to be a *reproducing kernel Hilbert space*(RKHS), if there exists a function K on $\mathbf{T} \times \mathbf{T}$ with the following two properties: for every t in $\mathbf{T}$ and g in H,

(i) $K(\cdot, t) \in H$,

(ii) $(g(\cdot), K(\cdot, t)) = g(t)$. (The reproducing property)

K is called the *reproducing kernel* of H.

The following basic properties can be found in Aronszajn([1]).

(1^o) If a reproducing kernel exists, then it is unique.

(2^o) If K is the reproducing kernel of a Hilbert space H, then $\{K(\cdot, t), t \in \mathbf{T}\}$ spans H.

(3^o) If K is the reproducing kernel of a Hilbert space H, then it is nonnegative definite in the sense that for all $t_1, \ldots, t_n$ in $\mathbf{T}$ and $a_1, \ldots, a_n \in \mathbb{R}$

$$\sum_{i,j=1}^{n} K(t_i, t_j) a_i a_j \ \geq \ 0.$$

The converse of (3^o), stated in Theorem 2.1 below, is fundamental towards understanding the RKHS representation of Gaussian processes. A proof of the theorem can be found in Aronszajn ([1]).

Theorem 2.1 (E. H. Moore) *A symmetric nonnegative definite function K on $\mathbf{T} \times \mathbf{T}$ generates a unique Hilbert space, which we denote by $H(K)$ or sometimes by $H(K, \mathbf{T})$, of which K is the reproducing kernel.*

Now suppose $K(s,t)$, $s, t \in \mathbf{T}$, is a nonnegative definite function. Then, by Theorem 2.1, there is a RKHS, $H(K, \mathbf{T})$, with K as its reproducing kernel. If we restrict K to $\mathbf{T}' \times \mathbf{T}'$ where $\mathbf{T}' \subset \mathbf{T}$, then K is still a nonnegative definite function. Hence K restricted to $\mathbf{T}' \times \mathbf{T}'$ will also correspond to a reproducing kernel Hilbert space $H(K, \mathbf{T}')$ of functions defined on $\mathbf{T}'$. The following result from Aronszajn ([1]; pp. 351) explains the relationship between these two.

Theorem 2.2 *Suppose $K_{\mathbf{T}}$, defined on $\mathbf{T} \times \mathbf{T}$, is the reproducing kernel of the Hilbert space $H(K_{\mathbf{T}})$ with the norm $\|\cdot\|$. Let $\mathbf{T}' \subset \mathbf{T}$, and $K_{\mathbf{T}'}$ be the restriction of $K_{\mathbf{T}}$ on $\mathbf{T}' \times \mathbf{T}'$. Then $H(K_{\mathbf{T}'})$ consists of all f in $H(K_{\mathbf{T}})$ restricted to $\mathbf{T}'$. Further, for such a restriction $f' \in H(K_{\mathbf{T}'})$ the norm $\|f'\|_{H(K_{\mathbf{T}'})}$ is the minimum of $\|f\|_{H(K_{\mathbf{T}})}$ for all $f \in H(K_{\mathbf{T}})$ whose restriction to $\mathbf{T}'$ is f'.*

If $K(s,t)$ is the covariance function for some zero mean process $Z_t, t \in \mathbf{T}$, then, by Theorem 2.1, there exists a unique RKHS, $H(K, \mathbf{T})$, for which K is the reproducing kernel. It is also easy to see (e.g., see Theorem 3D, [17]) that there exists a congruence (linear, one-to-one, inner product preserving map) between $H(K)$ and $\overline{\mathrm{sp}}^{L^2}\{Z_t, t \in \mathbf{T}\}$ which takes $K(\cdot, t)$ to Z_t. Let us denote by $< Z, h > \in \overline{\mathrm{sp}}^{L^2}\{Z_t, t \in \mathbf{T}\}$, the image of $h \in H(K, \mathbf{T})$ under the congruence.

We conclude the section with an important special case.

2.1 A useful example

Suppose the stochastic process Z_t is a Gaussian process given by

$$Z_t = \int_0^t F(t,u)dW_u, \; 0 \le t \le T,$$

where $\int_0^t F^2(t,u)du < \infty$ for all $0 \le t \le T$. Then the covariance function

$$K(s,t) \equiv EZ_sZ_t = \int_0^{t\wedge s} F(t,u)F(s,u)du \tag{2.1}$$

and the corresponding RKHS is given by

$$H(K) = \left\{ g : g(t) = \int_0^t F(t,u)g^*(u)du, 0 \le t \le T \right\} \tag{2.2}$$

for some (necessarily unique) $g^* \in \overline{\mathrm{sp}}^{L^2}\{F(t,\cdot)1_{[0,t]}(\cdot), 0 \le t \le T\}$, with the inner product

$$(g_1, g_2)_{H(K)} = \int_0^T g_1^*(u)g_2^*(u)du,$$

where

$$g_1(s) = \int_0^s F(s,u)g_1^*(u)du \quad \text{and} \quad g_2(s) = \int_0^s F(s,u)g_2^*(u)du.$$

For $0 \le t \le T$, by taking $K(\cdot,t)^*$ to be $F(t,\cdot)1_{[0,t]}(\cdot)$, we see, from (2.1) and (2.2), that $K(\cdot,t) \in H(K)$. To check the reproducing property suppose $h(t) = \int_0^t F(t,u)h^*(u)\,du \in H(K)$. Then

$$(h, K(\cdot,t))_{H(K)} = \int_0^T h^*(u)K(\cdot,t)^*\,du = \int_0^t h^*(u)F(t,u)\,du = h(t).$$

Also, in this case, it is very easy to check (cf. [16], Theorem 4D) that the congruence between $H(K)$ and $\overline{\mathrm{sp}}^{L^2}\{Z_t, t \in \mathbf{T}\}$ is given by

$$< Z, g > = \int_0^T g^*(u)dW_u. \tag{2.3}$$

3 Extension of the Kallianpur-Striebel formula

Suppose $X_t, 0 \le t \le T$, is a real-valued signal process and the observation process is given by

$$Y_t = \beta(t,X) + N_t, \quad 0 \le t \le T, \tag{3.1}$$

where $\beta : [0,T] \times \mathbb{R}^{[0,T]} \to \mathbb{R}$ is a non-anticipative function and the noise process (N_t) is independent of the signal process (X_t). We are interested in finding the best estimate of $f(X_t)$ based on $\mathcal{F}_t^Y$ which is given by the conditional expectation $E(f(X_t)|\mathcal{F}_t^Y)$. First we consider the one dimensional analog of the problem which captures the main idea of obtaining a Bayes type formula for $E(f(X_t)|\mathcal{F}_t^Y)$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Suppose Z is a standard normal random variable independent of X and $Y = X + Z$. Consider the problem of computing $E(X|Y)$. Suppose $P \ll Q$ and $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-field. Then

$$E_P(X|\mathcal{G}) = \frac{E_Q\left(X \, \frac{dP}{dQ}\middle|\, \mathcal{G}\right)}{E_Q\left(\frac{dP}{dQ}\middle|\, \mathcal{G}\right)}.$$

If we define

$$dQ = \exp\left\{-XY + \frac{1}{2}X^2\right\} dP,$$

then Q is a probability measure. Also, considering the joint characteristic function, under Q, of X and Y it is easy to see that under Q, Y is a standard normal random variable independent of X, and X has the same probability distribution as under P.

Now we consider the stochastic process version of this result. Let N_t be a Gaussian process with zero mean, i.e., $m_t \equiv E(N_t) = 0$, and with the covariance function $R(s,t) \equiv E(N_s N_t)$. Let $\{\xi_t, 0 \le t \le T\}$ be another process with values in a space $\mathcal{S}$ and be independent of $\{N_t, 0 \le t \le T\}$. Suppose

$$Y_t = f(t, \xi) + N_t, \; 0 \le t \le T,$$

where f is a measurable non-anticipative functional on $[0,T] \times \mathcal{S}^{[0,T]}$.

Let $H(R;t)$ denote the RKHS corresponding to $R|_{[0,t]\times[0,t]}$, with norm $\|\cdot\|_t$ and $H(R) = H(R;T)$. Also, let $< N, \cdot >_t$ denote the congruence between $H(R;t)$ and $\overline{\mathrm{sp}}^{L^2}\{N_s, 0 \le s \le t\}$ so that for $g, h \in H(R;t)$, the random variables $< N, g >_t$ and $< N, h >_t$ are normal random variables with mean zero and covariance $E(< N, g >_t < N, h >_t) = (g, h)_{H(R;t)}$. Then we have the following

Theorem 3.1 *Suppose $f(\cdot) \equiv f(\cdot, \xi) \in H(R)$ a.s. Define for each t, $(0 \le t \le T)$,*

$$dQ_t = e^{-<N,f>_t - \frac{1}{2}\|f\|_t^2} dP. \tag{3.2}$$

Then Q_t is a probability measure, and under Q_t,

(i) *$(Y_s)_{0\leq s\leq t}$ is a Gaussian process with zero mean and covariance function R, and is independent of $(\xi_s)_{0\leq s\leq T}$.*

(ii) *$(\xi_s)_{0\leq s\leq T}$ has the same distribution as under P.*

Proof: Fix $0 \leq t \leq T$. First note that since $f(\cdot) \in H(R)$ a.s., by Theorem 2.2, $f|_{[0,t]} \in H(R;t)$ a.s. That Q_t is a probability measure follows from the fact that N and ξ are independent and for $g \in H(R;t)$, $< N, g >_t$ is a zero mean normal random variable with variance $\|g\|_t^2$. Now suppose $0 \leq s_1, \ldots, s_m \leq t$, $0 \leq t_1, \ldots, t_n \leq T$, $g_1, \ldots, g_n : \mathcal{S} \to I\!R$ are measurable, and $\alpha_1, \ldots, \alpha_n, \gamma_1, \ldots, \gamma_m$ are real numbers. Consider the joint characteristic function

$$
\begin{aligned}
&E_{Q_t} e^{i(\alpha_1 g_1(\xi_{t_1})+\cdots+\alpha_n g_n(\xi_{t_n}))+i(\gamma_1 Y_{s_1}+\cdots+\gamma_m Y_{s_m})} \\
&= E_P e^{i\sum_{k=1}^n \alpha_k g_k(\xi_{t_k})+i\sum_{j=1}^m \gamma_j Y_{s_j}} e^{-<N,f>_t-\frac{1}{2}\|f\|_t^2} \\
&= E_P e^{i\sum_{k=1}^n \alpha_k g_k(\xi_{t_k})-\frac{1}{2}\|f\|_t^2+i\sum_{j=1}^m \gamma_j f(s_j)} e^{i\sum_{j=1}^m \gamma_j N_{s_j}-<N,f>_t} \\
&= E_P\left[e^{i\sum_{k=1}^n \alpha_k g_k(\xi_{t_k})-\frac{1}{2}\|f\|_t^2+i\sum_{j=1}^m \gamma_j f(s_j)} E_P\left(e^{i\sum_{j=1}^m \gamma_j N_{s_j}-<N,f>_t}\,\Big|\mathcal{F}_T^\xi\right)\right] \\
&= E_P\left[e^{i\sum_{k=1}^n \alpha_k g_k(\xi_{t_k})-\frac{1}{2}\|f\|_t^2+i\sum_{j=1}^m \gamma_j f(s_j)} e^{-\sum_{j,l=1}^m \gamma_j\gamma_l R(s_j,s_l)-\frac{1}{2}2i\sum_{j=1}^m \gamma_j f(s_j)+\frac{1}{2}\|f\|_t^2}\right] \\
&= E_P\left[e^{i\sum_{k=1}^n \alpha_k g_k(\xi_{t_k})}\right] e^{-\sum_{j,l=1}^m \gamma_j\gamma_l R(s_j,s_l)}.
\end{aligned}
$$

Hence the assertions (i) and (ii) follow. ■.

Let us now consider the observation process (Y_t) given by (3.1). It is easy to see, from (3.2) with $\mathcal{S} = I\!R$, $\xi = X$ and $f(\cdot,\xi) = \beta(\cdot, X)$, that

$$\frac{dP}{dQ_t} = \exp\left\{< Y, \beta(\cdot, X) >_t - \frac{1}{2}\|\beta(\cdot, X)\|_t^2\right\} \qquad \text{a.s. } [Q_t].$$

This is because if $\beta^n(\cdot) = \sum_{j=1}^{k_n} a_{nj} R(\cdot, t_j^n) \in H(R;t)$, $n = 1, 2, \ldots$ are such that $\beta^n \to \beta \equiv \beta(\cdot, X)$ in $H(R;t)$, then

$$
\begin{aligned}
< Y, \beta >_t &= \lim_{n\to\infty} < Y, \beta^n >_t \qquad (\ Q_t\text{-a.s. and hence } P\text{-a.s. }) \\
&= \lim_{n\to\infty} \sum_{j=1}^{k_n} a_{nj} Y_{t_j^n} = \lim_{n\to\infty} \sum_{j=1}^{k_n} a_{nj} N_{t_j^n} + \lim_{n\to\infty} \sum_{j=1}^{k_n} a_{nj} \beta_{t_j^n} \\
&= \lim_{n\to\infty} < N, \beta^n >_t + \lim_{n\to\infty} (\beta, \beta^n)_{H(R;t)} = < N, \beta >_t + \|\beta\|_t^2 \qquad P\text{-a.s.} \quad (3.3)
\end{aligned}
$$

Then for any $\mathcal{F}_T^X$ measurable integrable function $g(T,X)$, we have

$$\begin{aligned} E_P(g(T,X)|\mathcal{F}_t^Y) &= \frac{E_{Q_t}\left(g(T,X)\frac{dP}{dQ_t}\Big|\,\mathcal{F}_t^Y\right)}{E_{Q_t}\left(\frac{dP}{dQ_t}\Big|\,\mathcal{F}_t^Y\right)} \\ &= \frac{E_{Q_t}\left(g(T,X)e^{<Y,\beta(\cdot,X)>_t-\frac{1}{2}\|\beta(\cdot,X)\|_t^2}\Big|\,\mathcal{F}_t^Y\right)}{E_{Q_t}\left(e^{<Y,\beta(\cdot,X)>_t-\frac{1}{2}\|\beta(\cdot,X)\|_t^2}\Big|\,\mathcal{F}_t^Y\right)}. \end{aligned} \tag{3.4}$$

From Theorem 3.1, $\{Y_s, 0 \le s \le t\}$, under Q_t, is independent of $\{X_s, 0 \le s \le T\}$, and the distribution of X, under Q_t, is the same as that under P. Hence the conditional expectations of the form $E_{Q_t}(\phi(X,Y)|\mathcal{F}_t^Y)$ can be evaluated as

$$E_{Q_t}(\phi(X,Y)|\mathcal{F}_t^Y)(\omega) = \int_\Omega \phi(X(\omega'),Y(\omega))Q_t(d\omega') = \int \phi(x,Y(\omega))dP_X(x),$$

where P_X is the probability distribution of X.

Hence, from (3.4), we have the following

Theorem 3.2 *Suppose that the observation process Y_t is as in (3.1). If*

$$\beta(\cdot,X(\omega)) \in H(R) \quad \text{for almost all } \omega, \tag{3.5}$$

then for an $\mathcal{F}_T^X$-measurable and integrable function $g(T,X)$,

$$E\left(g(T,X)\,\Big|\,\mathcal{F}_t^Y\right) = \frac{\displaystyle\int g(T,x)e^{<Y,\beta(\cdot,x)>_t-\frac{1}{2}\|\beta(\cdot,x)\|_t^2}\,dP_X(x)}{\displaystyle\int e^{<Y,\beta(\cdot,x)>_t-\frac{1}{2}\|\beta(\cdot,x)\|_t^2}\,dP_X(x)}. \tag{3.6}$$

We next consider an important special case from which it can be easily shown that the formula (3.6) extends the Kallianpur-Striebel formula, as well as the one by Kunita.

3.1 An important special case

Suppose the noise N_t is of the form

$$N_t = \int_0^t F(t,u)dW_u, \tag{3.7}$$

where $\int_0^T \int_0^t F^2(t,u)dudt < \infty$. Denoting by $R(s,t)$, the covariance function of (N_t), from the example considered in section 2.1 we have

$$H(R;t) = \left\{ g : g(s) = \int_0^s F(s,u)g^*(u)du, g^* \in \overline{sp}^{L^2}\{F(s,\cdot)1_{[0,s]}(\cdot), 0 \le s \le t\} \right\} \tag{3.8}$$

with the inner product

$$(g_1, g_2)_{H(R;t)} = \int_0^t g_1^*(u)g_2^*(u)du,$$

where

$$g_1(s) = \int_0^s F(s,u)g_1^*(u)du \text{ and } g_2(s) = \int_0^s F(s,u)g_2^*(u)du.$$

Suppose the observation process is given by

$$Y_t = \int_0^t F(t,u)\tilde{h}(u, X_u)du + N_t, \tag{3.9}$$

such that

$$\tilde{h}(\cdot, X_{(\cdot)}) \in \overline{sp}^{L^2}\{F(s,\cdot)1_{[0,s]}(\cdot), 0 \le s \le t\}.$$

Then, by (2.3) and by an argument similar to the one used in (3.3), we have for $g(\cdot) = \int_0^{(\cdot)} F(\cdot, u)g^*(u)du \in H(R)$,

$$< Y, g >_t = \int_0^t g^*(u)\tilde{h}(u, X_u)du + \int_0^t g^*(u)dW_u = \int_0^t g^*(u)d\hat{Y}_u,$$

where

$$\hat{Y}_s = \int_0^s \tilde{h}(u, X_u)du + W_s, \quad 0 \le s \le T.$$

Hence the Bayes formula (3.6) becomes

$$E\left(g(T,X)\middle|\mathcal{F}_t^Y\right) = \frac{\int g(T,x)e^{\int_0^t \tilde{h}(u,x_u)d\hat{Y}_u - \frac{1}{2}\int_0^t |\tilde{h}(u,x_u)|^2 du} dP_X(x)}{\int e^{\int_0^t \tilde{h}(u,x_u)d\hat{Y}_u - \frac{1}{2}\int_0^t |\tilde{h}(u,x_u)|^2 du} dP_X(x)}. \tag{3.10}$$

Remark: It is now easy to see that the Bayes formula (3.6) is indeed an extension of the Kallianpur-Striebel formula. Take $F(t,u) \equiv 1$ in the model (3.7) and $\tilde{h}$ in the model (3.9) to be $h \in L^2[0,T] \equiv \overline{\mathrm{sp}}^{L^2}\{1_{[0,t]}(\cdot), 0 \le t \le T\}$, so that $N_t = W_t$ and the observation process satisfies the usual model

$$Y_t = \int_0^t h(u, X_u)du + W_t.$$

Note that, in this case, $\hat{Y}_t = Y_t$. Therefore the Bayes formula (3.6) reduces to the Kallianpur-Striebel formula

$$E\left(g(T,X)\,\middle|\mathcal{F}_t^Y\right) = \frac{\int g(T,x)e^{\int_0^t h(u,x_u)dY_u - \frac{1}{2}\int_0^t |h(u,x_u)|^2 du}\, dP_X(x)}{\int e^{\int_0^t h(u,x_u)dY_u - \frac{1}{2}\int_0^t |h(u,x_u)|^2 du}\, dP_X(x)}.$$

Our result also generalizes a similar result by Kunita. We show that in the next section.

4 Kunita's Result

In this section we shall derive Kunita's result ([12], Theorem 2.1), when $d = 1$, as a corollary of our result. Suppose the signal process (X_t) is a continuous process taking values in a complete metric space S. Suppose the observation process is given by

$$Y_t = \int_0^t h(X_s)\, ds + N_t, \quad 0 \le t \le T, \tag{4.1}$$

where h is a continuous map from S into $\mathbb{R}$ and the noise process (N_t) is given by

$$N_t = m_t + \int_0^t \psi(t,s)\, dW_s, \quad 0 \le t \le T, \tag{4.2}$$

with $\psi(t,s)$ and m_t satisfying the following three conditions.

Condition 1 : $\psi(t,s)$ is continuously differentiable in $(t,s) \in [0,T] \times [0,T]$.

Let $\mathbf{C}_0^r$ be the set of all r-times continuously differentiable functions from $[0,T]$ to $\mathbb{R}$ which vanish at zero. Define $\Psi : \mathbf{C}_0 \equiv \mathbf{C}_0^0 \to \mathbf{C}_0$ such that

$$(\Psi\phi)_t = \int_0^t \psi(t,s)\phi'(s)\, ds \tag{4.3}$$

for $\phi \in \mathbf{C}_0^1$. For general $\phi \in \mathbf{C}_0$, it is extended by integration by parts as

$$(\Psi\phi)_t = \psi(t,t)\phi(t) - \int_0^t \phi(s)\frac{\partial\psi}{\partial s}(t,s)\,ds. \tag{4.4}$$

Let $\mathcal{R}(\Psi) = \{\Psi\phi : \phi \in \mathbf{C}_0\}$. Note that for $f, g \in \mathbf{C}_0$ and $0 \le u \le t \le T$,

$$(\Psi f)_u - (\Psi g)_u = \psi(u,u)\,(f(u) - g(u)) - \int_0^u (f(s) - g(s))\,\frac{\partial\psi}{\partial s}(u,s)\,ds.$$

Hence Ψ is causal in the sense that

$$(\Psi f)_u = (\Psi g)_u \text{ holds for } u \le t, \text{ if } f(s) = g(s) \text{ holds for } s \le t. \tag{4.5}$$

Condition 2 : The transformation Ψ has a causal inverse transformation $K : \mathcal{R}(\Psi) \to \mathbf{C}_0$ such that $K\Psi\phi = \phi$ holds for all $\phi \in \mathbf{C}_0$. Further, Kg is differentiable whenever $g \in \mathbf{C}_0^1 \cap \mathcal{R}(\Psi)$ and the derivative is in $L^2[0,T]$.

Condition 3 : m_t is continuously differentiable in t and it belongs to $\mathcal{R}(\Psi)$.

Set

$$\dot{m}_t = \frac{dm_t}{dt}, \tag{4.6}$$

$$(Lf)_t = \frac{d}{dt}(Kg)_t, \quad \text{where} \quad g_t = \int_0^t f_s\,ds. \tag{4.7}$$

Since $R(s,t) = EN_sN_t = \int_0^{t\wedge s} \psi(t,u)\psi(s,u)\,du$, is as in the special case considered in section 2.1, from (3.8) we have

$$H(R) = \left\{ g : g(t) = \int_0^t g^*(u)\psi(t,u)\,du, g^* \in \overline{\mathrm{sp}}^{L^2}\{\psi(t,\cdot)1_{[0,t]}(\cdot) : 0 \le t \le T\} \right\} \tag{4.8}$$

With the help of the following Lemma 4.1 we can further simplify the form of $H(R)$.

Lemma 4.1 *If ψ satisfies Condition 1 and Condition 2, then*

$$\overline{sp}^{L^2}\{\psi(t,\cdot)1_{[0,t]}(\cdot) : 0 \le t \le T\} = L^2[0,T].$$

Proof : It suffices to show that if $f \in L^2[0,T]$ is such that $f \perp \psi(t,\cdot)1_{[0,t]}(\cdot)$ for all $t \in [0,T]$, then $f = 0$. So suppose $f \in L^2[0,T]$.

$$\int_0^t \psi(t,s)f(s)\,ds = 0 \;\forall t$$

$$\Rightarrow \Psi g = 0, \quad \text{where } g(t) = \int_0^t f(s)\, ds$$

$$\Rightarrow g = K\Psi g = 0 \;\Rightarrow \int_0^t f(s)\, ds = 0 \;\forall t \;\Rightarrow f = 0.$$

Hence the lemma is proved. ∎

Therefore, from Lemma 4.1 and from (4.8), we have

$$H(R) = \left\{ g : g(t) = \int_0^t g^*(u)\psi(t,u)\, du, \text{ for some } g^* \in L^2[0,T] \right\}. \tag{4.9}$$

The following proposition describes a relationship between the spaces $\mathcal{R}(\Psi)$ and $H(R)$.

Proposition 4.2 *Let $\mathcal{R}(\Psi)$ and $H(R)$ be as above. Then*

$$C_0^1 \cap \mathcal{R}(\Psi) \subseteq H(R) \subseteq \mathcal{R}(\Psi).$$

Furthermore, for $g \in H(R)$, $(Kg)_t = \int_0^t g^*(u)\, du$ and if $f \in C_0^1 \cap \mathcal{R}(\Psi)$, then $f^* = L(f')$, where L is given by (4.7).

Proof : Let $g \in H(R)$. From (4.9),

$$g(t) = \int_0^t \psi(t,s) g^*(s)\, ds. \tag{4.10}$$

Considering $\phi = \int_0^{(\cdot)} g^*(u)\, du$, we have $\phi \in \mathbf{C}_0$ and from (4.4),

$$\begin{aligned}
(\Psi\phi)_t &= \psi(t,t)\phi(t) - \int_0^t \frac{\partial\psi}{\partial s}(t,s)\phi(s)\, ds \\
&= \psi(t,t)\int_0^t g^*(u)\, du - \int_0^t \left\{ \frac{\partial\psi}{\partial s}(t,s) \int_0^s g^*(u)\, du \right\} ds \\
&= \int_0^t \psi(t,s) g^*(s)\, ds, \quad \text{using integration by parts} \\
&= g(t), \quad \text{by (4.10)}.
\end{aligned}$$

Hence $H(R) \subset \mathcal{R}(\Psi)$ and for $g \in H(R)$, $(Kg)_t = \int_0^t g^*(u)\, du$.

On the other hand, for $f \in \mathrm{C}_0^1 \cap \mathcal{R}(\Psi)$, letting $\phi \in \mathbf{C}_0$ to be such that $\Psi\phi = f$, by

Condition 2, we have $\phi = K\Psi\phi = Kf$ is differentiable with $\phi' = L(f') \in L^2[0,T]$. Now

$$\begin{aligned} f(t) &= \Psi\phi(t) = \psi(t,t)\phi(t) - \int_0^t \phi(s)\frac{\partial\psi}{\partial s}(t,s)\,ds \\ &= \int_0^t \psi(t,s)\phi'(s)\,ds \quad \text{using integration by parts.} \end{aligned} \tag{4.11}$$

Hence the proposition follows from (4.9). ■

We are now ready to derive the result of Kunita ([12], Theorem 2.1) as a corollary of our result, Theorem 3.2.

Theorem 4.3 (Kunita) *Suppose the noise process (N_t), given by (4.2), satisfies Conditions $1-3$, and the observation process (Y_t) is given by (4.1). Let P_X denote the probability distribution of X on $C[0,T]$. Assume further that $(\int_0^t h(X_s)\,ds)$ belongs to $\mathcal{R}(\Psi)$ a.s. Then for any measurable function g on S, the signal state space, such that $E|g(X_t)| < \infty$*

$$E(g(X_t)|\mathcal{F}_t^Y) = \frac{\int \alpha_t(x,Y)g(x(t))\,dP_X(x)}{\int \alpha_t(x,Y)\,dP_X(x)},$$

where

$$\alpha_t(x,Y) = \exp\left\{\int_0^t L(h(x)+\dot{m})_s\,d\hat{Y}_s - \frac{1}{2}\int_0^t |L(h(x)+\dot{m})_s|^2\,ds\right\}$$

and

$$\hat{Y}_t = \int_0^t Lh(x)_s\,ds + \int_0^t (L\dot{m})_s\,ds + W_t.$$

Remark : To check that $\alpha_t(x,Y)$ in the theorem is in fact $\mathcal{F}_t^Y$-measurable it has been shown in Kunita([12]) that $\hat{Y}_t = (KY)_t$ and then the causality of K is used. In the proof given below we show that $\alpha_t(x,Y) = < Y, \beta(\cdot,x) >_t$ which proves that it is indeed $\mathcal{F}_t^Y$-measurable.

Proof of Theorem 4.3 : Let Ω_0 with $P(\Omega_0) = 1$ be such that $\int_0^t h(X_s(\omega))\,ds \in \mathcal{R}(\Psi)$ for all $\omega \in \Omega_0$. Fix $\omega \in \Omega_0$. Since $h(X_s(\omega))$ is continuous in $s \in [0,T]$, $\int_0^{(\cdot)} h(X_s(\omega))\,ds \in \mathbf{C}_0^1 \cap \mathcal{R}(\Psi)$. So, by Proposition 4.2, $\int_0^{(\cdot)} h(X_s(\omega))\,ds$ belongs to

$H(R)$. Hence $(\int_0^t h(X_s)\,ds)$ belongs to $H(R)$ a.s. with $(\int_0^{\cdot} h(X_s)\,ds)^*(t) = (Lh(x))_t$. Similarly, since by Condition 3, $m \in \mathrm{C}_0^1 \cap \mathcal{R}(\Psi)$, we have $m \in H(R)$ with $m^* = L\dot{m}$. Rewriting the observation model (4.1) we have,

$$\begin{aligned} Y_t &= \int_0^t h(X_s)ds + m_t + \int_0^t \psi(t,s)dW_s \\ &= \int_0^t L(h(x) + \dot{m})_s \psi(t,s)ds + \int_0^t \psi(t,s)dW_s. \end{aligned}$$

The theorem then follows from the special case considered in section 3.1 with $F(t,s) = \psi(t,s)$ and $\tilde{h} = L(h(x) + \dot{m}) \in L^2[0,T]$. ∎

5 Fractional Brownian motion noise process

Suppose the observation process is given by

$$Y_t = \int_0^t h(X_u)du + B_H(t), \quad 0 \le t \le T \tag{5.1}$$

where $B_H(t)$ is a fractional Brownian motion (fBm) with Hurst parameter $H \in (\frac{1}{2}, 1)$ and is independent of the signal process (X_t). Assume that $h(u) \equiv h(X_u)$ is continuous a.s. To apply theorem 3.2 we shall need the following

Lemma 5.1 *Let $(B_H(t), 0 \le t \le T)$ be an fBm with $H \in (\frac{1}{2}, 1)$ and the covariance function $R(s,t)$. For any continuous function $c(\cdot)$ on $[0,\tau]$ $(\tau > 0)$, suppose $g_c^\tau(\cdot)$ satisfies the equation (see Carleman [6])*

$$\int_0^\tau g_c^\tau(u) H(2H-1)|v-u|^{2H-2} du = c(v), \quad 0 \le v \le \tau. \tag{5.2}$$

Suppose $a(\cdot)$ is continuous on $[0,T]$. Then $\int_0^{(\cdot)} a(u)du \in H(R)$ with

$$\left\langle \int_0^{(\cdot)} a(u)du, B_H \right\rangle_t = \int_0^t g_a^t(u)dB_H(u) \quad \textit{and} \quad \left\| \int_0^{(\cdot)} a(u)du \right\|_t^2 = \int_0^t g_a^t(u)a(u)du.$$

Proof: Recall that there exists a congruence between the RKHS, $H(R)$, and $\overline{\mathrm{sp}}^{L^2}\{B_H(s) : s \in [0,T]\}$ under which $R(\cdot,t) \mapsto B_H(t)$. Clearly, $\int_0^T g_a^T(u)dB_H(u) \in$

$\overline{\mathrm{sp}}^{L^2}\{B_H(s) : 0 \le s \le T\}$. Hence there exists $\tilde{g} \in H(R)$ such that the image of $\tilde{g}$, under the congruence, is $\int_0^T g_a^T(u)dB_H(u)$. Then for $0 \le s \le T$, by (5.2),

$$\begin{aligned}\tilde{g}(s) &= (R(\cdot,s),\tilde{g})_{H(R)} = E\ B_H(s)\int_0^T g_a^T(u)dB_H(u)\\ &= \int_0^s\int_0^T g_a^T(u)H(2H-1)|v-u|^{2H-2}dudv = \int_0^s a(v)dv.\end{aligned}$$

This proves that $\int_0^{(\cdot)} a(u)du \in H(R)$ and following the notation of section 2 we have $\left\langle \int_0^{(\cdot)} a(u)du, B_H\right\rangle = \int_0^T g_a^T(u)dB_H(u)$. Exactly in the same way it follows that $\left\langle \int_0^{(\cdot)} a(u)du, B_H\right\rangle_t = \int_0^t g_a^t(u)dB_H(u)$. Finally,

$$\begin{aligned}\left\|\int_0^{(\cdot)} a(u)du\right\|_t^2 &= E\int_0^t g_a^t(u)dB_H(u)\int_0^t g_a^t(u)dB_H(u)\\ &= \int_0^t\int_0^t g_a^t(u)g_a^t(v)H(2H-1)|u-v|^{2H-2}dvdu\\ &= \int_0^t g_a^t(u)a(u)du, \qquad \text{from (5.2)} \qquad \blacksquare\end{aligned}$$

Note that from theorem 3.1, under a suitable change of measure (Y_t) becomes a fractional Brownian motion. Therefore, from the Bayes formula (3.6) with $\beta(t,X) = \int_0^t h(u)du$, and $N_t = B_H(t)$, and from lemma 5.1, we have

$$E\left[f(X_t)\,\middle|\mathcal{F}_t^Y\right] = \frac{\int f(x_t)\exp\left\{\int_0^t g_h^t(u)dY(u) - \frac{1}{2}\int_0^t g_h^t(u)h(u)du\right\}dP_X(x)}{\int \exp\left\{\int_0^t g_h^t(u)dY(u) - \frac{1}{2}\int_0^t g_h^t(u)h(u)du\right\}dP_X(x)}. \tag{5.3}$$

When the signal process is actually a random variable η (independent of the noise process $B_H(t)$) such that $h(u) = \eta a(u)$ where a is a continuous (deterministic) function, then using the fact that for a constant k, $g_{ka}^t = kg_a^t$, from (5.3), we have

$$E\left[f(\eta)\,\middle|\mathcal{F}_t^Y\right] = \frac{\int f(x)\exp\left\{x\int_0^t g_a^t(u)dY(u) - \frac{1}{2}x^2\int_0^t g_a^t(u)a(u)du\right\}dP_\eta(x)}{\int \exp\left\{x\int_0^t g_a^t(u)dY(u) - \frac{1}{2}x^2\int_0^t g_a^t(u)a(u)du\right\}dP_\eta(x)}. \tag{5.4}$$

If we further assume that η is a Gaussian random variable with mean η_0 and variance γ_0, then η being independent of $(B_H(t))$, we have (η, Y) jointly Gaussian. Hence the conditional distribution of η given $\mathcal{F}_t^Y$ is also Gaussian with mean $E(\eta|\mathcal{F}_t^Y) = \hat{\eta}_t$, say, and variance $E\left((\eta - \hat{\eta}_t)^2 \,|\mathcal{F}_t^Y\right) = \hat{\gamma}_t$, say. Then

$$E\left(e^{\alpha\eta}\left|\mathcal{F}_t^Y\right.\right) = \exp\left\{\alpha\hat{\eta} + \frac{1}{2}\alpha^2\hat{\gamma}_t\right\}.$$

Now from (5.4), taking $f(x) = e^{\alpha x}$, we have

$$E\left[e^{\alpha\eta}\left|\mathcal{F}_t^Y\right.\right] = \frac{\int e^{\alpha x}\exp\left\{x\int_0^t g_a^t(u)dY(u) - \frac{1}{2}x^2\int_0^t g_a^t(u)a(u)du\right\}\phi(x;\eta_0,\gamma_0)dx}{\int \exp\left\{x\int_0^t g_a^t(u)dY(u) - \frac{1}{2}x^2\int_0^t g_a^t(u)a(u)du\right\}\phi(x;\eta_0,\gamma_0)dx}, \tag{5.5}$$

where $\phi(x;\eta_0,\gamma_0)$ is the density of a Gaussian random variable with mean η_0 and variance γ_0.

Let us consider the numerator of the right hand side of (5.5) :

$$\begin{aligned}
&\int e^{x\left(\alpha+\int_0^t g_a^t(u)dY(u)\right)-\frac{1}{2}x^2\int_0^t g_a^t(u)a(u)du}\,\frac{1}{\sqrt{2\pi\gamma_0}}\,e^{-\frac{1}{2\gamma_0}(x-\eta_0)^2}dx \\
&= \frac{1}{\sqrt{2\pi\gamma_0}}\int e^{-\frac{1}{2}\left[x^2\left(\gamma_0^{-1}+\int_0^t g_a^t(u)a(u)du\right)-2x\left(\alpha+\gamma_0^{-1}+\eta_0\int_0^t g_a^t(u)dY(u)\right)+\gamma_0^{-1}\eta_0^2\right]}dx \\
&= \frac{1}{\sqrt{2\pi\gamma_0}}\int e^{-\frac{1}{2\gamma_t}\left[x^2-2x(\alpha+m_t)\gamma_t+(\alpha+m_t)^2\gamma_t^2\right]}e^{-\frac{1}{2}\gamma_0^{-1}\eta_0^2+\frac{1}{2}\gamma_t(\alpha+m_t)^2}dx, \\
&\qquad \text{where } \gamma_t^{-1} = \gamma_0^{-1} + \int_0^t g_a^t(u)a(u)du \text{ and } m_t = \gamma_0^{-1}\eta_0 + \int_0^t g_a^t(u)dY(u) \qquad (5.6) \\
&= \sqrt{\gamma_0^{-1}\gamma_t}\; e^{-\frac{1}{2}\gamma_0^{-1}\eta_0^2+\frac{1}{2}\gamma_t(\alpha+m_t)^2}. \qquad (5.7)
\end{aligned}$$

Putting $\alpha = 0$ in (5.7) we get the denominator of the right hand side of (5.5) :

$$\text{Denominator } = \sqrt{\gamma_0^{-1}\gamma_t}\; e^{-\frac{1}{2}\gamma_0^{-1}\eta_0^2+\frac{1}{2}\gamma_t m_t^2}.$$

Therefore, from (5.5), we have

$$E\left[e^{\alpha\eta}\left|\mathcal{F}_t^Y\right.\right] = e^{\frac{1}{2}\left[\gamma_t(\alpha+m_t)^2-\gamma_t m_t^2\right]} = e^{\frac{1}{2}\gamma_t\alpha(\alpha+2m_t)}.$$

Collecting the coefficients of α and α^2 and using (5.6), we get

$$\hat{\eta}_t = \gamma_t m_t = \gamma_t\left(\gamma_0^{-1}\eta_0 + \int_0^t g_a^t(u)dY(u)\right),$$

$$\hat{\gamma}_t = \gamma_t = \left(\gamma_0^{-1} + \int_0^t g_a^t(u)a(u)du\right)^{-1}.$$

Note that these equations for the filter are exactly the same as those obtained by Breton ([5]).

Remark: Recently, Breton ([4]) considered the parametric estimation problem in a simple deterministic regression model setup with the fBm noise process . Our general Bayes formula can be used to study the parametric estimation problem in a more general setup with the fBm noise process, as done in Liptser and Shiryayev ([14]) in parameter estimation of the drift coefficient for diffusion type processes with the Wiener noise. We leave that for a future note.

6 Ornstein-Uhlenbeck noise process

Although the use of the Wiener process as noise produces elegant, powerful mathematical techniques to calculate the optimal filter, one of the main criticism against it (as expressed by Balakrishnan [2]) is from the practical point of view. Since the sample paths of a Wiener process are of unbounded variation with probability one, the actual data samples have zero probability of occurring and hence the results obtained cannot be instrumented. On the other hand, it has been argued by Nelson ([15]) that the Ornstein-Uhlenbeck (dispersion) process is natural to consider as an approximation to the Wiener process and the Ornstein-Uhlenbeck processes are realizable. In this section we consider the filtering problem corresponding to the Ornstein-Uhlenbeck noise process and show that it leads to the conventional theory with the Wiener noise process.

Suppose $v(t)$ is an Ornstein-Uhlenbeck velocity process satisfying the stochastic differential equation

$$dv(t) = -\beta v(t)dt + \sigma dW(t), \qquad (\beta > 0,\ \sigma > 0) \tag{6.1}$$

with the initial value $v(0) = 0$. Consider the Ornstein-Uhlenbeck (dispersion) process given by

$$\xi(t) = \int_0^t v(s)ds. \tag{6.2}$$

It is easy to see that if β and σ tend to infinity in such a way that $\sigma^2/\beta^2 \to 1$, then $\xi(t)$ converges in distribution to the standard Wiener process. See, for example, Theorem 9.5 of Nelson ([15]).

Now suppose the noise process (N_t) is given by an Ornstein-Uhlenbeck process so that, from (6.2) and (6.1), we have

$$N_t = \int_0^t \sigma \int_0^s \exp\{-\beta(s-r)\}dW_r ds = \int_0^t \frac{\sigma}{\beta}\left(1 - e^{-\beta(t-s)}\right) dW_s.$$

Also, suppose that the signal process X is independent of W and the observation process is given by

$$Y_t^{\beta,\sigma} = \int_0^t h(X_u)du + N_t, \tag{6.3}$$

where $h(u) \equiv h(X_u)$ is differentiable in $[0,T]$ and $h'(u) \in L^2[0,T]$.

Then, the covariance $R(t,s)$ of (N_t) is given by

$$R(s,t) = EN_sN_t = \int_0^{t\wedge s} F(t,u)F(s,u)du,$$

where

$$F(t,u) = \frac{\sigma}{\beta}\left(1 - e^{-\beta(t-u)}\right), \quad 0 \le u \le t \le T. \tag{6.4}$$

Also, it is easy to see that

$$\overline{sp}^{L^2}\{F(t,\cdot)1_{[0,t]}(\cdot), 0 \le t \le T\} = L^2[0,T].$$

This is because if $f \in L^2[0,T]$ such that $f \perp F(t,\cdot)1_{[0,t]}(\cdot)$, for all $0 \le t \le T$, then

$$\begin{aligned}
&\int_0^t f(u)F(t,u)du = 0 \quad \forall t \\
&\Rightarrow \int_0^t f(u)\frac{\sigma}{\beta}\left(1 - e^{-\beta(t-u)}\right) du = 0 \quad \forall t \\
&\Rightarrow \int_0^t f(u)du - e^{-\beta t}\int_0^t e^{\beta u} f(u)du = 0 \quad \forall t \\
&\Rightarrow f(t) + \beta e^{-\beta t}\int_0^t e^{\beta u} f(u)du - e^{-\beta t}e^{\beta t} f(t) = 0 \quad a.e.[t] \\
&\Rightarrow \int_0^t e^{\beta u} f(u)du = 0 \quad a.e.[t] \\
&\Rightarrow f(t) = 0 \quad a.e.[t]
\end{aligned}$$

Hence from (3.8) we have

$$H(R) = \left\{ g : g(s) = \int_0^s F(s,u) g^*(u) du, \text{ for some } g^* \in L^2[0,T] \right\}$$

It is also easy to check (assuming, without loss of generality, $h(X_0) = 0$) that

$$\int_0^t h(X_u) du = \int_0^t F(t,u) \left[\frac{\beta}{\sigma} h(X_u) + \frac{1}{\sigma} h'(u) \right] du.$$

Hence the noise process and the observation process are as in the special case considered in section 3.1, that is, N_t is of the form (3.7) with $F(t,s)$ given by (6.4) and $Y_t^{\beta,\sigma}$ is of the form (3.9) with

$$\tilde{h}(u, X_u) = \frac{\beta}{\sigma} h(X_u) + \frac{1}{\sigma} h'(u).$$

In this case, therefore, from (3.10), we have

$$\nu_t^{\beta,\sigma}(f)(Y^{\beta,\sigma}) := E(f(X_t) | \mathcal{F}_t^{Y^{\beta,\sigma}}) = \frac{\int f(x_t) \alpha_t^{\beta,\sigma}(x, Y^{\beta,\sigma}) P_X(dx)}{\int \alpha_t^{\beta,\sigma}(x, Y^{\beta,\sigma}) P_X(dx)},$$

where

$$\alpha_t^{\beta,\sigma}(x, Y^{\beta,\sigma}) = \exp\left\{ \int_0^t \left[\frac{\beta}{\sigma} h(x_u) + \frac{1}{\sigma} h'(u) \right] d\hat{Y}_u^{\beta,\sigma} - \frac{1}{2} \int_0^t \left[\frac{\beta}{\sigma} h(x_u) + \frac{1}{\sigma} h'(u) \right]^2 du \right\} \tag{6.5}$$

and

$$\hat{Y}_t^{\beta,\sigma} = \int_0^t \left[\frac{\beta}{\sigma} h(x_u) + \frac{1}{\sigma} h'(u) \right] du + W_t. \tag{6.6}$$

Now suppose that ν_t is the classical filter based on the observation process

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

Recall from the Kallianpur-Striebel formula that

$$\nu_t(f)(Y) := E(f(X_t) | \mathcal{F}_t^Y) = \frac{\int f(x_t) \alpha_t(x, Y) P_X(dx)}{\int \alpha_t(x, Y) P_X(dx)},$$

where

$$\alpha_t(x, Y) = \exp\left\{ \int_0^t h(x_u) dY_u - \frac{1}{2} \int_0^t h^2(x_u) du \right\}. \tag{6.7}$$

The following result shows that the conventional filter can be approximated by suitable filters corresponding to the Ornstein-Uhlenbeck noise process.

Theorem 6.1 *Suppose h satisfies the following condition*

$$E\left[\exp\left\{7\int_0^T h^2(X_u)du + \int_0^T (h'(u))^2\,du\right\}\right] < \infty. \tag{6.8}$$

Then for bounded function f, as $\beta,\ \sigma \to \infty$, with $\sigma^2/\beta^2 \to 1$,

$$\nu_t^{\beta,\sigma}(f)(Y^{\beta,\sigma}) \longrightarrow \nu_t(f)(Y) \quad a.s. \tag{6.9}$$

through an appropriate subsequence.

Proof: Denote by $a_t(\beta,\sigma)$ and a_t the expressions in the curly brackets in the equations (6.5) and (6.7), respectively. Then as $\beta \to \infty, \sigma \to \infty$ such that $\sigma^2/\beta^2 \to 1$, we have

$$\begin{aligned}
a_t(\beta,\sigma) &= \int_0^t \left[\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right] dW_u + \frac{1}{2}\int_0^t \left[\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right]^2 du \\
&= \frac{\beta}{\sigma}\int_0^t h(x_u)dW_u + \frac{1}{\sigma}\int_0^t h'(u)dW_u + \frac{1}{2}\int_0^t \left[\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right]^2 du \\
&\longrightarrow \int_0^t h(x_u)dW_u + \frac{1}{2}\int_0^t |h(x_u)|^2 du = a_t \ \text{ a.e. } x[P_X] \text{ and a.e.}[P]
\end{aligned} \tag{6.10}$$

Hence it is enough to show that

$$\int \alpha_t^{\beta,\sigma}(x, Y^{\beta,\sigma})P_X(dx) \longrightarrow \int \alpha_t(x,Y)P_X(dx) \quad \text{in } \ L^1, \tag{6.11}$$

for L^1-convergence will imply a.s. convergence through a subsequence and then the theorem will follow from Scheffe's theorem.

It is easy to check that for any numbers a and b,

$$|e^a - e^b| \le |a-b|\cdot \max\left(e^{|a|}, e^{|b|}\right).$$

Then

$$\begin{aligned}
&E\left|\int \alpha_t^{\beta,\sigma}(x, Y^{\beta,\sigma})P_X(dx) - \int \alpha_t(x,Y)P_X(dx)\right| \\
&\le E\int |\exp\{a_t(\beta,\sigma)\} - \exp\{a_t\}|P_X(dx) \\
&\le E\int |a_t(\beta,\sigma) - a_t|\cdot \max\left(e^{|a_t(\beta,\sigma)|}, e^{|a_t|}\right) P_X(dx) \\
&\le \left\{\int E|a_t(\beta,\sigma) - a_t|^2 P_X(dx)\cdot \int E\left(e^{2|a_t(\beta,\sigma)|} + e^{2|a_t|}\right) P_X(dx)\right\}^{1/2} \\
&\le \left(\int I_1 P_X(dx)\right)^{1/2}\left(\int I_2 P_X(dx)\right)^{1/2}, \quad \text{say.}
\end{aligned} \tag{6.12}$$

Then,

$$
\begin{aligned}
I_1 &= E|a_t(\beta,\sigma)\} - a_t|^2 \\
&= E\left|\int_0^t \left\{\left(\frac{\beta}{\sigma}-1\right)h(x_u) + \frac{1}{\sigma}h'(u)\right\} dW_u \right.\\
&\qquad \left. + \frac{1}{2}\int_0^t \left\{\left(\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right)^2 - h^2(x_u)\right\} du\right|^2 \\
&\le 2E\left|\int_0^t \left\{\left(\frac{\beta}{\sigma}-1\right)h(x_u) + \frac{1}{\sigma}h'(u)\right\} dW_u\right|^2 \\
&\qquad + 2\cdot\frac{1}{4}\left|\int_0^t \left(\frac{2\beta^2}{\sigma^2}-1\right)h^2(x_u)du + \int_0^t \frac{2}{\sigma^2}\left(h'(u)\right)^2 du\right|^2 \\
&\le 2\int_0^t \left\{\left(\frac{\beta}{\sigma}-1\right)h(x_u) + \frac{1}{\sigma}h'(u)\right\}^2 du \\
&\qquad + \frac{1}{2}\cdot 2\left\{\left(\frac{2\beta^2}{\sigma^2}-1\right)^2\left(\int_0^t h^2(x_u)du\right)^2 + \frac{4}{\sigma^4}\left(\int_0^t \left(h'(u)\right)^2 du\right)^2\right\} \\
&\le 4\left(\frac{\beta}{\sigma}-1\right)^2\int_0^t h^2(x_u)du + \frac{4}{\sigma^2}\int_0^t \left(h'(u)\right)^2 du \\
&\qquad + \left(\frac{2\beta^2}{\sigma^2}-1\right)^2\left(\int_0^t h^2(x_u)du\right)^2 + \frac{4}{\sigma^4}\left(\int_0^t \left(h'(u)\right)^2 du\right)^2.
\end{aligned}
$$

Hence from (6.8) it follows that

$$\int I_1 P_X(dx) \longrightarrow 0 \text{ as } \beta,\ \sigma\to\infty,\ \text{with } \frac{\sigma^2}{\beta^2}\to 1. \tag{6.13}$$

Now, using the fact that for a normal random variable Z with zero mean and variance σ^2, $Ee^{|Z|}\le 2e^{\sigma^2/2}$, we have

$$
\begin{aligned}
I_2 &= E\left(e^{2|a_t(\beta,\sigma)|} + e^{2|a_t|}\right) \\
&= E\exp\left\{\left|2\int_0^t \left[\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right] dW_u + \int_0^t \left[\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right]^2 du\right|\right\} \\
&\qquad + E\exp\left\{\left|2\int_0^t h(x_u)dW_u + \int_0^t |h(x_u)|^2 du\right|\right\} \\
&\le 2\exp\left\{3\int_0^t \left[\frac{\beta}{\sigma}h(x_u) + \frac{1}{\sigma}h'(u)\right]^2 du\right\} + 2\exp\left\{3\int_0^t |h(x_u)|^2 du\right\} \\
&\le 2\exp\left\{\frac{6\beta^2}{\sigma^2}\int_0^t h^2(x_u)du + \frac{6}{\sigma^2}\int_0^t \left(h'(u)\right)^2 du\right\} + 2\exp\left\{3\int_0^t h^2(x_u)du\right\}.
\end{aligned}
$$

Therefore, from (6.8), we have for large σ and β, $\int I_2 P_X(dx)$ is bounded and consequently, (6.11) follows from (6.12), and (6.13). ∎

Remark: Note that the condition (6.8) in theorem 6.1 will hold if one assumes that the functions $h(\cdot)$ and $h'(\cdot)$ are bounded.

Next we address the issue of implementation of the results obtained by considering the Ornstein-Uhlenbeck dispersion process as the observation noise process. We would like to obtain a Zakai-type evolution equation for the so called unnormalized conditional density of X_t given the observations up to time 't'. So let us assume that the signal process X_t is a Markov process.

First, we shall prove the following properties of $\hat{Y}_t \equiv \hat{Y}_t^{\beta,\sigma}$ and its relationship with $Y_t \equiv Y_t^{\beta,\sigma}$.

Lemma 6.2 *Suppose $\hat{Y}_t$ is given by (6.6). Suppose Q is defined by*

$$dP = \exp\left\{\int_0^T \left[\frac{\beta}{\sigma}h(X_u) + \frac{1}{\sigma}h'(u)\right] d\hat{Y}_u - \frac{1}{2}\int_0^T \left[\frac{\beta}{\sigma}h(X_u) + \frac{1}{\sigma}h'(u)\right]^2 du\right\} dQ.$$

Then

(i) *Under Q, $\hat{Y}_t$ is a Wiener process.*

(ii) $\overline{sp}^{L^2(Q)}\{Y_s, 0 \le s \le t\} = \overline{sp}^{L^2(Q)}\{\hat{Y}_s, 0 \le s \le t\}$.

(iii) $\mathcal{F}_t^Y = \mathcal{F}_t^{\hat{Y}}$.

Proof: Clearly (iii) follows from (ii) as under Q, (Y_t) and $(\hat{Y}_t)$ are Gaussian. (i) follows from Lemma 11.3.1 of [9], since (X_t) is independent of (W_t). For (ii) note that

$$\begin{aligned} Y_t &= \int_0^t h(X_u)du + \int_0^t F(t,u)dW_u \\ &= \int_0^t F(t,u)\left[\frac{\beta}{\sigma}h(X_u) + \frac{1}{\sigma}h'(u)\right] du + \int_0^t F(t,u)dW_u \\ &= \int_0^t F(t,u)d\hat{Y}_u. \end{aligned} \tag{6.14}$$

Hence

$$\overline{\mathrm{sp}}^{L^2(Q)}\{Y_s, 0 \le s \le t\} \subset \overline{\mathrm{sp}}^{L^2(Q)}\{\hat{Y}_s, 0 \le s \le t\}.$$

To show the reverse inclusion suppose $\xi \in \overline{\mathrm{sp}}^{L^2(Q)}\{\hat{Y}_s, 0 \le s \le t\}$ and $E_Q(\xi Y_s) = 0$, for all $0 \le s \le t$. Since $\hat{Y}_t$, under Q, is a Wiener process we can express ξ as an Ito integral, say, $\xi = \int_0^t \phi(u) d\hat{Y}_u$. Then

$$\begin{aligned}
& E_Q(\xi Y_s) = 0 \quad \text{for all } 0 \le s \le t \\
\Rightarrow \quad & E_Q\left(\int_0^t \phi(u) d\hat{Y}_u \int_0^s F(s,u) d\hat{Y}_u\right) = 0 \quad \text{for all } 0 \le s \le t \\
\Rightarrow \quad & \int_0^s \phi(u) F(s,u) du = 0 \quad \text{for all } 0 \le s \le t \\
\Rightarrow \quad & \int_0^s \phi(u) \frac{\sigma}{\beta}\left(1 - e^{-\beta(s-u)}\right) du = 0 \quad \text{for all } 0 \le s \le t \\
\Rightarrow \quad & \int_0^s \phi(u) du - e^{-\beta s} \int_0^s \phi(u) e^{\beta u} du = 0 \quad \text{for all } 0 \le s \le t \\
\Rightarrow \quad & \text{(by differentiating) } \beta e^{-\beta s} \int_0^s \phi(u) e^{\beta u} du = 0 \quad \text{a.e. } s \in [0,t] \\
\Rightarrow \quad & \phi(u) = 0 \quad \text{a.e. } u \in [0,t]
\end{aligned}$$

Hence

$$\overline{\mathrm{sp}}^{L^2(Q)}\{\hat{Y}_s, 0 \le s \le t\} \subset \overline{\mathrm{sp}}^{L^2(Q)}\{Y_s, 0 \le s \le t\}.$$

This completes the proof of part (ii). ■

Because of property (iii) of Lemma 6.2 the filter based on $\{Y_s, 0 \le s \le t\}$ will coincide with the filter based on $\{\hat{Y}_s, 0 \le s \le t\}$, where

$$\hat{Y}_t = \int_0^t \left[\frac{\beta}{\sigma} h(X_u) + \frac{1}{\sigma} h'(u)\right] du + W_t.$$

We shall consider the observation process to be given by

$$\hat{\hat{Y}}_t = \int_0^t \frac{\beta}{\sigma} h(X_u) du + W_t, \tag{6.15}$$

which, for large σ, will approximate $\hat{Y}_t$. We can then use the classical theory with the Wiener noise process to obtain the following result.

Suppose A_t with domain $\mathcal{D}$ is the generator of the Markov signal process (X_t). Denote by $\Phi(u,t)$ the unnormalized conditional density of X_t given $\mathcal{F}_t^Y$. Then

$$\Phi(u,t) = \Phi(u,0) + \int_0^t A_s^* \Phi(u,s) ds + \int_0^t \left[\frac{\beta}{\sigma} h(X_s)\right] \Phi(u,s) d\hat{\hat{Y}}_s, \tag{6.16}$$

where A_s^* is the formal adjoint of A_s.

Now note that from equations (6.14) and the form (6.4) of F we have

$$
\begin{aligned}
Y_t &= \frac{\sigma}{\beta}\int_0^t \left[1 - e^{-\beta(t-u)}\right] d\hat{Y}_u \;=\; \frac{\sigma}{\beta}\left[\hat{Y}_t - e^{-\beta t}\int_0^t e^{\beta u} d\hat{Y}_u\right] \\
&= \frac{\sigma}{\beta}\left[\hat{Y}_t - e^{-\beta t}\left\{e^{\beta t}\hat{Y}_t - \int_0^t \beta e^{\beta u}\hat{Y}_u du\right\}\right] \;=\; \sigma e^{-\beta t}\int_0^t e^{\beta u}\hat{Y}_u du.
\end{aligned}
$$

Hence,

$$
y_t \;:=\; \frac{d}{dt}Y_t \;=\; \sigma e^{-\beta t}(-\beta)\int_0^t e^{\beta u}\hat{Y}_u du + \sigma e^{-\beta t} e^{\beta t}\hat{Y}_t \;=\; \sigma\hat{Y}_t - \beta Y_t,
$$

that is,

$$
\hat{\hat{Y}}_t \;=\; \hat{Y}_t - \frac{1}{\sigma}\int_0^t h'(u)du \;=\; \frac{1}{\sigma}\left\{y_t - \int_0^t h'(u)du\right\} + \frac{\beta}{\sigma}Y_t.
$$

Therefore ignoring the first term in the expression for $\hat{\hat{Y}}_t$ above, which is of the order of σ^{-1}, we see that the solution of the Zakai equation (6.16), for large σ, can be approximated by the solution of the following ordinary partial differential equation

$$
\frac{d}{dt}\Phi(u,t) = A_t^*\Phi(u,t) + \left(\frac{\beta}{\sigma}\right)^2 h(X_t)\;\Phi(u,t)\;y_t.
$$

References

[1] N. Aronszajn, *Theory of reproducing kernels,* Trans. Amer. Math. Soc., 68 (1950), pp. 337-404.

[2] A. V. Balakrishnan, *Non-linear White Noise Theory,* in Multivariate Analysis V, P. R. Krishnaiah, ed., North Holland, Amsterdam, 1980, pp. 97-109.

[3] J. Beran, *Statistics for Long-Memory Processes,* Chapman & Hall, New York, 1994.

[4] Alain Le Breton, *Filtering and parameter estimation in a simple linear system driven by a fractional Brownian motion,* Statistics and Probability Letters, 38 (1998), pp. 263-274.

[5] ALAIN LE BRETON, *A Girsanov type approach to filtering in a simple linear system with fractional Brownian noise,* C. R. Acad. Sci. Paris, t. 326, Series I, 1998, pp. 997-1002.

[6] T. CARLEMAN, *Uber die Abelsche Integralgleichung mit konstanten Integrationsgrenzen,* Math. Z., 15, 1922, pp. 111-120.

[7] M. FUJISAKI, G. KALLIANPUR, AND H. KUNITA, *Stochastic differential equations for the nonlinear filtering problem,* Osaka J. Math., 9 (1972), pp. 19-40.

[8] M. HITSUDA, *Representation of Gaussian processes equivalent to Wiener process,* Osaka J. Math. 5 (1968), pp. 299-312.

[9] G. KALLIANPUR, *Stochastic Filtering Theory,* Springer Verlag, New York, Heidelberg, Berlin, 1980.

[10] G. KALLIANPUR, AND R. L. KARANDIKAR, *White Noise Theory of Prediction, Filtering and Smoothing,* Gordon Breach Science Publishers, 1988.

[11] G. KALLIANPUR, AND C. STRIEBEL, *Estimations of stochastic processes : Arbitrary system process with additive white noise observation errors,* Ann. Math. Statis., 39 (1968), pp. 785-801.

[12] H. KUNITA, *Representation and stability of nonlinear filters associated with Gaussian noises,* in Stochastic Processes : A Festschrift in Honour of Gopinath Kallianpur, Cambanis, Ghosh et al eds., Springer Verlag, 1992, pp. 201-210.

[13] H. KUSHNER, *Dynamical equations for optimal nonlinear filtering,* J. Differential Equations, 3 (1967), pp. 179-190.

[14] R. S. LIPTSER, AND A. N. SHIRYAYEV, *Statistics of Random Processes* Springer-Verlag, New York, 1978.

[15] EDWARD NELSON, *Dynamical Theories of Brownian Motion,* Princeton University Press, 1967.

[16] E. PARZEN, *Statistical inference on time series by Hilbert space methods, I,* in Time Series Analysis Papers, Holden-Day, San Francisco, Cambridge, London, Amsterdam, 1967, pp. 251-382.

[17] E. PARZEN, *Regression Analysis of Continuous Parameter Time Series,* in Proc. 4th Berkeley Sympos. Math. Statis. and Prob., Vol. I, Univ. California Press, Berkeley, Calif., 1961, pp. 469-489.

[18] PRANAB KUMAR MANDAL, *Topics in Stochastic Nonlinear Filtering,* Ph.D. dissertation, University of North Carolina at Chapel Hill, 1997.

[19] F. SMITHIES, *Integral Equations,* Cambridge University Press, London, 1958.

[20] M. ZAKAI, *On the optimal filtering of diffusion processes,* Z. Wahrsch. verw. Gebiete., 11 (1969), pp. 230-243.

Pranab Kumar Mandal
EURANDOM / LG 1.21
P O Box 513
5600 MB Eindhoven
Netherlands
mandal@eurandom.tue.nl

V. Mandrekar
Department of Statistics
and Probability
A-436 Wellls hall
Michigan State University
East Lansing, Michigan 48824
mandrekar@stt.msu.edu