Journal of the Korean Statistical Society (1999), **28**: *1, pp 93–106*

Bayes and Sequential Estimation in Hilbert Space Valued Stochastic Differential Equations

J.P.N.Bishwal[1]

Abstract

In this paper we consider estimation of a real valued parameter in the drift coefficient of a Hilbert space valued Itô stochastic differential equation. First we consider observation of the corresponding diffusion in a fixed time interval $[0, T]$ and prove the Bernstein - von Mises theorem concerning the convergence of posterior distribution of the parameter given the observation, suitably normalised and centered at the MLE, to the normal distribution as $T \to \infty$. As a consequence, the Bayes estimator of the drift parameter becomes asymptotically efficient and asymptotically equivalent to the MLE as $T \to \infty$. Next, we consider observation in a random time interval where the random time is determined by a predetermined level of precision. We show that the sequential MLE is better than the ordinary MLE in the sense that the former is unbiased, uniformly normally distributed and efficient but is latter is not so.

Keywords: Hilbert space valued diffusion process, Itô stochastic differential equation, Bernstein - von Mises theorem, posterior distribution, Bayes estimator, asymptotic normality, Sequential MLE, unbiasedness, efficiency

1. INTRODUCTION

Study of asymptotic properties of different estimators of the drift parameter in finite dimensional Itô stochastic differential equations (SDEs) has been paid a large amount of attention during the last three decades. See, e.g., Liptser and Shiryayev (1978), Basawa and Prakasa Rao (1980), Arato (1982) for long time asymptotics and sequential estimation, and Ibragimov and Hasminskii (1981), Kutoyants (1984, 1994) for small noise asymptotics. On the other hand, estimation problem for infinite dimensional stochastic differential equations has received very little amount of attention and in our opinion is very exciting. Loges (1984)

[1]Stat-Math Unit, Indian Statistical Institute, 203 B.T.Road, Calcutta-700 035, India e-mail : tsmv9705@isical.ac.in

initiated the study of asymptotic properties of maximum likelihood estimator in Hilbert-space valued SDEs. Koski and Loges (1985) applied the above theory to a stochastic heat flow problem. Koski and Loges (1986) studied the consistency and asymptotic normality of minimum contrast estimators in Hilbert space valued SDEs which include the MLE. Recently Kim (1996) also proved the consistency and asymptotic normality of MLE in a Hilbert space valued SDE using Fourier expansion of the solution as the observation time $T \to \infty$. Mohapl (1994) obtained strong consistency and asymptotic normality of MLE of a vector parameter in nuclear space valued SDEs from both time continuous observations as well as spatial observations. Huebner, Khasminskii and Rozovskii (1992) studied the asymptotic properties of the MLE of a parameter in the drift coefficient of parabolic stochastic partial differential equations (SPDEs) as the amplitude of the noise goes to zero. Huebner and Rozovskii (1995) introduced the spectral method and studied the asymptotics of the MLE of the drift parameter in parabolic SPDEs when the number of Fourier coefficients of the solutions of the SPDEs becomes large, both the observation time and the intensity of the noise remaining fixed. Piterbarg and Rozovskii (1996) used the last approach and studied the asymptotic properties of the MLE of a parameter in the SPDE which is used to model the upper oceon variability in physical oceonography using both continuous and discrete time sampling. Recently Piterbarg and Rozovskii (1997) gave necessary and sufficient conditions for the consistency, asymptotic normality and asymptotic efficiency of the MLE based on discrete time sampling when the number of observable Fourier coefficients of the random field governed by the SPDEs becomes large. Mohapl (1996) compared the least squares, optimal estimating function and maximum likelihood estimators of a planar (spatial and temporal) Ornstein-Uhlenbeck process satisfying an SPDE based on lattice sampling.

The MLEs were studied in a similar setting but with space dependent parameter by Bagchi and Borkar (1984), Aihara and Bagchi (1988, 1989, 1991), Aihara (1992, 1994, 1995) when continuous observation are available but since the observations were assumed to be corrupted by additional noise, i.e., only partial observations were available, only consistency of a specified function of the MLE under suitable conditions could be proved. All the above authors assumed linear equations whose solutions are infinite dimensional stationary Markov processes. As far as we know, no results are known about the properties of Bayes estimators and sequential estimators in Hilbert space vlaued SDEs. Our aim in the paper is to bridge this gap. Also we prove posterior asymptotic normality for Hilbert space valued diffusions.

For the previous work on Bernstein-von Mises theorem, concerning the convergence of posterior density to normal density, and the asymptotic properties of Bayes estimators in finite dimensional SDEs, see Prakasa Rao (1980, 1981), Bose (1983), Mishra (1989). For the previous work on sequential maximum likelihood estimation in finite dimensional SDEs, see Novikov (1972), Tikhov (1980), Le Breton and Musiela (1984).

This paper is organised as follows : Section 2 contains the model, assumptions and preliminaries, section 3 contains the convergence of the posterior distribution. In section 4 asymptotic properties of Bayes estimators are studied and in section 5 we study the properties of sequential MLE.

2. ASSUMPTIONS AND PRELIMINARIES

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a stochastic basis satisfying the usual hypotheses on which we define the infinite dimensional SDE

$$dX(t) = \theta AX(t)dt + dW(t), \; X(0) = X_0, t \geq 0 \tag{2.1}$$

where A is the infinitesimal generator of a strongly continuous semigroup acting on a real separable Hilbert space H with scalar product $< \cdot, \cdot >$ and norm $\| \cdot \|$, $\{W(t), t \geq 0\}$ is a H-valued Wiener process (see Curtain and Pritchard (1978)) and $\theta \in \Theta \subset \mathbf{R}$, which is to be estimated on the basis of H-valued process $\{X(t), 0 \leq t \leq T\}$. Let θ_0 be the true value of the unknown parameter θ.

It is well known that the covariance operator Q of $\{W(t), t \geq 0\}$ is nuclear and $W(t)$ has zero mean (see Itô (1984)). Let also $W(0) = 0$ and assume that the eigenvalues $\lambda_i, i \geq 1$ of Q are positive. One can write $W(t) = \sum_{i=1}^{\infty} \beta_i(t) e_i$ P-a.s. where $\{e_i, i \geq 1\}$ is a complete orthonormal basis for H consisting of eigenvectors $\{\lambda_i\}$ of Q and $\{\beta_i(t)\}_{t\geq 0} \equiv \{< W(t), e_i >\}_{t \geq 0}$ are mutually independent real valued Wiener processes with incremental covariances λ_i (see Curtain and Pritchard (1978)).

Let us assume that there exists a unique strong solution of (2.1). Sufficient conditions for this in terms of the operators in the SDE can be found in Curtain and Pritchard (1978).

Let P_θ^T be the measure in the space $(\mathcal{C}([0,T], H), \mathcal{B})$ induced by the solution $X(t) = X^\theta(t)$ of (2.1) and P_0^T corresponds to $(W(t) + X_0)$, $t \in [0,T]$. By

$\mathcal{C}([0,T],H)$ we mean the Banach space of continuous functions $f:[0,T]\to H$, which is equipped with the sup-norm. By $\mathcal{B}$, we denote the associated Borel σ-algebra. Let $X_0^T \equiv \{X(t), 0 \le t \le T\}$.

We assume the following conditions :

(A1) There exist two positive constants C and D s.t.

$$E\exp(D(\sum_{i=1}^{\infty}\frac{1}{\lambda_i} < AX(t), e_i >^2)) \le C \quad \text{for all } t \ge 0.$$

(A2) $\sum_{i=1}^{n_0}\frac{1}{\lambda_i}\int_0^{\infty} < AX(t), e_i >^2 dt = \infty$ P - a.s. for some $n_0 \in \mathbf{N}$.

(A3) $E(\sum_{i=1}^{\infty}\frac{1}{\lambda_i}\int_0^{l} < AX(t), e_i >^2 dt) < \infty$ for all $l \in (0,\infty)$.

(A4) $\lim_{T\to\infty}\frac{1}{T}\sum_{i=1}^{\infty}\frac{1}{\lambda_i}\int_0^{T} < AX(t), e_i >^2 dt = \Gamma$ in probability where Γ is a positive constant.

Conditions (A1) - (A4) are not stringent. Condition (A1) determines the absolute continuity of the measures generated by the diffusion. We need (A2) and (A3) for the strong consistency of the estimators. Condition (A4) determines the ergodic limit behaviour of the process $\sum_{i=1}^{\infty}\frac{1}{\lambda_i} < AX(t), e_i >^2$. We give an example where these conditions are satisfied.

Suppose $H = L_2[0,1]$ endowed with the usual scalar product. Let $Ah = \frac{\partial^2}{\partial x^2}h$ for h in the domain

$$\mathcal{D}(A) = \{h \in H : h', h'' \in H; h(0) = h(1) = 0; h'(0) = h'(1) = 1\}.$$

Let $\{k_i\}_{i\in\mathbf{N}}$ be the set of eigenvalues of A where $k_i = (\pi i)^2, i = 1, 2, \ldots.$ This implies the exponential stability of the strongly continuous semigroup corresponding to the infinisimal generator A. Thus the process $\{X_t\}$ is ergodic with an invariant distribution that of the initial random variable X_0. Hence $(A4)$ holds. Moreover, the Dirichlet boundary conditions make A selfadjoint. Conditions (A1) -(A3) are easily verified for this case.

Under the condition (A1), Loges (1984) showed that the measures P_θ^T and P_0^T are equivalent and the Radon-Nikodym derivative (likelihood) of P_θ^T with respect to P_0^T is given by

$$L_T(\theta) = \frac{dP_\theta^T}{dP_0^T}(X_0^T) = \exp\left\{ -\frac{1}{2}\theta^2 \sum_{i=1}^{\infty} \frac{1}{\lambda_i} \int_0^T < AX(t), e_i >^2 dt \right. \\ \left. + \theta\ L_2 - \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{\lambda_i} \int_0^T < AX(t), e_i > d(< X(t), e_i >). \right. \tag{2.2}$$

The maximum likelihood estimate (MLE) is given by

$$\theta_T = \frac{L_2 - \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{\lambda_i} \int_0^T < AX(t), e_i > d(< X(t), e_i >)}{\sum_{i=1}^{\infty} \frac{1}{\lambda_i} \int_0^T < AX(t), e_i >^2 dt} \tag{2.3}$$

if the denominator is greater than zero.

The following properties of MLE are known.

Theorem 2.1. *(Loges (1984). Under the assumptions (A2) - (A3), $\theta_T \to \theta_0$ a.s. $[P_{\theta_0}]$ and under the condition (A2) - (A4) θ_T is asymptotically normally distributed, i.e., $\sqrt{T}(\theta_T - \theta_0) \xrightarrow{\mathcal{D}} N(0, \Gamma^{-1})$ as $T \to \infty$.*

3. THE BERNSTEIN - VON MISES THEOREM

Suppose that Λ is a prior probability on (Θ, Ξ), where Ξ is the Borel σ-algebra of Θ. Assume that Λ has a density $\lambda(\cdot)$ with respect to the Lebesgue measure and the density is continuous and positive in an open neighbourhood of θ_0.

The posterior density of θ given X_0^T is given by

$$p(\theta|X_0^T) = \frac{\frac{dP_\theta^T}{dP_{\theta_0}^T}(X_0^T)\lambda(\theta)}{\int_\Theta \frac{dP_\theta^T}{dP_{\theta_0}^T}(X_0^T)\lambda(\theta)d\theta} \tag{3.1}$$

Let $u = T^{1/2}(\theta - \theta_T)$.

Then the posterior density of $\sqrt{T}(\theta - \theta_T)$ is given by

$$p^*(u|X_0^T) = T^{-1/2} p(\theta_T + uT^{-1/2}|X_0^T). \tag{3.2}$$

Let

$$\begin{aligned}\gamma_T(u) &= \frac{dP^T_{\theta_T+uT^{-1/2}}}{dP^T_{\theta_0}}(X_0^T)/\frac{dP^T_{\theta_T}}{dP^T_{\theta_0}}(X_0^T) \\ &= \frac{dP^T_{\theta_T+uT^{-1/2}}}{dP^T_{\theta_T}}(X_0^T),\end{aligned} \tag{3.3}$$

$$C_T = \int_{-\infty}^{\infty} \gamma_T(u)\lambda(\theta_T + uT^{-1/2})du. \tag{3.4}$$

Then

$$p^*(u|X_0^T) = C_T^{-1}\gamma_T(u)\lambda(\theta_T + uT^{-1/2}). \tag{3.5}$$

One can reduce equation (2.1) to

$$d(< X(t), e_i >) = \theta < AX(t), e_i > dt + d(< W(t), e_i >). \tag{3.6}$$

Hence we can write

$$\sqrt{T}(\theta_T - \theta) = \frac{L_2 - \lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{\lambda_i}T^{-1/2}\int_0^T < AX(t), e_i > d(< W(t), e_i >)}{\sum_{i=1}^{\infty}\frac{1}{\lambda_i}T^{-1}\int_0^T < AX(t), e_i >^2 dt}. \tag{3.7}$$

From (3.7) and (2.2), it is easy to check that

$$\log \gamma_T(u) = -\frac{1}{2}u^2\Gamma_T. \tag{3.8}$$

where $\Gamma_T = \frac{1}{T}\sum_{i=1}^{\infty}\frac{1}{\lambda_i}\int_0^T < AX(t), e_i >^2 dt.$

Let $K(\cdot)$ be a measurable function satisfying the following conditions.
(B1) There exists a number $\epsilon, 0 < \epsilon < \Gamma$ for which

$$\int_{-\infty}^{\infty} K(u)\exp\left\{-\frac{1}{2}u^2(\Gamma - \epsilon)\right\}du < \infty.$$

(B2) For every $h > 0$ and every $\delta > 0$, as $T \to \infty$

$$\exp(-T\delta)\int_{|u|>h} K(T^{1/2}u)\lambda(\theta_T + u)du \to 0 \ \text{ a.s. } [P_{\theta_0}].$$

Theorem 3.1. *Under the assumptions (A1) - (A3) and (B1) - (B2), we have*

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} K(u)\left|p^*(u|X_0^T) - \left(\frac{\Gamma}{2\pi}\right)^{1/2}\exp(-\frac{1}{2}\Gamma u^2)\right|du = 0 \ \ a.s.\ [P_{\theta_0}].$$

Proof: The proof of this theorem is analogous to that of Theorem 3.1 in Borwanker, Kallianpur and Prakasa Rao (1971) once one reduces the infinite dimensional equation (2.1) to a finite dimensional equation (3.6). Analogous proof for one dimensional diffusion processes is given in Prakasa Rao (1980), Bose (1983) and Mishra (1989). We omit the details. □

As a consequence of this theorem we obtain the following corollary.

Corollary 3.1. *If further* $\int_{-\infty}^{\infty} |\theta|^m \lambda(\theta) d\theta < \infty$ *for some positive integer* m, *then*

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} |u|^m \left| p^*(u|X_0^T) - \left(\frac{\Gamma}{2\pi}\right)^{1/2} \exp(-\frac{1}{2}\Gamma u^2) \right| du = 0 \quad a.s.\ [P_{\theta_0}].$$

Remark 3.1. The case $m = 0$ gives the classical Bernstein-von Mises theorem for Hilbert space valued SDEs in its simplest form.

4. BAYES ESTIMATION

We will study the asymptotic properties of the Bayes estimators in this section.

Suppose that $l(\theta, \phi)$ is a loss function defined on $\Theta \times \Theta$. Assume that $l(\theta, \phi) = l(|\theta - \phi|) \geq 0$ and $l(t)$ is non decreasing for $t \geq 0$. Suppose that $J(\cdot)$ is a non negative function and $K(\cdot)$ and $G(\cdot)$ are functions such that

(C1) $J(T) l(uT^{-1/2}) \leq G(u)$ for all $T \geq 0$.

(C2) $J(T) l(uT^{-1/2}) \to K(u)$ uniformly on bounded intervals of u as $T \to \infty$.

(C3) $\int_{-\infty}^{\infty} K(u + v) \exp(-\frac{1}{2}u^2\Gamma) du$ has a strict minimum at $v = 0$.

(C4) G satisfies (B1) and (B2).

A regular Bayes estimators $\tilde{\theta}_T$ of θ based on X_0^T is one which minimizes

$$B_T(\phi) = \int_{\Theta} l(\theta, \phi) p(\theta | X_0^T) d\theta.$$

Assume that a measurable Bayes estimator exists. We obtain the following results.

Theorem 4.1. *Under the condition (A1) - (A4) and (C1) - (C4), we have*

(i) $\sqrt{T}(\theta_T - \tilde{\theta}_T) \to 0$ *a.s.* $[P_{\theta_0}]$ *as* $T \to \infty$.

$$\begin{aligned}(ii)\ \lim_{T\to\infty} J(T)B_T(\theta_T) &= \lim_{T\to\infty} J(T)B_T(\widetilde{\theta}_T)\\ &= \left(\frac{\Gamma}{2\pi}\right)^{1/2}\int_{-\infty}^{\infty} K(u)\exp(-\frac{1}{2}\Gamma u^2)du.\end{aligned}$$

Proof: The proof is similar to that of Theorem 1 in Borwanker, Kallianpur and Prakasa Rao (1971). □

Combining Theorems 2.1 and 4.1 we obtain the following theorem.

Theorem 4.2. *Under the conditions (A1) - (A4) and (C1) - (C4), we have*
(i) $\widetilde{\theta}_T \to \theta_0$ *a.s.* $[P_{\theta_0}]$ *as* $T \to \infty$.
(ii) $\sqrt{T}(\widetilde{\theta}_T - \theta_0) \overset{\mathcal{D}[P_{\theta_0}]}{\to} \mathcal{N}(0, \Gamma^{-1})$ *as* $T \to \infty$.
In otherwords, Bayes estimators $\widetilde{\theta}_T$ *are strongly consistent and asymptotically normally distributed as* $T \to \infty$.

Theorem 4.1 asserts that Bayes and maximum likelihood estimators are asymptotically equivalent as $T \to \infty$. To distinguish between these two estimators one has to investigate the second order efficiency of these two estimators.

5. SEQUENTIAL ESTIMATION

Now we know that the MLE and the Bayes estimators have good asymptotic properties. However, in addition to asymptotic theory which certainly play a predominant role in statistical theory, sequential estimation has got certain advantages because in real life situations observation is always finite. In the finite dimensional linear SDEs , Novikov (1972) (see also Liptser and Shiryayev (1978)) studied the properties of sequential maximum likelihood estimate (SMLE) of the drift parameter which is the MLE based on observation on a random time interval. He showed that SMLE is better than the ordinary MLE in the sense that the former is unbiased, uniformly normally distributed and efficient (in the sense of having the least variance). His plan is to observe the process until the observed Fisher information exceeds a predetermined level of precision. Of course, this type of sampling plan dates back to Anscombe (1952) which has been used in many other situations, e.g., in autoregressive parameter estimation. Under the assumption that the mean duration of observation in the sequential plan and the ordinary (fixed time) plan are the same, Novikov (1972) showed that the SMLE is more efficient than the ordinary MLE. In this section, our aim is to extend the problem to Hilbert space valued SDE (2.1).

We assume that the process $\{X(t)\}$ is observed until the observed Fisher information of the process exceeds a predetermined level of precision H, i.e., we observe $\{X(t)\}$ over the random time interval $[0,\tau]$ where the stop time τ is defined as

$$\tau \equiv \tau_H := \inf\left\{t \geq 0 : \sum_{i=1}^{\infty} \frac{1}{\lambda_i} \int_0^t < AX(s), e_i >^2 ds = H\right\}, 0 < H < \infty. \quad (5.1)$$

Under the condition (A1) it is well known that the measures P_θ^τ and P_0^τ are equivalent (see Loges (1984), Liptser and Shiryayev (1977)) and the Radon-Nikodym derivative of P_θ^τ with respect to P_0^τ is given by

$$\begin{aligned} \frac{dP_\theta^\tau}{dP_0^\tau}(X_0^\tau) &= \exp\left\{-\frac{1}{2}\theta^2 \sum_{i=1}^{\infty} \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i >^2 dt \right. \\ &\left. + \theta L_2 - \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i > d(< X(t), e_i >)\right\}. \end{aligned} \quad (5.2)$$

Maximizing (5.2) with respect to θ provides the Sequential MLE

$$\begin{aligned} \theta_\tau &= \frac{L_2 - \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i > d < X(t), e_i >}{\sum_{i=1}^{\infty} \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i >^2 dt.} \quad (5.3) \\ &= \frac{1}{H} L_2 - \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i > d(< X(t), e_i >). \end{aligned}$$

We have the following properties of the SMLE θ_τ.

Theorem 5.1. *Under the conditions (A1) - (A4), we have*

(i) the sequential plan is closed,

i.e., $\tau < \infty$ *with prob. one.*

(ii) $E_{\theta_0}(\theta_\tau) = \theta_0$

i.e., the SMLE is unbiased.

(iii) θ_τ *is distributed normally with parameters* θ *and* $\frac{1}{H}$ *for all* $\theta \in \Theta$,

i.e., $\sqrt{H}(\theta_\tau - \theta_0) \sim N(0,1)$.
(iv) In the class of unbiased sequential plans $(\gamma, \hat{\theta}_\gamma)$*, satisfying the conditions*
(a) $P_{\theta_0}\left\{\sum_{i=1}^\infty \frac{1}{\lambda_i}\int_0^\gamma < AX(t), e_i >^2 dt < \infty\right\} = 1$,
(b) $E_{\theta_0}(\hat{\theta}_\gamma)^2 < \infty$,
(c) $E_{\theta_0}\left(\sum_{i=1}^\infty \frac{1}{\lambda_i}\int_0^\gamma < AX(t), e_i >^2 dt\right) \le H$,
the plan (τ, θ_τ) *is optimal in the mean square sense, i.e.,*

$$E_{\theta_0}(\theta_\tau - \theta_0)^2 \le E_{\theta_0}(\hat{\theta}_\gamma - \theta)^2.$$

Proof: From (5.1), we have

$$P_{\theta_0}(\tau > t) = P_{\theta_0}\left\{\sum_{i=1}^\infty \frac{1}{\lambda_i}\int_0^t < AX(s), e_i >^2 ds < H\right\}$$

from which, due to (A2), it follows that

$$P_{\theta_0}(\tau = \infty) = P_{\theta_0}\left\{\sum_{i=1}^\infty \frac{1}{\lambda_i}\int_0^\infty < AX(s), e_i >^2 ds < H\right\} = 0.$$

Hence $P_{\theta_0}(\tau < \infty) = 1$.

Because of

$$d(< X(t), e_i >) = \theta_0 < AX(t), e_i > + d(< W(t), e_i >)$$

we can write

$$\theta_\tau = \theta_0 + \frac{1}{H}\left[L_2 - \lim_{n\to\infty}\sum_{i=1}^n \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i > d(W(t), e_i >)\right].$$

Hence $E_{\theta_0}(\theta_\tau) = \theta$ and $\sqrt{H}(\theta_\tau - \theta)$ has standard normal distribution for all $\theta \in \Theta$ since $L_2 - \lim_{n\to\infty}\sum_{i=1}^n \frac{1}{\lambda_i}\int_0^\tau < AX(t), e_i > d(< W(t), e_i >)$ is a standard Wiener process (see Curtain and Pritchard (1978)).

Thus

$$E_{\theta_0}(\theta_\tau - \theta_0)^2 = \frac{1}{H} \text{ for all } \theta \in \Theta. \tag{5.4}$$

From the Cramer-Rao-Wolfowitz inequality (see Liptser and Shiryayev (1977)) it follows that

$$E_{\theta_0}(\hat{\theta}_\gamma - \theta_0)^2 \ge \left[E_{\theta_0}\left(\sum_{i=1}^\infty \frac{1}{\lambda_i}\int_0^\gamma < AX(t), e_i >^2 dt\right)\right]^{-1} \ge \frac{1}{H}, \tag{5.5}$$

for all $\theta \in \Theta$ by (c).

From (5.4) and (5.5), we obtain

$$E_{\theta_0}(\theta_\tau - \theta_0)^2 \leq E_{\theta_0}(\hat{\theta}_\gamma - \theta_0)^2.$$

Hence SMLE is efficient and the sequential plan (τ, θ_τ) is optimal in the mean square sense. □

Remark 5.2. Note that the conditions in Theorem 5.1 (iv) are not stringent. Condition (a) is required for the sampling plan to be closed almost surely. Condition (b), finiteness of the second moment of the sequential estimate is natural as we are comparing mean square error. Condition (c) merely says that the average information for the sequential plan can not exceed the total precesion. It is ideal for many unbiased sequential plans.

Remark 5.3. [Concluding Remarks] (1) It would be interesting to prove the asymptotic properties of the Bayes and maximum likelihood estimators by showing the local asymptotic normality property of the model as in Kutoyants (1984).

(2) The Berry-Esseen bound and large deviation results for the Bayes and maximum likelihood estimators and the rate of convergence of the posterior distribution remains open. Also bound on the asymptotic equivalence of the MLE and the Bayes estimator remains open.

(3) Asymptotic properties of MLE and the Bayes estimators, the Bernstein - von Mises theorem in nonlinear Hilbert space valued SDE remains to be investigated.

(4) Here we have studied the asymptotics based on continuous observation $\{X(t), 0 \leq t \leq T\}$ of the solution of the SDE. The study of asymptotic based on discrete observation of $X(t)$ at time points, $\{t_0, t_1, \dots, t_n\} \subset [0, T]$ from infinite dimensional SDE has been paid least amount of attention and it is a very fertile field to work with in our opinion.

(5) The SMLE definitely has certain advantages : the SMLE is unbiased but the ordinary MLE is not so. Ordinary MLE is not asymptotic normally distributed for all $\theta \in \Theta$, but the SMLE is uniformly normally distributed for all $\theta \in \Theta$. Hence the variance of SMLE is a constant. Since $\sqrt{H}(\theta_T - \theta)$ is Gaussian $\mathcal{N}(0, 1)$ one can construct confidence interval for θ.

(6) Here we have studied the properties of sequential maximum likelihood estimator. Another problem is the study the properties of sequential Bayes estimator.

ACKNOWLEDGEMENT

The author is grateful to the anonymous referee for careful reading of the paper and for giving valuable suggestions which improved the presentation of the paper.

REFERENCES

Aihara, S.I. (1994). "Maximum likelihood estimate for discontinuous parameter in stochastic hyperbolic systems," *Acta Applicandae Mathematicae* **35**, 131-151.

Aihara, S.I. (1995). "Identification of a discontinuous parameter in stochastic parabolic systems," To appear in *Applied Math. Optimization.*

Aihara, S.I. and Bagchi, A. (1989). "Infinite dimensional parameter identification for stochastic parabolic systems," *Statist. Probab. Letters* **8**, 279-287.

Aihara, S.I. and Bagchi, A. (1988). "Parameter identification for stochastic diffusion equations with unknown boundary conditions," *App. Math. Optimization* **17**, 15-36.

Aihara, S.I. and Bagchi, A. (1991). "Parameter identification for hyperbolic stochastic systems," *J. Math. Anal. Appl.* **160**, 485-499.

Aihara, S.I. (1992). "Regularized maximum likelihood estimate for an infinite dimensional parameter in stochastic parabolic systems," *SIAM J. Cont. Optim.* **30**, 745-764.

Anscombe, F.J. (1952). "Large sample theory of sequential estimation," *Proc. Camb. Phil. Soc.*, **48**, 600-607.

Arato, M. (1982). *Linear Stochastic Systems with Constant Coefficients : A Statistical Approach*, Lecture Notes in Control and Information Sciences **45**, Springer, New York.

Bagchi, A. and Borkar, V. (1984). "Parameter identification in infinite dimensional linear systems," *Stochastics* **12**, 201-213.

Basawa, I.V. and Prakasa Rao, B.L.S. (1980). *Statistical Inference for Stochastic Processes*, Academic Press, New York-London.

Borwanker, J.D., Kallianpur, G. and Prakasa Rao, B.L.S. (1971). "The Bernstein-von Mises theorem for Markov processes," *Ann. Math. Statist.* **42** (1971), 1241-1253.

Bose, A. (1983). "The Bernstein-von Mises theorem for a certain class of diffusion processes," *Sankhyā* Ser. A, **45**, 150-160.

Curtain,R. and Pritchard, A. (1978). *Infinite Dimensional Linear Systems Theory*, Springer, Berlin.

Huebner, M., Has'minskii, R.Z. and Rozovskii, B.L.. "Two examples of parameter estimation for stochastic partial differential equations," In : S. Cambanis, J.K. Ghosh, R.L. Karandikar, P.K. Sen (Eds.) *Stochastic Processes, Freschrift in honour of G. Kallianpur*, Springer, Berlin, p. 149-160.

Huebner, M. and Rozovskii, B.L. (1995). "On the asymptotic properties of maximum likelihood estimators for parabolic stochastic PDEs," *Prob. Theor. Rel. Fields*, **103**, 143-163.

Ibragimov, I.A. and Has'minskii, R.Z. (1981).*Statistical Estimation : Asymptotic Theory*, Springer-Verlag, Berlin.

Itô, K. (1984). *Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces*, CBMS-NSF Reginal Conference Series in Applied Mathematics, Vol. 49, SIAM, Philadeplphia, Pennsylvania.

Kim, Y.T. (1996). "Parameter estimation for an infinite dimensional stochastic differential equation," *J. Korean Statistical Society*, **25**, No. 2, 161-173.

Koski, T. and Loges, W. (1985). "Asymptotic statistical inference for a stochastic heat flow problem," *Statist. Prob. Letters*, **3**, 185-189.

Koski, T. and Loges, W. (1986). "On minimum contrast estimation for Hilbert space valued stochastic differential equations," *Stochastics*, **16**, 217-225.

Kutoyants, Yu. A. (1984). *Parameter Estimation for Stochastic Processes* (Translated and edited by B.L.S. Prakasa Rao) Heldermann-Verlag, Berlin.

Kutoyants, Yu. A. (1994). *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.

LeBreton, A. and Musiela, M. (1984). "A study of a one dimensional bilinear differential model for stochastic processes," *Prob. Math. Statist.*, **4**, 91-107.

Liptser, R.S. and Shiryayev, A.N. (1977, 1978). *Statistics of Random Processes I, II*, Springer-Verlag, Berlin.

Loges, W. (1984). "Girsanov theorem in Hilbert space and an application to the statistics of Hilbert space-valued stochastic differential equations," *Stoch. Proc. Appl.*, **17**, 243-263.

Mishra, M.N. (1989). "The Bernstein - von Mises theorem for a class of nonhomogeneous diffusion processes," *Statist. Decisions*, **7**, 153-165.

Mohapl, J. (1994). "Maximum likelihood estimation in linear infinite dimensional models," *Commun. Statist. - Stoch. Models*, **10**, 781-794.

Mohapl, J. (1997). "On estimation in planar Ornstein-Uhlenbeck process," *Commun. Statist. - Stoch. Models* **13**, 4325-455.

Novikov, A.A. (1972). "Sequential estimation of the parameters of diffusion type processes," *Mathematical Notes* **12**, 812-818.

Piterbarg, L. and Rozovskii, B.L. (1995)."Maximum likelihood estimators in the equations of physical oceanography," In: *Stochastic Modelling in Physical Oceanography*, (Eds.) R.Adler, P.Muller, and B.Rozovskii, Birkhauser, Boston.

Piterbarg, L. and Rozovskii, B.L. (1996). "On asymptotic problems of parameter estimation in stochastic PDEs : discrete time sampling," *Math. Methods Statist.* **6** (2), 200-223.

Prakasa Rao, B.L.S. (1980). "The Bernstein - von Mises theorem for a class of diffusion processes," *Theory of Random Processes*, **9**, 95-101 (in Russian).

Prakasa Rao, B.L.S. (1981). "On Bayes estimation for diffusion fields," *In : Statistics : Applications and New Directions*, *Proc. ISI Golden Jubilee Conferences*, ed. J.K. Ghosh and J. Roy, 504-511.

Tikhov, M.S. (1980). "On the sequential estimation of the drift coefficient of a diffusion process with quadratic and nonquadratic losses," *Theory Prob. Appl.*, **25**, 607-614.