Journal of the Korean Statistical Society (2002), **31**: *3, pp 329–342*

Parameter Estimation for a Hilbert Space-valued Stochastic Differential Equation II[†]

Yoon Tae Kim[1] and Hyun Suk Park[1]

Abstract

We deal with asymptotic properties of Maximum Likelihood Estimator(MLE) for the parameters appearing in a Hilbert space-valued Stochastic Differential Equation(SDE) and a Stochastic Partial Differential Equation (SPDE). In paractice, the available data are only the finite dimensional projections to the solution of the equation. Using these data we obtain MLE and consider the asymptotic properties as the dimension of projections increases. In particular we explore a relationship between the conditions for the solution and asymptotic properties of MLE.

Keywords. Stochastic differential equation, maximum likelihood estimator, Radon Nikodym derivative, statistical differential geometry.
AMS 2000 subject classifications. Primary 60H15; Secondary 62F12.

1. Introduction

When we deal with the problem of statistical parameter estimation for a Hilbert space-valued SDE or Stochastic Partial Differential Equation (SPDE) which depends on unknown parameters, the observation (or solution) $\{X_t, 0 \le t \le T\}$ usually has Fourier series expansion. Let the solution $\{X_t, 0 \le t \le T\}$ be observations with joint distribution $\{P_\theta, \theta \in \Theta\}$ on the set of continuous functions taking values in a Hilbert space. If $\{X_i, i = 1, 2, \ldots\}$ is a series representation for the process, then the finite dimensional projections $\{X_i, i = 1, \ldots, n\}$ of the solution induce the probability measure $P_\theta^n := P_\theta \mid \mathcal{F}_n$ where $\mathcal{F}_n := \sigma\{X_1, \ldots, X_n\}$. Then for each n, P_θ^n is absolutely continuous with respect to $P_{\theta_0}^n$. If only the finite dimensional projections $(X(t))$ of the solution can be used for parameter estimation, we simply consider the Radon-Nikodym derivative $dP_\theta^n/dP_{\theta_0}^n$. But

Received January 2002; accepted April 2002.
[†]This work was supported by Korea Research Foundation Grant(KRF-2000-003-D00040).
[1]Department of Statistics, Hallym University, Chunchon 200-702, Korea

with all the observations, it may happen that P_θ and P_{θ_0} are singular, although P_θ^n and $P_{\theta_0}^n$ are absolutely continuous for fixed n.

In this context Kallianpur and Kim (1998) found a curious result, in view of statistical differential geometry, from the following H-valued SDE's, H being a separable Hilbert apace, with unknown parameters θ and ξ:

$$\begin{aligned} dX(t) &= -\theta LX(t)dt + \xi AX(t)dt + BdW(t) \quad \text{for} \quad 0 < t \leq T, \\ X(0) &= 0, \end{aligned} \tag{1.1}$$

where $(W(t))$ is a cylindrical Brownian motion in H, and the operators L, A and B will be defined below.

When we deal with the problem of estimating θ and ξ in SDE (1.1), the Radon-Nikodym derivative is not available in the infinite dimensional case, but we may obtain the MLE $\hat{\theta}^n$ and $\hat{\xi}^n$ if we only use the finite dimensional projections $\{X_i, i = 1, \ldots, n\}$ of the solution as the observations. Kim and Lee (2000) studied the asymtotic properties of MLE $\hat{\theta}_n$ as $n \to \infty$ in the case when ξ is known : They address conditions on the operators appearing in the Hilbert space-valued SDE under which the unique continuous solution to the SDE exists, and then show that these conditions become the sufficient conditions for consistency of MLE. They also investigate conditions for the asymptotic normality of MLE.

In practice we can not observe $\{X(t) \in H, 0 \leq t \leq T\}$, but only a finite dimensional projection $\{X_i \in R, i \leq d_n, 0 \leq t \leq T\}$ of the solution of Equation (1.1). In this situation, the maximum likelihood method does not give a consistent estimator as the dimension n of a finite projection tends to infinity. Indeed as the dimension n goes to infinity the amount of information grows, but this information is not enough for MLE to have asymptotic properties such as consistency, asymptotic normality and asymptotic efficiency. The present paper treats these problems when estimating unknown parameters θ and ξ appearing in SDE (1.1). In particular we explore a relationship between the conditions for the existence and uniqueness of the solution and asymptotic properties of MLE.

As an example of this situation, in Section 5, we consider the following Dirichlet problem as an example of a parabolic SPDE

$$du(t,x) = (\theta\Delta u(t,x) + \xi u(t,x))dt + (I - \Delta)^{-1/2}dW(t,x), \quad 0 \leq t \leq T, \ 0 < x < \pi, \tag{1.2}$$

with initial and boundary conditions

$$u(0,x) = 0, \text{ for } \ 0 < x < \pi \ \text{ and } \ u(t,0) = u(t,\pi) = 0 \ , \ 0 \leq t \leq T, \tag{1.3}$$

where $W(t)$ is a cylindrical Brownian motion on $H = L^2(0,\pi)$, $\Delta = \partial^2/\partial x^2$ and $(\theta,\xi) \in \Theta \subseteq R^2$. We will show that among the MLE's $\hat{\theta}^{(n)}$ and $\hat{\xi}^{(n)}$ based on the observations $\{u_i(t), i = 1,\ldots,n\}$ the estimator $\hat{\theta}^{(n)}$ is strongly consist but the estimator $\hat{\xi}^{(n)}$ is not consistent as $n \to \infty$.

2. Preliminaries

We consider the estimation problem for the following H-valued SDE:

$$\begin{aligned} dX_t &= -\theta L X_t dt + \xi A X_t dt + B dW_t \quad \text{for} \quad 0 < t \le T, \\ X_0 &= 0. \end{aligned} \tag{2.1}$$

Here the operator L^{-1} is bounded self-adjoint with discrete spectrum. Let $\{\phi_i\}$ be the eigenfunctions of L^{-1} which constitues a complete orthonormal system in H and let $\{\lambda_i^{-1}\}$ be the corresponding eigenvalues. We also assume that $A : H \to H$ and $B : H \to H$ are bounded operators. Throughout what follows, it will be assumed that

(C1) $A^*\phi_i = a_i\phi_i$, $B^*\phi_i = b_i\phi_i$, $a_i, b_i > 0$ for all i where A^* and B^* are the adjoint of the operator A and B, respectively.

(C2) The sequences $\{a_i\}$ and $\{b_i\}$ satisfy

$$\sum_{i=1}^{\infty} \frac{a_i}{\lambda_i} < \infty, \quad \sum_{i=1}^{\infty} \frac{b_i^2}{\lambda_i^{1/2}} < \infty, \quad \text{and} \quad \sum_{i=1}^{\infty} \frac{a_i^2}{\lambda_i} < \infty.$$

(C3) The parameter set Θ is given by

$$\Theta = \{(\theta,\xi) | \theta > 0 \ \text{ and } \ \theta > c\xi\} \ \text{ where } \ c = \sup_i \frac{a_i}{\lambda_i} > 0.$$

(C4) As $n \to \infty$, the sequences $\{a_i\}$ and $\{\lambda_i\}$ satisfy

$$\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n \lambda_i} = o(n^{-1}) \quad \text{and} \quad \frac{(\sum_{i=1}^n a_i)^2}{\sum_{i=1}^n \lambda_i} \to d.$$

It follows by the assumptions (C1) and (C2) that there exists a unique, continuous solution X to the SDE (2.1) (see Bhatt *et al.*, 1993). The solution X_t is given by

$$X_t = \sum_{i=1}^{\infty} X_{t,i}\phi_i,$$

where for all i,

$$dX_i(t) = -(\theta\lambda_i - \xi a_i)X_i(t)dt + b_i dW(\phi_i)(t). \tag{2.2}$$

The solution of (2.2) is

$$X_i(t) = b_i \int_0^t e^{-(\theta\lambda_i - \xi a_i)(t-s)}\, dW(\phi_i)(s). \tag{2.3}$$

On the identifiability of $\{P^n_{\theta,\xi}, (\theta,\xi) \in \Theta\}$ and $\{P_{\theta,\xi}, (\theta,\xi) \in \Theta\}$, see Kallianpur and Kim (1998).

3. Estimation Based on Projection

Now we are interested in estimating (θ, ξ) on $[0,T]$ based on a sample $X_1, X_2, \ldots, X_n$ of the projection of $X(t) \in H$ on $[0,T]$. If $\{X_i, i = 1,2,\ldots\}$ is a series representation for the process, we consider the finite dimension n such that projections $\{X_i, i = 1,\ldots,n\}$ of the solution induce the probability measure $P^n_\theta := P_\theta|\mathcal{F}_n$ where $\mathcal{F}_n := \sigma\{X_1,\ldots,X_n\}$. The log likelihood is then

$$\begin{aligned} &l_n(X^n;\theta,\xi) := \log \frac{dP^n_{\theta,\xi}}{dP^n_{\theta_0,\xi_0}}(X^n) \\ &= \sum_{i=1}^n \frac{(\theta_0-\theta)\lambda_i + (\xi-\xi_0)a_i}{b_i^2} \int_0^T X_{t,i}\, dX_{t,i} \\ &\quad -\frac{1}{2}\sum_{i=1}^n \frac{(\theta^2-\theta_0^2)\lambda_i^2 + 2(\theta_0\xi_0 - \theta\xi)\lambda_i a_i + (\xi^2-\xi_0^2)a_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt. \end{aligned} \tag{3.1}$$

The Fisher information matrix is given by

$$\begin{aligned} g_{\theta\theta}(n) &= \sum_{i=1}^n \frac{\lambda_i^2}{b_i^2} \int_0^T E(X^2_{t,i})\, dt, \\ g_{\theta\xi}(n) &= \sum_{i=1}^n \frac{\lambda_i a_i}{b_i^2} \int_0^T E(X^2_{t,i})\, dt, \\ g_{\xi\xi}(n) &= \sum_{i=1}^n \frac{a_i^2}{b_i^2} \int_0^T E(X^2_{t,i})\, dt. \end{aligned}$$

Now we explicitly compute the Fisher information matrix.

$$g_{\theta\theta}(n) = \frac{T}{2}\sum_{i=1}^{n}\frac{\lambda_i^2}{(\theta\lambda_i - \xi a_i)} - \frac{1}{4}\sum_{i=1}^{n}\frac{\lambda_i^2}{(\theta\lambda_i - \xi a_i)^2} + \frac{1}{4}\sum_{i=1}^{n}\frac{\lambda_i^2 e^{-2(\theta\lambda_i - \xi a_i)T}}{(\theta\lambda_i - \xi a_i)^2},$$

$$g_{\theta\xi}(n) = \frac{T}{2}\sum_{i=1}^{n}\frac{\lambda_i a_i}{(\theta\lambda_i - \xi a_i)} - \frac{1}{4}\sum_{i=1}^{n}\frac{\lambda_i a_i}{(\theta\lambda_i - \xi a_i)^2} + \frac{1}{4}\sum_{i=1}^{n}\frac{\lambda_i a_i e^{-2(\theta\lambda_i - \xi a_i)T}}{(\theta\lambda_i - \xi a_i)^2},$$

$$g_{\xi\xi}(n) = \frac{T}{2}\sum_{i=1}^{n}\frac{a_i^2}{(\theta\lambda_i - \xi a_i)} - \frac{1}{4}\sum_{i=1}^{n}\frac{a_i^2}{(\theta\lambda_i - \xi a_i)^2} + \frac{1}{4}\sum_{i=1}^{n}\frac{a_i^2 e^{-2(\theta\lambda_i - \xi a_i)T}}{(\theta\lambda_i - \xi a_i)^2}.$$

Using the assumption (C4), as $n \to \infty$, we have

$$g_{\theta\theta}(n) = \frac{T}{2\theta}\sum_{i=1}^{n}\lambda_i + \frac{T\xi}{2\theta^2}\sum_{i=1}^{n}a_i - \frac{n}{4\theta^2} + O(1),$$

$$g_{\theta\xi}(n) = -\frac{T}{2\theta}\sum_{i=1}^{n}a_i + O(1),$$

$$g_{\xi\xi}(n) = O(1).$$

As mentioned in Introduction, the measures $P^n_{\theta,\xi}$ and $P^n_{\theta_0,\xi_0}$ are absolutely continuous, *i.e.*, $P_{\theta,\xi}$ is locally absolutely continuous with respect to P_{θ_0,ξ_0}. In the following Lemma, we show that the measures $P_{\theta,\xi}$ and P_{θ_0,ξ_0} corresponding to the solutions of the equation (2.1) are not absolutely continuous.

Lemma 3.1. *$P_{\theta,\xi} \perp P_{\theta_0,\xi_0}$, i.e., $P_{\theta,\xi}$ and P_{θ_0,ξ_0} are singular.*

Proof. Let $\rho^i_{\theta,\xi}$ be the induced measure on $C[0,T]$ by the process $(X_{t,i})$, $0 \leq t \leq T$. Then for $i = 1, 2, \cdots$,

$$\frac{d\rho^i_{\theta,\xi}}{d\rho^i_{\theta_0,\xi_0}}(X_i) = \exp\left\{\frac{(\theta_0\lambda_i - \xi_0 a_i) - (\theta\lambda_i - \xi a_i)}{b_i}\int_0^T X_{t,i}\, dW_t(\phi_i) - \frac{((\theta_0\lambda_i - \xi_0 a_i) - (\theta\lambda_i - \xi a_i))^2}{2b_i^2}\int_0^T (X_{t,i})^2\, dt\right\}.$$

The discrete Hellinger process is $h(1/2; P_{\theta,\xi}, P_{\theta_0,\xi_0})_n = \sum_{i=1}^n [1 - H(1/2; \rho^i_{\theta,\xi}, \rho^i_{\theta_0,\xi_0})]$, where the Hellinger integral is given by

$$H\left(\frac{1}{2}; \rho^i_{\theta,\xi}, \rho^i_{\theta_0,\xi_0}\right) = E_{\theta_0,\xi_0}\sqrt{\frac{d\rho^i_{\theta,\xi}}{d\rho^i_{\theta_0,\xi_0}}(X_i)}\,.$$

Note that $\sum_{i=1}^n [1-H(1/2;\rho_{\theta,\xi}^i,\rho_{\theta_0,\xi_0}^i)]$ converges or diverges according as $\sum_{i=1}^n |\log H(1/2;\rho_{\theta,\xi}^i,\rho_{\theta_0,\xi_0}^i)|$ does. From Jensen's inequality,

$$\begin{aligned}
&\sum_{i=1}^n \left|\log H\left(\frac{1}{2};\rho_{\theta,\xi}^i,\rho_{\theta_0,\xi_0}^i\right)\right| \\
&\geq \sum_{i=1}^n \left| E_{\theta_0,\xi_0}\left[\frac{(\theta_0\lambda_i-\xi_0 a_i)-(\theta\lambda_i-\xi a_i)}{X_{t,i}}\int_0^T b_i\, dW_t(\phi_i)\right.\right. \\
&\qquad \left.\left. -\frac{((\theta_0\lambda_i-\xi_0 a_i)-(\theta\lambda_i-\xi a_i))^2}{2b_i^2}\int_0^T (X_{t,i})^2\, dt\right]\right| \\
&= \sum_{i=1}^n \frac{((\theta_0\lambda_i-\xi_0 a_i)-(\theta\lambda_i-\xi a_i))^2}{2b_i^2}\int_0^T E_{\theta_0,\xi_0}(X_{t,i})^2\, dt \\
&= \sum_{i=1}^n ((\theta-\theta_0)\lambda_i+(\xi_0-\xi)a_i)^2\left(\frac{T}{2(\theta\lambda_i-\xi a_i)}-\frac{1-e^{-2(\theta\lambda_i-\xi a_i)T}}{4((\theta\lambda_i-\xi a_i))^2}\right) \\
&= c(\theta,\xi)\sum_{i=1}^n \lambda_i + o\left(\sum_{i=1}^n \lambda_i\right).
\end{aligned}$$

Since $\sum_{i=1}^\infty \lambda_i = \infty$, we have

$$\sum_{i=1}^\infty \left[1-H\big(\frac{1}{2};\rho_{\theta,\xi}^i,\rho_{\theta_0,\xi_0}^i\big)\right] = \infty.$$

By Theorem 2.37 in Jacod and Shiryaev (1987), the result follows. □

4. Asymptotic Properties of Estimator

The likelihood equation is given by

$$A^{(n)}\underline{\hat{\theta}}^{(n)} = B^{(n)}.$$

Here

$$A^{(n)} = \begin{pmatrix} \sum_{i=1}^n (\lambda_i^2/b_i^2)\int_0^T (X_{t,i})^2\, dt & -\sum_{i=1}^n (\lambda_i a_i/b_i^2)\int_0^T (X_{t,i})^2\, dt \\ -\sum_{i=1}^n (\lambda_i a_i/b_i^2)\int_0^T (X_{t,i})^2\, dt & \sum_{i=1}^n (a_i^2/b_i^2)\int_0^T (X_{t,i})^2\, dt \end{pmatrix},$$

$$B^{(n)} = \begin{pmatrix} -\sum_{i=1}^n (\lambda_i/b_i^2)\int_0^T X_{t,i}\, dX_{t.i} \\ \sum_{i=1}^n (a_i/b_i^2)\int_0^T X_{t,i}\, dX_{t,i} \end{pmatrix},$$

and $\underline{\hat{\theta}}^{(n)}$ is the vector with $(\hat{\theta}^{(n)}, \hat{\xi}^{(n)})$, where $\hat{\theta}^{(n)}$ and $\hat{\xi}^{(n)}$ are MLE's of θ and ξ based on the observations $\{X_i, i = 1, \dots, n\}$. Using the equation (2.2), we can write

$$A^{(n)} \underline{(\hat{\theta}^{(n)} - \theta)} = C^{(n)},$$

where $\underline{\hat{\theta}^{(n)} - \theta}$ is the vector with $\hat{\theta}^{(n)} - \theta$ and $\hat{\xi}^{(n)} - \xi$,

$$C^{(n)} = \begin{pmatrix} -\sum_{i=1}^n \frac{\lambda_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i) \\ \sum_{i=1}^n \frac{a_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i) \end{pmatrix}.$$

Therefore we have

$$\begin{aligned} &\hat{\theta}^{(n)} - \theta \\ &= \left(\sum_{i=1}^n \frac{\lambda_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt \sum_{i=1}^n \frac{a_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt - \left(\sum_{i=1}^n \frac{\lambda_i a_i}{b_i^2} \int_0^T (X_{t,i})^2\, dt \right)^2 \right)^{-1} \\ &\quad \times \left(\sum_{i=1}^n \frac{\lambda_i a_i}{b_i^2} \int_0^T (X_{t,i})^2\, dt \sum_{i=1}^n \frac{a_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i) \right. \\ &\quad \left. - \sum_{i=1}^n \frac{a_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt \sum_{i=1}^n \frac{\lambda_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i) \right). \end{aligned}$$

$$\begin{aligned} &\hat{\xi}^{(n)} - \xi \\ &= \left(\sum_{i=1}^n \frac{\lambda_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt \sum_{i=1}^n \frac{a_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt - \left(\sum_{i=1}^n \frac{\lambda_i a_i}{b_i^2} \int_0^T (X_{t,i})^2\, dt \right)^2 \right)^{-1} \\ &\quad \times \left(\sum_{i=1}^n \frac{\lambda_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt \sum_{i=1}^n \frac{a_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i) \right. \\ &\quad \left. - \sum_{i=1}^n \frac{\lambda_i a_i}{b_i^2} \int_0^T (X_{t,i})^2\, dt \sum_{i=1}^n \frac{\lambda_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i) \right). \end{aligned}$$

Let us introduce the following notations:

$$\begin{aligned} \alpha_n &= \sum_{i=1}^n \frac{\lambda_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt, \quad \gamma_n = \sum_{i=1}^n \frac{\lambda_i a_i}{b_i^2} \int_0^T (X_{t,i})^2\, dt, \\ \beta_n &= \sum_{i=1}^n \frac{a_i^2}{b_i^2} \int_0^T (X_{t,i})^2\, dt, \quad \delta_n = \sum_{i=1}^n \frac{\lambda_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i), \\ \rho_n &= \sum_{i=1}^n \frac{a_i}{b_i} \int_0^T X_{t,i}\, dW_t(\phi_i). \end{aligned}$$

Then we write

$$\hat{\theta}^{(n)} - \theta = \frac{\gamma_n \rho_n - \beta_n \delta_n}{\alpha_n \beta_n - (\gamma_n)^2}, \tag{4.1}$$

$$\hat{\xi}^{(n)} - \xi = \frac{\alpha_n \rho_n - \gamma_n \delta_n}{\alpha_n \beta_n - (\gamma_n)^2}. \tag{4.2}$$

Now we prove that MLE $\hat{\theta}^{(n)}$ converges to the true parameter θ $(a.s.)$ as $n \to \infty$. The notation $U_n \sim V_n$, used in the following Lemma, means that

$$\frac{U_n}{V_n} \to 1 \text{ as } n \to \infty.$$

in the sense of $a.s.$ Let $s_n = \sum_{i=1}^n \lambda_i$.

Lemma 4.1. *As $n \to \infty$, we have the following propositions.*

(a) $\alpha_n \sim g_{\theta\theta}(n)$ *if* $\sum_{n=1}^\infty (\lambda_n^2/s_n^2) < \infty$,

(b) $\gamma_n \sim g_{\theta\xi}(n)$,

(c) $\lim_{n\to\infty} \beta_n = \sum_{i=1}^\infty (a_i^2/b_i^2) \int_0^T (X_{t,i})^2\, dt < \infty$,

(d) $\lim_{n\to\infty} \rho_n = \sum_{i=1}^\infty (a_i/b_i) \int_0^T X_{t,i}\, dW_t(\phi_i) < \infty$,

(e) $\lim_{n\to\infty} \delta_n/\alpha_n = 0$.

Proof. (a). First note that $X_{t,i}$ is a Gaussain random variable mean zero and the variance

$$b_i^2 \int_0^t e^{2(\theta\lambda_i - \xi a_i)(t-s)} ds.$$

Let

$$U_i = \frac{\lambda_i^2}{b_i^2} \int_0^T (X_{t,i})^2 dt.$$

By Cauchy-Schwarz inequality we have

$$\begin{aligned}
&\mathrm{Var}(U_i) \\
&= \frac{\lambda_i^4}{b_i^4} E\left(\int_0^T \{(X_{t,i})^2 - E(X_{t,i})^2\} dt \right)^2 \\
&\le T \frac{\lambda_i^4}{b_i^4} \int_0^T E\{((X_{t,i})^2 - E((X_{t,i})^2\}^2 dt \\
&= 2T\lambda_i^4 \int_0^T \left(\int_0^t e^{2(\theta\lambda_i - \xi a_i)(t-s)} ds \right)^2 dt
\end{aligned}$$

$$= \frac{T\lambda_i^4}{2(\theta\lambda_i - \xi a_i)^2}$$
$$\times \left(T - \frac{1}{(\theta\lambda_i - \xi a_i)}(e^{-2(\theta\lambda_i - \xi a_i)T} - 1) - \frac{1}{4(\theta\lambda_i - \xi a_i)}(e^{-4(\theta\lambda_i - \xi a_i)T} - 1)\right).$$

Hence

$$\sum_{i=1}^{n} \frac{\mathrm{Var}(U_i)}{g_{\theta\theta}^2(i)} \sim \frac{\sum_{i=1}^{n} \lambda_i^2}{(\sum_{i=1}^{n} \lambda_i)^2},$$

which implies that

$$\sum_{i=1}^{\infty} \frac{\mathrm{Var}(U_i)}{g_{\theta\theta}^2(i)} < \infty.$$

Then by the strong law of large numbers, we have

$$\lim_{n\to\infty} \frac{\alpha_n}{g_{\theta\theta}(n)} = 1 \quad P_{\theta,\xi} - (a.s.).$$

(b). Let

$$U_i = \frac{\lambda_i a_i}{b_i^2} \int_0^T (X_{t,i})^2 dt.$$

Using the same arguments as in (a), we have

$$\sum_{i=1}^{n} \frac{\mathrm{Var}(U_i)}{g_{\theta\xi}^2(i)} \sim \sum_{i=1}^{n} \frac{a_i^2}{(\sum_{k=1}^{i} a_k)^2} \le \sup_i a_i \sum_{i=1}^{n} \frac{a_i}{(\sum_{k=1}^{i} a_k)^2}$$
$$\le \sup_i a_i \frac{1}{a_1} < \infty.$$

Hence

$$\lim_{n\to\infty} \frac{\gamma_n}{g_{\theta\xi}(n)} = 1 \quad P_{\theta,\xi} - (a.s.).$$

(c). From$\sum_{k=1}^{\infty} a_i^2/\lambda_i < \infty$ in the assumption (C1), it follows that we have

$$E(\beta_n) = \frac{T}{2}\sum_{i=1}^{\infty} \frac{a_i^2}{(\theta\lambda_i - \xi a_i)} - \frac{1}{4}\sum_{i=1}^{\infty} \frac{a_i^2}{(\theta\lambda_i - \xi a_i)^2}(1 - e^{-2(\theta_0\lambda_i - \xi_0 a_i)T}) < \infty.$$

(d). From (c), it is immediate that

$$\lim_{n\to\infty} E(\rho_n)^2 < \infty.$$

Hence the result follows.

(e). Note that $\{\delta_n, n = 1, 2, \cdots\}$ is a locally square integrable martingale with respect to a filtration $\mathcal{F}_n := \sigma\{X_1, \ldots, X_n\}$. The quadratic variation of the process $\{\delta_n\}$ is equal to α_n. It follows from (a) that $\lim_{n\to\infty} \alpha_n = \infty$. Hence by Martingale convergence theorem,

$$\lim_{n\to\infty} \frac{\delta_n}{\alpha_n} = \lim_{n\to\infty} \frac{\delta_n}{<\delta>_n} = 0,$$

where $<\delta>_n$ is the quadratic variation process of $\{\delta_n, n = 1, 2, \cdots\}$. □

Using the above lemma, we show that the MLE $\hat{\theta}^{(n)}$ is strongly consistent.

Theorem 4.1. *The MLE $\hat{\theta}^{(n)}$ is strongly consistent, i.e.,*

$$\lim_{n\to\infty} \hat{\theta}^{(n)} = \theta,$$

but $\hat{\xi}^{(n)}$ is not consistent,

$$\lim_{n\to\infty} \hat{\xi}^{(n)} = \xi + \frac{\rho_\infty - \sqrt{d}}{\beta_\infty - d} \quad \textit{in probability},$$

where $\rho_\infty = \lim_{n\to\infty} \rho_n$ and $\beta_\infty = \lim_{n\to\infty} \beta_n$.

Proof. Write $\hat{\theta}^{(n)} - \theta$ as follows:

$$\begin{aligned}\hat{\theta}^{(n)} - \theta &= \frac{\gamma_n \rho_n - \beta_n \delta_n}{\alpha_n \beta_n - (\gamma_n)^2} \\ &= \left(\frac{\gamma_n}{\alpha_n}\right) \frac{\rho_n}{\beta_n - (\gamma_n)^2/\alpha_n} - \left(\frac{\delta_n}{\alpha_n}\right) \frac{\beta_n}{\beta_n - (\gamma_n)^2/\alpha_n}.\end{aligned}$$

From Lemma 4.1, we have

$$\begin{aligned}&\lim_{n\to\infty} \frac{\gamma_n}{\alpha_n} = 0, \quad \lim_{n\to\infty} \frac{\rho_n}{\beta_n - (\gamma_n)^2/\alpha_n} < \infty, \\ &\lim_{n\to\infty} \frac{\delta_n}{\alpha_n} = 0, \quad \lim_{n\to\infty} \frac{\beta_n}{\beta_n - (\gamma_n)^2/\alpha_n} < \infty.\end{aligned}$$

Hence $\hat{\theta}^{(n)} \to \theta$ (a.s.). Now we show that $\hat{\xi}^{(n)}$ is not consistent. First we write

$$\frac{\gamma_n \delta_n}{\alpha_n} = \frac{\gamma_n}{\sqrt{\alpha_n}} \frac{\delta_n}{\sqrt{\alpha_n}} \sim \frac{\gamma_n}{g_{\theta\xi}(n)} \frac{g_{\theta\xi}(n)}{\sqrt{g_{\theta\theta}(n)}} \frac{\delta_n}{\sqrt{g_{\theta\theta}(n)}}. \tag{4.3}$$

From the condition (C4) and (b) in Lemma 4.1, it follows that

$$\frac{\gamma_n}{g_{\theta\xi}(n)} \to 1 \quad \text{and} \quad \frac{g_{\theta\xi}(n)}{\sqrt{g_{\theta\theta}(n)}} \to \sqrt{d}, \quad (a.s.). \tag{4.4}$$

By direct computation, we have that

$$E\left|\frac{\delta_n}{\sqrt{g_{\theta\theta}(n)}} - 1\right|^2 \to 0. \tag{4.5}$$

From (4.3), (4.4) and (4.5), it follows that as $n \to \infty$,

$$\frac{\gamma_n \delta_n}{\alpha_n} \to \sqrt{d} \text{ in probability.} \tag{4.6}$$

By the same arguments as before, it is easy to see that $(\gamma_n)^2/\alpha_n$ converges in probability to d as $n \to \infty$. Hence $\hat{\xi}^{(n)}$ converges in probability to

$$\xi + \frac{\sum_{i=1}^{\infty}(a_i/b_i)\int_0^T X_{t,i}\, dW_t(\phi_i) - \sqrt{d}}{\sum_{i=1}^{\infty}(a_i^2/b_i^2)\int_0^T (X_{t,i})^2\, dt - d}.$$

□

The following result will be used for studying the asymptotic distribution of MLE. Denote

$$Y_t^n = \sum_{i=1}^{n}\int_0^t (h_i^n(t))^2\, dW_t(\phi_i) \quad \text{and} \quad \mathcal{F}_t^n = \sigma\{X_1(s), \cdots, X_n(s), 0 \le s \le t\}.$$

Then $Y^n = (Y_t^n, \mathcal{F}_t^n)$, $n \ge 1$ is a martingale.

Lemma 4.2. *If $< Y^n >_T \to < Y >_T$ in probability as $n \to \infty$, then*

$$Y_T^n \to Y_T \quad \textit{in distribution as} \quad n \to \infty,$$

where $< Y >_T$ is the quadratic variation process of Y_T with the characteristic function.

$$Ee^{i\lambda Y_T} = E\exp\left\{-\frac{\lambda^2}{2} < Y >_T\right\}, \quad \lambda \in R. \tag{4.7}$$

One can derive this Lemma from Theorem 5.5.1 in Liptser and Shiryayev (1989) by taking the limiting process $X = (X_t, \mathcal{F}_t)$ where $X_t = \frac{1}{\sqrt{T}} < Y >_T W_t$ and $(W_t, \mathcal{F}_t)$ is a Wiener process, independent of $< Y >_T$. If $< Y^n >_T \to \sigma^2$ in probability as $n \to \infty$, where σ is a positive constant,

$$Y_T^n \to \zeta \text{ in distribution as } n \to \infty,$$

where ζ is a Gaussian random variable with parameter $(0, \sigma^2)$. In this case we take the limiting process $X_s = \zeta \mathbf{1}_{t \leq s}$ and $\mathcal{F}_s = \sigma\{X_u, 0 \leq u \leq s\}$.

Now we discuss the problem of asymptotic normality of the MLE $\hat{\theta}^{(n)}$. First we introduce the notation

$$\bar{g}_{\theta\theta}(n) = g_{\theta\theta}(n) - \frac{(g_{\theta\xi}(n))^2}{g_{\xi\xi}(n)},$$

which is called the orthogonalized Fisher information. This plays the role of the Fisher information in the presence of the nuisance parameter ξ.

Theorem 4.2. *If $\sum_{n=1}^{\infty}(\lambda_n^2/s_n^2) < \infty$, then*

$$\bar{g}_{\theta\theta}^{1/2}(n)(\hat{\theta}^{(n)} - \theta) \to Y_T \quad \textit{in distribution as} \quad n \to \infty,$$

where Y_T is a random variable with the characteristic function given in (4.7). In particular, if d=0 in the condition (C4),

$$\sqrt{\sum_{i=1}^{n} \lambda_i}(\hat{\theta}^{(n)} - \theta) \to \zeta \quad \textit{in distribution as} \quad n \to \infty,$$

where ζ is a Gaussian random variable with parameter (0,1).

Proof. Using the notations in Lemma 4.2, we set

$$h_i^n(t) := \bar{g}_{\theta\theta}^{1/2}(n)\frac{(\gamma_n(a_i/b_i)X_{t,i} - \beta_n(\lambda_i/b_i)X_{t,i})}{\alpha_n\beta_n - (\gamma_n)^2}.$$

Then from (4.1), we have

$$\bar{g}_{\theta\theta}^{1/2}(n)(\hat{\theta}^{(n)} - \theta) = \sum_{i=1}^{n}\int_0^T h_i^n(t)\, dW_t(\phi_i). \tag{4.8}$$

Denote

$$Y_t^n = \sum_{i=1}^{n}\int_0^t h_i^n(t)\, dW_t(\phi_i).$$

Then it is obvious that $Y^n = (Y_t^n, \mathcal{F}_t^n)$, $n \geq 1$, is a martingale. The quadratic variation of Y_T^n is given by

$$< Y^n >_T = \sum_{i=1}^{n}\int_0^T (h_i^n(t))^2 dt = \frac{\beta_n \bar{g}_{\theta\theta}(n)}{\alpha_n\beta_n - (\gamma_n)^2}.$$

From the condition (C4) and Lemma 4.1 it follows that

$$\frac{\beta_n \bar{g}_{\theta\theta}(n)}{\alpha_n\beta_n - (\gamma_n)^2} = \frac{g_{\theta\theta}(n)}{\alpha_n}\left(\frac{1-(g_{\theta\xi}(n))^2/(g_{\theta\theta}(n)g_{\xi\xi}(n))}{1-(\gamma_n)^2/(\alpha_n\beta_n)}\right) \to \eta \;\; P-(a.s.),$$

where

$$\eta = \frac{1-Td/(2\theta g_{\xi\xi}(\infty))}{1-Td/(2\theta\beta_\infty)}.$$

Hence from Lemma 4.2 the result follows. On the other hand, if $d = 0$, then $\eta = 1$. From $\bar{g}_{\theta\theta}(n) = g_{\theta\theta}(n) + o(g_{\theta\theta}(n))$ and Lemma 4.2,

$$\sqrt{\sum_{i=1}^{n}\lambda_i}(\hat{\theta}^{(n)} - \theta) \to \zeta \;\text{ in distribution as }\; n \to \infty,$$

where ζ is a Gaussian random variable with parameter (0,1). □

We give an example to illustrate our main results.

Example 4.1. Now we are going to consider the SPDE (1.2) given in Introduction which is a special case of the Hilbert space valued SDE (2.1), *i.e.*,

$$L = -\Delta, A = I, B = (I-\Delta)^{-1/2} \;\text{ and }\; X_0 = 0,$$

where I is an identity operator. In this case $a_i = 1$, $b_i = 1/\sqrt{i^2+1}$, $\lambda_i = i^2$, and $\phi_i(x) = \sin ix$, and the parameter space is given by

$$\Theta = \{(\theta,\xi)|\theta > 0 \;\text{ and }\; \theta > \xi\}.$$

The solution of the equation (1.2) can be defined as

$$u(t,x) = \sum_{i=1}^{\infty} u_i(t)\phi_i(x),$$

where the Fourier coefficients $u_i(t)$ satisfy the equations

$$du_i(t) = -(\theta i^2 - \xi)u_i(t)dt + \frac{1}{\sqrt{i^2+1}}dW_i(t). \tag{4.9}$$

with initial conditions $u_i(0) = 0$.

Then it is easy to check that the sequences $\{a_i\}$, $\{b_i\}$ and $\{\lambda_i\}$ satisfy the conditions (C1), (C2) and (C3). Moreover, since

$$n\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n \lambda_i} = \frac{6n}{(n+1)(2n+1)} \to 0 \;\text{ as }\; n \to \infty,$$

the first condition in (C4) is also satisfied. Note that $d = 0$ in the condition (C4), since

$$\frac{(\sum_{i=1}^{n} a_i)^2}{\sum_{i=1}^{n} \lambda_i} = \frac{6n^2}{n(n+1)(2n+1)} \to 0 \quad \text{as} \quad n \to \infty.$$

It is obvious that $\sum_{n=1}^{\infty}(\lambda_n^2/s_n^2) < \infty$. Hence it follows from Theorem 4.2 that

$$\sqrt{\frac{n(n+1)(2n+1)}{6}}(\hat{\theta}^{(n)} - \theta) \to \mathcal{N}(0,1) \quad \text{in distribution as} \quad n \to \infty,$$

where $\hat{\theta}^{(n)}$ is the MLE based on $\{u_i(t); i = 1, \dots, n\}$. □

References

Bhatt, A. G., Kallianpur, G., Karandikar R. L. and Xiong, J. (1993). "On interacting systems of Hilbert space valued diffusions", Technical Report 373, Center for Stochastic Processes, Department of Statistics, University of North Carolina.

Huebner, M., Khasminskii, R. and Rozovskii, B. (1992). "Two examples of parameter estimation", In *Stochastic Processes* (Cambanis, Ghosh, Karandikar and Sen, eds.), Springer, Berlin.

Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*, Amsterdam, North Holland.

Kallianpur, G. and Kim, Y. T. (1998). "A curious example from statistical differential geometry", *Theory of Probability and Its Applications*, **43**, 42–62.

Kim, Y. T. and Lee, K. W. (2000). "Parameter estimation for a Hilbert space-valued stochastic differential equation", *Journal of the Korean Statistical Society*, **29**, 375–384.

Liptser, R. Sh. and Shiryayev, A. N. (1989). *Theory of Martingales*, Kluwer Academic Publishers.

Shiryaev, A. N. (1984). *Probability*, Springer-Verlag, New York.