

# MARTINGALE ESTIMATING FUNCTIONS BASED ON EIGENFUNCTIONS FOR DISCRETELY OBSERVED SMALL DIFFUSIONS

Uchida, Masayuki  
Faculty of Mathematics, Kyushu University

<https://doi.org/10.5109/12597>

---

出版情報 : Bulletin of informatics and cybernetics. 38, pp.1-13, 2006-12. Research Association  
of Statistical Sciences

バージョン :

権利関係 :



MARTINGALE ESTIMATING FUNCTIONS BASED ON  
EIGENFUNCTIONS FOR DISCRETELY OBSERVED SMALL  
DIFFUSIONS

by

Masayuki UCHIDA

---

*Reprinted from the Bulletin of Informatics and Cybernetics  
Research Association of Statistical Sciences, Vol.38*

---

FUKUOKA, JAPAN  
2006

# MARTINGALE ESTIMATING FUNCTIONS BASED ON EIGENFUNCTIONS FOR DISCRETELY OBSERVED SMALL DIFFUSIONS

By

Masayuki UCHIDA\*

## Abstract

We consider asymptotic properties of an estimator of a drift parameter for a one-dimensional diffusion process with small dispersion parameter  $\varepsilon$ . For discrete data observed at equidistant times  $k/n$ ,  $k = 0, 1, \dots, n$ , we study consistency and asymptotic normality of an M-estimator derived from a martingale estimating function based on an eigenfunction as  $\varepsilon$  tends to 0 and  $n$  tends to  $\infty$  simultaneously.

*Key Words and Phrases:* Diffusion process with small noise, Discrete time observation, Parametric inference.

## 1. Introduction

We consider statistical inference for a class of one-dimensional diffusion processes defined as solutions of the stochastic differential equations

$$\begin{aligned} dX_t &= b(X_t, \theta)dt + \varepsilon\sigma(X_t)dw_t, \quad t \in [0, 1], \quad \varepsilon \in (0, 1], \\ X_0 &= x_0, \end{aligned} \tag{1}$$

where  $x_0$  and  $\varepsilon$  are known constants and  $w$  is an  $r$ -dimensional standard Wiener process. Further,  $b$  is an  $\mathbf{R}$ -valued function defined on  $\mathbf{R} \times \bar{\Theta}$ ,  $\sigma$  is an  $\mathbf{R}^r$ -valued function defined on  $\mathbf{R}$ ,  $\Theta$  is an open bounded convex subset of  $\mathbf{R}^p$  and  $\bar{\Theta}$  denotes the closure of  $\Theta$ . We assume that the drift  $b$  is known apart from the parameter  $\theta$ . The types of data considered in the present paper are observations of  $X$  at  $n$  regularly spaced time points:  $(X_{t_k})_{0 \leq k \leq n}$  with  $t_k = k/n$ . The asymptotics we treat is when  $\varepsilon \rightarrow 0$  and  $n \rightarrow \infty$  simultaneously.

Diffusion processes with small noise have a lot of applications in various fields. Applications to mathematical finance were given in Yoshida (1992c), Kunitomo and Takahashi (2001) and Uchida and Yoshida (2004). The asymptotic theory of parametric inference for diffusion processes with small noise is well developed. For continuous-time observations, see Kutoyants (1984, 1994) and Yoshida (1992a, 2003). For discrete observations, see Genon-Catalot (1990), Laredo (1990), Sørensen (2000), Sørensen and Uchida (2003) and Uchida (2003). Recently, Uchida (2004) derived an explicit estimating function by using the expansion in Florens-Zmirou (1989). His estimating function is asymptotically equivalent to the martingale estimating function proposed by Bibby and

---

\* Faculty of Mathematics, Kyushu University Ropponmatsu 4-2-1 Chuo-Ku Fukuoka 810-8560 Japan.  
tel +81-92-726-4785 uchida@math.kyushu-u.ac.jp

Sørensen (1995) under  $(\varepsilon n^l)^{-1} \rightarrow 0$  for a positive integer  $l$ . He also proved that the estimator obtained from the approximate martingale estimating function has asymptotic efficiency under  $(\varepsilon n^l)^{-1} \rightarrow 0$  for a positive integer  $l$ . However, there is a disadvantage that the estimator may not be obtained explicitly when the approximate martingale estimating function is complicated. In order to overcome this difficulty, Matsuzaki and Uchida (2004) proposed one-step estimators based on the method of Newton-Raphson in numerical analysis. Using an initial estimator, they derived a one-step estimator with asymptotic efficiency. A sufficient condition for the initial estimator is that the initial estimator has asymptotic normality as  $\varepsilon \rightarrow 0$  and  $n \rightarrow \infty$ . Note that the initial estimator does not have to be asymptotically efficient. Therefore, we investigate explicit estimating functions to obtain estimators with asymptotic normality as  $\varepsilon \rightarrow 0$  and  $n \rightarrow \infty$ . The present paper is a contribution in this direction.

Kessler and Sørensen (1999) proposed explicit martingale estimating functions based on eigenfunctions for discretely observed ergodic diffusion processes. Though the estimators obtained from the estimating function may not be asymptotically efficient, they have very good asymptotic properties. We will apply the martingale estimating functions based on eigenfunctions to small diffusion models defined by (1). Another difference between Kessler and Sørensen (1999) and the present paper is asymptotics. Note that the asymptotics in the present paper is when  $\varepsilon \rightarrow 0$  and  $t_k - t_{k-1} = 1/n \rightarrow 0$  simultaneously, while the asymptotics in Kessler and Sørensen (1999) is when  $t_k - t_{k-1} = \Delta$  is fixed and  $n\Delta \rightarrow \infty$ .

In section 2, we construct the martingale estimating functions based on eigenfunctions. We also present asymptotic properties of M-estimators obtained from the martingale estimating functions. Section 3 gives two examples and simulation studies. Section 4 concludes. In section 5, we prove the results stated in section 2.

## 2. Martingale estimating functions based on eigenfunctions

We first state notation used in this paper. Let  $\theta_0 \in \Theta$  denote the true value of  $\theta$ . Let  $X_t^0$  be the solution of the ordinary differential equation corresponding to  $\varepsilon = 0$ , i.e.  $dX_t^0 = b(X_t^0, \theta_0)dt$ ,  $X_0^0 = x_0$ . For a matrix  $A$ ,  $A^*$  is the transpose of  $A$ . We denote by  $C_{\uparrow}^{\infty, k}(\mathbf{R} \times \Theta \times (0, 1]; \mathbf{R})$  the space of all functions  $f$  satisfying the following two conditions: (i)  $f(x, \theta, \varepsilon)$  is an  $\mathbf{R}$ -valued function on  $\mathbf{R} \times \Theta \times (0, 1]$  that is infinitely differentiable with respect to  $x$  and continuously differentiable with respect to  $\theta$  up to order  $k$ , (ii) for  $n \geq 0$  and  $0 \leq |\nu| \leq k$ , there exists  $C > 0$  such that  $\sup_{\theta \in \Theta} |\delta^\nu \partial_x^n f| \leq C(1 + |x|)^C$  for all  $x, \varepsilon$ . Here  $\partial_x = \partial/\partial x$  and  $\nu = (\nu^1, \dots, \nu^p)$  is a multi-index,  $|\nu| = \nu^1 + \dots + \nu^p$ ,  $\delta^\nu = \delta_1^{\nu^1} \dots \delta_p^{\nu^p}$ ,  $\delta_j = \partial/\partial \theta^j$ ,  $j = 1, \dots, p$ . Likewise, let  $C_{\uparrow}^{\infty, k}(\mathbf{R} \times \Theta; \mathbf{R})$  be the space of all functions  $f$  satisfying the following two conditions: (i)  $f(x, \theta)$  is an  $\mathbf{R}$ -valued function on  $\mathbf{R} \times \Theta$  that is infinitely differentiable with respect to  $x$  and continuously differentiable with respect to  $\theta$  up to order  $k$ , (ii) for  $n \geq 0$  and  $0 \leq |\nu| \leq k$ , there exists  $C > 0$  such that  $\sup_{\theta \in \Theta} |\delta^\nu \partial_x^n f| \leq C(1 + |x|)^C$  for all  $x$ . Let  $C_{\uparrow}^{\infty}(\mathbf{R}; \mathbf{R}^r)$  be the set of all functions  $f$  of class  $C^{\infty}(\mathbf{R}; \mathbf{R}^r)$  such that  $f$  and all of its derivatives have polynomial growth. Let  $C_b^k(\Theta \times (0, 1]; \mathbf{R})$  be the space of all functions  $f$  satisfying the following two conditions: (i)  $f(\theta, \varepsilon)$  is an  $\mathbf{R}$ -valued function on  $\Theta \times (0, 1]$  that is continuously differentiable with respect to  $\theta$  up to order  $k$ , (ii) for  $0 \leq |\nu| \leq k$ , there exists  $C > 0$  such that  $|\delta^\nu f| \leq C$  for all  $\theta, \varepsilon$ . We denote by  $\xrightarrow{p}$  and  $\xrightarrow{d}$  the convergence in probability and the convergence in distribution as  $\varepsilon \rightarrow 0$  and  $n \rightarrow \infty$ , respectively. Let  $P_\theta$  be the law of the solution of

(1). Let  $L_\theta$  be the infinitesimal generator of the diffusion (1): For  $g \in C^2(\mathbf{R})$ ,

$$L_\theta g(x) = b(x, \theta) \partial_x g(x) + \frac{1}{2} \varepsilon^2 [\sigma \sigma^*](x) \partial_x^2 g(x).$$

For a twice continuously differentiable function  $\phi(x, \theta, \varepsilon)$  with respect to  $x$ ,  $\phi(x, \theta, \varepsilon)$  is called an eigenfunction for  $L_\theta$  with eigenvalue  $\lambda(\theta, \varepsilon)$  if

$$L_\theta \phi(x, \theta, \varepsilon) = -\lambda(\theta, \varepsilon) \phi(x, \theta, \varepsilon)$$

for all  $x$  in the state space of  $X$  under  $P_\theta$ . For more details, see Kessler and Sørensen (1999).

In this paper we make the following assumptions.

- A1. For every  $\theta \in \bar{\Theta}$ , there exists a constant  $K_\theta > 0$  such that  $b^2(x, \theta) + |\sigma(x)|^2 \leq K_\theta(1 + x^2)$ .
- A2.  $b$  is Lipschitz on  $\mathbf{R} \times \bar{\Theta}$ .
- A3.  $b(x, \theta) \in C_{\uparrow}^{\infty, 3}(\mathbf{R} \times \bar{\Theta}; \mathbf{R})$ ,  $\sigma(x) \in C_{\uparrow}^{\infty}(\mathbf{R}; \mathbf{R}^r)$ .
- A4. There exists an eigenfunction  $\phi(x, \theta, \varepsilon) \in C_{\uparrow}^{\infty, 2}(\mathbf{R} \times \bar{\Theta} \times (0, 1]; \mathbf{R})$  with eigenvalue  $\lambda(\theta, \varepsilon) \in C_b^2(\bar{\Theta} \times (0, 1]; \mathbf{R})$  such that  $L_\theta \phi(x, \theta, \varepsilon) = -\lambda(\theta, \varepsilon) \phi(x, \theta, \varepsilon)$  for all  $x$  in the state space of  $X$  under  $P_\theta$ .
- A5.  $\inf_x [\sigma \sigma^*](x) > 0$ .

REMARK. (i) By a localization argument, the results which will be presented in this paper still hold under mild regularity conditions about  $b$ ,  $\sigma$  and  $\phi$  in the neighborhood of the path of  $X_t^0$  instead of A3–A5. (ii) By A4, we set  $\phi(x, \theta, 0) = \lim_{\varepsilon \rightarrow 0} \phi(x, \theta, \varepsilon)$  and  $\lambda(\theta, 0) = \lim_{\varepsilon \rightarrow 0} \lambda(\theta, \varepsilon)$ .

Under some regularity conditions, we have

$$E_\theta[\phi(X_{t_k}, \theta, \varepsilon) | X_{t_{k-1}}] = e^{-\lambda(\theta, \varepsilon)/n} \phi(X_{t_{k-1}}, \theta, \varepsilon).$$

For details of the regularity conditions, see section 5 in Kessler and Sørensen (1999). We then obtain the following martingale estimating function: for  $i = 1, 2, \dots, p$ ,

$$\begin{aligned} M_{\varepsilon, n}^i(\theta) &= \sum_{k=1}^n (\delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) (\phi(X_{t_k}, \theta, \varepsilon) - E_\theta[\phi(X_{t_k}, \theta, \varepsilon) | X_{t_{k-1}}]) \\ &= \sum_{k=1}^n (\delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) \left[ \phi(X_{t_k}, \theta, \varepsilon) - e^{-\lambda(\theta, \varepsilon)/n} \phi(X_{t_{k-1}}, \theta, \varepsilon) \right]. \end{aligned}$$

For results put later on, we introduce several functions. Let  $K_{\varepsilon, n}(\theta) = (K_{\varepsilon, n}^{ij}(\theta))_{1 \leq i, j \leq p}$ , where  $K_{\varepsilon, n}^{ij}(\theta) = \delta_j M_{\varepsilon, n}^i(\theta)$ . Let  $K(\theta) = (K^{ij}(\theta))_{1 \leq i, j \leq p}$  and  $A(\theta) = (A^{ij}(\theta))_{1 \leq i, j \leq p}$ , where

$$K^{ij}(\theta) = \int_0^1 (\delta_j \delta_i b)(X_s^0, \theta) [\sigma \sigma^*]^{-1}(X_s^0) b(X_s^0, \theta_0) \partial_x \phi(X_s^0, \theta, 0) ds$$

$$\begin{aligned}
& + \int_0^1 (\delta_j \delta_i b)(X_s^0, \theta) [\sigma \sigma^*]^{-1}(X_s^0) \lambda(\theta, 0) \phi(X_s^0, \theta, 0) ds \\
& + \int_0^1 (\delta_i b)(X_s^0, \theta) [\sigma \sigma^*]^{-1}(X_s^0) b(X_s^0, \theta_0) \partial_x \delta_j \phi(X_s^0, \theta, 0) ds \\
& + \int_0^1 (\delta_i b)(X_s^0, \theta) [\sigma \sigma^*]^{-1}(X_s^0) \lambda(\theta, 0) \delta_j \phi(X_s^0, \theta, 0) ds \\
& + \int_0^1 (\delta_i b)(X_s^0, \theta) [\sigma \sigma^*]^{-1}(X_s^0) (\delta_j \lambda)(\theta, 0) \phi(X_s^0, \theta, 0) ds,
\end{aligned}$$

and

$$A^{ij}(\theta_0) = \int_0^1 (\delta_i b)(X_s^0, \theta_0) [\sigma \sigma^*]^{-1}(X_s^0) (\partial_x \phi)^2(X_s^0, \theta_0, 0) (\delta_j b)(X_s^0, \theta_0) ds.$$

In order to show asymptotic results of M-estimators, we need the following two lemmas.

LEMMA 2.1. *Under A1–A5,  $K_{\varepsilon,n}(\theta) \xrightarrow{p} K(\theta)$  uniformly in  $\theta \in \bar{\Theta}$ .*

LEMMA 2.2. *Under A1–A5,  $\varepsilon^{-1} M_{\varepsilon,n}(\theta_0) \xrightarrow{d} N(0, A(\theta_0))$ .*

Let  $\hat{\theta}_{\varepsilon,n}$  be an M-estimator defined as a solution of  $M_{\varepsilon,n}(\theta) = 0$ . The main theorem is as follows.

THEOREM 2.3. *Let  $\gamma \in (0, 1)$ . Suppose that A1–A5 hold true. Moreover, suppose that there exists an open set  $\tilde{\Theta}$  including  $\theta_0$  such that*

$$\inf_{\theta_1, \theta_2 \in \tilde{\Theta}, |x|=1} \left| x^* \left( \int_0^1 K(\theta_1 + s(\theta_2 - \theta_1)) ds \right) \right| > 0.$$

Then,

$$P_{\theta_0}[(\exists_1 \hat{\theta}_{\varepsilon,n} \in \tilde{\Theta} \text{ such that } M_{\varepsilon,n}(\hat{\theta}_{\varepsilon,n}) = 0) \text{ and } (|\hat{\theta}_{\varepsilon,n} - \theta_0| \leq \varepsilon^\gamma)] \rightarrow 1$$

and

$$\varepsilon^{-1}(\hat{\theta}_{\varepsilon,n} - \theta_0) \xrightarrow{d} N(0, K(\theta_0)^{-1} A(\theta_0) (K^*(\theta_0))^{-1}).$$

### 3. Examples

In this section, we consider two kinds of non-linear models and examine the asymptotic behaviour of the estimator proposed in section 2 through simulations. We generate 1000 independent sample paths with  $\theta = \theta_0$  (true parameter value) and the initial value  $x_0$ . In order to evaluate the estimator  $(\hat{\theta}_{\varepsilon,n})$ , we also examine the estimator in corollary 1 of Genon-Catalot (1990) and the estimator in Laredo (1990), which are called the simple estimator  $(\hat{\theta}_{\varepsilon,n}^{(S)})$  and Laredo's estimator  $(\hat{\theta}_{\varepsilon,n}^{(L)})$  respectively in this section. For each of the three estimators, its mean and standard deviation are computed. The simulations are done for each  $\varepsilon = 0.1, 0.05, 0.01$  and  $n = 5, 10, 50$ , with the Euler-Maruyama scheme, see Kloeden and Platen (1992).

### 3.1. Non-linear model

Consider a non-linear model defined by the following stochastic differential equation

$$\begin{aligned} dX_t &= -\theta \tan X_t dt + \varepsilon dw_t, \\ X_0 &= x_0, \end{aligned}$$

where  $x_0$  and  $\varepsilon$  are known constants and  $\theta$  is an unknown parameter. Here we assume that  $X_t \neq m\pi/2$  ( $m = 0, 1, \dots$ ). The non-trivial eigenfunction is  $\phi(x, \theta, \varepsilon) = \sin(x)$  and its eigenvalue is  $\lambda(\theta, \varepsilon) = \theta + \varepsilon^2/2$ . The martingale estimating function based on an eigenfunction is given by

$$M_{\varepsilon,n}(\theta) = -\sum_{k=1}^n \tan X_{t_{k-1}} \left[ \sin X_{t_k} - \exp \left\{ -\frac{1}{n} \left( \theta + \frac{\varepsilon^2}{2} \right) \right\} \sin X_{t_{k-1}} \right],$$

and one has the estimator

$$\hat{\theta}_{\varepsilon,n} = n \left\{ \log \left( \sum_{k=1}^n \tan X_{t_{k-1}} \sin X_{t_k} \right) - \log \left( \sum_{k=1}^n \tan X_{t_{k-1}} \sin X_{t_{k-1}} \right) \right\} - \frac{\varepsilon^2}{2}.$$

By Theorem 2.3, the asymptotic variance of  $\varepsilon^{-1}(\hat{\theta}_{\varepsilon,n} - \theta_0)$  is as follows.

$$K(\theta_0)^{-1} A(\theta_0) (K^*(\theta_0))^{-1} = \frac{\int_0^1 \sin^2 X_s^0 ds}{\left( \int_0^1 \frac{\sin X_s^0}{\cos X_s^0} ds \right)^2},$$

where  $X_s^0$  is a solution of the ordinary differential equation:  $dX_s^0 = -\theta_0 \tan X_s^0 ds$ .

Table 1 shows the means and standard deviations of simulated values. We set  $\theta = 2$  and  $x_0 = 1/2$ . When  $n \leq 10$ ,  $\hat{\theta}_{\varepsilon,n}^{(S)}$  has a considerable bias and  $\hat{\theta}_{\varepsilon,n}^{(L)}$  has a little small bias, while  $\hat{\theta}_{\varepsilon,n}$  is better than the others. Compared with those of  $\hat{\theta}_{\varepsilon,n}^{(S)}$  and  $\hat{\theta}_{\varepsilon,n}^{(L)}$ , the standard deviations of  $\hat{\theta}_{\varepsilon,n}$  get a little greater because  $\hat{\theta}_{\varepsilon,n}$  is not asymptotically efficient. However,  $\hat{\theta}_{\varepsilon,n}$  gives satisfactory approximation in all cases.

### 3.2. The radial Ornstein-Uhlenbeck process

We consider another non-linear model defined by the stochastic differential equation

$$\begin{aligned} dX_t &= \left( \frac{\theta}{X_t} - X_t \right) dt + \varepsilon dw_t, \\ X_0 &= x_0, \end{aligned}$$

where  $x_0$  and  $\varepsilon$  are known constants and  $\theta$  is an unknown parameter. Here we assume that the state space is the positive real line. This model is called the radial Ornstein-Uhlenbeck process. The eigenfunctions are  $\phi_i(x, \theta, \varepsilon) = L_i^{(\theta/\varepsilon^2 - 1/2)}(x^2/\varepsilon^2)$ , where  $L_i^{(\nu)}$  is the  $i$ th-order Laguerre polynomial with parameter  $\nu$ , and the associated eigenvalues are  $\{2i : i = 1, 2, \dots\}$  (Karlin and Taylor (1981), p.333). The first non-trivial eigenfunction is  $\phi(x, \theta, \varepsilon) = (\theta + \varepsilon^2/2) - x^2$  with eigenvalue  $\lambda(\theta, \varepsilon) = 2$ . We obtain the following martingale estimating function.

$$M_{\varepsilon,n}(\theta) = \sum_{k=1}^n \frac{1}{X_{t_{k-1}}} \left[ \left( \theta + \frac{\varepsilon^2}{2} \right) - X_{t_k}^2 - e^{-\frac{2}{n}} \left\{ \left( \theta + \frac{\varepsilon^2}{2} \right) - X_{t_{k-1}}^2 \right\} \right].$$

Table 1: (Non-linear model) The mean and standard deviation (s.d.) of the three estimators determined from 1000 independent simulated sample paths for  $\theta = 2$ ,  $x_0 = 1/2$ .

| $n$ | $\varepsilon$ | $\hat{\theta}_{\varepsilon,n}^{(S)}$ |          | $\hat{\theta}_{\varepsilon,n}^{(L)}$ |          | $\hat{\theta}_{\varepsilon,n}$ |          |
|-----|---------------|--------------------------------------|----------|--------------------------------------|----------|--------------------------------|----------|
|     |               | mean                                 | s.d.     | mean                                 | s.d.     | mean                           | s.d.     |
| 5   | 0.10          | 1.619732                             | 0.244242 | 1.916941                             | 0.350186 | 2.046181                       | 0.399759 |
| 5   | 0.05          | 1.604015                             | 0.122962 | 1.927802                             | 0.176971 | 2.009593                       | 0.196568 |
| 5   | 0.01          | 1.601109                             | 0.024705 | 1.934346                             | 0.035514 | 2.001343                       | 0.039226 |
| 10  | 0.10          | 1.820309                             | 0.306519 | 1.963445                             | 0.373302 | 2.050758                       | 0.393158 |
| 10  | 0.05          | 1.793241                             | 0.153053 | 1.976303                             | 0.188003 | 2.011086                       | 0.193620 |
| 10  | 0.01          | 1.786957                             | 0.030637 | 1.983474                             | 0.037636 | 2.001493                       | 0.038591 |
| 50  | 0.10          | 2.005406                             | 0.368952 | 1.974965                             | 0.377853 | 2.053692                       | 0.388012 |
| 50  | 0.05          | 1.965942                             | 0.183091 | 1.991505                             | 0.190903 | 2.011772                       | 0.191976 |
| 50  | 0.01          | 1.956236                             | 0.036568 | 2.000019                             | 0.038265 | 2.001479                       | 0.038305 |

We then have the estimator,

$$\hat{\theta}_{\varepsilon,n} = \frac{\sum_{k=1}^n X_{t_k}^2 X_{t_{k-1}}^{-1} - e^{-\frac{2}{n}} \sum_{k=1}^n X_{t_{k-1}}}{(1 - e^{-\frac{2}{n}}) \sum_{k=1}^n X_{t_{k-1}}^{-1}} - \frac{\varepsilon^2}{2}.$$

The asymptotic variance of  $\varepsilon^{-1}(\hat{\theta}_{\varepsilon,n} - \theta_0)$  is given by

$$K(\theta_0)^{-1} A(\theta_0) (K^*(\theta_0))^{-1} = \left( \int_0^1 \frac{1}{X_s^0} ds \right)^{-2},$$

where  $X_s^0$  is a solution of the ordinary differential equation,  $dX_t^0 = (\theta_0/X_t^0 - X_t^0) dt$ .

The results with  $\theta = 10$  and  $x_0 = 3$  are given in Table 2. Even if  $n = 5$ , both  $\hat{\theta}_{\varepsilon,n}^{(L)}$  and  $\hat{\theta}_{\varepsilon,n}$  are unbiased, while  $\hat{\theta}_{\varepsilon,n}^{(S)}$  has a small bias. Although  $\hat{\theta}_{\varepsilon,n}$  has a little greater standard deviation than the others, as the theory shows, it gives as good an approximation as  $\hat{\theta}_{\varepsilon,n}^{(L)}$ . When  $n = 50$ , there seems no difference between the three estimators.

#### 4. Conclusions

This paper has proposed a martingale estimating function based on an eigenfunction for a small diffusion process using the same technique as in Kessler and Sørensen (1999). Moreover, it has been shown that an estimator obtained from the estimating function has consistency and asymptotic normality. As seen from the assumption A4, in order to obtain the martingale estimating function, it is important to get an eigenfunction  $\phi(x, \theta, \varepsilon)$  with an eigenvalue  $\lambda(\theta, \varepsilon)$  such that  $L_\theta \phi(x, \theta, \varepsilon) = -\lambda(\theta, \varepsilon) \phi(x, \theta, \varepsilon)$ . Since this is the second order differential equation, in general, it is not easy to solve this equation. However, for several diffusion models, we can refer section 13 of chapter 15 in Karlin and Taylor (1981). See also Kessler and Sørensen (1999). From the viewpoint of theory, an extension of this paper to multivariate case is straightforward. In practice,



Table 2: (The radial Ornstein-Uhlenbeck process) The mean and standard deviation (s.d.) of the three estimators determined from 1000 independent simulated sample paths for  $\theta = 10$ ,  $x_0 = 3$ .

| $n$ | $\varepsilon$ | $\hat{\theta}_{\varepsilon,n}^{(S)}$ |          | $\hat{\theta}_{\varepsilon,n}^{(L)}$ |          | $\hat{\theta}_{\varepsilon,n}$ |          |
|-----|---------------|--------------------------------------|----------|--------------------------------------|----------|--------------------------------|----------|
|     |               | mean                                 | s.d.     | mean                                 | s.d.     | mean                           | s.d.     |
| 5   | 0.10          | 9.915631                             | 0.271403 | 10.003568                            | 0.299242 | 10.010793                      | 0.301986 |
| 5   | 0.05          | 9.908870                             | 0.135665 | 9.998501                             | 0.149627 | 10.005073                      | 0.150860 |
| 5   | 0.01          | 9.904811                             | 0.027127 | 9.994905                             | 0.029927 | 10.001282                      | 0.030152 |
| 10  | 0.10          | 9.966824                             | 0.286189 | 10.008460                            | 0.300001 | 10.011606                      | 0.301243 |
| 10  | 0.05          | 9.959136                             | 0.143081 | 10.003445                            | 0.150052 | 10.005410                      | 0.150511 |
| 10  | 0.01          | 9.954612                             | 0.028614 | 9.999736                             | 0.030019 | 10.001342                      | 0.030086 |
| 50  | 0.10          | 10.005790                            | 0.298375 | 10.010167                            | 0.300978 | 10.012232                      | 0.301648 |
| 50  | 0.05          | 9.997270                             | 0.149186 | 10.005154                            | 0.150546 | 10.005693                      | 0.150739 |
| 50  | 0.01          | 9.992323                             | 0.029837 | 10.001322                            | 0.030118 | 10.001394                      | 0.030135 |

however, it can be difficult to obtain an eigenfunction with an eigenvalue satisfying the multivariate version of the assumption A4. For a multivariate diffusion process with a small noise, we will need to consider a more convenient procedure.

## 5. Proofs

Let  $R$  be a function  $\bar{\Theta} \times (0, 1] \times \mathbf{R} \rightarrow \mathbf{R}$  for which there exists a constant  $C > 0$  such that  $|R(\theta, a, x)| \leq aC(1 + |x|)^C$  for all  $\theta, a, x$ . We define  $\mathcal{G}_k^n = \sigma(w_s; s \leq t_k)$ . Let  $\tilde{C}_{\uparrow}^{1,1}(\mathbf{R} \times \Theta; \mathbf{R})$  be the space of all functions  $f : \mathbf{R} \times \Theta \rightarrow \mathbf{R}$  such that (i)  $f$  is continuously differentiable with respect to  $x$  and  $\theta$ , (ii)  $f$  and its derivatives are of polynomial growth in  $x$  uniformly in  $\theta$ . Let  $Q(x, \theta) = \phi(x, \theta, \varepsilon) - \phi(X_{t_{k-1}}, \theta, \varepsilon)$ . In order to prove Lemmas 2.1 and 2.2, we will use the following two lemmas.

LEMMA 5.1. *Suppose that A1–A4 hold true. Then, for  $i = 1, \dots, p$ ,*

$$\begin{aligned} E_{\theta_0}[Q(X_{t_k}, \theta) | \mathcal{G}_{k-1}^n] &= \frac{1}{n} b(X_{t_{k-1}}, \theta_0) \partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon) \\ &\quad + R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right), \end{aligned} \quad (2)$$

$$\begin{aligned} E_{\theta_0}[\delta_i Q(X_{t_k}, \theta) | \mathcal{G}_{k-1}^n] &= \frac{1}{n} b(X_{t_{k-1}}, \theta_0) \partial_x \delta_i \phi(X_{t_{k-1}}, \theta, \varepsilon) \\ &\quad + R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right), \end{aligned} \quad (3)$$

$$\begin{aligned} E_{\theta_0}[(Q(X_{t_k}, \theta))^2 | \mathcal{G}_{k-1}^n] &= \frac{\varepsilon^2}{n} [\sigma \sigma^*](X_{t_{k-1}}) (\partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon))^2 \\ &\quad + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right), \end{aligned} \quad (4)$$

$$E_{\theta_0}[(\delta_i Q(X_{t_k}, \theta))^2 | \mathcal{G}_{k-1}^n] = \frac{\varepsilon^2}{n} [\sigma \sigma^*](X_{t_{k-1}}) (\partial_x \delta_i \phi(X_{t_{k-1}}, \theta, \varepsilon))^2 \quad (5)$$

$$+R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right).$$

LEMMA 5.2. Let  $f \in \tilde{C}_\uparrow^{1,1}(\mathbf{R} \times \bar{\Theta}; \mathbf{R})$ . Assume A1–A4. Then, uniformly in  $\theta \in \bar{\Theta}$ ,

$$(i) \quad \frac{1}{n} \sum_{k=1}^n f(X_{t_{k-1}}, \theta) \xrightarrow{p} \int_0^1 f(X_s^0, \theta) ds,$$

$$(ii) \quad \sum_{k=1}^n f(X_{t_{k-1}}, \theta) Q(X_{t_k}, \theta) \xrightarrow{p} \int_0^1 f(X_s^0, \theta) b(X_s^0, \theta_0) \partial_x \phi(X_s^0, \theta, 0) ds,$$

(iii) for  $i = 1, \dots, p$ ,

$$\sum_{k=1}^n f(X_{t_{k-1}}, \theta) (\delta_i Q)(X_{t_k}, \theta) \xrightarrow{p} \int_0^1 f(X_s^0, \theta) b(X_s^0, \theta_0) \partial_x \delta_i \phi(X_s^0, \theta, 0) ds.$$

**Proof of Lemma 5.1.** By the same method as the proof of Lemma 7 in Kessler (1997), we prove Lemma 5.1. An easy computation enables us to get

$$\begin{aligned} L_{\theta_0} Q(x, \theta) &= b(x, \theta_0) \partial_x Q(x, \theta) + \frac{1}{2} \varepsilon^2 [\sigma \sigma^*](x) \partial_x^2 Q(x, \theta) \\ &= b(x, \theta_0) \partial_x \phi(x, \theta, \varepsilon) + \frac{1}{2} \varepsilon^2 [\sigma \sigma^*](x) \partial_x^2 \phi(x, \theta, \varepsilon). \end{aligned}$$

It then follows that

$$\begin{aligned} E_{\theta_0}[Q(X_{t_k}, \theta) | \mathcal{G}_{k-1}^n] &= Q(X_{t_{k-1}}, \theta) + \frac{1}{n} L_{\theta_0} Q(X_{t_{k-1}}, \theta) \\ &\quad + \int_0^{\frac{1}{n}} \int_0^{u_1} E_{\theta_0}[L_{\theta_0}^2 Q(X_{t_{k-1}+u_2}, \theta) | \mathcal{G}_{k-1}^n] du_2 du_1 \\ &= \frac{1}{n} b(X_{t_{k-1}}, \theta_0) \partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon) \\ &\quad + R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right). \end{aligned}$$

This completes the proof of (2). Similarly, we can prove (3).  $\square$

Next, we see that

$$\begin{aligned} &E_{\theta_0}[(Q(X_{t_k}, \theta))^2 | \mathcal{G}_{k-1}^n] \\ &= (Q(X_{t_{k-1}}, \theta))^2 + \frac{1}{n} L_{\theta_0}(Q(X_{t_{k-1}}, \theta))^2 \\ &\quad + \int_0^{\frac{1}{n}} \int_0^{u_1} E_{\theta_0}[L_{\theta_0}^2 (Q(X_{t_{k-1}+u_2}, \theta))^2 | \mathcal{G}_{k-1}^n] du_2 du_1 \\ &= \frac{\varepsilon^2}{n} [\sigma \sigma^*](X_{t_{k-1}}) (\partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon))^2 + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right). \end{aligned}$$

This completes the proof of (4). Similarly, (5) can be proved.  $\square$

**Proof of Lemma 5.2.** (i) In the same way as in the proof of Lemma 8 in Kessler (1997), we can prove (i). See also Lemma 4-(i) in Uchida (2004).  $\square$

(ii) Let  $\xi_k(\theta) = f(X_{t_{k-1}}, \theta) \{Q(X_{t_k}, \theta) - b(X_{t_{k-1}}, \theta_0) \partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon)/n\}$ . It follows from Lemmas 5.1 and 5.2-(i) that

$$\sum_{k=1}^n E[\xi_k(\theta) | \mathcal{G}_{k-1}^n] = \sum_{k=1}^n \left\{ R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right) \right\} \xrightarrow{p} 0,$$

$$\sum_{k=1}^n E[(\xi_k(\theta))^2 | \mathcal{G}_{k-1}^n] = \sum_{k=1}^n \left\{ R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right) \right\} \xrightarrow{p} 0.$$

By Lemma 9 in Genon-Catalot and Jacod (1993), it can be proved that  $\sum_{k=1}^n \xi_{k,\varepsilon}(\theta) \xrightarrow{p} 0$ . In order to show the tightness of  $\sum_{k=1}^n \xi_{k,\varepsilon}(\cdot)$ , it suffices to prove the following inequalities (cf. Theorem 20 in Appendix I in Ibragimov and Has'minskii (1981) or Lemma 3.1 in Yoshida (1990)): there exists a constant  $m$  such that  $m > p/2$ , and for any  $\theta, \theta_1, \theta_2 \in \bar{\Theta}$ ,

$$E_{\theta_0} \left[ \left( \sum_{k=1}^n \xi_{k,\varepsilon}(\theta) \right)^{2m} \right] \leq C, \quad (6)$$

$$E_{\theta_0} \left[ \left( \sum_{k=1}^n \xi_{k,\varepsilon}(\theta_2) - \sum_{k=1}^n \xi_{k,\varepsilon}(\theta_1) \right)^{2m} \right] \leq C |\theta_2 - \theta_1|^{2m}, \quad (7)$$

where  $C$  is a constant independent of  $\theta, \theta_1, \theta_2, \varepsilon$  and  $n$ . By Ito's formula, one has

$$\begin{aligned} \xi_k(\theta) &= f(X_{t_{k-1}}, \theta) \int_{t_{k-1}}^{t_k} L_{\theta_0} \phi(X_{t_k}, \theta, \varepsilon) ds \\ &\quad + \varepsilon f(X_{t_{k-1}}, \theta) \int_{t_{k-1}}^{t_k} \sum_{j=1}^r (\partial_x \phi)(X_s, \theta, \varepsilon) \sigma^j(X_s) dw_s^j \\ &\quad - \frac{1}{n} f(X_{t_{k-1}}, \theta) b(X_{t_{k-1}}, \theta_0) \partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon) \\ &=: A_{k,1}(\theta) + A_{k,2}(\theta) - A_{k,3}(\theta) \quad (\text{say}). \end{aligned}$$

In the same way as the proof of Lemma 1 in Yoshida (1992b), we will estimate  $A_{k,1}(\theta)$ ,  $A_{k,2}(\theta)$  and  $A_{k,3}(\theta)$ . By Lemma 6 in Kessler (1997),

$$\begin{aligned} &E_{\theta_0} \left[ \left| \sum_{k=1}^n A_{k,1}(\theta) \right|^{2m} \right] \\ &\leq n^{2m-1} \sum_{k=1}^n E_{\theta_0} \left[ \left( \int_{t_{k-1}}^{t_k} |f(X_{t_{k-1}}, \theta) L_{\theta_0} \phi(X_{t_k}, \theta, \varepsilon)| ds \right)^{2m} \right] \\ &\leq \sum_{k=1}^n \int_{t_{k-1}}^{t_k} E_{\theta_0} [|f(X_{t_{k-1}}, \theta)|^{2m} E_{\theta_0} [|L_{\theta_0} \phi(X_{t_k}, \theta, \varepsilon)|^{2m} | \mathcal{G}_{k-1}^n]] ds \\ &\leq n \cdot \frac{1}{n} \cdot C. \end{aligned}$$

It follows from the Burkholder-Davis-Gundy inequality that

$$\begin{aligned}
& E_{\theta_0} \left[ \left| \sum_{k=1}^n A_{k,2}(\theta) \right|^{2m} \right] \\
& \leq C_{2m} \varepsilon^{2m} E_{\theta_0} \left[ \left( \sum_{k=1}^n \int_{t_{k-1}}^{t_k} f(X_{t_{k-1}}, \theta)^2 (\partial_x \phi)(X_s, \theta, \varepsilon) [\sigma \sigma^*](X_s) ds \right)^m \right] \\
& \leq C_{2m} \varepsilon^{2m} \sum_{k=1}^n \int_{t_{k-1}}^{t_k} E_{\theta_0} [f(X_{t_{k-1}}, \theta)^{2m} E_{\theta_0} [((\partial_x \phi)(X_s, \theta, \varepsilon) [\sigma \sigma^*](X_s))^m | \mathcal{G}_{k-1}^n]] ds \\
& \leq C_{2m} \varepsilon^{2m} C,
\end{aligned}$$

where  $C_m$  denotes a constant which depends on an integer  $m$ . Moreover, we see that

$$\begin{aligned}
E_{\theta_0} \left[ \left| \sum_{k=1}^n A_{k,3}(\theta) \right|^{2m} \right] & \leq \frac{1}{n} \sum_{k=1}^n E_{\theta_0} \left[ |f(X_{t_{k-1}}, \theta) b(X_{t_{k-1}}, \theta_0) \partial_x \phi(X_{t_{k-1}}, \theta, \varepsilon)|^{2m} \right] \\
& \leq C.
\end{aligned}$$

Consequently, we obtain the inequality (6) and similarly the inequality (7) can be proved. This completes the proof of (ii). By the same fashion, we can prove (iii).  $\square$

**Proof of Lemma 2.1.** It follows from an easy computation that

$$\begin{aligned}
K_{\varepsilon, n, l}^{ij}(\theta) &= \sum_{k=1}^n (\delta_j \delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) [\phi(X_{t_k}, \theta, \varepsilon) - \phi(X_{t_{k-1}}, \theta, \varepsilon)] \\
&\quad + \sum_{k=1}^n (\delta_j \delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) (1 - e^{-\lambda(\theta, \varepsilon)/n}) \phi(X_{t_{k-1}}, \theta, \varepsilon) \\
&\quad + \sum_{k=1}^n (\delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) [\delta_j \phi(X_{t_k}, \theta, \varepsilon) - \delta_j \phi(X_{t_{k-1}}, \theta, \varepsilon)] \\
&\quad + \sum_{k=1}^n (\delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) (1 - e^{-\lambda(\theta, \varepsilon)/n}) \delta_j \phi(X_{t_{k-1}}, \theta, \varepsilon) \\
&\quad + \sum_{k=1}^n (\delta_i b)(X_{t_{k-1}}, \theta) [\sigma \sigma^*]^{-1}(X_{t_{k-1}}) \frac{(\delta_j \lambda)(\theta, \varepsilon)}{n} e^{-\lambda(\theta, \varepsilon)/n} \phi(X_{t_{k-1}}, \theta, \varepsilon).
\end{aligned}$$

By Lemma 5.2, we complete the proof.  $\square$

**Proof of Lemma 2.2.** The predictable quadratic variation of the martingale  $\varepsilon^{-1} M_{\varepsilon, n}(\theta_0)$  is

$$\begin{aligned}
& \varepsilon^{-2} \langle M^i(\theta_0), M^j(\theta_0) \rangle_n \\
&= \varepsilon^{-2} \sum_{k=1}^n (\delta_i b)(X_{t_{k-1}}, \theta_0) [\sigma \sigma^*]^{-2}(X_{t_{k-1}}) v(X_{t_{k-1}}, \theta_0) (\delta_j b)(X_{t_{k-1}}, \theta_0),
\end{aligned}$$

where  $v(X_{t_{k-1}}, \theta_0) = E_{\theta_0}[(\phi(X_{t_k}, \theta_0, \varepsilon) - e^{-\lambda(\theta_0, \varepsilon)/n} \phi(X_{t_{k-1}}, \theta_0, \varepsilon))^2 | X_{t_{k-1}}]$ . Note that

$$L_{\theta_0}(\phi(x, \theta_0, \varepsilon))^2 = 2(-\lambda(\theta_0, \varepsilon))(\phi(x, \theta_0, \varepsilon))^2 + \varepsilon^2[\sigma\sigma^*](x)(\partial_x \phi)^2(x, \theta_0, \varepsilon),$$

and that for  $l \geq 2$ ,

$$\begin{aligned} L_{\theta_0}^l(Q(x, \theta_0))^2 &= 2(-\lambda(\theta_0, \varepsilon))^l \phi(x, \theta_0, \varepsilon) Q(x, \theta_0) + (2^l - 2)(-\lambda(\theta_0, \varepsilon))^l (\phi(x, \theta_0, \varepsilon))^2 \\ &\quad + R(\theta_0, \varepsilon^2, x). \end{aligned}$$

By using the same method as the proof of Lemma 5.1, it follows that for any  $m \geq 1$ ,

$$\begin{aligned} v(X_{t_{k-1}}, \theta_0) &= E_{\theta_0}[(Q(X_{t_k}, \theta_0) + (1 - e^{-\lambda(\theta_0, \varepsilon)/n})\phi(X_{t_{k-1}}, \theta_0, \varepsilon))^2 | X_{t_{k-1}}] \\ &= E_{\theta_0}[(Q(X_{t_k}, \theta_0))^2 | X_{t_{k-1}}] - ((1 - e^{-\lambda(\theta_0, \varepsilon)/n})\phi(X_{t_{k-1}}, \theta_0, \varepsilon))^2 \\ &= \sum_{l=1}^{2m} \frac{1}{l!n^l} L_{\theta_0}^l(Q(X_{t_{k-1}}, \theta_0))^2 \\ &\quad - (1 - 2e^{-\lambda(\theta_0, \varepsilon)/n} + e^{-2\lambda(\theta_0, \varepsilon)/n})(\phi(X_{t_{k-1}}, \theta_0, \varepsilon))^2 \\ &\quad + R\left(\theta_0, \frac{1}{n^{2m+1}}, X_{t_{k-1}}\right) \\ &= \frac{\varepsilon^2}{n} [\sigma\sigma^*](X_{t_{k-1}})(\partial_x \phi)^2(X_{t_{k-1}}, \theta_0, \varepsilon) + \sum_{l=2}^{2m} \frac{1}{l!n^l} L_{\theta_0}^l(Q(X_{t_{k-1}}, \theta_0))^2 \\ &\quad - \sum_{l=2}^{2m} \frac{(2^l - 2)(-\lambda(\theta_0, \varepsilon))^l}{l!n^l} (\phi(X_{t_{k-1}}, \theta_0, \varepsilon))^2 + R\left(\theta_0, \frac{1}{n^{2m+1}}, X_{t_{k-1}}\right) \\ &= \frac{\varepsilon^2}{n} [\sigma\sigma^*](X_{t_{k-1}})(\partial_x \phi)^2(X_{t_{k-1}}, \theta_0, \varepsilon) \\ &\quad + R\left(\theta_0, \frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right) + R\left(\theta_0, \frac{1}{n^{2m+1}}, X_{t_{k-1}}\right). \end{aligned}$$

Therefore,  $\varepsilon^{-2} < M^i, M^j >_n \xrightarrow{p} A^{ij}(\theta_0)$ . By the central limit theorem for martingales, one has that  $\varepsilon^{-1}M_{\varepsilon,n}(\theta_0) \xrightarrow{d} N(0, A(\theta_0))$ . This completes the proof.  $\square$

**Proof of Theorem 2.3.** Now that Lemmas 2.1 and 2.2 have been already proved, it is not difficult to prove Theorem 2.3. Using the same method as the proof of Theorem 6.1 in Sakamoto and Yoshida (1999, 2004), we can show the existence, the uniqueness and the consistency of  $\hat{\theta}_{\varepsilon,n}$ . For the asymptotic normality of  $\hat{\theta}_{\varepsilon,n}$ , we can prove it along the same lines as the proof of Theorem 1 in Uchida (2004). This completes the proof.  $\square$

### Acknowledgement

The author wishes to thank the referee and the editor for their valuable comments. This work is partially supported by the Research Fund for Scientists of the Ministry of Education, Culture, Sports, Science and Technology, and by Kyushu University 21st Century COE Program, Development of Dynamic Mathematics with High Functionality, of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

- Bibby, B. M. and Sørensen, M. (1995). Martingale estimating functions for discretely observed diffusion processes. *Bernoulli* **1**, 17–39.
- Florens-Zmirou, D. (1989). Approximate discrete time schemes for statistics of diffusion processes. *Statistics* **20**, 547–557.
- Genon-Catalot, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations. *Statistics* **21**, 99–116.
- Genon-Catalot, V. and Jacod, J. (1993). On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Ann. Inst. Henri Poincaré Probab. Statist.* **29**, 119–151.
- Ibragimov, I. A. and Has'minskii, R. Z. (1981). *Statistical estimation*. Springer Verlag, New York.
- Karlin, S. and Taylor, H. M. (1981). *A second course in stochastic processes*. Academic Press, New York.
- Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24**, 211–229.
- Kessler, M. and Sørensen, M. (1999). Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli* **5**, 299–314.
- Kloeden, P. E. and Platen, E. (1991). *Numerical solution of stochastic differential equations*. Springer-Verlag, New York.
- Kunitomo, N. and Takahashi, A. (2001). The asymptotic expansion approach to the valuation of interest rate contingent claims. *Mathematical Finance* **11**, 117–151.
- Kutoyants, Yu. A. (1984). Parameter estimation for stochastic processes. Prakasa Rao, B.L.S. (ed.) Helder mann, Berlin.
- Kutoyants, Yu. A. (1994). *Identification of dynamical systems with small noise*. Kluwer, Dordrecht.
- Laredo, C. F. (1990). A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process. *Ann. Statist.* **18**, 1158–1171.
- Matsuzaki, R. and Uchida, M. (2005). One-step estimators for diffusion processes with small dispersion parameters from discrete observations. MHF Preprint Series, MHF 2005-10, Kyushu University. (submitted.)
- Sakamoto, Y. and Yoshida, N. (1999). Higher order asymptotic expansions for a functional of a mixing process and applications to diffusion functionals. (unpublished manuscript).
- Sakamoto, Y. and Yoshida, N. (2004). Asymptotic expansion formulas for functionals of  $\epsilon$ -Markov processes with a mixing property. *Ann. Inst. Statist. Math.* **56**, 545–597.
- Sørensen, M. (2000). Small dispersion asymptotics for diffusion martingale estimating functions. Preprint No. 2000-2, Department of Statistics and Operations Research, University of Copenhagen.
- Sørensen, M. and Uchida, M. (2003). Small diffusion asymptotics for discretely sampled stochastic differential equations. *Bernoulli* **9**, 1051–1069.

- Uchida, M. (2003). Estimation for dynamical systems with small noise from discrete observations. *J. Japan Statist. Soc.* **33**, 157–167.
- Uchida, M. (2004). Estimation for discretely observed small diffusions based on approximate martingale estimating functions. *Scand. J. Statist.* **31**, 553–566.
- Uchida, M. and Yoshida, N. (2004). Asymptotic expansion for small diffusions applied to option pricing. *Statist. Infer. Stochast. Process.* **7**, 189–223.
- Yoshida, N. (1990). Asymptotic behavior of M-estimator and related random field for diffusion process. *Ann. Inst. Statist. Math.* **42**, 221–251.
- Yoshida, N. (1992a). Asymptotic expansion of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe. *Probab. Theory Relat. Fields* **92**, 275–311.
- Yoshida, N. (1992b). Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.* **41**, 220–242.
- Yoshida, N. (1992c). Asymptotic expansion for statistics related to small diffusions. *J. Japan Statist. Soc.* **22**, 139–159.
- Yoshida, N. (2003). Conditional expansions and their applications. *Stochastic Process. Appl.* **107**, 53–81.

*Received March 26, 2004*

*Revised May 30, 2005*