J. Japan Statist. Soc.
Vol. 22 No. 2 1992 139-159

ASYMPTOTIC EXPANSION FOR STATISTICS RELATED TO SMALL DIFFUSIONS*

Nakahiro Yoshida**

Using Malliavin calculus we derived asymptotic expansions of mean values of statistics related to small diffusions. Applications to statistics and economics are discussed.

1. Introduction

Let $V_0: \boldsymbol{R}^d \to \boldsymbol{R}^d$ and $V: \boldsymbol{R}^d \to \boldsymbol{R}^d \otimes \boldsymbol{R}^r$ be smooth functions with bounded derivatives. For $\alpha = 1, 2, \cdots, r$, we denote by V_α the αth column vector of V. Consider a family of d-dimensional diffusion processes

$$(1.1) \qquad \begin{aligned} dX_t^\epsilon &= V_0(X_t^\epsilon)dt + \epsilon V(X_t^\epsilon)dw_t \,, \\ X_0 &= x_0 \,, \quad t \in [0, T] \,, \quad \epsilon \in (0, 1] \,, \end{aligned}$$

where w is an r-dimensional standard Wiener process. T and x_0 are constants and ϵ is a parameter. Many statistical problems associated with this small diffusion are related to functionals of the form

$$(1.2) \qquad F_T^\epsilon = F_0^\epsilon + \int_0^T f_0^\epsilon(X_t^\epsilon)dt + \epsilon \int_0^T f^\epsilon(X_t^\epsilon)dw_t \,,$$

where $F_0^\epsilon \in \boldsymbol{R}^k$, and $\boldsymbol{R}^k$-valued function $f_0^\epsilon(x)$ and $\boldsymbol{R}^k \otimes \boldsymbol{R}^r$-valued $f^\epsilon(x)$ are smooth.

In the problem of pricing path dependent options in economics, the price X_t of underlying security is supposed to satisfy the one-dimensional stochastic differential equation:

$$(1.3) \qquad \begin{aligned} dX_t^\epsilon &= cX_t^\epsilon dt + \epsilon X_t^\epsilon dw \,, \\ X_t^\epsilon &= x_0 \,, \end{aligned}$$

where c and x_0 are constants. To price average options at time $t=0$ we have to caclulate the expectation

$$(1.4) \qquad E[\mathrm{Max}\{Z_T^\epsilon - K, 0\}] \,,$$

where

$$Z_T^\epsilon = \frac{1}{T}\int_0^T X_t^\epsilon dt$$

and K is a striking price (see Kunitomo and Takahashi 1990). The time-average Z_T^ϵ takes the form of (1.2). It is difficult to express this expectation explicitly, so several methods involving FFT or the numerical analysis for partial differential equations have

Received December, 1991. Revised April, 1992. Accepted August, 1992.
* AMS 1980 subject classifications. 62E20, 62F12, 60H10.
Key words and phrases. Small diffusion, Malliavin calculus, asymptotic expansion.
** The Institute of Statistical Mathematics, 4-6-7, Minamiazabu Minato-ku, Tokyo 106, Japan.

been proposed. Kunitomo-Takahashi proposed that the log normal approximation to the distribution of Z_T^ϵ is applicable when ϵ is small. Here one problem is to derive the asymptotic expansion of (1.4).

A statistical problem occurs if the equation describing the observation process involves unknown parameters. Let X_t^ϵ satisfy the stochastic differential equation

$$\begin{aligned} dX_t^\epsilon &= V_0(X_t^\epsilon, \theta)dt+\epsilon V(X_t^\epsilon)dw_t\,, \\ X_0 &= x_0\,, \quad t\in[0,T]\,, \quad \epsilon\in(0,1]\,, \end{aligned} \tag{1.5}$$

where V_0 and V are smooth functions and θ is an unknown parameter. The consistency and the first order properties of the maximum likelihood estimator and Bayes estimator when $\epsilon\downarrow 0$ are known (Tsitovich 1977, Kutoyants 1978, 1984 and Genon-Catalot 1990). As for the higher order properties, the asymptotic expansions for distributions of the maximum likelihood estimator and the Bayes estimator were proved in Yoshida (1990a, b, 1991a). When discussing the second order optimal property of these estimators, we need to derive asymptotic expansions of statistics, such as log likelihood ratio statistics, Akahira-Takeuchi (1981). The log likelihood function and estimating equations of the M-estimators in the parameter estimation problem are examples of the functional (1.2). So it is necessary to show the asymptotic expansion for F_T^ϵ of (1.2).

As $\epsilon\downarrow 0$, the process X_t^ϵ defined by (1.1) converges to X_t^0 satisfying the differential equation

$$\frac{dX_t^0}{dt}=V_0(X_t^0)\,,\quad t\in[0,T]\,,\quad X_0^0=x_0\,.$$

For example it is known that

$$\sup_{0\le t\le T}|X_t^\epsilon-X_t^0|\to 0$$

a.s. as $\epsilon\downarrow 0$. Hence, under some regularity conditions, it follows that

$$F_T^\epsilon\to f_{-1}$$

a.s. as $\epsilon\downarrow 0$, where

$$f_{-1}=F_0^0+\int_0^T f_0^0(X_t^0)dt\,.$$

Therefore it is more convenient to treat

$$\tilde{F}^\epsilon=\frac{F_T^\epsilon-f_{-1}}{\epsilon}$$

instead of F_T^ϵ itself. In view of the above two examples, we consider the functional defined by

$$\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)\,, \tag{1.6}$$

where $\varphi^\epsilon(x)$, $\epsilon\in[0,1]$, are smooth functions on $\boldsymbol{R}^k$, and derive the asymptotic expansion of the mean value

$$E[\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)] \tag{1.7}$$

of this functional of F_T^ϵ. This covers the asymptotic expansions of the distribution and moments of $\tilde{F}^{\epsilon}$ over some regions of the sample space. Then it is a simple matter to obtain the asymptotic expansion of the mean value (1.4). Moreover, from these results, we can obtain the asymptotic expansions of the likelihood ratio statistic, which give the bounds of the probability of concentration for second order asymptotically median unbiased estimators. These bounds serve to show the second order efficiency of the maximum likelihood estimator and the Bayes estimator.

Let us discuss this from a mathematical point of view. Assume that a family of random variables F_ϵ has the asymptotic expansion:

$$F_\epsilon \sim f_0 + \epsilon f_1 + \cdots$$

as $\epsilon \downarrow 0$ in some sense. If function $T(x)$ satisfies a certain regularity, we then have the stochastic expansion

$$T(F_\epsilon) \sim \Phi_0 + \epsilon \Phi_1 + \cdots \tag{1.7}$$

as $\epsilon \downarrow 0$, where $\Phi_0, \Phi_1, \cdots$ are determined by formal Taylor expansion and in particular $\Phi_0 = T(f_0)$ and $\Phi_1 = f_1(\partial T/\partial x)(f_0)$. Expectation of (1.7) yields the asymptotic expansion

$$E[T(F_\epsilon)] \sim E[\Phi_0] + \epsilon E[\Phi_1] + \cdots \tag{1.8}$$

as $\epsilon \downarrow$. If we are able to take $T(x) = I_A(x)$, the indicator function of the Borel set A, (1.7) formally gives

$$I_A(F_\epsilon) \sim I_A(f_0) + \epsilon f_1 \frac{\partial I_A}{\partial x}(f_0) + \cdots \tag{1.9}$$

and hence,

$$P(F_\epsilon \in A) \sim P(f_0 \in A) + \epsilon E\left[f_1 \frac{\partial I_A}{\partial x}(f_0) \right] + \cdots \tag{1.10}$$

as $\epsilon \downarrow 0$. Thus we can obtain the asymptotic expansion of the distribution of F_ϵ quite directly.

Unfortunately two difficulties arise. The second term of the right-hand side of (1.9) is a composite function of the random variable f_0 and the Schwartz distribution $\partial I_A/\partial x$. There is no usual meaning of a random variable as a measurable function on a probability space. So we have the problem of how to define or interpret such composite functionals. After removing this difficulty we still meet the question of how to justify the formal expansion (1.9).

These difficulties have been solved by S. Watanabe (1983, 1987) in the Malliavin calculus. He introduced the notion of the pull-back of Wiener functionals under a Schwartz distribution to justify the composite functionals in certain Sobolev spaces as generalized Wiener functionals (i.e. the Schwartz distribution on the probability space). He also exploited the method of the asymptotic expansion of generalized Wiener functionals to apply the heat kernels of diffusion processes. This theory turned out to be successfully applicable to some problems of the higher order statistical inference, Yoshida (1990a, b, 1991a, b).

To use this theory, the crucial step is to show the nondegeneracy of the Malliavin covariance of functionals. However, it dose not seem easy to do this even for a simple statistical estimator, whose Malliavin covariance is given by an integration of some non-adaptive process, (Yoshida, 1990a). The Malliavin covariance corresponding to (1.2) is also written in a similar manner. Moreover, as for estimators appearing in parameter estimation, such as maximum likelihood estimators, we can not ensure their existence on the whole sample space in general. So we will need a modification of this theory with truncation. For the convenience of reference, in Section 4 we summarize several results used later. From these results we derive asymptotic expansions of mean values of functionals of F_T^ϵ.

The organization of this paper is as follows. In Section 2 we state our main result. Section 3 presents the illustrative examples of our main result. In Section 4 we sum up fundamental results about the Malliavin calculus with truncation. Finaly, in Section 5, we prove our result stated in Section 2.

Our method can be applied to statistical models with a discrete time parameter as well as models with a continuous one. For example, we have similar results for an AR(p) model with small noise (this is the case where the model fitting does well) if it is realized on a Wiener space, (Yoshida, 1991b).

2. Main results: Asymptotic expansions of the mean values of functionals of F_T^ϵ

Let an $\boldsymbol{R}^d \otimes \boldsymbol{R}^d$-valued process $Y_t^\epsilon(w)$ be the solution of the stochastic differential equation

$$dY_t^\epsilon = \partial V_0(X_t^\epsilon) Y_t^\epsilon dt + \epsilon \textstyle\sum_{\alpha=1}^r \partial V_\alpha(X_t^\epsilon) Y_t^\epsilon dw_t^\alpha , \quad t \in [0, T] ,$$
$$Y_0^\epsilon = I_d ,$$

where $[\partial V_\alpha]^{ij} = \partial_j V_\alpha^i$, $\partial_j = \partial/\partial x^j$, $i, j = 1, \cdots, d$, $\alpha = 0, 1, \cdots r$. Then $Y_t := Y_t^0$ is a non-singular deterministic $\boldsymbol{R}^d \otimes \boldsymbol{R}^d$-valued process. For function g^ϵ, $g^{(j)}$ denotes its j-th derivative in ϵ at $\epsilon = 0$. Let,

$$x^0 = F_0^{(1)} + \int_0^T f_0^{(1)}(X_t^0) dt$$

and

$$a_t = \int_t^T \partial f_0^0(X_s^0) Y_s ds Y_t^{-1} V(X_t^0) + f^0(X_t^0) .$$

We write $D_t = X_t^{(1)}$ and $E_t = X_t^{(2)}$. For $\epsilon \in [0, 1]$, let $\varphi^\epsilon(x)$ be smooth functions on $\boldsymbol{R}^k$. In this paper we assume the following conditions for F_T^ϵ given in (1.2):

(1) The function F_0^ϵ is smooth in ϵ. The functions $\varphi^\epsilon(x)$, $f_0^\epsilon(x)$ and $f^\epsilon(x)$ are smooth in (ϵ, x) on $[0, 1] \times \boldsymbol{R}^d$, and all derivatives are of polynomial growth order in x uniformly in ϵ.

(2) The matrix $\Sigma = \int_0^T a_t a_t' dt$ is positive definite.

We will use Einstein's rule for repeated indices. For matrix A, $[A]^{ij}$ denotes its (i, j)-element. Moreover, $[A]^{i\cdot}$ and $[A]^{\cdot j}$ are the i-th row vector and the j-th column vector of A, respectively. For vector a, a^i is its i-th element. Put $\sigma^{ij} = [\Sigma]^{ij}$ and $\sigma_{ij} = [\Sigma^{-1}]^{ij}$. As before, let

$$\tilde{F}^{\epsilon}=\frac{F_T^{\epsilon}-f_{-1}}{\epsilon}$$

with

$$f_{-1}=F_0^0+\int_0^T f_0^0(X_t^0)dt\ .$$

Define several functions as follows.

$$\lambda_{t,s}=Y_tY_s^{-1}V(X_s^0)\ ,$$

$$\lambda_{t,s}^i=[Y_tY_s^{-1}V(X_s^0)]^{i\cdot}\ ,\qquad i=1,\cdots,d\ ,$$

$$\mu_{i,t,s}=Y_tY_s^{-1}\varrho_iV(X_s^0)\ ,\qquad i=1,\cdots,d\ ,$$

$$\mu_{i,t,s}^j=[Y_tY_s^{-1}\partial_iV(X_s^0)]^{j\cdot}\ ,\qquad i,j=1,\cdots,d\ ,$$

$$\nu_{i,j,t,s}=Y_tY_s^{-1}\partial_i\partial_jV_0(X_s^0)\ ,\qquad i,j=1,\cdots,d\ ,$$

$$\nu_{i,j,t,s}^l=[Y_tY_s^{-1}\partial_i\partial_jV_0(X_s^0)]^l\ ,\qquad i,j,l=1,\cdots,d\ .$$

For $\boldsymbol{R}^1\otimes\boldsymbol{R}^r$-valued function h_t,

$$C_0(h)_T=\int_0^T a_th_t'dt\ .$$

For $\boldsymbol{R}$-valued function h_t,

$$C_1(h)_T=\int_0^T\int_0^t h_t\lambda_{t,s}a_s'dsdt\ .$$

For $\boldsymbol{R}^1\otimes\boldsymbol{R}^r$-valued function h_t,

$$C_2^i(h)_T=\int_0^T\int_0^t a_th_t'\lambda_{t,s}^ia_s'dsdt\ ,\qquad i=1,\cdots,d\ .$$

For $\boldsymbol{R}^1\otimes\boldsymbol{R}^r$-valued functions b_t and c_t, put

$$C_2(b,c)_T=\frac{1}{2}\int_0^T\int_0^T a_t[b_t'c_s+c_t'b_s]a_s'dsdt\ .$$

Let

$$C_2^{i,j}(t)=C_2(\lambda_{t,\cdot}^iI_{\{\cdot\le t\}},\lambda_{t,\cdot}^jI_{\{\cdot\le t\}})_T\ .$$

Define constants $A^{0,\alpha}$, $A_l^{1,\alpha}$ and $A_{p,q}^{2,\alpha}$ by

$$\begin{aligned}A^{0,\alpha}=&\frac{1}{2}F_0^{(2),\alpha}+\frac{1}{2}\int_0^T f_0^{(2),\alpha}(X_t^0)dt\\&+\frac{1}{2}\int_0^T\partial_\beta f_0^{0,\alpha}(X_t^0)\int_0^t\int_0^s\nu_{i,j,t,s}^\beta\lambda_{s,u}^i(\lambda_{s,u}^j)'dudsdt\\&+\frac{1}{2}\int_0^T\int_0^t\partial_i\partial_jf_0^{0,\alpha}(X_t^0)\lambda_{t,s}^i(\lambda_{t,s}^j)'dsdt\ ,\end{aligned}$$

$$A_l^{1,\alpha}=[C_1(\partial_if_0^{(1),\alpha}(X_\cdot^0))_T]^{ij}\sigma_{jl}+[C_0([f^{(1)}(X_\cdot^0)]^{\alpha\cdot})_T]^j\sigma_{jl}\ ,$$

$$A_{p,q}^{2,\alpha}=\frac{1}{2}\int_0^T\partial_i\partial_jf_0^{0,\alpha}(X_t^0)[C_2^{i,j}(t)]^{mn}dt\sigma_{pm}\sigma_{qn}$$

$$
\begin{aligned}
&+\frac{1}{2}\int_0^T \partial_\beta f_0^{0,\alpha}(X_t^0)\int_0^t \nu_{i,j,t,s}^\beta [C_2^{i,j}(s)]^{mn}\, dsdt\sigma_{pm}\sigma_{qn} \\
&+\int_0^T \partial_\beta f_0^{0,\alpha}(X_t^0)[C_2^i(\mu_{i,t,\cdot}^\beta)_t]^{mn} dt\sigma_{pm}\sigma_{qn} \\
&+[C_2^i([\partial_i f^0(X_\cdot^0)]^{\alpha\cdot})_T]^{mn}\sigma_{pm}\sigma_{qn} .
\end{aligned}
$$

Now, the following theorem is our main result ($\boldsymbol{B}^k$ denotes the Borel σ-field of $\boldsymbol{R}^k$).

THEOREM 2.1. *Under Assumptions* (1) *and* (2), *the following asymptotic expansion holds*:

$$
E[\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)]\sim\int_A p_0(x)dx+\epsilon\int_A p_1(x)dx+\cdots ,
$$

as $\epsilon\downarrow 0$. *This expansion is uniform in* $A\in\boldsymbol{B}^k$. *In particular,*

$$
\begin{aligned}
p_0(x) &= \varphi^0(x)\phi(x; x^0, \Sigma) , \\
p_1(x) &= \varphi^{(1)}(x)\phi(x; x^0, \Sigma) \\
&\quad +\varphi^0(x)\phi(x; x^0, \Sigma)[-A_\alpha^{1,\alpha} \\
&\quad +(A^{0,\alpha}\sigma_{\alpha p}-A_{q,l}^{2,\alpha}\sigma^{ql}\sigma_{\alpha p}-A_{p,\alpha}^{2,\alpha}-A_{\alpha,p}^{2,\alpha})[x-x^0]^p \\
&\quad +A_p^{1,\alpha}\sigma_{\alpha q}[x-x^0]^p[x-x^0]^q+A_{p,q}^{2,\alpha}\sigma_{\alpha l}[x-x^0]^p[x-x^0]^q[x-x^0]^l] ,
\end{aligned}
$$

where $\phi(x; x^0, \Sigma)$ *is the density of the normal distribution with mean* x^0 *and covariance matrix* Σ.

3. Examples

3.1 Applications to economics

In the problem of pricing path dependent options, the price X_t of underlying security is supposed to satisfy the one-dimensional stochastic differential equation

$$
\begin{aligned}
dX_t^\epsilon &= cX_t^\epsilon dt+\epsilon X_t^\epsilon dw_t , \\
X_0^\epsilon &= x_0 ,
\end{aligned}
\tag{3.2}
$$

where c and x_0 are constants. To price average options at time $t=0$ we have to calculate the expectation

$$
E[\mathrm{Max}\{Z_T^\epsilon-K, 0\}] ,
$$

where

$$
Z_T^\epsilon=\frac{1}{T}\int_0^T X_t^\epsilon dt
$$

and K is a striking price (see Kunitomo and Takahashi, 1990 and its references). It is difficult to express this expectation explicitly, so several methods involving FFT or the numerical analysis for partial differential equations have been proposed. Kunitomo-Takahashi proposed that the log normal approximation to the distribution of Z_T^ϵ is applicable when ϵ is small. We can derive asymptotic expansions for the diffusion defined by (1.1). Here we will only consider the asymptotic expansion of the distribution of Z_T^ϵ though we can treat transforms of Z_T^ϵ in the same manner.

Let X^ϵ satisfy the stochastic differential equation (1.1). For $F_T^\epsilon = Z_T^\epsilon$ and $\varphi^\epsilon(x)=1$, Theorem 2.1 gives the asymptotic expansion

$$P\left[\epsilon^{-1}\left(Z_T^\epsilon - \frac{1}{T}\int_0^T X_t^0 dt\right) \in A\right] \sim \int_A p_0^z(x)dx + \epsilon\int_A p_1^z(x)dx + \cdots ,$$

as $\epsilon \downarrow 0$ uniformly in $A \in \boldsymbol{B}^d$. In particular,

$$p_0^z(x) = \phi(x; 0, \Sigma) ,$$

$$p_1^z(x) = \phi(x; 0, \Sigma)[(A^{0,a}\sigma_{ai} - A_{j,l}^{2,a}\sigma^{jl}\sigma_{ai} - A_{i,a}^{2,a} - A_{a,i}^{2,a})x^i + A_{i,j}^{2,a}\sigma_{a,l}x^i x^j x^l] ,$$

where

$$A^{0,a} = \frac{1}{2T}\int_0^T\int_0^t\int_0^s \nu_{i,j,t,s}^a \lambda_{s,u}^i (\lambda_{s,u}^j)' du ds dt$$

$$A_{p,q}^{2,a} = \frac{1}{2T}\int_0^T\int_0^t \nu_{i,j,t,s}^a [C_2^{i,j}(s)]^{mn} ds dt \sigma_{pm}\sigma_{qn} + \frac{1}{T}\int_0^T [C_2^i(\mu_{i,t,\cdot}^a)_t]^{mn} dt \sigma_{pm}\sigma_{qn} .$$

Here $x^0 = 0$ and

$$a_t = \int_t^T \frac{1}{T} Y_s ds Y_t^{-1} V(X_t^0) .$$

For example, when X is the geometric Brownian motion (3.2),

$$p_1^z(x) = \phi(x; 0, \Sigma) A_{1,1}^{2,1}(-3x + \Sigma^{-1}x^3) ,$$

where

$$\Sigma = x_0^2 (cT)^{-2}\int_0^T (e^{cT} - e^{ct})^2 dt$$

and

$$A_{1,1}^{2,1} = \frac{1}{T}\int_0^T\int_0^t\int_0^s \Sigma^{-2} x_0^3 (cT)^{-2} e^{ct}(e^{cT} - e^{cs})(e^{cT} - e^{cu}) du ds dt . \tag{3.3}$$

We note that the asymptotic expansions given $\{X_s^\epsilon; 0 \le s \le t\}$, $t < T$, reduce to the unconditional case because of the Markov property.

We may derive the asymptotic expansion of $E[\mathrm{Max}\{Z_T^\epsilon - K, 0\}]$ directly in this context. Suppose that a one-dimensional diffusion X satisfies (1.1) for $d=1$. Let $F_T^\epsilon = Z_T^\epsilon$, $\varphi^\epsilon(x) = \epsilon x + f_{-1} - K$, $f_{-1} = \frac{1}{T}\int_0^T X_t^0 dt$ and $A_\epsilon' = \{Z_T^\epsilon \ge K\}$. Then, by Theorem 2.1, we see

$$\begin{aligned} &E[\mathrm{Max}\{Z_T^\epsilon - K, 0\}] \\ &= E[(Z_T^\epsilon - K) I_{A_\epsilon'}] \\ &\sim (f_{-1} - K)\int_{x \ge \frac{K - f_{-1}}{\epsilon}} \phi(x; 0, \Sigma) dx \end{aligned} \tag{3.4}$$

$$+\epsilon\int_{x\geq\frac{K-f_{-1}}{\epsilon}} p_1(x)dx$$
$$+\epsilon^2\int_{x\geq\frac{K-f_{-1}}{\epsilon}} p_2(x)dx+\cdots,$$

where

$$p_1(x)=x\phi(x;0,\Sigma)+(f_{-1}-K)\phi(x;0,\Sigma)[(A^{0,1}\Sigma^{-1}-3A_{1;1}^{2;1})x+A_{1;1}^{2;1}\Sigma^{-1}x^3]$$

and some smooth function $p_2(x)$. For (3.2),

$$p_1(x)=x\phi(x;0,\Sigma)+(f_{-1}-K)A_{1;1}^{2;1}(-3x+\Sigma^{-1}x^3)\phi(x;0,\Sigma)$$

with $A_{1;1}^{2;1}$ given in (3.3). If we want expansions in ϵ-power, we have from (3.4) that when $K-f_{-1}<0$:

$$E[\mathrm{Max}\{Z_T^\epsilon-K,0\}]$$
$$\sim(f_{-1}-K)+\epsilon^2\int p_2(x)dx+\cdots;$$

and when $K-f_{-1}>0$:

$$E[\mathrm{Max}\{Z_T^\epsilon-K,0\}]\sim O\epsilon(^n)$$

for $n=1,2,\cdots$; and when $K-f_{-1}=0$:

$$E[\mathrm{Max}\{Z_T^\epsilon-K,0\}]$$
$$\sim\epsilon\int_{x\geq 0} x\phi(x;0,\Sigma)dx+\epsilon^2\int_{x\geq 0} p_2'(x)dx+\cdots$$

for some function $p_2'(x)$. Finally, we note that $p_2(x)$ in (3.4) can be specified by

$$\int_A p_2(x)dx=E[f_1 I_A(f_0)]+E[f_0)f_1\partial I_A(f_0)]$$
$$+(f_{-1}-K)E\left[f_2\partial I_A(f_0)+\frac{1}{2}(f_1)^2\partial^2 I_A(f_0)\right]$$

for $A\in \boldsymbol{B}^1$, where $f_0, f_1, \cdots$ are the coefficients appearing in the expansion of $\tilde{F}^\epsilon$. It is possible to compute the right-hand side from conditional expectations of fourfold Wiener integrals given a Wiener integral. If $K-f_{-1}=0$ (this is an important case pointed out in Kunitomo-Takahashi), then Lemma 5.7 is sufficient to get $p_2'(x)$.

Using (3.4) we obtained similar numerical results for an example given in Kunitomo-Takahashi (1990).

3.2 Likelihood ratio statistic

Consider a parametric model of the d-dimensional small diffusions defined by

$$dX_t^\epsilon=V_0(X_t^\epsilon,\theta)dt+\epsilon V(X_t^\epsilon)dw_t,$$
$$X_0=x_0,\quad t\in[0,T],\quad \epsilon\in(0,1], \tag{1.5}$$

where θ is an unknown parameter in Θ, a bounded convex domain in $\boldsymbol{R}^k$. The likelihood function is given by the formula:

$$\Lambda_\epsilon(\theta;X)=\exp\left\{\int_0^T \epsilon^{-2}V_0'(VV')^+(X_t,\theta)dX_t-\frac{1}{2}\int_0^T \epsilon^{-2}V_0'(VV')^+V_0(X_t,\theta)dt\right\},$$

where A^+ denotes the Moore-Penrose generalized inverse matrix of the matrix A (Liptser and Shiryayev 1977). Let $\theta_0 \in \Theta$ denote the true value of the unknown parameter θ. For $h \in \boldsymbol{R}^k$, the log likelihood ratio is given by:

$$
\begin{aligned}
l_{\epsilon,h}(w;\theta_0) &= \log \Lambda_\epsilon(\theta_0+\epsilon h; X) - \log \Lambda_\epsilon(\theta_0; X) \\
&= \int_0^T \epsilon^{-1}[V_0(X_t^\epsilon, \theta+\epsilon h) - V_0(X_t^\epsilon, \theta_0)]'(VV')^+ V(X_t^\epsilon) dw_t \\
&\quad -\frac{1}{2}\int_0^T \epsilon^{-2}[V_0(X_t^\epsilon, \theta_0+\epsilon h) - V_0(X_t^\epsilon, \theta_0)]'(VV')^+(X_t^\epsilon)[V_0(X_t^\epsilon, \theta_0+\epsilon h) \\
&\quad - V_0(X_t^\epsilon, \theta_0)]dt\,,
\end{aligned}
$$

where X_t^ϵ is the solution of the above stochastic differential equation for $\theta=\theta_0$. Here we assume that $V_0(x,\theta)-V_0(x,\theta_0) \in M\{V(x)\}$: the linear manifold generated by column vectors of $V(x)$, for each x and θ. The asymptotic expansion of the distribution of the log likelihood ratio plays an important role in the higher order asymptotic theory to derive bounds of probability of concentration of statistics (see Akahira and Takeuchi 1981). We can obtain the asymptotic expansion from the result in Section 2.

We shall prepare several notations. Let $\delta_i = \partial/\partial\theta^i$ and denote $\delta=(\delta_1, \cdots, \delta_k)$. The Fisher information matrix $I(\theta_0)=(I_{ij}(\theta_0))$ is defined by:

$$
I_{ij}(\theta_0) = \int_0^T \delta_i V_0(X_t^0, \theta_0)'(VV')^+(X_t^0)\delta_j V_0(X_t^0, \theta_0) dt
$$

for $i, j=1, \cdots, k$. Denote $I(\theta_0)$ by $I=(I_{ij})$ and $I(\theta_0)^{-1}$ by $I^{-1}=(I^{ij})$. We use the multi-index: $|\boldsymbol{n}|=n_1+n_2+\cdots+n_d$ and $\partial^{\boldsymbol{n}}=\partial_1^{n_1}\partial_2^{n_2}\cdots\partial_d^{n_d}$ for $\boldsymbol{n}=(n_1, n_2, \cdots, n_d)$; $|\nu|=\nu_1+\nu_2+\cdots+\nu_k$ and $\delta^\nu=\delta_1^{\nu_1}\delta_2^{\nu_2}\cdots\delta_k^{\nu_k}$ for $\nu=(\nu_1, \nu_2, \cdots, \nu_k)$.

In this subsection we assume the following conditions.

(1) V_0, V and $(VV')^+$ are smooth in $(x,\theta) \in \boldsymbol{R}^d \times \Theta$.

(2) For $|\boldsymbol{n}| \geq 1$, $F=V_0, V, (VV')^+$, $\sup_{x,\theta}|\partial^{\boldsymbol{n}} F| < \infty$.

(3) For $|\nu| \geq 1$ and $|\boldsymbol{n}| \geq 0$, a constant $C_{\nu,\boldsymbol{n}}$ exists and

$$
\sup_\theta |\partial^{\boldsymbol{n}}\delta^\nu V_0| \leq C_{\nu,\boldsymbol{n}}(1+|x|)^{Q_{\nu,\boldsymbol{n}}}\,,
$$

for all x.

(4) $I(\theta)$, $\theta \in \Theta$, are positive definite.

Let

$$
A_{i,j,n} = \frac{1}{2}\int_0^T\int_0^t \partial_l\{\delta_i V_0'(VV')^+\delta_j V_0\}(X_t^0, \theta_0)[Y_t]^{lm}[g_s]^{mn} ds dt
$$

and

$$
B_{i,j,l} = \int_0^T \delta_i\delta_j V_0'(VV')^+\delta_l V_0(X_t^0, \theta_0) dt\,,
$$

where

$$
\begin{aligned}
g_s &= Y_s^{-1}VV'(VV')^+(X_s^0)\delta V_0(X_s^0, \theta_0) \\
&= Y_s^{-1}\delta V_0(X_s^0, \theta_0)\,.
\end{aligned}
$$

Then the following theorem, which was originally proved in Yoshida (1990b), is a corollary of Theorem 2.1.

THEOREM 3.1. *The probability distribution of the log likelihood ratio* $l_{\epsilon,h}(w;\theta_0)$, $h\neq 0$, *has the asymptotic expansion*

$$P[l_{\epsilon,h}(w;\theta_0)\in A]\sim\int_A p_0^L(x)dx+\epsilon\int_A p_1^L(x)dx+\cdots,\qquad as\ \epsilon\downarrow 0,\ A\in\boldsymbol{B}^1 .$$

The expansion is uniform in $A\in\boldsymbol{B}^1$. *In particular,*

$$p_0^L(x)=\phi(\bar{x};0,J),$$
$$p_1^L(x)=[A_{i,j,l}h^ih^jh^l]J^{-3}[\bar{x}^3-J\bar{x}^2-3J\bar{x}+J^2]\phi(\bar{x};0,J)$$
$$+\frac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-2}[\bar{x}^2-J\bar{x}-J]\phi(\bar{x};0,J),$$

where $J=h'I(\theta_0)h$ *and* $\bar{x}=x+(1/2)J$. *The probability distribution function of* $l_{\epsilon,h}(w;\theta_0)$ *has the asymptotic expansion*

$$P[l_{\epsilon,h}(w;\theta_0)\leq x]\sim\Phi(\bar{x};0,J)$$
$$-\epsilon\left\{[A_{i,j,l}h^ih^jh^l]J^{-2}[\bar{x}^2-J\bar{x}-J]\right.$$
$$\left.+\frac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-1}[\bar{x}-J]\right\}\phi(\bar{x};0,J)+\cdots,$$

where $\Phi(x;\mu,\sigma^2)$ *is the probability distribution function of the one-dimensional normal distribution with mean* μ *and variance* σ^2.

PROOF. We use Theorem 2.1 for $F_T^\epsilon=\epsilon l_{\epsilon,h}(w;\theta_0)$ and $\varphi^\epsilon(x)=1$. It is then easy to show that

$$x^0=-\frac{J}{2},$$
$$a_t=h^i\delta_iV_0'(X_t^0,\theta_0)(VV')^+V(X_t^0),$$
$$p_{f_0}(x)=\phi\left(x;-\frac{J}{2},J\right)$$
$$A^{0,1}=-\frac{1}{2}B_{i,j,l}h^ih^jh^l,$$
$$A_1^{1,1}=-A_{i,j,l}J^{-1}h^ih^jh^l+\frac{1}{2}B_{i,j,l}J^{-1}h^ih^jh^l,$$
$$A_{1,1}^{2,1}=A_{i,j,l}J^{-2}h^ih^jh^l .$$

Thus we have the result. □

In this case, non-singularity of the Fisher information corresponds to the non-degeneracy of the Malliavin covariance. We can also obtain the asymptotic expansion of the likelihood ratio statistic under the contiguous alternative $\theta_0+\epsilon h$. By the argument of Yoshida (1990b), or by the results in Section 4, the following theorem can be proved.

THEOREM 3.2. *The probability distribution of the log likelihood ratio* $l_{\epsilon,h}(w;\theta_0+\epsilon h)$ *has the asymptotic expansion*

$$P[l_{\epsilon,h}(w;\theta_0+\epsilon h)\in A]\sim\int_A p_0^{Lc}(x)dx+\epsilon\int_A p_1^{Lc}(x)dx+\cdots,\qquad as\ \epsilon\downarrow 0,\ A\in \boldsymbol{B}^1.$$

The expansion is uniform in $A\in\boldsymbol{B}^1$. *In particular,*

$$p_0^{Lc}(x)=\phi(\underline{x};0,J),$$
$$p_1^{Lc}(x)=[A_{i,j,l}h^ih^jh^l]J^{-3}[\underline{x}^3+2J\underline{x}^2-(3J-J^2)\underline{x}-2J^2]\phi(\underline{x};0,J)$$
$$+\frac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-2}[\underline{x}^2+J\underline{x}-J]\phi(\underline{x};0,J),$$

where $\underline{x}=x-(1/2)J$. *The probability distribution function of* $l_{\epsilon,h}(w;\theta_0+\epsilon h)$ *has the asymptotic expansion*

$$P[l_{\epsilon,h}(w;\theta_0+\epsilon h)\le x]\sim\Phi(\underline{x};0,J)$$
$$+\epsilon\Big\{[A_{i,j,l}h^ih^jh^l]J^{-2}[-\underline{x}^2-2J\underline{x}+J-J^2]$$
$$+\frac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-1}[-\underline{x}-J]\Big\}\phi(\underline{x};0,J)+\cdots,$$

Now let us discuss the second order efficiency of estimators. Here we adopt the criterion based on concentration of the distribution of an estimator in the neighbourhood of the true value due to Takeuchi and Akahira. See Akahira-Takeuchi (1981), also see Taniguchi (1983, 1987, 1991) for time series. When $k=1$, it follows from Theorems 3.1 and 3.2, that the second order distributions:

$$\Phi(J;0,J)+\epsilon\left[A_{1,1,1}h^3+\frac{1}{2}B_{1,1,1}h^3\right]\phi(J;0,J)$$

for $h>0$ and

$$\Phi(-J;0,J)-\epsilon\left[A_{1,1,1}h^3+\frac{1}{2}B_{1,1,1}h^3\right]\phi(J;0,J)$$

for $h<0$ are the bounds of second order distributions for any second order asymptotically median unbiased (AMU) estimator. A second order AMU estimator attaining these bounds for any $h>0$ and $h<0$ is said to be second order efficient.

Here we consider the maximum likelihood estimator and the Bayes estimators with respect to the quadratic loss functions and discuss their second order properties. When $\theta=\theta_0$ is true, the maximum likelihood estimator is denoted by $\hat{\theta}_\epsilon(w;\theta_0)$ and the Bayes estimator by $\tilde{\theta}_\epsilon(w;\theta_0)$. Let $\pi(\theta)$ be the Bayes prior which is a positive smooth function on Θ with all derivatives of polynomial growth order and $\inf_{\theta\in\Theta}\pi(\theta)>0$. It is known that maximum likelihood estimators and Bayes estimators have consistency and asymptotic normality, and they are efficient in the first order (see, Kutoyants, 1984). For function $f(x,\theta)$ of x and θ, $f_t^\epsilon(\theta)$ denotes $f(X_t^\epsilon,\theta)$. We assume that

(5) For any $\theta_0\in\Theta$, there exits a positive constant a_0 such that

$$\int_0^T [V_{0,t}^0(\theta)-V_{0,t}^0(\theta_0)]'(VV')_t^{+0}[V_{0,t}^0(\theta)-V_{0,t}^0(\theta_0)]dt\ge a_0|\theta-\theta_0|^2$$

for $\theta\in\Theta$.

In the context of the higher order statistical asymptotic theory we need to modify the estimators to obtain second-order efficient estimators. We call an estimator $\hat{\theta}_\epsilon^*$ a bias corrected maximum likelihood estimator if:

$$\hat{\theta}_\epsilon^* = \hat{\theta}_\epsilon - \epsilon^2 b(\hat{\theta}_\epsilon) ,$$

where $b(\theta)$ is a bounded smooth function with bounded derivatives. Similarly, we call an estimator $\tilde{\theta}_\epsilon^*$ a bias corrected Bayes estimator if:

$$\tilde{\theta}_\epsilon^* = \tilde{\theta}_\epsilon - \epsilon^2 \tilde{b}(\tilde{\theta}_\epsilon) ,$$

where $\tilde{b}(\theta)$ is a bounded smooth function with bounded derivatives. We have the asymptotic expansions of the distributions of the maximum likelihood estimator and the Bayes estimator (Yoshida 1990b, 1991a). From them it follows that the maximum likelihood estimator and the Bayes estimator are second order AMU if their biases are corrected by

$$b(\theta_0) = -I(\theta_0)^{-2} A_{1,1,1}$$

and

$$\tilde{b}(\theta_0) = -I(\theta_0)^{-2} A_{1,1,1} - \frac{3}{2} I(\theta_0)^{-2} B_{1,1,1} + I(\theta_0)^{-1} \pi(\theta_0)^{-1} \delta\pi(\theta_0) ,$$

respectively. Then we can prove that they are second order efficient, comparing their second order distributions with the bounds of second order distributions.

4. Preparations: Malliavin calculus with truncation

We begin with preparing notations used in the Malliavin calculus. For details see Malliavin (1976, 1978), Watanabe (1983, 1984, 1987), Ikeda and Watanabe (1984, 1989), and Kusuoka and Stroock (1984). Kusuoka and Stroock (1991) also present a method for deriving asymptotic expansions.

Let (W, P) be the r-dimensional Wiener space and let H be the Cameron-Martin subspace of W endowed with the norm

$$|h|_H^2 = \int_0^T |\dot{h}_t|^2 dt$$

for $h \in H$. For a Hilbert space E, $||\cdot||_p$ denotes the $L^p(E)$-norm of E-valued Wiener functional. Define $||f||_{p,s}$ for E-valued Wiener functional f, $s \in \boldsymbol{R}$, $p \in (1, \infty)$, by:

$$||f||_{p,s} = ||(I-L)^{s/2} f||_p ,$$

where L is the Ornstein-Uhlenbeck operator. An E-valued function f: $W \to E$ is called an E-valued polynomial if

$$f = \sum_{i=1}^n p_i([h_1](w), \cdots, [h_m](w)) e_i ,$$

where $n \in \boldsymbol{Z}^+$, $h_i \in H$, $e_i \in E$, $p_i(x^1, \cdots, x^m)$ are polynomials and $[h](w) = \sum_{i=1}^r \int_0^T \dot{h}_t^i dw^i(t)$ for $h \in H$. Let P(E) denote the totality of E-valued polynomials on the Wiener space (W, P). The Banach space $D_p^s(E)$ is the completion of P(E) with respect to $||\cdot||_{p,s}$. Dual of $D_p^s(E)$ is $D_q^{-s}(E)$, where $s \in \boldsymbol{R}$, $p>1$ and $1/p+1/q=1$. The space $D^\infty(E)=$

$\cap_{s>0}\cap_{1<p<\infty}D_p^s(E)$ is the set of Wiener test functionals and $\tilde{D}^{-\infty}(E)=\cup_{s>0}\cap_{1<p<\infty}D_p^{-s}(E)$ is a space of generalized Wiener functionals (see Watanabe, 1984). We suppress $\boldsymbol{R}$ when $E=\boldsymbol{R}$. For $F\in \mathrm{P}(E)$ and $h\in H$, the derivative of F in the direction of h is defined by

$$\langle D_hF(w), e\rangle=\left(\frac{d}{d\epsilon}\right)_0\langle F(w+\epsilon h), e\rangle$$

for $e\in E$, and $DF\in \mathrm{P}(H\otimes E)$ is called the H-derivative of F. It is known that for $k=0, 1, \cdots$ and $p>1$, the norm $||\cdot||_{p,k}$ is equivalent to the norm $\sum_{l=0}^{k}||D^l\cdot||_p$.

The Fréchet space $\mathrm{S}(\boldsymbol{R}^d)$ is the totality of rapidly decreasing smooth functions on $\boldsymbol{R}^d$, and $\mathrm{S}'(\boldsymbol{R}^d)$ is its dual. Let $A=1+|x|^2-(1/2)\Delta$, and then A^{-1} is an integral operator.

In order to apply the ideas of Malliavin and Watanabe to statistical problems, we need a version of their theory with truncation. Let $F\in D^\infty(\boldsymbol{R}^d)$, $G\in D^\infty$ and $\xi\in D^\infty$. Let ψ: $\boldsymbol{R}\to\boldsymbol{R}$ be a smooth function such that $0\leq\psi(x)\leq1$ for $x\in\boldsymbol{R}$, $\psi(x)=1$ for $|x|\leq1/2$ and $\psi(x)=0$ for $|x|\geq1$. Suppose, for any $p\in(1,\infty)$, the Malliavin convariance σ_F of F satisfies

$$E[1_{\{|\xi|\leq1\}}(\det\sigma_F)^{-p}]<\infty\ . \tag{4.1}$$

Then the composite functional $\psi(\xi)GT\circ F\in\tilde{D}^{-\infty}$ is well-defined for any $T\in S'(\boldsymbol{R}^d)$. For ψ, ξ, F, G given as above and any measurable function $f(x)$ of polynomial growth order

$$\psi(\xi)Gf\circ F=\psi(\xi)Gf(F)$$

in $\tilde{D}^{-\infty}$.

Let us consider a family of E-valued Wiener functionals (or generalized Wiener functionals) $\{F_\epsilon(w)\}$, $\epsilon\in(0,1]$. For $k>0$ if

$$\limsup_{\epsilon\downarrow0}\frac{||F_\epsilon||_{p,s}}{\epsilon^k}<\infty\ ,$$

we say $F_\epsilon(w)=O(\epsilon^k)$ in $D_p^s(E)$ as $\epsilon\downarrow0$. Following Watanabe (1987), we say that $F_\epsilon(w)\in D^\infty(E)$ has the asymptotic expansion:

$$F_\epsilon(w)\sim f_0+\epsilon f_1+\cdots$$

in $D^\infty(E)$ as $\epsilon\downarrow0$ with $f_0, f_1, \cdots\in D^\infty(E)$, if for every $p>1$, $s>0$ and $k=1, 2, \cdots$

$$F_\epsilon(w)-(f_0+\epsilon f_1+\cdots+\epsilon^{k-1}f_{k-1})=O(\epsilon^k)$$

in $D_p^s(E)$ as $\epsilon\downarrow0$. Similarly, we say that $F_\epsilon(w)\in\tilde{D}^{-\infty}(E)$ has the asymptotic expansion

$$F_\epsilon(w)\sim f_0+\epsilon f_1+\cdots$$

in $\tilde{D}^{-\infty}(E)$ as $\epsilon\downarrow0$ with $f_0, f_1, \cdots\in\tilde{D}^{-\infty}(E)$, if for every $k=1, 2, \cdots$ there exists $s>0$ such that, for every $p>1$, $F_\epsilon(w), f_0, f_1, \cdots\in D_p^{-s}(E)$ and

$$F_\epsilon(w)-(f_0+\epsilon f_1+\cdots+\epsilon^{k-1}f_{k-1})=O(\epsilon^k)$$

in $D_p^{-s}(E)$ as $\epsilon\downarrow0$. The generalized means of these expansions yield the ordinary asymptotic expansions.

The following theorem is a version of Theorem 2.3 of Watanabe (1987).

THEOREM 4.1. *Let Λ be an index set. Suppose that families $\{F_\epsilon(w); \epsilon \in (0, 1]\} \subset D^\infty(\boldsymbol{R}^d)$, $\{\xi_\epsilon(w); \epsilon \in (0, 1]\} \subset D^\infty$ and $\{T_\lambda; \lambda \in \Lambda\} \subset S'(\boldsymbol{R}^d)$ satisfy the following conditions.*

1) *For any $p \in (1, \infty)$*

$$\sup_{\epsilon \in (0,1]} E[_{[|\xi_\epsilon| \le 1]}(\det \sigma_{F_\epsilon})^{-p}] < \infty \, .$$

2) *$\{F_\epsilon(w); \epsilon \in (0, 1]\}$ has the asymptotic expansion*

$$F_\epsilon(w) \sim f_0 + \epsilon f_1 + \cdots \ in \ D^\infty(\boldsymbol{R}^d) \ as \ \epsilon \downarrow 0$$

with $f_i \in D^\infty(\boldsymbol{R}^d)$.

3) $\xi_\epsilon(w) = 0(1)$ in D^∞ *as* $\epsilon \downarrow 0$.

4) *For any $n = 1, 2, \cdots,$*

$$\lim_{\epsilon \downarrow 0} \epsilon^{-n} P\left\{|\xi_\epsilon| > \frac{1}{2}\right\} = 0 \, .$$

5) *For any $n = 0, 1, 2, \cdots,$ there exists a nonnegative integer m such that $A^{-m}T_\lambda \in C_B^n(\boldsymbol{R}^d)$ for all $\lambda \in \Lambda$ and*

$$\sup_{\lambda \in \Lambda} \sum_{|\boldsymbol{n}| \le n} ||\partial^{\boldsymbol{n}} A^{-m} T_\lambda||_\infty < \infty \, .$$

Let $\{G_{\mu,\epsilon}(w); \mu \in M, \epsilon \in (0, 1]\}$ be a family of Wiener functionals, where $\mu \in M$ is an index set, and suppose that $G_{\mu,\epsilon}(w)$ has the asymptotic expansion

$$G_{\mu,\epsilon}(w) \sim g_{\mu,0} + \epsilon g_{\mu,1} + \cdots$$

in D^∞ as $\epsilon \downarrow 0$ uniformly in $\mu \in M$ with $g_{\mu,0}, g_{\mu,1} \cdots \in D^\infty$. Then the composite functional $\psi(\xi_\epsilon) G_{\mu,\epsilon} T_\lambda \circ F_\epsilon \in \tilde{D}^{-\infty}$ is well-defined and has the asymptotic expansion

$$\psi(\xi_\epsilon) G_{\mu,\epsilon} T_\lambda \circ F_\epsilon \sim \Phi_{\lambda,\mu,0} + \epsilon \Phi_{\lambda,\mu,1} + \cdots \ in \ \tilde{D}^{-\infty} \ as \ \epsilon \downarrow 0$$

uniformly in $(\lambda, \mu) \in \Lambda \times M$ with $\Phi_{\lambda,\mu,0}, \Phi_{\lambda,\mu,1}, \cdots \in \tilde{D}^{-\infty}$. Coefficients $\Phi_{\lambda,\mu,0}, \Phi_{\lambda,\mu,1}, \cdots$ are *determined by the formal Taylor expansion*

$$\begin{aligned} G_{\mu,\epsilon} T_\lambda(f_0 + [\epsilon f_1 + \epsilon^2 f_2 + \cdots]) &= (g_{\mu,0} + \epsilon g_{\mu,1} + \cdots) \sum_{\boldsymbol{n}} \frac{1}{\boldsymbol{n}!} \partial^{\boldsymbol{n}} T_\lambda(f_0)[\epsilon f_1 + \epsilon^2 f_2 + \cdots]^{\boldsymbol{n}} \\ &= \Phi_{\lambda,\mu,0} + \epsilon \Phi_{\lambda,\mu,1} + \cdots \, , \end{aligned}$$

where $\boldsymbol{n} = (n_1, \cdots, n_d)$ is a multi-index, $\boldsymbol{n}! = n_1! \cdots n_d!$, $a^{\boldsymbol{n}} = a_1^{n_1} \cdots a_d^{n_d}$ for $a \in \boldsymbol{R}^d$. In particular,

$$\begin{aligned} \Phi_{\lambda,\mu,0} &= g_{\mu,0} T_\lambda(f_0) \, , \\ \Phi_{\lambda,\mu,1} &= g_{\mu,0} \textstyle\sum_{i=1}^d f_1^i \partial_i T_\lambda(f_0) + g_{\mu,1} T_\lambda(f_0) \, , \\ \Phi_{\lambda,\mu,2} &= g_{\mu,0} \left\{ \textstyle\sum_{i=1}^d f_2^i \partial_i T_\lambda(f_0) + \dfrac{1}{2} \textstyle\sum_{i,j=1}^d f_1^i f_1^j \partial_i \partial_j T_\lambda(f_0) \right\} \\ &\quad + g_{\mu,1} \textstyle\sum_{i=1}^d f_1^i \partial_i T_\lambda(f_0) + g_{\mu,2} T_\lambda(f_0), \cdots \, . \end{aligned}$$

For $n = 0, 1, 2, \cdots$, there exists a positive integer m such that

$$\sup_{B \in \boldsymbol{B}^d} \sum_{|\boldsymbol{n}| \le n} ||\partial^{\boldsymbol{n}} A^{-m} I_B||_\infty < \infty \, .$$

5. Proof of the main result

We have the stochastic expansion of the functional F_T^ϵ.

LEMMA 5.1. *$F_T^\epsilon \in D^\infty(\boldsymbol{R}^k)$ and it has the asymptotic expansion*

$$F_T^\epsilon \sim f_{-1}+\epsilon f_0+\epsilon^2 f_1+\cdots \text{ in } D^\infty(\boldsymbol{R}^k) \text{ as } \epsilon\downarrow 0$$

with $f_{-1}, f_0, f_1, \cdots \in D^\infty(\boldsymbol{R}^k)$. In particular,

$$f_{-1}=F_0^0+\int_0^T f_0^0(X_t^0)dt\,,$$

$$\begin{aligned} f_0=F_0^{(1)}+\int_0^T f_0^{(1)}(X_t^0)dt+\int_0^T \partial_i f_0^0(X_t^0)D_t^i dt\\ +\int_0^T f^0(X_t^0)dw_t\,, \end{aligned}$$

$$\begin{aligned} f_1=\frac{1}{2}F_0^{(2)}+\frac{1}{2}\int_0^T f_0^{(2)}(X_t^0)dt\\ +\int_0^T \partial_i f_0^{(1)}(X_t^0)D_t^i dt+\int_0^T f^{(1)}(X_t^0)dw_t\\ +\frac{1}{2}\int_0^T [\partial_i\partial_j f_0^0(X_t^0)D_t^i D_t^j+\partial_i f_0^0(X_t^0)E_t^i]dt\\ +\int_0^T \partial_i f^0(X_t^0)D_t^i dw_t\,. \end{aligned}$$

The proof is done by the Taylor formula. We will only calculate the first order H-derivative of F_T^ϵ. For $h\in H$, the H-derivative of $\tilde{F}^\epsilon$ in the direction of h is given by

$$\begin{aligned} D_h\tilde{F}^\epsilon &=\epsilon^{-1}\int_0^T \partial_i f_0^\epsilon(X_t^\epsilon)D_h X_t^{\epsilon,i}dt+\int_0^T \partial_i f^\epsilon(X_t^\epsilon)D_h X_t^{\epsilon,i}dw_t\\ &\quad+\int_0^T f^\epsilon(X_t^\epsilon)\dot{h}_t dt\\ &=\int_0^T \partial_i f_0^\epsilon(X_t^\epsilon)\int_0^t [Y_t^\epsilon Y_s^{\epsilon-1}V(X_s^\epsilon)\dot{h}_s]^i ds dt\\ &\quad+\epsilon\int_0^T \partial_i f^\epsilon(X_t^\epsilon)\int_0^t [Y_t^\epsilon Y_s^{\epsilon-1}V(X_s^\epsilon)\dot{h}_s]^i ds dw_t\\ &\quad+\int_0^T f^\epsilon(X_t^\epsilon)\dot{h}_t dt\\ &=\int_0^T C_s^\epsilon \dot{h}_s ds\,, \end{aligned}$$

where $C_s^\epsilon=(C_{s,ij}^\epsilon)$, $i=1,\cdots,k$, $j=1,\cdots,r$, are defined by

$$\begin{aligned} C_{s,ij}^\epsilon=\int_s^T \partial_l f_0^{\epsilon,i}(X_t^\epsilon)[Y_t^\epsilon Y_s^{\epsilon-1}V(X_s^\epsilon)]^{lj}dt\\ +\epsilon\int_s^T [\partial_l f^\epsilon(X_t^\epsilon)]^{im}[Y_t^\epsilon Y_s^{\epsilon-1}V(X_s^\epsilon)]^{lj}dw_t^m\\ +[f^\epsilon(X_s^\epsilon)]^{ij}\,. \end{aligned}$$

From this formula we can estimate DF_T^ϵ by using

$$|D\tilde{F}^{\epsilon}|^2_{H\otimes R^k}=\int_0^T |C_s^{\epsilon}|^2 ds .$$

Similarly, we can show the boundedness of higher order H-derivatives with L^p estimates of solutions of stochastic differential equations. Details are omitted. Thus we obtain the above lemma.

The Malliavin covariance $\sigma_{\tilde{F}^{\epsilon}}$ of $\tilde{F}^{\epsilon}$ is given by

$$\sigma_{\tilde{F}^{\epsilon}}=\int_0^T C_s^{\epsilon} C_s^{\epsilon\prime} ds .$$

Let

$$\xi_c^{\epsilon}=c\sum_{l,j}\int_0^T \left|\epsilon\int_s^T [\partial_l f^{\epsilon}(X_t^{\epsilon})]^{im}[Y_t^{\epsilon}Y_s^{\epsilon-1}V(X_s^{\epsilon})]^{lj}dw_t^m\right|^2 ds \tag{5.1}$$
$$+c\int_s^T |a_t^{\epsilon}-a_t|^2 dt$$

for $c>0$, where

$$a_t^{\epsilon}=\int_t^T \partial f_0^{\epsilon}(X_u^{\epsilon})Y_u^{\epsilon}du Y_t^{\epsilon-1}V(X_t^{\epsilon})+f^{\epsilon}(X_t^{\epsilon}) .$$

It is then easy to show the following lemma by Assumption (2).

LEMMA 5.2. *The Malliavin covariance $\sigma_{\tilde{F}^{\epsilon}}$ of $\tilde{F}^{\epsilon}$ is uniformly non-degenerate, i.e., $c_0>0$ exists such that for $c>c_0$ and any $p>1$*

$$\sup_{\epsilon} E[1_{(|\xi_c^{\epsilon}|\le 1)}(\det \sigma_{\tilde{F}^{\epsilon}})^{-p}]<\infty .$$

We will use the following large deviation inequalities.

LEMMA 5.3. (1) *There exist positive constants a_i, $i=1, 2$, independent of ϵ such that*

$$P[\sup_{0\le t\le T}|X_t^{\epsilon}-X_t^0|>a_0]\le\frac{a_1}{a_0}(a_0+1)\exp\left\{-\frac{a_2a_0^2}{a_0^2+1}\epsilon^{-2}\right\}$$

for all $a_0>0$.

(2) *There exist positive constants a_i, $i=1, 2, 3$, independent of ϵ such that*

$$P[\sup_{0\le t\le T}|Y_t^{\epsilon}-Y_t|>a_0]\le\frac{a_1}{a_0}(a_0+1)\exp\left\{-\frac{a_2a_0^2}{a_0^2+1}\epsilon^{-2}\right\}$$

for $a_0>0$.

(3) *Let g_t^{ϵ} be a predictable process such that*

$$E\left[\int_0^T |g_t^{\epsilon}|^2 dt\right]<\infty$$

for each ϵ and

$$P[\sup_{0\le t\le T}|g_t^{\epsilon}-g_t^0|>\epsilon^{c_0}]\le c_1\exp\{-c_2\epsilon^{-c_3}\}$$

for some positive c_i, $i=0, \cdots, 3$, and a function g_t^0. Then, for $a_0>0$, there exist positive constants a_i, $i=1, 2, 3$, such that

$$P\left[\sup_{0\le t\le T}\left|\int_0^t g_s^\epsilon dB_s-\int_0^t g_s^0 dB_s\right|>a_0\right]\le a_1\exp\{-a_2\epsilon^{-a_3}\}$$

for any ϵ, *where* B *is a Brownian motion process.*

PROOF. For simplicity, we assume that $d=r=1$. From the Lipschitz continuity of V_0 we have

$$|X_t^\epsilon-X_t^0|\le\int_0^t C|X_s^\epsilon-X_s^0|ds+\sup_{0\le s\le t}\left|\int_0^s \epsilon V(X_u^\epsilon)dw_u\right| .$$

Let $\tau=\inf\{t;\, |X_t^\epsilon-X_t^0|>a_0\}$. For each ϵ, there exists a Browman motion process b such that $\int_0^s \epsilon V(X_u^\epsilon)dw_u=b(A_s)$, $s\ge 0$, where $A_s=\int_0^s \epsilon^2 V(X_u^\epsilon)^2du$. Then, by Gronwall's lemma,

$$\begin{aligned}
&P[\sup_{0\le t\le T}|X_t^\epsilon-X_t^0|>a_0]\\
&=P[\tau<T]\\
&=P[\tau<T,\ a_0\le e^{CT}\sup_{s\le\tau}|b(A_s)|]\\
&\le P[a_0\le e^{CT}\sup_{u\le\epsilon^2 TC_1(a_0+1)^2}|b(u)|]\\
&\le\frac{a_1}{a_0}(a_0+1)\exp\left\{-\frac{a_2a_0^2}{a_0^2+1}\epsilon^{-2}\right\}
\end{aligned}$$

for some C_1, a_1, $a_2>0$ and any $a_0>0$.

Next, by Gronwall's lemma, we have

$$\sup_{0\le t\le u}|Y_t^\epsilon-Y_t|\le C_1\left[\sup_{0\le t\le u}|X_t^\epsilon-X_t^0|+\sup_{0\le t\le u}\left|\int_0^t \epsilon\partial V(X_s^\epsilon)Y_s^\epsilon dw\right|\right]$$

for some $C_1>0$. Let $\delta=a_0/2C_1$ and let

$$\sigma=\inf\{t;\, |Y_t^\epsilon-Y_t|>a_0 \text{ or } |X_t^\epsilon-X_t^0|>\delta\} .$$

Then there exists a Brownian motion process $\bar{b}_t$ and

$$\begin{aligned}
&P[\sup_{0\le t\le T}|Y_t^\epsilon-Y_t|>a_0]\\
&\le P[\sigma<T,\ \sup_{0\le t\le T}|X_t^\epsilon-X_t^0|\le\delta]+P[\sup_{0\le t\le T}|X_t^\epsilon-X_t^0|>\delta]\\
&\le P[\sigma<T,\ a_0\le C_1[\delta+\sup_{0\le u\le C_2(a_0^2+1)\epsilon^2}|\bar{b}_u|]+P[\sup_{0\le t\le T}|X_t^\epsilon-X_t^0|>\delta]\\
&\le\frac{a_1}{a_0}(a_0+1)\exp\left\{-\frac{a_2a_0^2}{a_0^2+1}\epsilon^{-2}\right\}
\end{aligned}$$

for some a_1, $a_2>0$ and any $a_0>0$.

By a similar argument we obtain (3). □

The effect of truncation $\psi(\xi_c^\epsilon)$ is negligible.

LEMMA 5.4. *For* $c>0$, ξ_c^ϵ *has an asymptotic expansion in* D^∞ *and for* $c_0>0$, *there exist positive constants* c_i, $i=1, 2, 3$, *such that*

$$P[|\xi_c^\epsilon|>c_0]\le c_1\exp\{-c_2\epsilon^{-c_3}\} .$$

PROOF. For the second term on the right-hand side of (5.1), we use (1) and (2) of Lemma 5.3. For the first term, we use (1), (2) and (3) of Lemma 5.3 for $c_0=1/2$. This yields the result. □

From Lemmas 5.2, 5.1, 5.4 and Theorem 4.1, we see that the composite functional

$\psi(\xi_c^\epsilon)\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)$ is well-defined for $A \in \boldsymbol{B}^k$ and obtain lemma 5.5.

LEMMA 5.5. *For large $c>0$, $\psi(\xi_c^\epsilon)\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)$ has the asymptotic expansion*

$$\psi(\xi_c^\epsilon)\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)\sim\Phi_0+\epsilon\Phi_1+\cdots$$

in $\tilde{D}^{-\infty}$ as $\epsilon\downarrow 0$ uniformly in $A \in \boldsymbol{B}^k$. In particular,

$$\Phi_0=\varphi^0(f_0)I_A(f_0)\ ,$$

$$\Phi_1=(f_1^i\partial_i\varphi^0(f_0)+\varphi^{(1)}(f_0))I_A(f_0)+\varphi^0(f_0)f_1^i\partial_iI_A(f_0)$$

$$\begin{aligned}\Phi_2=&\Big(f_2^i\partial_i\varphi^0(f_0)+\frac{1}{2}f_1^if_1^j\partial_i\partial_j\varphi^0(f_0)+f_1^i\partial_i\varphi^{(1)}(f_0)+\frac{1}{2}\varphi^{(2)}(f_0)\Big)I_A(f_0)\\&+(f_1^i\partial_i\varphi^0(f_0)+\varphi^{(1)}(f_0))f_1^j\partial_jI_A(f_0)\\&+\varphi^0(f_0)(f_2^i\partial_iI_A(f_0)+\frac{1}{2}f_1^if_1^j\partial_i\partial_jI_A(f_0))\ .\end{aligned}$$

Let $e^i(x)=E[f_1^i|f_0=x]$, $i=1,\cdots,k$.

LEMMA 5.6. *Let $p_{f_0}(x)$ be the density of f_0. For $A \in \boldsymbol{B}^k$,*

(1)
$$E[\varphi^0(f_0)f_1^i\partial_iI_A(f_0)]=\int_A p_1'(x)dx\ ,$$

where

$$p_1'(x)=-\partial_i\{\varphi^0(x)e^i(x)p_{f_0}(x)\}\ .$$

(2)
$$E[(f_1^i\partial_i\varphi^0(f_0)+\varphi^{(1)}(f_0))I_A(f_0)]=\int_A p_1''(x)dx\ ,$$

where

$$p_1''(x)=(e^i(x)\partial_i\varphi^0(x)+\varphi^{(1)}(x))p_{f_0}(x)\ .$$

(3)
$$\begin{aligned}p_1(x):&=p_1'(x)+p_1''(x)\\&=\varphi^{(1)}(x)p_{f_0}(x)-\varphi^0(x)\partial_ie^i(x)\cdot p_{f_0}(x)\\&\quad-\varphi^0(x)e^i(x)\partial_ip_{f_0}(x)\ .\end{aligned}$$

PROOF. (1) Using the integration-by-parts firmula, a smooth functional $G(w)$ exists and

$$\begin{aligned}&E[\varphi^0(f_0)f_1^i\partial_iI_A(f_0)]\\&\quad=E[G(w)I_A(f_0)]\\&\quad=\int_A p_1'(x)dx\ ,\end{aligned}$$

where

$$p_1'(x)=E[G(w)|f_0=x]p_{f_0}(x)$$

is a smooth function. To obtain $p_1'(x)$, let $A=(x_1,\infty)\times\cdots\times(x_k,\infty)$, then

$$\begin{aligned}p_1'(x)&=(-1)^k\partial_1\cdots\partial_kE[\varphi^0(f_0)f_1^i\partial_iI_A(f_0)]\\&=-\partial_i\{E[\varphi^0(f_0)f_1^i\delta_x(f_0)]\}\\&=-\partial_i\{\varphi^0(x)e^i(x)p_{f_0}(x)\}\ .\end{aligned}$$

To show (2) and (3) is easy. □

For second order statistical inference, it is necessary to calculate $e^i(x)$, $i=1, \cdots, k$, explicitly. For this purpose we prepare two lemmas.

LEMMA 5.7. *Let w be an r-dimensional Wiener process and let functions a_t, b_t, c_t on $[0, T]$ be deterministic. Let $\Sigma=\int_0^T a_t a_t' dt$.*

(1) *Let a_t be $\boldsymbol{R}^k \otimes \boldsymbol{R}^r$-valued and let b_t be $\boldsymbol{R}^r$-valued. Then*

$$E\left[\int_0^T b_t' dw_t \,\middle|\, \int_0^T a_t dw_t = x\right] = x' \Sigma^{-1} \int_0^T a_t b_t dt .$$

(2) *Let a_t, b_t and c_t be $\boldsymbol{R}^k \otimes \boldsymbol{R}^r$, $\boldsymbol{R}^m \otimes \boldsymbol{R}^r$ and $\boldsymbol{R}^m \otimes \boldsymbol{R}^r$-valued, respectively. Then*

$$E\left[\int_0^T \left(\int_0^t b_s dw_s\right)' c_t dw_t \,\middle|\, \int_0^T a_t dw_t = x\right]$$
$$= Trace \int_0^T \int_0^t \Sigma^{-1} a_t c_t' b_s a_s' \Sigma^{-1} (xx' - \Sigma) ds dt .$$

(3) *Let a_t be $\boldsymbol{R}^k \otimes \boldsymbol{R}^r$-valued and let b_t and c_t be $\boldsymbol{R}^1 \otimes \boldsymbol{R}^r$-valued. Then*

$$E\left[\int_0^T b_t dw_t \int_0^T c_t dw_t \,\middle|\, \int_0^T a_t dw_t = x\right]$$
$$= \frac{1}{2} Trace \int_0^T \int_0^T \Sigma^{-1} a_t [c_t' b_s + b_t' c_s] a_s' \Sigma^{-1} (xx' - \Sigma) ds dt + \int_0^T b_t c_t' dt .$$

The proof is easy and is omitted.

LEMMA 5.8. (1) *For $\boldsymbol{R}$-valued function h_t,*

$$E\left[\int_0^T h_t D_t^i dt \,\middle|\, \int_0^T a_t dw_t = x\right] = [C_1(h)_T]^{ij} \sigma_{jl} x^l , \qquad i=1, \cdots, d .$$

(2) *For $\boldsymbol{R}^1 \otimes \boldsymbol{R}^r$-valued function h_t,*

$$E\left[\int_0^T h_t D_t^i dw_t \,\middle|\, \int_0^T a_t dw_t = x\right] = [C_2^i(h)_T]^{mn} \sigma_{jm} \sigma_{ln} (x^j x^l - \sigma^{jl}) , \qquad i=1, \cdots, d .$$

(3) *For $\boldsymbol{R}$-valued function h_t,*

$$E\left[\int_0^T h_t D_t^i D_t^j dt \,\middle|\, \int_0^T a_t dw_t = x\right]$$
$$= \int_0^T h_t [C_3^{i,j}(t)]^{mn} dt \sigma_{pm} \sigma_{qn} (x^p x^q - \sigma^{pq})$$
$$+ \int_0^T \int_0^t h_t \lambda_{t,s}^i (\lambda_{t,s}^j)' ds dt , \qquad i, j=1, \cdots, d .$$

(4) *For $\boldsymbol{R}$-valued function h_t,*

$$E\left[\int_0^T h_t E_t^a dt \,\middle|\, \int_0^T a_t dw_t = x\right]$$
$$= \int_0^T h_t \int_0^t \nu_{i,j,t,s}^a [C_2^{i,j}(s)]^{mn} ds dt \sigma_{pm} \sigma_{qn} (x^p x^q - \sigma^{pq})$$
$$+ \int_0^T h_t \int_0^t \int_0^s \nu_{i,j,t,s}^a \lambda_{s,u}^i (\lambda_{s,u}^j)' du ds dt$$

$$+\int_0^T h_t[C_2^i(2\mu_{i,t,\cdot}^a)_t]^{mn}\sigma_{jm}\sigma_{ln}dt(x^j x^l-\sigma^{jl}) , \qquad \alpha=1, \cdots, d .$$

PROOF. In view of the representation

$$D_t=\int_0^t Y_t Y_s^{-1} V(X_s^0)dw_s ,$$

(1), (2) and (3) follow from (1), (2) and (3) of Lemma 5.7, respectively. (2) and (3) imply (4) if we note that

$$E_t=\int_0^t Y_t Y_s^{-1}[\partial_i\partial_j V_0(X_s^0)D_s^i D_s^j ds+2\partial_i V(X_s^0)D_s^i dw_s] .$$

□

From Lemmas 5.1 and 5.8, we have

LEMMA 5.9. *For* $\alpha=1, \cdots, k$,

$$\begin{aligned} e^\alpha(x) &= E[f_1^\alpha | f_0=x] \\ &= A^{0,\alpha}+A_l^{1,\alpha}[x-x^0]^l \\ &\quad +A_{p,q}^{2,\alpha}([x-x^0]^p[x-x^0]^q-\sigma^{pq}) , \end{aligned}$$

where $A^{0,\alpha}$, $A_l^{1,\alpha}$ *and* $A_{p,q}^{2,\alpha}$ *are defined in Section* 2.

PROOF OF THEOREM 2.1. By uniform integrability of $\{|\varphi^\epsilon(\tilde{F}^\epsilon)|^p; \epsilon\in(0, 1]\}$, $p>1$, Lemmas 5.4 and 5.5, we have

$$\begin{aligned} &E[\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)] \\ &\sim E[\psi(\xi_c^\epsilon)\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)] \\ &\sim E[\psi(\xi_c^\epsilon)\varphi^\epsilon(\tilde{F}^\epsilon)I_{A^\circ}(\tilde{F}^\epsilon)] \\ &\sim E[\Phi_0]+\epsilon E[\Phi_1]+\epsilon^2 E[\Phi_2]+\cdots \end{aligned}$$

Using the integration-by-parts formula we see that each term on the right-hand side is represented by an integration of some smooth function over A. In particular, $p_0(x)$ is easy and Lemmas 5.6 and 5.9 give $p_1(x)$. □

Acknowledgments

The author is grateful to Professor S. Watanabe of Kyoto University and Dr. N. Kunitomo of the University of Tokyo for their extensive help and advice, and also wishes to thank the editor and the referees for their valuable comments.

REFERENCES

[1] Akahira, M. and Takeuchi, K. (1981). Asymptotic efficiency of statistical estimators: concepts and higher order asymptotic efficiency, *Lecture Notes in Statistics*, **7**, Springer, Berlin Heidelberg New York.

[2] Genon-Catalot, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations, *Statistics*, **21**, 99–116.

[3] Ikeda, N. and Watanabe, S. (1984). An introduction to Malliavin's calculus, Stochastic Analysis, Proc. Taniguchi Intern. Symp. on Stochastic Analysis, Katata and Kyoto 1982, K. Itô (ed.), 1–52, Kinokuniya/North-Holland, Tokyo.

[4] Ikeda, N. and Watanabe, S. (1989). Stochastic differential equations and diffusion processes,

second edition, North-Holland/Kodansha, Tokyo.

[5] Kunitomo, N. and Takahashi, A. (1990). Pricing average options, discussion paper, Faculty of Economics, University of Tokyo. (in Japanese)

[6] Kusuoka, S. and Stroock, D. (1984). Applications of the Malliavin calculus I, Stochastic Analysis, Proc. Taniguchi Internatl. Symp. Katata and Kyoto, 1982, K. Itô (ed.), 271–306, Kinokuniya, Tokyo.

[7] Kusuoka, S. and Stroock, D. W. (1991). Precise asymptotics of certain Wiener functionals, *J. Functional Analysis*, **99**, 1–74.

[8] Kutoyants, Yu. A. (1978). Estimation of a parameter of a diffusion process, *Theory Probab. and Appl.*, **23**, 641–649.

[9] Kutoyants, Yu. A. (1984). Parameter estimation for stochastic processes, Translated and edited by B. L. S. Prakasa Rao, Heldermann, Berlin.

[10] Liptser, R. S. and Shiryayev, A. N. (1977). Statistics of random processes I, Springer, Berlin Heidelberg New York.

[11] Malliavin, P. (1976). Stochastic calculus of variation and hypoelliptic operators, Proc. Int. Symp. S.D.E. Kyoto, Itô, K. (ed.), Kinokuniya, 195–263.

[12] Malliavin, P. (1978). C^k-hypoellipticity with degeneracy, Stochastic Analysis, Friedman, A., Pinsky, M. (eds.), Academic Press, New York. 199–214, 327–340.

[13] Swe, M. and Taniguchi, M. (1989). Higher order asymptotic properties of Bayes estimators for Gaussian ARMA processes, Technical Report No. 237, Hiroshima University.

[14] Taniguchi, M. (1983). On the second order asymptotic efficiency of estimators of Gaussian ARMA processes, *Ann. Statist.*, **11**, 157–169.

[15] Taniguchi, M. (1987). Validity of Edgeworth expansion of minimum contrast estimators for Gaussian ARMA processes, *J. Multivariate Analysis*, **21**, 1–28.

[16] Taniguchi, M. (1991). Higher order asymptotic theory for time series analysis, *Lect. Notes in Statistics*, **68**, Springer, Berlin Heidelberg New York.

[17] Tsitovich, I. I. (1977). On estimating the drift of a diffusion process from observations in a domain, *Theory Probab. and Appl.*, **22**, 851–858.

[18] Watanabe, S. (1983). Malliavin's calculus in terms of generalized Wiener functionals, *Lecture Notes in Control and Inf. Sci.*, **49**, 284–290, Springer, Berlin Heidelberg New York.

[19] Watanabe, S. (1984). Lectures on stochastic differential equations and Malliavin calculus, Tata Institute of Fundamental Research, Springer, Berlin Heidelberg New York.

[20] Watanabe, S. (1987). Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels, *Ann. Probab.*, **15**, 1–39.

[21] Yoshida, N. (1990a). Asymptotic expansion for small diffusion—an application of Malliavin-Watanabe theory, *Research Memorandum*, **383**, Institute of Statistical Mathematics.

[22] Yoshida, N. (1990b). Asymptotic expansion for small diffusion. An application of the theory of Malliavin-Watanabe. General case and second order efficiency, *Research Memorandum*, **392**, Institute of Statistical Mathematics.

[23] Yoshida, N. (1991a). Asymptotic expansion for statistics related to small diffusion, *Research Memorandum*, **406**, Institute of Statistical Mathematics.

[24] Yoshida, N. (1991b). Asymptotic expansions for small noise system on Wiener space I: maximum likelihood estimator, *Research Memorandum*, **422**, Institute of Statistical Mathematics.