Probab. Theory Relat. Fields 95, 429–450 (1993)

Probability Theory and Related Fields

Asymptotic expansion of Bayes estimators for small diffusions

Nakahiro Yoshida

The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku, Tokyo 106, Japan

Received January 2, 1992; in revised form October 2, 1992

Summary. Using the Malliavin calculus we derived asymptotic expansion of the distributions of the Bayes estimators for small diffusions. The second order efficiency of the Bayes estimator is proved.

Mathematics Subject Classification (1980): 62M05, 62F12, 60H10

1 Introduction

Consider a family of d-dimensional diffusion processes defined by the stochastic differential equation

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

where w is an r-dimensional standard Wiener process, V_0 and $V=(V_1, \dots, V_r)$ are $\mathbf{R}^d$-valued and $\mathbf{R}^d \otimes \mathbf{R}^r$-valued smooth functions with (bounded) derivatives defined on $\mathbf{R}^d \times \Theta$ (Θ is a bounded convex domain of $\mathbf{R}^k$) and $\mathbf{R}^d$ respectively. T and x_0 are constants and $\varepsilon \in (0, 1]$ is a parameter. The parameter θ requires to be estimated from the observation $\{X_t^{\varepsilon,\theta}; t \in [0, T]\}$. It is known that the maximum likelihood estimator and Bayes estimator have consistency, asymptotic normality and first-order optimality when $\varepsilon \downarrow 0$. See Kutoyants [6]. To refine the normal approximation and to examine higher order properties of these estimators it is necessary to derive their asymptotic expansions. The asymptotic expansion for the distribution of the maximum likelihood estimator and its second order optimal property were proved in [21, 22]. It this paper, we show the asymptotic expansion for the Bayes estimator, from which we prove that it is optimal in the second order if its bias is appropriately corrected.

The underlying mathematical tool used here is the Malliavin calculus advanced by Watanabe [20]. This theory has been proved to be successfully applicable to the problems of the higher order statistical inference, [21, 22, 23]. Namely, it enables us to obtain asymptotic expansion of distributions of various statistics quite easily and intuitively by simple computation with the Taylor formula if some regularity condition is verified. When we use this theory, the crucial step is to show the nondegeneracy of the Malliavin covariance of Wiener functionals. However, it does not seem easy to do this even for a simple statistical estimator, whose Malliavin covariance is given by an integration of some anticipative process, cf. [21]. The Malliavin covariance corresponding to the Bayes estimator 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 on the Wiener space.

This paper is organized as follows. In Sect. 2, we state our main results. The second order efficiency of a bias corrected Bayes estimator is proved in Sect. 3. There we adopt the criterion by probability of concentration of estimators introduced and established by Takeuchi, Akahira and Pfanzagl. In Sect. 4, for convenience of reference, we prepare several notation and results about the above mentioned modification of the Malliavin calculus used later to prove the asymptotic expansions. Finally, Sect. 5 presents the proof of the results stated in Sect. 2.

2 Main results: the asymptotic expansions of Bayes estimators

Consider a parametric model of the d-dimensional small diffusions defined by (1.1). Let $P_{\varepsilon,\theta}$ be the distribution on $C([0,T],\mathbf{R}^d)$ of $X^{\varepsilon,\theta}$, the solution of (1.1) for ε and θ. The Radon-Nikodym derivative of $P_{\varepsilon,\theta}$ with respect to P_{ε,θ_0} is given by the formula (Liptser and Shiryayev [7])

$$\Lambda_\varepsilon(\theta;X)\,\Lambda_\varepsilon(\theta_0;X)^{-1},$$

where

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

Here A^+ denotes the Moore-Penrose generalized inverse matrix of matrix A and 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.

Let $\theta_0\in\Theta$ denote the true value of the unknown parameter θ. For $h\in\mathbf{R}^k$, the log likelihood ratio is defined by

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

where X_t^ε is the solution of the stochastic differential equation (1.1) for $\theta=\theta_0$. The function X_t^0 is defined by the ordinary differential equation (1.1) for $\varepsilon=0$ and $\theta=\theta_0$.

Let $\delta_i=\partial/\partial\theta^i$. 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 $i,j=1,\ldots,k$. From now, for simplicity, denote $I(\theta_0)$ by $I=(I_{ij})$ and $I(\theta_0)^{-1}$ by $I^{-1}=(I^{ij})$.

For function $f(x,\theta)$ of x and θ, $f_t^\varepsilon(\theta)$ denotes $f(X_t^\varepsilon,\theta)$. For $\mathbf{n}=(n_1,\ldots,n_d)$, let $\partial^{\mathbf{n}}=\partial_1^{n_1}\ldots\partial_d^{n_d}$, where $\partial_i=\dfrac{\partial}{\partial x^i}$, and let $|\mathbf{n}|=n_1+\ldots+n_d$. Moreover, for $v=(v_1,\ldots,v_k)$, let $\delta^v=\delta_1^{v_1}\ldots\delta_k^{v_k}$ and let $|v|=v_1+\ldots+v_k$.

In this article we assume the following conditions.

(1) V_0, V and $(VV')^+$ are smooth in $(x,\theta)\in\mathbf{R}^d\times\Theta$.
(2) For $|\mathbf{n}|\geqq 1$, $F=V_0, V, (VV')^+$, $\sup_{x,\theta}|\partial^{\mathbf{n}}F|<\infty$.
(3) For $|v|\geqq 1$ and $|\mathbf{n}|\geqq 0$, a constant $C_{v,\mathbf{n}}$ exists and

$$\sup_\theta|\partial^{\mathbf{n}}\delta^v V_0|\leqq C_{v,\mathbf{n}}(1+|x|)^{C_{v,\mathbf{n}}}$$

for all x.
(4) For any $\theta_0\in\Theta$, there exits a positive constant a_0 such that

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

for $\theta\in\Theta$.

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

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

where $[\partial V_\alpha]^{ij}=\partial_j V_\alpha^i$, $i,j=1,\ldots,d$, $\alpha=0,1,\ldots,r$. Then $Y_t:=Y_t^0$ is a nonsingular deterministic $\mathbf{R}^d\otimes\mathbf{R}^d$-valued process. For function g^ε, $g^{(j)}$ denotes its j-th derivative in ε at $\varepsilon=0$. We write $D_t=X_t^{(1)}$. Then D_t is represented by

$$D_t=\int_0^t Y_t Y_s^{-1} V_s^0\,dw_s,\qquad t\in[0,T].$$

We will use Einstein's rule for repeated indices. For matrix A, $[A]^{ij}$ denotes its (i,j)-element. For vector a, a^i is its i-th element.

Let

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

and

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

where

$$g_s=Y_s^{-1}[VV'(VV')^+\delta V_0]_s^0(\theta_0)=Y_s^{-1}\,\delta V_{0,s}^0(\theta_0).$$

Here we consider Bayes estimators with respect to the quadratic loss function and derive asymptotic expansions for their distributions. Let

$$G_\varepsilon(w)=\int_{B_\varepsilon}\exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0+\varepsilon h)\,dh$$

and

$$\bar{G}_\varepsilon(w)=\int_{B_\varepsilon} h\exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0+\varepsilon h)\,dh,$$

where π denotes the Bayes prior which is a smooth positive function on Θ whose derivatives are bounded and $\inf_{\theta\in\Theta}\pi(\theta)>0$, and $B_\varepsilon=\{h\in\mathbf{R}^k;\ \theta_0+\varepsilon h\in\Theta\}$. The Bayes estimator under $\theta=\theta_0$ is denoted by $\tilde{\theta}_\varepsilon(w;\theta_0)$ and defined by

$$\tilde{\theta}_\varepsilon(w;\theta_0)=\frac{\int_\Theta \theta\Lambda_\varepsilon(\theta;X^\varepsilon)\,\pi(\theta)\,d\theta}{\int_\Theta \Lambda_\varepsilon(\theta;X^\varepsilon)\,\pi(\theta)\,d\theta}.$$

For Borel set $A\subset B_\varepsilon$, let

$$\tilde{p}_\varepsilon(h;A)=\frac{\exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0+\varepsilon h)}{\int_A \exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0+\varepsilon h)\,dh}.$$

Then

$$\varepsilon^{-1}(\tilde{\theta}_\varepsilon(w;\theta_0)-\theta_0)=\int_{B_\varepsilon} h\tilde{p}_\varepsilon(h;B_\varepsilon)\,dh=\frac{\bar{G}_\varepsilon(w)}{G_\varepsilon(w)}.$$

In the context of the higher order statistical asymptotic theory we need to modify the Bayes estimator to obtain an efficient estimator. We call an estimator $\tilde{\theta}_\varepsilon^*$ a bias corrected Bayes estimator if

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

where $\tilde{b}(\theta)$ is a bounded smooth function with bounded derivatives. Let $\phi(x;\mu,\Sigma)$ denote the density of the normal distribution $N(\mu,\Sigma)$ on $\mathbf{R}^k$. The Borel σ-field of $\mathbf{R}^k$ is denoted by $\mathbf{B}^k$. Now, we have the following result.

Theorem 2.1 *The probability distribution of the bias corrected Bayes estimator* $\tilde{\theta}_\varepsilon^*(w;\theta_0)$ *has the asymptotic expansion*

$$P\left[\frac{\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0}{\varepsilon}\in A\right]\sim\int_A \tilde{p}_0(x)\,dx+\varepsilon\int_A \tilde{p}_1(x)\,dx+\dots$$

as $\varepsilon\downarrow 0$, $A\in\mathbf{B}^k$, *where* $\tilde{p}_0$, $\tilde{p}_1$, ... *are smooth functions. The expansion is uniform in Borel sets* $A\in\mathbf{B}^k$. *In particular,*

$$\begin{aligned}\tilde{p}_0(x)&=\phi(x;0,I^{-1}),\\ \tilde{p}_1(x)&=[I^{ij}A_{i,j,l}x^l-\tfrac{1}{2}I^{ij}B_{i,j,l}x^l-\tilde{b}(\theta_0)^j I_{jl}x^l+\pi(\theta_0)^{-1}\delta_l\pi(\theta_0)x^l\\ &\quad -A_{i,j,l}x^i x^j x^l-\tfrac{1}{2}B_{i,j,l}x^i x^j x^l]\,\phi(x;0,I^{-1}).\end{aligned}$$

More generally, we can show the asymptotic expansion of distribution of the bias corrected Bayes estimator under the contiguous alternative $\theta_0+\varepsilon h$, $h\in\mathbf{R}^k$. This is important from a statistical point of view.

Theorem 2.2 *The probability distribution of the bias corrected Bayes estimator* $\tilde{\theta}_\varepsilon^*(w;\theta_0+\varepsilon h)$ *under the contiguous alternative* $P_{\varepsilon,\theta_0+\varepsilon h}$ *has the asymptotic expansion*

$$P\left[\frac{\tilde{\theta}_\varepsilon^*(w;\theta_0+\varepsilon h)-(\theta_0+\varepsilon h)}{\varepsilon}\in A\right]\sim\int_A \tilde{p}_0^c(y)\,dy+\varepsilon\int_A \tilde{p}_1^c(y)\,dy+\dots$$

as $\varepsilon\downarrow 0$, $A\in\mathbf{B}^k$, $h\in\mathbf{R}^k$, *where* $\tilde{p}_0^c$, $\tilde{p}_1^c$, ... *are smooth functions. The expansion is uniform in* $A\in\mathbf{B}^k$ *and* $h\in K$, *where* K *is any compact set in* $\mathbf{R}^k$. *In particular,*

$$\begin{aligned}\tilde{p}_0^c(y)&=\phi(y;0,I^{-1}),\\ \tilde{p}_1^c(y)&=A_{i,j,l}[-y^i y^j y^l-h^l y^i y^j+I^{ij}y^l+I^{ij}h^l]\,\phi(y;0,I^{-1})\\ &\quad+B_{i,j,l}[-\tfrac{1}{2}y^i y^j y^l-h^i y^j y^l-\tfrac{1}{2}I^{ij}y^l+I^{il}h^j]\,\phi(y;0,I^{-1})\\ &\quad-\tilde{b}(\theta_0)^j I_{jl}y^l\phi(y;0,I^{-1})+\pi(\theta_0)^{-1}\delta_l\pi(\theta_0)y^l\phi(y;0,I^{-1}).\end{aligned}$$

3 Application: second order efficiency of a Bayes estimator

It is known that maximum likelihood estimators and Bayes estimators are asymptotically efficient for regular statistical experiments induced mainly from independent observations. As for the small diffusions, they have consistency and asymptotic normality and are efficient in the first order. See, e.g., Kutoyants [6]. Here we are interested in their second order efficiency. The notions of the second order efficiency of estimators have been introduced by Fisher, Rao [12, 13], Takeuchi, Akahira, Pfanzagl and others. Here we adopt the criterion by probability of concentration of estimators introduced and established by Pfanzagl, Takeuchi and Akahira. It is possible to show that these estimators are optimal in the second order in this criterion in various cases. See for example Akahira and Takeuchi [1], Pfanzagl [10, 11]. For time series see Taniguchi [15, 16, 17], Swe and Taniguchi [14]. The second order efficiency of the maximum likelihood estimator was proved in [22].

To avoid meaningless super-efficiency, an invariant condition is imposed on estimators in question. For simplicity we consider the case $k=1$.

Definition 3.1 An estimator T_ε is *second order asymptotically median unbiased* (*second order* AMU) if for any $\theta_0 \in \Theta$ and any $c>0$

$$\lim_{\varepsilon \downarrow 0} \sup_{\theta \in \Theta: |\theta-\theta_0|<\varepsilon c} \varepsilon^{-1} |P_{\varepsilon,\theta}[T_\varepsilon - \theta \leqq 0] - \tfrac{1}{2}| = 0$$

and

$$\lim_{\varepsilon \downarrow 0} \sup_{\theta \in \Theta: |\theta-\theta_0|<\varepsilon c} \varepsilon^{-1} |P_{\varepsilon,\theta}[T_\varepsilon - \theta \geqq 0] - \tfrac{1}{2}| = 0.$$

Then we have the following theorem from Neyman-Pearson's fundamental lemma. See [22, 23] for details. For $h \in \mathbf{R}$, let $J = I h^2$. The distribution function of the normal distribution $N(\mu, \sigma^2)$ is denoted by $\Phi(x; \mu, \sigma^2)$.

Theorem 3.1 *For any second order* AMU *estimator* T_ε, *for* $h>0$

$$\liminf_{\varepsilon \downarrow 0} \varepsilon^{-1} \{\Phi(J; 0, J) + \varepsilon[A_{1,1,1} h^3 + \tfrac{1}{2} B_{1,1,1} h^3] \phi(J; 0, J) - P_{\varepsilon,\theta_0}[\varepsilon^{-1}(T_\varepsilon - \theta_0) \leqq h]\} \geqq 0,$$

and for $h<0$

$$\limsup_{\varepsilon \downarrow 0} \varepsilon^{-1} \{\Phi(-J; 0, J) - \varepsilon[A_{1,1,1} h^3 + \tfrac{1}{2} B_{1,1,1} h^3] \phi(J; 0, J) - P_{\varepsilon,\theta_0}[\varepsilon^{-1}(T_\varepsilon - \theta_0) \leqq h]\} \leqq 0.$$

Definition 3.2

$$\Phi(J; 0, J) + \varepsilon[A_{1,1,1} h^3 + \tfrac{1}{2} B_{1,1,1} h^3] \phi(J; 0, J)$$

and

$$\Phi(-J; 0, J) - \varepsilon[A_{1,1,1} h^3 + \tfrac{1}{2} B_{1,1,1} h^3] \phi(J; 0, J)$$

are called the *bounds of second order distributions*. A second order AMU estimator is said to be *second order efficient* if it attains these bounds for any $h>0$ and any $h<0$.

From Theorems 2.1, 2.2 and 3.1, we obtain

Corollary 3.1 *The Bayes estimator is second order* AMU *and second order efficient if its bias is corrected by*

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

4 Preparations: the Malliavin calculus with truncation

We begin with preparing notations used in the Malliavin calculus. For details see Malliavin [8, 9], Watanabe [18, 19, 20], Ikeda and Watanabe [3, 4], and

Kusuoka and Stroock [5]. 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 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, $s \in \mathbf{R}$, $p \in (1, \infty)$, by

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

where L is the Ornstein-Uhlenbeck operator. The Banach space $D_p^s(E)$ is the completion of the totality of E-valued polynomials on the Wiener space (W, P) with respect to $\|\cdot\|_{p,s}$. The space $D^\infty(E) = \bigcap_{s>0} \bigcap_{1<p<\infty} D_p^s(E)$ is the set of Wiener test functionals and $\tilde{D}^{-\infty}(E) = \bigcup_{s>0} \bigcap_{1<p<\infty} D_p^{-s}(E)$ is a space of generalized Wiener functionals. See Watanabe [19]. We suppress $\mathbf{R}$ when $E = \mathbf{R}$. The Fréchet space $\mathsf{S}(\mathbf{R}^d)$ is the totality of rapidly decreasing smooth functions on $\mathbf{R}^d$ and $\mathsf{S}'(\mathbf{R}^d)$ is its dual. Let $A = 1 + |x|^2 - \frac{1}{2}\Delta$.

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

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

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

$$\psi(\xi)\, G f \circ F = \psi(\xi)\, G f(F)$$

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

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

$$\limsup_{\varepsilon \downarrow 0} \frac{\|F_\varepsilon\|_{p,s}}{\varepsilon^k} < \infty,$$

we say $F_\varepsilon(w) = O(\varepsilon^k)$ in $D_p^s(E)$ as $\varepsilon \downarrow 0$. It is said that $F_\varepsilon(w) \in D^\infty(E)$ has the asymptotic expansion

$$F_\varepsilon(w) \sim f_0 + \varepsilon f_1 + \dots$$

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

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

in $D_p^s(E)$ as $\varepsilon\downarrow 0$. Similarly, we say that $F_\varepsilon(w)\in\tilde{D}^{-\infty}(E)$ has the asymptotic expansion

$$F_\varepsilon(w)\sim f_0+\varepsilon f_1+\dots$$

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

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

in $D_p^{-s}(E)$ as $\varepsilon\downarrow 0$. The generalized means of these expansions yield the ordinary asymptotic expansions.

The following theorem is a version of Theorem 2.3 of Watanabe [20].

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

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

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

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

$$F_\varepsilon(w)\sim f_0+\varepsilon f_1+\dots \textit{ in } D^\infty(\mathbf{R}^d) \textit{ as } \varepsilon\downarrow 0$$

with $f_i\in D^\infty(\mathbf{R}^d)$.
(3) *$\xi_\varepsilon(w)=O(1)$ in D^∞ as $\varepsilon\downarrow 0$.*
(4) *For any $n=1, 2, \dots,$*

$$\lim_{\varepsilon\downarrow 0}\varepsilon^{-n}P\{|\xi_\varepsilon|>\tfrac{1}{2}\}=0.$$

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

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

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

$$G_{\mu,\varepsilon}(w)\sim g_{\mu,0}+\varepsilon g_{\mu,1}+\dots$$

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

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

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

$$G_{\mu,\varepsilon}\, T_\lambda(f_0 + [\varepsilon f_1 + \varepsilon^2 f_2 + \ldots]) = (g_{\mu,0} + \varepsilon g_{\mu,1} + \ldots) \sum_{\mathbf{n}} \frac{1}{\mathbf{n}!} \partial^{\mathbf{n}}\, T_\lambda(f_0)[\varepsilon f_1 + \varepsilon^2 f_2 + \ldots]^{\mathbf{n}}$$
$$= \Phi_{\lambda,\mu,0} + \varepsilon \Phi_{\lambda,\mu,1} + \ldots,$$

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

$$\Phi_{\lambda,\mu,0} = g_{\mu,0}\, T_\lambda(f_0),$$
$$\Phi_{\lambda,\mu,1} = g_{\mu,0} \sum_{i=1}^{d} f_1^i\, \partial_i\, T_\lambda(f_0) + g_{\mu,1}\, T_\lambda(f_0),$$
$$\Phi_{\lambda,\mu,2} = g_{\mu,0} \left\{ \sum_{i=1}^{d} f_2^i\, \partial_i\, T_\lambda(f_0) + \tfrac{1}{2} \sum_{i,j=1}^{d} f_1^i f_1^j\, \partial_i\, \partial_j\, T_\lambda(f_0) \right\}$$
$$+ g_{\mu,1} \sum_{i=1}^{d} f_1^i\, \partial_i\, T_\lambda(f_0) + g_{\mu,2}\, T_\lambda(f_0), \ldots.$$

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

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

5 Proof of the main results

Let

$$m((i_1, \ldots, i_p)/0) = \int_0^T [\delta_{i_1} \ldots \delta_{i_p} V_0'(VV')^+ V]_t^0(\theta_0)\, dw_t,$$
$$m((i_1, \ldots, i_p)/1) = \int_0^T [\partial_l \{\delta_{i_1} \ldots \delta_{i_p} V_0'(VV')^+ V\}]_t^0(\theta_0)\, D_t^l\, dw_t,$$
$$n((i_1, \ldots, i_p)(j_1, \ldots, j_q)/0) = \int_0^T [\delta_{i_1} \ldots \delta_{i_p} V_0'(VV')^+ \delta_{j_1} \ldots \delta_{j_q} V_0]_t^0(\theta_0)\, dt,$$

and

$$n((i_1, \ldots, i_p)(j_1, \ldots, j_q)/1) = \int_0^T \partial_l [\delta_{i_1} \ldots \delta_{i_p} V_0'(VV')^+ \delta_{j_1} \ldots \delta_{j_q} V_0]_t^0(\theta_0)\, D_t^l\, dt.$$

We know that for any $a_0 > 0$, there exist positive constants a_i, $i = 1, 2$, independent of ε such that

$$P[\sup_{0 \leqq t \leqq T} |X_t^\varepsilon - X_t^0| > a_0] \leqq a_1 \exp\{-a_2\, \varepsilon^{-2}\}$$

for $\varepsilon\in(0,1]$. Therefore the Bayes estimator obtained from the stopped data $\{X^\varepsilon_{t\wedge\tau};\ 0\leqq t\leqq T\}$, where $\tau=\inf\{t\geqq 0;\ |V(X^\varepsilon_t)|>2\alpha\}\wedge T$, $\alpha=\sup_{0\leqq t\leqq T}|V(X^0_t)|$, has the same asymptotic expansion as $\tilde{\theta}_\varepsilon(w;\theta_0)$. So to show Theorems 2.1 and 2.2 we may assume that V is bounded without loss of generality.

Lemma 5.1 *There exists $\varepsilon_0>0$ such that for any $p\geqq 0$*

$$\sup_{\varepsilon\leqq\varepsilon'<\varepsilon_0} E\Big[\int_{B_{\varepsilon'}} |h|^p\,\tilde{p}_\varepsilon(h;B_{\varepsilon'})\,dh\Big]<\infty.$$

Proof. Let $\varphi(x)$: $\mathbf{R}^d\to[0,1]$ be a smooth function such that $\varphi(x)=1$ for $|x|\leqq 2a$ and $\varphi(x)=0$ for $|x|\geqq 3a$, where $a=\sup_{0\leqq t\leqq T}|X^0_t|$. Let $W(x)=\varphi(x)(VV')^+(x)$. Then we see that

$$\begin{aligned}
J^\varepsilon(h):=&\int_0^T [V^\varepsilon_{0,t}(\theta_0+\varepsilon h)-V^\varepsilon_{0,t}(\theta_0)]'(VV')^{+\varepsilon}_t[V^\varepsilon_{0,t}(\theta_0+\varepsilon h)-V^\varepsilon_{0,t}(\theta_0)]\,dt\\
\geqq&\int_0^T \Delta^0(\theta_0+\varepsilon h,\theta_0)'(VV')^{+0}_t\,\Delta^0(\theta_0+\varepsilon h,\theta_0)\,dt\\
&-\int_0^T |\Delta^0(\theta_0+\varepsilon h,\theta_0)'\{W^\varepsilon_t-W^0_t\}\,\Delta^0(\theta_0+\varepsilon h,\theta_0)|\,dt\\
&-2\int_0^T |\{\Delta^\varepsilon(\theta_0+\varepsilon h,\theta_0)-\Delta^0(\theta_0+\varepsilon h,\theta_0)\}'\,W^\varepsilon_t\,\Delta^0(\theta_0+\varepsilon h,\theta_0)|\,dt,
\end{aligned}$$

where $\Delta^\varepsilon(\theta,\theta_0)=V_0(X^\varepsilon_t,\theta)-V_0(X^\varepsilon_t,\theta_0)$. We can find, by a representation theorem for continuous martingales, Brownian motions w^ε_i, $i=1,2,\dots,d$, for some filtrations, and a positive constant C independent of ε such that

$$\sup_{0\leqq t\leqq T}|X^\varepsilon_t-X^0_t|\leqq C\varepsilon w^*,$$

where $w^*=\sum_{i=1}^d \sup_{0\leqq t\leqq T}|w^\varepsilon_i(t)|$. Since $\sup_{\theta\in\Theta}|\Delta^0(\theta,\theta_0)|<\infty$,

$$-\int_0^T |\Delta^0(\theta_0+\varepsilon h,\theta_0)'\{W^\varepsilon_t-W^0_t\}\,\Delta^0(\theta_0+\varepsilon h,\theta_0)|\,dt\geqq -C_1\,\varepsilon^2\,w^*|h|$$

for some $C_1>0$. From condition (4) we obtain

$$J^\varepsilon(h)\geqq C_2\,\varepsilon^2|h|^2-C_3\,\varepsilon^2\,w^*|h|$$

for some C_2, $C_3>0$. Then we can show the large deviation inequality

$$E\left[\exp\left\{-\frac{p}{\varepsilon^2}J^\varepsilon(h)\right\}\right]\leqq C_4\,e^{-C_5|h|^2},\qquad h\in B_\varepsilon,\qquad \varepsilon\in(0,1]$$

for some constants $p, C_4, C_5 > 0$. See p. 91 of Kutoyants [6]. Then we have in standard argument

$$P[\exp(l_{\varepsilon,h}(w;\theta_0)) > \exp(-C_6|h|^2)] \leqq C_7 \exp(-C_6|h|^2), \qquad h \in B_\varepsilon, \qquad \varepsilon \in (0,1]$$

for some $C_6, C_7 > 0$. Following the proof of Lemma 5.2 of Chap. 1 of Ibragimov and Has'minskii [2], we obtain the result. □

Lemma 5.2 $\dfrac{\bar{G}_\varepsilon(w)}{G_\varepsilon(w)}$ *is well-defined on W and in* $D^\infty(\mathbf{R}^k)$ for each $\varepsilon < \varepsilon_0$.

Proof. For any $p > 1$,

$$E\left(\left|\frac{\bar{G}_\varepsilon(w)}{G_\varepsilon(w)}\right|^p\right) = E\Big(\Big|\int_{B_\varepsilon} h\tilde{p}_\varepsilon(h; B_\varepsilon)\,dh\Big|^p\Big)$$
$$\leqq E\Big(\int_{B_\varepsilon} |h|^p\,\tilde{p}_\varepsilon(h; B_\varepsilon)\,dh\Big) < \infty$$

by Lemma 5.1. For $\eta \in H$, we have

$$D_\eta G_\varepsilon(w) = \int_{B_\varepsilon} \exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0 + \varepsilon h)\,D_\eta l_{\varepsilon,h}(w;\theta_0)\,dh$$
$$= \int_{B_\varepsilon} \exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0 + \varepsilon h)\left(\int_0^T \sum_{i=1}^r \xi^i_{\varepsilon,h}(t)\,\dot{\eta}^i_t\,dt\right)dh$$
$$= \int_0^T \sum_{i=1}^r \Big[\int_{B_\varepsilon} \exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0 + \varepsilon h)\,\xi^i_{\varepsilon,h}(t)\,dh\Big]\,\dot{\eta}^i_t\,dt,$$

where $\xi^i_{\varepsilon,h}(t)$ is an anticipative process for which there exist $m > 0$ and a random variable $Q_\varepsilon(w)$ such that

$$\sup_{0 \leqq t \leqq T} |\xi^i_{\varepsilon,h}(t)| \leqq (1 + |h|^m)\,Q_\varepsilon(w)$$

and $\sup_\varepsilon E[Q^p_\varepsilon] < \infty$ for any $p > 1$. Here we used the moment inequalitites for random fields (or Sobolev's inequality and Burkholder-Davis-Gundy's inequality) to estimate the moments of the form

$$E\left(\sup_{\substack{\theta \in \Theta \\ t \in [0,T]}} \left|\int_0^t f(w, \varepsilon, s, \theta)\,dw_s\right|^p\right).$$

Then

$$\left|\frac{DG_\varepsilon(w)}{G_\varepsilon(w)}\right|^2_H = G_\varepsilon^{-2} \int_0^T \sum_i \Big[\int_{B_\varepsilon} \exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0 + \varepsilon h)\,\xi^i_{\varepsilon,h}(t)\,dh\Big]^2\,dt.$$

Therefore we have $E\left[\left|\frac{DG_\varepsilon(w)}{G_\varepsilon(w)}\right|_H^p\right] < \infty$ for $p > 1$ from Lemma 5.1. We obtain similar estimates for $\frac{D\bar{G}_\varepsilon(w)}{G_\varepsilon}$ and hence $E\left[\left|D\frac{\bar{G}_\varepsilon(w)}{G_\varepsilon(w)}\right|_H^p\right] < \infty$ for $p > 1$. In a similar way we can estimate higher order H-derivatives, which completes the proof. □

The log likelihood ratio $l_{\varepsilon,h}(w;\theta_0)$ is in D^∞ and has the asymptotic expansion

$$l_{\varepsilon,h}(w;\theta_0) \sim f_0^L + \varepsilon f_1^L + \dots$$

in D^∞ as $\varepsilon \downarrow 0$ with $f_0^L, f_1^L, \dots \in D^\infty$. In particular,

$$\begin{aligned} f_0^L &= h'B - \tfrac{1}{2}h'I(\theta_0)h, \qquad B^i = m((i)/0), \\ f_1^L &= h^i m((i)/1) + \tfrac{1}{2}h^i h^j m((ij)/0) \\ &\quad - \tfrac{1}{2}h^i h^j n((i)(j)/1) - \tfrac{1}{2}h^i h^j h^m n((ij)(m)/0). \end{aligned}$$

The following lemma gives an expansion formula for the bias corrected Bayes estimator.

Lemma 5.3 *There exist Wiener functionals $\psi_\varepsilon(w) \in D^\infty$ satisfying the following conditions.*
(1) $0 \leqq \psi_\varepsilon(w) \leqq 1$ and $\psi_\varepsilon(w) = 1 - O(\varepsilon^n)$ in D^∞ as $\varepsilon \downarrow 0$ for any $n \in \mathbf{N}$.
(2) $\psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0) - \theta_0) \in D^\infty(\mathbf{R}^k)$ has the asymptotic expansion

$$\psi_\varepsilon(w)\frac{\tilde{\theta}_\varepsilon^*(w;\theta_0) - \theta_0}{\varepsilon} \sim \tilde{f}_0 + \varepsilon\tilde{f}_1 + \dots \text{ in } D^\infty(\mathbf{R}^k) \text{ as } \varepsilon \downarrow 0$$

with $\tilde{f}_0, \tilde{f}_1, \dots \in D^\infty(\mathbf{R}^k)$. In particular,

$$\begin{aligned} \tilde{f}_0 &= I(\theta_0)^{-1}B, \\ \tilde{f}_1 &= \tfrac{1}{2}I(\theta_0)^{-1}\Gamma + \tfrac{1}{2}I(\theta_0)^{-1}QI(\theta_0)^{-1}B - \tilde{b}(\theta_0) \\ &\quad + \pi(\theta_0)^{-1}I(\theta_0)^{-1}\delta\pi(\theta_0)' + \tfrac{1}{2}I(\theta_0)^{-1}R, \end{aligned}$$

where $B = (B^i)$, $\Gamma = (\Gamma^i)$, $Q = (Q_{i,j})$ and $R = (R_i)$ are defined as follows.

$$\begin{aligned} B^i &= m((i)/0), \qquad i = 1, \dots, k, \\ \Gamma^i &= 2m((i)/1), \qquad i = 1, \dots, k, \\ Q_{i,j} &= \sum_{m=1}^{k}[I(\theta_0)^{-1}B]^m N_{i,j,m} + 2\Delta_{i,j}, \qquad i,j = 1, \dots, k, \\ R_i &= \sum_{j,m=1}^{k}[I(\theta_0)^{-1}]^{jm}N_{i,j,m}, \qquad i = 1, \dots, k, \end{aligned}$$

where

$$\begin{aligned} N_{i,j,m} &= -[n((ji)(m)/0) + n((im)(j)/0) + n((mj)(i)/0)], \qquad i,j,m = 1, \dots, k, \\ \Delta_{i,j} &= m((ji)/0) - n((j)(i)/1), \qquad i,j = 1, \dots, k. \end{aligned}$$

Proof. For Borel set $A \subset B_\varepsilon$, let

$$U(\varepsilon; A) = \int_A h \tilde{p}_\varepsilon(h; A)\, dh.$$

For each $\varepsilon' < \varepsilon_0$, the function $\varepsilon \to U(\varepsilon; B_{\varepsilon'})$ is smooth in $[0, \varepsilon']$. Here we consider the right derivatives and left derivatives at $\varepsilon = 0$ and $\varepsilon = \varepsilon'$, respectively. It is possible to show that for $j = 0, 1, \ldots$ and $\varepsilon < \varepsilon_0$ there exist nonnegative random variables $Q_{j,\varepsilon}(w)$ which satisfy the following two conditions.

(1) For each j and any $p > 1$, $\sup_{\varepsilon < \varepsilon_0} E[Q_{j,\varepsilon}(w)^p] < \infty$.

(2) If $\varepsilon \leqq \varepsilon' < \varepsilon_0$, $\dfrac{\partial^j U}{\partial \varepsilon^j}(\varepsilon; B_{\varepsilon'})$ is dominated by a polynomial in $Q_{j,\varepsilon}(w)$ and $\int_{B_{\varepsilon'}} |h|^p \tilde{p}_\varepsilon(h; B_{\varepsilon'})\, dh$, $p = 0, 1, \ldots$.

Therefore, by Lemma 5.1, we have

$$\sup_{\varepsilon \leqq \varepsilon' < \varepsilon_0} E\left[\left|\frac{\partial^j U}{\partial \varepsilon^j}(\varepsilon; B_{\varepsilon'})\right|^p\right] < \infty$$

for $p > 1$. From the Taylor formula, for $l \in \mathbf{N}$ and $0 \leqq \varepsilon \leqq \varepsilon' < \varepsilon_0$,

$$U(\varepsilon; B_{\varepsilon'}) = \sum_{j=0}^{l-1} \frac{\varepsilon^j}{j!} \frac{\partial^j U}{\partial \varepsilon^j}(0; B_{\varepsilon'}) + \frac{\varepsilon^l}{(l-1)!} \int_0^1 (1-u)^{l-1} \frac{\partial^l U}{\partial \varepsilon^l}(u\varepsilon; B_{\varepsilon'})\, du.$$

Hence we can show that for $l \in \mathbf{N}$, $p > 1$ and $s = 0$ there exists $C_{l,p,s} > 0$ such that

$$\left\| U(\varepsilon; B_\varepsilon) - \sum_{j=0}^{l-1} \frac{\varepsilon^j}{j!} \frac{\partial^j U}{\partial \varepsilon^j}(0; B_\varepsilon) \right\|_{p,s} \leqq C_{l,p,s}\, \varepsilon^l \tag{5.1}$$

if $\varepsilon < \varepsilon_0$, and also, as in the proof of Lemma 5.2 using Lemma 5.1, that (5.1) holds for $l \in \mathbf{N}$, $p > 1$ and $s > 0$. For $j = 0, 1, \ldots$, define $\dfrac{\partial^j U}{\partial \varepsilon^j}(0; \mathbf{R}^k)$ by

$$\frac{\partial^j U}{\partial \varepsilon^j}(0; \mathbf{R}^k) = \lim_{R \to \infty} \frac{\partial^j U}{\partial \varepsilon^j}(0; \{h; |h| \leqq R\}).$$

Let $\psi_\varepsilon(w) = \psi(c_1 |\varepsilon I^{-1} B|^2)$, where c_1 is any positive constant. Then $\psi_\varepsilon(w) = 1 - O(\varepsilon^n)$ in D^∞ as $\varepsilon \downarrow 0$ for any $n \in \mathbf{N}$. For any $p > 1$ there exist positive numbers b_1 and b_2 such that

$$E[\psi_\varepsilon(w)(\int_{|h| \geqq H} \tilde{p}_0(h; B_\varepsilon)\, dh)^p] < b_1\, e^{-b_2 H^2}$$

for $H \geqq 0$ and $\varepsilon < \varepsilon_0$. Indeed, with $c_0 = (2\pi)^{-\frac{k}{2}}(\det I)^{\frac{1}{2}}$,

$$\begin{aligned}
&E[\{(\int_{\mathbf{R}^k} \exp(h'B - \tfrac{1}{2}h'Ih)\,dh)^{-1} \int_{|h|\geqq H} \exp(h'B - \tfrac{1}{2}h'Ih)\,dh\}^p] \\
&= E[\{c_0 \int_{|h|\geqq H} \exp(-\tfrac{1}{2}(h - I^{-1}B)'I(h - I^{-1}B))\,dh\}^p] \\
&\leqq P(|I^{-1}B| \geqq \tfrac{1}{2}H) + E[1\{|I^{-1}B| < \tfrac{1}{2}H\} \\
&\quad \cdot c_0 \int_{|h|\geqq H} \exp(-\tfrac{1}{2}(h - I^{-1}B)'I(h - I^{-1}B))\,dh] \\
&\leqq b_1' e^{-b_2' H^2},
\end{aligned}$$

$H \geqq 0$, for some positive numbers b_1' and b_2'; and we can find $\delta > 0$ for which

$$\begin{aligned}
&\psi_\varepsilon(w)\{[\int_{B_\varepsilon} \exp(h'B - \tfrac{1}{2}h'Ih)\,dh]^{-1} \int_{\mathbf{R}^k} \exp(h'B - \tfrac{1}{2}h'Ih)\,dh\}^p \\
&= \psi_\varepsilon(w)\left[\frac{c_0}{\varepsilon^k} \int_{\Theta - \theta_0} \exp\left(-\frac{1}{2\varepsilon^2}(t - \varepsilon I^{-1}B)'I(t - \varepsilon I^{-1}B)\right) dt\right]^{-p} \\
&\leqq \delta^{-p}
\end{aligned}$$

for $\varepsilon < \varepsilon_0$, where $\Theta - \theta_0 = \{\theta - \theta_0\,;\, \theta \in \Theta\}$. We can show that, for $j = 0, 1, 2, \ldots, p > 1$ and $s \geqq 0$, there exist positive constants $a_j(p, s)$ and $c_j(p, s)$ such that

$$\left\| \psi_\varepsilon(w)\left(\frac{\partial^j U}{\partial \varepsilon^j}(0; B_\varepsilon) - \frac{\partial^j U}{\partial \varepsilon^j}(0; \mathbf{R}^k)\right)\right\|_{p,s} \leqq c_j(p, s)\, e^{-a_j(p,s)\varepsilon^{-2}} \tag{5.2}$$

for $\varepsilon < \varepsilon_0$. In fact, since

$$\begin{aligned}
U(0; B_\varepsilon) - U(0; \mathbf{R}^k) = &- \int_{\mathbf{R}^k - B_\varepsilon} h\tilde{p}_0(h; B_\varepsilon)\,dh \\
&+ \int_{\mathbf{R}^k} h\tilde{p}_0(h; \mathbf{R}^k)\,dh \cdot \int_{\mathbf{R}^k - B_\varepsilon} \tilde{p}_0(h; B_\varepsilon)\,dh,
\end{aligned}$$

it follows that, for $p > 1$ and $s \geqq 0$, there exist positive constants $a_0(p, s)$ and $c_0(p, s)$ such that (5.2) holds true for $j = 0$. In a similar manner, we can show (5.2) for $j = 1, 2, \ldots$. From (5.1) and (5.2), it follows that, for $l \in \mathbf{N}$, $p > 1$ and $s \geqq 0$, there exists positive $c(l, p, s)$ such that

$$\left\| \psi_\varepsilon(w)\, U(\varepsilon; B_\varepsilon) - \sum_{j=0}^{l-1} \frac{\varepsilon^j}{j!} \frac{\partial^j U}{\partial \varepsilon^j}(0; \mathbf{R}^k)\right\|_{p,s} \leqq c(l, p, s)\varepsilon^l$$

for $\varepsilon < \varepsilon_0$. Therefore $\psi_\varepsilon(w)\, U(\varepsilon; B_\varepsilon) = \psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon(w; \theta_0) - \theta_0)$ has an asymptotic expansion

$$\psi_\varepsilon(w)\varepsilon^{-1}(\tilde{\theta}_\varepsilon(w; \theta_0) - \theta_0) \sim g_0 + \varepsilon g_1 + \cdots$$

in $D^\infty(\mathbf{R}^k)$ as $\varepsilon \downarrow 0$ with $g_0, g_1, \ldots \in D^\infty(\mathbf{R}^k)$. In particular, the first two terms are given by $g_0 = U(0; \mathbf{R}^k)$ and $g_1 = \dfrac{\partial U}{\partial \varepsilon}(0; \mathbf{R}^k)$.

Next, we determine these two terms. Put

$$\mu=(2\pi)^{\frac{k}{2}} \det I(\theta_0)^{-\frac{1}{2}} \exp(\tfrac{1}{2} B' I(\theta_0)^{-1} B)$$

and

$$\mu_{i_1 i_2 \ldots i_p}=\int_{\mathbf{R}^k} h^{i_1} h^{i_2} \ldots h^{i_p} e^{f_0^L} dh,$$

where $f_0^L=h' B-\frac{1}{2} h' I(\theta_0) h$. Define $q(\varepsilon, h)$ by $q(\varepsilon, h)=\exp(l_{\varepsilon,h}(w;\theta_0))\,\pi(\theta_0+\varepsilon h)$. Then we have Eqs. (5.3):

$$\lim_{R\to\infty} \int_{|h|\leqq R} q(0,h)\,dh=\pi(\theta_0)\mu,$$

$$\lim_{R\to\infty} \int_{|h|\leqq R} \frac{\partial q}{\partial \varepsilon}(0,h)\,dh=\pi(\theta_0)[\mu_i\, m((i)/1)+\frac{1}{2}\mu_{ij}\, m((ij)/0)-\frac{1}{2}\mu_{ij}\, n((i)(j)/1)$$

$$-\frac{1}{2}\mu_{ijm}\, n((ij)(m)/0)+\pi(\theta_0)^{-1}\delta_i \pi(\theta_0)\,\mu_i],$$

$$\lim_{R\to\infty} \int_{|h|\leqq R} h^l q(0,h)\,dh=\pi(\theta_0)\mu_l,$$

$$\lim_{R\to\infty} \int_{|h|\leqq R} h^l \frac{\partial q}{\partial \varepsilon}(0,h)\,dh=\pi(\theta_0)[\mu_{li} m((i)/1)+\frac{1}{2}\mu_{lij}\, m((ij)/0)-\frac{1}{2}\mu_{lij}\, n((i)(j)/1)$$

$$-\frac{1}{2}\mu_{lijm}\, n((ij)(m)/0)+\pi(\theta_0)^{-1}\delta_i \pi(\theta_0)\,\mu_{li}].$$

Let $u=I(\theta_0)^{-1} B$. Then we can obtain

$$\begin{aligned}
\text{(5.4)}\qquad \mu_{i_1}&=\mu[u]^{(i_1)},\\
\mu_{i_1 i_2}&=\mu([uu]^{(i_1 i_2)}+I^{(i_1 i_2)}),\\
\mu_{i_1 i_2 i_3}&=\mu([uuu]^{(i_1 i_2 i_3)}+[Iu]^{(i_1 i_2 i_3)}),\\
\mu_{i_1 i_2 i_3 i_4}&=\mu([uuuu]^{(i_1 i_2 i_3 i_4)}+[Iuu]^{(i_1 i_2 i_3 i_4)}+[II]^{(i_1 i_2 i_3 i_4)}),
\end{aligned}$$

where $[\;]^{(i_1\ldots i_p)}$ denotes the sum of the terms corresponding to all partitions of $(i_1 \ldots i_p)$, e.g., $[Iu]^{(i_1 i_2 i_3)}=I^{i_1 i_2} u^{i_3}+I^{i_2 i_3} u^{i_1}+I^{i_3 i_1} u^{i_2}$. From (5.3) and (5.4),

$$\begin{aligned}
g_0^l=U(0;\mathbf{R}^k)^l&=\int_{\mathbf{R}^k} h^l \tilde{p}_0(h;\mathbf{R}^k)\,dh\\
&=\lim_{R\to\infty}\Big[\int_{|h|\leqq R} q(0,h)\,dh\Big]^{-1} \int_{|h|\leqq R} h^l q(0,h)\,dh\\
&=u^l=[I(\theta_0)^{-1}B]^l.
\end{aligned}$$

We note that

$$\text{(5.5)}\qquad (I^{li} u^j+I^{lj} u^i)\,\Delta_{i,j}=[2I^{-1}\Delta I^{-1} B]^l,$$

where $\Delta=(\Delta_{i,j})$, and that

$$\begin{aligned}
\text{(5.6)}\qquad &-\tfrac{1}{2}(I^{li} u^j u^m+I^{lj} u^i u^m+I^{lm} u^i u^j+[II]^{(lijm)})\, n((ij)(m)/0)\\
&=\tfrac{1}{2} I^{li} u^j u^m N_{i,j,m}+\tfrac{1}{2} I^{li} I^{jm} N_{i,j,m}.
\end{aligned}$$

Since

$$
\begin{aligned}
g_1^l &= \frac{\partial U^l}{\partial \varepsilon}(0; \mathbf{R}^k) \\
&= \lim_{R\to\infty} \left(\frac{\partial}{\partial \varepsilon}\right)_0 \{[\int_{|h|\leq R} q(\varepsilon, h)\, dh]^{-1} \int_{|h|\leq R} h^l q(\varepsilon, h)\, dh\} \\
&= \lim_{R\to\infty} [\int_{|h|\leq R} q(0, h)\, dh]^{-1} \int_{|h|\leq R} h^l \frac{\partial q}{\partial \varepsilon}(0, h)\, dh \\
&\quad - \lim_{R\to\infty} [\int_{|h|\leq R} q(0, h)\, dh]^{-2} \int_{|h|\leq R} h^l q(0, h)\, dh \int_{|h|\leq R} \frac{\partial q}{\partial \varepsilon}(0, h)\, dh,
\end{aligned}
$$

if follows, from (5.3) and (5.4), that the right-hand side equals

$$
\begin{aligned}
&\mu^{-1}[\mu_{li}\, m((i)/1) + \tfrac{1}{2}\mu_{lij}\, m((ij)/0) - \tfrac{1}{2}\mu_{lij}\, n((i)(j)/1) \\
&\quad - \tfrac{1}{2}\mu_{lijm}\, n((ij)(m)/0) + \pi(\theta_0)^{-1}\delta_i \pi(\theta_0)\mu_{li}] \\
&\quad - \mu^{-2}\mu_l[\mu_i\, m((i)/1) + \tfrac{1}{2}\mu_{ij}\, m((ij)/0) - \tfrac{1}{2}\mu_{ij}\, n((i)(j)/1) \\
&\quad - \tfrac{1}{2}\mu_{ijm}\, n((ij)(m)/0) + \pi(\theta_0)^{-1}\delta_i \pi(\theta_0)\mu_i] \\
&= I^{li}\, m((i)/1) + \tfrac{1}{2}(I^{li}u^j + I^{lj}u^i)\Delta_{i,j} \\
&\quad - \tfrac{1}{2}(I^{li}u^j u^m + I^{lj}u^i u^m + I^{lm}u^i u^j + [II]^{(lijm)})\, n((ij)(m)/0) \\
&\quad + \pi(\theta_0)^{-1}\delta_i \pi(\theta_0) I^{li}.
\end{aligned}
$$

Therefore, by (5.5) and (5.6), we obtain

$$
\begin{aligned}
g_1^l &= I^{li}\, m((i)/1) + [I^{-1}\Delta I^{-1}B]^l + (\tfrac{1}{2}I^{li}u^j u^m + \tfrac{1}{2}I^{li}I^{jm})N_{i,j,m} \\
&\quad + \pi(\theta_0)^{-1}\delta_i \pi(\theta_0) I^{li} \\
&= \tfrac{1}{2}[I^{-1}\Gamma]^l + [I^{-1}\Delta I^{-1}B]^l + \tfrac{1}{2}I^{li}u^j(Q_{i,j} - 2\Delta_{i,j}) \\
&\quad + \tfrac{1}{2}I^{li}I^{jm}N_{i,j,m} + \pi(\theta_0)^{-1}\delta_i \pi(\theta_0) I^{li} \\
&= \tfrac{1}{2}[I^{-1}\Gamma]^l + \tfrac{1}{2}[I^{-1}QI^{-1}B]^l + \tfrac{1}{2}[I^{-1}R]^l + \pi(\theta_0)^{-1}[I^{-1}\delta\pi(\theta_0)']^l.
\end{aligned}
$$

Since $\tilde{f}_0 = g_0$ and $\tilde{f}_1 = g_1 - \tilde{b}(\theta_0)$, we have completed the proof of Lemma 5.3. □

Lemma 5.4 (1) *For* $\theta_0 \in \Theta$ *and* $h \in \mathbf{R}^k$, *there exist functionals* ϕ_ε^h, $\varepsilon \in (0, 1)$, *on* $C([0, T], \mathbf{R}^d)$ *satisfying the following conditions for any compact set* $K \subset \mathbf{R}^k$.
(i) $0 \leq \phi_\varepsilon^h(X) \leq 1$, $X \in C([0, T], \mathbf{R}^d)$.
(ii) $\phi_\varepsilon^h(X^{\varepsilon,\theta_0}(w)) = 1 - O(\varepsilon^n)$ *in* D^∞ *as* $\varepsilon \downarrow 0$ *uniformly in* $h \in K$ *for* $n = 1, 2, \ldots$.
(iii) $\phi_\varepsilon^h(X^{\varepsilon,\theta_0+\varepsilon h}(w)) = 1 - O(\varepsilon^n)$ *in* D^∞ *as* $\varepsilon \downarrow 0$ *uniformly in* $h \in K$ *for* $n = 1, 2, \ldots$.
(iv) *For all* $p > 1$,

$$
\sup_{\substack{\varepsilon \in (0,1) \\ h \in K}} E[1_{\{\phi_\varepsilon^h(X^{\varepsilon,\theta_0}(w)) > 0\}} \exp\{p\, l_{\varepsilon,h}(w; \theta_0)\}] < \infty.
$$

(2) *Let* $\theta_0 \in \Theta$ *and let* K *be any compact set of* $\mathbf{R}^k$. *Then* $\phi_\varepsilon^h(X^{\varepsilon,\theta_0}) \exp\{l_{\varepsilon,h}(w; \theta_0)\}$ *has the asymptotic expansion*

$$
\phi_\varepsilon^h(X^{\varepsilon,\theta_0}) \exp\{l_{\varepsilon,h}(w; \theta_0)\} \sim e^{f_0^L}(1 + \varepsilon\Psi_1 + \varepsilon^2\Psi_2 + \ldots)
$$

in D^∞ *as* $\varepsilon\downarrow 0$ *with* $\Psi_1, \Psi_2, \ldots \in D^\infty$ *determined by the formal Tayler expansion*

$$\exp\{\varepsilon f_1^L + \varepsilon^2 f_2^L + \ldots\} = 1 + \varepsilon\Psi_1 + \varepsilon^2 \Psi_2 + \ldots.$$

This expansion is uniform in $h\in K$.

For the proof of this lemma see Lemma 4.5 of [22]. Let $A+h=\{z\in\mathbf{R}^k;\ z-h\in A\}$ for $A\subset\mathbf{R}^k$ and $h\in\mathbf{R}^k$.

Lemma 5.5 (1) *Define* $R_\varepsilon^{ij}(w)$ *by the equation*

$$\langle \varepsilon^{-1} D[\psi_\varepsilon(w)(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0]^i, \varepsilon^{-1} D[\psi_\varepsilon(w)(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0]^j\rangle_H = I^{ij} + \varepsilon R_\varepsilon^{ij}(w).$$

Let

$$\xi_\varepsilon(w) = c\varepsilon \sum_{i,j} |R_\varepsilon^{ij}(w)|^2$$

for $c>0$. *Then, for large* $c>0$, *the conditions* (1), (3), (4) *of Theorem* 4.1 *are satisfied for* $F_\varepsilon = \psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0)$ *and* $\xi_\varepsilon(w)$.
(2) *Let* $\theta_0\in\Theta$ *and let* K *be any compact set of* $\mathbf{R}^k$. *Then, for* $A\in\mathbf{B}^k$,

$$\begin{aligned}&\phi_\varepsilon^h(X^{\varepsilon,\theta_0}) \exp\{l_{\varepsilon,h}(w;\theta_0)\}\, \psi(\xi_\varepsilon)\, I_{A+h}\circ(\psi_\varepsilon(w)\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0))\\ &\quad\sim \Phi_{h,A+h,0} + \varepsilon\Phi_{h,A+h,1} + \ldots\end{aligned}$$

in $\tilde{D}^{-\infty}$ *as* $\varepsilon\downarrow 0$ *uniformly in* $A\in\mathbf{B}^k$ *and* $h\in K$ *with* $\Phi_{h,A+h,0}, \Phi_{h,A+h,1}, \ldots \in \tilde{D}^{-\infty}$. *In particular,*

$$\begin{aligned}\Phi_{h,A+h,0} &= \Lambda^0\,\Phi_{0,A+h,0},\\ \Phi_{h,A+h,1} &= \Lambda^0\,\Phi_{0,A+h,1} + \Lambda^0 f_1^L\,\Phi_{0,A+h,0},\end{aligned}$$

where

$$\begin{aligned}\Phi_{0,A+h,0} &= I_{A+h}(\tilde{f}_0),\\ \Phi_{0,A+h,1} &= \tilde{f}_1^i\,\partial_i I_{A+h}(\tilde{f}_0),\\ \Lambda^0 &= e^{f_0^L},\\ f_0^L &= h'B - \tfrac{1}{2}h' I(\theta_0)h,\\ f_1^L &= h^i m((i)/1) + \tfrac{1}{2} h^i h^j m((ij)/0)\\ &\quad - \tfrac{1}{2} h^i h^j n((i)(j)/1) - \tfrac{1}{2} h^i h^j h^m n((ij)(m)/0).\end{aligned}$$

Proof. (1) From Lemma 5.3 we have

$$\varepsilon^{-1} D[\psi_\varepsilon(w)(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0]^i \sim D\tilde{f}_0^i + \varepsilon D\tilde{f}_1^i + \ldots$$

in $D^\infty(H)$ as $\varepsilon\downarrow 0$ and hence

$$\begin{aligned}&\langle \varepsilon^{-1} D[\psi_\varepsilon(w)(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0]^i, \varepsilon^{-1} D[\psi_\varepsilon(w)(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0]^j\rangle_H\\ &\quad= \langle D\tilde{f}_0^i, D\tilde{f}_0^j\rangle_H + \varepsilon R_\varepsilon^{ij}(w),\end{aligned}$$

where

$$R_\varepsilon^{ij}(w) = \langle D\tilde{f}_0^i, D\tilde{f}_1^j\rangle_H + \langle D\tilde{f}_1^i, D\tilde{f}_0^j\rangle_H + O(\varepsilon)$$

in D^∞ as $\varepsilon\downarrow 0$. It is then easy to show (1), because $\langle D\tilde{f}_0^i, D\tilde{f}_0^j\rangle_H = I^{ij}$ and $\sup_\varepsilon E(|R_\varepsilon^{ij}(w)|^n)<\infty$ for any $n\in\mathbf{N}$.

(2) From Theorem 4.1, Lemmas 5.3, 5.4 and (1), we obtain (2). □

Let

$$A^*_{i,j,n}=\int_0^T\int_0^t [\partial_l\{\delta_i V_0'(VV')^+ V\}]_t^0(\theta_0)[V'(VV')^+\delta_j V_0]_t^0(\theta_0)[Y_t]^{lm}[g_s]^{mn}\,ds\,dt.$$

Lemma 5.6 *Let w be an r-dimensional Wiener process and let functions a_t, b_t, c_t on $[0,T]$ be deterministic. Let $\Sigma=\int_0^T a_t a_t'\,dt$. Then*

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

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

Here Tr *stands for the trace.*

(3)

$$E[m((i)/1)|\tilde{f}_0=x]=A^*_{i,j,l}(x^j x^l-I^{jl}),$$
$$E[m((j\,i)/0)|\tilde{f}_0=x]=B_{i,j,l}\,x^l$$

and

$$E[n((i)(j)/1)|\tilde{f}_0=x]=2A_{i,j,l}\,x^l.$$

(4)

$$\begin{aligned}E[\tilde{f}_1^i|\tilde{f}_0=x]=&-I^{im}A^*_{j,m,l}x^jx^l-\tfrac12 I^{im}B_{l,j,m}x^jx^l\\&-\tilde{b}(\theta_0)^i-I^{im}I^{jl}A^*_{m,j,l}+\pi(\theta_0)^{-1}I^{ij}\delta_j\pi(\theta_0)\\&-I^{im}I^{jl}B_{j,m,l}-\tfrac12 I^{im}I^{jl}B_{l,j,m}.\end{aligned}$$

(5)

$$\begin{aligned}E[f_1^L|\tilde{f}_0=x]=&A^*_{i,j,l}h^i(x^jx^l-I^{jl})+\tfrac12 B_{i,j,l}h^ih^jx^l\\&-A_{i,j,l}h^ih^jx^l-\tfrac12 B_{i,j,l}h^ih^jh^l.\end{aligned}$$

Proof. It is easy to show (1) and (2). Using (1) and (2) we obtain (3) and hence (4) and (5) in view of the definition of $\tilde{f}_1$ and f_1^L, respectively. □

Proof of Theorems 2.1 and 2.2 It suffices to show Theorem 2.2 only. Since $\psi(\xi_\varepsilon(w))=1-O(\varepsilon^n)$ in D^∞ as $\varepsilon\downarrow 0$ for any $n\in\mathbf{N}$, from Lemma 5.4 and Lemma 5.5 (2) we have

$$\begin{aligned}
&P(\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0+\varepsilon h)-\theta_0-\varepsilon h)\in A)\\
&\quad\sim E[\phi_\varepsilon^h(X^{\varepsilon,\theta_0+\varepsilon h})\,I_{A+h}(\psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0+\varepsilon h)-\theta_0))]\\
&\quad=E[\phi_\varepsilon^h(X^{\varepsilon,\theta_0})\exp\{l_{\varepsilon,h}(w;\theta_0)\}\,I_{A+h}(\psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0))]\\
&\quad\sim E[\phi_\varepsilon^h(X^{\varepsilon,\theta_0})\exp\{l_{\varepsilon,h}(w;\theta_0)\}\,\psi(\xi_\varepsilon)\,I_{A+h}(\psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0))]\\
&\quad=E[\phi_\varepsilon^h(X^{\varepsilon,\theta_0})\exp\{l_{\varepsilon,h}(w;\theta_0)\}\,\psi(\xi_\varepsilon)\,I_{A+h^\circ}(\psi_\varepsilon(w)\,\varepsilon^{-1}(\tilde{\theta}_\varepsilon^*(w;\theta_0)-\theta_0))]\\
&\quad\sim E[\Phi_{h,A+h,0}]+\varepsilon E[\Phi_{h,A+h,1}]+\dots
\end{aligned}$$

as $\varepsilon\downarrow 0$ uniformly in $A\in\mathbf{B}^k$ and $h\in K$. Each $\Phi_{h,A+h,i}$ is represented by a sum of the terms of the form $g\,\partial^{\mathbf{n}}\,I_{A+h}(\tilde{f}_0)$, $g\in D^\infty$. Hence, we see that

$$E[\Phi_{h,A+h,i}]=\int_A \tilde{p}_i^c(y)\,dy$$

for some smooth function $\tilde{p}_i^c(y)$ depending on h from the integration-by-parts formula in the Malliavin calculus. We shall find $\tilde{p}_0^c$ and $\tilde{p}_1^c$. First, we have by definition

$$\begin{aligned}
E[\Phi_{h,A+h,0}]&=E[A^0\,\Phi_{0,A+h,0}]=E[e^{f_0^L}\,I_{A+h}(\tilde{f}_0)]\\
&=\int_{A+h}\exp\{h'Ix-\tfrac{1}{2}h'Ih\}\,\phi(x;0,I^{-1})\,dx\\
&=\int_{A+h}\phi(x-h;0,I^{-1})\,dx\\
&=\int_A\phi(y;0,I^{-1})\,dy.
\end{aligned}$$

Therefore

$$\tilde{p}_0^c(y)=\phi(y;0,I^{-1}).$$

We note that for any Wiener functional $G\in D^\infty$ and any nondegenerate Wiener functional $f\in D^\infty(\mathbf{R}^k)$,

$$E[G\,\partial_i\,I_A(f)]=\int_A p(x)\,dx, \tag{5.7}$$

where

$$p(x)=-\partial_i\{E[G\,|\,f=x]\,p_f(x)\}.$$

Here $p_f(x)$ denotes the density of f. In fact, we know the existence of $p(x)$ satisfying (5.7) from the integration-by-parts formula. If we take $A=A_x=[x_1,\infty)\times\dots\times[x_k,\infty)$, we have

$$\begin{aligned}
p(x)&=(-1)^k\,\partial_1\dots\partial_k\,E[G\,\partial_i\,I_{A_x}(f)]\\
&=-\partial_i\,E[G\,\delta_x(f)]\\
&=-\partial_i\{E[G\,|\,f=x]\,p_f(x)\}.
\end{aligned}$$

From this fact, it follows that

$$E[\Lambda^0 \Phi_{0,A+h,1}] = E[e^{f_0^L} \tilde{f}_1^i \partial_i I_{A+h}(\tilde{f}_0)] = \int_{A+h} \tilde{p}^{c1}(x)\, dx,$$

where

$$\begin{aligned}\tilde{p}^{c1}(x) &= -\partial_i \{E[e^{f_0^L} \tilde{f}_1^i | \tilde{f}_0 = x]\, p_{\tilde{f}_0}(x)\} \\ &= -\partial_i \{\exp(h' I x - \tfrac{1}{2} h' I h)\, E[\tilde{f}_1^i | \tilde{f}_0 = x]\, p_{\tilde{f}_0}(x)\} \\ &= -\partial_i \{E[\tilde{f}_1^i | \tilde{f}_0 = x]\, \phi(x-h; 0, I^{-1})\}.\end{aligned}$$

By Lemma 5.6 (4), we obtain

$$\begin{aligned}(5.8)\quad \tilde{p}^{c1}(x) = \{&I^{im} A^*_{i,m,l} x^l + I^{im} A^*_{j,m,i} x^j + \tfrac{1}{2} I^{im} B_{l,i,m} x^l + \tfrac{1}{2} I^{im} B_{i,j,m} x^j \\ &+ (-I^{im} A^*_{j,m,l} x^j x^l - \tfrac{1}{2} I^{im} B_{l,j,m} x^j x^l - \tilde{b}(\theta_0)^i + \pi(\theta_0)^{-1} I^{ij} \delta_j \pi(\theta_0) \\ &- I^{im} I^{jl} A^*_{m,j,l} - I^{im} I^{jl} B_{j,m,l} \\ &- \tfrac{1}{2} I^{im} I^{jl} B_{l,j,m}) I_{ip}(x^p - h^p)\}\, \phi(x-h; 0, I^{-1}).\end{aligned}$$

On the other hand, by Lemma 5.6 (5), we have

$$E[\Lambda^0 f_1^L \Phi_{0,A+h,0}] = E[e^{f_0^L} f_1^L I_{A+h}(\tilde{f}_0)] = \int_{A+h} \tilde{p}^{c2}(x)\, dx,$$

where

$$\begin{aligned}(5.9)\qquad \tilde{p}^{c2}(x) &= E[f_1^L | \tilde{f}_0 = x]\, \phi(x-h; 0, I^{-1}) \\ &= \{A^*_{i,j,l} h^i (x^j x^l - I^{jl}) + \tfrac{1}{2} B_{i,j,l} h^i h^j x^l \\ &\quad - A_{i,j,l} h^i h^j x^l - \tfrac{1}{2} B_{i,j,l} h^i h^j h^l\}\, \phi(x-h; 0, I^{-1}).\end{aligned}$$

From (5.8), (5.9) and the facts that $A_{i,j,l} = \frac{1}{2}(A^*_{i,j,l} + A^*_{j,i,l})$ and that $B_{i,j,l} = B_{j,i,l}$, we obtain by tedious calculation

$$\begin{aligned}(5.10)\qquad &\tilde{p}^{c1}(x) + \tilde{p}^{c2}(x) \\ &= \{A_{i,j,l}[I^{ij} x^l - h^i h^j x^l + 2 h^i x^j x^l - x^i x^j x^l] \\ &\quad + B_{i,j,l}[-\tfrac{1}{2} I^{ij} x^l + \tfrac{1}{2} h^i h^j x^l \\ &\quad + \tfrac{1}{2} h^l x^i x^j - \tfrac{1}{2} x^i x^j x^l + I^{il} h^j + \tfrac{1}{2} I^{ij} h^l - \tfrac{1}{2} h^i h^j h^l] \\ &\quad - I_{ij} \tilde{b}(\theta_0)^i (x^j - h^j) + \pi(\theta_0)^{-1} \delta_i \pi(\theta_0)(x^i - h^i)\}\, \phi(x-h; 0, I^{-1}).\end{aligned}$$

Since

$$\begin{aligned}\int_A \tilde{p}_1^c(y)\, dy = E[\Phi_{h,A+h,1}]& \\ &= E[\Lambda^0 \Phi_{0,A+h,1}] + E[\Lambda^0 f_1^L \Phi_{0,A+h,0}] \\ &= \int_{A+h} [\tilde{p}^{c1}(x) + \tilde{p}^{c2}(x)]\, dx \\ &= \int_A [\tilde{p}^{c1}(y+h) + \tilde{p}^{c2}(y+h)]\, dy,\end{aligned}$$

it follows that

$$\tilde{p}_1^c(y) = \tilde{p}^{c1}(y+h) + \tilde{p}^{c2}(y+h).$$

Then we obtain the desired form of $\tilde{p}_1^c(y)$ by substituting $y+h$ for x in (5.10). $\square$

Acknowledgements. I am grateful to Prof. S. Watanabe, Dr. H. Sugita and Dr. H. Uemura for their helpful advice and valuable suggestions. Prof. S. Watanabe informed me of the ideas of truncation. I also wish to thank Dr. Y. Ogata and Dr. S. Konishi for discussion with them. Many thanks go to the editor, the associate editor and the referee for their careful reading of this manuscript and valuable comments.

References

1. Akahira, M., Takeuchi, K.: Asymptotic efficiency of statistical estimators: concepts and higher order asymptotic efficiency. (Lect. Notes Stat., vol. 7) Berlin Heidelberg New York: Springer 1981
2. Ibragimov, I.A., Has'minskii, R.Z.: Statistical estimation. Berlin Heidelberg New York: Springer 1981
3. Ikeda, N., Watanabe, S.: An introduction to Malliavin's calculus. In: Itô, K. (ed.) Stochastic analysis. Proc. Taniguchi Int. Symp. on Stochastic Analysis, Katata and Kyoto 1982, pp. 1–52. Tokyo: Kinokuniya and North-Holland 1984
4. Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes, 2nd edn. Tokyo: North-Holland and Kodansha 1989
5. Kusuoka, S., Stroock, D.: Applications of the Malliavin calculus, part I. In: Itô, K. (ed.) Stochastic analysis. Proc. Taniguchi Int. Symp. on Stochastic Analysis, Katata and Kyoto 1982, pp. 271–306. Tokyo: Kinokuniya and North-Holland 1984
6. Kutoyants, Yu.A.: Parameter estimation for stochastic processes. Prakasa Rao, B.L.S. (ed.) Berlin: Heldermann 1984
7. Liptser, R.S., Shiryayev, A.N.: Statistics of random processes I. Berlin Heidelberg New York: Springer 1977
8. Malliavin, P.: Stochastic calculus of variation and hypoelliptic operators. In: Itô, K. (ed.) Proc. Int. Symp. S.D.E. Kyoto 1976, pp. 195–263. Tokyo: Kinokuniya 1978
9. Malliavin, P.: C^k-hypoellipticity with degeneracy. In: Friedman, A., Pinsky, M. (eds.) Stochastic analysis, pp. 199–214, 327–340. New York: Academic Press 1978
10. Pfanzagl, J.: Contributions to a general asymptotic statistical theory. (Lect. Notes Stat., vol. 13) Berlin Heidelberg New York: Springer 1982
11. Pfanzagl, J.: Asymptotic expansions for general statistical models. (Lect. Notes Stat., vol. 31) Berlin Heidelberg New York: Springer 1985
12. Rao, C.R.: Asymptotic efficiency and limiting information. In: Proc. Fourth Berkeley Symp. on Math. Statist. Prob., vol. 1, pp. 531–545. Berkeley Los Angeles: University of California Press 1961
13. Rao, C.R.: Criteria of estimation in large samples. Sankhyā **25**, 189–206 (1963)
14. Swe, M., Taniguchi, M.: Higher order asymptotic properties of Bayes estimators for Gaussian ARMA processes. Technical Report No. 237. Hiroshima University 1989
15. Taniguchi, M.: On the second order asymptotic efficiency of estimators of Gaussian ARMA processes. Ann. Stat. **11**, 157–169 (1983)
16. Taniguchi, M.: Validity of Edgeworth expansions of minimum contrast estimators for Gaussian ARMA processes. J. Multivariate Anal. **21**, 1–28 (1987)
17. Taniguchi, M.: Higher order asymptotic theory for time series analysis. (Lect. Notes Stat., vol. 68) Berlin Heidelberg New York: Springer 1991
18. Watanabe, S.: Malliavin's calculus in terms of generalized Wiener functionals. In: Kallianpur, G. (ed.) Theory and application of random fields. (Lect. Notes Control Inf. Sci., vol. 49, pp. 284–290) Berlin Heidelberg New York: Springer 1983
19. Watanabe, S.: Lectures on stochastic differential equations and Malliavin calculus. Tata Institute of Fundamental Research. Berlin Heidelberg New York: Springer 1984
20. Watanabe, S.: Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. Ann. Probab. **15**, 1–39 (1987)
21. Yoshida, N.: Asymptotic expansion for small diffusion – an application of Malliavin-Watanabe theory. Research Memorandum, vol. 383. Institute of Statistical Mathematics 1990

22. Yoshida, N.: Asymptotic expansion for small diffusion. An application of the theory of Malliavin-Watanabe. General case and second order efficiency. Research Memorandum, vol. 392. Institute of Statistical Mathematics 1990
23. Yoshida, N.: Asymptotic expansion for statistics related to small diffusion II. Research Memorandum, vol. 406. Institute of Statistical Mathematics 1991