Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE @ DIRECT®

Stochastic Processes and their Applications 107 (2003) 53 – 81

stochastic processes and their applications

www.elsevier.com/locate/spa

Conditional expansions and their applications☆

Nakahiro Yoshida

Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153, Japan

Received 26 September 2001; received in revised form 25 February 2003; accepted 19 March 2003

Abstract

In the present article, we will consider a conditional limit theorem and conditional asymptotic expansions. Our discussion will be based on the Malliavin calculus. First, we treat a problem of lifting limit theorems to their conditional counterparts. Next, we provide asymptotic expansions in a general setting including the so-called small σ-models. In order to give a basis to the asymptotic expansion scheme for perturbed jump systems, we will build an extension to the Watanabe theory in part. Finally, we derive the asymptotic expansions (*double Edgeworth expansions*) of conditional expectations.

MSC: 60F99; 60H07; 60H35; 60J60; 60J75; 62E20; 62M20; 62P05

Keywords: Conditional distribution; Asymptotic expansion; Double Edgeworth expansion; Malliavin calculus; Perturbation; Watanabe theory; Hidden Markov model; Stochastic differential equation; Filtering

1. Introduction

The Malliavin calculus is nowadays recognized as an important instrument from a practical computational point of view in theoretical statistics, stochastic numerical analysis and mathematical finance as well as probability theory. It enables us to apply a usual differential calculus to *irregular* functionals, which very often appear, for example, as coverage probabilities, non-differentiable payoff functions, and so on.

The conditional expectation may be one of the most irregular functionals. For a continuously distributed conditioning variable, it requires the analysis over a null set.

☆ This is an extended version of Yoshida (2001), including generalized results and new examples. This work was in part supported by the Research Fund for Scientists of the Ministry of Science, Education and Culture, and by Cooperative Research Program of the Institute of Statistical Mathematics.

E-mail address: nakahiro@ms.u-tokyo.ac.jp (N. Yoshida).

0304-4149/03/$ - see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/S0304-4149(03)00046-2

Without doubt, the conditional stochastic calculus features in statistics: sufficient statistics in unbiased estimation and testing hypotheses (e.g., Lehmann-Scheffé theorem, Rao-Blackwell theorem, Neyman structure), conditional likelihood and conditional inference, conditionally Gaussian experiments as limits in LAMN situations, approximation formulas connected with the conditional distribution such as the p^* (magic) formula of Barndorff-Nielsen, filtering problems, recently introduced partial mixing, etc. In spite of the importance, conditional asymptotics does not seem to be so well founded as to fulfill the practical purpose.

In the present article, we will consider conditional limit theorems and conditional asymptotic expansions. Our discussion will be based on the Malliavin calculus, i.e., integration-by-parts (IBP) formulas, since it would be the most possible way to develop a theory applicable to functionals in practice, especially to stochastic differential equations. In Section 2, we will treat a problem of lifting limit theorems to their conditional counterparts. In non-ergodic statistical theory, a conditional limit law (i.e., a mixture of normal distributions or more generally a mixture of infinitely divisible laws) is usually deduced from the stable convergence of the limit theorems. It would be deeply related; however, it is *not* a conditional limit theorem. Indeed, previously, Prof. Sweeting (1986) showed his great skill to derive a conditional limit theorem for a branching process. So it may be a natural question when unconditional limit theorems can be lifted to conditional ones.

In the later sections, we will confine our attention to the so-called small σ-theory.

Section 5 provides, in a general setting, asymptotic expansions under small perturbations. The small σ-theory has been well developed in statistics. Kutoyants (1994) thoroughly investigated inference for diffusion-type processes with small noises. Asymptotic expansions were presented by Yoshida (1992a,b, 1993) by means of the Malliavin calculus and Prof. Watanabe's theory. See also Dermoune and Kutoyants (1995); Sakamoto and Yoshida (1996), Yoshida (1996), and Uchida and Yoshida (1999) for more statistical applications. As a byproduct, the asymptotic expansion scheme to compute the values of options was provided in Yoshida (1992b). There are many studies thereafter in this direction: Kunitomo and Takahashi (1998, 2001), Takahashi (1995, 1999), Kim and Kunitomo (1999), Sørensen and Yoshida (2000), Takahashi and Yoshida (2001), Kashiwakura and Yoshida (2001).

Recently, modeling with Lévy processes is attracting attention in financial statistics. In order to give a basis to the asymptotic expansion scheme for perturbed jump systems, we will in Section 5 build an extension to the Watanabe theory. We adopted the Malliavin calculus formulated by Bichteler et al. (1987). Differently from the original form of Watanabe's theory (Watanabe, 1987) for Wiener functionals (also see Watanabe, 1983; Ikeda and Watanabe, 1990), we do not use (have) Sobolev spaces of generalized functionals in our setting. For this reason, we will go through by the generalized integral operator for Schwartz distributions.

After preparing asymptotic expansions for generalized expectations, it is straightforward to obtain our main results. The asymptotic expansion of the conditional expectation will be derived in Section 6 together with several variants. They are called the double Edgeworth expansions. In the present article, we only treat most simple double expansions. We will present other variants (e.g., Edgeworth-saddlepoint approximation) elsewhere by applying Schilder-type expansions of densities (cf. Kusuoka and Stroock, 1991; Takanobu and Watanabe, 1993).

As for the role of the asymptotic expansion in the theoretical statistics, we refer the reader for example to Barndorff-Nielsen and Cox (1994), Ghosh (1994). An introduction to the Malliavin calculus from statistics is Yoshida (1999).

2. Lifting of limit theorems to conditional laws

Let $(\Omega,\mathscr{F},P)$ be a complete probability space. Let $F:\Omega\to\mathbb{R}^{d_1}$ be a d_1-dimensional random variable. For $k\in\mathbb{N}$, define $\mathscr{G}_k^F$ by

$$\begin{aligned}\mathscr{G}_k^F=\{\psi\in L^1;\ &\text{there exist functionals } \mathscr{I}^F_{(i_1,\dots,i_k)}(\psi)\in L^1(P)\\ &((i_1,\dots,i_k)\in\{1,\dots,d_1\}^k) \text{ such that}\\ &P[(\partial_{i_1}\cdots\partial_{i_k}f)\circ F\psi]=P[f\circ F\mathscr{I}^F_{(i_1,\dots,i_k)}(\psi)]\\ &\text{for any } f\in C_B^\infty(\mathbb{R}^{d_1}) \text{ and } (i_1,\dots,i_k)\in\{1,\dots,d_1\}^k\}.\end{aligned}$$

Lemma 1. *Suppose that there exists an integer $k>d_1$ such that $\cos(u\cdot Z)$ and $\sin(u\cdot Z)\in\mathscr{G}_k^F$ for all $u\in\mathbb{R}^d$, and that $\mathscr{I}^F_{(i_1,\dots,i_k)}(\cos(u\cdot Z))$ and $\mathscr{I}^F_{(i_1,\dots,i_k)}(\sin(u\cdot Z))$ are u-locally L^1-bounded. Then*

(a) *F has a continuous probability density function p^F with respect to the Lebesgue measure.*
(b) *For every $u\in\mathbb{R}^d$ and $x\in\mathbb{R}^{d_1}$, let*

$$\varphi_{Z/F}(u;x)=\frac{1}{(2\pi)^{d_1}}\int_{\mathbb{R}^{d_1}}\mathrm{e}^{-\mathrm{i}v\cdot x}P[\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]\,\mathrm{d}v,$$

then it is a $\mathbb{C}$-valued continuous function of (u,x), and

$$\varphi_{Z/F}(u;x)=P[\mathrm{e}^{\mathrm{i}u\cdot Z}|F=x]\,p^F(x)\,\mathrm{d}x\text{-a.e.}$$

Proof. Though this lemma is more or less well known, we shall give a proof for convenience of reference. The existence and the continuity of $\varphi_{Z/F}(u;x)$ follow from the assumption, since

$$\sup_{v\in\mathbb{R}^{d_1}}|iv_{i_1}\cdots iv_{i_{d_1+1}}P[\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]|<\infty$$

for every $u\in\mathbb{R}^d$ and every $(i_1,\dots,i_{d_1+1})\in\{1,\dots,d_1\}^{d_1+1}$. In particular, by Fourier inversion, we know that $p^F(x)=\varphi_{Z/F}(0;x)$ is a probability density of $\mathscr{L}\{F\}$. Set $h_u(x)=P[\mathrm{e}^{\mathrm{i}u\cdot Z}|F=x]\,p^F(x)$. Then $h_u\in L^1(\mathrm{d}x;\mathbb{C})$ and

$$P[\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]=\int_{\mathbb{R}^{d_1}}\mathrm{e}^{\mathrm{i}v\cdot x}h_u(x)\,\mathrm{d}x=\mathscr{F}[h_u](v).$$

Since the mapping $v \mapsto \mathscr{F}[h_u](v) = P[\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]$ is in $L^1(\mathrm{d}v;\mathbb{C})$, $h_u(x) = \mathscr{F}^{-1}\mathscr{F}[h_u](x) = \varphi_{Z/F}(u;x)\,\mathrm{d}x$-a.e. □

Remark 1. For $x \in S := \{x \in \mathbb{R}^{d_1};\, p^F(x) > 0\}$, define $\phi(u;x)$ by

$$\phi(u;x) = \frac{\varphi_{Z/F}(u;x)}{p^F(x)}.$$

For arbitrary $\alpha_1,\dots,\alpha_n \in \mathbb{C}$ and $u_1,\dots,u_n \in \mathbb{R}^d$,

$$\sum_{a,b=1}^{n} \alpha_a\bar{\alpha}_b\varphi_{Z/F}(u_a-u_b;x) = \frac{1}{(2\pi)^{d_1}}\int_{\mathbb{R}^{d_1}} \mathrm{e}^{-\mathrm{i}v\cdot x}P\left[\left|\sum_{a=1}^{n}\alpha_a\mathrm{e}^{\mathrm{i}u_a\cdot Z}\right|^2 \mathrm{e}^{\mathrm{i}v\cdot F}\right]\mathrm{d}v. \tag{1}$$

Since $|\sum_{a=1}^{n}\alpha_a\mathrm{e}^{\mathrm{i}u_a\cdot z}|^2$ is a sum of exponential functions of iz, it follows from our assumption that $P[\,\cdots]$ is integrable. Applying an elementary property in Fourier analysis, we see that

$$\sum_{a,b=1}^{n} \alpha_a\bar{\alpha}_b\varphi_{Z/F}(u_a-u_b;x) = P\left[\left|\sum_{a=1}^{n}\alpha_a\mathrm{e}^{\mathrm{i}u_a\cdot Z}\right|^2 \middle| F=x\right] p^F(x) \geqslant 0\,\mathrm{d}x\text{-a.e.} \tag{2}$$

The left-hand side is continuous in x because of (1). This together with (2) implies

$$\sum_{a,b=1}^{n} \alpha_a\bar{\alpha}_b\varphi_{Z/F}(u_a-u_b;x) \geqslant 0 \quad (\forall x \in \mathbb{R}^{d_1}).$$

Thus, $\phi(\cdot;x)$ has positivity for $x \in S$. Since $\phi(0;x) = 1$, $\phi(\cdot;x)$ is a characteristic function of some probability measure μ_x for each $x \in S$.[1] We define μ_x adequately on S^c, then μ_x is nothing but a regular conditional probability of Z given F. In the sequel, we choose a version of the conditional law $\mathscr{L}\{Z|F=x\}$ as

$$\mathscr{L}\{Z|F=x\} = \mu_x.$$

Thus, its characteristic function completely coincides with $\phi(u;x)$ for all u and x.

Let (Z_n) be a sequence of d-dimensional random variables and (F_n) be a sequence of d_1-dimensional random variables. Put

$$\mathscr{C}_n(u) = \{\cos(u\cdot Z_n), \sin(u\cdot Z_n)\}.$$

Theorem 1. *Let $k > d_1$. Suppose that the following conditions are satisfied*:

(i) $(Z_n, F_n) \to^d (Z, F)$ *as* $n \to \infty$.

[1] In place of the positivity, one may use a result on the limits of characteristic functions in the above passage.

(ii) $\mathscr{C}_n(u) \subset \mathscr{G}_k^{F_n}$ *for all* $u \in \mathbb{R}^d$, *and* $\{\mathscr{I}^{F_n}_{(i_1,\ldots,i_k)}(\psi);\ \psi \in \mathscr{C}_n(u),\ (i_1,\ldots,i_k) \in \{1,\ldots,d_1\}^k,\ u \in K,\ n \in \mathbb{N}\}$ *is bounded with respect to the* L^1*-norm for every compact set* $K \subset \mathbb{R}^d$.

Then (a) F_n *and* F *have continuous densities* p^{F_n} *and* p^F, *respectively, and*

$$p^{F_n}(x) \to p^F(x) \quad (n \to \infty)$$

for every $x \in \mathbb{R}^{d_1}$.

(b) *If* $p^F(x) > 0$, *then*

$$\mathscr{L}\{Z_n \,|\, F_n = x\} \Rightarrow \mathscr{L}\{Z \,|\, F = x\} \quad (n \to \infty).$$

Proof. By (ii), there exists a constant C such that

$$|P[\mathrm{e}^{\mathrm{i}v\cdot F_n}]| \leqslant \frac{C}{|v|^k}$$

for any $v \in \mathbb{R}^{d_1}$ and $n \in \mathbb{N}$. [2] Since $F_n \to^d F$, the representation of the density

$$p^{F_n}(x) = \frac{1}{(2\pi)^{d_1}} \int_{\mathbb{R}^{d_1}} \mathrm{e}^{-\mathrm{i}v\cdot x} P[\mathrm{e}^{\mathrm{i}v\cdot F_n}]\,\mathrm{d}v$$

yields

$$p^{F_n}(x) \to p^F(x) := \frac{1}{(2\pi)^{d_1}} \int_{\mathbb{R}^{d_1}} \mathrm{e}^{-\mathrm{i}v\cdot x} P[\mathrm{e}^{\mathrm{i}v\cdot F}]\,\mathrm{d}v$$

as $n \to \infty$, with the help of Lebesgue's theorem. In the same fashion, we see

$$\varphi_{Z_n/F_n}(u;x) \to \varphi_{Z/F}(u;x) \quad (n \to \infty).$$

Therefore, when $p^F(x) > 0$, the conditional characteristic function $\varphi_{Z_n/F_n}(u;x)/p^{F_n}(x)$ converges to $\varphi_{Z/F}(u;x)/p^F(x)$ for every $u \in \mathbb{R}^d$. $\square$

Example 1. If

$$Z_n \xrightarrow{d} Z \ (stably)$$

and if F is non-degenerate (in Malliavin's sense), then from Theorem 1, we obtain

$$\mathscr{L}\{Z_n \,|\, F = x\} \Rightarrow \mathscr{L}\{Z \,|\, F = x\}.$$

For example, if there are random variables Z_n, B, C on $(\Omega, \mathscr{F}, P)$ with $C > 0$ a.s., and if Z_n converges stably to $Z = C^{1/2}\zeta$ with ζ independent of $\mathscr{F}$, then

$$\mathscr{L}\{(B, Z_n) \,|\, F = x\} \Rightarrow \mathscr{L}\{(B, Z) \,|\, F = x\},$$

suppose that $F = f(B, C)$ is non-degenerate.

It is also possible to consider conditioning by F_n if (F_n) is uniformly non-degenerate.

[2] For this proof, it is sufficient to use (ii) for $(i, i, \ldots, i)$ among $(i_1, i_2, \ldots, i_k)$.

Genon-Catalot and Jacod (1993), and Jacod (1996) treated stable convergences for estimation of volatility parameters. This example is a model of the asymptotics in *non-ergodic* statistical inference. B corresponds to the *observed information* and C to the *energy* of the score.

Remark 2. Condition (ii) of Theorem 1 may seem to be difficult to verify for the reader unfamiliar with the Malliavin calculus. However, a sufficient condition for it is the uniform non-degeneracy of the Malliavin covariance of F_n, plus the boundedness of the Soborev norms of (Z_n, F_n). Those properties have been thoroughly investigated. See Ikeda and Watanabe (1990), Bichteler et al. (1987), Nualart (1995), and Malliavin (1997). The form of Condition (ii) is a minimal sufficient one for our use in this paper. It is easy to give a more smart (but restrictive) sufficient condition to (Z_n, F_n) on a certain probability space. For instance, if (Z_n, F_n) are functionals defined on a Wiener space, Condition (ii) follows from the non-degeneracy of the Malliavin covariance (i.e., the boundedness of L^p-norms of $\det \sigma_{F_n}^{-1}$) and the boundedness of the $\mathbf{D}_{s,p}$-norms $\|(Z_n, F_n)\|_{s,p}$. It is also the case for a Wiener–Poisson space. For functionals stemming from stochastic differential equations, the non-degeneracy is a consequence of non-degeneracy of either diffusion part or jump part, and the boundedness of norms comes from coefficients' regularity such as smoothness. Thus, for it is just an exercise, we shall not rephrase here those known sufficient conditions to give possible corollaries to our results. The same remark holds in Sections 5 and 6.

3. A class of smooth functionals and IBP

In the following sections, we will confine our attention to perturbed models and derive conditional asymptotic expansions (double expansions). First, we will extend Watanabe's methodology (Watanabe, 1987) to include jump-type processes. Let L be a Malliavin operator on a probability space $(\Omega, \mathscr{F}, P)$, cf. Bichteler et al. (1987). Let $\mathbf{D}_{2,p}$ be the completion of the domain $\mathscr{D}(L)$ by the norm

$$\|F\|_{2,p} = \|F\|_p + \|\Gamma(F,F)^{1/2}\|_p + \|LF\|_p$$

for $F \in \mathscr{D}(L)$, $p \geqslant 2$. Denote $\mathbf{D}_{2,\infty} = \bigcap_{p \geqslant 2} \mathbf{D}_{2,p}$.

Let $C_{\uparrow}^{\infty}(\mathbb{R}^d)$ denote the set of smooth functions $q : \mathbb{R}^d \to \mathbb{R}$ such that q and all derivatives $\partial^{\nu} q$ are of at most polynomial growth. We shall consider a linear space $\hat{\mathbf{D}} \subset \mathbf{D}_{2,\infty}$ such that

(i) If $F, G \in \hat{\mathbf{D}}$, then $L(FG) \in \hat{\mathbf{D}}$.
(ii) For any $\varphi \in C_{\uparrow}^{\infty}(\mathbb{R}^n)$ and $F_1, \ldots, F_n \in \hat{\mathbf{D}}$, $\varphi(F_1, \ldots, F_n) \in \hat{\mathbf{D}}$.

It is then easily seen that

(iii) $1 \in \hat{\mathbf{D}}$.
(iv) $F \in \hat{\mathbf{D}} \Rightarrow L^k F \in \hat{\mathbf{D}}$ $(\forall k \in \mathbb{N})$.
(v) $F, G \in \hat{\mathbf{D}} \Rightarrow FG \in \hat{\mathbf{D}}$ and $\Gamma(F, G) \in \hat{\mathbf{D}}$.

Also, (i) is obviously equivalent to

(vi) If $F \in \hat{\mathbf{D}}$, then $L(F^2) \in \hat{\mathbf{D}}$.

For example, in Wiener case, Watanabe's space $\mathbf{D}^\infty$ (cf. Watanabe, 1984; Ikeda and Watanabe, 1990) serves as $\hat{\mathbf{D}}$. In Wiener–Poisson case, we may take $\hat{\mathbf{D}}$ as $\hat{\mathbf{D}} = \bigcap_{m\in\mathbb{N}} \mathbf{D}_{2m,\infty}$, where algebras $\mathbf{D}_{2m,\infty}$ are inductively defined by $\mathbf{D}_{2(m+1),\infty} = \{F \in \mathbf{D}_{2m,\infty};\ LF \in \mathbf{D}_{2m,\infty}, \Gamma(F,F) \in \mathbf{D}_{2m,\infty}\}$. It is a routine job to verify that a given functional for a stochastic differential equation with jumps belongs to $\hat{\mathbf{D}}$.

Let $F \in \hat{\mathbf{D}}(\mathbb{R}^d) = (\hat{\mathbf{D}})^d$, $\xi \in \hat{\mathbf{D}}$. Then $\sigma_F \in \hat{\mathbf{D}}(\mathbb{R}^d \otimes \mathbb{R}^d)$, where $\sigma_F = (\Gamma(F^i, F^j))$, and $\Delta := \det\sigma_F \in \hat{\mathbf{D}}$. Let $\psi \in C^\infty(\mathbb{R};[0,1])$ such that $\psi(x) = 1$ for $|x| \leqslant 1/2$ and $\psi(x) = 0$ for $|x| \geqslant 1$. Identify σ_F with the set $\{\sigma_F^{ij}\}$, F with $\{F^i\}$, LF with $\{LF^i\}$, and so on.

For $S \subset \hat{\mathbf{D}}$, define $\mathscr{G}_i(S;F)$ $(i \in \mathbb{Z}_+)$ as follows:

$$\mathscr{G}_0(S;F) = \sigma_F \cup S,$$

$$\mathscr{G}_1(S;F) = LF \cup \mathscr{G}_0(S;F) \cup \Gamma(\mathscr{G}_0(S;F), F),$$

$$\mathscr{G}_i(S;F) = \mathscr{G}_{i-1} \cup \Gamma(\mathscr{G}_{i-1}(S;F), F) \ (i \geqslant 2).$$

Note that

$$\mathscr{G}_i(S;F) \subset \hat{\mathbf{D}}.$$

For example, if $S = \{\xi\}$, then

$$\mathscr{G}_0(\{\xi\};F) = \{\sigma_F^{ij}, \xi\},$$

$$\mathscr{G}_1(\{\xi\};F) = \{LF^i, \sigma_F^{ij}, \xi, \Gamma(\sigma_F^{ij}, F^k), \Gamma(\xi, F^k)\},$$

$$\begin{aligned}\mathscr{G}_2(\{\xi\};F) = \{&LF^i, \sigma_F^{ij}, \xi, \Gamma(\sigma_F^{ij}, F^k), \Gamma(\xi, F^k),\\ &\Gamma(LF^i, F^k), \Gamma(\sigma_F^{ij}, F^k), \Gamma(\xi, F^k), \Gamma(\Gamma(\sigma_F^{ij}, F^k), F^l), \Gamma(\Gamma(\xi, F^k), F^l)\}.\end{aligned}$$

The following IBP formula is known.

Proposition 1. *Suppose that $F \in \hat{\mathbf{D}}(\mathbb{R}^d)$ and $G, \xi \in \hat{\mathbf{D}}$. Let $\Delta = \det \sigma_F$. Suppose that*

$$P[1_{\{|\xi| \leqslant 1\}} \Delta^{-p}] < \infty$$

for every $p > 1$. Then, for any $f \in C_{\uparrow}^\infty(\mathbb{R}^d)$,

$$P[(\partial_{i_1} \cdots \partial_{i_k} f) \circ F \psi(\xi) G] = P[f \circ F \mathscr{I}^F_{(i_1,\ldots,i_k)}(G;\xi,\psi)].$$

The functional $\mathscr{I}^F_{(i_1,\ldots,i_k)}(G;\xi,\psi)$ has a representation:

$$\mathscr{I}^F_{(i_1,\ldots,i_k)}(G;\xi,\psi) = \sum_{j,l} \Delta^{-K(k)} \psi^{(l)}(\xi) G_{j,l} \quad (\textit{finite sum})$$

for some $K(k) \in \mathbb{N}$, where

$$G_{j,l} \in \mathscr{A}lg[\mathscr{G}_k(\{\xi, G\};F)]$$

(more precisely, $G_{j,l}$ is G-linear).

Proof. Extend the original probability space to the product space with a d-dimensional Wiener space. We attach the Ornstein-Uhlenbeck operator to this Wiener space. Then, L can be extended to this product space. If we replace F by $F_\varepsilon = F + \varepsilon W_1$, then the IBP formula is valid with Δ replaced by $\Delta_\varepsilon = \det(\sigma_F + \varepsilon^2 I_d)$. This new Malliavin covariance matrix is uniformly positive definite, so that we can obtain a similar representation of $\mathscr{I}^F_{(i_1,\dots,i_k)}(G;\xi,\psi)$ as the given one. Note that at this stage, it includes factors involving W_1. For example, $P \otimes W[\partial_i f(F_\varepsilon)\psi(\xi)G] = P \otimes W[f(F_\varepsilon)\mathscr{I}^{F_\varepsilon}_{(i)}(G;\xi,\psi)]$ with

$$\mathscr{I}^{F_\varepsilon}_{(i)}(G;\xi,\psi) = -\sum_{i'=1}^d \{2\Delta_\varepsilon^{-1}\psi(\xi)G\sigma^{F_\varepsilon}_{[i,i']}LF_\varepsilon^{i'} + \Gamma(\Delta_\varepsilon^{-1}\psi(\xi)G\sigma^{F_\varepsilon}_{[i,i']}, F_\varepsilon^{i'})\}.$$

Here $\sigma^F_{[i,i']}$ denotes the (i,i')-cofactor of σ_F. Higher-order IBP formulas can be written in a similar manner. Letting $\varepsilon \downarrow 0$ (it is possible, because of the non-degeneracy under truncation), we obtain the desired IBP formula and the representation. □

4. Generalized integral operator for Schwartz distributions

Define a second-order differential operator A by:

$$A = 1 + |x|^2 - \frac{1}{2}\sum_{i=1}^d \partial_i^2.$$

Let

$$p(t,x,y) = \Pi_{i=1}^d (2\pi(\sinh\sqrt{2}t)2^{-1/2})^{-1/2} \times \exp\left\{-\frac{\sqrt{2}}{2}(\coth\sqrt{2}t)[(x^i)^2 - 2x^i y^i \operatorname{sech}\sqrt{2}t + (y^i)^2]\right\}.$$

$\mathscr{S}$ denotes the Schwartz space and $\mathscr{S}'$ its dual. Let $m \in \mathbb{N}$. For $g \in \mathscr{S}'(\mathbb{R}^d)$ for which the function $t \mapsto t^{m-1}\mathrm{e}^{-t}_{\mathscr{S}'(\mathbb{R}^d)}\langle g, p(t,x,\cdot)\rangle_{\mathscr{S}(\mathbb{R}^d)}$ is absolutely integrable on $(0,\infty)$ for any $x \in \mathbb{R}^d$, the integral operator $A^{-m}: g \to A^{-m}g$ is defined by

$$A^{-m}g(x) = \int_0^\infty \frac{t^{m-1}\mathrm{e}^{-t}}{\Gamma(m)}\,{}_{\mathscr{S}'(\mathbb{R}^d)}\langle g, p(t,x,\cdot)\rangle_{\mathscr{S}(\mathbb{R}^d)}\,\mathrm{d}t.$$

See Ikeda and Watanabe (1989), or Sakamoto and Yoshida (1996) for details. It is known that

$$A^m A^{-m} = A^{-m}A^m = I_d \quad \text{on } \mathscr{S}(\mathbb{R}^d).$$

Let

$$\tilde{C}^{-2m}(\mathbb{R}^d) = \Big\{g \in \mathscr{S}'(\mathbb{R}^d);\ A^{-m}g \in \hat{C}(\mathbb{R}^d),\ \text{and there exists } g_n \in \mathscr{S}(\mathbb{R}^d \text{ such that } \lim_{n\to\infty}\|A^{-m}g_n - A^{-m}g\|_\infty = 0\Big\},$$

where $\hat{C}(\mathbb{R}^d)$ is the set of continuous functions $f(x)$ tending to zero when $|x| \to \infty$. Then it is also known that

$$\bigcup_{m \in \mathbb{N}} \tilde{C}^{-2m}(\mathbb{R}^d) = \mathscr{S}'(\mathbb{R}^d).$$

By Proposition 1, it is easy to obtain:

Proposition 2. *Suppose that* $F \in \hat{\mathbf{D}}(\mathbb{R}^d)$ *and* $G, \xi \in \hat{\mathbf{D}}$. *Let* $\Delta = \det \sigma_F$. *Suppose that*

$$P[1_{\{|\xi| \leqslant 1\}} \Delta^{-p}] < \infty \tag{3}$$

for every $p > 1$. *Then, for any* $f \in \mathscr{S}(\mathbb{R}^d)$,

$$P[A^m f \circ F \psi(\xi) G] = P[f \circ F \Psi^F_{2m}(G; \xi, \psi)].$$

The functional $\Psi^F_{2m}(G; \xi, \psi)$ *takes the form*:

$$\Psi^F_{2m}(G; \xi, \psi) = \sum_{j,l} \Delta^{-M(m)} \psi^{(l)}(\xi) \tilde{G}_{j,l} \quad (\textit{finite sum})$$

for some $M(m) \in \mathbb{N}$, *where*

$$\tilde{G}_{j,l} \in \mathscr{A}lg[\mathscr{G}_{2m}(\{F, \xi, G\}; F)]$$

(*more precisely*, $\tilde{G}_{j,l}$ *is* G*-linear*).

Notice that F appears in $\tilde{G}_{j,l}$ as $\{F, \xi, G\}$ because of the multiplication of F in the operator A.

Fix $g \in \mathscr{S}'(\mathbb{R}^d)$ arbitrarily. There exist a number $m \in \mathbb{Z}_+$ and a sequence $\tilde{g}_n \in \mathscr{S}(\mathbb{R}^d)$ ($n \in \mathbb{N}$) such that

$$\|A^{-m}\tilde{g}_n - A^{-m}g\|_\infty \to 0 \tag{4}$$

as $n \to \infty$. Then the sequences

$$\begin{aligned} P[\psi(\xi)\tilde{g}_n(F)G] &= P[A^{-m}\tilde{g}_n(F)\Psi^F_{2m}(G; \xi, \psi)] \\ &\to P[A^{-m}g(F)\Psi^F_{2m}(G; \xi, \psi)] \end{aligned}$$

as $n \to \infty$, for all $G \in \hat{\mathbf{D}}$, simultaneously. Clearly, this limit does not depend on the choice of the number m and the sequence $\tilde{g}_n$ satisfying (4), therefore, a linear functional $I_{(\psi(\xi)G,F)} : \tilde{C}_{-2m} \to \mathbb{R}$ is well defined by

$$I_{(\psi(\xi)G,F)}(g) = P[A^{-m}g(F)\Psi^F_{2m}(G; \xi, \psi)].$$

Because of the compatibility for different m's, this linear functional can be extended to the one over $\mathscr{S}'(\mathbb{R}^d)$. Intuitively speaking, we may regard $I_{(\psi(\xi)G,F)}(g)$ as "$P[\psi(\xi) g(F)G]$" for $g \in \mathscr{S}'(\mathbb{R}^d)$. But of course, this expectation itself in general does not make sense.

We denote by $\mathscr{F}_{\uparrow}(\mathbb{R}^d)$ the set of measurable functions on $\mathbb{R}^d$ of at most polynomial growth.

Proposition 3. *If* $F \in \hat{\mathbf{D}}(\mathbb{R}^d)$ *is non-degenerate (without truncation, i.e., (3) with* $\xi \equiv 0$*), then it has a smooth density* p^F *and for* $g \in \mathscr{F}_{\uparrow}(\mathbb{R}^d)$ *and* $G \in \hat{\mathbf{D}}$,

$$I_{(G,F)}(\partial^\nu g) = \int_{\mathbb{R}^d} g(z)(-\partial)^\nu (P[G|F=z]p^F(z))\,\mathrm{d}z \quad (\nu \in \mathbb{Z}_+^d).$$

The differentiability and the integrability in the above expression hold true.

Proof. It follows from a modification of Lemma 1 that the measure $P[G|F=x]P^F(\mathrm{d}x)$ has a density $\bar{p}$ given by:

$$\bar{p}(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}v\cdot x} P[G\mathrm{e}^{\mathrm{i}v\cdot F}]\,\mathrm{d}v.$$

From the fast decay of $P[G\mathrm{e}^{\mathrm{i}v\cdot F}]$ together with the IBP, we see that $\bar{p} \in \mathscr{S}(\mathbb{R}^d)$.

Now, by definition, for each $g \in \mathscr{S}'(\mathbb{R}^d)$ and $\nu \in \mathbb{Z}_+^d$, there exists an $m \in \mathbb{N}$ such that

$$I_{(G,F)}(\partial^\nu g) = P[A^{-m}\partial^\nu g(F)\Psi_{2m}^F(G;0,1)].$$

The same proof as that of Lemma 4 of Sakamoto and Yoshida (1996), taking sufficiently large m, shows that there exists a sequence $g_n \in \mathscr{S}(\mathbb{R}^d)$ for which

$$\|A^{-m}\partial^\nu g_n - A^{-m}\partial^\nu g\|_\infty \to 0$$

and simultaneously,

$$\|A^{-m} g_n - A^{-m} g\|_\infty \to 0$$

as $n \to \infty$. We will later use similar approximations in a few places. The proofs of them are elementary and omitted. But formulas convenient for proofs will be presented in Appendix. We see

$$\begin{aligned} I_{(G,F)}(\partial^\nu g) &\leftarrow P[A^{-m}\partial^\nu g_n(F)\Psi_{2m}^F(G;0,1)] \\ &= P[\partial^\nu g_n(F)G] \text{ (IBP-formula)} \\ &= \int_{\mathbb{R}^d} \partial^\nu g_n(x)\,\bar{p}(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} g_n(x)(-\partial)^\nu \bar{p}(x)\,\mathrm{d}x. \end{aligned}$$

Since $g_n \to g$ in $\tilde{C}_{-2m}$, $g_n \to g$ weakly in $\mathscr{S}'(\mathbb{R}^d)$. Due to the fact that $\bar{p} \in \mathscr{S}(\mathbb{R}^d)$, the last integral converges to

$$\int_{\mathbb{R}^d} g(x)(-\partial)^\nu \bar{p}(x)\,\mathrm{d}x,$$

and this completes the proof. □

In the sequel, we will use more convenient, and more intuitive notation than $I_{(\psi(\xi)G,F)}(g)$:

$$\mathbf{P}[\psi(\xi)g(F)G] = I_{(\psi(\xi)G,F)}(g) \quad (g \in \mathscr{S}'(\mathbb{R}^d)).$$

5. Asymptotic expansion

Define $\mathscr{K}_i(S)$ by

$$\mathscr{K}_0(S) = S,$$

$$\mathscr{K}_i(S) = \mathscr{K}_{i-1}(S) \cup L(S) \cup \Gamma(\mathscr{K}_{i-1}(S), \mathscr{K}_{i-1}(S)) \quad (i \geqslant 1).$$

Put

$$\mathscr{K}_\infty(S) = \bigcup_i \mathscr{K}_i(S).$$

Definition 1. (i) Let $S \subset \hat{\mathbf{D}}$. We say that $\mathscr{K}_\infty(S)$ is σ-finite if there exists an increasing sequence of subsets S_j in S such that $S = \bigcup_j S_j$ and $\mathscr{K}_i(S_j)$ is L_p-bounded for every $i, j, p \in \mathbb{N}$.

(ii) For a family of sequences $S_\varepsilon = (S_\varepsilon^{(1)}, S_\varepsilon^{(2)}, \ldots) \subset \hat{\mathbf{D}}$, we say that $\{\mathscr{K}_\infty(S_\varepsilon)\}_{\varepsilon \in (0,1]}$ is uniformly σ-finite if for every $n, p \in \mathbb{N}$, $\mathscr{K}_n(\{S_\varepsilon^{(1)}, \ldots, S_\varepsilon^{(n)}\})$ is L^p-bounded uniformly in ε.

(iii) We say that $(\xi_\varepsilon)_{\varepsilon \in (0,1]}$ is *uniformly bounded* in $\hat{\mathbf{D}}$ if for every $n, p \in \mathbb{N}$, the family $\{\mathscr{K}_n(\{\xi_\varepsilon\})\}_{\varepsilon \in (0,1]}$ is L^p-bounded.

For one sequence $S = (S^{(1)}, S^{(2)}, \ldots)$ (independent of ε), the σ-finiteness of $\mathscr{K}_\infty(S)$ is equivalent to the uniform σ-finiteness of $\{\mathscr{K}_\infty(S)\}$. Indeed, if S is σ-finite, for any $n \in \mathbb{N}$, there exists $m \in \mathbb{N}$ such that $\{S^{(1)}, \ldots, S^{(n)}\} \subset S_m$ and so $\mathscr{K}_n(\{S^{(1)}, \ldots, S^{(n)}\})$ is L^p-bounded (i.e., all elements are L^p-finite) for every $p \in \mathbb{N}$. Converse direction is obvious if one sets $S_j = \{S^{(1)}, \ldots, S^{(j)}\}$.

Define $r_k(\varepsilon)$ by

$$r_k(\varepsilon) = \varepsilon^{-k}(F_\varepsilon - [f_0 + \varepsilon f_1 + \cdots + \varepsilon^{k-1} f_{k-1}]).$$

Definition 2. Let $(\xi_\varepsilon)_{\varepsilon \in (0,1]} \subset \hat{\mathbf{D}}$. We say that $(F_\varepsilon)_{\varepsilon \in (0,1]}$ has *a smooth stochastic expansion associated with* ξ_ε

$$F_\varepsilon \sim f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \cdots \quad (\varepsilon \downarrow 0)$$

if $F_\varepsilon, f_0, f_1, \ldots \in \hat{\mathbf{D}}(\mathbb{R}^{d_0})$ and if $\{\mathscr{K}_\infty(S_\varepsilon)\}_{\varepsilon \in (0,1]}$ is uniformly σ-finite for

$$S_\varepsilon = (\xi_\varepsilon, F_\varepsilon, f_0, r_1(\varepsilon), f_1, r_2(\varepsilon), f_2, \ldots).$$

When $\xi_\varepsilon \equiv 1$, we simply say that $(F_\varepsilon)_{\varepsilon \in (0,1]}$ has *a smooth stochastic expansion.*

In spite of its apparent complexity, a smooth stochastic expansion is easy to derive and validate. A strong solution X_t^ε admits an expansion if the stochastic differential equation has smooth coefficients depending on ε smoothly. It is also the case for smooth functionals like a stochastic integral involving X_t^ε. See Remark 2 for references.

Theorem 2. *Suppose that*

(i) *the* $d_0 + d_2$*-dimensional sequence* $(F_\varepsilon, H_\varepsilon)_{\varepsilon\in(0,1]}$ *has a smooth stochastic expansion*

$$F_\varepsilon \sim f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \cdots,$$
$$H_\varepsilon \sim h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \cdots \quad (\varepsilon \downarrow 0).$$

(ii) $\limsup_{\varepsilon\downarrow 0} P[\Delta_{F_\varepsilon}^{-p}] < \infty$ *for every* $p > 1$.

Then, for every $g \in \mathscr{S}'(\mathbb{R}^{d_0})$ *and* $q \in C_\uparrow^\infty(\mathbb{R}^{d_2})$, $\mathbf{P}[g(F_\varepsilon)q(H_\varepsilon)]$ *has an (ordinary) asymptotic expansion*:

$$\mathbf{P}[g(F_\varepsilon)q(H_\varepsilon)] \sim \mathbf{P}[\Phi_0] + \varepsilon\mathbf{P}[\Phi_1] + \varepsilon^2\mathbf{P}[\Phi_2] + \cdots,$$

where Φ_i *are determined by the formal Taylor expansion of* $g(F_\varepsilon)q(H_\varepsilon)$ *around* (f_0, h_0). *In particular*,

$$\Phi_0 = g(f_0)q(h_0),$$
$$\Phi_1 = \sum_{a=1}^{d_0} \partial_a g(f_0) f_1^{(a)} q(h_0) + \sum_{b=1}^{d_2} \partial_b q(h_0) h_1^{(b)} g(f_0).$$

Here we dared to use d_0 for the dimension of F_ε, for later convenience. The following theorem is an extension of Theorem 2 to the case of non-degeneracy under truncation, cf. Takanobu and Watanabe (1993), Yoshida (1992b) for Wiener spaces. The truncation technique makes the proof of the non-degeneracy essentially easy; moreover, in many cases in statistics, the non-degeneracy does not hold without such localization.

Theorem 3. *Let* $\xi_\varepsilon \in \hat{\mathbf{D}}(\varepsilon \in (0,1])$. *Suppose that the following conditions are satisfied*:

(i) *The* $d_0 + d_2$*-dimensional sequence* $(F_\varepsilon, H_\varepsilon)_{\varepsilon\in(0,1]}$ *has a smooth stochastic expansion associated with* (ξ_ε):

$$F_\varepsilon \sim f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \cdots,$$
$$H_\varepsilon \sim h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \cdots \quad (\varepsilon \downarrow 0).$$

(ii) *For every* $p > 1$,

$$\limsup_{\varepsilon\downarrow 0} P[1_{\{|\xi_\varepsilon|\leqslant 1\}} \Delta_{F_\varepsilon}^{-p}] < \infty,$$

and for every $k\in\mathbb{N}$,

$$P\left[|\xi_\varepsilon|>\frac{1}{2}\right]=\mathrm{O}(\varepsilon^k).$$

Then, for every $g\in\mathscr{S}'(\mathbb{R}^{d_0})$ *and* $q\in C^\infty_\uparrow(\mathbb{R}^{d_2})$, *the generalized integral* $\mathbf{P}[\psi(\xi_\varepsilon)g(F_\varepsilon)q(H_\varepsilon)]$ *has an asymptotic expansion*:

$$\mathbf{P}[\psi(\xi_\varepsilon)g(F_\varepsilon)q(H_\varepsilon)]\sim\mathbf{P}[\Phi_0]+\varepsilon\mathbf{P}[\Phi_1]+\varepsilon^2\mathbf{P}[\Phi_2]+\cdots,$$

where Φ_i *are determined by the formal Taylor expansion of* $g(F_\varepsilon)q(H_\varepsilon)$ *around* (f_0,h_0) *as Theorem* 2.

Proof. Theorem 2 follows from Theorem 3 if we take $\xi_\varepsilon=0$, so we shall prove Theorem 3. Let $k\in\mathbb{Z}_+$. For $g\in\mathscr{S}'(\mathbb{R}^{d_0})$, there exists $m\in\mathbb{N}$ such that $A^{-m}g\in C_B^{k+1}(\mathbb{R}^{d_0})$. Here we use A with the dimension d_0. By definition and Taylor's formula, we obtain:

$$\begin{aligned}\mathbf{P}[\psi(\xi_\varepsilon)g(F_\varepsilon)q(H_\varepsilon)]&=P[A^{-m}g(F_\varepsilon)\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon);\xi_\varepsilon,\psi)]\\&=\sum_{v:|v|\leqslant k}P[(v!)^{-1}\partial^vA^{-m}g(f_0)[\varepsilon f_1+\cdots+\varepsilon^kf_k+\varepsilon^{k+1}r_{k+1}(\varepsilon)]^v\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon);\xi_\varepsilon,\psi)]\\&\quad+\sum_{v:|v|=k+1}P[R_{k+1,v}(\varepsilon)[\varepsilon f_1+\cdots+\varepsilon^kf_k+\varepsilon^{k+1}r_{k+1}(\varepsilon)]^v\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon);\xi_\varepsilon,\psi)],\end{aligned}$$

where

$$R_{k+1,v}(\varepsilon)=c(k,v)\int_0^1(1-\theta)^k\partial^vA^{-m}g((1-\theta)f_0+\theta F_\varepsilon)\,\mathrm{d}\theta.$$

We will show that $\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon);\xi_\varepsilon,\psi)$ has an asymptotic expansion in ε-power in L^p sense for every $p>1$. First,

$$\begin{aligned}\Gamma(F_\varepsilon,F_\varepsilon)&=\Gamma\left(\sum_{i=0}^k\varepsilon^if_i+\varepsilon^{k+1}r_{k+1}(\varepsilon),\sum_{i=0}^k\varepsilon^if_i+\varepsilon^{k+1}r_{k+1}(\varepsilon)\right)\\&=\Gamma(f_0,f_0)+\sum_{i=1}^k\varepsilon^i\Xi_i+\varepsilon^{k+1}R_{k+1}(\varepsilon),\end{aligned}$$

where $R_{k+1}(\varepsilon)$ is written out by $\Gamma(f_i^a,f_j^b)$, $\Gamma(f_i^a,r_{k+1}^b(\varepsilon))$ and $\Gamma(r_{k+1}^a(\varepsilon),r_{k+1}^b(\varepsilon))$. Ξ_i have similar expressions. By assumption, Ξ_i and $R_{k+1}(\varepsilon)$ are bounded in L^p uniformly in ε for every p. In particular, the truncation non-degeneracy condition for F_ε and Fatou's lemma, we see that the limit f_0 is non-degenerate without any truncation. $\psi^{(l)}(\xi_\varepsilon)=\delta_{l,0}-\mathrm{O}(\varepsilon^K)$ in L^p sense for every p and $K>0$, therefore, by applying

Taylor's formula to $1/x$, we obtain an L^p-expansion:

$$\Delta_{F_\varepsilon}^{-M(m)}\psi^{(l)}(\xi_\varepsilon) = \sum_{i=0}^{k} \varepsilon^i \gamma_i + \varepsilon^{k+1}\rho_{k+1}(\varepsilon),$$

$\gamma_i, \rho_{k+1}(\varepsilon)$ being uniformly L^p-bounded. In Proposition 2, we saw that $\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon); \xi_\varepsilon, \psi)$ takes the form:

$$\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon); \xi_\varepsilon, \psi) = \sum_{j,l} \Delta_{F_\varepsilon}^{-M(m)}\psi^{(l)}(\xi_\varepsilon)\tilde{G}_{j,l}(\varepsilon) \quad \text{(finite sum)} \tag{5}$$

and

$$\tilde{G}_{j,l}(\varepsilon) \in \mathscr{A}lg[\mathscr{G}_{2m}(\{F_\varepsilon, \xi_\varepsilon, q(H_\varepsilon)\}; F_\varepsilon)],$$

precisely, $\tilde{G}_{j,l}(\varepsilon)$ is $q(H_\varepsilon)$-linear. It is trivially seen that each $\tilde{G}_{j,l}(\varepsilon)$ is uniformly (in ε) L^p-bounded from the assumption. Therefore, the terms with $l \geqslant 1$ on the right-hand side of (5) are all $O(\varepsilon^{k+1})$ in L^p-sense. For $l = 0$, $\tilde{G}_{j,l}(\varepsilon)$ does not include "derivatives" of ξ_ε, and we may write

$$\tilde{G}_{j,0}(\varepsilon) \in \mathscr{A}lg[\mathscr{G}_{2m}(\{F_\varepsilon, q(H_\varepsilon)\}; F_\varepsilon)].$$

Thus, we can expand it by using expansion of $(F_\varepsilon, H_\varepsilon)$:

$$\tilde{G}_{j,0}(\varepsilon) = \sum_{i=0}^{k} \varepsilon^i \pi_{j,i} + \varepsilon^{k+1}\tilde{\pi}_{j,k+1}(\varepsilon)$$

with $\pi_{j,i}, \tilde{\pi}_{j,k+1}(\varepsilon)$ being uniformly bounded in L^p. Consequently, we obtain L^p-asymptotic expansion of $\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon); \xi_\varepsilon, \psi)$.

After expanding $\Psi_{2m}^{F_\varepsilon}(q(H_\varepsilon); \xi_\varepsilon, \psi)$ into an asymptotic expansion and rearranging terms, we see that $\mathbf{P}[\psi(\xi_\varepsilon)g(F_\varepsilon)q(H_\varepsilon)]$ has an asymptotic expansion:

$$\mathbf{P}[\psi(\xi_\varepsilon)g(F_\varepsilon)q(H_\varepsilon)] \sim c_0(g) + \varepsilon c_1(g) + \varepsilon^2 c_2(g) + \cdots \quad (\varepsilon \downarrow 0).$$

We note that each $c_i(g)$ takes the form of

$$c_i(g) = \sum_{v:|v| \leqslant k} P[\partial^v A^{-m} g(f_0) G_{i,v}] \tag{6}$$

for some $G_{i,v}$ described by $f_1, f_2, \ldots, f_k$, $h_1, \ldots, h_k$, $\partial^v q(h_0)$ ($|v| \leqslant k$) and their derivatives. We here used the fact that $\psi(\xi_\varepsilon) - 1$ and its derivatives (i.e., Γ's involving ξ_ε) are of $O(\varepsilon^{k+1})$ in L^p. For $\tilde{g} \in \mathscr{S}(\mathbb{R}^{d_0})$,

$$\mathbf{P}[\psi(\xi_\varepsilon)\tilde{g}(F_\varepsilon)q(H_\varepsilon)] \equiv P[\psi(\xi_\varepsilon)\tilde{g}(F_\varepsilon)q(H_\varepsilon)] \quad \text{(usual expectation)},$$

and $c_i(\tilde{g})$ is of the same form as $c_i(g)$.

If we expand $P[\psi(\xi_\varepsilon)\tilde{g}(F_\varepsilon)q(H_\varepsilon)]$ without taking the IBP formula, we obtain an asymptotic expansion

$$P[\psi(\xi_\varepsilon)\tilde{g}(F_\varepsilon)q(H_\varepsilon)] = \sum_{i=0}^{k} \varepsilon^i P[\Phi_i(\tilde{g},q)] + \mathrm{O}(\varepsilon^{k+1}).$$

Here each $\Phi_i(\tilde{g},q)$ takes the form:

$$\Phi_i(\tilde{g},q) = \sum_{v:|v|\leqslant k} \partial^v \tilde{g}(f_0)H_{i,v}$$

for some

$$H_{i,v} \in \mathscr{A}lg(f_0, f_1, \ldots, f_k, h_1, \ldots, h_k, \partial^v q(h_0) \quad (|v| \leqslant k)).$$

Consequently, we see that

$$\begin{aligned} c_i(\tilde{g}) &= P[\Phi_i(\tilde{g},q)] \\ &= \sum_{v:|v|\leqslant k} P[A^{-m}\partial^v \tilde{g}(f_0)\Psi_{2m}^{f_0}(H_{i,v};0,1)] \quad (\tilde{g} \in \mathscr{S}(\mathbb{R}^{d_0})). \end{aligned}$$

The last line is of course due to the IBP.

On the other hand, it is known that for every $g \in \mathscr{S}'(\mathbb{R}^{d_0})$ and $k \in \mathbb{N}$, if we take large m, then there exists a sequence $(g_n) \in \mathscr{S}(\mathbb{R}^{d_0})$ such that

$$\partial^v A^{-m} g_n \to \partial^v A^{-m} g \quad \text{in } C_B(\mathbb{R}^{d_0}) \quad (\forall v : |v| \leqslant k)$$

and simultaneously

$$A^{-m}\partial^v g_n \to A^{-m}\partial^v g \quad \text{in } C_B(\mathbb{R}^{d_0}) \quad (\forall v : |v| \leqslant k)$$

as $n \to \infty$ (cf. Sakamoto and Yoshida, 1996). Thus, $c_i(g_n) = P[\Phi_i(g_n,q)]$ converges and by definition,

$$\lim_{n\to\infty} c_i(g_n) = \mathbf{P}[\Phi_i(g,q)].$$

Furthermore, since $c_i(g_n) \to c_i(g)$ by (6), we finally obtain

$$c_i(g) = \mathbf{P}[\Phi_i(g,q)].$$

This was what we wanted to prove. $\square$

6. Double Edgeworth expansion

Theorem 4. *Let $x \in \mathbb{R}^{d_1}$ and let $g \in \mathscr{F}_\uparrow(\mathbb{R}^d)$ and $q \in C_\uparrow^\infty(\mathbb{R}^{d_2})$. Assume the following conditions*:

(i) $(Z_\varepsilon, F_\varepsilon, H_\varepsilon)_{\varepsilon\in(0,1]}$ *has a smooth stochastic expansion associated with* $(\xi_\varepsilon)_{\varepsilon\in(0,1]} \subset \hat{\mathbf{D}}$:

$$Z_\varepsilon \sim \zeta_0 + \varepsilon\zeta_1 + \varepsilon^2\zeta_2 + \cdots,$$

$$F_\varepsilon \sim f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \cdots \qquad (\varepsilon \downarrow 0),$$
$$H_\varepsilon \sim h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \cdots; \tag{7}$$

(ii) $\limsup_{\varepsilon\downarrow 0} P[\Delta_{F_\varepsilon}^{-p}] < \infty$ *for every* $p > 1$.

(iii) (α) $\limsup_{\varepsilon\downarrow 0} P[1_{\{|\xi_\varepsilon|\leqslant 1\}}\Delta_{(Z_\varepsilon,F_\varepsilon)}^{-p}] < \infty$ *for every* $p > 1$.

(β) $P[|\xi_\varepsilon| > \frac{1}{2}] = O(\varepsilon^k)$ $(\varepsilon \downarrow 0)$ *for every* $k \in \mathbb{N}$,

(γ) $P[(1-\psi(\xi_\varepsilon))g(Z_\varepsilon)q(H_\varepsilon) \,|\, F_\varepsilon = x] = O(\varepsilon^k)$ *for every* $k \in \mathbb{N}$.

Then, if (*a smooth version*) $p^{f_0}(x) > 0$ *at a point* x, *then*

$$P[g(Z_\varepsilon)q(H_\varepsilon)|F_\varepsilon = x] \sim c_0(x;g) + \varepsilon c_1(x;g) + \varepsilon^2 c_2(x;g) + \cdots \qquad (\varepsilon \downarrow 0). \tag{8}$$

In particular,

$$c_0(x;g) = P[g(\zeta_0)q(h_0)|f_0 = x],$$

$$\begin{aligned} c_1(x;g) = &\int_{\mathbb{R}^d} g(z)(-\partial_z)\cdot(P[q(h_0)\zeta_1|(\zeta_0,f_0) = (z,x)]p^{\zeta_0,f_0}(z,x))\,\mathrm{d}z/p^{f_0}(x) \\ &+ \int_{\mathbb{R}^d} g(z)(-\partial_x)\cdot(P[q(h_0)f_1|(\zeta_0,f_0) = (z,x)]p^{\zeta_0,f_0}(z,x))\,\mathrm{d}z/p^{f_0}(x) \\ &+ \int_{\mathbb{R}^d} g(z)P[\partial q(h_0)[h_1]|(\zeta_0,f_0) = (z,x)]p^{\zeta_0,f_0}(z,x)\,\mathrm{d}z/p^{f_0}(x) \\ &- P[g(\zeta_0)q(h_0)|f_0 = x]\ (-\partial_x)\cdot(P[f_1|f_0 = x]p^{f_0}(x))/p^{f_0}(x), \end{aligned} \tag{9}$$

where dot stands for the inner product (*i.e., divergence*), *and differentiability and integrability of appearing functions are implied by the assumptions. Asymptotic expansion* (8) *is uniformly valid on every compact set in* $\{x : p^{f_0}(x) > 0\}$ *if* (iii) (γ) *is uniform on it*.

Remark 3. Condition (iii) (γ) of the above theorem more precisely means that there exists a smooth (in x) version $P[\psi(\xi_\varepsilon)g(Z_\varepsilon)q(H_\varepsilon)|F_\varepsilon = x]$ under the assumptions, and for this version, we assume the existence of a version of $P[g(Z_\varepsilon)q(H_\varepsilon)|F_\varepsilon = x]$ for which

$$P[\psi(\xi_\varepsilon)g(Z_\varepsilon)q(H_\varepsilon)|F_\varepsilon = x] - P[g(Z_\varepsilon)q(H_\varepsilon)|F_\varepsilon = x] = O(\varepsilon^k)$$

for every $k \in \mathbb{N}$. Theorem 4 asserts the validity of (8) for such a nice version of $P[g(Z_\varepsilon)q(H_\varepsilon)|F_\varepsilon = x]$; any assertion would be meaningless unless selecting a nice version of the conditional expectation.

Remark 4. When Z_ε and F_ε are independent and $q \equiv 1$, the second and the last terms on the right-hand side of (9) cancel out, and also the third term vanishes; thus the expansion coincides with the unconditional expansion for $P[g(Z_\varepsilon)]$.

Remark 5. The principle of *double expansions* is very simple and various extensions are possible. For example, if p^{F_ε} admits a saddlepoint expansion, then we can immediately obtain an Edgeworth-saddlepoint expansion.

Lemma 2. *Let $x \in \mathbb{R}^{d_1}$ and let $g \in \mathscr{F}_\uparrow(\mathbb{R}^d)$. For g, take L sufficiently large, then there exists a sequence $(g_n) \in \mathscr{S}(\mathbb{R}^d)$ such that*

$$g_n \to g \quad \text{in } L^1((1+|z|^2)^{-L}\,\mathrm{d}z)$$

as $n \to \infty$, and that for some number m,

$$g_n \otimes \delta_x \to g \otimes \delta_x \quad \text{in } \tilde{C}^{-2m}(\mathbb{R}^{d+d_1}).$$

Indeed, it follows in a similar way as Lemma 4 of Sakamoto and Yoshida (1996, p. 51).

Proof of Theorem 4. (a) Let $Z \in \hat{\mathbf{D}}(\mathbb{R}^d)$ and $F \in \hat{\mathbf{D}}(\mathbb{R}^{d_1})$ such that F is non-degenerate without truncation and that (Z,F) is non-degenerate under truncation by $\psi(\xi) \in \hat{\mathbf{D}}$. Let $g \in \mathscr{F}_\uparrow(\mathbb{R}^d)$. Take $g_n \in \mathscr{S}(\mathbb{R}^d)$ as in Lemma 2. Then by definition, for $G \in \hat{\mathbf{D}}$,

$$I_{(\psi(\xi)G,(Z,F))}(g_n \otimes \delta_x) \to I_{(\psi(\xi)G,(Z,F))}(g \otimes \delta_x)$$

as $n \to \infty$.

We will show that for $g_n \in \mathscr{S}(\mathbb{R}^d)$,

$$I_{(\psi(\xi)G,(Z,F))}(g_n \otimes \delta_x) = P[\psi(\xi)Gg_n(Z)|F = x]\,p^F(x) \tag{10}$$

and that

$$P[\psi(\xi)Gg_n(Z)|F = x]\,p^F(x) \to P[\psi(\xi)Gg(Z)|F = x]\,p^F(x) \tag{11}$$

as $n \to \infty$. Once those relations are established, we obviously have

$$P[\psi(\xi)Gg(Z)|F = x]\,p^F(x) = I_{(\psi(\xi)G,(Z,F))}(g \otimes \delta_x).$$

Applying it to $\xi_\varepsilon, F_\varepsilon, Z_\varepsilon, H_\varepsilon$, one has

$$P[\psi(\xi_\varepsilon)q(H_\varepsilon)g(Z_\varepsilon)|F_\varepsilon = x] = \frac{I_{(\psi(\xi_\varepsilon)q(H_\varepsilon),(Z_\varepsilon,F_\varepsilon))}(g \otimes \delta_x)}{I_{(1,F_\varepsilon)}(\delta_x)},$$

and it is then not difficult to prove the theorem by expanding both numerator and denominator first with Theorems 3 and 2, and next by expanding the fractional expression, and finally by using Proposition 3 for each resulting term.

(b) We denote by p_1 the kernel "$p(t,x,y)$" of dimension d_1. Similarly, $\check{\Psi}_{2m}^{(Z,F)}$ denotes "$\Psi_{2m}^{\cdot}$"-operator for (Z,F) of dimension $d_0 = d + d_1$. Let $H \in \mathscr{S}(\mathbb{R}^{d_1})$. Denote

g_n by f. Then

$$\begin{aligned}
&\int I_{(\psi(\xi)G,(Z,F))}(f\otimes\delta_x)H(x)\,\mathrm{d}x\\
&\quad=\int P[A^{-m}(f\otimes\delta_x)(Z,F)\check{\Psi}_{2m}^{(Z,F)}(G;\xi,\psi)]\,\mathrm{d}x\\
&\quad=\int \mathrm{d}xH(x)P\left[\int_0^\infty \mathrm{d}t\Gamma(m)^{-1}\mathrm{e}^{-t}t^{m-1}\right.\\
&\qquad\left.\times\int p(t,Z,y')f(y')\,\mathrm{d}y'\,p_1(t,F,x)\check{\Psi}_{2m}^{(Z,F)}(G;\xi,\psi)\right]\\
&\quad=P\left[\int_0^\infty \mathrm{d}t\Gamma(m)^{-1}\mathrm{e}^{-t}t^{m-1}\right.\\
&\qquad\left.\times\int p(t,Z,y')f(y')\,\mathrm{d}y'\int p_1(t,F,x)H(x)\,\mathrm{d}x\check{\Psi}_{2m}^{(Z,F)}(G;\xi,\psi)\right]\\
&\quad=P[A^{-m}(f\otimes H)(Z,F)\check{\Psi}_{2m}^{(Z,F)}(G;\xi,\psi)]\\
&\quad=P[\psi(\xi)Gf(Z)H(F)]\\
&\quad=\int P[\psi(\xi)Gf(Z)|F=x]H(x)p^F(x)\,\mathrm{d}x.
\end{aligned}$$

Therefore,

$$I_{(\psi(\xi)G,(Z,F))}(f\otimes\delta_x)=P[\psi(\xi)Gf(Z)|F=x]p^F(x)$$

and we obtained (10).

(c) For a while, let us assume that g is a bounded measurable function with compact support. Define $q(z,x)$ by

$$q(z,x)=\frac{1}{(2\pi)^{d+d_1}}\iint \mathrm{e}^{-\mathrm{i}(u\cdot z+v\cdot x)}P[\psi(\xi)G\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]\,\mathrm{d}u\,\mathrm{d}v.$$

Under the assumption, $q(z,x)$ is well defined and all derivatives of q are integrable. We denote by $\vee$ the Fourier inversion. For $h\in C_K^\infty(\mathbb{R}^{d_1})$,

$$\begin{aligned}
&\int h(x)g(z)q(z,x)\,\mathrm{d}z\,\mathrm{d}x=\int \mathrm{d}z\int \mathrm{d}xh(x)g(z)\\
&\quad\times\frac{1}{(2\pi)^{d+d_1}}\iint \mathrm{e}^{-\mathrm{i}(u\cdot z+v\cdot x)}P[\psi(\xi)G\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]\,\mathrm{d}u\,\mathrm{d}v\\
&\quad=\iint (g\otimes h)^\vee(u,v)P[\psi(\xi)G\mathrm{e}^{\mathrm{i}u\cdot Z+\mathrm{i}v\cdot F}]\,\mathrm{d}u\,\mathrm{d}v
\end{aligned}$$

$$= \int\int (g\otimes h)^{\vee}(u,v)(P[\psi(\xi)G|(Z,F)=(z,x)]\cdot P^{(Z,F)})^{\wedge}(u,v)\,\mathrm{d}u\,\mathrm{d}v$$

$$= \int\int g(z)h(x)P[\psi(\xi)G|(Z,F)=(z,x)]P^{(Z,F)}(\mathrm{d}z,\mathrm{d}x).$$

Disintegrate $P^{(Z,F)}$ by the regular conditional probability $P^{Z/F}(\mathrm{d}z|x)$ of Z given F:

$$P^{(Z,F)}(\mathrm{d}z,\mathrm{d}x) = P^{Z/F}(\mathrm{d}z|x)p^F(x)\,\mathrm{d}x.$$

From the above equation,

$$\int g(z)P[\psi(\xi)G|(Z,F)=(z,x)]P^{Z/F}(\mathrm{d}z|x) = \int g(z)\frac{q(z,x)}{p^F(x)}\,\mathrm{d}z,$$

and hence, if $p^F(x)>0$, then

$$P[\psi(\xi)G|(Z,F)=(z,x)]P^{Z/F}(\mathrm{d}z|x) = \frac{q(z,x)}{p^F(x)}\,\mathrm{d}z$$

as measures. As a result, we see that for bounded measurable g,

$$\begin{aligned} P[\psi(\xi)Gg(Z)|F=x] &= P[P[\psi(\xi)G|Z,F]g(Z)|F=x] \\ &= \int P[\psi(\xi)G|(Z,F)=(z,x)]g(z)P^{Z/F}(\mathrm{d}z|x) \end{aligned}$$

and hence

$$P[\psi(\xi)Gg(Z)|F=x] = \int g(z)\frac{q(z,x)}{p^F(x)}\,\mathrm{d}z.$$

From the L^p-integrability of Z and $q(z,x)$ in z, the last relation holds for $g\in\mathscr{F}_{\uparrow}(\mathbb{R}^d)$. It follows from this and the integrability of $q(z,x)$ in z that

$$P[\psi(\xi)Gg_n(Z)|F=x] \to P[\psi(\xi)Gg(Z)|F=x]$$

if $g_n\in\mathscr{S}(\mathbb{R}^d)$ satisfy

$$g_n \to g \quad \text{in } L^1((1+|z|^2)^{-L}\,\mathrm{d}z),$$

which implies (11), and completes the proof. □

If the joint random vector $(Z_\varepsilon,F_\varepsilon)$ is completely non-degenerate, then we can obtain an expansion for Schwartz distributions $g\in\mathscr{S}'(\mathbb{R}^d)$. The proof of the following theorem is easier than Theorem 4, so omitted.

Theorem 5. *Let $x\in\mathbb{R}^{d_1}$. Assume the following conditions hold*:

(i) $(Z_\varepsilon,F_\varepsilon,H_\varepsilon)_{\varepsilon\in(0,1]}$ *has a smooth stochastic expansion* (7);

(ii) $\limsup_{\varepsilon\downarrow 0} P[\Delta^{-p}_{(Z_\varepsilon,F_\varepsilon)}]<\infty$ *for every* $p>1$.

Then, if (a smooth version) $p^{f_0}(x)>0$ at the point x, then for any $q\in C^\infty_\uparrow(\mathbb{R}^{d_2})$ and $g\in\mathscr{S}'(\mathbb{R}^d)$, the double Edgeworth expansion (8) *is valid suppose that* $P[g(Z_\varepsilon)$

$q(H_\varepsilon)|F_\varepsilon = x]$ is interpreted as $\mathbf{P}[(g\otimes\delta_x)(Z_\varepsilon,F_\varepsilon)q(H_\varepsilon)]/\mathbf{P}[\delta_x(F_\varepsilon)]$. The coefficients $c_i(x;g)$ have the same expressions as Theorem 4 if one interprets $\int_{\mathbb{R}^d} g(z)\phi(z)\,\mathrm{d}z$ as the coupling ${}_{\mathscr{S}'(\mathbb{R}^d)}\langle g,\phi\rangle_{\mathscr{S}(\mathbb{R}^d)}$ and $P[g(\zeta_0)q(h_0)|f_0 = x]$ as $\mathbf{P}[g\otimes\delta_x(\zeta_0,f_0)q(h_0)]/\mathbf{P}[\delta_x(f_0)]$.

If g is smooth function, we do not need non-degeneracy condition for Z_ε, that is, we easily obtain the following result.

Theorem 6. *Let $x\in\mathbb{R}^{d_1}$. Assume the following conditions hold*:

(i) $(Z_\varepsilon,F_\varepsilon,H_\varepsilon)_{\varepsilon\in(0,1]}$ *has a smooth stochastic expansion* (7);

(ii) $\limsup_{\varepsilon\downarrow 0} P[\Delta_{F_\varepsilon}^{-p}]<\infty$ *for every* $p>1$.

Then, if (a smooth version) $p^{f_0}(x)>0$ at the point x, then for any $q\in C_\uparrow^\infty(\mathbb{R}^{d_2})$ and $g\in C_\uparrow^\infty(\mathbb{R}^d)$, the double Edgeworth expansion (8) is valid. The coefficients $c_i(x;g)$ $(i=0,1)$ are in particular given by

$$c_0(x;g) = P[g(\zeta_0)q(h_0)|f_0 = x],$$

$$\begin{aligned} c_1(x;g) = {} & \{(-\partial_x)\cdot(P[g(\zeta_0)q(h_0)f_1|f_0 = x]p^{f_0}(x))\}/p^{f_0}(x) \\ & + P[\partial(gq)(\zeta_0,h_0)[(\zeta_1,h_1)]|f_0 = x] \\ & - P[g(\zeta_0)q(h_0)|f_0 = x]\ \{(-\partial_x)\cdot(P[f_1|f_0 = x]p^{f_0}(x))\}/p^{f_0}(x). \end{aligned}$$

Proof. We take $(Z_\varepsilon,H_\varepsilon)$ for H_ε in Theorem 4 and prepare a new variable $Z_\varepsilon = Z_0$ which is non-degenerate and independent of all given random variables. For the proof, it suffices to apply Theorem 4 with $\xi_\varepsilon = 0$. □

7. Examples

Illustrative simple examples will be presented in this section.

Example 2. Consider a system of stochastic differential equations

$$\mathrm{d}X_t^\varepsilon = A_0(\theta_{t-}^\varepsilon)\,\mathrm{d}t + A_1(\theta_{t-}^\varepsilon)\,\mathrm{d}L_t^* + A_2(\theta_{t-}^\varepsilon)\,\mathrm{d}L_t^\dagger,$$

$$X_0^\varepsilon = x_0$$

and

$$\mathrm{d}Y_t^\varepsilon = B_0(\theta_{t-}^\varepsilon)\,\mathrm{d}t + B_1(\theta_{t-}^\varepsilon)\,\mathrm{d}L_t^* + B_2(\theta_{t-}^\varepsilon)\,\mathrm{d}L_t^\dagger,$$

$$Y_0^\varepsilon = y_0,$$

where L^* and $L^\dagger$ are independent Lévy processes with nice regular distributions. The process θ_t^ε is a hidden (Markov) process, and we assume that it satisfies a stochastic differential equation:

$$d\theta_t^\varepsilon = C_0(\theta_t^\varepsilon)\,dt + \varepsilon C(\theta_{t-}^\varepsilon)\,dL_t,$$

$$\theta_0^\varepsilon = \vartheta_0,$$

where L_t is a Lévy process independent of L^* and $L^\dagger$, for simplicity. We consider the Malliavin operators corresponding to the shifts of L^* and $L^\dagger$. Then under non-degeneracy of A and B, $Z_\varepsilon = X_1^\varepsilon$ and $F_\varepsilon = Y_1^\varepsilon$ are non-degenerate. It is possible to treat more complicated, non-linear stochastic differential equations with jumps. In such cases, the sufficient conditions for non-degeneracy presented by Bichteler et al. (1987) are useful.

Computations of coefficients would be not so complicated. If noises are Wiener, then there is no problem: there are formulas for the conditional expectation of a multiple Wiener integral given a Wiener integral (Yoshida, 1992a; Kunitomo and Takahashi, 1998; Takahashi, 1995, 1999). If the conditioning variables are written out by Wiener integrals, then computations will not be difficult.

Only as an illustration, let us consider a simple case. In particular, the case where A_i and B_i $(i = 1, 2)$ are constants is specially easy. In this case,

$$\zeta_0 = X_1^0 = \int_0^1 A_0(\theta_t^0)\,dt + A_1 L_1^* + A_2 L_1^\dagger,$$

$$f_0 = Y_1^0 = \int_0^1 B_0(\theta_t^0)\,dt + B_1 L_1^* + B_2 L_1^\dagger$$

and the derivatives $X_t^{(1)} = (\partial_\varepsilon)_0 X_t^\varepsilon$ and $Y_t^{(1)} = (\partial_\varepsilon)_0 Y_t^\varepsilon$ are determined by

$$\zeta_1 = X_1^{(1)} = \int_0^1 \partial A_0(\theta_t^0)\theta_t^{(1)}\,dt,$$

$$f_1 = Y_1^{(1)} = \int_0^1 \partial B_0(\theta_t^0)\theta_t^{(1)}\,dt,$$

where θ_t^0 is a deterministic process satisfying the ordinary differential equation

$$d\theta_t^0 = C_0(\theta_t^0)\,dt,$$

$$\theta_0^0 = \vartheta_0$$

and $\theta_t^{(1)} = (\partial_\varepsilon)_0\theta_t^\varepsilon$ is determined by

$$d\theta_t^{(1)} = \partial C_0(\theta_t^0)\ \theta_t^{(1)}\,dt + C(\theta_t^0)\,dL_t,$$

$$\theta_0^{(1)} = 0.$$

Independence yields

$$P[\zeta_1|(\zeta_0, f_0) = (z, x)] = P[\zeta_1] = \int_0^1 \partial A_0(\theta_t^0)P[\theta_t^{(1)}]\,dt,$$

$$P[f_1|(\zeta_0, f_0) = (z,x)] = P[f_1|f_0 = x] = P[f_1] = \int_0^1 \partial B_0(\theta_t^0)P[\theta_t^{(1)}]\,\mathrm{d}t,$$

which simplify computations in the second order. The deterministic process $P[\theta_t^{(1)}]$ is a solution of an ordinary differential equation, and non-zero in general unless the mean of L equals zero.

Finally, in order to compare the asymptotic expansion scheme with Monte-Carlo simulation, we shall reduce the model to a simplest one, namely,

$$A_0(\theta) = A\theta, \quad A_1(\theta) = A_1, \quad A_2(\theta) = A_2,$$

$$B_0(\theta) = B\theta, \quad B_1(\theta) = B_1, \quad B_2(\theta) = B_2,$$

$$C_0(\theta) = 0, \quad C(\theta) = 1,$$

and let $L^*, L^\dagger, W$ be independent standard Wiener processes, and take $L_t = C(mt + W_t)$. Here is an outcome of a numerical study among several studies.[3] We chose the parameter values as follows: $C=2.0$, $A=1.3$, $A_1=0.5$, $A_2=1.0$, $B=-0.5$, $B_1=1.0$, $B_2=0.8$, $m = 0.5$, $x = 0.05$, $\varepsilon = 0.1$. For $g(z) = z$, the closed form of $\mathbf{V} := E[g(Z_\varepsilon)|F_\varepsilon = x]$ is given by

$$\mathbf{V} = \frac{A_1B_1 + A_2B_2 + \varepsilon^2 ABC^2/3}{B_1^2 + B_2^2 + \varepsilon^2 B^2C^2/3}\left(x - \varepsilon\frac{mBC}{2}\right) + \varepsilon\frac{mAC}{2},$$

and for parameters given above, $\mathbf{V} = 0.123935$.

The time interval $[0,1]$ was divided into 1000 subintervals of equal length to generate approximate stochastic processes. We set a window of length 0.05 including the point x, and the simulated F_ε's hit this interval 15 475 times among 1000000 repetitions. The estimated value of $\mathbf{V}$ by the Monte-Carlo simulation based on those 15 475 hits was 0.128308. The Monte-Carlo simulation with 2 GHz consumed 25 minutes.

On the other hand, we obtained

$$c_0(x;g) = \frac{A_1B_1 + A_2B_2}{B_1^2 + B_2^2}\,x = 0.039634,$$

$$\varepsilon c_1(x;g) = \varepsilon\frac{mC}{2}\left[A - B\frac{A_1B_1 + A_2B_2}{B_1^2 + B_2^2}\right] = 0.084817,$$

therefore the estimated value of $\mathbf{V}$ by the asymptotic expansion up to the second-order was 0.124451. This value was better than the Monte-Carlo estimate. It may be said that the second-order term fairly improved the accuracy, compared with the first-order term. The asymptotic expansion scheme has an advantage that we can obtain a new estimate immediately even when coefficients of the equations are changed. Multi-dimensional conditioning will make the difference between them clearer. According to Masuda's studies of Lévy cases with jumps and of multi-dimensional conditioning cases, our approach achieved satisfactory precision and reduced computational time, for example,

[3] I owe the numerical study and the comparison with the true value to Mr. Hiroki Masuda.

from about 300 min for each by Monte-Carlo methods on 2 GHz PC (10000000 repetitions for comparable precision) to almost zero second.

Example 3. Let us consider a system of $D+D_1$ stochastic differential equations

$$\mathrm{d}X_t^\varepsilon = \sum_{\alpha=0}^{r} V_\alpha(X_{t-}^\varepsilon, Y_{t-}^\varepsilon, \varepsilon)\,\mathrm{d}L_t^\alpha, \qquad X_0^\varepsilon = x_0,$$

$$\mathrm{d}Y_t^\varepsilon = \sum_{\alpha=0}^{r} V'_\alpha(X_{t-}^\varepsilon, Y_{t-}^\varepsilon, \varepsilon)\,\mathrm{d}L_t^\alpha, \qquad Y_0^\varepsilon = y_0, \tag{12}$$

where $L_t^0 = t$ and $L_t^1,\dots,L_t^r$ are Lévy processes whose Lévy measures have moments of any order on a region apart from the origin. ε is parameter in $(0,1]$, and x_0 and y_0 are constants independent of ε: x_0 is here observable only for simplicity. $L^1,\dots,L^r$ may be dependent or have non-zero means. Time-dependent equations are included in this model if one takes t as an argument. Moreover, since coefficients can degenerate, noises driving X^ε and Y^ε can be independent, as it is rather often assumed in practical situations. Suppose that $V_\alpha \in C_b^\infty(\mathbb{R}^{D+D_1}\times(0,1];\mathbb{R}^D)$ and $V'_\alpha \in C_b^\infty(\mathbb{R}^{D+D_1}\times(0,1];\mathbb{R}^{D_1})$. Here C_b^∞ stands for the set of smooth functions such that all derivatives of order $\geqslant 1$ are bounded.

In this example, V_α and V'_α are general non-linear functions; however, we will put the deterministic limit condition:

$$V_\alpha(\cdot,\cdot,0) = 0 \quad \text{and} \quad V'_\alpha(\cdot,\cdot,0) = 0 \quad (\alpha = 1,\dots,r).$$

Model (12) forms a filtering model with a system process X^ε and an observation process Y^ε. Several authors dealt with filtering problems in small diffusion settings (cf. Picard, 1991, Zeitouni, 1988). Recently, Del Moral et al. (2001) studied filtering with discrete-time observations and presented error bounds for certain Monte-Carlo filtering schemes. Here we will view the filtering problem based on discrete-time observations or more generally finite dimensional functionals from a small-σ-theoretical aspect.

For fixed maturity $T>0$, let $\mathbf{T} = [0,T]$. Let $\mathbf{t} = (t_1,\dots,t_m) \in \mathbf{T}^m$ and put $X_{\mathbf{t}} = (X_{t_1},\dots,X_{t_m})$. Similarly, $Y_{\mathbf{s}} = (Y_{s_1},\dots,Y_{s_n})$ for $\mathbf{s} = (s_1,\dots,s_n) \in \mathbf{T}^n$. Among many possibilities, we here consider the following two functionals:

$$\check{Z}(X.) = \left(\int_{\mathbf{T}^m} Z_l(X_{\mathbf{t}})\mu_l(\mathrm{d}\mathbf{t})\right)_{l=1,\dots,d}$$

and

$$\check{F}(Y.) = \left(\int_{\mathbf{T}^n} F_k(Y_{\mathbf{s}})\nu_k(\mathrm{d}\mathbf{s})\right)_{k=1,\dots,d_1},$$

where $Z_l \in C_\uparrow^\infty(\mathbb{R}^{Dm})$, $F_k \in C_\uparrow^\infty(\mathbb{R}^{D_1 n})$, μ_l are finite signed-measures on $\mathbf{T}^m$, and ν_k are finite signed-measures on $\mathbf{T}^n$. Functionals $\check{Z}$ and $\check{F}$ are very simple, but they include, e.g., finite samples $(Y_{s_1^0},\dots,Y_{s_n^0})$, anticipative integrals like $\int_0^{T/2} X_t X_{T-t}\,\mathrm{d}t$, and so on. Our results can apply if the functional admits a von Mises-type expansion the residual term of which has a representation compatible with L-operations.

Let

$$Z^\varepsilon = \varepsilon^{-1}(\check{Z}(X^\varepsilon_\cdot) - \check{Z}(X^0_\cdot)) = \left(\int_{\mathbf{T}^m} \varepsilon^{-1}(Z_l(X^\varepsilon_{\mathbf{t}}) - Z_l(X^0_{\mathbf{t}}))\mu_l(\mathrm{d}\mathbf{t})\right)_{l=1,\dots,d}$$

and

$$F^\varepsilon = \varepsilon^{-1}(\check{F}(Y^\varepsilon_\cdot) - \check{F}(Y^0_\cdot)) = \left(\int_{\mathbf{T}^n} \varepsilon^{-1}(F_k(Y^\varepsilon_{\mathbf{s}}) - F_k(Y^0_{\mathbf{s}}))\nu_k(\mathrm{d}\mathbf{s})\right)_{k=1,\dots,d_1}.$$

It is then easy to show that Z^ε and F^ε admit smooth stochastic expansions $Z^\varepsilon \sim \zeta_0 + \varepsilon\zeta_1 + \cdots$ and $F^\varepsilon \sim f_0 + \varepsilon f_1 + \cdots$ with

$$\zeta_0 = \left(\int_{\mathbf{T}^m} \partial Z_l(X^0_{\mathbf{t}})[X^{[1]}_{\mathbf{t}}]\mu_l(\mathrm{d}\mathbf{t})\right)_{l=1,\dots,d},$$

$$\zeta_1 = \left(\int_{\mathbf{T}^m} \frac{1}{2}\{\partial^2 Z_l(X^0_{\mathbf{t}})[(X^{[1]}_{\mathbf{t}})^{\otimes 2}] + \partial Z_l(X^0_{\mathbf{t}})[X^{[2]}_{\mathbf{t}}]\}\mu_l(\mathrm{d}\mathbf{t})\right)_{l=1,\dots,d},$$

$$f_0 = \left(\int_{\mathbf{T}^n} \partial F_k(Y^0_{\mathbf{s}})[Y^{[1]}_{\mathbf{s}}]\nu_k(\mathrm{d}\mathbf{s})\right)_{k=1,\dots,d_1},$$

$$f_1 = \left(\int_{\mathbf{T}^n} \frac{1}{2}\{\partial^2 F_k(Y^0_{\mathbf{s}})[(Y^{[1]}_{\mathbf{s}})^{\otimes 2}] + \partial F_k(Y^0_{\mathbf{t}})[Y^{[2]}_{\mathbf{s}}]\}\nu_k(\mathrm{d}\mathbf{s})\right)_{k=1,\dots,d_1}.$$

Here the component limit processes and derivatives X^0_t, Y^0_t, $X^{[1]}_t$, $Y^{[1]}_t$, $X^{[2]}_t$, $Y^{[2]}_t$ are determined by

$$\mathrm{d}X^0_t = V_0(X^0_t, Y^0_t, 0)\,\mathrm{d}t, \quad X^0_0 = x_0,$$

$$\mathrm{d}Y^0_t = V'_0(X^0_t, Y^0_t, 0)\,\mathrm{d}t, \quad Y^0_0 = x_0,$$

$$\mathrm{d}X^{[1]}_t = \partial V_0(X^0_t, Y^0_t, 0)[(X^{[1]}_t, Y^{[1]}_t, 1)]\,\mathrm{d}t + \sum_{\alpha=1}^r \partial_\varepsilon V_\alpha(X^0_t, Y^0_t, 0)\,\mathrm{d}L^\alpha_t, \quad X^{[1]}_0 = 0,$$

$$\mathrm{d}Y^{[1]}_t = \partial V'_0(X^0_t, Y^0_t, 0)[(X^{[1]}_t, Y^{[1]}_t, 1)]\,\mathrm{d}t + \sum_{\alpha=1}^r \partial_\varepsilon V'_\alpha(X^0_t, Y^0_t, 0)\,\mathrm{d}L^\alpha_t, \quad Y^{[1]}_0 = 0,$$

$$\begin{aligned}\mathrm{d}X^{[2]}_t &= \partial_{(x,y)} V_0(X^0_t, Y^0_t, 0)[(X^{[2]}_t, Y^{[2]}_t)]\,\mathrm{d}t + \partial^2 V_0(X^0_t, Y^0_t, 0)[(X^{[1]}_t, Y^{[1]}_t, 1)^{\otimes 2}]\,\mathrm{d}t \\ &\quad + \sum_{\alpha=1}^r (\partial^2 V_\alpha)(X^0_t, Y^0_t, 0)[(X^{[1]}_{t-}, Y^{[1]}_{t-}, 1)^{\otimes 2}]\,\mathrm{d}L^\alpha_t, \quad X^{[2]}_0 = 0,\end{aligned}$$

$$\begin{aligned}\mathrm{d}Y^{[2]}_t &= \partial_{(x,y)} V'_0(X^0_t, Y^0_t, 0)[(X^{[2]}_t, Y^{[2]}_t)]\,\mathrm{d}t + \partial^2 V'_0(X^0_t, Y^0_t, 0)[(X^{[1]}_t, Y^{[1]}_t, 1)^{\otimes 2}]\,\mathrm{d}t \\ &\quad + \sum_{\alpha=1}^r (\partial^2 V'_\alpha)(X^0_t, Y^0_t, 0)[(X^{[1]}_{t-}, Y^{[1]}_{t-}, 1)^{\otimes 2}]\,\mathrm{d}L^\alpha_t, \quad Y^{[2]}_0 = 0.\end{aligned}$$

In the above equations, the terms $\partial^2_{(x,y)}V_\alpha$ and $\partial^2_{(x,y)}V_\alpha$ in $\partial^2 V_\alpha$ and $\partial^2 V'_\alpha$ vanish for $\alpha = 1,\dots,r$ as a matter of fact.

If one uses Θ_t defined by

$$\mathrm{d}\Theta_t = \partial_{(x,y)}\begin{bmatrix} V_0 \\ V'_0 \end{bmatrix}(X^0_t, Y^0_t, 0)\Theta_t\,\mathrm{d}t \quad \text{(matrix product)},$$

$$\Theta_0 = I_{D+D_1},$$

then the derivatives are expressed in

$$\begin{bmatrix} X^{[1]}_t \\ Y^{[1]}_t \end{bmatrix} = \sum_{\alpha=0}^r \int_0^t \Theta_t\Theta_s^{-1}\partial_\varepsilon \begin{bmatrix} V_\alpha \\ V'_\alpha \end{bmatrix}(X^0_s, Y^0_s, 0)\,\mathrm{d}L^\alpha_s \quad (L^0_s = s),$$

and

$$\begin{bmatrix} X^{[2]}_t \\ Y^{[2]}_t \end{bmatrix} = \int_0^t \Theta_t\Theta_s^{-1}\left\{\partial^2 \begin{bmatrix} V_0 \\ V'_0 \end{bmatrix}(X^0_s, Y^0_s, 0)[(X^{[1]}_s, Y^{[1]}_s, 1)^{\otimes 2}]\,\mathrm{d}s\right.$$
$$\left. + \sum_{\alpha=1}^r \left(\partial^2 \begin{bmatrix} V_\alpha \\ V'_\alpha \end{bmatrix}\right)(X^0_s, Y^0_s, 0)[(X^{[1]}_{s-}, Y^{[1]}_{s-}, 1)^{\otimes 2}]\,\mathrm{d}L^\alpha_s\right\}.$$

Therefore, under non-degeneracy condition, we can obtain an expansion of $P[g(Z^\varepsilon)|F^\varepsilon = x]$ (filtering and smoothing).

Functionals ζ_0 and f_0 are linear Lévy integrals, and ζ_1 and f_1 are double Lévy integrals. The second-order terms may involve computational problems. If L^α are Gaussian, there is no problem. In this case, all terms appearing in the expansion have closed forms. Otherwise, for general Lévy processes, it is not necessarily easy to give explicit expressions of the conditional expectations of double Lévy integrals given linear Lévy integrals. However, in such cases, we can still apply rough Monte-Carlo simulations to the second-order terms. This is called the *hybrid I method* investigated in Kashiwakura and Yoshida (2001).

Example 4. Consider a $(D+D_1)$-dimensional stochastic process $(X^\varepsilon, Y^\varepsilon)$ satisfying the stochastic differential equation:

$$\mathrm{d}X^\varepsilon_t = A^X_0(\theta^\varepsilon_t)X^\varepsilon_t\,\mathrm{d}t + A^Y_0(\theta^\varepsilon_t)Y^\varepsilon_t\,\mathrm{d}t + \sum_{\alpha=1}^r A_\alpha(\theta^\varepsilon_t)\,\mathrm{d}w^\alpha_t, \quad X^\varepsilon_0 = x_0,$$

$$\mathrm{d}Y^\varepsilon_t = B^X_0(\theta^\varepsilon_t)X^\varepsilon_t\,\mathrm{d}t + B^Y_0(\theta^\varepsilon_t)Y^\varepsilon_t\,\mathrm{d}t + \sum_{\alpha=1}^r B_\alpha(\theta^\varepsilon_t)\,\mathrm{d}w^\alpha_t, \quad Y^\varepsilon_0 = y_0,$$

where $A^X_0, A^Y_0, A_\alpha, B^X_0, B^Y_0, B_\alpha$ are matrix-valued smooth functions, θ^ε_t is a vector-valued stochastic process, and w^α are Wiener processes independent of θ^ε. θ^ε_t and X^ε_t are latent variables and Y^ε_t are observable. We assume that θ^ε_t converges to a deterministic process θ_t^0 as $\varepsilon \downarrow 0$, and moreover that θ_t^ε is smooth in ε. This is a *partial non-Gaussian state space model* which recently attracts time series analysts; see Shephard (1994), also Kitagawa (1987).

As the preceding example, we consider random variables

$$Z^\varepsilon = \left(\int_{\mathbf{T}^m} Z_l(X_{\mathbf{t}}^\varepsilon)\mu_l(\mathrm{d}\mathbf{t}) \right)_{l=1,\dots,d}$$

and

$$F^\varepsilon = \left(\int_{\mathbf{T}^n} F_k(Y_{\mathbf{s}}^\varepsilon)\nu_k(\mathrm{d}\mathbf{s}) \right)_{k=1,\dots,d_1}.$$

In particular,

$$\zeta_0 = \left(\int_{\mathbf{T}^m} Z_l(X_{\mathbf{t}}^0)\mu_l(\mathrm{d}\mathbf{t}) \right)_{l=1,\dots,d}, \quad \zeta_1 = \left(\int_{\mathbf{T}^m} \partial Z_l(X_{\mathbf{t}}^0)[X_{\mathbf{t}}^{[1]}]\mu_l(\mathrm{d}\mathbf{t}) \right)_{l=1,\dots,d},$$

$$f_0 = \left(\int_{\mathbf{T}^n} F_k(Y_{\mathbf{s}}^0)\nu_k(\mathrm{d}\mathbf{s}) \right)_{k=1,\dots,d_1}, \quad f_1 = \left(\int_{\mathbf{T}^n} \partial F_k(Y_{\mathbf{s}}^0)[Y_{\mathbf{s}}^{[1]}]\nu_k(\mathrm{d}\mathbf{s}) \right)_{k=1,\dots,d_1}.$$

It is easy to see

$$\mathrm{d}X_t^0 = A_0^X(\theta_t^0)X_t^0\,\mathrm{d}t + A_0^Y(\theta_t^0)Y_t^0\,\mathrm{d}t + \sum_{\alpha=1}^r A_\alpha(\theta_t^0)\,\mathrm{d}w_t^\alpha,$$

$$\mathrm{d}Y_t^0 = B_0^X(\theta_t^0)X_t^0\,\mathrm{d}t + B_0^Y(\theta_t^0)Y_t^0\,\mathrm{d}t + \sum_{\alpha=1}^r B_\alpha(\theta_t^0)\,\mathrm{d}w_t^\alpha,$$

$$\begin{aligned}\mathrm{d}X_t^{[1]} = {} & A_0^X(\theta_t^0)X_t^{[1]}\,\mathrm{d}t + A_0^Y(\theta_t^0)Y_t^{[1]}\,\mathrm{d}t + \partial A_0^X(\theta_t^0)[\theta_t^{[1]}]X_t^0\,\mathrm{d}t + \partial A_0^Y(\theta_t^0)[\theta_t^{[1]}]Y_t^0\,\mathrm{d}t \\ & + \sum_{\alpha=1}^r \partial A_\alpha(\theta_t^0)[\theta_t^{[1]}]\,\mathrm{d}w_t^\alpha,\end{aligned}$$

$$\begin{aligned}\mathrm{d}Y_t^{[1]} = {} & B_0^X(\theta_t^0)X_t^{[1]}\,\mathrm{d}t + B_0^Y(\theta_t^0)Y_t^{[1]}\,\mathrm{d}t + \partial B_0^X(\theta_t^0)[\theta_t^{[1]}]X_t^0\,\mathrm{d}t + \partial B_0^Y(\theta_t^0)[\theta_t^{[1]}]Y_t^0\,\mathrm{d}t \\ & + \sum_{\alpha=1}^r \partial B_\alpha(\theta_t^0)[\theta_t^{[1]}]\,\mathrm{d}w_t^\alpha\end{aligned}$$

with

$$X_0^0 = x_0, \quad Y_0^0 = y_0, \quad X_0^{[1]} = 0, \quad Y_0^{[1]} = 0.$$

Thus, $\bar{X}_t = [X_t^{0\prime}, Y_t^{0\prime}, X_t^{[1]\prime}, Y_t^{[1]\prime}]'$ forms a linear system admitting a representation:

$$\mathrm{d}\bar{X}_t = K_t\bar{X}_t\,\mathrm{d}t + L_t\,\mathrm{d}\bar{w}_t, \quad \bar{X}_t = \bar{x}_0$$

for $\bar{w}_t = (w_t^\alpha)$. K_t and L_t are random matrices described by θ_t^0 and $\theta_t^{[1]}$. Trivially,

$$\bar{X}_t = \mathrm{e}^{\int_0^t K_s\,\mathrm{d}s}\left[\int_0^t \mathrm{e}^{-\int_0^s K_u\,\mathrm{d}u} L_s\,\mathrm{d}\bar{w}_s + \bar{x}_0\right].$$

It will be possible to give an explicit expression, to some extent, of conditional expectations of ζ_1, f_1 given (ζ_0, f_0). If Z_l and F_k are linear, it is easy to do by conditioning first with $\theta_\cdot^0$ and $\theta_\cdot^{[1]}$, and by integration next.

We do not go into details but we may use the partial Malliavin calculus which shifts only w^α and leaves θ^ε unchanged. Then, under a non-degeneracy condition, we finally obtain an expansion of $P[g(Z^\varepsilon)|F^\varepsilon = x]$. See Masuda and Yoshida (2002) for a practical numerical scheme.

Acknowledgements

I express my gratitude to the editor and the referees for careful reading and valuable comments. Also, I appreciate Mr. Hiroki Masuda's assistance.

Appendix

We will here list formulas for the symmetric transition kernel $p(t,x,y)$ given in Section 4. Those were used in Sakamoto and Yoshida (1996). For $\mathbf{n} \in \mathbb{Z}_+^d$, $|\mathbf{n}| = n_1 + \cdots + n_d$, and $\partial^{\mathbf{n}} = \partial_1^{n_1} \cdots \partial_d^{n_d}$, $\partial_i = \partial/\partial x_i$. Denote by $\phi(x)$ the d-dimensional standard normal density function, and let

$$H_{\mathbf{n}}(x) = \phi(x)^{-1}(-\partial)^{\mathbf{n}}\phi(x), \qquad \sigma(t) = \left(\frac{\tanh\sqrt{2}t}{\sqrt{2}}\right)^{1/2},$$

$$g(x,t) = \frac{\exp(-\sigma^2(t)|x|^2)}{\sqrt{\cosh\sqrt{2}t}^{\,d}}.$$

Then $p(t,x,y)$ has representations:

$$p(t,x,y) = \frac{g(x,t)}{\sigma(t)^d}\phi\left(\frac{y - x\,\mathrm{sech}\sqrt{2}t}{\sigma(t)}\right) = \frac{g(y,t)}{\sigma(t)^d}\phi\left(\frac{x - y\,\mathrm{sech}\sqrt{2}t}{\sigma(t)}\right).$$

Moreover,

$$\partial_x^{\mathbf{n}} p(t,x,y) = \frac{1}{(-\sigma(t))^{|\mathbf{n}|}} H_{\mathbf{n}}\left(\frac{x - y\,\mathrm{sech}\sqrt{2}t}{\sigma(t)}\right) p(t,x,y)$$

and the derivatives of $g(x,t)$ are given by

$$\partial_x^{\mathbf{n}} g(x,t) = (-\sqrt{2}\sigma(t))^{|\mathbf{n}|} H_n(\sqrt{2}\sigma(t)x)\ g(x,t).$$

References

Barndorff-Nielsen, O.E., Cox, D.R., 1994. Inference and Asymptotics. Chapman & Hall, London.

Bichteler, K., Gravereaux, J.-B., Jacod, J., 1987. Malliavin Calculus for Processes with Jumps. Gordon and Breach Science Publishers, New York, London, Paris, Montreux, Tokyo.

Del Moral, P., Jacod, J., Protter, Ph., 2001. The Monte-Carlo method for filtering with discrete-time observations. Probab. Theory Relat. Fields 120, 346–368.

Dermoune, A., Kutoyants, Yu.A., 1995. Expansion of distribution function of maximum likelihood estimate for misspecified diffusion type observations. Stochastics and Stochastic Reports 52, 121–145.

Genon-Catalot, V., Jacod, J., 1993. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Ann. Inst. Henri Poincaré 29 (1), 119–151.

Ghosh, J.K., 1994. Higher order asymptotics. IMS, California.

Ikeda, N., Watanabe, S., 1990. Stochastic Differential Equations and Diffusion Processes, 2nd Edition. North-Holland/Kodansha, Tokyo.

Jacod, J., 1996. On continuous conditional Gaussian martingales and stable convergence in law. Prépublication No. 339 du Laboratoire de Probabilités, Université Paris 6.

Kashiwakura, K., Yoshida, N., 2001. Asymptotic and hybrid expansions. In preparation.

Kim, Y.J., Kunitomo, N., 1999. Pricing options under stochastic interest rates. Asia Pacific Financial Markets 6, 49–70.

Kitagawa, G., 1987. Non-Gaussian state space modelling of non-stationary time series. J. Amer. Statist. Assoc. 82, 503–514.

Kunitomo, N., Takahashi, A., 1998. On validity of the asymptotic expansion approach in contingent claim analysis. Discussion Paper No.98-F-6, Faculty of Economics, University of Tokyo.

Kunitomo, N., Takahashi, A., 2001. The aysmptotic expansion approach to the valuation of interest rate contingent claims. Math. Finance 11, 117–151.

Kusuoka, S., Stroock, D.W., 1991. Precise asymptotics of certain Wiener functionals. J. Funct. Anal. 99, 1–74.

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

Malliavin, P., 1997. Stochastic Analysis. Springer, Berlin, Heidelberg, New York.

Masuda, H., Yoshida, N., 2002. The double Edgeworth expansion for the partial non-Gaussian state space models. Preprint

Nualart, D., 1995. The Malliavin Calculus and Related Topics. Springer, Berlin.

Picard, J., 1991. Efficiency of the extended Kalman filter for nonlinear systems with small noise. SIAM J. Appl. Math. 51, 843–885.

Sakamoto, Y., Yoshida, N., 1996. Expansion of perturbed random variables based on generalized Wiener functionals. J. Multivariate Anal. 59 (1), 34–59.

Shephard, N., 1994. Partial non-Gaussian state space. Biometrika 81, 115–131.

Sørensen, M., Yoshida, N., 2000. Random limit expansion for small diffusion processes. Unpublished manuscript.

Sweeting, T.J., 1986. Asymptotic conditional inference for the offspring mean of a supercritical Galton-Watson process. Ann. Statist. 14, 925–933.

Takahashi, A., 1995. Essays on the valuation problems of contingent claims. Ph.D. Dissertation, Haas School of Business, University of California, Berkeley.

Takahashi, A., 1999. An asymptotic expansion approach to pricing contingent claims. Asia-Pacific Financial Markets 6, 115–151.

Takahashi, A., Yoshida, N., 2001. Asymptotic expansion scheme for optimal portfolio. Submitted for publication.

Takanobu, S., Watanabe, S., 1993. Asymptotic expansion formulas of the Schilder type for a class of conditional Wiener functional integrations. In: Elworthy, K.D., Ikeda, N. (Eds.), Asymptotic problems in probability theory: Wiener functionals and asymptotics. Proceedings of the Taniguchi International Symposium, Sanda and Kyoto, 1990, Longman, UK, pp. 194–241.

Uchida, M., Yoshida, N., 1999. Information criteria for small diffusions via the theory of Malliavin–Watanabe. Submitted for publication.

Watanabe, S., 1983. Malliavin's calculus in terms of generalized Wiener functionals. In: Kallianpur, G. (Ed.), Theory and Applications of Random Fields, Proc. IFIP Conference Bangalore 1982, Lecture Notes in Control Information Science, Vol. 49. Springer, Berlin, Heidelberg, New York, pp. 284–290.

Watanabe, S., 1984. Lectures on Stochastic Differential Equations and Malliavin Calculus. Tata Institute of Fundamental Research, Springer, Bombay.

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

Yoshida, N., 1992a. Asymptotic expansion for small diffusions via the theory of Malliavin–Watanabe. Prob. Theory Related Fields 92, 275–311.

Yoshida, N., 1992b. Asymptotic expansion for statistics related to small diffusions. J. Jpn. Statist. Soc. 22 (2), 139–159.

Yoshida, N., 1993. Asymptotic expansion of Bayes estimators for small diffusions. Probab. Theory Relat. Fields 95, 429–450.

Yoshida, N., 1996. Asymptotic expansions for perturbed systems on Wiener space: maximum likelihood estimators. J. Multivariate Anal. 57, 1–36.

Yoshida, N., 1999. Malliavin calculus and statistics. In: S. Kotz et. al. (Eds.), Encyclopedia of Statistical Sciences, Update Vol. 3. Wiley, New York.

Yoshida, N., 2001. Conditional asymptotics. Preprint.

Zeitouni, O., 1988. Approximate and limit results for nonlinear filters with small observation noise: the linear sensor and constant diffusion coefficient case. IEEE Trans. Automat. Control 33, 595–599.