

# ASYMPTOTIC EXPANSION UNDER DEGENERACY

Yuji Sakamoto\* and Nakahiro Yoshida\*\*

We will consider a stochastic expansion described by random variables whose covariance matrix is asymptotically degenerate. Though the conventional approach with Bhattacharya-Ghosh's transform requires the nondegeneracy of the covariance matrix, it is known that this method still works even in degenerate cases with the help of the so-called global approach. In this paper, we explain this fact and also mention, as an example, the third order asymptotic expansion of the maximum likelihood estimator for the O-U process.

*Key words and phrases:* Asymptotic expansion, diffusion functional,  $M$ -estimator, mixing process.

## 1. Introduction

The fundamentals of asymptotic expansion for semimartingales (such as jump diffusions) and the theory of asymptotic expansions for statistics in the higher-order inference for semimartingales both have already been established by a series of papers.

Yoshida (1992a, 1993, 1996a) presented asymptotic expansions for a diffusion process with small noises, and for a general perturbed model. Yoshida (1996b, 1997) derived distributional asymptotic expansions for martingales with applications, e.g., non-linear ergodic diffusions, volatility estimation for discretely observed diffusion processes, and long-memory time series. This approach is called the **global approach**. In this direction, Sakamoto and Yoshida (1998a) presented expansions of  $M$ -estimators for ergodic diffusion processes. They treated non-linear diffusions, to say nothing of linear diffusions, and (exact and conditional) maximum likelihood methods, and more generally  $M$ -estimators (robust and non-robust; here “robust” means the control theoretic robustness that the estimating functional is continuous with respect to the local uniform norms for the input process).

Kusuoka and Yoshida (2000) derived and proved the validity of asymptotic expansions for an  $\epsilon$ -Markov process with mixing property. It covers diffusions even with jumps and point processes. This approach is nowadays called the **local approach**<sup>1</sup>. Due to this work, the next interest of “expanders” in this field focused on derivation of expansions for statistical estimators and test statistics.

---

Received April 1, 2003. Revised July 23, 2003. Accepted September 18, 2003.

\*Faculty of Human Environment, Hiroshima International University, 555-36 Gakuendai, Kurose-cho, Kamo-Gun, Hiroshima 724-0695, Japan.

\*\*Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8914, Japan.

<sup>1</sup> It is called so because one only estimates conditional characteristic functionals over local short time intervals, differently from the martingale expansion. For details, see Yoshida(2000).

Sakamoto and Yoshida applied Bhattacharya and Ghosh's theory on transformation of asymptotic expansions, and presented expansion formulae, of course, with validity when adopting the mixing  $\epsilon$ -Markov process as the underlying stochastic process constructing functionals. Sakamoto and Yoshida (1998b) gave the expansion of the maximum likelihood estimator for general diffusion processes (i.e., multidimensional non-linear and non-linearly parametrized). They extended it to  $M$ -estimators for general statistical model and presented key expansion formulas for numerous statistics in Sakamoto and Yoshida (1999)<sup>2</sup>.

Many applications have been reported: Sakamoto (2000) presented a transformation formula for chi-square type expansions tailored to mixing  $\epsilon$ -Markov processes, and gave the precise expression of expansions in testing diffusion processes; model selection problem through asymptotic expansion methods by Uchida and Yoshida (2001, 2003a, 2003b); Kutoyants and Yoshida (2001) studied for nonparametric estimation; Sakamoto and Yoshida (2001) worked for cluster processes. Other applications of related expansions are for example: general expansion under random limits by Kashiwakura *et al.* (2002), application to shrinkage estimators by Sakamoto and Yoshida (1996), conditional expansions for Lévy driving SDE and filtering by Yoshida (2001b). Option price expansion by Yoshida (1992b) and a series of papers by Kunitomo and Takahashi (1998, 2001) and others. Details will appear in Yoshida (2003).

Bhattacharya and Ghosh's transform method requires that the random vector which is used to express a stochastic expansion of a statistic in question satisfies a kind of non-degeneracy. Only the O-U (Ornstein-Uhlenbeck) process may seem exceptional because of the complete linearity. Concerning the asymptotic expansion of its maximum likelihood estimator and, more generally,  $M$ -estimators, the second order expansion does not have any problem: the cases are explained in Yoshida (1997) and Sakamoto and Yoshida (1998a). Though it may apparently have a problem for the third order expansion, it is also covered by the existing general third order result if we use a simple result presented in this paper. Since, as the reader will find below, the idea is very simple (in some sense similar to the technique used in Yoshida (1997)) but quite useful, this fact has been common among experts in this field, as Remark of Uchida and Yoshida (2001) mentioned it, since the annual meeting of the Japan Statistical Society in 2000's summer at latest.

It was backing the validity of the third order expansion for the O-U process as a simple corollary, and in fact it is nothing but the formula obtained by the formal application of the general result.

Thus far, we briefly reminded the reader of progress in *distributional* asymptotic expansion theory for semimartingales. The Malliavin calculus features in that theory. On the other hand, we should comment that Professor Mykland's expansions are not distributional but his work (Mykland, 1992) inspired one of the author to find a martingale distributional expansion by means of the Malli-

---

<sup>2</sup> This manuscript is still unpublished, but has already been well circulated among researchers of this field and has been applied to various problems.

avin calculus.

The aim of this note is to introduce a simple but useful result mentioned above. Moreover, it extends the global approach, one of the fundamental methodologies in the theory of asymptotic expansion for semimartingales. Also, we hope this note will help the reader recognize the real front of this field.

## 2. Result

Let  $(W, H, P)$  be a (partial)  $r$ -dimensional Wiener space :  $W = W^{(1)} \times W^{(2)}$  and  $P = P^{(1)} \otimes P^{(2)}$ , where  $(W^{(1)}, H, P^{(1)})$  is a usual Wiener space and  $(W^{(2)}, \mathcal{B}, P^{(2)})$  is a probability space. Denote by  $D_{p,s}$  the Sobolev space of Wiener functionals equipped with a Sobolev norm  $\|\cdot\|_{p,s}$ , and denote by  $\sigma_F$  the Malliavin covariance of a Wiener functional  $F$  on  $W$ . For any  $M, \gamma > 0$ , let  $\mathcal{E}(M, \gamma)$  be the set of measurable functions  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $|f(x)| \leq M(1 + |x|^\gamma)$  for all  $x \in \mathbb{R}^d$ . For any  $T > 0$ , let  $X_T$  be an  $\mathbb{R}^d$ -valued Wiener functional admitting an asymptotic expansion, more precisely, suppose that there exists a signed measure  $\Psi_T$  such that for any  $f \in \mathcal{E}(M, \gamma)$ ,  $E[f(X_T)] = \Psi_T[f] + \bar{o}(s_T)$ , where  $s_T$  is a given sequence tending to 0 as  $T \rightarrow \infty$  and  $\bar{o}(s_T) = o(s_T)$  uniformly in  $\mathcal{E}(M, \gamma)$ . Define  $S_T$  by  $S_T = X_T + s_T Y_T$  for another  $\mathbb{R}^d$ -valued Wiener functional  $Y_T$ . For the basic notions of the Malliavin calculus, see Ikeda and Watanabe (1989).

For notational simplicity, we here focus our attention to Wiener functionals, however it is also possible to show a similar result for Wiener-Poisson functionals.

**THEOREM 2.1.** *For  $m \in \mathbb{N}$  and positive numbers  $M, \gamma, \ell_1$  and  $\ell_2$ , suppose that  $m > \gamma/2 + 1$ ,  $\ell_1 \geq \ell_2 - 1$  and  $\ell_2 \geq (2m + 1) \vee (d + 3)$ . Assume that (1)  $\sup_T \|Y_T\|_{p,\ell_2} + \sup_T \|X_T\|_{p,\ell_2} < \infty$  for any  $p > 1$ , (2)  $(X_T, Y_T) \xrightarrow{d} (X_\infty, Y_\infty)$  for some random variables  $X_\infty$  and  $Y_\infty$ . In addition, assume that there exists a functional  $\xi_T$  such that (3)  $\sup_T \|\xi_T\|_{p,\ell_1} < \infty$  for any  $p > 1$ , (4)  $P[|\xi_T| > 1/2] = O(s_T^\alpha)$  for some  $\alpha > 1$ , (5)  $\sup_T E[1_{\{|\xi_T| < 1\}}(\det \sigma_{X_T})^{-p}] < \infty$  for any  $p > 1$ . Then for any  $f \in \mathcal{E}(M, \gamma)$ ,*

$$(2.1) \quad E[f(S_T)] = \Psi_T[f] + s_T \int_{\mathbb{R}^d} f(x) g_\infty(x) dx + \bar{o}(s_T),$$

where

$$g_\infty(x) = -\partial_x (\mathbb{E}[Y_\infty \mid X_\infty = x] p^{X_\infty}(x)).$$

**PROOF.** Let  $\eta_T = s_T(|\sigma_{X_T}|^{2d} + |\sigma_{Y_T}|^{2d})(\det \sigma_{X_T})^{-2}$ , and  $\psi_T = \psi(\xi_T)\psi(\eta_T)$  for a function  $\psi \in C^\infty$  such that  $\psi(x) = 1$ ,  $|x| < 1/2$ ,  $\psi(x) = 0$ ,  $|x| > 1$ ,  $0 \leq \psi(x) \leq 1$  for any  $x \in \mathbb{R}$ . Then it follows that for  $q = (1 + \alpha)/2$  and  $q' > 1$

satisfying  $1/q + 1/q' = 1$ ,

$$\begin{aligned} |E[f(S_T)] - E[\psi_T f(S_T)]| &\leq \|1 - \psi_T\|_q \|f(S_T)\|_{q'} \\ &\leq \left( P[|\xi_T| > 1/2]^{1/q} + P[|\xi_T| \leq 1/2, \eta_T > 1/2]^{1/q} \right) \\ &\quad \times M \left( 1 + \|S_T\|_{\gamma q'}^\gamma \right) \\ &= \bar{O}\left(s_T^{\alpha/q}\right) + \bar{O}\left(s_T^2\right) = \bar{o}(s_T). \end{aligned}$$

For any  $f \in \mathcal{E}(M, \gamma)$ , one can choose a sequence  $\phi_n \in \mathcal{S}$  such that  $\partial^j \phi_n \rightarrow \partial^j f$  in  $C_{-2m}$ ,  $j = 0, 1, 2$ . Put  $F_{T,u} = X_T + u s_T Y_T$ ,  $0 \leq u \leq 1$ . Taylor's formula yields

$$\begin{aligned} E[\psi_T \phi_n(S_T)] &= E[\psi_T(\phi_n(X_T) + s_T \partial \phi_n(X_T) Y_T)] \\ &\quad + s_T^2 \int_0^1 du (1-u) E[\psi_T \partial^2 \phi_n(F_{T,u}) [Y_T^{\otimes 2}]]. \end{aligned}$$

If  $\eta_T < 1$ , then for some constant  $c$  independent of  $T$ ,

$$\det \sigma_{F_{T,u}} \geq \det \sigma_{X_T} - c s_T \left( |\sigma_{X_T}|^{2d} + |\sigma_{Y_T}|^{2d} \right)^{1/2} > \left( 1 - c s_T^{1/2} \right) \det \sigma_{X_T}.$$

Therefore, for sufficiently large  $T$ , the IBP formula can be applied to  $F_{T,u}$  with the truncation functional  $\psi_T$ , and it holds that there exists  $T_0 > 0$  such that

$$\begin{aligned} &\sup_{T > T_0, n \in \mathbb{N}} \left| E \left[ \psi_T \partial_i \partial_j \phi_n(F_{T,u}) Y_T^i Y_T^j \right] \right| \\ &= \sup_{T > T_0, n \in \mathbb{N}} \left| E \left[ A^{-2m} \partial_i \partial_j \phi_n(F_{T,u}) \Psi_{2m}^{F_{T,u}} \left( \psi_T Y_T^i Y_T^j \right) \right] \right| \leq C_f < \infty \end{aligned}$$

for some positive constant  $C_f$  depending only on  $f$ , which implies

$$s_T^2 \int_0^1 du (1-u) E[\psi_T \partial^2 \phi_n(F_{T,u}) [Y_T^{\otimes 2}]] = \bar{O}(s_T^2).$$

Since  $E[\psi_T \phi_n(X_T)] \rightarrow E[\psi_T f(X_T)]$ ,  $n \rightarrow \infty$  and  $E[\psi_T \partial \phi_n(X_T) Y_T] \rightarrow E[\psi_T \partial f(X_T) Y_T]$ , where two limits are generalized expectations, we see that

$$E[f(S_T)] = \Psi_T[f] + s_T E[\psi_T \partial f(X_T) Y_T] + \bar{o}(s_T).$$

Therefore, it remains to show that

$$E[\psi_T \partial f(X_T) Y_T] \rightarrow \int_{\mathbb{R}^d} f(x) g_\infty(x) dx.$$

From the IBP formula, it is easy to show

$$|E[\psi_T i u e^{iu X_T} Y_T]| \leq (1 + |u|)^{-(\ell_2 - 2)}.$$

Therefore, we can define

$$g_T(x) := \frac{1}{(2\pi)^d} \int e^{-iux} E [\psi_T iue^{iuX_T} Y_T] du$$

and

$$g_\infty(x) := \frac{1}{(2\pi)^d} \int e^{-iux} \mathbb{E} [iue^{iuX_\infty} Y_\infty] du.$$

Moreover, by using the IBP formula over  $\mathbb{R}^d$ , we can also show that for any multi-index  $\mathbf{n}$ ,

$$\sup_{x,T} |x^{\mathbf{n}} g_T(x)| < \infty, \quad \sup_x |x^{\mathbf{n}} g_\infty(x)| < \infty.$$

Combining this and Lebesgue's theorem, we see that

$$\int f(x) g_T(x) dx \rightarrow \int f(x) g_\infty(x) dx, \quad \text{as } T \rightarrow \infty.$$

From the IBP formula, it follows that

$$E [\psi_T \partial f(X_T) Y_T] = \int f(x) \mu(dx)$$

for some finite signed measure  $\mu$ . Substituting  $e^{iu \cdot x}$  into  $f$  and using the uniqueness of the Fourier transform, it is found that  $\mu$  has density  $g_T$ , and therefore we see that

$$E [\psi_T \partial f(X_T) Y_T] \rightarrow \int f(x) g_\infty(x) dx, \quad \text{as } T \rightarrow \infty.$$

Since

$$|\mathbb{E} [e^{iuX_\infty}]| \leq (1 + |u|)^{-(\ell_2 - 1)},$$

we see that  $X_\infty$  has a differentiable density  $p^{X_\infty}$ . In the same way, it is easily shown that  $\mathbb{E}[Y_\infty | X_\infty = x] p^{X_\infty}(x)$  is differentiable. Thus we have that

$$g_\infty(x) = -\partial_x (\mathbb{E}[Y_\infty | X_\infty = x] p^{X_\infty}(x)). \quad \square$$

*Remark 2.1.* In the next section, we will apply this theorem to the third-order asymptotic expansion of the maximum likelihood estimator for the Ornstein-Uhlenbeck process, but this result might be applicable to other higher order expansions with such degeneracy problems.

### 3. Example

Let  $\Theta$  be an open bounded subset in  $\mathbb{R}$ , and for any  $\theta \in \Theta$  let  $X^\theta = (X_t^\theta : t \in \mathbf{R}_+)$  be a one-dimensional stationary ergodic diffusion process satisfying

$$(3.1) \quad dX_t = b(X_t, \theta)dt + dw_t,$$

with stationary distribution  $\nu_\theta$  given by

$$\nu_\theta(dx) = \frac{n(x, \theta)}{\int_{-\infty}^{\infty} n(u, \theta)du} dx,$$

where

$$n(x, \theta) = \exp\left(2 \int_0^x b(u, \theta)du\right).$$

We denote by  $\theta_0$  the true value of  $\theta$ , and omit  $\theta_0$  in functions of  $\theta$  when they are evaluated at  $\theta = \theta_0$ , e.g.,  $X = X^{\theta_0}$ ,  $\nu = \nu_{\theta_0}$ . Assume that for any  $T > 0$  the likelihood function  $\ell_T$  of  $\theta$  based on observations  $X = (X_t : t \in [0, T])$  is given by  $\ell_T(\theta) = \log \frac{d\nu}{dx}(X_0, \theta) + \Lambda_T(\theta)$  for some reference measure,

$$\Lambda_T(\theta) = \int_0^T b(X_t, \theta)dX_t - \frac{1}{2} \int_0^T b^2(X_t, \theta)dt.$$

Let  $\hat{\theta}_T$  be the maximum likelihood estimator solving  $\delta\ell_T(\hat{\theta}_T) = 0$ ,  $\delta = \partial/\partial\theta$  (see Theorem 7.1 in Sakamoto and Yoshida (1999) for the existence and consistency of the multi-dimensional  $M$ -estimator). As in the i.i.d. setting or time series, the third order stochastic expansion of  $\hat{\theta}_T$  is given by

$$\begin{aligned} \sqrt{T}(\hat{\theta}_T - \theta_0) &= S_T + \frac{1}{T\sqrt{T}}R_3, \\ S_T &= Z_1 + \frac{1}{\sqrt{T}}\left(Z_1Z_2 + \frac{1}{2}\bar{\nu}_3Z_1^2\right) \\ &\quad + \frac{1}{T}\left(\frac{1}{6}(\bar{\nu}_4 + 3\bar{\nu}_3^2)Z_1^3 + \frac{3}{2}\bar{\nu}_3Z_1^2Z_2 + Z_1Z_2^2 + \frac{1}{2}Z_1^2Z_3\right), \end{aligned}$$

where  $Z_1$ ,  $Z_2$ ,  $Z_3$  and  $\bar{\nu}_3$ ,  $\bar{\nu}_4$  are random variables and constants defined by

$$\begin{aligned} Z_1 &= \frac{1}{\sqrt{T}}\bar{\nu}_2^{-1}\delta\ell_T(\theta_0), & \bar{\nu}_2 &= -\frac{1}{T}E_{\theta_0}[\delta^2\ell_T(\theta_0)], \\ Z_2 &= \frac{1}{\sqrt{T}}\bar{\nu}_2^{-1}(\delta^2\ell_T(\theta_0) - E_{\theta_0}[\delta^2\ell_T(\theta_0)]), \\ Z_3 &= \frac{1}{\sqrt{T}}\bar{\nu}_2^{-1}(\delta^3\ell_T(\theta_0) - E_{\theta_0}[\delta^3\ell_T(\theta_0)]), \\ \bar{\nu}_3 &= \frac{1}{T}\bar{\nu}_2^{-1}E_{\theta_0}[\delta^3\ell_T(\theta_0)], & \bar{\nu}_4 &= \frac{1}{T}\bar{\nu}_2^{-1}E_{\theta_0}[\delta^4\ell_T(\theta_0)]. \end{aligned}$$

The remainder term  $R_3/(T\sqrt{T})$  can be neglected in the third order asymptotic expansion of  $\sqrt{T}(\hat{\theta}_T - \theta_0)$  under some moment conditions (the Delta-method).

In general, the joint distribution of  $Z = (Z_1, Z_2, Z_3)$  admits an asymptotic expansion, from which that of  $S_T$  is obtained through the so-called Bhattacharya-Ghosh's transform  $Z \rightarrow S_T$ . In the paper of Bhattacharya and Ghosh (1978), this approach was adopted for the mathematically rigorous derivation of the asymptotic expansion of the statistical estimator in the i.i.d. setting, and it was also used for the time series models in Götze and Hipp (1983). Kusuoka and Yoshida (2000) obtained the asymptotic expansion of the  $S_T$ -type functional of the continuous-time stochastic process, which includes the (jump) diffusion process as a typical example, and Sakamoto and Yoshida (1999) presented the formula for the asymptotic expansion of the  $M$ -estimator, which was there applied to the general diffusion model.

As a special case of Theorem 7.2 in Sakamoto and Yoshida (1999), we here present the asymptotic expansion of the maximum likelihood estimator for the one-dimensional diffusion process defined by (3.1). In order to express explicitly the coefficients in the expansion, we prepare some notations. For the measurable function  $f$  satisfying  $\nu(f) = 0$ , denote by  $G_f$  the Green function  $G$  solving  $\mathcal{A}G = f$ , where

$$\mathcal{A} = b(x, \theta_0) \frac{\partial}{\partial x} + \frac{1}{2} \frac{\partial^2}{\partial x^2}, \quad \nu(f) = \int_{-\infty}^{\infty} f(x) \nu(x) dx.$$

Moreover, for each measurable function  $f$ , let

$$[f](x) = -\frac{\partial}{\partial x} G_{f-\nu(f)}(x),$$

and put

$$\begin{aligned} F_{i,j} &= \nu(b_i \cdot b_j), & F_{[i,j],k} &= \nu([b_i \cdot b_j] \cdot b_k), \\ F_{[[i,j],k],m} &= \nu([[b_i \cdot b_j] \cdot b_k] \cdot b_m), & F_{[i,j],[k,m]} &= \nu([b_i \cdot b_j] \cdot [b_k \cdot b_m]), \end{aligned}$$

where  $b_i(x) = \delta^i b(x, \theta_0)$ . Let  $\rho = F_{1,1}$ ,

$$\tau = \text{Var} \left[ \delta \log \frac{d\nu_{\theta_0}}{dx} \right], \quad \zeta = E \left[ \delta^2 \log \frac{d\nu_{\theta_0}}{dx} \right], \quad \Delta = \nu \left( \delta \log \frac{d\nu_{\theta_0}}{dx} \right).$$

The Hermit polynomials  $h_j(z; g)$  are defined by

$$h_j(z; g) = (-g)^{-j} e^{gz^2/2} \frac{d^j}{dz^j} e^{-gz^2/2}.$$

**THEOREM 3.1** (A specialized version of Theorem 7.2 in Sakamoto and Yoshida (1999)). *Let  $M, \gamma > 0$ ,  $f \in \mathcal{E}(M, \gamma)$ , and define  $\hat{\theta}_T^*$  by  $\hat{\theta}_T^* = \hat{\theta}_T - \beta(\hat{\theta}_T)/T$  for a given function  $\beta$ . Then, under the regularity conditions of Theorem 7.2 in Sakamoto and Yoshida (1999) for the one-dimensional maximum*

likelihood estimator  $\hat{\theta}_T$  in the one-dimensional diffusion setting, there exist positive constants  $c, \tilde{C}, \tilde{\epsilon}, \hat{g} > \rho$  such that

$$(3.2) \quad \left| E \left[ f \left( \sqrt{T} \left( \hat{\theta}_T^* - \theta \right) \right) \right] - \int dz f(z) q_{T,2}(z) \right| \leq c\omega \left( f, \tilde{C}T^{-(\tilde{\epsilon}+2)/2}, \hat{g} \right) + o(T^{-1}),$$

where

$$q_{T,2}(z) = \phi(z; \rho^{-1}) \left( 1 - \frac{1}{2\sqrt{T}} \Gamma^{(-1/3)} h_3(z; \rho) - \frac{1}{\sqrt{T}} \check{\beta} h_1(z; \rho) + \frac{1}{2T} A_2^* h_2(z; \rho) + \frac{1}{24T} c_4 h_4(z; \rho) + \frac{1}{8T} \Gamma^{(-1/3)} \Gamma^{(-1/3)} h_6(z; \rho) \right),$$

$$\Gamma^{(\alpha)} = F_{2,1} - F_{[1,1],1} + 3 \frac{1-\alpha}{2} F_{[1,1],1}, \quad \check{\beta} = \rho \beta - \Delta + \frac{1}{2} \rho^{-1} \Gamma^{(-1)},$$

$$A_2^* = \tau + 2\zeta - \rho^{-1} (F_{3,1} + 5F_{[2,1],1} + 4F_{[1,1],2} + 4F_{[[1,1],1],1} - F_{[1,1],[1,1]}) + \rho^{-2} \left( \frac{5}{2} \Gamma^{(-1)} \Gamma^{(-1)} - \Gamma^{(1)} \Gamma^{(1)} + 2\Gamma^{(-1)} \Gamma^{(1)} \right) + \check{\beta}^2 - 2\rho \left( \rho^{-2} \Delta \left( \Gamma^{(1)} + \Gamma^{(-1)} \right) + \delta\beta \right)$$

$$c_4 = -12 (F_{[[1,1],1],1} + F_{[1,1],2} + F_{[2,1],1}) + 3F_{[1,1],[1,1]} - 4F_{3,1} + 12\rho^{-1} \Gamma^{(-1)} \left( \Gamma^{(-1)} + \Gamma^{(1)} \right) + 12\Gamma^{(-1/3)} \check{\beta}.$$

Note that this theorem embraces the following three cases: (i) the joint distribution of  $(Z_1, Z_2, Z_3)$  admits the asymptotic expansion, (ii) the joint distribution of  $(Z_1, Z_2)$  or  $(Z_1, Z_3)$  admits the asymptotic expansion and  $Z_2$  and  $Z_3$  are linearly dependent for each  $T > 0$ , (iii) the distribution of  $Z_1$  admits an asymptotic expansion and both  $Z_2$  and  $Z_3$  vanish. Therefore, Theorem 3.1 covers many cases including almost all non-linear models, but we can adduce an exceptional case as follows.

Let us consider the case where  $\theta > 0$ .  $b(x, \theta) = \theta m(x)$  for a given function  $m$ . We then have

$$\begin{aligned} Z_1 &= \frac{1}{\sqrt{T}} g^{-1} \left( \bar{\ell}_1^{(0)}(X_0) + \int_0^T m(X_t) dw_t \right), \\ Z_2 &= \frac{1}{\sqrt{T}} g^{-1} \left( \bar{\ell}_2^{(0)}(X_0) - \int_0^T (m^2(X_t) - \nu(m^2)) dt \right), \\ Z_3 &= \frac{1}{\sqrt{T}} g^{-1} \bar{\ell}_3^{(0)}(X_0). \end{aligned}$$

where

$$\begin{aligned} \bar{\ell}_1^{(0)}(x) &= \delta \log \frac{d\nu_{\theta_0}}{dx}(x), \\ \bar{\ell}_i^{(0)}(x) &= \delta^i \log \frac{d\nu_{\theta_0}}{dx}(x) - E_{\theta_0} \left[ \delta^i \log \frac{d\nu_{\theta_0}}{dx}(X_0) \right], \quad i \geq 2. \end{aligned}$$



From Itô's formula, it follows that if  $m$  is integrable and differentiable,

$$\mu(X_T) = \mu(X_0) + \int_0^T m(X_t)dw_t + \theta \int_0^T m^2(X_t)dt + \frac{1}{2} \int_0^T m'(X_t)dt,$$

where  $\mu(x) = \int_0^x m(u)du$ . Combining this and stationarity of  $X$  yields  $\nu(m^2) = -\nu(m')/(2\theta)$ , and therefore we have

$$\begin{aligned} Z_1 - \theta Z_2 &= \frac{1}{\sqrt{T}}g^{-1} \left( \bar{\ell}_1^{(0)}(X_0) - \theta \bar{\ell}_2^{(0)}(X_0) + \mu(X_T) - \mu(X_0) \right. \\ &\quad \left. - \frac{1}{2} \int_0^T (m'(X_t) - \nu(m')) dt \right) \\ &= o_p(1) + \frac{1}{2\sqrt{T}}g^{-1} \int_0^T [m'](X_t)dt. \end{aligned}$$

In addition to the linearity of the drift  $b$  in  $\theta$ , assuming that  $m$  is also linear in  $x$ , i.e.,  $X$  is the Ornstein-Uhlenbeck process defined by

$$dX_t = -\theta X_t dt + dw_t,$$

we see that

$$\begin{aligned} Z_1 &= \frac{1}{\sqrt{T}}g^{-1} \left( -X_0^2 + \frac{1}{2\theta_0} - \int_0^T X_t dw_t \right), \quad Z_2 = -\frac{1}{\sqrt{T}}g^{-1} \int_0^T \left( X_t^2 - \frac{1}{2\theta_0} \right) dt, \\ Z_3 &= 0, \quad g = \frac{1}{2\theta_0} + \frac{1}{T} \frac{1}{2\theta_0^2}, \end{aligned}$$

and that  $Z_1$  and  $Z_2$  are linearly independent for each  $T > 0$ , but asymptotically linearly dependent because

$$Z_1 - \theta_0 Z_2 = \frac{1}{\sqrt{T}}v(X_0, X_T) \rightarrow 0 \quad \text{as } T \rightarrow \infty,$$

where

$$v(x, y) = g^{-1} \left( -x^2 + \frac{1}{2\theta_0} - \frac{1}{2}y^2 + \frac{1}{2}x^2 \right).$$

This implies that the Ornstein-Uhlenbeck process is an exception to Theorem 3.1, but the asymptotic expansion of  $\hat{\theta}_T$  for the Ornstein-Uhlenbeck process can be easily obtained by applying Theorem 2.1. In fact,  $S_T$  can be represented as

$$S_T = \tilde{X}_T + \frac{1}{T}\tilde{Y}_T + \frac{1}{T\sqrt{T}}\tilde{R}_3,$$

where

$$\begin{aligned} \tilde{X}_T &= Z_1 + \frac{1}{\sqrt{T}} \left( \frac{1}{\theta_0} + \frac{1}{2}\bar{\nu}_3 \right) Z_1^2 + \frac{1}{T} \left( \frac{1}{6}(\bar{\nu}_4 + 3\bar{\nu}_3^2) + \frac{3}{2\theta_0}\bar{\nu}_3 + \frac{1}{\theta_0^2} \right) Z_1^3, \\ \tilde{Y}_T &= -\frac{1}{\theta_0}v(X_0, X_T)Z_1, \end{aligned}$$

and  $\tilde{R}_3/(T\sqrt{T})$  can be neglected by the Delta-method. The asymptotic expansion of  $\tilde{X}_T$  can be derived from Theorem 7.2 of Sakamoto and Yoshida (1999), and in the same fashion as in Lemma 11 of Yoshida (1997), it can be shown that for some positive constant  $c$ ,  $P(\sigma_{\tilde{X}_T} < c) = O(T^{-2})$ . Combining these facts, we can easily check the conditions of Theorem 2.1 with  $\tilde{X}_T$  and  $\tilde{Y}_T$  in place of  $X_T$  and  $Y_T$ , and consequently it is shown that the asymptotic expansion given in Theorem 3.1 still holds true for the Ornstein-Uhlenbeck process and the coefficients are given by

$$\rho = \frac{1}{2\theta_0}, \quad \tau = \frac{1}{2\theta_0^2}, \quad \zeta = -\frac{1}{2\theta_0^2}, \quad \Delta = 0, \quad \check{\beta} = \frac{1}{2\theta_0}(\beta - 2), \quad \Gamma^{(\alpha)} = -\frac{1 - 3\alpha}{4\theta_0^2},$$

$$A_2^* = \frac{3}{2\theta_0^2} + \frac{1}{4\theta_0^2}(\beta - 2)^2 - \frac{1}{\theta_0}\delta\beta, \quad c_4 = \frac{15}{2\theta_0^3} - \frac{3}{\theta_0^3}(\beta - 2),$$

which imply that

$$q_{T,2}(y) = \phi(y; 2\theta) \left( 1 + \frac{1}{4\theta^2\sqrt{T}}x(x^2 - 2(\beta + 1)\theta) \right. \\ \left. + \frac{1}{32\theta^4T} \left( x^6 - 4\theta x^4(3 + \beta) + 4\theta^2 x^2(1 + 8\beta + \beta^2 - 4\theta\delta\beta) \right. \right. \\ \left. \left. + 8\theta^3(2 - 2\beta - \beta^2) + 32\theta^4\delta\beta \right) \right).$$

This result was obtained in 1999 as a corollary of the result in Sakamoto and Yoshida (1998b, 1999).

If we consider the conditional MLE  $\hat{\theta}_T^{(c)}$  defined by  $\delta\Lambda_T(\hat{\theta}_T^{(c)}) = 0$  for the O-U process, we encounter the same degeneracy problem. But we can settle the problem in exactly the same approach, and we can prove the validity of the third order asymptotic expansion of  $\hat{\theta}_T^{(c)}$  with

$$q_{T,2}(y) = \phi(y; 2\theta) \left( 1 + \frac{1}{4\theta^2\sqrt{T}}x(x^2 - 2(\beta + 1)\theta) \right. \\ \left. + \frac{1}{32\theta^4T} \left( x^6 - 4\theta x^4(3 + \beta) + 4\theta^2 x^2(3 + 8\beta + \beta^2 - 4\theta\delta\beta) \right. \right. \\ \left. \left. - 8\theta^3\beta(2 + \beta) + 32\theta^4\delta\beta \right) \right).$$

*Remark 3.1.* Professor Kakizawa checked our third order formula for O-U case<sup>3</sup> and he recently successfully showed in Kakizawa (2002) that fourth order expansion and beyond, i.e.,  $O(T^{-r/2})$  ( $r \geq 3$ ) order in the expansion of  $P[\sqrt{T}(\hat{\theta}_T - \theta) \leq x]$ , are possible for O-U process including a non-stationary

---

<sup>3</sup> The third order formula (with validity overcoming degeneracy problem) for the maximum likelihood estimator for O-U process was supplied in 1999 and 2000. The form of the general third order formula for diffusions itself goes back to Sakamoto and Yoshida (1998b) (rigorously the cooperative research meeting of the Institute of Statistical Mathematics in 1997), according to our best knowledge.

case, by direct computation that is possible owing to both the linearity and the gaussianity.

*Remark 3.2.* As for the case where the drift  $b$  is linear in  $\theta$  but non-linear in  $x$ , it is usual that  $Z_1, Z_2, Z_3$  are asymptotically linearly dependent, but  $(Z_1, Z_2)$  admits the asymptotic expansion. Therefore, even in this case, the asymptotic expansion of MLE can be obtained in the same way as for the Ornstein-Uhlenbeck process.

*Remark 3.3.* We can also obtain the asymptotic expansion of the conditional maximum likelihood estimator defined as a solution of  $\delta\Lambda_T(\theta)$ , and the resultant expansion becomes the same one as the (exact) maximum likelihood estimator except for  $\tau = 0$ .

### Acknowledgements

We express our gratitude to the editor, the associate editor and the referees for careful reading and valuable comments. 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.

### REFERENCES

- Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.*, **6**, 434–451.
- Götze, F. and Hipp, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors, *Z. Wahr.*, **64**, 211–239.
- Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland/Kodansha, Tokyo.
- Kakizawa, Y. (2002). Edgeworth approximation in the O-U process, *Sankhyā*, **64**, 16–41.
- Kashiwakura, K., Sørensen, M. and Yoshida, N. (2002). Random limit expansion for small diffusion processes, in preparation.
- Kunitomo, N. and 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. and Takahashi, A. (2001). The asymptotic expansion approach to the valuation of interest rate contingent claims, *Mathematical Finance*, **11**, 117–151.
- Kusuoka, S. and Yoshida, N. (2000). Malliavin calculus, geometric mixing, and expansion of diffusion functionals, *Probab. Theory Related Fields*, **116**, 457–484.
- Kutoyants, Yu. A. and Yoshida, N. (2001). On moments estimation for ergodic diffusion processes, preprint.
- Mykland, P. A. (1992). Asymptotic expansions and bootstrapping distributions for dependent variables: a martingale approach, *Ann. Statist.*, **20**, No. 2, 623–654.
- Sakamoto, Y. (2000). Asymptotic expansions of test statistics for mixing processes, preprint.
- Sakamoto, Y. and Yoshida, N. (1996). Expansion of perturbed random variables based on generalized Wiener functionals, *J. Multivariate Anal.*, **59**, No. 1, 34–59.
- Sakamoto, Y. and Yoshida, N. (1998a). Asymptotic expansions of  $M$ -estimator over Wiener space, *Statistical Inference for Stochastic Processes*, **1**, 85–103.
- Sakamoto, Y. and Yoshida, N. (1998b). Third order asymptotic expansions for diffusion processes, Cooperative Research Report, **107**, 53–60, The Institute of Statistical Mathematics, Tokyo.

- Sakamoto, Y. and Yoshida, N. (1999). Higher order asymptotic expansions for a functional of a mixing process and applications to diffusion functionals, unpublished.
- Sakamoto, Y. and Yoshida, N. (2001). Asymptotic expansion for cluster processes, preprint.
- Uchida, M. and Yoshida, N. (2001). Information criteria in model selection for mixing processes, *Statistical Inference for Stochastic Processes*, **4**, 73–98.
- Uchida, M. and Yoshida, N. (2003a). Information criteria for small diffusions via the theory of Malliavin-Watanabe, *Statistical Inference for Stochastic Processes*, to appear.
- Uchida, M. and Yoshida, N. (2003b). Asymptotic expansion and information criteria, submitted (1999).
- Yoshida, N. (1992a). Asymptotic expansion for small diffusions via the theory of Malliavin-Watanabe, *Probab. Theory Related Fields*, **92**, 275–311.
- Yoshida, N. (1992b). Asymptotic expansion for statistics related to small diffusions, *J. Japan Statistical Society*, **22**, No. 2, 139–159.
- Yoshida, N. (1993). Asymptotic expansion of Bayes estimators for small diffusions, *Probab. Theory Related Fields*, **95**, 429–450.
- Yoshida, N. (1996a). Asymptotic expansions for perturbed systems on Wiener space: maximum likelihood estimators, *J. Multivariate Analysis*, **57**, 1–36.
- Yoshida, N. (1996b). Asymptotic expansion for martingales with jumps, Research Memorandum, **601**, The Institute of Statistical Mathematics.
- Yoshida, N. (1997). Malliavin calculus and asymptotic expansion for martingales, *Probab. Theory Related Fields*, **109**, 301–342.
- Yoshida, N. (2000). Partial mixing and conditional Edgeworth expansion for diffusions with jumps, *Probab. Theory Related Fields*, to appear.
- Yoshida, N. (2001). Conditional asymptotics, preprint.
- Yoshida, N. (2003). *Stochastic Analysis and Statistics*, Springer, in preparation.