REVSTAT – Statistical Journal
Volume 5, Number 3, November 2007, 249–267

LARGE DEVIATIONS AND BERRY–ESSEEN INEQUALITIES FOR ESTIMATORS IN NONLINEAR NONHOMOGENEOUS DIFFUSIONS

Authors: Jaya P.N. Bishwal
– Department of Mathematics and Statistics,
University of North Carolina at Charlotte,
376 Fretwell Bldg., 9201 University City Blvd.,
Charlotte, NC 28223, U.S.A.
J.Bishwal@uncc.edu

Received: February 2007 Revised: June 2007 Accepted: September 2007

Abstract:

- Bounds on the large deviations probability of the maximum likelihood estimator and regular Bayes estimators, and Berry–Esseen type bound for the suitably normalized maximum likelihood estimator of a parameter appearing nonlinearly in the nonhomogeneous drift coefficient of Itô stochastic differential equation are obtained under some regularity conditions. Berry–Esseen results are illustrated for nonhomogeneous Ornstein–Uhlebeck process.

Key-Words:

- *nonhomogeneous diffusion processes; Itô stochastic differential equation; drift parameter; maximum likelihood estimator; Bayes estimators; large deviations probability; Berry–Esseen type inequality.*

AMS Subject Classification:

- 62M05, 62F12, 60H10, 60J60.

1. INTRODUCTION

Nonhomogenous diffusions are useful for modeling term structure of interest rates in finance and other fields. Asymptotic properties such as weak consistency, asymptotic normality and convergence of moments of maximum likelihood estimator (MLE) and Bayes estimators (BEs) of the drift parameter in the nonlinear nonhomogeneous Itô stochastic differential equations having nonstationary solutions were first studied by Kutoyants (1978) for the small noise asymptotics case and Kutoyants (1984) for the general case which includes both small noise and long time asymptotics. The approach was through Ibragimov and Khasminskii (1981). Later on, strong consistency and asymptotic normality for large sample case were studied by Borkar and Bagchi (1982), Mishra and Prakasa Rao (1985a) and Levanony, Shwartz and Zeitouni (1994) using the martingale approach under stronger regularity conditions. Asymptotic normality of BEs was studied by Mishra (1989) and Harison (1992) as a consequence of Bernstein-von Mises theorem. Slightly weaker assumptions than those used in Kutoyants (1984) were used by Yoshida (1990) to obtain the asymptotic behaviour of M-estimator. For first order theory in general nonergodic stochastic models through the LAMN (defined below) approach, see Basawa and Scott (1983). See the monograph Bishwal (2007) for recent results on likelihood asymptotics and Bayesian asymptotics for drift estimation of finite and infinite dimensional stochastic differential equations.

All the above results are on first order asymptotics. Beyond the first order asymptotics in consistency, Florens and Pham (1999) obtained large deviations for MLE and a minimum contrast estimator for the Ornstein–Uhlenbeck process. For the nonlinear stationary homogeneous diffusions a large deviations upper bound for the MLE and Bayes estimators was obtained by Bishwal (1999). For the nonhomogeneous diffusions, Levanony (1994) obtained the conditional large deviations upper and lower bounds for the MLE through the martingale approach following Dupuis and Kushner (1989). He obtained unconditional large deviations lower bounds following Bahadur *et al.* (1980). We obtain unconditional large deviations upper bounds following Ibragimov and Khasminskii (1981).

Beyond the first order results in asymptotic normality, Berry–Esseen type bounds in the linear homogeneous case were obtained by Mishra and Prakasa Rao (1985b) which were sharpened to the Ornstein–Uhlenbeck process by Bose (1986), Bishwal and Bose (1995) and Bishwal (2000a) respectively in order. Sharp Berry–Esseen bound for the Bayes estimators and minimum contrast estimator were obtained in Bishwal (2000b) and Bishwal (2005) respectively. In the above works on Ornstein–Uhlenbeck process, stationarity was not assumed. For nonlinear stationary ergodic diffusion, Edgeworth expansion of the distribution of the MLE was obtained by Yoshida (1997) and that for M-estimator by Sakamoto and Yoshida (1998) through the Malliavin calculus approach. As far as we know, no result is known on the rate of convergence to normality of the MLE in the nonergodic case. We obtain a Berry–Esseen type bound for the MLE following Michel and Pfanzagl (1971). Finally Berry–Esseen results are illustrated for a nonhomogeneous Ornstein–Uhlenbeck process.

2. MODEL, ASSUMPTIONS AND PRELIMINARIES

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0} P)$ be a stochastic basis satisfying the usual hypotheses on which is defined a diffusion process $\{X_t,\, t \geq 0\}$ satisfying the Itô stochastic differential equation

$$dX_t = f(\theta, t, X_t)\, dt + dW_t \;, \qquad t \geq 0, \quad X_0 = 0 \tag{2.1}$$

where $\{W_t,\, t\geq 0\}$ is a standard Wiener process, $f(\theta, t, x)$ is a known real valued function continuous on $\Theta \times [0, T] \times \mathbb{R}$ where Θ is a closed interval of the real line and the parameter θ is unknown, which is to be estimated on the basis of observation of the process $\{X_t,\, 0 \leq t \leq T\} =: X_0^T$. Let θ_0 be the true value of the parameter which lies inside the parameter space Θ.

Let P_θ^T be the measure generated by the process X_0^T on the space (C_T, B_T) of continuous functions on $[0, T]$ with the associated Borel σ-algebra B_T associated to the sup-norm topology of C_T. Let E_θ^T be the expectation with respect to the measure P_θ^T. Suppose P_θ^T is absolutely continuous with respect to $P_{\theta_0}^T$. Then it is well known that (see Liptser and Shiryayev (1977, p. 239)

$$\begin{aligned} L_T(\theta) \;&:=\; \frac{dP_\theta^T}{dP_{\theta_0}^T}(X_0^T) \\ &=\; \exp\left\{ \int_0^T \Big[f(\theta, s, X_s) - f(\theta_0, s, X_s) \Big]\, dW_s \right. \\ &\quad \left. -\, \frac{1}{2} \int_0^T \Big[f(\theta, s, X_s) - f(\theta_0, s, X_s) \Big]^2 ds \right\} . \end{aligned} \tag{2.2}$$

is the Radon–Nikodym derivative (likelihood) of P_θ^T with respect to $P_{\theta_0}^T$. The MLE θ_T of θ based on X_0^T is defined as

$$\theta_T := \arg\max_{\theta \in \Theta} L_T(\theta) \;.$$

Throughout the paper prime denotes derivative with respect to θ. Let us denote the log-likelihood function by $l_t(\theta) \equiv \log L_T(\theta)$, and let $l_t'(\theta) \equiv U_T(\theta)$, $l_T''(\theta) \equiv H_T(\theta)$ and $l_T'''(\theta) \equiv Q_T(\theta)$.

If $L_T(\theta)$ is continuous in θ, it can be shown that there exists a measurable MLE by using Lemma 3.3 in Schmetterer (1974). Hereafter, we assume the existence of such a measurable MLE. We assume the following regularity conditions on $f(\theta, t, x)$.

(**A1**) $P_{\theta_1} \neq P_{\theta_2}$ for $\theta_1 \neq \theta_2$ in Θ.

(**A2**) $\{X_t\}$ is the unique strong solution of (2.1) with

$$P_\theta\left(\int_0^T f^2(\theta, t, X_t)\, dt < \infty\right) = 1 \qquad \text{for all } \theta \in \Theta, \quad T < \infty \; . \tag{2.3}$$

The condition (2.3) ensures that $P_\theta^T \ll P_W^T$ for all θ where P_W^T is the standard Wiener measure and likelihood function is given by

$$\frac{dP_\theta^T}{dP_W^T} = \exp\left\{\int_0^T f(\theta, t, X_t)\, dX_t - \frac{1}{2}\int_0^T f^2(\theta, t, X_t)\, dt\right\} . \tag{2.4}$$

(**A3**) (**i**) $f(\theta, t, x)$ is differentiable in t and x.
The log-likelihood with respect to P_W^T can be written as

$$\log \frac{dP_\theta^T}{dP_W^T} = \int_0^T f(\theta, t, X_t)\, dX_t - \frac{1}{2}\int_0^T f^2(\theta, t, X_t)\, dt \; . \tag{2.5}$$

(**ii**) The integrals in (2.4) and (2.5) can be differentiated twice under the integral sign with respect to θ.
Let $I_T(\theta) := \int_0^T f'^2(\theta, t, X_t)\, dt$ and $Y_T(\theta) := \int_0^T f''^2(\theta, t, X_t)\, dt$.

(**iii**) l''_T is continuous in a neighborhood V_θ of θ for every $\theta \in \Theta$ and

$$n_T = n_T(\theta) := E_\theta(I_T(\theta)) < \infty \; , \qquad E_\theta(Y_T(\theta)) < \infty$$

with $n_T \to \infty$ as $T \to \infty$ and there exists a constant C_0 such that for any $\theta, \theta_1, \theta_2 \in \Theta$

$$\frac{E_\theta(I_T(\theta_2))}{n_T(\theta_1)} < C_0 \; .$$

(**iv**) $\dfrac{I_T(\theta)}{n_T} \xrightarrow{P_\theta} 1$ as $T \to \infty$.

(**A4**) Suppose there exists $\gamma \geq 2$ and $C > 0$ such that for all $\theta \in \Theta$

$$E_\theta \exp\left\{-\frac{1}{3}\int_0^T \Big[f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)\Big]^2 dt\right\} \leq \exp\big(-C|u|^\gamma\big) \, .$$

(**A5**) Suppose that there exists $m_T = m_T(\theta) \uparrow \infty$ as $T \to \infty$ such that

(**i**) $\dfrac{I_T(\theta)}{m_T} \xrightarrow{P_\theta} \eta(\theta)$ as $T \to \infty$ where $P_\theta\big(\eta(\theta) > 0\big) > 0$ and $E(\eta^{-1}(\theta)) < \infty$.

(**ii**) $\dfrac{Y_T(\theta)}{m_T} \xrightarrow{P_\theta} \xi(\theta)$ as $T \to \infty$.

Some of regularity conditions (A1)–(A5) can be found in the literature, for example, Borkar and Bagchi (1982) and Levanony *et al.* (1994). However both proved strong consistency and Levanony *et al.* (1994) proved asymptotic normality. We need stronger regularity condition (A4) in order to prove large deviations.

Let us introduce the Bayes estimator. Let Λ be a prior probability measure on $(\Theta, \mathcal{B})$ where $\mathcal{B}$ is the σ-algebra of Borel subsets of Θ. Suppose that Λ has a density $\lambda(\cdot)$ with respect to the Lebesgue measure on $\mathbb{R}$, which is continuous and positive on Θ and possesses a polynomial majorant in Θ.

Let $p(\theta|X_0^T)$ be the posterior density of θ given X_0^T. By Bayes theorem $p(\theta|X_0^T)$ is given by

$$p(\theta|X_0^T) = \frac{L_T(\theta)\,\lambda(\theta)}{\int_\Theta L_T(\theta)\,\lambda(\theta)\,d\theta} .$$

Let $l(\cdot,\cdot)\colon \Theta\times\Theta \to \mathbb{R}$ be a loss function as defined in Ibragimov and Khasminskii (1981) which satisfies the following conditions:

(**B1**) $\psi(u,v) = \psi(u-v)$.

(**B2**) $\psi(u)$ is defined and nonnegative on $\mathbb{R}, \psi(0) = 0$ and $\psi(u)$ is continuous at $u = 0$ but is not identically equal to 0.

(**B3**) ψ is symmetric, i.e., $\psi(u) = \psi(-u)$.

(**B4**) $\{u\colon \psi(u) < c\}$ are convex sets and are bounded for all $c > 0$ sufficiently small.

(**B5**) There exists numbers $\gamma > 0$, $h_0 \geq 0$ such that for $h \geq h_0$

$$\sup\Big\{\psi(u)\colon\ |u| \leq h^\gamma\Big\} \leq \inf\Big\{\psi(u)\colon\ |u| \geq h\Big\} .$$

Clearly, all power loss functions of the form $|u-v|^r$, $r > 0$, satisfy the condition (B1)–(B5). In particular, quadratic loss function $|u-v|^2$ satisfies these conditions.

A Bayes estimator $\widetilde{\theta}_T$ of θ with respect to the loss function $\psi(\theta,\phi)$ and prior density $\lambda(\theta)$ is one which minimizes the posterior risk and is given by

$$\widetilde{\theta}_T := \arg\min_{\phi\in\Theta} \int_\Theta l(\phi,\theta)\,p(\theta|X_0^T)\,d\theta .$$

In particular, for the quadratic loss function $\psi(u,v) = |u-v|^2$, the Bayes estimator $\widetilde{\theta}_T$ becomes the posterior mean given by

$$\widetilde{\theta}_T = \frac{\int_\Theta \phi\ p(\phi|X_0^T)\,du}{\int_\Theta p(\phi|X_0^T)\,d\phi} .$$

Let us consider the likelihood ratio process

$$Z_T(u) := \frac{dP_{\theta+un_T^{-1/2}}}{dP_\theta}(X_0^T) .$$

By (2.2) with $g_t(u) := f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)$, we have

$$\begin{aligned} Z_T(u) &= \exp\Bigg\{ \int_0^T \Big[f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)\Big]\, dW_t \\ &\quad - \frac{1}{2}\int_0^T \Big[f\big(\theta + u n_T^{-1/2}, t, X_t\big) - f(\theta, t, X_t)\Big]^2 dt \Bigg\} \\ &= \exp\left\{ \int_0^T g_t(u)\, dW_t - \frac{1}{2}\int_0^T g_t^2(u)\, dt \right\} . \end{aligned}$$

We define the LAMN condition below:

Definition (Le Cam and Yang (1990), Jeganathan (1982, 1995))**.** Let $\mathcal{E}_n = \big(\Omega_n, \mathcal{F}_n, (P_\theta^n, \theta \in \Theta)\big)$, $n \geq 1$, be a sequence of statistical experiments, where Θ is an open subset of $\mathbb{R}$. We denote by

$$\Lambda_{\eta\theta}^n = \log\left(\frac{dP_\eta^n}{dP_\theta^n}\right)$$

the log-likelihood between η and θ at stage n.

We say that the sequence $\mathcal{E}_n$ satisfies the *local asymptotically quadratic* (LAQ) condition at a point $\theta \in \Theta$ if there are random variables Δ_n and Γ_n defined on $(\Omega_n, \mathcal{F}_n)$, $\Gamma_n > 0$ a.s. $[P_\theta^n]$ and a positive numerical sequence $\phi_n \downarrow 0$ such that for each bounded sequence of numbers u_n,

$$\Lambda_{\theta+\phi_n u_n, \theta}^n - \Big(u_n \Delta_n - \frac{1}{2} u_n^2 \Gamma_n\Big) \overset{P_\theta^n}{\to} 0$$

and

$$(\Delta_n, \Gamma_n) \to (\Delta, \Gamma) \qquad \text{in } \ P_\theta^n\text{-distribution}$$

where Δ and Γ are random variables on a measurable space $(\Omega, \mathcal{F}, P)$ with $\Gamma > 0$ a.s. (P) and

$$E_P \exp\Big(u\Delta - \frac{1}{2}u^2\Gamma\Big) = 1 \ .$$

The sequence of experiments is called *locally asymptotically Brownian functional* (LABF) if $\Delta = \int_0^1 F_s\, dW_s$ and $\Gamma = \int_0^1 F_s^2\, ds$ with W a standard Brownian motion and F a predictable process with respect to some filtration in $\mathcal{F}$. It is called *locally asymptotically mixed normal* (LAMN) if $\Delta = \Gamma^{1/2} W_1$ with W_1 standard normal variable independent of Γ and *locally asymptotically normal* (LAN) if, in addition, Γ is nonrandom.

Let $\Phi(.)$ denote the standard normal distribution function and C denote a generic positive constant. We shall use the following lemmas to prove our main results. The first lemma is a revised version of Theorem 19 of Ibragimov and Khasminskii (1981, p. 372) from Kallianpur and Selukar (1993, p. 330).

Lemma 2.1. *Let $\zeta(t)$ be a real valued random function defined on a closed subset F of the Euclidean space $\mathbb{R}^k$. We shall assume that the random process $\zeta(t)$ is measurable and separable. Assume that the following condition is fulfilled: there exist numbers $m \geq r > k$ and a function $H(x)\colon \mathbb{R}^k \to \mathbb{R}^1$ bounded on compact sets such that for all $x, h \in F$, $x+h \in F$,*

$$\begin{aligned} E\,|\zeta(x)|^m &\leq H(x)\ , \\ E\big|\zeta(x+h) - \zeta(x)\big|^m &\leq H(x)\,|h|^r\ . \end{aligned}$$

Then with probability one the realizations of $\zeta(t)$ are continuous functions of F. Moreover, set

$$w(\delta, \zeta, L) \;=\; \sup_{\substack{x,y\in F \\ |x|,|y|\leq L \\ |x-y|\leq\delta}} \big|\zeta(x) - \zeta(y)\big|\ ,$$

then

$$E\big(w(h;\zeta,L)\big) \;\leq\; B_0\Big(\sup_{|x|<L} H(x)\Big)^{1/m} L^k\, h^{(r-k)/m} \log(h^{-1})$$

where $B_0 = 64^k\big(1 - 2^{-(r-k)m}\big)^{-1} + \big(2^{(m-r)/m} - 1\big)^{-1}$.

Lemma 2.2 (Ibragimov and Khasminskii (1981, p. 45))**.** *Let $Z_{\epsilon,\theta}(u)$ be the likelihood ratio function corresponding to the points $\theta + \phi(\epsilon)u$ and θ where $\phi(\epsilon)$ denotes a normalizing factor such that $|\phi(\epsilon)| \to 0$ as $\epsilon \to 0$. Thus $Z_{\epsilon,\theta}$ is defined on the set $U_\epsilon = (\phi(\epsilon))^{-1}(\Theta - \theta)$. Let $Z^{\theta}_{\epsilon,0}(u)$ possesses the following properties: given a compact set $K \subset \Theta$ there exist numbers $M_1 > 0$ and $m_1 \geq 0$ and functions $g^K_\epsilon(y) = g_\epsilon(y)$ correspond such that*

(**1**) *For some $\alpha > 0$ and all $\theta \in K$,*

$$\sup_{\substack{|u_1|\leq R \\ |u_2|\leq R}} |u_2 - u_1|^{-\alpha}\, E^{(\epsilon)}_\theta \big|Z^{1/2}_{\epsilon,\theta}(u_2) - Z^{1/2}_{\epsilon,\theta}(u_1)\big|^2 \;\leq\; M_1(1+R^{m_1})\ .$$

(**2**) *For all $\theta \in K$ and $u \in U_\epsilon$, $E^{(\epsilon)}_\theta Z^{1/2}_{\epsilon,\theta}(u) \leq e^{-g_\epsilon(u)}$.*

(**3**) *$g_\epsilon(u)$ is a monotonically increasing to ∞ function of y*

$$\lim_{\substack{y\to\infty \\ \epsilon\to 0}} y^N e^{-g_\epsilon(y)} \;=\; 0\ .$$

Let $\{\widetilde{\theta}_\epsilon\}$ be a family of Bayes estimators with respect to the prior density q, which is continuous and positive on K and possesses in Θ a polynomial majorant and a loss function $\omega_\epsilon(u,v) := \psi\big((\phi(\epsilon))^{-1}(u-v)\big)$ where ψ satisfies (B1)–(B5). Then for all N,

$$\lim_{\substack{h\to\infty \\ \epsilon\to 0}} h^N \sup_{\theta \epsilon K} P^{(\epsilon)}_\theta \Big\{ \big|(\phi(\epsilon))^{-1}(\widetilde{\theta}_\epsilon - \theta)\big| > h \Big\} \;=\; 0\ .$$

If in addition, $\psi(u) = \tau(|u|)$, then for all ϵ sufficiently small, $0 < \epsilon < \epsilon_0$,

$$\sup_{\theta\in K} P_\theta^{(\epsilon)}\Big\{\big|(\phi(\epsilon))^{-1}(\widetilde{\theta}_\epsilon - \theta)\big| > h\Big\} \;\le\; B_0\, e^{-b_0 g_\epsilon(h)} \;.$$

Lemma 2.3. *Under the assumptions (A1)–(A4),*

(**a**) $\displaystyle\sup_{\theta\in\Theta} E_\theta^T\Big[Z_T^{1/2}(u_1) - Z_T^{1/2}(u_2)\Big]^2 \le \frac{C_0^2}{4}\,(u_2-u_1)^2\,;$

(**b**) $\displaystyle\sup_{\theta\in\Theta} E_\theta^T\Big[Z_T^{1/2}(u)\Big] \le C\exp(-C|u|^\gamma)\,.$

Proof: Observe that

$$\begin{aligned} &E_\theta^T\Big[Z_T^{1/2}(u_1) - Z_T^{1/2}(u_2)\Big]^2 = \\ &\qquad = E_\theta^T\big[Z_T(u_1)\big] + E_\theta^T\big[Z_T(u_2)\big] - 2\,E_\theta^T\Big[Z_T^{1/2}(u_1)\,Z_T^{1/2}(u_2)\Big] \\ &\qquad \le 2 - 2\,E_\theta^T\Big[Z_T^{1/2}(u_1)\,Z_T^{1/2}(u_2)\Big]\,. \end{aligned} \tag{2.6}$$

From Gikhman and Skorohod (1972, p. 82), for all u, we have

$$E_\theta^T\big[Z_T(u)\big] \;=\; E_\theta^T\left[\exp\left\{\int_0^T g_t(u)\,dW_t - \frac{1}{2}\int_0^T g_t^2(u)\,dt\right\}\right] \le 1\;. \tag{2.7}$$

Let

$$\begin{aligned} &\theta_1 := \theta + u_1 n_T^{-1/2}\;, \qquad \theta_2 := \theta + u_2\, n_T^{-1/2}\,, \\ &\delta_t := f(\theta_2, t, X_t) - f(\theta_1, t, X_t)\;, \\ &J(\theta_1,\theta_2) := E_{\theta_1}\big(I_T(\theta_2)\big)\;, \\ &V_T := \exp\left\{\frac{1}{2}\int_0^T \delta_t\,dW_t - \frac{1}{4}\int_0^T \delta_t^2\,dt\right\} = \left(\frac{dP_{\theta_2}^T}{dP_{\theta_1}^T}\right)^{1/2}\,. \end{aligned} \tag{2.8}$$

By Itô formula, V_T can be represented as

$$V_T \;=\; 1 - \frac{1}{8}\int_0^T V_t\,\delta_t^2\,dt + \frac{1}{2}\int_0^T V_t\,\delta_t\,dW_t\;. \tag{2.9}$$

The random process $\big\{V_t^2, \mathcal{F}_t, P_\theta^T,\, 0 \le t \le T\big\}$ is a martingale and from the

$\mathcal{F}_t$-measurability of δ_t for each $t \in [0,T]$,

$$
\begin{aligned}
E^T_{\theta_1} \int_0^T V_t^2\, \delta_t^2\, dt &= E^T_{\theta_1} \int_0^T E^T_{\theta_1}(V_t^2 \,|\, \mathcal{F}_t)\, \delta_t^2\, dt \\
&= E^T_{\theta_1} V_T^2 \int_0^T \delta_t^2\, dt \\
&= \int V_T^2 \left(\int_0^T \delta_t^2\, dt \right) dP_{\theta_1} \\
&= \int \left(\int_0^T \delta_t^2\, dt \right) dP^T_{\theta_2} \qquad (2.10)\\
&= E^T_{\theta_2} \left(\int_0^T \delta_t^2\, dt \right) \\
&= E^T_{\theta_2} \int_0^T \big| f(\theta_2, t, X_t) - f(\theta_1, t, X_t) \big|^2\, dt \\
&= E^T_{\theta_2} \int_0^T \left(\int_{\theta_1}^{\theta_2} f'(y, t, X_t)\, dy \right)^2 dt \qquad \text{(by (A1))} \\
&\le (\theta_2 - \theta_1)\, E^T_{\theta_2} \int_0^T \int_{\theta_1}^{\theta_2} f'^2(y, t, X_t)\, dy\, dt \qquad (2.11)\\
&= (\theta_2 - \theta_1) \int_{\theta_1}^{\theta_2} J(\theta_2, y)\, dy \;<\; \infty\, .
\end{aligned}
$$

Hence $E^T_{\theta_1} \int_0^T V_t\, \delta_t\, dW_t = 0$. Therefore, using $|a\,b| \le \frac{a^2+b^2}{2}$, we obtain from (2.10)

$$
\begin{aligned}
E^T_{\theta_1}(V_T) &= 1 - \frac{1}{8} \int_0^T E^T_{\theta_1}(\delta_t V_t \,.\, \delta_t)\, dt \\
&\ge 1 - \frac{1}{16} \int_0^T E^T_{\theta_1} \delta_t^2\, dt - \frac{1}{16} \int_0^T E^T_{\theta_1} V_t^2\, \delta_t^2\, dt \qquad (2.12)\\
&= 1 - \frac{1}{16}\, E^T_{\theta_1} \int_0^T \delta_t^2\, dt - \frac{1}{16}\, E^T_{\theta_2} \int_0^T \delta_t^2\, dt \qquad \text{(by (2.11))}\, .
\end{aligned}
$$

Now

$$
\begin{aligned}
E^T_\theta \Big[Z_T^{1/2}(u_1)\, Z_T^{1/2}(u_2) \Big] &= E^T_\theta \left[\frac{dP^T_{\theta + u_1 n_T^{-1/2}}}{dP^T_\theta} \right]^{1/2} \left[\frac{dP^T_{\theta + u_2 n_T^{-1/2}}}{dP^T_\theta} \right]^{1/2} \\
&= \int \left[\frac{dP^T_{\theta_1}}{dP^T_\theta} \right]^{1/2} \left[\frac{dP^T_{\theta_2}}{dP^T_\theta} \right]^{1/2} dP^T_\theta \qquad (2.13)\\
&= \int \left[\frac{dP^T_{\theta_2}}{dP^T_{\theta_1}} \right]^{1/2} dP^T_{\theta_1} \;=\; E^T_{\theta_1}(V_T)\, .
\end{aligned}
$$

Substituting (2.13) into (2.6) and using (2.12), we obtain

$$
\begin{aligned}
E_\theta\Big[Z_T^{1/2}(u_1) - Z_T^{1/2}(u_2)\Big]^2 \le \\
&\le\ 2 - 2\,E_{\theta_1}(V_T) \\
&\le\ \frac{1}{8}\,E_{\theta_1}\int_0^T \delta_t^2\,dt + \frac{1}{8}\,E_{\theta_2}\int_0^T \delta_t^2\,dt \\
&\le\ \frac{1}{8}\,(\theta_2-\theta_1)\int_{\theta_1}^{\theta_2}\Big[J(\theta_1,y)+J(\theta_2,y)\Big]\,dy \\
&\qquad\qquad \text{(by using arguments similar to (2.10))} \\
&\le\ \frac{1}{4}\,(\theta_2-\theta_1)^2\,\sup_{\theta,y} J(\theta,y) \\
&=\ \frac{(u_2-u_1)^2}{4\,n_T}\,\sup_{\theta,y} J(\theta,y) \\
&\le\ \frac{C_0}{4}\,(u_2-u_1)^2 \qquad \text{(by (A3)(iii))}\ .
\end{aligned}
$$

This completes the proof of (a).

Let us now prove (b). By Hölder inequality,

$$
\begin{aligned}
E_\theta\Big[Z_T^{1/2}(u)\Big] \ = \\
&=\ E_\theta\left[\exp\left\{\frac{1}{2}\int_0^T g_t(u)\,dW_t - \frac{1}{4}\int_0^T g_t^2(u)\,dt\right\}\right] \\
&=\ E_\theta\left[\exp\left\{\frac{1}{2}\int_0^T g_t(u)\,dW_t - \frac{1}{6}\int_0^T \big(g_t(u)\big)^2 dt\right\}\exp\left\{-\frac{1}{12}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right] \\
&\le\ \left\{E_\theta\left[\exp\left\{\frac{1}{2}\int_0^T g_t(u)\,dW_t - \frac{1}{6}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right]^{4/3}\right\}^{3/4} \qquad (2.14) \\
&\quad\times\left\{E_\theta\left[\exp\left\{-\frac{1}{12}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right]^4\right\}^{1/4} \\
&\le\ \left[E_\theta\,\exp\left\{\frac{2}{3}\int_0^T g_t(u)\,dW_t - \frac{2}{9}\int_0^T \big(g_t^2(u)\big)\,dt\right\}\right]^{3/4} \\
&\quad\times\left[E_\theta\,\exp\left\{-\frac{1}{3}\int_0^T \big(g_t(u)\big)^2 dt\right\}\right]^{1/4}.
\end{aligned}
$$

Assumption (A5) implies that

$$
\begin{aligned}
E\,\exp\left\{-\frac{1}{3}\int_0^T \big(g_t(u)\big)^2 dt\right\} = \\
&=\ E\,\exp\left\{-\frac{1}{3}\int_0^T \Big[f\big(\theta+u n_T^{-1/2},t,X_t\big) - f(\theta,t,X_t)^2\Big]^2 dt\right\} \qquad (2.15) \\
&\le\ \exp\big(-C\,|u|^\gamma\big)\ .
\end{aligned}
$$

On the other hand, from Gikhman and Skorohod (1972, p. 82)

$$E_\theta\left[\exp\left\{\int_0^T \frac{2}{3}\,g_t(u)\,dW_t-\frac{1}{2}\int_0^T\Big(\frac{2}{3}\,g_t(u)\Big)^2 dt\right\}\right]\le 1\;. \tag{2.16}$$

Combination of (2.14)–(2.16) completes the proof of (b). □

Lemma 2.4 (Michel and Pfanzagl (1971)). *Let Y and Z be two random variables on some probability space with $P(Z>0)=1$. Then for all $\epsilon>0$, we have*

$$\sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y}{Z}\le x\right\}-\Phi(x)\right| \;=\; \sup_{x\in\mathbb{R}}\Big|P\big\{Y\le x\big\}-\Phi(x)\Big| \;+\; P\big\{|Z-1|>\epsilon\big\} \;+\; \epsilon\;.$$

The following is the generalization of the above lemma from non-random η to random η.

Lemma 2.5 (Oblakova (1989)). *Let Y, Z and η be three random variables on some probability space with $P(Z>0)=1$, and η is a positive random variable with $P\{0<\eta^2<\infty\}=1$, $E(\eta^{-1})<\infty$. Then for all $\epsilon>0$, we have*

$$\sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y}{Z}\le x\right\}-\Phi(x)\right| = E\sup_{x\in\mathbb{R}}\Big|P\big\{Y\le x\,|\,\mathcal{G}\big\}-\widetilde{\Phi}(x)\Big|+2P\big\{|Z-\eta|>\epsilon\big\}+\epsilon E(\eta^{-1})\;.$$

where $\widetilde{\Phi}(x)=P(\zeta\eta\le x\,|\,\eta)$, $\mathcal{G}=\sigma(\eta)\subset\mathcal{F}_0$ *and* ζ *is* $\mathcal{N}(0,1)$ *random variable independent of* η.

3. MAIN RESULTS

We obtain the following large deviations upper bound for the MLE.

Theorem 3.1. *Under the assumptions (A1)–(A4), for $\rho>0$, we have*

$$\sup_{\theta\in\Theta} P_\theta^T\left\{n_T^{1/2}\,|\theta_T-\theta|\ge\rho\right\}\;\le\; B\exp\big(-b|\rho|^\gamma\big)$$

for some positive constants b and B independent of ρ and T.

Proof: Let

$$S_T := \left\{u\colon\; \theta+u\,n_T^{-1/2}\in\Theta\right\}\,,$$

$$
\begin{aligned}
P_\theta^T\left\{n_T^{1/2}|\theta_T-\theta|>\rho\right\} &= P_\theta^T\left\{|\theta_T-\theta|>\rho\, n_T^{-1/2}\right\}\\
&\le P_\theta^T\left\{\sup_{\substack{|u|\ge\rho\\ u\in S_T}} L_T\big(\theta+uT^{-1/2}\big)\ge L_T(\theta)\right\}\\
&= P_\theta^T\left\{\sup_{|u|\ge\rho}\frac{L_T\big(\theta+uT^{-1/2}\big)}{L_T(\theta)}\ge 1\right\} \qquad (3.1)\\
&= P_\theta^T\left\{\sup_{|u|\ge\rho} Z_T(u)\ge 1\right\}\\
&\le \sum_{r=0}^{\infty} P_\theta^T\left\{\sup_{u\in\Gamma_r} Z_T(u)\ge 1\right\},
\end{aligned}
$$

where $\Gamma_r=[\rho+r,\rho+r+1]$. Applying Lemma 2.1 with $\zeta(u)=Z_T^{1/2}(u)$, we obtain from Lemma 2.3 that there exists a constant $B>0$ such that

$$
\sup_{\theta\in\Theta} E_\theta^T\left\{\sup_{\substack{|u_1-u_2|\le h\\ |u_1|,|u_2|\le l}}\left[Z_T^{1/2}(u_1)-Z_T^{1/2}(u_2)\right]\right\}\le B\, l^{1/2}\, h^{1/2}\,\log h^{-1}\,. \tag{3.2}
$$

Divide Γ_r into subintervals of length at most $h>0$. The number n of subintervals is clearly less than or equal to $[\frac{1}{h}]+1$. Let $\Gamma_r^{(j)}$, $1\le j\in n$ be the subintervals chosen. Choose $u_j\in\Gamma_r^{(j)}$. Then

$$
\begin{aligned}
&P_\theta^T\left[\sup_{u\in\Gamma_r} Z_T(u)\ge 1\right]\le\\
&\quad\le \sum_{j=1}^{n} P_\theta^T\left[Z_T^{1/2}(u_j)\ge\frac12\right]+P_\theta^T\left\{\sup_{\substack{|u-v|\le h\\ |u|,|v|\le\rho+r+1}}\left|Z_T^{1/2}(u)-Z_T^{1/2}(v)\right|\ge\frac12\right\}\\
&\quad\le 2\sum_{j=1}^{n} E_\theta^T\left[Z_T^{1/2}(u_j)\right]+2\,B(\rho+r+1)^{1/2}\,h^{1/2}\,\log(h^{-1})\\
&\qquad\qquad\text{(by Markov inequality and (3.2))}\\
&\quad\le 2\,C\sum_{j=1}^{n}\exp\big(-C\,|u_j|^\gamma\big)+2\,B(\rho+r+1)^{1/2}\,h^{1/2}\,\log(h^{-1})\quad\text{(by Lemma 2.2)}\\
&\quad\le 2\,C\left(\left[\frac1h\right]+1\right)\exp\big\{-C(\rho+r)^\gamma\big\}+2\,B(\rho+r+1)^{1/2}\,h^{1/2}\,\log(h^{-1})\,.
\end{aligned}
$$

Let us now choose $h=\exp\left\{\frac{-C(\rho+r)^\gamma}{2}\right\}$. Then

$$
\begin{aligned}
\sup_{\theta\in\Theta} P_\theta^T\left\{\sup_{u>\rho} Z_T(u)\ge 1\right\} &\le B\sum_{r=0}^{\infty}(\rho+r+1)^{1/2}\exp\left\{\frac{-C(\rho+r)^\gamma}{4}\right\}\\
&\le B\exp(-b\rho^\gamma)\,,
\end{aligned}\tag{3.3}
$$

where B and b are positive generic constants independent of ρ and T. Similarly it can be shown that

$$
\sup_{\theta\in\Theta} P_\theta^T\left[\sup_{u<-\rho} Z_T(u)\ge 1\right]\le B\exp(-b\rho^\gamma)\,. \tag{3.4}
$$

Combining (3.3) and (3.4), we obtain

$$\sup_{\theta\in\Theta} P_\theta^T \left[\sup_{|u|>\rho} Z_T(u) \geq 1 \right] \leq B \exp(-b\rho^\gamma) \ . \tag{3.5}$$

The theorem follows from (3.2) and (3.5). $\square$

By substituting $\rho = n_T^{1/2}\epsilon$ in Theorem 3.1, the following result is obtained.

Corollary 3.1. *Under the conditions of Theorem 3.1, for arbitrary $\epsilon > 0$ and all $T > 0$, we have*

$$\sup_{\theta\in\Theta} P_\theta^T \left\{ |\theta_T - \theta| > \epsilon \right\} \leq B \exp(-b\, n_T\, \epsilon^\gamma)$$

where B and b are positive constants independent of ϵ and T.

We obtain the following large deviations bound for the Bayes estimator $\widetilde{\theta}_T$.

Theorem 3.2. *Suppose (A1)–(A4) and (B1)–(B5) hold. For $\rho > 0$, the Bayes estimator $\widetilde{\theta}_T$ with respect to the prior $\lambda(\cdot)$ and a loss function $l(\cdot,\cdot)$ with $l(u) = l(|u|)$ satisfies*

$$\sup_{\theta\in\Theta} P_\theta^T \left\{ \sqrt{T} \left| \widetilde{\theta}_T - \theta \right| \geq \rho \right\} \leq B \exp(-b\rho^2)$$

for some positive constants B and b independent of ρ and T.

Proof: Using Lemma 2.3, conditions (1), (2) and (3) of Lemma 2.2 are satisfied with $\alpha = 2$ and $g(u) = u^2$. Hence the result follows from Lemma 2.2. $\square$

Corollary 3.2. *Under the conditions of Theorem 3.3, for arbitrary $\epsilon > 0$ and all $T > 0$, we have*

$$\sup_{\theta\in\Theta} P_\theta^T \left\{ \left| \widetilde{\theta}_T - \theta \right| > \epsilon \right\} \leq B \exp(-C\, T\, \epsilon^2) \ .$$

As another application of Theorem 3.3 we obtain the following result.

Theorem 3.3. *Under the assumptions (A1)–(A4), for all N, we have for the Bayes estimator $\widetilde{\theta}_T$ with respect to the prior $\lambda(\cdot)$ and loss function $\psi(\cdot,\cdot)$ satisfying the conditions (B1)–(B5),*

$$\lim_{\substack{H\to\infty \\ T\to\infty}} H^N \sup_{\theta\epsilon\Theta} P_\theta^T \left\{ \sqrt{T} \left| \widetilde{\theta}_T - \theta \right| > H \right\} = 0 \ .$$

We establish the following Berry–Esseen type inequality for the MLE.

Theorem 3.4. *Under the assumptions (A2), (A3) and (A5),*

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\Big|P_\theta\Big\{I_T^{-1/2}(\theta)\,(\theta_T-\theta)\le x\Big\}-\Phi(x)\Big| \le \\
&\quad\le E_\theta^{1/3}\left|\frac{I_T(\theta)}{m_T}-\eta(\theta)\right| + P_\theta\left\{\left|\frac{H_T(\theta)}{I_T}-1\right|>\frac{\epsilon_T}{2}\right\} + P_\theta\left\{\sup_{\theta\in\Theta}\big|I_T^{-1}(Q_T(\theta))\big|>\frac{\epsilon_T}{2\delta}\right\} \\
&\qquad + C\exp\big(-b\,n_T\,\delta^2\big)\ .
\end{aligned}$$

for any $\delta>0$ and some $\epsilon_T\downarrow 0$ as $T\to\infty$ and $b>0$ is a constant independent of T.

Proof: Recall that $l_t'(\theta)\equiv U_T(\theta)$, $l_T''(\theta)\equiv H_T(\theta)$ and $l_T'''(\theta)\equiv Q_T(\theta)$.

By a Taylor expansion of $U_T(\theta)$ around θ, we have

$$0 = U_T(\theta_T) = U_T(\theta) + (\theta_T-\theta)\,H_T(\overline{\theta}_T) \qquad \text{where } \ |\overline{\theta}_T-\theta| < |\theta_T-\theta|\ .$$

Hence

$$\begin{aligned}
I_T^{1/2}(\theta)\,(\theta_T-\theta) &= -I_T^{1/2}(\theta)\,\frac{U_T(\theta)}{H_T(\overline{\theta}_T)} \\
&= -\left(\frac{I_T(\theta)}{m_T}\right)^{1/2}\frac{m_T^{-1/2}\,U_T(\theta)}{m_T^{-1}\,H_T(\overline{\theta}_T)}\ .
\end{aligned}$$

Thus by Lemma 2.3, we have

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\Big|P_\theta\Big\{I_T^{1/2}(\theta)\,(\theta_T-\theta)\le x\Big\}-\Phi(x)\Big| = \\
&\qquad = \sup_{x\in\mathbb{R}}\left|P_\theta\left\{\frac{-\left(\frac{I_T(\theta)}{m_T}\right)^{-1/2}m_T^{-1/2}\,U_T(\theta)}{\left(\frac{I_T(\theta)}{m_T}\right)^{-1}m_T^{-1}\,H_T(\overline{\theta}_T)}\le x\right\}-\Phi(x)\right| \\
&\qquad = \sup_{x\in\mathbb{R}}\left|P_\theta\left\{\frac{-m_T^{-1/2}\,U_T(\theta)}{\left(\frac{I_T(\theta)}{m_T}\right)^{1/2}}\le x\right\}-\Phi(x)\right| \\
&\qquad\quad + P_\theta\Big\{\big|I_T^{-1}(\theta)\,H_T(\overline{\theta}_T)-1\big|>\epsilon_T\Big\}+\epsilon_T \\
&\qquad =: J_1+J_2+\epsilon_T\ .
\end{aligned}$$

Let $M_T(\theta) = -m_T^{-1/2}\,U_T(\theta) = m_T^{-1/2}\int_0^T f'(\theta,t,X_t)\,dW_t$, a normalized continuous martingale with respect to $\mathcal{F}_T$ and $\langle M(\theta)\rangle_T = m_T^{-1}I_T(\theta) = m_T^{-1}\int_0^T f'^2(\theta,t,X_t)\,dt$ be its corresponding increasing process. Let $\widetilde{\Phi}(x) = P\big(G\eta\le x|\eta\big)$, $G\sim\mathcal{N}(0,1)$

and $\mathcal{G} \equiv \sigma(\eta) \subset \mathcal{F}_0$. Then by Lemma 2.5

$$
\begin{aligned}
J_1 &= \sup_{x\in\mathbb{R}} \left| P_\theta \left\{ \frac{-m_T^{-1/2}\, U_T(\theta)}{\left(\frac{I_T(\theta)}{m_T}\right)^{1/2}} \le x \right\} - \Phi(x) \right| \\
&= \sup_{x\in\mathbb{R}} \left| P_\theta \left\{ \frac{M_T(\theta)}{\sqrt{\langle M(\theta)\rangle_T}} \le x \right\} - \Phi(x) \right| \\
&\le E_\theta \sup_{x\in\mathbb{R}} \left| P_\theta\big(M_T(\theta) \le x \,|\, \mathcal{G}\big) - \widetilde{\Phi}(x) \right| \\
&\quad + 2\, P_\theta \left\{ \left| \sqrt{\langle M(\theta)\rangle_T} - \eta(\theta) \right| > \epsilon_T \right\} + \epsilon_T\, E\big(\eta^{-1}(\theta)\big) \\
&\le C_1\, E_\theta^{1/3} \left| \frac{I_T(\theta)}{m_T} - \eta(\theta) \right| \qquad \text{(by Lemma 2.5)}
\end{aligned}
$$

where C_1 depends only on $E(\eta^{-1}(\theta))$. Further,

$$
\begin{aligned}
J_2 &= P_\theta \left\{ \left| I_T^{-1} H_T(\overline{\theta}_T) - 1 \right| > \epsilon_T \right\} \\
&\le P_\theta \left\{ \left| I_T^{-1} \right| \left| H_T(\overline{\theta}_T) - H_T(\theta) \right| > \frac{\epsilon_T}{2} \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&= P_\theta \left\{ I_T^{-1} \left| (\overline{\theta}_T - \theta)\, Q_T(\theta_T^\star) \right| > \frac{\epsilon_T}{2} \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \big(\text{where } \ |\theta_T^\star - \theta| < |\overline{\theta}_T - \theta|\big) \\
&\le P_\theta \left\{ \left| I_T^{-1} \big(Q_T(\theta_T^\star)\big) \right| > \frac{\epsilon_T}{2\,\delta} \right\} + P_\theta \left\{ \left| \overline{\theta}_T - \theta \right| > \delta \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\le P_\theta \left\{ \left| I_T^{-1} \big(Q_T(\theta_T^\star)\big) \right| > \frac{\epsilon_T}{2\,\delta} \right\} + P_\theta \left\{ \left| \theta_T - \theta \right| > \delta \right\} + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\le P_\theta \left\{ \sup_{\theta\in\Theta} \left| I_T^{-1} \big(Q_T(\theta)\big) \right| > \frac{\epsilon_T}{2\delta} \right\} + C \exp\big(-b\, n_T\, \delta^\gamma\big) + P_\theta \left\{ \left| I_T^{-1} H_T(\theta) - 1 \right| > \frac{\epsilon_T}{2} \right\} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Corollary 3.2)}\ .
\end{aligned}
$$

This completes the proof of the theorem. □

Remark. We used the splitting technique developed by Michel and Pfanzagl (1971) for the i.i.d. case. The upper bound in the Berry–Esseen type inequality obtained here contains four terms. The first term is cube root of the absolute moment, the second and the third term are moderate deviations type probabilities of the second and the third derivatives of log-likelihood respectively, and the fourth term is decays exponentially. The bound is quite sharp as seen in the linear case in the following example.

4. NONHOMOGENEOUS ORNSTEIN–UHLENBECK PROCESS

We apply the Berry–Esseen results for the MLE in the nonhomogeneous Ornstein–Uhlenbeck process satisfying the stochastic differential equation

$$dX_t = \theta\, t\, X_t\, dt + dW_t\ , \qquad t \geq 0, \quad X_0 = 0 \tag{4.1}$$

where $\theta > 0$. Note that the solution is a nonstationary and nonergodic process. Here the MLE based on $\{X_t,\, 0 \leq t \leq T\}$ is given by

$$\theta_T = \frac{\int_0^T t\, X_t\, dX_t}{\int_0^T t^2 X_t^2\, dt}\ ,$$

and $I_T(\theta) = \int_0^T t^2 X_t^2\, dt$. Let us choose $m_T = \int_0^T t^2\, e^{\theta t^2} dt$. Note that $m_T^{1/2}(\theta)(\theta_T - \theta)$ converges to Cauchy distribution with parameters $(0,1)$ as $T \to \infty$. Here $\frac{I_T}{m_T} \to \Delta^2$ a.s. where Δ has $\mathcal{N}\big(0, (\frac{\pi}{4\theta})^{1/2}\big)$ distribution and $\eta^2(\theta) = \Delta^2$. Directly from the calculation of J_1 in Theorem 3.6, we have

$$\begin{aligned}
\sup_{x \in \mathbb{R}} \left| P_\theta \left\{ I_T^{1/2}(\theta)\,(\theta_T - \theta) \leq x \right\} - \Phi(x) \right| &\leq E_\theta^{1/3} \left| \frac{\int_0^T t^2 X_t^2\, dt}{\int_0^T t^2\, e^{\theta t^2} dt} - \Delta^2 \right| \\
&\leq C\, T^{1/2} \exp\left(\frac{-\theta\, T^4}{12} \right) \\
&\leq C \exp\left(\frac{-\theta\, T^4}{24} \right) .
\end{aligned}$$

This shows that rate of weak convergence can be faster in the nonergodic processes than in ergodic processes in which case the sharpest possible rate is $O(T^{-1/2})$.

Remarks.

(**1**) Levanony *et al.* (1994) (see also Trofimov (1982)) showed that, for large enough t, MLE θ_t is a continuous semimartingale satisfying a stochastic differential equation. One could use the Berry–Esseen bound for semimartingales (see, e.g., Liptser and Shiryayev (1982, 1989)) to obtain a Berry–Esseen bound for the MLE θ_t. However, it would not give sharp bounds. Hence we follow the method of Michel and Pfanzagl (1971)) developed for the independent observations case.

(**2**) Large deviations for M-estimator remains to be investigated.

(**3**) It would be interesting if one can improve the Berry–Esseen bound in the above example by applying the characteristic function technique used in Bishwal (2000a).

(**4**) Berry–Esseen type bounds for Bayes estimators remain open.

(**5**) Large deviations and Berry–Esseen results for diffusions based on discrete observations remains to be investigated which would be more interesting in view of applications in finance.

REFERENCES

[1] Bahadur, R.R.; Zabell, S.L. and Gupta, J.C. (1980). *Large deviations, tests and estimates*. In "Asymptotic Theory of Statistical Tests" (I.M. Chakravarti, Ed.), Academic Press, New York, 33–64.

[2] Basawa, I.V. and Scott, D.J. (1983). *Asymptotic Optimal Inference for Non-ergodic Models*, Lecture Notes in Statistics **17**, Springer, Berlin.

[3] Bishwal, J.P.N. (1999). Large deviations inequalities for the maximum likelihood estimator and the Bayes estimators in nonlinear stochastic differential equations, *Statist. Probab. Letters*, **43**(2), 207–215.

[4] Bishwal, J.P.N. (2000a). Sharp Berry–Esseen bound for the maximum likelihood estimator in the Ornstein–Uhlenbeck process, *Sankhyā Ser. A*, **62**(1), 1–10.

[5] Bishwal, J.P.N. (2000b). Rates of convergence of the posterior distributions and the Bayes estimators in the Ornstein–Uhlenbeck process, *Random Operators and Stochastic Equations*, **8**(1), 51–70.

[6] Bishwal, J.P.N. (2005). Uniform rate of weak convergence for the minimum contrast estimator in the Ornstein–Uhlenbeck process, *Methodology and Computing in Applied Probability* (to appear).

[7] Bishwal, J.P.N. (2007). *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, Vol. **1923**, Springer-Verlag.

[8] Bishwal, J.P.N. and Bose, A. (1995). Speed of convergence of the maximum likelihood estimator in the Ornstein–Uhlenbeck process, *Calcutta Statist. Assoc. Bulletin*, **45**, 245–251.

[9] Borkar, V. and Bagchi, A. (1982). Parameter estimation in continuous time stochastic processes, *Stochastics*, **8**, 193–212.

[10] Bose, A. (1986). Berry–Esseen bound for the maximum likelihood estimator in the Ornstein–Uhlenbeck process, *Sankhyā Ser. A*, **48**, 181–187.

[11] Dupuis, P. and Kushner, H.J. (1989). Stochastic approximations and large deviations: upper bounds and w.p.1 convergence, *SIAM J. Control Optim.*, **27**, 1108–11035.

[12] Florens, D. and Pham, H. (1999). Large deviations in estimation an Ornstein–Uhlenbeck model, *J. Appl. Probab.*, **36**, 60–77.

[13] Gikhman, I.I. and Skorohod, A.V. (1972). *Stochastic Differential Equations*, Springer, Berlin.

[14] Harison, V. (1992). *The Bernstein-von Mises theorem for a certain class of Gaussian Diffusion processes*, Publications du Service de Mathematiques, Vol. 6, Universite d'Antananarivo, 1–7.

[15] Ibragimov, I.A. and Khasminskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*, Springer-Verlag, Berlin.

[16] Jeganathan, P. (1982). On the asymptotic theory of statistical inference when the limit of the log-likelihood ratios is mixed normal, *Sankhyā Ser. A*, **44**, 173–212.

[17] Jeganathan, P. (1995). Some aspects of asymptotic theory with applications to time series models, *Econometric Theory*, **11**, 818–887.

[18] KALLIANPUR, G. and SELUKAR, R.S. (1993). *Estimation of Hilbert space valued parameter by the method of sieves.* In "Statistics and Probability: A Raghu Raj Bahadur Festschrift" (J.K. Ghosh, S.K. Mitra, K.R. Parthasarathy and B.L.S. Prakasa Rao, Eds.), Wiley Eastern, 325–347.

[19] KUTOYANTS, YU.A. (1978). Estimation of parameter of a diffusion type process, *Theor. Probab. Applic.*, **23**, 641–649.

[20] KUTOYANTS, YU.A. (1984). *Parameter Estimation for Stochastic Processes* (translated and edited by B.L.S. Prakasa Rao), Heldermann-Verlag, Berlin.

[21] KUTOYANTS, YU.A. (1994). *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.

[22] LE CAM, L. and YANG, G.L. (1990). *Asymptotics in Statistics*, Springer-Verlag, New York.

[23] LEVANONY, D. (1994). Conditional tail probabilities in continuous-time martingale LLN with application to parameter estimation in diffusions, *Stochastic Process. Appl.*, **51**, 117–134.

[24] LEVANONY, D.; SHWARTZ, A. and ZEITOUNI, O. (1994). Recursive identification in continuous-time stochastic processes, *Stochastic Process Appl.*, **49**, 245–275.

[25] LIPTSER, R.S. and SHIRYAYEV, A.N. (1977). *Statistics of Random Processes I*, Springer-Verlag, Berlin.

[26] LIPTSER, R.S. and SHIRYAYEV, A.N. (1982). On the rate of convergence in the central limit theorem for semimartingales, *Theory Probab. Appl.*, **27**, 1–13.

[27] LIPTSER, R.S. and SHIRYAYEV, A.N. (1989). *Theory of Martingales*, Kluwer, Doredrecht.

[28] MISHRA, M.N. (1989). The Bernstein- von Mises theorem for a class of nonhomogeneous diffusion processes, *Statist. Decisions*, **7**, 153–165.

[29] MISHRA, M.N. and PRAKASA RAO, B.L.S. (1985a). Asymptotic study of maximum likelihood estmation for nonhomogeneous diffusion processes, *Statist. Decisions*, **3**, 193–203.

[30] MICHEL, R. and PFANZAGL, J. (1971). The accuracy of the normal approximation for minimum contrast estimate, *Zeit Wahr. Verw. Gebiete*, **18**, 73–84.

[31] MISHRA, M.N. and PRAKASA RAO, B.L.S. (1985b). On the Berry–Esseen theorem for maximum likelihood estimator for linear homogeneous diffusion processes, *Sankhyā, Ser. A*, **47**, 392–398.

[32] OBLAKOVA, T.V. (1989). On the rate of convergence in the central limit theorem for stochastic integrals, *Russian Math. Surveys*, **44**(2), 289–290.

[33] SAKAMOTO, Y. and YOSHIDA, N. (1998). Asymptotic expansion of M-estimator over Wiener space, *Statistical Inference for Stochastic Processes*, **1**, 85–103.

[34] SCHMETTERER, L. (1974). *Introduction to Mathematical Statistics*, Springer-Verlag, Berlin.

[35] TROFIMOV, E.I. (1982). A stochastic equation for the maximum likelihood estimate of the parameter of a diffusion type process, *Russian Math. Surveys*, **37**(4), 221–222.

[36] YOSHIDA, N. (1990). Asymptotic behaviour of M-estimator and related random field for diffusion process, *Ann. Inst. Statist. Math.*, **42**, 221–251.

[37] YOSHIDA, N. (1997). Malliavin calculus and asymptotic expansions of martingales, *Probab. Theory Relat. Fields*, **109**, 301–342.