Statistics & Decisions 19, 121 – 136 (2001)

ASYMPTOTIC MINIMAX ESTIMATION IN NONLINEAR STOCHASTIC DIFFERENTIAL EQUATIONS FROM DISCRETE OBSERVATIONS

M.N. Mishra and B.L.S. Prakasa Rao

Received: Revised version: *June 27, 2000*

Abstract

Asymptotic optimality in the minimax sense of an approximate maximum likelihood estimator and an approximate Bayes estimator of a parameter appearing nonlinearly in the drift coefficient of an Ito stochastic differential equation has been established when observations are made at regularly spaced but dense time points.

1 Introduction

The study of statistical inference for diffusion processes which arise as the solutions of Ito stochastic differential equations (SDE) is of great importance in view of its large number of applications (cf. Prakasa Rao (1999 a,b)). There are a lot of contributions to the problem of drift estimation in diffusion processes observed in continuous time. But the assumption that the process can be observed continuously throughout a time interval is actually impractical. In view of this, it is of utmost importance to know the asymptotic behaviour of estimators of drift and diffusion parameters, when the process is observed at a discrete set of time points.

Let $t_0 = 0 < t_1 < t_2 < \ldots < t_n = T$ and suppose the process $\{X_t, 0 \leq t \leq T\}$

AMS 1991 Subject Classification: 62M05.

Key Words and phrases : Stochastic differential equation, Diffusion process, Approximate maximum likelihood estimator, Approximate Bayes estimator, Local asymptotic normality.

is observed at the time points $t_i, 0 \leq i \leq n$. An approximation of the maximum likelihood estimator (MLE) of the drift parameter in a linear SDE when T is fixed, based on the process $\{X_t\}$ observed over $[0,T]$, by an estimator based on $X_{t_i}, 0 \leq i \leq n$ and $\Delta_n = \max\{|t_{i-1} - t_i|, 1 \leq i \leq n\} \to 0$ as $n \to \infty$ has been studied in Le Breton (1976) and Mishra and Bishwal (1995). The weak consistency of the least squares estimator (LSE) when $\Delta_n \to 0$ and $T \to \infty$ has been studied in Dorogovcev (1976). The asymptotic normality and the asymptotic efficency of the LSE have been studied in Prakasa Rao (1983) when $T \to \infty$ and $\frac{T}{\sqrt{n}} \to 0$. Consistency and asymptotic efficency of a minimum contrast estimator of the drift parameter in nonlinear SDE when $T \to \infty$ and $\frac{T}{\sqrt{n}} \to 0$ have been studied in Dacunha-Castelle and Florens-Zmirou (1986). They have also studied the expansion of transition probability density function suitable for their purpose. A comprehensive discussion on parametric and nonparametric inference for stochastic process from sampled data is given in Prakasa Rao (1988) and more recently in Prakasa Rao (1999a,b).

In this paper we study the asymptotic properties of an approximate maximum likelihood estimator (AMLE) and an approximate Bayes estimator (ABE) of a parameter in the nonlinear drift coefficient of an Ito SDE when the process is observed at equidistant time points. The main tool of investigation is a suitable expansion of the transition probability density function of the Markov chain $\{X_{t_i}, 0 \leq i \leq n\}$. We prove the local asymptotic normality (LAN) property of the model under some conditions.

The existence of LAN property implies consistency, asymptotic normality and asymptotic efficiency of the estimator. It is worth mentioning that the asymptotic properties of the MLE and the BE of the drift parameter in homogenous nonlinear SDE have been studied in Kutoyants (1977) using the LAN property of the model when the process is observed continuously through out the time interval.

Through out the paper we shall write c for a generic positive constant. The paper is organised as follows. Section 2 contains preliminaries, definitions and notations. Section 3 describes the main results. Section 4 gives the proof of the main results.

2 Notations, Definitions and Preliminaries

Let $\Omega = \boldsymbol{R}^T$ where $\boldsymbol{R}$ denotes the real line and $T = [0, \infty)$. Let $\mathcal{B}$ denote the σ-algebra generated by the class of all finite dimensional cylinder sets. Let P_θ be a probability measure defined on the measurable space $(\Omega, \mathcal{B})$ such that the coordinate process $\boldsymbol{X}$ is such that it is stationary and satisfies the stochastic differential equation

$$dX_t = b(\theta, X_t)dt + dW_t, t > 0, X_0 = \eta, \tag{2.1}$$

where $b(\theta, .)$ is the drift coefficient, $\theta \in \Theta$ open in $\boldsymbol{R}$ is the unknown parameter and $\{W_t, t \geq 0\}$ is the standard Wiener process under the probability measure P_θ.. Let θ_0 be an arbitrary but fixed value of the parameter.

In this paper, we are concerned with the inference based on the process $\{X_t, t \geq 0\}$ when it is observed at equidistant time points $t_k = kn^{-1/2}$, $0 \leq k \leq n$. Thus denoting $\Delta t_k = t_k - t_{k-1}$, we have $n\Delta t_n \to \infty$ and $\Delta t_n \to 0$. All the limit statements seen here after are to be considered as $n \to \infty$. Let $\boldsymbol{X}^{(n)} = \{X_{t_k}, k = 0, 1, 2, \ldots n\}$ be the stationary Markov chain that we observe.

On $\{\Omega_n, \zeta^{(n)}\}$, denote the measure generated by the chain $\boldsymbol{X}^{(n)}$ by $P_\theta^{(n)}$ where $\Omega_n = \boldsymbol{R}^n, \zeta^{(n)}$ is the σ-algebra generated by $\{X_{(k-1)n^{-1/2}}, k = 1, 2, \ldots, n+1\}$ and $\theta \in \Theta$. Let $\zeta_k^{(n)}$ be the σ-algebra generated by $\{X_{(i-1)n^{-\frac{1}{2}}}, i = 1, 2, \ldots, k\}$. We assume that the following conditions hold:

(A1) $b(\theta, x)$ is a known real valued function continuous on $\Theta \times \boldsymbol{R}$, such that $\frac{\partial b(\theta,x)}{\partial x}$, $\frac{\partial b(\theta,x)}{\partial \theta}$, $\frac{\partial^2 b(\theta,x)}{\partial \theta^2}$ and $\frac{\partial^2 b_x(\theta,x)}{\partial \theta^2}$ denoted as $b_x, b_\theta, b_{\theta\theta}$ and $b_{x\theta\theta}$ respectively exist and are differentiable with respect to θ on $\Theta \times \boldsymbol{R}$ having uniformly bounded derivatives for $\theta \in \Theta$.

(A2) b_θ and $b_{x\theta\theta}$ are Lipschitz in θ with respective Lipschitzan functions $g_1(x)$ and $g_2(x)$ satisfying $E_\theta g_1^2(X) < \infty$ and $E_\theta g_2^2(X) < \infty$.

(A3) For all θ and $\theta_1 \neq \theta_2$ in $\Theta, E_\theta(b(\theta_1, X_{kn^{-\frac{1}{2}}}) - b(\theta_2, X_{kn^{-\frac{1}{2}}}))^2 \neq 0$.

(A4) For the stationary Markov chain $X^{(n)} = \{X_{t_k}, 0 \le k \le n\}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} b_\theta^2(\theta, X_{t_{k-1}}) = E_\theta b_\theta^2(\theta,\eta) = \gamma > 0$$

a.s. under P_θ-measure as $n \to \infty$ uniformly for $\theta \in \Theta$.

In addition to the conditions (A1)-(A4), we suppose that sufficient conditions hold (cf. Lemma 3 of Dacunha-Castelle and Florens-Zmirou (1986)) to ensure the expansion of the transition density function of the Markov chain $\{X_{j/\sqrt{n}}, 0 \le j \le n\}$. Let $q_\theta(.,.)$ be the transition density function of this Markov chain.

Let

$$Z_{n,\theta}(h, X^{(n)}) = \frac{dP^{(n)}_{\theta+hn^{-1/4}\gamma^{-1/2}}}{dP^{(n)}_\theta}(X^{(n)}),$$

Recall that the sequence of families of probability measures $\{P_\theta^{(n)}, \theta \in \Theta, n \ge 1\}$ is locally asymptotically normal (LAN) at a point $\theta_0 \in \Theta$ if the following conditions are satisfied:

(i) there exists a sequence $\{\Delta_{n,\theta_0}, n \ge 1\}$ of $\zeta^{(n)}$-measurable random variables such that

$$\log Z_{n,\theta_0}(h, X^{(n)}) - h\Delta_{n,\theta_0} + \frac{1}{2}h^2 \xrightarrow{P} 0 \text{ in } P_{\theta_0} \text{ -measure}$$

as $n \to \infty$, where

$$\Delta_{n,\theta_0} = \frac{1}{n^{1/4}\gamma^{\frac{1}{2}}} \sum_{k=1}^{n} b_\theta(\theta_0, X_{(k-1)n^{-\frac{1}{2}}})\Delta W_k,$$

and

$$\Delta W_k = W_{\frac{k}{\sqrt{n}}} - W_{\frac{k-1}{\sqrt{n}}}$$

(ii) the sequence $\{\Delta_{n,\theta_0}\}$ converges in distribution under $P_{\theta_0}^{(n)}$-measure to $N(0,1)$ as $n \to \infty$.

If the family $\{P_\theta^{(n)}, \theta \in \Theta\}$ is LAN at every point $\theta \in \Theta$, then it is LAN in Θ and if the conditions (i) and (ii) are satisfied uniformly in $\theta \in \Theta$, then it is uniformly LAN in Θ.

Using the expansion of the transition probability density function of a Markov chain as in Lemma 3 of Dacunha-Castelle and Florens-Zmirou (1986), we write the expansion of the transition probability density function $q_\theta(.,.)$ of the Markov chain $X^{(n)}$ as

$$\begin{aligned} q_\theta\left(X_{\frac{k-1}{\sqrt{n}}}, X_{\frac{k}{\sqrt{n}}}\right) &= \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{\sqrt{n}}{2}\left(X_{\frac{k}{\sqrt{n}}} - X_{\frac{k-1}{\sqrt{n}}}\right)^2 + G\left(X_{\frac{k}{\sqrt{n}}}\right) - G\left(X_{\frac{k-1}{\sqrt{n}}}\right)\right\} \\ &\times \left[1 + \frac{S_1(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}})}{\sqrt{n}} + \frac{S_2(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}})}{n}\right] + \frac{1}{n^{\frac{5}{4}}} R\left(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}}\right) \end{aligned} \tag{2. 2}$$

with $|R(x,y)| < c\tilde{\phi}_\tau^4(x,y)$ uniformly with respect to θ, where $\phi_\tau(x) = \sup\{1, |b(\theta_0, x)|^\tau)$ for $\tau > 0, \tilde{\phi}_\tau(x,y) = \int_0^1 \phi_\tau((1-u)x + uy)du, S_1, S_2$ are as defined below and

$$G(x) = \int_0^x b(\theta_0, u)du.$$

This can be seen as follows. By the Ito formula

$$\begin{aligned} G\left(X_{\frac{k}{\sqrt{n}}}\right) - G\left(X_{\frac{k-1}{\sqrt{n}}}\right) &= \int_{(k-1)n^{-\frac{1}{2}}}^{kn^{-\frac{1}{2}}} b(\theta, X_t)dX_t + \frac{1}{2}\int_{(k-1)n^{-\frac{1}{2}}}^{kn^{-\frac{1}{2}}} b_x(\theta, X_t)dt \\ &\simeq b(\theta, X_{\frac{k-1}{\sqrt{n}}})\Delta X_k + \frac{1}{2\sqrt{n}} b_x(\theta, X_{\frac{k-1}{\sqrt{n}}}) \end{aligned} \tag{2. 3}$$

Following Dacuna-Castelle and Florens-Zmirou ((1986), page 272), we write

$$\frac{1}{\sqrt{n}} S_1\left(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}}\right) = -\frac{1}{2\sqrt{n}} b^2(\theta, X_{\frac{k-1}{\sqrt{n}}}) - \frac{1}{2\sqrt{n}} b_x(\theta, X_{\frac{k-1}{\sqrt{n}}}), \tag{2.4}$$

and

$$\frac{1}{n}S_2\left(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}}\right) = -\frac{1}{12n}\left\{b_\theta^2(\theta, X_{\frac{k-1}{\sqrt{n}}}) + b(\theta, X_{\frac{k-1}{\sqrt{n}}})b_{\theta\theta}(\theta, X_{\frac{k-1}{\sqrt{n}}})\right\} -\frac{1}{24n}b_{x\theta\theta}(\theta, X_{\frac{k-1}{\sqrt{n}}}) \tag{2. 5}$$

Substituting the approximations (2.3)-(2.5) in (2.2), we obtain an approximation of the transition density function $q_\theta(\cdot,\cdot)$. We assume that sufficient regularity conditions hold justifying the validity of this approximation of the transition density function $q_\theta(\cdot,\cdot)$.

Using (2.3) - (2.5), we write the log likelihood function of $X^{(n)}$ as

$$\begin{aligned}
&\log \pi_{k=1}^n q_\theta\left(X_{\frac{k-1}{\sqrt{n}}}, X_{\frac{k}{\sqrt{n}}}\right) \\
&= \frac{n}{2}\log\left(\frac{n}{2\pi}\right) - \frac{\sqrt{n}}{2}\sum_{k=1}^n\left(X_{\frac{k}{\sqrt{n}}} - X_{\frac{k-1}{\sqrt{n}}}\right)^2 + \sum_{k=1}^n b(\theta, X_{(k-1)n^{-\frac{1}{2}}})\Delta X_k \\
&-\frac{1}{2\sqrt{n}}\sum_{k=1}^n b^2(\theta, X_{(k-1)n^{-\frac{1}{2}}}) \\
&-\frac{1}{12n}\sum_{k=1}^n\{b_\theta^2(\theta, X_{(k-1)n^{-\frac{1}{2}}}) + b(\theta, X_{(k-1)n^{-\frac{1}{2}}})b_{\theta\theta}(\theta, X_{(k-1)n^{-\frac{1}{2}}})\} \\
&-\frac{1}{24n}\sum_{k=1}^n b_{x\theta\theta}(\theta, X_{(k-1)n^{-\frac{1}{2}}}) + \frac{1}{n^{\frac{5}{4}}}\sum R_\theta(X_{kn^{-\frac{1}{2}}}, X_{(k-1)n^{-\frac{1}{2}}})
\end{aligned} \tag{2. 6}$$

The log likelihood ratio is given by

$$\begin{aligned}
&\log \pi_{k=1}^n q_{\theta_2}(X_{(k-1)n^{-\frac{1}{2}}}, X_{kn^{-\frac{1}{2}}}) - \log \pi_{k=1}^n q_{\theta_1}(X_{(k-1)n^{-\frac{1}{2}}}, X_{kn^{-\frac{1}{2}}}) \\
&= \sum_{k=1}^n\{b(\theta_2, X_{(k-1)n^{-\frac{1}{2}}}) - b(\theta_1, X_{(k-1)n^{-\frac{1}{2}}})\}\Delta W_k \\
&-\frac{1}{2\sqrt{n}}\sum_{k=1}^n\{b(\theta_2, X_{(k-1)n^{-\frac{1}{2}}}) - b(\theta_1, X_{(k-1)n^{-\frac{1}{2}}})\}^2 \\
&-\frac{1}{12n}\{\sum_{k=1}^n(b_\theta^2(\theta_2, X_{(k-1)n^{-\frac{1}{2}}}) - b_\theta^2(\theta_1, X_{(k-1)n^{-\frac{1}{2}}})) \\
&\quad+\sum_{k=1}^n(b_\theta(\theta_2, X_{(k-1)n^{-\frac{1}{2}}})b_{\theta\theta}(\theta_2, X_{(k-1)n^{-\frac{1}{2}}})
\end{aligned}$$

$$-\sum_{k=1}^{n}(b_\theta(\theta_1, X_{(k-1)n^{-\frac{1}{2}}})b_{\theta\theta}(\theta_1, X_{(k-1)n^{-\frac{1}{2}}}))\}$$
$$-\frac{1}{24n}\sum_{k=1}^{n}\{b_{x\theta\theta}(\theta_2, X_{(k-1)n^{-\frac{1}{2}}}) - b_{x\theta\theta}(\theta_1, X_{(k-1)n^{-\frac{1}{2}}})\}$$
$$+\frac{1}{n^{\frac{5}{4}}}\sum_{k=1}^{n} R_{\theta_2}\left(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}}\right) - \frac{1}{n^{\frac{5}{4}}}\sum_{k=1}^{n} R_{\theta_1}\left(X_{\frac{k}{\sqrt{n}}}, X_{\frac{k-1}{\sqrt{n}}}\right) \qquad (2.\ 7)$$

Using the condition (A1), let us denote

$$(i) \quad \sup_\theta |b_\theta(\theta,x)b_{\theta\theta}(\theta,x)| = \phi_1(x)$$
$$(ii) \quad \sup_\theta |b_{\theta\theta}^2(\theta,x) + b_\theta(\theta,x)b_{\theta\theta\theta}(\theta,x)| = \phi_2(x)$$
$$(iii) \quad \sup_\theta |b_{x\theta\theta\theta}(\theta,x)| = \phi_3(x)$$

and $\max(\phi_1(x), \phi_2(x), \phi_3(x) = \phi_4(x)$. Assume that

(A5) $\frac{1}{n^{\frac{5}{4}}}\sum_{k=1}^{n} E\{\phi_4(X_{\frac{k-1}{\sqrt{n}}})|\zeta_{k-2}^{(n)}\} = o(1)$ as $n \to \infty$.

3 Main results

Theorem 3.1 Let the conditions (A1)-(A5) hold, and $\{X_t, t \geq 0\}$ satisfy (2.1). The sequence of probability measures $\{P_\theta^{(n)}, \theta \in \Theta, n \geq 1\}$ generated by $X^{(n)}$ satisfy the LAN condition uniformly in $\theta \in \Theta$ with the normalizing sequence $n^{-\frac{1}{4}}\gamma^{-\frac{1}{2}}$.

Let the loss function $L(x), x \in R$ be continuous at zero, symmetric, $L(0) = 0$, the subset $\{x : L(x) < c\}$ be convex for all $c > 0$ and suppose that for any $h > 0, L(x)$ as $|x| \to \infty$ does not increase faster than $\exp(h|x|^2)$. We state the Hajek inequality in Theorem 3.2 below for the scheme of observations $X^{(n)}$ and the loss function $L(u)$. The proof of this result follows as in Ibragimov and Hasminskii (1981) due to the property of LAN for $X^{(n)}$ as shown earlier.

Theorem 3.2 Let the conditions (A1)-(A5) hold. Then for any family of estimators

$\{\hat{\theta}_n, n \geq 1\}$ and the loss function $L(\cdot)$ and for any $\delta > 0$, the inequality.

$$\liminf_{n\to\infty} \sup_{|\theta-\theta_0|<\delta} E_\theta\{L(n^{1/4}\gamma^{1/2}(\hat{\theta}_n - \theta_0))\} \geq \frac{1}{\sqrt{2\pi}} \int_R L(x)e^{-x^2/2}dx$$

holds.

Remark 3.1 For $\Theta \subset R$, the inequality in Theorem 3.2 given above becomes equality if and only if $\gamma^{\frac{1}{2}} n^{\frac{1}{4}}(\hat{\theta}_n - \theta_0) - \Delta_{n,\theta_0} \to 0$ in $P_{\theta_0}^{(n)}$-measure as $n \to \infty$. Note that this condition holds for an approximate MLE and an approximate Bayes estimator of the parameter θ based on the discrete set of observations , $\{X_{t_k}, t_k = kn^{-\frac{1}{2}}, 0 \leq k \leq n\}$,.

In discussing the properties of the Bayes estimators, we shall consider a quadratic loss function and a prior density $\pi(\theta)$ which is continuous, bounded and positive on Θ.

We shall use the following lemmas for the proof of our main result.

Lemma 3.1 Let the conditions (A1)-(A5) hold. Then there exists a constant $c > 0$ such that

$$\sup_{\theta\in\Theta} E_\theta(Z_{n,\theta}(h_1, X^{(n)})^{\frac{1}{2}} - Z_{n,\theta}(h_2, X^{(n)})^{\frac{1}{2}})^2 \leq c(h_2 - h_1)^2$$

whenever $h_i \in H_{\theta,n} = \{h : \theta + hn^{-\frac{1}{4}}\gamma^{-\frac{1}{2}} \in \Theta\}$ for $i = 1,2$.

Lemma 3.2 Let the conditions (A1)-(A5) hold. Then

$$\sup_{\theta\in\Theta} P_\theta\{Z_{n,\theta}(h, X^{(n)}) > e^{-\alpha h^2}\} \leq e^{-\alpha h^2}$$

where $h \in H_{\theta,n} = \{h : \theta + hn^{-\frac{1}{4}}\gamma^{-\frac{1}{2}} \in \Theta\}, \alpha = \beta/6\gamma$ and $\beta = \inf_{\theta,x} b_\theta(\theta, x)$.

Definition 3.1 : Let the family of measures $\{P_\theta^{(n)}, \theta \in \Theta\}$ be LAN in Θ. Then an estimator θ_n^* is called asymptotically efficient if for every $\theta_0 \in \Theta$,

$$\lim_{\delta\to 0}\lim_{n\to\infty} \sup_{|\theta^*-\theta_0|<\delta} E_\theta\left(\frac{\theta_n^* - \theta_0}{n^{-1/4}\gamma^{-1/2}}\right)^2 = 1.$$

Theorem 3.3 Let the assumptions (A_1)-(A_5) hold. Let $\hat{\theta}_n$ be the approximate maximum likelihood estimator of the parameter θ. Then $\hat{\theta}_n$ possesses the following properties uniformly for $\theta \in \Theta$ as $n \to \infty$:

(i) $\hat{\theta}_n$ is consistent in P_θ-measure;

(ii) $n^{1/4}\gamma^{1/2}(\hat{\theta}_n - \theta)$ is asymptotically standard normal ;

(iii) all the moments of the random variable $n^{1/4}\gamma^{1/2}(\hat{\theta}_n - \theta)$ converge to the corresponding moments of the standard normal distribution and

(iv) $\lim_{n\to\infty} \sup_{\theta\in\Theta} |n^{1/4}\gamma^{1/2}(\hat{\theta}_n - \theta) - \Delta_{n,\theta}| = 0$ in $P_\theta^{(n)}$-measure.

Further from above discussions, it is easy to prove the asymptotic efficiency of $\hat{\theta}_n$.

Remark : For a detailed proof of these results, see the proof of Theorems 1.1, 1.2 and 1.3 of Ibragimov and Hasminskii (1981, Chapter III) .

Theorem 3.4 Let the conditions (A_1)-(A_5) hold. Let $\tilde{\theta}_n$ be the approximate Bayes estimator of the parameter θ with quadratic loss function and a prior density function $\pi(\theta)$. Then $\tilde{\theta}_n$ possesses all the properties of $\hat{\theta}_n$ enumerated above as in Theorem 3.3.

Remark : For detailed proof of these results, see Theorem 2.2 and Theorem 2.1 of Ibragimov and Hasminskii (1981, Chapter III).

4 Proofs of main results :

We now give a proof of Theorem 3.1. From (2.7) we have the log likelihood ratio

$$Z_{n,\theta}(h, X^{(n)}) = \frac{dP^{(n)}_{\theta+h\gamma^{-1/2}n^{-1/4}}}{dP_\theta^{(n)}}$$

$$\begin{aligned}
= \; & h\gamma^{-1/2}n^{-1/4}\sum_{k=1}^{n} b_\theta(\theta, X_{(k-1)n^{-1/2}})\Delta W_k \\
& -\frac{h^2\gamma^{-1}}{2n}\sum_{k=1}^{n} b_\theta^2(\theta, X_{(k-1)n^{-1/2}}) + \frac{1}{2}h^2\gamma^{-1}n^{-1/2}\sum_{k=1}^{n} b_{\theta\theta}(\tilde{\theta}, X_{(k-1)n^{-1/2}})\Delta W_k \\
& -\frac{h^3\gamma^{-3/2}}{n^{5/4}}\sum_{k=1}^{n} b_\theta(\tilde{\theta}, X_{(k-1)n^{-1/2}})b_{\theta\theta}(\tilde{\theta}, X_{(k-1)n^{-1/2}}) \\
& -\frac{h\gamma^{-1/2}}{12n^{5/4}}\sum_{k=1}^{n}\{b_{\theta\theta}(\theta, X_{(k-1)n^{-1/2}})b_{\theta\theta}(\tilde{\theta}, X_{(k-1)n^{-1/2}}) \\
& \qquad + b_\theta(\theta, X_{(k-1)n^{-1/2}})b_{\theta\theta\theta}(\tilde{\theta}, X_{(k-1)n^{-1/2}})\} \\
& -\frac{h\gamma^{-1/2}}{24n^{5/4}}\sum_{k=1}^{n} b_{x\theta\theta\theta}(\tilde{\theta}, X_{(k-1)n^{-1/2}}) + \frac{c}{n^{5/4}}\tilde{\phi}_T^{\,4}(X_{(k-1)n^{-1/2}}, X_{kn^{-1/2}}). \qquad (4.1)
\end{aligned}$$

The above relation is a consequence of the Taylor's expansion of

$$b(\theta + h\gamma^{-1/2}n^{-1/4}, \quad X_{(k-1)n^{-1/2}})$$

at θ and note that $|\tilde{\theta} - \theta| < h\gamma^{-1/2}n^{-1/4}$.

In the above expression (4.1), the third term tends to zero in probability uniformly for $\theta \in \Theta$ as $n \to \infty$ by Prakasa Rao (1983). All the other terms except the first and second are bounded by $\phi_4(X_{(k-1)}n^{-1/2})$ uniformly for $\theta \in \Theta$. Again

$$\begin{aligned}
\frac{1}{n^{5/4}}\sum_{k=1}^{n}\phi_4(X_{(k-1)n^{-\frac{1}{2}}}) &= \frac{1}{n^{5/4}}\sum_{k=1}^{n}\{\phi_4(X_{(k-1)n^{-\frac{1}{2}}}) - E(\phi_4(X_{(k-1)n^{-\frac{1}{2}}})|\zeta_{k-2}^{(n)}) \\
& \quad + \frac{1}{n^{5/4}}\sum_{k=1}^{n} E(\phi_4(X_{\frac{k-1}{\sqrt{n}}})|\zeta_{k-2}^{(n)}) \\
&= J_1 + J_2 \quad \text{(say)}.
\end{aligned}$$

Note that $EJ_1 = 0$ and $\text{Var}\, J_1 = n^{-5/2}\sum_{k=1}^{n} \text{Var}\,(\phi_4(X_{\frac{k-1}{\sqrt{n}}})) = \frac{1}{n^{5/2}}n\,\text{Var}\,(\phi_4(X_0)) \to 0$ as $n \to \infty$ by stationarity of $\{X_t, t \geq 0\}$. Hence $J_1 = o_p(1)$.

By the condition (A_5), $J_2 = o_p(1)$ as $n \to \infty$. Therefore $J_1 + J_2 = o_p(1)$ as $n \to \infty$ uniformly for $\theta \in \Theta$. Hence

$$
\begin{aligned}
Z_{n,\theta}(h, X^{(n)}) &= h\gamma^{-1/2}n^{-1/4}\sum_{k=1}^{n} b_\theta(\theta, X_{(k-1)}n^{-1/2})\Delta W_k \\
&\quad \frac{-h^2\gamma^{-1}}{2n}\sum_{k=1}^{n} b_\theta^2(\theta, X_{(k-1)n^{-1/2}}) + o_p(1).
\end{aligned}
$$

This verifies the first condition of LAN. In order to verify the second condition, we shall prove that $\Delta_{n,\theta} \xrightarrow{\mathcal{L}} \Delta$ uniformly for $\theta \in \Theta$, where Δ is a standard normal random variable. Note that

$$
\begin{aligned}
\Delta_{n,\theta} &= n^{-1/4}\gamma^{-1/2}\sum_{k=1}^{n} b_\theta(\theta, X_{(k-1)n^{-1/2}})\Delta W_k \\
&= \sum_{k=1}^{n} Y_{k,n} \quad \text{(say)}.
\end{aligned}
$$

Here $\{Y_{k,n}, \ \zeta_{k-1}^{(n)}, \ 1 \le k \le n\}$ is a martingale difference sequence for every $n \ge 1$. In order to use the martingale central limit theorem, we verify the following conditions mentioned in Brown and Hewitt (1975) and in a most convenient form in Brown and Eagleson (1971). Observe that

$$E(Y_{k,n}|\zeta_{k-1}^{(n)}) = 0 \ \text{ almost surely}$$

and

$$E(Y_{k,n}^2|\zeta_{k-1}^{(n)}) = \frac{1}{n\gamma}b_\theta^2(\theta, X_{(k-1)n^{-1/2}}) \ \text{ almost surely.}$$

Therefore, uniformly for $\theta \in \Theta$, $\lim_{n\to\infty}\sum_{k=1}^{n} E(Y_{k,n}^2|\zeta_{k-1}^{n}) = 1$ almost surely by (A4)

Also, writing $M(x)$ for the probability that a χ_1^2- random variable exceeds x, we have,

$$E(Y_{k,n}^2 I_{[|Y_{k,n}| \ge \epsilon]}|\zeta_{k-1}^{(n)}) = \frac{1}{n\gamma}b_\theta^2(\theta, X_{(k-1)n^{-1/2}})M(\frac{n\epsilon}{b^2(\theta, X_{(k-1)n^{-1/2}})}).$$

Then, by condition (A4) and the fact that $M(x) \downarrow 0$ as $x \uparrow \infty$, we obtain that , uniformly for $\theta \in \Theta$,

$$\lim_{n\to\infty} \sum_{K=1}^{n} E(Y_{k,n}^2 I_{[|Y_{k,n}|\geq\varepsilon]}|\zeta_{k-1}^{(n)}) = 0 \text{ almost surely} \tag{4.2}$$

for $\varepsilon > 0$.

These conditions ensure that $\sum_{k=1}^{n} Y_{k,n} \xrightarrow{\mathcal{L}} N(0,1)$ as $n \to \infty$ uniformly for $\theta \in \Theta$ where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution.

In the proofs of the Lemmas 3.1 and 3.2, we have followed the procedure adopted in Kutoyants (1977).

Proof of Lemma 3.1. We know that $P_{\theta_1}^{(n)}, P_{\theta_2}^{(n)}$ and $P_{\theta}^{(n)}$ are equivalent from (A2). Let us write

$$V_n = \exp -(\frac{1}{4\sqrt{n}} \sum_{k=1}^{n} \delta_k^2 - \frac{1}{2} \sum_{k=1}^{n} \delta_k \Delta W_{t_k} + o_p(1))$$

where

$$\delta_k = b(\theta_2, X_{(k-1)n^{-(1/2)}}) - b(\theta_1, X_{(k-1)n^{-(1/2)}}).$$

Therefore,

$$\begin{aligned} & E_\theta((Z_{n,\theta}(h_1, X^{(n)})^{1/2} - Z_{n,\theta}(h_2, X^{(n)})^{1/2})^2 \\ = \; & E_\theta(\frac{dP_{\theta_1}^{(n)}}{dP_{\theta}^{(n)}}) + E_\theta(\frac{dP_{\theta_2}^{(n)}}{dP_{\theta}^{(n)}}) - 2E_{\theta_1}(\frac{dP_{\theta_2}^{(n)}}{dP_{\theta_1}^{(n)}})^{1/2} \\ \leq \; & 2 - 2E_{\theta_1} V_n. \end{aligned}$$

If

$$\frac{1}{4\sqrt{n}} \sum_{k=1}^{n} \delta_k^2 - \frac{1}{2} \sum_{k=1}^{n} \delta_k \Delta W_k + o_p(1)) \geq 0,$$

then

$$\begin{aligned}2(1-E_{\theta_1}V_n) &\leq 2E_{\theta_1}\{(\frac{1}{4\sqrt{n}}\sum_{k=1}^{n}\delta_k^2-\frac{1}{2}\sum_{k=1}^{n}\delta_k\Delta W_k)+o_p(1)\}\\ &\quad (\text{since } 1-e^x\leq x \text{ when } x\geq 0),\\ &= \frac{1}{2\sqrt{n}}E_{\theta_1}\{\sum_{k=1}^{n}\delta_k^2+o_p(1)\} \qquad (4.\ 3)\end{aligned}$$

If

$$\frac{1}{4\sqrt{n}}\sum_{k=1}^{n}\delta_k^2-\frac{1}{2}\sum_{k=1}^{n}\delta_k\Delta W_k+o_p(1)<0,$$

then

$$\begin{aligned}2(1-E_{\theta_1}V_n) &= 2(1-E_{\theta_1}\exp(\frac{1}{2}\sum_{k=1}^{n}\delta_k\Delta W_k-\frac{1}{4\sqrt{n}}\sum_{k=1}^{n}\delta_k^2+o_p(1))\\ &\leq 2(1-\exp(\frac{1}{4\sqrt{n}}E_{\theta_1}\{\sum_{k=1}^{n}\delta_k^2+o_p(1)\}\\ &\quad (\text{by Jensen's inequlity })\\ &\leq \frac{1}{2\sqrt{n}}E_{\theta_1}(\sum_{k=1}^{n}\delta_k^2+o_p(1)). \qquad (4.\ 4)\end{aligned}$$

From (4.3), (4.4) and using (A2), we get

$$\begin{aligned}2E_{\theta_1}(1-V_n) &\leq \frac{1}{2\sqrt{n}}E_{\theta_1}(\sum_{k=1}^{n}\delta_k^2+o_p(1))\\ &\leq \frac{1}{2\sqrt{n}}(\theta_1-\theta_2)^2\sum_{k=1}^{n}E_{\theta_1}(\phi_4^2(X_{(k-1)n^{-(1/2)}}))\\ &\leq c(h_2-h_1)^2.\end{aligned}$$

Thus

$$\sup_{\theta\in\Theta}2E_{\theta_1}(1-V_n)\leq c(h_2-h_1)^2,$$

Proof of Lemma 3.2. Let us write

$$\eta_k = b(\theta + hn^{-1/4}\gamma^{-1/2}, X_{(k-1)n^{-(1/2)}}) - b(\theta, X_{(k-1)n^{-(1/2)}}).$$

Let $\alpha = \frac{\beta}{6\gamma}$ and $\beta = \inf_{\theta,x} b_\theta^2(\theta, x)$ as before. From (4.4), we obtain that

$$\begin{aligned}
&P_\theta^{(n)}\{Z_{n\theta}(h, X^{(n)}) > e^{-\alpha h^2}\} \\
&= P_\theta^{(n)}\{(\exp(\sum_{k=1}^n \eta_k \Delta W_k - \frac{1}{2\sqrt{n}}\sum_{k=1}^n \eta_k^2 + o_p(1))) > e^{-\alpha h^2}\} \\
&= P_\theta^{(n)}\{\frac{1}{2}(\sum_{k=1}^n \eta_k \Delta W_k - \frac{1}{2\sqrt{n}}\sum_{k=1}^n \eta_k^2) + o_p(1) > -\frac{\alpha h^2}{2}\}.
\end{aligned}$$

Again,

$$\begin{aligned}
&\sum_{k=1}^n \{(b(\theta + hn^{-(1/4)}\gamma^{-1/2}, X_{(k-1)n^{-(1/2)}}) - b(\theta, X_{(k-1)n^{-(1/2)}})\}^2 \\
&= \frac{h^2}{\gamma\sqrt{n}}\sum_{k=1}^n b_\theta^2(\tilde{\theta}, X_{(k-1)n^{-(1/2)}}) \quad \text{(by Taylor's formula)} \\
&\geq h^2\sqrt{n}\frac{\beta}{\gamma}. \qquad (4.5)
\end{aligned}$$

Using the Chebyshev's inequality and (4.5) we get,

$$\begin{aligned}
&P_\theta^{(n)}\{(\frac{1}{2}\sum_{k=1}^n \eta_k \Delta W_k - \frac{1}{4\sqrt{n}}\sum_{k=1}^n \eta_k^2 - \frac{1}{4\gamma}h^2\beta) + o_p(1))) > -\frac{\alpha h^2}{2}\} \\
&\leq \exp\{-h^2(\frac{\beta}{4\gamma} - \frac{\alpha}{2})\} E_\theta \exp\{\frac{1}{2}\sum_{k=1}^n \eta_k \Delta W_k - \frac{1}{\sqrt{n}}\sum_{k=1}^n (\frac{\eta_k}{2})^2 + o_p(1))\} \\
&= \exp\{-h^2(\frac{\beta}{4\gamma} - \frac{\alpha}{2})\} \\
&\quad (\text{since } E_\theta \exp\{\tfrac{1}{2}\eta_k \Delta W_k - \tfrac{1}{\sqrt{n}}\textstyle\sum_{k=1}^n (\frac{\eta_k}{2})^2 + o_p(1))\} \leq 1) \\
&= e^{-\alpha h^2}
\end{aligned}$$

References

[1] Brown, B.M. and Eagleson, G.K., Martingale convergence to infinitely divisible laws with finite variance, *Trans. Amer. Math. Soc.* **162**, (1971), 449-453.

[2] Brown, B.M. and Hewitt, J.L., Asymptotic likelihood theory for diffusion processes, *Journal of Applied Probability* **12**, (1975), 228-238.

[3] Dacunha - Castelle, D. and Florens - Zmirou, D., Estimation of the coefficients of a diffusion from discrete observation. *Stochastics* **19** (1986), 263 - 284.

[4] Dorogovcev, A.Ja., The consistency of an estimate of a parameter of a stochastic differential equation, *Theory of Probability and Math. Statist.* **10** (1976), 73 - 82.

[5] Ibragimov, I.A. and Hasminskii, R.Z., *Statistical Estimation : Asymptotic Theory,* Springer-Verlag, New York (1981).

[6] Kutoyants, Yu.A., Estimation of drift coefficient parameter of a diffusion in the smooth case, *Theory Prob. Appl.* **22** (1977), 399-406.

[7] Kutoyants, Yu.A., *Parameter Estimation for Stochastic Processes* (Translated and edited by B.L.S.Prakasa Rao)(1984), Heldermann Verlag, Berlin.

[8] Le Breton, A., On continuous and discrete sampling for parameter estimation in diffusion type process, *Mathematical Programming Study* **5** (1976), 124-144.

[9] Mishra, M.N. and Bishwal, J.P.N., Approximate maximum likelihood estimation, *Stochastics and Stochastic Reports,* **52** (1995), 1-13.

[10] Prakasa Rao, B.L.S., Asymptotic theory for nonlinear least squares estimator for diffusion process, *Math. Operationsforsch. Statist. Ser. Statistics* **12** (2) (1983), 195- 156.

[11] Prakasa Rao, B.L.S., Statistical inference from sampled data for stochastic process, In Contemporary Mathematics: *Statistical Inference from Stochastic Processes* (Ed. N. U. Prabhu) *Amer. Math. Soc.* **80**, 249-284 (1988).

[12] Prakasa Rao, B.L.S. *Statistical Inference for Diffusion Type Processes*, Arnold, London and Oxford University Press, New York(1999a).

[13] Prakasa Rao, B.L.S. *Semimartingales and their Statistical Inference*, CRC Press, Boca Raton, Florida and Chapman and Hall, London (1999b).

M.N.Mishra
Department of Mathematics
Ravenshaw College
Cuttack
INDIA

B.L.S.Prakasa Rao
Indian Statistical Institute
7, S.J.S.Sansanwal Marg
New Delhi 110 016
INDIA

e-mail: blsp@isid.ac.in