STOCHASTIC ANALYSIS AND APPLICATIONS
Vol. 20, No. 6, pp. 1309–1329, 2002

APPROXIMATION OF MAXIMUM LIKELIHOOD ESTIMATOR FOR DIFFUSION PROCESSES FROM DISCRETE OBSERVATIONS

M. N. Mishra[1] and B. L. S. Prakasa Rao[2,3,*]

[1]Sambalpur University, Sambalpur, India
[2]Indian Statistical Institute, New Delhi, India
[3]University of Iowa, Iowa City, Iowa

ABSTRACT

This paper is concerned with the approximation of the maximum likelihood estimator of parameter in the nonlinear drift coefficient of an Itô stochastic differential equation. Trapezoidal rule of approximation and rectangular rule of approximation have been compared when the observations are made at regularly spaced discrete but dense time points.

Key Words: Stochastic differential equation; Diffusion process; Maximum likelihood estimator; Trapezoidal rule of approximation; Rectangular rule of approximation; Regularly spaced discrete time points

*Corresponding author. Indian Statistical University, 7 S.J.S. Sansanwal Marg, New Delhi, India. E-mail: blsp@isid.isid.ac.in

1. INTRODUCTION

The study of statistical inference for diffusion type processes which arise as the solutions of Ito stochastic differential equations (SDE) is of extreme importance in view of its large number of applications (cf. [1]). The problem of drift estimation for diffusion processes observed over a continuous period of time is extensively discussed in the literature. But the assumption that the process can be observed continuously throughout a time interval is generally impossible in actual practice. In view of this, it is of importance and of interest to know the asymptotic behaviour of the estimators of the drift and the diffusion parameters when the process is observed at a discrete set of time points. Let us consider a subdivision of the interval $[0, T]$ with $t_0 = 0 \leq t_1 \leq t_2 \leq \cdots \leq t_n = T$. The weak consistency of the least squares estimator (LSE) when $\Delta_n = \max\{|t_{i+1} - t_i|, 0 \leq i \leq n\} \to 0$ and $T \to \infty$ has been studied in Ref. [2]. Strong consistency of the LSE under the condition $n \to \infty$ and $T = O(\sqrt{n})$ has been studied in Ref. [3]. The asymptotic normality and asymptotic efficiency of the LSE have been studied in Ref. [4] when $T \to \infty$ and $T/\sqrt{n} \to 0$. Consistency and asymptotic efficiency of minimum contrast estimator of the drift parameter in a nonlinear SDE, when $T \to \infty$ and $T/n^{2/3} \to 0$, have been studied in Ref. [5]. Maximum contrast estimator of the unknown parameter in the nonlinear drift coefficient has been studied in Ref. [6]. The estimation of a parameter appearing linearly in the diffusion coefficient of the SDE has been studied in Ref. [7] using the local asymptotic mixed normality property of the model when $n \to \infty$ and T is fixed. Prakasa Rao and Rubin[8] have studied the limiting properties of a process related to the LSE of the drift parameter and discussed the asymptotic properties of the maximum likelihood estimator (MLE) derived from the limiting process by using Fourier analytic methods. Yoshida[9] studied the maximum likelihood estimation of the unknown parameter in the drift coefficient of a diffusion process based on an approximate likelihood given by the discrete observations. For a more extensive discussion, see Ref. [1].

Approximation of the MLE of the drift parameter based on continuous observations in a linear SDE by estimation based on discrete data when T is fixed and $\Delta_n \to 0$ as $n \to \infty$ has been studied in Ref. [10] by using that Ito type approximation for a stochastic integral and the "rectangular rule" of approximation for an ordinary integral. Mishra and Bishwal[11] studied the same problem by using the "trapezoidal rule" of approximation for both the stochastic and the ordinary integrals when the observations are taken at equally spaced time points $0 = t_0, t_1, t_2, \ldots, t_N = T$. When T is fixed, $t_K = K\delta_N$, $K = 0, 1, 2, \ldots N$. Obviously $\delta_N = t_K - t_{K-1}$ and $\delta_N \to 0$ as $N \to \infty$.

In this paper we consider the approximation of the MLE of the parameter in the nonlinear drift coefficient of an Ito SDE. We have transformed the Ito integral to a Stratonovich integral and used the "trapezodial rule" of approximation for both the stochastic and the ordinary integrals. Further we have shown that the approximate MLE by the "trapezoidal rule" of approximation converges in probability to the MLE based on the continuous observation of the process over $[0, T]$ and their difference is bounded in probability with order $O_p(\delta_N)$. The approximate maximum likelihood estimator for the estimation of a parameter in a linear SDE obtained by Le Breton[10] by using the "rectangular rule" of approximation converges in probability to the MLE based on the continuous observation over $[0, T]$ and the error is bounded in probability with order $O_p(\delta_N^{1/2})$. This shows that our method provides a sharper rate of convergence. The paper is organised as follows. Section 2 contains assumptions, definitions of some basic notions and preliminaries. Section 3 deals with the connection between the Stratonovich integral and Ito integral. In Section 4 we describe the properties of the approximate log-likelihood function based on discrete sampling. In Section 5 the main results of the paper dealing with the approximation of a MLE have been discussed. Proofs of lemmas and theorems have been given in the Section 6.

2. ASSUMPTIONS AND PRELIMINARIES

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\{\mathcal{F}_t, t \geq 0\}$ be a family of sub σ-algebra of $\mathcal{F}$ increasing in t so that $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$. Let $\{X_t, t \geq 0\}$ be a real valued stochastic process adapted to $\{\mathcal{F}_t, t \geq 0\}$ satisfying the stochastic differential equation

$$dx(t) = f(\theta, X(t))dt + dW(t), \quad X(0) = X_0, \quad t \geq 0 \tag{2.1}$$

where $EX_0^4 < \infty, f(\theta, \cdot)$ is the drift coefficient and $\theta \in \Theta \subset [-1, 1]$. Let θ_0 be the true value of the parameter. Here $\{W(t), t \geq 0\}$ is the standard Wiener process adapted to $\{\mathcal{F}_t, t \geq 0\}$ such that, for $0 \leq s \leq t$, $\{W(t) - W(s)\}$ is independent of $\mathcal{F}_s$. Let us denote by P_θ the measure generated, on the measurable space $(C_T, B_T\}$ of continuous functions on $[0, T]$ with the associated Borel σ-algebra B_T generated under the supremum norm, by the process $X_0^T = \{X(t), 0 \leq t \leq T\}$ satisfying (2.1). Let E_θ be the expectation with respect to the measure P_θ and P_W be the measure induced by the standard Wiener process on (C_T, B_T). Through out the paper we shall use C_1, C_2, C_3 etc. for positive constants and assume that the following conditions hold.

(A1) (i) $|f(\theta, x)| \leq L(\theta)(1 + |x|)$ for $\theta \in \Theta$, $x \in \mathbf{R}$ and

(ii) $|f(\theta, x) - f(\theta, y)| \leq L(\theta)|x - y|$ for $x, y \in \mathbf{R}$ and $\theta \in \Theta$.

It is well known that the equation (2.1) has a unique solution $\{X(t), 0 \leq t \leq T\}$ under these two conditions. Further suppose that the process is a stationary process and sufficient conditions for the existence of a stationary solution given in [12] are satisfied. In addition, suppose that the following conditions hold.

(A2) For all $\theta_1 \neq \theta_2$ in Θ, $E_\theta\{f(\theta_1, X(0)) - f(\theta_2, X(0))\}^2 \neq 0$.

(A3) (i) $f(\theta, x)$ is a known real valued continuous function on $\Theta \times \mathbf{R}$;

(ii) The function $f(\theta, x)$ is thrice differentiable with respect to x. In addition, $f^{(i)}_{\theta}$, $f^{(i)}_{x\theta}$; and $f^{(i)}_{xx\theta}$ denoting the partial derivatives with respect to θ of f, $\partial f/\partial x$ and $\partial^2 f/\partial x^2$ respectively of order $i = 0, 1, 2, 3$ exist, where the differentiation of order zero gives the function itself.

(A4) $E_\theta[\sup_{|t|\leq h} f^2(\theta, X(t))] = O(1)$ for all $\theta \in \Theta$.

Further assume that the following Lipschitz conditions are satisfied for $i = 1, 2, 3$, and θ and $\phi \in \Theta$. There exists $\alpha > 0$, such that

(A5) (i) $|f^{(i)}_{x\theta}(\theta, x) - f^{(i)}_{x,\theta}(\phi, x) \leq p_i(x)|\theta - \phi|^\alpha$ and $E_\theta p_i^2(X_0) < \infty$;

(ii) $|f^{(i)}_{xx\theta}(\theta, x) - f^{(i)}_{xx\theta}(\phi, x) \leq q_i(x)|\theta - \phi|^\alpha$ and $E_\theta q_i^2(X_0) < \infty$; and denoting $f^2(\theta, x)$ by $g(\theta, x)$,

(iii) $|g^{(i)}_{x\theta}(\theta, x) - g^{(i)}_{x,\theta}(\phi, x)| \leq r_i(x)|\theta - \phi|^\alpha$ and $E_\theta r_i^2(X_0) < \infty$.

Let the Fourier expansions of $f_x(\theta, x)$, $f_{xx}(\theta, x)$ and $g_x(\theta, x)$ in $L_2[-1,1]$ and $x \in \mathbf{R}$ be given by

$$f_x(\theta, x) = \sum_{n=1}^{\infty} a_n(x) e^{\pi i n \theta},$$

$$f_{xx}(\theta, x) = \sum_{n=1}^{\infty} b_n(x) e^{\pi i n \theta},$$

and

$$g_x(\theta, x) = \sum_{n=1}^{\infty} d_n(x) e^{\pi i n \theta}.$$

Under the Lipschitz condition (A5) and using Lemma 2 (appendix) of Ref. [8], it can be shown that for some $0 < \gamma < \alpha$,

$$\sum_{n=1}^{\infty} E|a_n(X(0))|^2 n^{2(1+\gamma)} < \infty. \tag{2.2}$$

Therefore, using the Borel-Cantelli Lemma, we get that, for sufficiently large n,

$$|a_n(X(0))| \leq \frac{C_1}{n^{1+\gamma}} \quad \text{a.s.} \tag{2.3}$$

Similarly it can be shown that,

$$|b_n(X(0))| \leq \frac{C_2}{n^{1+\gamma}} \quad \text{a.s.} \tag{2.4}$$

and

$$|d_n(X(0))| \leq \frac{C_3}{n^{1+\gamma}} \quad \text{a.s.} \tag{2.5}$$

Under the assumption (A4), it is obvious that $E_\theta \int_0^T f^2(\theta, X(t))dt < \infty$ for $T > 0$ and all $\theta \in \Theta$ by the stationarity of the process. Therefore the measures P_θ and P_W are equivalent and the Radon-Nikodym derivative of P_θ with respect to P_W is given by

$$\begin{aligned}\frac{dP_\theta}{dP_W}(X_0^T) &= \exp\left\{\int_0^T f(\theta, X(t))dX(t) - \frac{1}{2}\int_0^T f^2(\theta, X(t))dt\right\} \\ &= \exp\left\{\int_0^T f(\theta, X(t))dW(t) + \frac{1}{2}\int_0^T f^2(\theta, X(t))dt\right\}\end{aligned} \tag{2.6}$$

by using (2.1) and the log likelihood function $L_T(\theta) = \log \frac{dP_\theta}{dP_W}(X_0^T)$ is a well defined measurable function (cf. [13] or [14]).

3. CONNECTION BETWEEN ITO INTEGRAL AND STRATONOVICH INTEGRAL

Let π_N be a partition such that $\pi_N = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ for fixed $T, K = 0, 1, 2, \ldots N$ and $t_u = u\delta_N$. Obviously $\delta_N \to 0$ as $N \to \infty$. The Ito integral is defined in P_θ-measure as

$$\int_0^T f(\theta, X(t))dW(t) = \lim_{\delta_N \to 0} \sum_{i=1}^N f(\theta, X(t_{i-1}))(W(t_i) - W(t_{i-1})) \tag{3.1}$$

whenever it exists and the Rubin–Fisk–Stratonovich integral is defined as

$$\begin{aligned}&\oint_0^T f(\theta, X(t))\, dW(t) \\ &\quad = \lim_{\delta_N \to 0} \sum_{i=1}^N \frac{f(\theta, X(t_{i-1})) + f(\theta, X(t_i))}{2}(W(t_i) - W(t_{i-1}))\end{aligned} \tag{3.2}$$

in P_θ-measure. Here, the limit is interpreted as convergence in probability. Under the condition (A4) and the existence of $f_x(\theta, x)$, Stratonovich[15] has shown that

$$\oint_0^T f(\theta, X(t))dW(t) = \int_0^T f(\theta, X(t))dW(t) + \frac{1}{2}\int_0^T f_x(\theta, X(t))dt. \tag{3.3}$$

4. APPROXIMATION OF THE LOG-LIKELIHOOD FUNCTION

Let $\{\pi_N, N \geq 1\}$ be a sequence of partitions, $\pi_N = \{0 = t_0 < t_1 < t_2 < \cdots < t_N = T\}$ such that for fixed $T, K = 1, 2, \ldots N$, $t_K = K\delta_N$, $t_K - t_{K-1} = \delta_N = T/N \to 0$ as $N \to \infty$ as defined above. Let

$$L^{(i)}_{N,T}(\theta) = \sum_{K=1}^{N} \left\{ \frac{f^{(i)}(\theta, X(t_K)) + f^{(i)}(\theta, X(t_{K-1}))}{2} (W(t_K) - W(t_{K-1})) \right\}$$

$$+ \frac{1}{2} \sum_{K=1}^{N} \left\{ \frac{g^{(i)}(\theta, X(t_K)) + g^{(i)}(\theta, X(t_{K-1}))}{2} \right\} (t_K - t_{K-1})$$

$$- \frac{1}{2} \sum_{K=1}^{N} \left\{ \frac{f^{(i)}_x(\theta, X(t_K)) + f^{(i)}_x(\theta, X(t_{K-1}))}{2} \right\} (t_K - t_{K-1}) \tag{4.1}$$

for $i = 0, 1, 2, 3$. The function $L^{(i)}_{N,T}(\theta)$ can be considered as an approximation of the function

$$L^{(i)}_T(\theta) = \oint_0^T f^{(i)}(\theta, X(t)) dW(t) + \frac{1}{2} \int_0^T g^{(i)}(\theta, X(t)) dt - \frac{1}{2} \int_0^T f^{(i)}_x(\theta, X(t)) dt \tag{4.2}$$

by using (3.3). We now investigate the difference between $L^{(i)}_{N,T}(\theta)$ and $L^{(i)}_T(\theta)$ and study their properties in the following theorem.

Theorem 4.1. *Let the conditions* (*A*1)–(*A*5) *hold. Then, for* $i = 0, 1, 2$, *there exists a sequence of positive random variables* $\{\Delta^{(i)}_{N,T}, N > 1\}$ *such that for all* $\theta^* \in \Theta$,

(i) $\{\Delta^{(i)}_{N,T}, N > 1\}$ *is bounded in* P_{θ^*}*-measure, and*
(ii) $\sup_{\theta \in \Theta} |L^{(i)}_{N,T}(\theta) - L^{(i)}_T(\theta)| \leq \delta_N^2 \Delta^{(i)}_{N,T}$, P_{θ^*}*-almost surely.*

We need following lemma to prove the theorem. Let θ_0 be the true parameter.

Lemma 4.1. *Let the assumptions* (A1)–(A5) *hold. Then, in* P_{θ_0}*-measure, the following results are true:*

(i) $$\sup_{\theta\in\Theta}\sum_{K=1}^{N}\left|\left\{\frac{f^{(i)}(\theta,X(t_K))+f^{(i)}(\theta,X(t_{K-1}))}{2}\right\}(W(t_K)-W(t_{K-1})) - \oint_0^T f^{(i)}(\theta,X(t))dW(t)\right| \le \delta_N^2 U_N^{(i)},$$

(ii) $$\sup_{\theta\in\Theta}\sum_{K=1}^{N}\left|\left\{\frac{f_x^{(i)}(\theta,X(t_K))+f_x^{(i)}(\theta,X(t_{K-1}))}{2}\right\}(t_K-t_{K-1}) - \int_0^T f_x^{(i)}(\theta,X(t))dt\right| \le \delta_N^2 V_n^{(i)}, \quad and$$

(iii) $$\sup_{\theta\in\Theta}\sum_{K=1}^{N}\left|\left\{\frac{g^{(i)}(\theta,X(t_K))+g^{(i)}(\theta,X(t_{K-1}))}{2}\right\}(t_K-t_{K-1}) - \int_0^T g^{(i)}(\theta,X(t))dt\right| \le \delta_N^2 Z_N^{(i)}$$

where $\{U_N^{(i)}, N\ge 1\}$, $\{V_N^{(i)}, N\ge 1\}$ *and* $\{Z_N^{(i)}, N\ge 1\}$ *are sequence of random variables which are bounded in* P_{θ_0}*-probability.*

Proof of this lemma is given in Section 6.

Proof of Theorem 4.1. Using (4.1) and (4.2) and Lemma 4.1 we have,

$$\begin{aligned}
&\sup_{\theta\in\Theta}|L_{N,T}^{(i)}(\theta)-L_T^{(i)}(\theta)| \\
&= \sup_{\theta\in\Theta}\left|\sum_{K=1}^{N}\left\{\left(\frac{f^{(i)}(\theta,X(t_K))+f^{(i)}(\theta,X(t_{K-1}))}{2}\right)(W(t_K)-W(t_{K-1}))\right.\right. \\
&\quad +\frac{1}{2}\left(\frac{g^{(i)}(\theta,X(t_K))+g^{(i)}(\theta,X(t_{K-1}))}{2}\right)(t_K-t_{K-1}) \\
&\quad \left. -\frac{1}{2}\left(\frac{f^{(i)}(\theta,X(t_K))+f^{(i)}(\theta,X(t_{K-1}))}{2}\right)(t_K-t_{K-1})\right\} \\
&\quad \left. -\left\{\oint_0^T f^{(i)}(\theta,X(t))dW(t)+\frac{1}{2}\int_0^T g^{(i)}(\theta,X(t))dt-\frac{1}{2}\int_0^T f_x(\theta,X(t))dt\right\}\right| \\
&\le \delta_N^2(U_N^{(i)}+V_N^{(i)}+Z_N^{(i)}) \le \delta_N^2\Delta_{N,T}^{(i)}
\end{aligned}$$

where $U_N^{(i)}+V_N^{(i)}+Z_N^{(i)}=\Delta_{N,T}^{(i)}$ is bounded in $P\theta_0^*$-probability. □

5. APPROXIMATION OF THE MAXIMUM LIKELIHOOD ESTIMATOR

We now state the main results of the paper. In addition to the conditions (A0) to (A5), assume that

(A6) Differentiation twice under the stochastic integral sign with respect to θ is valid in

$$\int_0^T f(\theta, X(t))dW(t)$$

and in the integral

$$\int_0^T f(\theta, X(t))dt.$$

Let $\hat{\theta}_T$ be the MLE of the parameter θ based on the observations of the process $\{X(t), 0 \leq t \leq T\}$ described by (2.1) over $[0, T]$.

Theorem 5.1. *Let us assume the conditions* (A1)–(A6)*. In addition, suppose that* (A7) *there exists a unique solution* $\hat{\theta}_T$ *to the equation* $L_T^{(1)}(\hat{\theta}_T) = 0$ *and the function* $L_T^{(2)}(\tilde{\theta}_T)$ *is strictly negative* $P_{\theta*}$*-almost surely whenever* $|\tilde{\theta}_T - \hat{\theta}_T| \leq 1$. *Then there exists a sequence* $\{\hat{\theta}_{N,T}, N \geq 1\}$ *of random variables satisfying*

(i) $\hat{\theta}_{N,T}$ is ζ_N*-measurable where* ζ_N *denotes the* σ*-algebra generated by* $(X(t_1), X(t_2), \ldots X(t_N))$,
(ii) $lim_{N\to\infty} P_{\theta^*}\ \{L_{N,T}^{(1)}(\hat{\theta}_{N,T}) = 0\} = 1$, *and*
(iii) $P_{\theta*} - lim_{N\to\infty} \hat{\theta}_{N,T} = \hat{\theta}_T$

for all $\theta^* \in \Theta$. *Further if* $\{\tilde{\theta}_{N,T}, N \geq 1\}$ *is another sequence of random variables satisfying* (i), (ii) *and* (iii) *above, then for all* $\theta^* \in \Theta$,

$$\lim_{N\to\infty} P_{\theta^*}\{\hat{\theta}_{N,T} = \tilde{\theta}_{N,T}\} = 1.$$

Lastly if $\{\hat{\theta}_{N,T}, N \geq 1\}$ *is any sequence of random variables satisfying* (i), (ii) *and* (iii) *above, then the sequence* $\{\delta_N^{-1}(\hat{\theta}_{N,T} - \hat{\theta}_T), N \geq 1\}$ *is bounded in* $P_{\theta*}$*-measure.*

Proof. We write for $|x| \leq 1$,

$$L_{N,T}^{(1)}(\hat{\theta}_T + x) = L_{N,T}^{(1)}(\hat{\theta}_T) + xL_T^{(2)}(\theta'_T) + x\{L_{N,T}^{(2)}(\theta'_T) - L_T^{(2)}(\theta'_T)\},$$

where $\theta'_T = \hat{\theta}_T + \lambda x$, $\lambda \in [0, 1]$.

Let $\bar{K} > 1$ be a given constant. The function $L^{(1)}_{N,T}(\hat{\theta}_T + x)$ is continuous and differentiable on $|x| \leq 1/\bar{K}$ and

$$xL^{(1)}_{N,T}(\hat{\theta}_T + x) = xL^{(1)}_{N,T}(\hat{\theta}_T) + x^2 L^{(2)}_T(\theta'_T) + x^2\{L^{(2)}_{N,T}(\theta'_T) - L^{(2)}_T(\theta'_T)\}.$$

Note that $L^{(2)}_T(\tilde{\theta}_T)$ is almost surely negative by (A7) whenever $|\tilde{\theta}_T - \hat{\theta}_T| \leq 1$. Hence $\inf_{|x|=1}\{-x^2 L^{(2)}_T(\theta'_T)\} = q_T > 0$ almost surely and $-x^2 L^{(2)}_T(\theta'_T) \geq q_T x^2$. Therefore $xL^{(1)}_{N,T}(\hat{\theta}_T + x)$ shall be strictly negative for $|x| = 1/\bar{K}$ if $|L^{(1)}_{N,T}(\hat{\theta}_T)| < (q_T)/2\bar{K}$ and

$$\sup_{|x| \leq \frac{1}{\bar{K}}} |L^{(2)}_{N,T}(\theta'_T) - L^{(2)}_T(\theta'_T)| < \frac{q_T}{2}.$$

Let us consider, for $\bar{K} > 1$,

$$\Omega_N(\bar{K}) = \left\{\omega \in \Omega : \delta_N^2 \Delta^{(1)}_{N,T}(\omega) < \frac{q_T(\omega)}{2\bar{K}}, \delta_N^2 \Delta^{(2)}_{N,T}(\omega) < \frac{q_T(\omega)}{2} \quad \text{and} \right.$$
$$\left. |\hat{\theta}_T(\omega)| < 1 - \frac{1}{\bar{K}}\right\}$$

For $\omega \in \Omega_N(\bar{K})$, we have

$$\sup_{|\theta| \leq 1 - \frac{1}{\bar{K}}} |L^{(1)}_{N,T}(\theta, \omega) - L^{(1)}_T(\theta, \omega)| < \frac{q_T(\omega)}{2\bar{K}}$$

and since $|\hat{\theta}_T| \leq 1 - 1/\bar{K}$, we have

$$|L^{(1)}_{N,T}(\hat{\theta}_T, \omega) - L^{(1)}_T(\hat{\theta}_T, \omega)| < \frac{q_T(\omega)}{2\bar{K}}.$$

Further more,

$$\sup_{|\theta| \leq 1} |L^{(2)}_{N,T}(\theta, \omega) - L^{(2)}_T(\theta, \omega)| < \frac{q_T(\omega)}{2}.$$

Since $|\hat{\theta}_T| \leq 1 - 1/\bar{K}$, it follows that, for $|x| \leq 1/\bar{K}$,

$$\sup_{|x| \leq \frac{1}{\bar{K}}} |L^{(2)}_{N,T}(\hat{\theta}_T + x, \omega) - L^{(2)}_T(\hat{\theta}_T + x, \omega)| < \frac{q_T(\omega)}{2}.$$

Moreover, for $\omega \in \Omega_N(\bar{K})$, $L^{(1)}_{N,T}(\hat{\theta}_T + x)$ is continuous on $|x| \leq 1/\bar{K}$ and for $|x| = 1/\bar{K}$,

$$xL^{(1)}_{N,T}(\hat{\theta}_T + x) < 0.$$

By using a lemma of Aitchinson and Silvey,[16] for $\omega \in \Omega_N(\bar{K})$, we obtain, that the equation $L^{(1)}_{N,T}(\hat{\theta}_T + x, \omega) = 0$ has a solution $X_N(\omega)$satisfying$|X_N(\omega)| < 1/\bar{K}$.

Following the procedure in Ref. [10] and considering the changes in the scale of x and the parameter θ as mentioned above, we can prove the following results.

(i) For $|x| \leq 1/\bar{K}$ and $\omega \in \Omega_N(\bar{K})$,

$$L_{N,T}(\hat{\theta}_T + x, \omega) < L_{N,T}(\hat{\theta}_T + X_N, \omega).$$

(ii) $\lim_{\bar{K}\to\infty} \liminf_{N\to\infty} P_{\theta^*}(\Omega_N(\bar{K})) = 1$.

(iii) There exists a sequence of random variables $\{\hat{\theta}_{N,T}, N \geq 1\}$ such that

$$\lim_{N\to\infty} P_{\theta^*}\{L^{(1)}_{N,T}(\hat{\theta}_{N,T}) = 0\} = 1.$$

(iv) If there is another sequence of random variably $\{\hat{\theta}'_{N,T}, N \geq 1\}$ satisfying the above condition (iii), then

$$\lim_{N\to\infty} P_{\theta^*}(\hat{\theta}'_{N,T} = \hat{\theta}_{N,T}) = 1$$

(v) $\lim_{A\to\infty} \limsup_{N\to\infty} P_{\theta^*}\{\delta_N^{-2}|\hat{\theta}_{N,T} - \hat{\theta}_T| > A\} = 0.$

□

6. PROOF OF LEMMA 4.1

We now prove the Lemma 4.1 following the techniques in Ref. [17]. Let θ_0 be the true parameter. First of all we prove assertion (i). Let

$$S(\theta) = \oint_0^T f(\theta, X(t))dW(t)$$

and

$$S^{\theta}_{\pi_N} = \sum_{K=1}^{N} \left\{\frac{f(\theta, X(t_K)) + f(\theta, X(t_{K-1}))}{2}\right\}(W(t_K) - W(t_{K-1})) \tag{6.1}$$

where π_N is a partition as defined previously. Let $\pi_{N^{(1)}}$ be a partition finer than π_N obtained by choosing the midpoints $\tilde{t}_K$ from each of the intervals $t_{K-1} < \tilde{t}_{K-1} < t_K$, $K = 1, 2, \ldots N$. Let $0 = t'_0 < t'_1 < t'_2 < \cdots < t'_{2N} = T$ be the points of subdivision of the refined partition $\pi_{N^{(1)}}$. Let us define the approximating sum $S^{\theta}_{\pi_N^{(1)}}$ as

$$S^{\theta}_{\pi_N^{(1)}} = \sum_{K=1}^{2N} \frac{f(\theta, X(t'_K)) + f(\theta, X(t'_{K-1}))}{2}(W(t'_K) - W(t'_{K-1})). \tag{6.2}$$

Let $0 = t_0^* \leq t_1^* \leq t_2^* \leq T$ be three equally spaced points over $[0, T]$ and let us denote $X(t_K^*)$ and $W(t_K^*)$ by X_K and W_K respectively for $K = 0, 1, 2$.

Define

$$\begin{aligned} Z=&\left\{\frac{f(\theta, X_2)+f(\theta, X_0)}{2}\right\}(W_2 - W_0) \\ &-\left\{\left(\frac{f(\theta, X_2)+f(\theta, X_1)}{2}\right)(W_2 - W_1)+\left(\frac{f(\theta, X_1)+f(\theta, X_0)}{2}\right)(W_1 - W_0)\right\} \\ =&\frac{W_1 - W_0}{2}\{f(\theta, X_2)-f(\theta, X_1)\}+\frac{W_2 - W_1}{2}\{f(\theta, X_0)-f(\theta, X_1)\}. \end{aligned} \tag{6.3}$$

We know that

$$f(\theta, X_2)-f(\theta, X_1)=(X_2 - X_1)\frac{\partial f(\theta, X_1)}{\partial X_1}+\frac{(X_2 - X_1)^2}{2}\frac{\partial^2 f(\theta, \mu)}{\partial X_1^2},$$

and

$$f(\theta, X_0)-f(\theta, X_1)=(X_0 - X_1)\frac{\partial f(\theta, X_1)}{\partial X_1}+\frac{(X_0 - X_1)^2}{2}\frac{\partial^2 f}{\partial X_1^2}(\theta, \nu)$$

where $|X_1 - \mu| < |X_2 - X_1|$ and $|X_1 - \nu| < |X_1 - X_0|$. Substituting these values in (6.3), we have,

$$\begin{aligned} Z=&\frac{(W_1 - W_0)}{2}(X_2 - X_1)\frac{\partial f(\theta, X_1)}{\partial X_1}+\frac{(W_1 - W_0)}{2}\frac{(X_2 - X_1)^2}{2}\frac{\partial^2 f(\theta, \mu)}{\partial X_1^2} \\ &+\frac{(W_2 - W_1)}{2}(X_0 - X_1)\frac{\partial f(\theta, X_1)}{\partial X_1} \\ &+\frac{(W_2 - W_1)}{2}\frac{(X_2 - X_1)^2}{2}\frac{\partial^2 f(\theta, \nu)}{\partial X_1^2}. \end{aligned} \tag{6.4}$$

Note that

$$dX(t)=f(\theta_0, X(t))dt+dW(t)$$

since θ_0 is the true parameter. Let

$$I_2^{(2K+1)}=\int_{t_{2K}}^{t_{2K+1}} f(\theta_0, X(t))dt.$$

Then

$$X_{2K+1}-X_{2K}=\int_{t_{2K}}^{t_{2K+1}} f(\theta_0,X(t))dt+(W_{2K+1}-W_{2K})$$
$$=I_2^{(2K+1)}+(W_{2K+1}-W_{2K}) \tag{6.5}$$

and

$$E_{\theta_0}(I_2^{(2K+1)})^2 \leq C_4(t_{2K+1}-t_{2K})^2$$

by (A4). Similarly we write

$$X_{2K-1}-X_{2K}=\int_{t_{2K}}^{t_{2K-1}} f(\theta_0,X(t))dt+(W_{2K-1}-W_{2K})$$
$$=I_1^{(2K)}+(W_{2K-1}-W_{2K}) \tag{6.6}$$

and

$$E_{\theta_0}(I_1^{(2K)})^2 \leq C_5(t_{2K-1}-t_{2K})^2.$$

by (A4). Now let us define a random variable Z as in (6.4) for each sub interval of the subdivision $\pi_N=\{0=t_0<t_1<\cdots<t_{2N}=T\}$ and denote this for the interval $[t_{K-1},t_K)$ by Z_K, $K=1,2,\ldots 2N$. Then

$$Z_K$$
$$=\left(\frac{W_{2K}-W_{2K-1}}{2}\right)\left\{I_2^{(2K+1)}f_x(\theta,X_{2K})+\left(\frac{X_{2K+1}-X_{2K}}{2}\right)^2 f_{xx}(\theta,\mu_{2K})\right\}$$
$$+\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\left\{I_1^{(2K)}f_x(\theta,X_{2K})+\frac{(X_{2K-1}-X_{2K})^2}{2}f_{xx}(\theta,\nu_{2K})\right\}$$
$$=Z_K^{(1)}+Z_K^{(2)} \text{ (say)} \tag{6.7}$$

where $Z_K^{(1)}$ and $Z_K^{(2)}$ are the first and the second expressions in (6.7). Consider

$$Z_K^{(2)}=\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\left\{I_1^{(2K)}f_x(\theta,X_{2K})+\frac{(X_{2K-1}-X_{2K})^2}{2}f_{xx}(\theta,\nu_{2K})\right\}$$

and the following Fourier expansions:

$$f_X(\theta,x)=\sum_{n=1}^{\infty}a_n(x)e^{\pi in\theta},$$

and

$$f_{xx}(\theta,x)=\sum_{K=1}^{n}b_n(x)e^{\pi in\theta}.$$

Then

$$\sum_{K=1}^{2N}Z_K^{(2)}$$

$$=\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\left\{I_1^{(2K)}f_x(\theta,X_{2K})+\frac{(X_{2K-1}-X_{2K})^2}{2}f_{xx}(\theta,\nu_{2K})\right\}$$

$$=\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)I_1^{(2K)}f_x(\theta,X_{2K})$$

$$+\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\left(\frac{X_{2K-1}-X_{2K}}{2}\right)^2 f_{xx}(\theta,\nu_{2K})$$

for some sequence ν_{2K}, $K\geq 1$. Then

$$\sum_{K=1}^{2N}Z_K^{(2)}=\sum_{k=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\left\{I_1^{(2K)}\left(\sum_{n=1}^{\infty}a_n(X_{2K})e^{\pi in\theta}\right)\right\}$$

$$+\sum_{k=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\frac{(X_{2K-1}-X_{2K})^2}{2}\left(\sum_{n=1}^{\infty}b_n(\nu_{2K})e^{\pi in\theta}\right)$$

$$=R_N^{(1)}+R_N^{(2)}\text{ (say).}$$

Now, for any sequence $\varepsilon_n > 0$ such that $\sum_{n=1}^{\infty} \varepsilon_n < \infty$,

$$
\begin{aligned}
&P_{\theta_0}\left\{\sup_{\theta\in\Theta}|R_N^{(1)}| \geq \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \\
&= P_{\theta_0}\left\{\sup_{\theta\in\Theta}\left|\sum_{k=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\left\{I_1^{(2K)}\left(\sum_{n=1}^{\infty}a_n(X_{2K})e^{\pi in\theta}\right)\right\}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \\
&= P_{\theta_0}\left\{\sup_{\theta\in\Theta}\left|\sum_{n=1}^{\infty}a_n(X_{2K})e^{\pi in\theta}\left(\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)I_1^{(2K)}\right)\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \\
&\leq P_{\theta_0}\left\{\sup_{\theta\in\Theta}\sum_{n=1}^{\infty}\left|a_n(X_{2K})e^{\pi in\theta}\left(\sum_{k=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)I_1^{(2K)}\right)\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \\
&\leq P_{\theta_0}\left\{\sum_{n=1}^{\infty}|a_n(X_{2K})|\left|\sum_{k=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)I_1^{(2K)}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \\
&\leq P_{\theta_0}\left\{\sum_{n=1}^{\infty}\frac{C_1}{n^{1+\gamma}}\left|\sum_{k=1}^{2N}(W_{2K+1}-W_{2K})I_1^{(2K)}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \text{(by using (2.3))} \\
&\leq \sum_{n=1}^{\infty}P_{\theta_0}\left\{\frac{C_1}{n^{1+\gamma}}\left|\sum_{k=1}^{2N}(W_{2K+1}-W_{2K})I_1^{(2K)}\right| > \frac{\varepsilon_n}{2}\right\} \\
&\leq \sum_{n=1}^{\infty}\frac{4C_1^2}{(n^{1+\gamma})^2\varepsilon_n^2}\left(\sum_{K=1}^{2N}E(W_{2K+1}-W_{2K})^2E(I_1^{(2K)})^2\right) \\
&\leq \sum_{n=1}^{\infty}\frac{4C_1^2}{n^{2+2\gamma}\varepsilon_n^2}\left(\sum_{K=1}^{2N}(t_{2K+1}-t_{2K})(t_{2K}-t_{2K-1})\right)^2 = C_6\left(\sum_{n=1}^{\infty}\frac{1}{n^{2+2\gamma}\varepsilon_n^2}\right)\frac{T^3}{N^2}.
\end{aligned}
\tag{6.8}
$$

Again

$$
\begin{aligned}
&P_{\theta_0}\left\{\sup_{\theta\in\Theta}|R_N^{(2)}| \geq \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\} \\
&= P_{\theta_0}\left\{\sup_{\theta}\left|\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\frac{(X_{2K-1}-X_{2K})^2}{2}\left(\sum_{n=1}^{\infty}b_n(\nu_{2K})e^{\pi in\theta}\right)\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\}
\end{aligned}
$$

$$= P\left\{\sup_{\theta\in\Theta}\left|\sum_{n=1}^{\infty} b_n(\nu_{2K})e^{\pi i n\theta}\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\frac{(X_{2K-1}-X_{2K})^2}{2}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\}$$

$$\le P\left\{\sup_{\theta\in\Theta}\sum_{n=1}^{\infty}\left|b_n(\nu_{2K})e^{\pi i n\theta}\sum_{K=1}^{2N}\left(\frac{W_{2K+1}-W_{2K}}{2}\right)\frac{(X_{2K-1}-X_{2K})^2}{2}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\}$$

$$\le P\left\{\sup_{\theta\in\Theta}\sum_{n=1}^{\infty}|b_n(\nu_{2K})|\left|\sum_{K=1}^{2N}\frac{(W_{2K+1}-W_{2K})}{2}\frac{(X_{2K-1}-X_{2K})^2}{2}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\}$$

$$\le P\left\{\sum_{n=1}^{\infty}\frac{C_2}{n^{1+\gamma}}\left|\sum_{K=1}^{2N}(W_{2K+1}-W_{2K})\frac{(X_{2K-1}-X_{2K})^2}{2}\right| > \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right\}\text{(by using 2.4)}$$

$$\le C_7\sum_{n=1}^{\infty}\frac{1}{(n^{1+\gamma})^2\varepsilon_n^2}\sum_{K=1}^{2N}(t_{2K+1}-t_{2K})(t_{2K-1}-t_{2K})^2.$$

Applying the following inequality of Gihman and Skorohod,[12]

$$E_{\theta_0}(X_{2K-1}-X_{2K})^4 = E_{\theta_0}(X_1-X_0)^4 \le C_7(E_{\theta_0}X_0^4+1)\left(\frac{T}{2N}\right)^2$$

where C is a constant, it follows that

$$P_\theta\left(\sup_{\theta\in\Theta}|R_N^{(2)}| \ge \sum_{n=1}^{\infty}\frac{\varepsilon_n}{2}\right) \le C_8\sum_{n=1}^{\infty}\frac{1}{n^{2+2\gamma}\varepsilon_n^2}\frac{T^3}{N^2}. \tag{6.9}$$

Therefore, combining (6.8) and (6.9), we have

$$P\left\{\left|\sum_{K=1}^{2N}Z_K^{(2)}\right| > \sum_{n=1}^{\infty}\varepsilon_n\right\} \le C_9\frac{T^3}{N^2}\sum_{n=1}^{\infty}\frac{1}{n^{2+2\gamma}\varepsilon_n^2}. \tag{6.10}$$

Similar estimate can be obtained for the term $\sum_{K=1}^{2N}Z_K^{(1)}$ by using the reverse martingale property of the differences $Z_K^{(1)}$ and the stationarity of the process $\{X(t),\ 0\le t\le T\}$. Thus

$$P_{\theta_0}\left\{\left|\sum_{K=1}^{2N}Z_K^{(1)}\right| > \sum_{n=1}^{\infty}\varepsilon_n\right\} \le C_{10}\frac{T^3}{(2N)^2}\sum_{n=1}^{\infty}\frac{1}{n^{2+2\gamma}\varepsilon_n^2}. \tag{6.11}$$

Therefore, combining (6.10) and (6.11), we have

$$P_{\theta_0}\left\{\sup_{\theta\in\Theta}|S_{\pi_N}^{\theta}-S_{\pi_{N(1)}}^{\theta}| \ge \sum_{n=1}^{\infty}2\varepsilon_n\right\} \le C_{11}\frac{T^3}{N^2}\sum_{n=1}^{\infty}\frac{1}{(n^{1+\gamma})^2\varepsilon_n^2} \le C_{12}\frac{T^3}{N^2\varepsilon^2}.$$

Choosing $\varepsilon_n = \varepsilon(1/n^{1+\gamma})^{\beta}$, where $1/2 < \beta < 3/4$ and $\varepsilon > 0$, is arbitrary, we get,

$$P_{\theta_0}\left\{\sup_{\theta\in\Theta} |S^{\theta}_{\pi_N} - S^{\theta}_{\pi_{N(1)}}| > \varepsilon A\right\} \le C_{13}\frac{T^3}{N^2\varepsilon^2}$$

where $A = \sum_{n=1}^{\infty}\frac{1}{n^{\beta(1+\gamma)}}$.

Let $\pi_N(p)$ be a sequence of partitions of $[0, T]$ as defined earlier. Then we have,

$$P_{\theta_0}\left\{\sup_{\theta\in\Theta} |S^{\theta}_{\pi_N} - S^{\theta}_{\pi_{N(p+1)}}| > \rho\right\} \le C_{14}\sum_{K=0}^{p}\frac{1}{\rho_K^2(2^K N)^2}$$

where $\sum_{K=1}^{\infty}\rho_K \le \rho$. Choosing ρ_K suitably,

$$P_{\theta_0}\left\{\sup_{\theta\in\Theta} |S^{\theta}_{\pi_N} - S^{\theta}_{\pi_{N(p+1)}}| > \rho\right\} \le C_{15}N^{-2}\rho^{-2}, p \ge 0.$$

Letting $p \to \infty$, we have $S^{\theta}_{\pi_N} \to S^{\theta}$ uniformly in θ in P_{θ_0}-measure.

Let us now prove the assertion (iii). Let $f^2 = g$. Let π_N and $\pi_N(1)$ be partitions defined exactly as before.

Define

$$M^{\theta} = \int_0^T g(\theta, X(t))dt,$$

$$M^{\theta}_{\pi_N} = \sum_{K=1}^{N}\left\{\frac{g(\theta, X(t_K)) + g(\theta, X(t_{K-1}))}{2}\right\}(t_K - t_{K-1}), \tag{6.12}$$

and

$$M^{\theta}_{\pi_N(1)} = \sum_{K=1}^{2N}\left\{\frac{g(\theta, X(t'_K)) + g(\theta, X(t'_{K-1}))}{2}\right\}(t'_K - t'_{K-1}). \tag{6.13}$$

We shall first compute bounds on $P\{\sup_{\theta\in\Theta}|M^{\theta}_{\pi_N} - M^{\theta}_{\pi_N(1)}| \ge \sum 2\varepsilon_n\}$. Let $0 \le t_0^* \le t_1^* \le t_2^* \le T$ be three equally spaced points in $[0, T]$ and let us denote

$X(t_K)$ by X_K, $K = 0, 1, 2$. Define

$$Y = \left(\frac{g(\theta,X_2)+g(\theta,X_0)}{2}\right)(t_2^* - t_0^*)$$
$$- \left\{\left(\frac{g(\theta,X_2)+g(\theta,X_1)}{2}\right)(t_2^* - t_1^*) + \left(\frac{g(\theta,X_1)+g(\theta,X_0)}{2}\right)(t_1^* - t_0^*)\right\}$$
$$= \frac{t_1^* - t_0^*}{2}\{g(\theta,X_2) - g(\theta,X_1)\} + \frac{t_2^* - t_1^*}{2}\{g(\theta,X_0) - g(\theta,X_1)\}. \tag{6.14}$$

Then

$$g(\theta,X_2) - g(\theta,X_1) = (X_2 - X_1)\bar{g}(\theta,\mu)$$

$$g(\theta,X_0) - g(\theta,X_1) = (X_0 - X_1)\bar{g}(\theta,\nu)$$

where $|X_1 - \mu) < |X_2 - X_1|$, $|X_1 - \nu| < |X_0 - X_1|$ and $\bar{g}(\theta,x) = \partial g(\theta,x)/\partial x$.

Substituting these values in (6.14), it follows that

$$Y = \frac{t_1^* - t_0^*}{2}(X_2 - X_1)\bar{g}(\theta,\mu) + \frac{t_2^* - t_1^*}{2}(X_0 - X_1)\bar{g}(\theta,\nu). \tag{6.15}$$

Now let us define a random variable analogous to Y as in (6.14) for each interval of the subdivision of $\pi_N = \{0 = t_0 < t_1 < \cdots < t_{2N} = T\}$ and denote this for the interval $[t_{K-1}, t_K)$ by Y_K, $K = 1, 2, \ldots, 2N$. Then

$$Y_K = \frac{t_{2K} - t_{2K-1}}{2}(X_{2K+1} - X_{2K})\bar{g}(\theta,\mu_{2K}) + \frac{t_{2K+1} - t_{2K}}{2}$$
$$\times (X_{2K-1} - X_{2K})\bar{g}(\theta,\nu_{2K}) = Y_K^{(1)} + Y_K^{(2)} \quad \text{(say)} \tag{6.16}$$

where $Y_K^{(1)}$ and $Y_K^{(2)}$ are the first and second expressions in (6.16).

Consider $Y_k^{(1)} = (t_{2K} - t_{2K-1}/2)(X_{2K+1} - X_{2K})g_x(\theta,\mu_{2K})$ and the Fourier expansion of

$$g_x(\theta,x) = \sum_{n=1}^{\infty} d_n(x)e^{\pi i n\theta}.$$

Then

$$\sum_{K=1}^{2N} Y_K^{(1)} = \sum_{K=1}^{2N} \frac{t_{2K} - t_{2K-1}}{2}(X_{2K+1} - X_{2K})\sum_{n=1}^{\infty} d_n(\mu_{2K})e^{\pi in\theta}$$

$$= \sum_{K=1}^{2n} \left(\frac{t_{2K} - t_{2K-1}}{2}\right)(I_2^{(2K+1)} + (W_{2K+1} - W_{2K}))\sum_{n=1}^{\infty} d_n(\mu_{2K})e^{\pi in\theta}$$

$$= \sum_{K=1}^{2N} \left(\frac{T}{2N}\right) I_2^{(2K+1)} \left\{\sum_{n=1}^{\infty} d_n(\mu_{2K})e^{\pi in\theta}\right\}$$

$$+ \sum_{K=1}^{2N} \left(\frac{T}{2N}\right)(W_{2K+1} - W_{2K})\left(\sum_{n=1}^{\infty} d_n(\mu_{2K})e^{\pi in\theta}\right)$$

$$= Q_N^{(1)} + Q_N^{(2)} \quad \text{(say).}$$

Now

$$P_{\theta_0}\left\{\sup_{\theta\in\Theta}|Q_N^{(1)}| > \frac{\varepsilon_n}{2}\right\} = P\left\{\sup_{\theta\in\Theta}\left|\sum_{K=1}^{2N}\frac{T}{2N}I_2^{(2K+1)}\left(\sum_{n=1}^{\infty} d_n(\mu_{2K})e^{\pi in\theta}\right)\right| > \sum_{K=1}^{\infty}\varepsilon_n\right\}$$

$$= P_{\theta_0}\left\{\sup_{\theta\in\Theta}\left|\sum_{n=1}^{\infty} d_n(\mu_{2K})e^{\pi in\theta}\sum_{K=1}^{2N}\frac{T}{2N}I_2^{(2K+1)}\right| > \sum_{n=1}^{\infty}\varepsilon_n\right\}$$

$$\leq P_{\theta_0}\left\{\sup_{\theta\in\Theta}\sum_{n=1}^{\infty}\left|d_n(\mu_{2K})e^{\pi in\theta}\sum_{K=1}^{2N}\frac{T}{2N}I_2^{(2K+1)}\right| > \sum\varepsilon_n\right\}$$

$$\leq P_{\theta_0}\left\{\sum_{n=1}^{\infty}\left|d_n(\mu_{2K})\frac{T}{2N}\right|\left|\sum_{K=1}^{2N}I_2^{(2K+1)}\right| > \sum_{n=1}^{\infty}\varepsilon_n\right\}$$

$$\leq P_{\theta_0}\left\{\sum_{n=1}^{\infty}\frac{C_3}{n^{1+\gamma}}\frac{\mathrm{T}}{2N}\left|\sum_{K=1}^{2N}I_2^{(2K+1)}\right| > \sum_{n=1}^{\infty}\varepsilon_n\right\} \text{(by using 2.5)}$$

$$\leq \sum_{n=1}^{\infty}P\left\{\frac{C_3}{n^{1+\gamma}}\frac{T}{2N}\left|\sum_{K=1}^{2N}I_2^{(2K+1)}\right| > \varepsilon_n\right\}$$

$$\leq \sum_{n=1}^{\infty} \frac{C_3^2}{n^{2+2\gamma}} \frac{T^2}{4N^2} E \left| \sum_{K=1}^{2N} I_2^{(2K+1)} \right|^2$$

$$\leq C_{16} \sum_{n=1}^{\infty} \frac{1}{n^{2+2\gamma} \varepsilon_n^2} \frac{T^2}{(2N)^2} \left\{ \sum_{K=1}^{2N} E(I_2^{(2K+1)})^2 + \sum_{\substack{K,\ell=1 \\ K \neq \ell}}^{2N} \sqrt{E(I_2^{(2K+1)})^2 E(I_2^{(2\ell+1)})^2} \right\}$$

$$= C_{17} \sum_{n=1}^{\infty} \frac{1}{n^{2+2\gamma} \varepsilon_n^2} \frac{T^2}{(2N)^2} \left\{ \left(\frac{T}{2N}\right)^2 + (2N)(2N-1) \left(\frac{T}{2N}\right)^2 \right\}$$

$$\leq C_{18} \sum_{n=1}^{\infty} \frac{1}{n^{2+2\gamma} \varepsilon_n^2} \frac{T^4}{(2N)^2}. \tag{16.7}$$

Again,

$$P_{\theta_0} \left\{ \sup_{\theta \in \Theta} |Q_N^{(2)}| > \frac{\varepsilon_n}{2} \right\}$$

$$= P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \left| \sum_{k=1}^{2N} \frac{T}{2N} (W_{2K+1} - W_{2K}) \left(\sum_{n=1}^{\infty} d_n(\mu_{2K}) e^{\pi i n \theta} \right) \right| > \sum_{n=1}^{\infty} \varepsilon_n \right\}$$

$$= P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \left| \sum_{n=1}^{\infty} d_n(\mu_{2K}) e^{\pi i n \theta} \left(\sum_{K=1}^{2N} (W_{2K+1} - W_{2K}) \frac{T}{2N} \right) \right| > \sum_{n=1}^{\infty} \varepsilon_n \right\}$$

$$\leq P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \sum_{n=1}^{\infty} \left| d_n(\mu_{2K}) e^{\pi i n \theta} \left(\sum_{K=1}^{2N} (W_{2K+1} - W_{2K}) \frac{T}{2N} \right) \right| > \sum_{n=1}^{\infty} \varepsilon_n \right\}$$

$$\leq P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \sum_{n=1}^{\infty} |d_n(\mu_{2K})| \left| \sum_{K=1}^{2N} (W_{2K+1} - W_{2K}) \frac{T}{2N} \right| > \sum_{n=1}^{\infty} \varepsilon_n \right\}$$

$$\leq P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \sum_{n=1}^{\infty} \left(\frac{C_3}{n^{1+\gamma}} \right) \left| \sum_{K=1}^{2N} (W_{2K+1} - W_{2K}) \frac{T}{2N} \right| > \sum_{n=1}^{\infty} \varepsilon_n \right\}$$

$$\leq \sum_{n=1}^{\infty} P_{\theta_0} \left\{ \left(\frac{C_3}{n^{1+\gamma}} \right) \left| \sum_{k=1}^{2N} (W_{2K+1} - W_{2K}) \frac{T}{2N} \right| > \sum_{n=1}^{\infty} \varepsilon_n \right\}$$

$$\text{(by using 2.5)} \leq \sum_{n=1}^{\infty} \frac{C_3^2}{(n^{1+\gamma})^2 \varepsilon_n^2} \left(\sum_{k=1}^{2N} E(W_{2K+1} - W_{2K})^2 \right) \left(\frac{T}{2N} \right)^2$$

$$\leq C_{19} \sum_{n=1}^{\infty} \frac{1}{n^{2+2\gamma} \varepsilon_n^2} \frac{T^3}{(2N)^2}. \tag{6.18}$$

(2.5) Therefore, combining (6.17) and (6.18) we get,

$$P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \left| \sum_{K=1}^{2N} Y_K^{(1)} \right| > \varepsilon_n \right\} \leq C_{20} \sum_{n=1}^{\infty} \frac{1}{(n^{2+2\gamma}) \varepsilon_n^2} \frac{T^3}{(2N)^2}. \tag{6.19}$$

Proceeding as above, it is easy to obtain the estimate

$$P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \left| \sum_{K=1}^{2N} Y_K^{(2)} \right| > \varepsilon_n \right\} \leq C_{21} \sum_{n=1}^{\infty} \frac{1}{n^{2+2\gamma} \varepsilon_n^2} \frac{T^3}{(2N)^2}. \tag{6.20}$$

Therefore combining these two estimates we get,

$$p_{\theta_0} \left\{ \sup_{\theta \in \Theta} \left| M_{\pi_n}^{\theta} - M_{\pi_{N^{(1)}}} \right| \geq \sum_{n=1}^{\infty} 2\varepsilon_n \right\} \leq \frac{C_{22}}{(2N)^2} \sum_{n=1}^{\infty} \frac{1}{(n^{1+\gamma})^2 \varepsilon_n^2}.$$

Now proceeding as above and defining ε_n and ρ in similar manner as above, it is easy to show that

$$P_{\theta_0} \left\{ \sup_{\theta \in \Theta} \left| M_{\pi_N}^{\theta} - M_{\pi_{N^{(1)}}}^{\theta} \right| > \rho \right\} \leq C_{23} N^{-2} \rho^{-2}$$

and $M_{\pi_N}^{\theta} \to M^{\theta}$ uniformly in θ in P_{θ_0}-measure. Considering the Fourier expansion of $f_x(\theta, x)$ and $f_{xx}(\theta, x)$ and replacing g by f_x, we can prove assertion (ii).

REFERENCES

1. Prakasa Rao, B.L.S. *Statistical Inference for Diffusion Type Processes*; Arnold: London, 1999.
2. Dorogovcev, A.Ja. The Consistency of an Estimate of a Parameter of a Stochastic Differential Equation. Theory Probab. Math. Statist. **1976**, *10*, 73–82.

3. Kasonga, R.A. The Consistency of a Nonlinear Least Squares Estimate from Diffusion Processes. Stoch. Proc. Appl. **1988**, *30*, 263–275.
4. Prakasa Rao, B.L.S. Asymptotic Theory for Nonlinear Least Squares Estimator for Diffusion Process. Math. Oper. Forch. Statist. Ser. Statist. **1983**, *12* (2), 195–210.
5. Dacunha-Castelle, D.; Florens-Zmirou, D. Estimation of the Coefficients of a Diffusion from Discrete Observations. Stochastics **1986**, *19*, 263–284.
6. Genon Catalot, V. Maximum Contrast Estimation for Diffusion Process from Discrete Observations. Statistics **1990**, *21*, 99–116.
7. Dohnal, G. On Estimating the Diffusion Co-efficient. J. Appl. Probab. **1987**, *24*, 105–114.
8. Prakasa Rao, B.L.S.; Rubin, H. Asymptotic Theory of Estimation in Nonlinear Stochastic Differential Equations. Sankhya Ser. A **1981**, *43*, 170–189.
9. Yoshida, N. Estimation for Diffusion Process from Discrete Observations. J. Multivariate Anal. **1992**, *41*, 220–242.
10. Le Breton, A. On Continuous and Discrete Sampling for Parameter Estimation in Diffusion Type Process. Math. Program. Study **1976**, *5*, 124–144.
11. Mishra, M.N.; Bishwal, J.P.N. Approximate Maximum Likelihood Estimation. Stochastics and Stochastics Rep. **1995**, *52*, 1–13.
12. Gihman, I.I.; Skorohod, A.V. *Stochastic Differential Equations*; Springer: Berlin, 1972.
13. Lipster, R.S.; Shiryayev, A.N. *Statistics of Random Process I, General Theory*; Springer: Berlin, 1978.
14. Basawa, I.V.; Prakasa Rao, B.L.S. *Statistical Inference for Stochastic Processes*; Academic Press: London, 1980.
15. Stratonovich, R.L. A New Representation for Stochastic Integrals and Equations. SIAM J. Control **1966**, *4*, 362–371.
16. Aitchinson, J.; Silvey, S.D. Maximum Likelihood Estimation of Parameters Subject to Constraint. Ann. Math. Statist. **1995**, *29*, 813–828.
17. Prakasa Rao, B.L.S.; Rubin, H. Uniform Approximation for Families of Stochastic Integrals. Sankhya Ser. A **1998**, *60*, 57–73.