Theory of Stochastic Processes
Vol. 15 (31), no. 2, 2009, pp. 62–83

JAYA P. N. BISHWAL

M-ESTIMATION FOR DISCRETELY SAMPLED DIFFUSIONS

We study the estimation of a parameter in the nonlinear drift coefficient of a stationary ergodic diffusion process satisfying a homogeneous Itô stochastic differential equation based on discrete observations of the process, when the true model does not necessarily belong to the observer's model. Local asymptotic normality of M-ratio random fields are studied. Asymptotic normality of approximate M-estimators based on the Itô and Fisk–Stratonovich approximations of a continuous M-functional are obtained under a moderately increasing experimental design condition through the weak convergence of approximate M-ratio random fields. The derivatives of an approximate log-M functional based on the Itô approximation are martingales, but the derivatives of a log-M functional based on the Fisk–Stratonovich approximation are not martingales, but the average of forward and backward martingales. The averaged forward and backward martingale approximations have a faster rate of convergence than the forward martingale approximations.

1. Introduction

Drift estimation in diffusion processes based on continuous observations inside a time interval is now classical, see e.g., Liptser and Shiryaev (1978), Basawa and Prakasa Rao (1980), Arato (1982), Prakasa Rao (1999), Kutoyants (1984, 2003), and Bishwal (2008) for the long-time asymptotics and Ibragimov and Has'minskii (1980) and Kutoyants (1984, 1994) for the small-noise asymptotics of different parametric and nonparametric estimators. On the other hand, the drift estimation for discretely observed diffusions is the recent trend of investigations due to the difficulty in observing the diffusion process continuously throughout a time interval. Several approaches are used for the parametric estimation based on discrete observations *viz.*, conditional least squares (cf. Dorogovcev (1976), Kasonga (1988)), approximating the continuous Girsanov likelihood (cf. Le Breton (1976), Florens-Zmirou (1989), Genon-Catalot (1990), Yoshida (1992), Mishra and Bishwal (1995), Harison (1996), Kloeden *et al.* (1996), Kessler (1997), and Bishwal (2006, 2007)), approximating the transition densities (cf. Dacunha-Castelle and Florens-Zmirou, Pedersen (1995a,b), Aït-Sahalia (2002)), martingale estimation function (cf. Bibby and Sørensen (1995a,b), Kessler and Sørensen (1995)) and generalized method of moments (cf. Clement (1993, 1995, 1997), Duffie and Glynn (1997)), and indirect inference method (cf. Gourieroux and Monfort (1995), Gourieroux *et al.* (1996), Broze *et al.* (1998)). Several approaches are used for the nonparametric estimation based on discrete observations, cf. Pham (1981), Nguyen and Pham (1981), Coutin (1994), Arfi and Lecoutre (1994), and Arfi (1995). Statistical inference for ergodic diffusions for continuous time data is studied in detail by Kutoyants (2003). See the new monograph by Bishwal (2008) for recent results on approximate likelihood asymptotics and approximate Bayes asymptotics for the drift estimation of discretely observed diffusions based on high-frequency data.

2000 *Mathematics Subject Classification.* Primary 62F12, 62F15, 62M05, 62F35; Secondary 60F05, 60F10, 60H05, 60H10.

Key words and phrases. Itô stochastic differential equations, diffusion processes, model misspecification, discrete observations, moderately increasing experimental design, approximate M-estimators, local asymptotic normality, robustness, weak convergence of random fields.

However, a gap always exists between the *ideal* mathematical models and the real data. One can hardly get clean data generated from the model due to the contamination by some noises and misspecification of the true model. We assume that the statistician/econometrician does not know the true model and uses a parametric model with an unknown parameter. The true model does not necessarily belong to the observer's parametric model. In such a situation, the robust estimation of a parameter was studied by McKeague (1984). The robust estimation in diffusion processes based on discrete observations is the focus of this paper. We study the asymptotic behavior of the approximate M-estimators which maximize two different approximate M-functionals based on discrete observations. The motivation of using two different approximate M-functionals comes from the fact that, for the correctly specified model, the estimators based on one approximate log-likelihood (Fisk–Stratonovich) are known to have faster rates of convergence than those based on another approximate log-likelihood (Itô) in the Ornstein–Uhlenbeck process, as shown by Bishwal and Bose (2001), and the approximants of the Fisk–Stratonovich integral converge to the corresponding integral faster than the approximants of the Itô integral converge to the corresponding integral for nonlinear integrators, as shown in Section 7. Our method of proof is through the weak convergence of approximate M-ratio random fields.

The organization of the paper is as follows : In Section 2, we prepare notations, assumptions, and preliminaries. In Section 3, we study the weak convergence of approximate M-ratio random fields. In section 4, we study the asymptotic normality of approximate M-estimators.

2. Model, Assumptions, and Preliminaries

Let the true process follow the homogeneous nonlinear Itô stochastic differential equation

$$\begin{aligned} dX_t &= g(X_t)dt + dW_t,\ t \geq 0 \\ X_0 &= \eta, \end{aligned} \tag{2.1}$$

where $\{W_t, t \geq 0\}$ is a one-dimensional standard Wiener process, g is a known real-valued function defined on $\mathbb{R}$. We assume that the process $\{X_t, t \geq 0\}$ is observed at $0 = t_0 < t_1 < \ldots < t_n = T$ with $t_i - t_{i-1} = \frac{T}{n} = h$.

We assume the parametric form of the Itô stochastic differential equation

$$\begin{aligned} dX_t &= \mu(\theta, X_t)dt + dW_t,\ t \geq 0 \\ X_0 &= \eta \end{aligned} \tag{2.2}$$

estimate $\theta \in \Theta$ from the observations $\{X_{t_0}, X_{t_1}, \ldots, X_{t_n}\} \equiv X_0^{n,h}$. Let θ^* be the quasi-true parameter defined as

$$\theta^* := \arg\inf_{\theta} \frac{1}{2} \int_{\mathbb{R}} (g(x) - \mu(\theta, x))^2 \nu(dx),$$

where ν is the invariant measure of the ergodic diffusion process. Suppose that θ^* lies in the interior of Θ.

We use the following notations throughout the paper: $\Delta X_i = X_{t_i} - X_{t_{i-1}}$, $\Delta W_i = W_{t_i} - W_{t_{i-1}}$, C is a generic constant independent of h, n, and other variables (perhaps, it may depend on θ). *A prime denotes the derivative with respect to θ and a dot denotes the derivative with respect to x.*

Looking back at the Girsanov likelihood, if the continuous observations of $\{X_t\}$ on the interval $[0, T]$ were available, then the M-functional of θ would be

$$M_T(\theta) = \exp\{\int_0^T a(\theta, X_t)dX_t - \frac{1}{2}\int_0^T b(\theta, X_t)dt\}, \tag{2.3}$$

and the M-estimate would be

$$\theta_T = \arg\max_{\theta \in \Theta} M_T(\theta),$$

where a and b are two given functions on $\mathbb{R} \times \Theta$ with Θ a closed subset of $\mathbb{R}$. The functions a and b arise from the model misspecification. In practice, we have discrete data, and we have to approximate the likelihood to get an M-estimate. First, we use an Itô type approximation of the stochastic integral and the rectangular approximation of the ordinary integral in (2.3) and obtain an approximate M-functional

$$M_{n,T}(\theta) = \exp\left\{\sum_{i=1}^{n} a(\theta, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) - \frac{h}{2}\sum_{i=1}^{n} b(\theta, X_{t_{i-1}})\right\}. \tag{2.4}$$

The approximate M-estimate (AME1) based on $M_{n,T}$ is defined as

$$\theta_{n,T} := \arg\max_{\theta\in\Theta} M_{n,T}(\theta).$$

Next, we transform the Itô integral in (2.3) to the Fisk–Stratonovich (FS) integral and obtain

$$M_T(\theta) = \exp\{\oint_0^T a(\theta, X_t)dX_t - \frac{1}{2}\int_0^T \dot{a}(\theta, X_t)dt - \frac{1}{2}\int_0^T b(\theta, X_t)dt\}, \tag{2.5}$$

where $\oint_0^T a(\theta, X_t)dX_t$ is the Fisk–Stratonovich (FS) integral.

We apply the Fisk–Stratonovich approximation of a stochastic integral and the rectangular approximation of ordinary integrals and obtain another approximate M-functional

$$\widetilde{M}_{n,T}(\theta) = \exp\left\{\frac{1}{2}\sum_{i=1}^{n}\left[a(\theta, X_{t_{i-1}}) + a(\theta, X_{t_i})\right](X_{t_i} - X_{t_{i-1}}) - \frac{h}{2}\sum_{i=1}^{n}\dot{a}(\theta, X_{t_{i-1}}) - \frac{h}{2}\sum_{i=1}^{n} b(\theta, X_{t_{i-1}})\right\}. \tag{2.6}$$

The approximate M-estimate (AME2) based on $\widetilde{M}_{n,T}$ is defined as

$$\tilde{\theta}_{n,T} := \arg\max_{\theta\in\Theta} \widetilde{M}_{n,T}(\theta).$$

We assume that the following conditions are satisfied:

(A1) There exists the constants K and K_1 such that

$|g(x) - g(y)| \le K(|x - y|)$,
$|a(\theta, x)| \le K_1(\theta)(1 + |x|)$,
$|a(\theta, x) - a(\theta, y)| \le K_1(\theta)|x - y|$.

(A2) The diffusion process X is stationary and ergodic with invariant measure ν, i.e., for any ϕ with $E[\phi(\cdot)] < \infty$,

$$\frac{1}{n}\sum_{i=1}^{n}\phi(X_{t_i}) \xrightarrow{P} E_\nu[\phi(X_0)] \text{ as } T \to \infty \text{ and } h \to 0.$$

(A3) For each $p > 0, \sup_t E|X_t|^p < \infty$.

(A4) $a(\theta, x)$ and $b(\theta, x)$ are twice continuously differentiable in $\theta \in \Theta$ and, for some $\gamma > 0$,

$$|a'(\theta, x)| + |a''(\theta, x)| \le C(\theta)(1 + |x|^\gamma),$$

$$|a'(\theta, x) - a'(\theta, y)| \le C(\theta)|x - y|.$$

(A5) The functions $a, b,$ and a' are smooth in x, and their derivatives are of polynomial growth order in x uniformly in θ.

(A6) We assume

$$\Gamma \equiv \Gamma(\theta^*) := \int_{\mathbb{R}} \rho''(\theta^*, x)d\nu(x) > 0,$$

where

$$\rho(\theta, x) := \frac{1}{2}b(\theta, x) - a(\theta, x)g(x).$$

(A7) $m(\theta)$ has its unique maximum at $\theta = \theta^*$ in Θ,
where

$$m(\theta) := \int_{\mathbb{R}} [\rho(\theta, x) - \rho(\theta^*, x)]\nu(dx).$$

(A8) There exists a twice continuously differentiable function $\kappa(x)$ satisfying the partial differential equation

$$\frac{1}{2}\ddot{\kappa}(x) + g(x)\dot{\kappa}(x) = \rho'(\theta, x)$$

and

$$\frac{1}{\sqrt{T}}\kappa(X_T) \xrightarrow{P} 0 \text{ as } T \to \infty.$$

(A9) We assume

$$\beta(\theta^*) := \int_{\mathbb{R}} [a'(\theta^*, x) + \dot{\kappa}(x)]^2 g^2(x) d\nu(x) > 0.$$

(A10) a is twice continuously differentiable function in x with bounded derivatives up to the second order.

Remark: Using (2.1), we have

$$\begin{aligned} M_T(\theta) &= \exp\{\int_0^T a(\theta, X_t)dW_t - \int_0^T [\frac{1}{2}b(\theta, X_t) - a(\theta, X_t)g(X_t)]dt\} \\ &= \exp\{\int_0^T a(\theta, X_t)dW_t - \int_0^T \rho(\theta, X_t)dt.\} \end{aligned}$$

3. Weak Convergence of Approximate M-Ratio Random Fields

Let $\theta = \theta^* + T^{-1/2}u, \ u \in \mathbb{R}$. Consider the approximate M-ratio (AMR) random fields

$$Z_{n,T}(u) := \frac{M_{n,T}(\theta)}{M_{n,T}(\theta^*)}, \quad \widetilde{Z}_{n,T}(u) := \frac{\widetilde{M}_{n,T}(\theta)}{\widetilde{M}_{n,T}(\theta^*)}. \tag{3.1}$$

Let

$$A_{\alpha,T} := \{u \in \mathbb{R} : |u| \le \alpha, \theta^* + T^{-1/2}u \in \Theta\}, \ \alpha > 0.$$

$$m_{n,T}(\theta) := \frac{1}{T}\log M_{n,T}(\theta), \qquad \widetilde{m}_{n,T}(\theta) := \frac{1}{T}\log \widetilde{M}_{n,T}(\theta).$$

$$Da(\theta^*, X_{t_{i-1}}, u) := a(\theta, X_{t_{i-1}}) - a(\theta^*, X_{t_{i-1}}) - T^{-1/2}ua'(\theta^*, X_{t_{i-1}}).$$

Below, we prove the weak convergence of the random fields $Z_{n,T}(\cdot)$ and $\widetilde{Z}_{n,T}(\cdot)$. We start with some lemmas.

Lemma 3.1 Under assumptions (A1) - (A9), we have

$$\sup_{\theta\in\Theta} |m_{n,T}(\theta) - m(\theta)| \xrightarrow{P} 0 \text{ as } T \to \infty, n \to \infty \text{ and } \frac{T}{n} \to 0.$$

Proof. Note that

$$\begin{aligned} & m_{n,T}(\theta) \\ = \ & T^{-1}\sum_{i=1}^n a(\theta, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) - \frac{1}{2}T^{-1}h\sum_{i=1}^n b(\theta, X_{t_{i-1}}) \\ = \ & T^{-1}\sum_{i=1}^n a(\theta, X_{t_{i-1}})\Delta W_i + \frac{1}{2}T^{-1}\sum_{i=1}^n \int_{t_{i-1}}^{t_i} a(\theta, X_{t_{i-1}})g(X_t)dt \\ & -\frac{1}{2}T^{-1}h\sum_{i=1}^n b(\theta, X_{t_{i-1}}) \\ =: \ & H_1 + H_2 + H_3. \end{aligned}$$

Note that $E|H_1|^2 \le C\frac{T}{n}$ and by the Burkholder–Davis–Gundy inequality $E|H_2|^2 \le CT^{-1}, E|F_2|^2 \le C, E|H_3|^2 \le C$.
Thus,

$$E|m_{n,T}(\theta)|^2 \le C,$$

$$E|m_{n,T}(\theta_2) - m_{n,T}(\theta_1)|^2 \le C|\theta_2 - \theta_1|$$

for $\theta, \theta_1, \theta_2 \in \Theta$ by (A5). Therefore, the family of distributions of $M_{n,T}(\cdot)$ on the Banach space $C(\Theta)$ with sup-norm is tight. Since $m(\cdot)$ is a point of $C(\Theta)$ and since, by ergodic property, $m_{n,T}(\theta) \xrightarrow{P} m(\theta)$ as $T \to \infty$ and $\frac{T}{n} \to 0$, this completes the proof of the lemma using Lemma 3.1 in Yoshida (1990). □

The next lemma is a generalization of local asymptotic normality (LAN) for ergodic diffusions using the random field $Z_{n,T}(u)$.

Lemma 3.2 Under assumptions (A1) - (A9), for all $u \in \mathbb{R}$,

$$\log Z_{n,T}(u) = u\Delta_{n,T}(\theta^*) - \frac{1}{2}u^2\Gamma(\theta^*) + r_{n,T}(u),$$

where

$$\Delta_{n,T}(\theta^*) \xrightarrow{D} \Delta(\theta^*), \quad \Delta \sim N(0, \beta(\theta^*))$$

and $r_{n,T}(u) \xrightarrow{P} 0$ as $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$.

The next two lemmas give the tightness of the distributions of the AMR random field $Z_{n,T}(u)$.

Lemma 3.3 Under assumptions (A1) - (A9), for each $\epsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{T\to\infty, \frac{T}{n^{2/3}} \to 0} P\left\{ \sup_{u_1, u_2 \in A_{\alpha,T}, |u_2 - u_1| \le \delta} |\log Z_{n,T}(u_2) - \log Z_{n,T}(u_1)| > \epsilon \right\} = 0.$$

Lemma 3.4 Under assumptions (A1) - (A9), for each $\epsilon > 0$,

$$\lim_{\alpha \to \infty} \limsup_{T\to\infty, \frac{T}{n^{2/3}} \to 0} P\left\{ \sup_{|u| \ge \alpha} Z_{n,T}(u) > \epsilon \right\} = 0.$$

Lemma 3.5. Under assumptions (A1) - (A9),

$$\sup_{\theta \in \Theta} |\widetilde{m}_{n,T}(\theta) - m(\theta)| \xrightarrow{P} 0 \quad \text{as } T \to \infty \quad \text{and } \frac{T}{n} \to 0.$$

Proof : Note that

$$\begin{aligned}
\widetilde{m}_{n,T}(\theta) &= T^{-1}\sum_{i=1}^{n} \frac{a(\theta, X_{t_{i-1}}) + a(\theta, x_{t_i})}{2}(X_{t_i} - X_{t_{i-1}}) - \frac{1}{2}T^{-1}h\sum_{i=1}^{n} \dot{a}(\theta, X_{t_{i-1}}) \\
&\quad - \frac{1}{2}T^{-1}h\sum_{i=1}^{n} b(\theta, X_{t_{i-1}}). \\
&= \left\{ T^{-1}\sum_{i=1}^{n} \frac{a(\theta, X_{t_{i-1}}) + a(\theta, X_{t_i})}{2}\Delta W_i - \frac{1}{2}T^{-1}h\sum_{i=1}^{n} \dot{a}(\theta, X_{t_{i-1}}) \right\} \\
&\quad + \frac{1}{2}T^{-1}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \left[a(\theta, X_{t_{i-1}}) + a(\theta, X_{t_i}) \right] g(X_t)dt \\
&\quad - \frac{1}{2}T^{-1}h\sum_{i=1}^{n} b(\theta, X_{t_{i-1}}) \\
&=: F_1 + F_2 + F_3.
\end{aligned}$$

Note that

$$\begin{aligned} F_1 &= \Bigg[\Bigg\{T^{-1}\sum_{i=1}^{n}\frac{a(\theta,X_{t_{i-1}})+a(\theta,X_{t_i})}{2}\Delta W_i - \frac{1}{2}T^{-1}h\sum_{i=1}^{n}\dot{a}(\theta,X_{t_{i-1}})\Bigg\} \\ &\quad -T^{-1}\sum_{i=1}^{n}a(\theta,X_{t_{i-1}})\Delta W_i\Bigg] + T^{-1}\sum_{i=1}^{n}a(\theta,X_{t_{i-1}})\Delta W_i \\ &=: G_1+G_2. \end{aligned}$$

From Theorem 7.1, it follows that $E|G_1|^2 \le C\frac{T}{n}$ and, by the Burkholder–Davis–Gundy inequality, $E|G_2|^2 \le CT^{-1}, E|F_2|^2 \le C, E|F_3|^2 \le C$. Thus,

$$E|\widetilde{m}_{n,T}(\theta)|^2 \le C$$

and

$$E|\widetilde{m}_{n,T}(\theta_2)-\widetilde{m}_{n,T}(\theta_1)|^2 \le C|\theta_2-\theta_1|$$

for $\theta,\theta_1,\theta_2 \in \Theta$ by (A5). Now using arguments similar to Lemma 3.1 completes the proof of the lemma. □

The proofs of Lemmas 3.2 - 3.4 are similar and much simpler respectively to the proofs of three next lemmas. Hence, we omit the details.

The next lemma is a generalization of LAN for ergodic diffusions using the random field $\widetilde{Z}_{n,T}(u)$.

Lemma 3.6 Under assumptions (A1) - (A10), for all $u \in \mathbb{R}$,

$$\log\ \widetilde{Z}_{n,T}(u) = u\widetilde{\Delta}_{n,T}(\theta^*) - \frac{1}{2}u^2\Gamma(\theta^*) + \widetilde{\gamma}_{n,T}(u),$$

where $\widetilde{\Delta}_{n,T}(\theta^*)\xrightarrow{D}\Delta(\theta^*)$, $\Delta \sim N(0,\beta(\theta^*))$ and $\widetilde{\gamma}_{n,T}(u)\xrightarrow{P}0$ as $T\to\infty$ and $\frac{T}{n^{2/3}}\to 0$.

Proof: For $\theta = \theta^* + T^{-1/2}u$, we have

$$\begin{aligned} &\log\widetilde{Z}_{n,T}(u) = \log\frac{\widetilde{M}_{n,T}(\theta)}{\widetilde{M}_{n,T}(\theta^*)} \\ &= \sum_{i=1}^{n}\left[\frac{a(\theta,X_{t_i})+a(\theta,X_{t_{i-1}})}{2} - \frac{a(\theta^*,X_{t_i})+a(\theta^*,X_{t_{i-1}})}{2}\right]\Delta X_i \\ &\quad -\frac{h}{2}\sum_{i=1}^{n}\left[\dot{a}(\theta,X_{t_{i-1}})-\dot{a}(\theta^*,X_{t_{i-1}})\right] - \frac{h}{2}\sum_{i=1}^{n}\left[b(\theta,X_{t_{i-1}})-b(\theta^*,X_{t_{i-1}})\right] \\ &= \frac{1}{2}\sum_{i=1}^{n}\left[\{a(\theta,X_{t_i})-a(\theta^*,X_{t_i})\}+\left\{a(\theta,X_{t_{i-1}})-a(\theta^*,X_{t_{i-1}})\right\}\right]\Delta W_i \\ &\quad +\frac{1}{2}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\left[\{a(\theta,X_{t_i})-a(\theta^*,X_{t_i})\}+\left\{a(\theta,X_{t_{i-1}})-a(\theta^*,X_{t_{i-1}})\right\}\right]g(X_t)dt \\ &\quad -\frac{h}{2}\sum_{i=1}^{n}\left[\dot{a}(\theta,X_{t_{i-1}})-\dot{a}(\theta^*,X_{t_{i-1}})\right] - \frac{h}{2}\sum_{i=1}^{n}\left[b(\theta,X_{t_{i-1}})-b(\theta^*,X_{t_{i-1}})\right] \end{aligned}$$

(by using (2.1))

$$\begin{aligned} &= T^{-1/2}u\sum_{i=1}^{n}\frac{a'(\theta^*,X_{t_{i-1}})+a'(\theta^*,X_{t_i})}{2}\Delta W_i \\ &\quad + \frac{1}{2}\sum_{i=1}^{n}\left[Da(\theta^*,X_{t_{i-1}},u)+Da(\theta^*,X_{t_i},u)\right]\Delta W_i \\ &\quad - \frac{h}{2}T^{-1/2}u\sum_{i=1}^{n}a'(\theta^*,X_{t_{i-1}}) - \frac{h}{2}\sum_{i=1}^{n}Da(\theta^*,X_{t_{i-1}},u) \end{aligned}$$

$$+\frac{1}{2}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\left[\{a(\theta,X_{t_i})-a(\theta^*,X_{t_i})\}+\left\{a(\theta,X_{t_{i-1}})-a(\theta^*,X_{t_{i-1}})\right\}\right]g(X_t)dt$$
$$-\frac{h}{2}\sum_{i=1}^{n}\left[a(\theta,X_{t_{i-1}})-a(\theta^*,X_{t_{i-1}})\right]^2$$
$$-\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\left[a(\theta,X_{t_{i-1}})-a(\theta^*,X_{t_{i-1}})\right]a(\theta^*,X_{t_{i-1}})dt$$
$$=u\widetilde{\Delta}_{n,T}-\frac{1}{2}\Gamma_{n,T}+\widetilde{\rho}_{n,T}(u)$$

where

$$\widetilde{\Delta}_{n,T}=:T^{-1/2}\left[\sum_{i=1}^{n}\frac{a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(X_{t_{i-1}})+a'(\theta^*,X_{t_i})+\dot{\kappa}(X_{t_i})}{2}\Delta W_i\right.$$
$$\left.-\tfrac{h}{2}\sum_{i=1}^{n}\dot{a}'(\theta^*,X_{t_{i-1}})+\ddot{\kappa}(X_{t_{i-1}})\right],$$
$$\Gamma_{n,T}=:h\sum_{i=1}^{n}\left[\rho(\theta,X_{t_{i-1}})-\rho(\theta^*,X_{t_{i-1}})\right]^2$$

and

$$\widetilde{\rho}_{n,T}(u)$$
$$=:\quad\left\{\sum_{i=1}^{n}\frac{Da(\theta^*,X_{t_{i-1}},u)+Da(\theta^*,X_{t_i},u)}{2}\Delta W_i-\frac{h}{2}\sum_{i=1}^{n}Da(\theta^*,X_{t_{i-1}},u)\right\}$$
$$+T^{-1/2}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}a'(\theta^*,X_{t_{i-1}})\left[a(\theta^*,X_t)-a(\theta^*,X_{t_{i-1}})\right]dt$$
$$+\sum_{i=1}^{n}Da(\theta^*,X_{t_{i-1}},u)\left[a(\theta^*,X_t)-a(\theta^*,X_{t_{i-1}})\right]dt$$
$$+\frac{1}{2}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\left\{\left[a(\theta,X_{t_i})-a(\theta^*,X_{t_i})\right]-\left[a(\theta,X_{t_{i-1}})-a(\theta^*,X_{t_{i-1}})\right]\right\}g(X_t)dt$$
$$=:\quad S_1(u)+S_2(u)+S_3(u)+S_4(u).$$

Thus,

$$\begin{aligned}\log\widetilde{Z}_{n,T}(u)&=u\widetilde{\Delta}_{n,T}-\frac{1}{2}\Gamma_{n,T}+\widetilde{\rho}_{n,T}(u)\\&=u\widetilde{\Delta}_{n,T}-\frac{1}{2}u^2\Gamma-\frac{1}{2}(\Gamma_{n,T}-u^2\Gamma)+\widetilde{\rho}_{n,T}(u)\ ,\\&=u\widetilde{\Delta}_{n,T}-\frac{1}{2}u^2\Gamma+\widetilde{\gamma}_{n,T}(u)\end{aligned}$$

where

$$\widetilde{\gamma}_{n,T}(u)=\widetilde{\rho}_{n,T}(u)-\frac{1}{2}(\Gamma_{n,T}-u^2\Gamma).$$

Due to the mean-value theorem and ergodicity, we have

$$\Gamma_{n,T}-u^2\Gamma\xrightarrow{P}0\quad\text{as } T\to\infty\quad\text{and}\quad\frac{T}{n}\to0.$$

Notice that

$$\widetilde{\Delta}_{n,T}=\left\{T^{-1/2}\left[\sum_{i=1}^{n}\frac{a'(\theta^*,X_{t_{i-1}}+\dot{\kappa}(\theta^*,X_{t_{i-1}})\ +a'(\theta^*,X_{t_i})+\dot{\kappa}(\theta^*,X_{t_i})}{2}\Delta W_i\right.\right.$$
$$\left.-\frac{h}{2}\sum_{i=1}^{n}[\dot{a}'(\theta^*,X_{t_{i-1}})+\ddot{\kappa}(\theta^*,X_{t_{i-1}})]\right]$$

$$-T^{-1/2}\sum_{i=1}^{n}[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})]\Delta W_i\Bigg\}$$

$$+T^{-1/2}\sum_{i=1}^{n}[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})]\Delta W_i$$

$$=: H_3+\Delta_{n,T}.$$

From Theorem 7.1, it follows that

$$H_3\xrightarrow{P}0 \quad \text{as } \frac{T}{n^{2/3}}\to 0.$$

Notice that for

$$\zeta_i(t)=a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})-a'(\theta^*,X_t)-\dot{\kappa}(\theta^*,X_t)$$

with $t_{i-1}\le t<t_i$, $i=1,2,\ldots,n$,

$$\begin{aligned}
&E|T^{-1/2}\sum_{i=1}^{n}[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})]\Delta W_i \\
&-T^{-1/2}\textstyle\int_0^T[a'(\theta^*,X_t)+\dot{\kappa}(\theta^*,X_t)]dW_t| \\
=\;& T^{-1/2}E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})]dW_t \\
&-\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}[a'(\theta^*,X_t)+\dot{\kappa}(\theta^*,X_t)]dW_t| \\
=\;& T^{-1/2}E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\Big[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})-a'(\theta^*,X_t)-\dot{\kappa}(\theta^*,X_t)\Big]\,dW_t| \\
=\;& T^{-1/2}E|\int_0^T\zeta_i(t)dW_t|\le T^{-1/2}\left\{E|\int_0^T\zeta_i(t)dW_t|^2\right\}^{1/2} \\
=\;& T^{-1/2}\left\{\int_0^T E|\zeta_i(t)|^2dt\right\}^{1/2}\le T^{-1/2}\left\{\int_0^T CE|X_{t_{i-1}}-X_t|^2dt\right\}^{1/2}\quad \text{(by (A4))} \\
\le\;& CT^{-1/2}\left\{\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(t-t_{i-1})dt\right\}^{1/2} \\
\le\;& CT^{-1/2}\left\{n\left(\frac{T}{n}\right)^2\right\}^{1/2}\le C\left(\frac{T}{n}\right)^{1/2}.
\end{aligned}$$

Thus,

$$T^{-1/2}\sum_{i=1}^{n}[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})]\Delta W_i-T^{-1/2}\int_0^T[a'(\theta^*,X_t)+\dot{\kappa}(\theta^*,X_t)]dW_t\xrightarrow{P}0$$

as $T/n\to 0$.

On the other hand, using condition (A2) by the Central Limit Theorem for stochastic integrals (see Basawa and Prakasa Rao (1980, Theorem 2.1, Appendix 2)), we have

$$T^{-1/2}\int_0^T[a'(\theta^*,X_t)+\dot{\kappa}(\theta^*,X_t)]dW_t\xrightarrow{D}N(0,\beta(\theta^*)) \quad \text{as } T\to\infty.$$

Hence,

$$\Delta_{n,T}=T^{-1/2}\sum_{i=1}^{n}[a'(\theta^*,X_{t_{i-1}})+\dot{\kappa}(\theta^*,X_{t_{i-1}})]\Delta W_i\xrightarrow{D}N(0,\beta(\theta^*)) \quad \text{as } T\to\infty \text{ and } \frac{T}{n}\to 0.$$

Thus, to complete the proof of the lemma, we have to show that

$$\widetilde{\rho}_{n,T}(u)\xrightarrow{P}0.$$

Let us first estimate $S_1(u)$.

$$S_1(u) = \left\{ \left[\sum_{i=1}^{n} \frac{Da(\theta^*, X_{t_{i-1}}, u) + Da(\theta^*, X_{t_i}, u)}{2} \Delta W_i - \frac{h}{2} \sum_{i=1}^{n} D\dot{a}(\theta^*, X_{t_{i-1}}, u) \right] - \sum_{i=1}^{n} Da(\theta^*, X_{t_{i-1}}, u)\Delta W_i \right\} + \sum_{i=1}^{n} Da(\theta^*, X_{t_{i-1}}, u)\Delta W_i$$
$$=: H_4 + r_1(u).$$

From Theorem 7.1, it follows that $H_4 \xrightarrow{P} 0$ as $\frac{T}{n^{2/3}} \to 0$.
Next,

$$\begin{aligned} E(r_1^2(u)) &= E\left[\sum_{i=1}^{n} Da(\theta^*, X_{t_{i-1}}, u)\Delta W_i\right]^2 \\ &= \sum_{i=1}^{n} E|Da(\theta^*, X_{t_{i-1}}, u)|^2 E|\Delta W_i|^2 \\ &= h\sum_{i=1}^{n} E|Da(\theta^*, X_{t_{i-1}}, u)|^2 \end{aligned} \tag{3.9}$$

But

$$\begin{aligned} Da(\theta^*, X_t, u) &= (\theta - \theta^*)a'(\theta^*, X_t) - T^{-1/2}ua'(\theta^*, X_{t_{i-1}}) \\ &\qquad \text{(where } |\theta^* - \theta^*| < T^{-1/2}u) \\ &= T^{-1/2}u\left[a'(\theta^*, X_t) - a'(\theta^*, X_t)\right]. \end{aligned}$$

Hence

$$\begin{aligned} E|Da(\theta^*, X_{t_{i-1}}, u)|^2 &= T^{-1}u^2E|a'(\theta^*, X_{t_{i-1}}) - a'(\theta^*, X_{t_{i-1}})|^2 \\ &\le T^{-1}u^2E|J(X_{t_{i-1}})(\theta^* - \theta^*)|^2 \\ &\le T^{-2}u^4E[J^2(X_0)] \\ &\le CT^{-2}u^4. \end{aligned} \tag{3.10}$$

Substituting (3.10) into (3.9), we obtain

$$E(r_1^2(u)) \le CT^{-2}u^4nh \le CT^{-1}.$$

Thus, $r_1(u) \xrightarrow{P} 0$ as $T \to \infty$. Hence,

$$S_1(u) \xrightarrow{P} 0, \quad \text{as } T \to \infty \ \text{ and } \frac{T}{n^{2/3}} \to 0.$$

Next, let us estimate $S_2(u)$. We have, by the Itô formula,

$$\begin{aligned} & a(\theta^*, X_t) - a(\theta^*, X_{t_{i-1}}) \\ =\ & \textstyle\int_{t_{i-1}}^{t} \dot{a}(\theta^*, X_u)dX_u + \frac{1}{2}\int_{t_{i-1}}^{t} \ddot{a}(\theta^*, X_u)du \\ =\ & \int_{t_{i-1}}^{t} \dot{a}(\theta^*, X_u)dW_u + \int_{t_{i-1}}^{t} [\dot{a}(\theta^*, X_u)\dot{a}(\theta^*, X_u) + \frac{1}{2}\ddot{a}(\theta^*, X_u)]du \\ =:\ & \int_{t_{i-1}}^{t} \dot{a}(\theta^*, X_u)dW_u + \int_{t_{i-1}}^{t} A(\theta^*, X_u)du. \end{aligned}$$

Thus

$$\begin{aligned} E|S_2(u)|^2 &= E|T^{-1/2}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} a'(\theta^*, X_{t_{i-1}})\left[a(\theta^*, X_t) - a(\theta^*, X_{t_{i-1}})\right]dt|^2 \\ &= E|T^{-1/2}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \Big[a'(\theta^*, X_{t_{i-1}})\dot{a}(\theta^*, X_u)dW_u \\ &\quad + \int_{t_{i-1}}^{t} a'(\theta^*, X_{t_{i-1}})A(\theta^*, X_u)du\Big]dt|^2 \end{aligned}$$

$$\leq 2T^{-1}u^2 \left\{ E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t} a'(\theta^*, X_{t_{i-1}})\dot{a}(\theta^*, X_u)dW_u dt|^2 \right.$$
$$\left. + E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t} a'(\theta^*, X_{t_{i-1}})A(\theta^*, X_u)du dt|^2 \right\}$$
$$=: 2T^{-1}u^2(N_1 + N_2).$$

Note that, for $B_{i,t} := \int_{t_{i-1}}^{t} a'(\theta^*, X_{t_{i-1}}\dot{a}(\theta^*, X_u)dW_u$,

$$\begin{aligned} N_1 &= \sum_{i=1}^{n} E\left(\int_{t_{i-1}}^{t_i} B_{i,t}dt\right)^2 + \sum_{j\neq i=1}^{n} E\left(\int_{t_{i-1}}^{t_i} B_{i,t}dt\right)\left(\int_{t_{j-1}}^{t_j} B_{j,t}dt\right) \\ &\leq \sum_{i=1}^{n}(t_i - t_{i-1})\int_{t_{i-1}}^{t_i} E(B_{i,t}^2)dt \end{aligned}$$

(the last term being zero due to the orthogonality of the integrals)

$$\begin{aligned} &\leq \sum_{i=1}^{n}(t_i - t_{i-1})\int_{t_{i-1}}^{t_i}\left\{\int_{t_{i-1}}^{t} E\left[a'(\theta^*, X_{t_{i-1}})\dot{a}(\theta^*, X_u)\right]^2 du\right\} dt \\ &\leq C\frac{T}{n}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(t - t_{i-1})dt \quad \text{(by (A4) and (A3))} \\ &\leq C\frac{T}{n}\sum_{i=1}^{n}(t_i - t_{i-1})^2 = C\frac{T^3}{n^2}. \end{aligned}$$

On the other hand, for $R_{i,t} = \int_{t_{i-1}}^{t} a'(\theta^*, X_{t_{i-1}})a(\theta^*, X_u)du$,

$$\begin{aligned} N_2 &= E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t} a'(\theta^*, X_{t_{i-1}})A(\theta^*, X_u)du dt|^2 \\ &= E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} R_{i,t}dt|^2 = \sum_{i=1}^{n}(\int_{t_{i-1}}^{t_i} R_{i,t}dt)^2 + \sum_{j\neq i=1}^{n} E\left(\int_{t_{i-1}}^{t_i} R_{i,t}dt\right)\left(\int_{t_{j-1}}^{t_j} R_{j,t}dt\right) \\ &\leq \sum_{i=1}^{n}(t_i - t_{i-1})E\left(\int_{t_{i-1}}^{t_i} R_{i,t}dt\right)^2 + \sum_{j\neq i=1}^{n}\left\{E(\int_{t_{i-1}}^{t_i} R_{i,t}dt)^2 E(\int_{t_{j-1}}^{t_j} R_{j,t}dt)^2\right\}^{1/2} \\ &\leq \sum_{i=1}^{n}(t_i - t_{i-1})\int_{t_{i-1}}^{t_i} E(R_{i,t}^2)dt \\ &\quad + \sum_{j\neq i=1}^{n}\left\{(t_i - t_{i-1})\int_{t_{i-1}}^{t_i} E(R_{i,t}^2)dt(t_j - t_{j-1})\int_{t_{j-1}}^{t_j} E(R_{j,t}^2)dt\right\}^{1/2} \end{aligned}$$

But $E(R_{i,t}^2) \leq C(t - t_{i-1})^2$ using (A4) and (A3). On substitution, the last term is bounded by

$$\begin{aligned} & C\sum_{i=1}^{n}(t_i - t_{i-1})^4 + C\sum_{j\neq i=1}^{n}(t_i - t_{i-1})^2(t_j - t_{j-1})^2 \\ =\; & Cn.\frac{T^4}{n^4} + C\frac{n(n-1)}{2}.\frac{T^4}{n^4} \leq C\frac{T^4}{n^2}. \end{aligned}$$

Thus, $E|S_2(u)|^2$

$$\begin{aligned} &\leq 2T^{-1}u(A_1 + A_2) \leq CT^{-1}u(\frac{T^3}{n^2}) + CT^{-1}u\frac{T^4}{n^2} \\ &\leq C\frac{T^3}{n^2} = C(\frac{T}{n^{2/3}})^3. \end{aligned}$$

Thus, $S_2(u) \xrightarrow{p} 0$ as $\frac{T}{n^{2/3}} \to 0$. Next, let us estimate $S_3(u)$.

$$\begin{aligned} E|S_3(u)| &= E|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} Da(\theta^*, X_{t_{i-1}}, u)\left[a(\theta^*, X_t) - a(\theta^*, X_{t_{i-1}})\right] dt| \\ &\leq \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} E|Da(\theta^*, X_{t_{i-1}}, u)||a(\theta^*, X_t) - a(\theta^*, X_{t_{i-1}})|dt \\ &\leq \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \left\{E|Da(\theta^*, X_{t_{i-1}}, u)|^2 E|a(\theta^*, X_t) - a(\theta^*, X_{t_{i-1}})|^2\right\}^{1/2} dt \\ &\leq \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \left\{CT^{-2}u^2.E|X_t - X_{t_{i-1}}|^2\right\}^{1/2} dt \text{ (by (3.10) and (A1))} \\ &\leq CT^{-1}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} (t - t_{i-1})^{1/2} dt \end{aligned}$$

(since $E(X_t - X_{t_{i-1}})^2 \leq C(t - t_{i-1})$ by Gikhman and Skorohod (1972, p. 48))

$$\begin{aligned} &\leq CT^{-1}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \sup_{t_{i-1}\leq t\leq t_i} (t_i - t_{i-1})^{1/2} dt \\ &\leq CT^{-1}un.\left(\frac{T}{n}\right)^{3/2} \leq C\left(\frac{T}{n}\right)^{1/2} \end{aligned}$$

Thus, $S_3(u) \xrightarrow{P} 0$ as $\frac{T}{n} \to 0$. Next, let us estimate $S_4(u)$.

$$\begin{aligned} 2S_4(u) &= \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \left\{[a(\theta, X_{t_i}) - a(\theta^*, X_{t_i})] - \left[a(\theta, X_{t_{i-1}}) - a(\theta^*, X_{t_{i-1}})\right]\right\} g(X_t)dt \\ &= T^{-1/2}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \left[a'(\theta^*, X_{t_i}) - a'(\theta^*, X_{t_{i-1}})\right] g'(X_t)dt \\ &\qquad \text{where } |\theta^* - \theta^*| < T^{-1/2}u. \end{aligned}$$

Now proceeding similarly as in the proof of convergence of $S_2(u)$ to zero in probability, it can be shown that

$$S_4(u) \xrightarrow{P} 0 \quad \text{as } \frac{T}{n^{2/3}} \to 0.$$

This completes the proof of the lemma □

Lemma 3.7 Under assumptions (A1) - (A10), for each $\epsilon > 0$,

$$\lim_{\delta\to 0}\limsup_{T\to\infty, \frac{T}{n^{2/3}}\to 0} P\left\{\sup_{u_1,u_2\in A_{\alpha,T}, |u_2-u_1|\leq\delta} |\log\widetilde{Z}_{n,T}(u_2) - \log\widetilde{Z}_{n,T}(u_1)| > \epsilon\right\} = 0.$$

Proof : From Lemma 3.6, we have

$$\begin{aligned} &\left|\log\widetilde{Z}_{n,T}(u_2) - \log\widetilde{Z}_{n,T}(u_1)\right| \\ = &\left|(u_2 - u_1)\widetilde{\Delta}_{n,T} - \frac{1}{2}(u_2^2 - u_1^2)\Gamma + \widetilde{\gamma}_{n,T}(u_2) - \widetilde{\gamma}_{n,T}(u_1)\right| \\ \leq & |u_2 - u_1||\widetilde{\Delta}_{n,T}| + C|u_2 - u_1| + |\widetilde{\gamma}_{n,T}(u_2)| + |\widetilde{\gamma}_{n,T}(u_1)|, \end{aligned}$$

where C is a positive constant.

Therefore,

$$P\left\{\sup_{u_1,u_2,\in A_{\alpha,T}, |u_2-u_1|\leq\delta} |\log\widetilde{Z}_{n,T}(u_2) - \log\widetilde{Z}_{n,T}(u_1)| > \epsilon\right\}$$

$$\leq P\left\{|\widetilde{\Delta}_{n,T}| + B > \frac{\epsilon}{3\delta}\right\} + 2P\left\{\sup_{u\in A_{\alpha,T}} |\widetilde{\gamma}_{n,T}(u_1)| > \frac{\epsilon}{3}.\right\}$$

Note that

$$\begin{aligned}
& P\left\{\sup_{u\in A_{\alpha,T}} |\widetilde{\gamma}_{n,T}(u)| > \frac{\epsilon}{3}\right\} \\
= \;& P\left\{\sup_{|u|\leq\alpha} |S_1(u)+S_2(u)+S_3(u)+S_4(u)| > \frac{\epsilon}{6}\right\} + P\left\{\sup_{u\in A_{\alpha,T}} |\Gamma_{n,T} - u^2\Gamma| > \frac{\epsilon}{3}\right. \\
\leq \;& P\left\{\sup_{|u|\leq\alpha} |S_1(u)| > \frac{\epsilon}{24}\right\} + P\left\{\sup_{|u|\leq\alpha} |S_2(u)| > \frac{\epsilon}{24}\right\} + P\left\{\sup_{|u|\leq\alpha} |S_3(u)| > \frac{\epsilon}{24}\right\} \\
& +P\left\{\sup_{|u|\leq\alpha} |S_4(u)| > \frac{\epsilon}{24}\right\} - P\left\{\sup_{u\in A_{\alpha,T}} |\Gamma_{n,T} - u^2\Gamma| > \frac{\epsilon}{3}\right\} \\
\to \;& 0 \text{ as } T\to\infty \text{ and } \frac{T}{n^{2/3}} \to 0.
\end{aligned}$$

Since $\widetilde{\Delta}_{n,T}$ converges in distribution to $N(0,\Gamma)$, hence

$$\lim_{\delta\to 0} \lim_{T\to\infty \frac{T}{n^{2/3}}\to 0} P\left\{\sup_{u_1,u_2\in A_{\alpha,T}, |u_2-u_1|\leq\delta} |\log \widetilde{Z}_{n,T}(u_2) - \log \widetilde{Z}_{n,T}(u_1)| > \epsilon\right\} = 0. \quad \square$$

Lemma 3.8. Under assumptions (A1) - (A10), we have, for each $\epsilon > 0$,

$$\lim_{\alpha\to\infty} \limsup_{T\to\infty, \frac{T}{n^{2/3}}\to 0} P\left\{\sup_{|u|\geq\alpha} \widetilde{Z}_{n,T}(u) > \epsilon\right\} = 0.$$

Proof : Since Γ is positive, there exists a number η such that

$$\eta u^2 \leq \frac{1}{4}\Gamma u^2, \quad u \in \mathbb{R}.$$

By Lemma 3.6,

$$\log \widetilde{Z}_{n,T}(u) = u\widetilde{\Delta}_{n,T} - \frac{1}{2}u^2\Gamma + \widetilde{\gamma}_{n,T}(u).$$

Let

$$\overline{S}_i(u) = \frac{1}{1+u^2} S_i(u), \quad i = 1,2,3,4,$$

and

$$U_1 = \left\{u : |u| \leq \delta T^{-1/2}\right\} \quad \text{for any } \delta > 0.$$

For $p > 1$,

$$\begin{aligned}
&E|\overline{S}_1(u)|^{2p} \leq CT^{-p}, \\
&E|\overline{S}_1(u_2) - \overline{S}_1(u_1)|^{2p} \leq CT^{-p}|u_2-u_1|^{2p}.
\end{aligned}$$

Therefore,

$$\sup_{u_1\in U_1} |\overline{S}_1(u)| \xrightarrow{P} 0 \text{ as } T\to\infty.$$

It is obvious that

$$\sup_{u\in U_1} |\overline{S}_3(u)| \xrightarrow{P} 0 \text{ as } \frac{T}{n} \to 0.$$

Next,

$$\begin{aligned} S_2(u) &= T^{-1/2}u\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} a'(\theta^*, X_{t_{i-1}})\left[a(\theta^*, X_t) - a(\theta^*, X_{t_{i-1}})\right]dt \\ &= T^{-1/2}u\sum_{i=1}^{n} a'(\theta^*, X_{t_{i-1}})\int_{t_{i-1}}^{t_i}\left(\int_{t_{i-1}}^{t}\dot{a}(\theta^*, X_u)dW_u\right)dt \\ &\quad +T^{-1/2}u\sum_{i=1}^{n} a'(\theta^*, X_{t_{i-1}})\int_{t_{i-1}}^{t_i}\left(\int_{t_{i-1}}^{t}[a(\theta^*, X_u)\dot{a}(\theta^*, X_u) \right. \\ &\quad \left. +\frac{1}{2}\ddot{a}(\theta^*, X_u)]du\right)dt \quad \text{(by the Itô formula)} \\ &=: S_{21}(u) + S_{22}(u). \end{aligned}$$

Define

$$\overline{S}_{2j}(u) = \frac{1}{1+u^2}S_{2j}(u), \quad j = 1, 2.$$

It is easy to show that

$$\sup_{u\in U_1}|\overline{S}_{22}(u)|\xrightarrow{P}0 \ \text{ as } \frac{T}{n^{2/3}} \to 0.$$

As in the estimation of $\overline{S}_1$, we can show that, for $p \geq 1$,

$$\begin{aligned} &E|\overline{S}_{21}(u)|^{2p} \leq C(\frac{T}{n})^p, \\ &\overline{S}_{21}(u_2) - \overline{S}_{21}(u_1)|^{2p} \leq C(\frac{T}{n})^{2p}|u_2 - u_1|^{2p} \end{aligned}$$

Hence,

$$\sup_{u\in U_1}|\overline{S}_{21}(u)|\xrightarrow{P}0 \ \text{ as } \frac{T}{n} \to 0.$$

Thus,

$$\sup_{u\in U_1}\frac{|\widetilde{\rho}_{n,T}(u)|}{1+u^2}\xrightarrow{P}0 \ \text{ as } T\to\infty \ \text{ and } \frac{T}{n^{2/3}} \to 0.$$

On the other hand, for any $\epsilon > 0$, if $\delta > 0$ is small enough, (A2) yields

$$\lim_{T\to\infty,\frac{T}{n}\to 0} P\left\{\sup_{u\in U_1}\frac{1}{1+u^2}|u^2\Gamma - \Gamma_{n,T}| < \epsilon\right\} = 1.$$

Hence, for any $\epsilon > 0$ for small $\delta > 0$,

$$\lim_{n,T} P\{V_{n,T}\} = 0,$$

where

$$V_{n,T} = \{\sup_{u\in U_1}\frac{1}{1+u^2}|\widetilde{\gamma}_{n,T}(u)| < \epsilon\}.$$

Let $\epsilon < \eta$. On the event $V_{n,T}$, ia $|u| \leq \delta T^{1/2}$,

$$\begin{aligned} \log\widetilde{Z}_{n,T}(u) &\leq |u||\widetilde{\Delta}_{n,T}| - \frac{1}{2}u^2\Gamma + |\widetilde{\gamma}_{n,T}(u)| \\ &\leq |u||\widetilde{\Delta}_{n,T}| - \frac{1}{2}u^2\Gamma + \epsilon(1+u^2) \\ &\leq |u||\widetilde{\Delta}_{n,T}| - \frac{1}{2}u^2\Gamma + \eta u^2 + \epsilon \\ &\leq |u||\widetilde{\Delta}_{n,T}| - \frac{1}{4}u^2\Gamma + \epsilon \\ &\leq |u||\widetilde{\Delta}_{n,T}| - \eta u^2 + \epsilon. \end{aligned}$$

Let $U_2 := \{u : q \le |u| \le \delta T^{1/2}\}$, where q is a positive number. We have

$$
\begin{aligned}
& P\left\{\sup_{u\in U_2} \widetilde{Z}_{n,T}(u) \ge \exp\left(-\frac{\eta q^2}{2}\right)\right\} \\
\le\ & P(V^c_{n,T}) + P\left\{\sup_{u\in u_2}\left(|u||\widetilde{\Delta}_{n,T}| - \eta u^2\right) + \epsilon \ge -\frac{\eta q^2}{2}\right\} \\
\le\ & o(1) + P\left\{q|\widetilde{\Delta}_{n,T}| - \eta q^2 + \epsilon \ge -\eta q^2/2\right\} + P\left\{|\widetilde{\Delta}_{n,T}| > 2\eta q\right\} \\
\le\ & 2P\left\{|\widetilde{\Delta}_{n,T}| > \eta q/2 - \epsilon/q\right\} + o(1).
\end{aligned}
$$

Let χ and τ be arbitrary positive numbers. For large q, $\exp(-\eta q^2/2) < \tau$ and

$$\limsup_{T,n}, P\left\{|\widetilde{\Delta}_{n,T}| > \eta q/2 - \epsilon/q\right\} < \frac{\chi}{3}.$$

Then

$$\limsup_{T,n} P\left\{\sup_{u\in U_2} \widetilde{Z}_{n,T}(u) \ge \tau\right\} \le \chi.$$

Define

$$U_3 := \left\{u : |u| \ge \delta T^{1/2}\right\} \quad \text{and } H_1 := \{y : |y| \ge \delta\}.$$

Then

$$
\begin{aligned}
& \limsup_{T\to\infty,\frac{T}{n^{2/3}}\to 0} P\left\{\sup_{u\in U_3} \widetilde{Z}_{n,T}(u) \ge \tau\right\} \\
=\ & \limsup_{T\to\infty,\frac{T}{n^{2/3}}\to 0} P\left\{\sup_{y\in H_1}\left[\widetilde{m}_{n,T}(\theta^*+y) - \widetilde{m}_{n,T}(\theta^*)\right] \ge T^{-1} log\tau\right\} \\
\le\ & \limsup_{T\to\infty,\frac{T}{n^{2/3}}\to 0} P\left\{\sup_{y\in H_1} \widetilde{m}_{n,T}(\theta^*+y) - l(\theta^*+y) + \widetilde{m}_{n,T}(\theta^*) - m(\theta^*) \ge q\right\} \\
& + \limsup_{T\to\infty,\frac{T}{n^{2/3}}\to 0} P\left\{\sup_{y\in H_1} m(\theta^*+y) - m(\theta^*) \ge T^{-1}\log\tau - q\right\}
\end{aligned}
$$

If q is small, the second term on the r.h.s. is zero. The first term tends to zero by Lemma 3.5. Therefore, for $\tau > 0$ and $\chi > 0$, if q is large,

$$\limsup_{T\to\infty,\frac{T}{n^{2/3}}\to 0} P\left\{\sup_{|u|\ge q} \widetilde{Z}_{n,T}(u) > \tau\right\} \le \chi. \qquad \square$$

4. Properties of Approximate M-Estimators

Let $C_0(\mathbb{R})$ be the Banach space of real-valued continuous functions on $\mathbb{R}$ vanishing at the infinity with sup-norm. Let

$$U_{n,T} = \left\{u : \theta^* + T^{-1/2}u \in \Theta\right\}.$$

For $u \in U_{n,T}, Z_{n,T}(u)$ and $\widetilde{Z}_{n,T}(u)$ have been defined and extend it to an element of $C_0(\mathbb{R})$, whose maximal points are contained in $U_{n,T}$. Using the weak convergence of the random field $Z_{n,T}(u)$ (Lemma 3.2, Lemma 3.3 and Lemma 3.4), we obtain the following results.

Theorem 4.1 Under conditions (A1) - (A9), we have

$$Z_{n,T}(\cdot) \stackrel{D}{\to} Z(\cdot) \quad \text{in } C_0(\mathbb{R})$$

as $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$, where $Z(\cdot) = \exp(u\Delta - \frac{1}{2}u^2\Gamma)$,
i.e., for any continuous functional ψ on $C_0(\mathbb{R})$,

$$E\left[\psi(Z_{n,T}(\cdot))\right] \to E[\psi(Z(\cdot))] \text{ as } T \to \infty \text{ and } \frac{T}{n^{2/3}} \to 0.$$

In particular, for the AME1,

$$T^{1/2}(\theta_{n,T}-\theta^*) \xrightarrow{D} \Gamma^{-1}\Delta \sim N(0,\ \beta(\theta^*)\Gamma^{-2}(\theta^*)) \text{ as } T\to\infty \text{ and } \frac{T}{n^{2/3}}\to 0.$$

Proof follows easily by the application of Theorem 5.3.1 in Kutoyants (1984).

Theorem 4.2 Under conditions (A1) - (A10), we have

$$\widetilde{Z}_{n,T}(\cdot) \xrightarrow{D} Z(\cdot) \text{ in } C_0(\mathbb{R})$$

as $T\to\infty$ and $\frac{T}{n^{2/3}}\to 0$ where $Z(u)=\exp(u\Delta-\frac{1}{2}u^2\Gamma)$,
i.e, for any continuous functional ψ on $C_0(\mathbb{R})$,

$$E\left[\psi(\widetilde{Z}_{n,T}(\cdot))\right] \to E\left[\psi(Z(\cdot))\right] | \text{ as } T\to\infty \text{ and } \frac{T}{n^{2/3}}\to 0.$$

In particular, for the AME2,

$$T^{1/2}(\tilde{\theta}_{n,T}-\theta^*) \xrightarrow{D} \Gamma^{-1}\Delta \sim N(0,\beta(\theta)\Gamma^{-2}(\theta)) \text{ as } T\to\infty \text{ and } \frac{T}{n^{2/3}}\to 0.$$

Proof: Using the weak convergence of the AMR random field $\widetilde{Z}_{n,T}(u)$ (Lemmas 3.6–3.8), the theorem follows easily by applying Theorem 5.3.1 in Kutoyants (1984).

5. Approximate Maximum Likelihood Estimation for Misspecified Diffusion Models

If the parametric model contains the true one, then the maximum likelihood estimator is efficient. But the maximum likelihood estimator is sensitive to a contamination of data and a misspecification of the true model. In this section, we study how the difference between the true model and the parametric model affects the asymptotic behavior of the approximate maximum likelihood estimators McKeague (1984) and asymptotic properties of the maximum likelihood estimator in a misspecified diffusion model. As a special case of our M-estimators, we obtain two approximate maximum likelihood estimators, when $a=\mu$ and $b=\mu^2$.

6. Comparison of the AMEs for the Ornstein–Uhlenbeck Process

In this section, we compare the rates of convergence to normality of the Itô AME and the Stratonovich AME in the Ornstein–Uhlenbeck model, when the model may not be correctly specified. Bishwal and Bose (2001) showed that, for the correctly specified Ornstein–Uhlenbeck model, Stratonovich approximate maximum likelihood estimators have faster rate of convergence than the Itô approximate maximum likelihood estimator.

Suppose the true model is the ergodic and stationary

$$dX_t = -\frac{1}{2}X_t dt + dW_t, X_0=\xi,\ \xi\sim N(0,1)$$

which is used to estimate the unknown parameter $\theta<0$ in the observer's parametric model

$$dX_t = -\theta(X_t-1)dt + dW_t, X_0=\xi,$$

where $\theta>0$.

The quasi-true parameter is

$$\theta^* = \arg\inf \int_{\mathbb{R}} [\theta(x-1)-\frac{1}{2}x]^2\nu(dx) = \frac{3}{4},$$

$$\theta_{n,T} = -\frac{\sum_{i=1}^n (X_{t_i-1}-1)(X_{t_i}-X_{t_{i-1}})}{\sum_{i=1}^n (X_{t_i-1}-1)^2\Delta_{t_i}},$$

$$\tilde{\theta}_{n,T} = -\frac{1}{2}\frac{\sum_{i=1}^n((X_{t_i-1}-1)+(X_{t_i-1}-1))(X_{t_i}-X_{t_{i-1}})-\sum_{i=1}^n\Delta t_i}{\sum_{i=1}^n (X_{t_i-1}-1)^2\Delta t_i}$$

$$= -\frac{1}{2}\frac{X_T^2-T-2}{\sum_{i=1}^n (X_{t_i-1}-1)^2\Delta t_i}.$$

Following Bishwal and Bose (2001), it can be shown that $\tilde{\theta}_{n,T}$ has a faster Berry–Esseen bound than the $\theta_{n,T}$.

7. Rates of Convergence of Approximate Integrals

In this section, we show the mean square rate of convergence of the approximations of stochastic integrals. We establish the L^2 rates of convergence of the Itô approximant and the Fisk–Stratonovich approximant to the corresponding integrals, when the integrator is the standard Wiener process.

Throughout the section, we assume that C is a generic constant. We use the following notations:
$f_x := \frac{\partial f}{\partial x}, f_t := \frac{\partial f}{\partial t}, f_{xx} := \frac{\partial^2 f}{\partial x^2}, f_{tt} := \frac{\partial^2 f}{\partial t^2}, f_{tx} := \frac{\partial^2 f}{\partial t \partial x}$.

We assume the following conditions:

(B1) $f(\cdot,\cdot)$ satisfies the Lipschitz and growth conditions:
$$|f(t,x) - f(t,y)| \le K|x-y|, \text{ and } |f(t,x)| \le K(1+|x|) \text{ for all } t \in [0,T]$$
for some constant $K > 0$,

(B2) $f_x(\cdot,\cdot)$ satisfies the Lipschitz and growth conditions:
$$|f_x(t,x) - f_x(t,y)| \le K_1|x-y| \text{ and } |f_x(t,x)| \le K_1(1+|x|) \text{ for all } t \in [0,T]$$
for some constant $K_1 > 0$,

(B3) $f(\cdot,\cdot)$ is a real-valued function satisfying
$$E_{\theta^*}\left\{\int_0^T f^2(t,X_t)dt\right\} < \infty \quad \text{for all } T > 0,$$

$(B4)_j$ $f(\cdot,\cdot)$ is j times continuously differentiable in x for $j = 1,2,\ldots,6$ and $\sup_{0\le t\le T} E|f_x(t,X_t)|^8 < \infty, \quad \sup_{0\le t\le T} E|f_{xx}(t,X_t)|^8 < \infty$,

$(B5)_k$ $f(\cdot,\cdot)$ is k times continuously differentiable in t for $k = 1,2,\ldots,6$ and $\sup_{0\le t\le T} E|f_t(t,X_t)|^8 < \infty, \quad \sup_{0\le t\le T} E|f_{tt}(t,X_t)|^8 < \infty$,

(B6) $\sup_{0\le t\le T} E|f_{tx}(t,X_t)|^8 < \infty$.

Theorem 7.1 Under assumptions (B1) - (B3), $(B4)_1$, (B5) and (B6), we have

(a) $E\left|\sum_{i=1}^n f(t_{i-1},X_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}) - \int_0^T f(t,X_t)dW_t\right|^2 \le C\frac{T^2}{n}$.

Under assumptions (B1) - (B3), $(B4)_2$, (B5) and (B6), we have

(b) $E\left|\sum_{i=1}^n \frac{f(t_{i-1},X_{t_{i-1}}) + f(t_i,X_{t_i})}{2}(W_{t_i} - W_{t_{i-1}}) - \oint_0^T f(t,X_t)dW_t\right|^2 \le C\frac{T^3}{n^2}$.

Proof. Note that the Itô integral is a forward martingale. On the other hand, the FS integral is the arithmetic average of a forward martingale and a backward martingale. Let π_n be the partition as defined in the previous section. Define F and F_{π_n} as
$$F := \oint_0^T f(t,X_t)dW_t \tag{7.1}$$
and
$$F_{\pi_n} := \sum_{i=1}^n \frac{f(t_{i-1},X_{t_{i-1}}) + f(t_i,X_{t_i})}{2}(W_{t_i} - W_{t_{i-1}}). \tag{7.2}$$

Let π_n' be a partition, finer than π_n, obtained by choosing the mid-point $\tilde{t}_{i-1}$ from each of the intervals $t_{i-1} < \tilde{t}_{i-1} < t_i$, $i = 1,2,\ldots,n$. Let $0 = t_0' < t_1' < \ldots < t_{2n}' = T$ be the points of a subdivision of the refined partition π_n'. Define the approximating sum

$F_{\pi_n'}$ as before. First, we will obtain bounds on $E|F_{\pi_n} - F_{\pi_n'}|^2$ in order to get bounds on $E|F_{\pi_n} - F|^2$.

Let $0 \le t_0^* < t_1^* < t_2^* \le T$ be three equally spaced points on $[0,T]$, and let us denote $X_{t_i^*}$ by X_i and $W_{t_i^*}$ by $W_i, i = 0,1,2$. Define

$$\begin{aligned} Z &:= \frac{f(t_2^*, X_2) + f(t_0^*, X_0)}{2}(W_2 - W_0) \\ &\quad - \left\{ \frac{f(t_2^*, X_2) + f(t_1^*, X_1)}{2}(W_2 - W_1) + \frac{f(t_1^*, X_1) + f(t_0^*, X_0)}{2}(W_1 - W_0) \right\} \\ &= (\frac{W_1 - W_0}{2})(f(t_2^*, X_2) - f(t_1^*, X_1)) + (\frac{W_2 - W_1}{2})(f(t_0^*, X_0) - f(t_1^*, X_1)). \end{aligned} \tag{7.3}$$

Let

$$I_1 := \int_{t_0^*}^{t_1^*} f(t, X_t)dt, \ I_2 := \int_{t_1^*}^{t_2^*} f(t, X_t)dt. \tag{7.4}$$

Clearly, by the Taylor expansion,

$$\begin{aligned} &f(t_2^*, X_2) - f(t_1^*, X_1) \\ &= (X_2 - X_1)f_x(t_1^*, X_1) + (t_2^* - t_1^*)f_t(t_1^*, X_1) + \frac{1}{2}(X_2 - X_1)^2 f_{xx}(\tau_1, \nu_1) \\ &\quad + \frac{1}{2}(t_2^* - t_1^*)^2 f_{tt}(\tau_1, \nu_1) + (t_2^* - t_1^*)(X_2 - X_1)f_{tx}(\tau_1, \nu_1) \\ &= (W_2 - W_1 + I_2)f_x(t_1^*, X_1) + (t_2^* - t_1^*)f_t(t_1^*, X_1) + \frac{1}{2}(X_2 - X_1)^2 f_{xx}(\tau_1, \nu_1) \\ &\quad + \frac{1}{2}(t_2^* - t_1^*)^2 f_{tt}(\tau_1, \nu_1) + (t_2^* - t_1^*)(X_2 - X_1)f_{tx}(\tau_1, \nu_1), \end{aligned} \tag{7.5}$$

where $|X_1 - \nu_1| < |X_2 - X_1|, |t_1^* - \tau_1| \le |t_2^* - t_1^*|$, and

$$\begin{aligned} &f(t_0^*, X_0) - f(t_1^*, X_1) \\ &= (X_0 - X_1)f_x(t_1^*, X_1) + (t_0^* - t_1^*)f_t(t_1^*, X_1) + \frac{1}{2}(X_0 - X_1)^2 f_{xx}(\tau_2, \nu_2) \\ &\quad + \frac{1}{2}(t_0^* - t_1^*)^2 f_{tt}(\tau_2, \nu_2) + (t_0^* - t_1^*)(X_0 - X_1)f_{tx}(\tau_2, \nu_2) \\ &= -(W_1 - W_0 + I_1)f_x(t_1^*, X_1) + (t_0^* - t_1^*)f_t(t_1^*, X_1) + \frac{1}{2}(X_0 - X_1)^2 f_{xx}(\tau_2, \nu_2) \\ &\quad + \frac{1}{2}(t_0^* - t_1^*)^2 f_{tt}(\tau_2, \nu_2) + (t_0^* - t_1^*)(X_0 - X_1)f_{tx}(\tau_2, \nu_2), \end{aligned} \tag{7.6}$$

where $|X_1 - \nu_2| < |X_0 - X_1|, |t_1^* - \tau_2| < |t_0^* - t_1^*|$.

Relations (7.3) to (7.6) show that

$$\begin{aligned} Z &= \left(\frac{W_1 - W_0}{2}\right) I_2 f_x(t_1^*, X_1) + \left(\frac{W_1 - W_0}{2}\right)(t_2^* - t_1^*)f_t(t_1^*, X_1) \\ &\quad + \left(\frac{W_1 - W_0}{4}\right)(X_2 - X_1)^2 f_{xx}(\tau_1, \nu_1) + \left(\frac{W_1 - W_0}{4}\right)(t_2^* - t_1^*)^2 f_{tt}(\tau_1, \nu_1) \\ &\quad + \left(\frac{W_1 - W_0}{2}\right)(X_2 - X_1)(t_2^* - t_1^*)f_{tx}(\tau_1, \nu_1) - \left(\frac{W_2 - W_1}{2}\right) I_1 f_x(t_1^*, X_1) \\ &\quad + \left(\frac{W_2 - W_1}{2}\right)(t_0^* - t_1^*)f_t(t_1^*, X_1) + \left(\frac{W_2 - W_1}{4}\right)(X_0 - X_1)^2 f_{xx}(\tau_2, \nu_2) \\ &\quad + \left(\frac{W_2 - W_1}{4}\right)(t_0^* - t_1^*)^2 f_{tt}(\tau_2, \nu_2) \\ &\quad + \left(\tfrac{W_2 - W_1}{2}\right)(X_0 - X_1)(t_0^* - t_1^*)f_{tx}(\tau_2, \nu_2) \\ &\equiv M_1 + M_2, \end{aligned} \tag{7.7}$$

where

$$\begin{aligned} M_1 &:= (W_1 - W_0)\left\{\frac{I_2}{2} f_x(t_1^*, X_1) + \frac{(t_2^* - t_1^*)}{2} f_t(t_1^*, X_1) + \frac{(X_2 - X_1)^2}{4} f_{xx}(\tau_1, \nu_1)\right. \\ &\quad \left. + \frac{(t_2^* - t_1^*)^2}{4} f_{tt}(\tau_1, X_1) + \frac{1}{2}(X_2 - X_1)(t_2^* - t_1^*) f_{tx}(\tau_1, \nu_1)\right\}, \end{aligned} \tag{7.8}$$

$$\begin{aligned} M_2 &:= (W_2 - W_1)\left\{-\frac{I_1}{2} f_x(t_1^*, X_1) + \frac{(t_0^* - t_1^*)}{2} f_t(t_1^*, X_1) + \frac{(X_1 - X_0)^2}{4} f_{xx}(\tau_2, \nu_2)\right. \\ &\quad \left. + \frac{(t_0^* - t_1^*)^2}{4} f_{tt}(\tau_2, X_2) + \frac{1}{2}(X_0 - X_1)(t_0^* - t_1^*) f_{tx}(\tau_2, \nu_2)\right\}. \end{aligned} \tag{7.9}$$

Clearly, $E(Z^2) \leq 2\left[E(M_1^2) + E(M_2^2)\right]$. Notice that M_2 corresponding to different subintervals of $[0,T]$-generated by π_n form a martingale difference sequence, and M_1 corresponding to different subintervals of $[0,T]$ generated by π_n form a reverse martingale difference sequence.

Observe that

$$\begin{aligned} &E|M_2|^2 \\ = \;& E(W_2 - W_1)^2 E\left\{-\frac{I_1}{2} f_x(t_1^*, X_1) + \frac{(t_0^* - t_1^*)}{2} f_t(t_1^*, X_1) + \frac{(X_1 - X_0)^2}{4} fxx(\tau_2, \nu_2)\right. \\ & \left. + \frac{(t_0^* - t_1^*)^2}{4} f_{tt}(\tau_2, \nu_2) + \frac{1}{2}(X_0 - X_1)(t_0^* - t_1^*) f_{tx}(\tau_2, \nu_2)\right\}^2 \\ \leq \;& 4(t_2^* - t_1^*)\left\{E(-\frac{I_1}{2} f_x(t_1^*, X_1))^2 + \frac{(t_0^* - t_1^*)^2}{4} E(f_t(t_1^*, X_1))^2\right. \\ & + E\left\{\tfrac{(X_1 - X_0)^4}{16}(f_{xx}(\tau_2, \nu_2))^2\right\} + \tfrac{(t_0^* - t_1^*)^2}{16} E(f_{tt}(\tau_2, \nu_2))^2 \\ & \left. + \frac{1}{4}(t_0^* - t_1^*)^2 E\{(X_0 - X_1) f_{tx}(\tau_2, \nu_2)\}^2\right\} \\ \leq \;& 4(t_2^* - t_1^*)\left\{\frac{E(I_1)^4}{16} E(f_x(t_1^*, X_1))^4\right\}^{1/2} + \frac{(t_0^* - t_1^*)^2}{4} E(f_t(t_1^*, X_1))^2 \\ & + \left\{\tfrac{E(X_1 - X_0)^8}{256} E(f_{xx}(\tau_2, \nu_2))^2\right\}^{1/2} + \tfrac{(t_0^* - t_1^*)^2}{16} E(f_{tt}(\tau_2, \nu_2))^2 \\ & + \frac{1}{4}(t_0^* - t_1^*)^2 E(X_0 - X_1)^4 E(f_{tx}(\tau_2, \nu_2))^4\Big\}^{1/2} \\ \leq \;& C\ (t_2^* - t_1^*)\left[\{E(I_1)^4\}^{1/2} + \{E(X_1 - X_0)^8\}^{1/2}\right] \end{aligned} \tag{7.10}$$

by $(B4)_2$.

By Theorem 4 of Gikhman and Skorohod (1972, p. 48), there exists $C > 0$ such that

$$E(X_1 - X_0)^{2p} \leq C(E(X_0^{4p}) + 1)(t_1^* - t_0^*)^p, p \geq 1 \tag{7.11}$$

and, by $(B4)_2$,

$$\begin{aligned} E(I_1^4) &= E\left(\int_{t_0^*}^{t_1^*} f(t, X_t) dt\right)^4 \\ &\leq K^4 E\left(\int_{t_0^*}^{t_1^*} (1 + |X_t|) dt\right)^4 \\ &\leq 4K^4 (t_1^* - t_0^*)^4 \sup_{0 \leq t \leq T} E(1 + |X_t|)^4) \\ &\leq C(t_1^* - t_0^*)^4. \end{aligned} \tag{7.12}$$

Relations (7.10) - (7.12) prove that

$$E(M_2^2) \leq C(t_2^* - t_1^*)(t_1^* - t_0^*)^2 \tag{7.13}$$

for some constant $C > 0$ independent of t_0^*, t_1^* and t_2^*. Let us now estimate $E(M_1^2)$. Note that

$$
\begin{aligned}
&E(M_1^2)\\
= \; & E\Bigg[(W_1 - W_0)\left\{\frac{I_2}{2} f_x(t_1^*, X_1) + \frac{(t_2^* - t_1^*)}{2} f_t(t_1^*, X_1) + \frac{(X_2 - X_1)^2}{4} f_{xx}(\tau_1, \nu_1)\right.\\
&\left. + \frac{(t_2^* - t_1^*)^2}{4} f_{tt}(\tau_1, X_1) + \frac{1}{2}(X_2 - X_1)(t_2^* - t_1^*) f_{tx}(\tau_1, \nu_1)\right\}\Bigg]^2\\
\le \; & E\Bigg[(W_1 - W_0)^4 E\left\{\frac{I_2}{2} f_x(t_1^*, X_1) + \frac{(t_2^* - t_1^*)}{2} f_t(t_1^*, X_1) + \frac{(X_2 - X_1)^2}{4} f_{xx}(\tau_1, \nu_1)\right.\\
&\left. + \frac{(t_2^* - t_1^*)^2}{4} f_{tt}(\tau_1, X_1) + \frac{1}{2}(X_2 - X_1)(t_2^* - t_1^*) f_{tx}(\tau_1, \nu_1)\right\}^4\Bigg]^{1/2}\\
& \text{(by the Cauchy-Schwarz inequality)}\\
\le \; & C(t_1^* - t_0^*)^2 \left\{\frac{E(I_2 f_x(t_1^*, X_1)^4}{16} + \frac{(t_2^* - t_1^*)^4}{16} E(f_t(t_1^*, X_1))^4\right.\\
& + \frac{E((X_2 - X_1) f_{xx}(\tau_1, \nu_1))^4}{256} + \frac{(t_2^* - t_1^*)^8}{256} E(f_{tt}(\tau_1, X_1))^4\\
&\left. + \frac{1}{16}(t_2^* - t_1^*) E((X_2 - X_1) f_{tx}(\tau_1, \nu_1))^4\right\}^{1/2}\\
& \text{(by } C_r \text{ - inequality and the fact that } E(W_1 - W_0)^4 = 3(t_1^* - t_0^*)^2 \text{)}\\
\le \; & C(t_1^* - t_0^*)^2 \left\{\frac{E(I_2 f_x(t_1^*, X_1))^4}{16} + \frac{(t_2^* - t_1^*)^4}{16} E(f_t(t_1^*, X_1))^4\right.\\
& + \frac{E((X_2 - X_1) f_{xx}(\tau_1, \nu_1))^4}{256} + \frac{(t_2^* - t_1^*)^8}{256} E(f_{tt}(\tau_1, X_1))^4\\
&\left. + \frac{1}{16}(t_2^* - t_1^*)^4 E((X_2 - X_1) f_{tx}(\tau_1, \nu_1))^4\right\}^{1/2}\\
\le \; & C(t_1^* - t_0^*)^2 \left\{\frac{\{E(I_2)^8 E(f_x(t_1^*, X_1))^8\}^{1/2}}{16} + \frac{(t_2^* - t_1^*)^4}{16} E(f_t(t_1^*, X_1))^4\right.\\
& + \frac{\{E(X_2 - X_1))^8 E(f_{xx}(\tau_1, \nu_1))^8\}^{1/2}}{256} + \frac{(t_2^* - t_1^*)^8}{256} E(f_{tt}(\tau_1, X_1))^4\\
&\left. + \frac{1}{16}(t_2^* - t_1^*)^4 \{E((X_2 - X_1)^8 E(f_{tx}(\tau_1, \nu_1))^8\}^{1/2}\right\}^{1/2}\\
& \text{(by the Cauchy–Schwarz inequality)}
\end{aligned}
\tag{7.14}
$$

Note that there exists a constant $C > 0$ such that

$$E(X_2 - X_1)^8 \le C(t_2^* - t_1^*)^4 \tag{7.15}$$

by Theorem 4 of Gikhman and Skorohod (1972, p. 48).

Furthermore, by (A5)

$$
\begin{aligned}
E(I_2)^8 \; &= \; E\left[\int_{t_1^*}^{t_2^*} f(t, X_t) dt\right]^8 \le CE\left[\int_{t_1^*}^{t_2^*} (1 + |X_t|) dt\right]^8 \text{ by (A1)}\\
&\le \; CE\left[\left\{\int_{t_1^*}^{t_2^*} (1 + |X_t|) dt\right\}^2\right]^4 \le CE\left[(t_2^* - t_1^*)\left\{\int_{t_1^*}^{t_2^*} (1 + |X_t|)^2 dt\right\}\right]^4\\
&\le \; C(t_2^* - t_1^*)^4 E\left[\int_{t_1^*}^{t_2^*} (1 + |X_t|^2) dt\right]^4\\
&\le \; C(t_2^* - t_1^*)^7 \int_{t_1^*}^{t_2^*} E(1 + |X_t|^8) dt\\
&= \; C(t_2^* - t_1^*)^8 \text{ (by (AB)}_2\text{)}.
\end{aligned}
\tag{7.16}
$$

Relations (7.14) - (7.16) prove that

$$E(M_1^2) \leq C(t_1^* - t_0^*)(t_2^* - t_1^*)^2 \tag{3.17}$$

for some constant $C > 0$ independent of t_0^*, t_1^* and t_2^*. Inequalities (7.13) and (7.17) prove that there exists a constant $C > 0$ independent of t_0^*, t_1^*, and t_2^* such that

$$E(M_i^2) \leq C(t_2^* - t_1^*)^3, \quad i = 1, 2. \tag{7.18}$$

Using the property that M_2 corresponding to different subintervals form a martingale difference sequence, and M_1 form a reverse martingale difference sequence, it follows that

$$E|F_{\pi_n} - F_{\pi'_n}|^2 \leq C\frac{T^3}{n^2} \tag{7.19}$$

for some constant $C > 0$.

Let $\{\pi_n^{(q)}, q \geq 0\}$ be the sequence of partitions such that $\pi_n^{(i+1)}$ is a refinement of $\pi_n^{(i)}$ by choosing the mid-points of the subintervals generated by $\pi_n^{(i)}$. Note that $\pi_n^{(0)} = \pi_n$ and $\pi_n^{(1)} = \pi'_n$. The analysis given above proves that

$$E|F_{\pi_n}(q) - F_{\pi_n}(q+1)|^2 \leq C\frac{T^3}{2^q n^2}, \tag{7.20}$$

where $F_{\pi_n}(q)$ is the approximant corresponding to $\pi_n^{(q)}$ and $F_{\pi_n}^{(0)} = F_{\pi_n}$. Therefore,

$$\begin{aligned} E|F_{\pi_n} - F_{\pi_n}(q+1)|^2 &= E\left\{\sum_{k=0}^{q}[F_{\pi_n}(k) - F_{\pi_n}(k+1)]\right\}^2 \\ &\leq \left\{\sum_{k=0}^{q}\left(E|F_{\pi_n}(k) - F_{\pi_n}(k+1)|^2\right)^{\frac{1}{2}}\right\}^2 \\ &\leq \left\{\sum_{k=0}^{q}\left(\frac{CT^3}{2^k n^2}\right)^{1/2}\right\}^2 \leq C\frac{T^3}{n^2} \end{aligned}$$

for all $q \geq 0$. Let $q \to \infty$. Since the integral F exists, $F_{\pi_n}(q+1)$ converges in L_2 to F as $q \to \infty$. Note that $\{\pi_n(q+1), q \geq 0\}$ is a sequence of partitions such that the mesh of the partition tends to zero as $q \to \infty$ for any fixed n. Therefore,

$$E|F_{\pi_n} - F|^2 \leq C\frac{T^3}{n^2} \leq Ch^2, \tag{7.21}$$

where

$$F = \lim_{n\to\infty} F_{\pi_n} = \oint_0^T f(t, X_t)dW_t.$$

To prove (a), let π_n be the partition as defined previously, and let I_{π_n} and I be defined by

$$I_{\pi_n} := \sum_{i=1}^{n} f(t_{i-1}, X_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}), \qquad I := \int_0^T f(t, X_t)dW_t.$$

By arguments used to establish (3.2) and by noting that $\{I_{\pi_n}, n \geq 1\}$ is a martingale, it can be easily shown that (in this case, we need the conditions of existence and finite moments for the first derivative of f only)

$$E|I_{\pi_n} - I|^2 \leq C\frac{T^2}{n}. \tag{7.22}$$

This completes the proof of the theorem. □

Remarks

(1) The partial differential equation defined in (A8) is an elliptic type equation. If the model is correctly specified, condition (A8) can be eliminated for approximate maximum likelihood estimators since, in this case, $\rho'(\theta, x) = 0$. Moreover, the approximate maximum likelihood estimators (AMLEs) are consistent with asymptotic variance $\Gamma^{-1}(\theta^*)$.

(2) The problem of obtaining the rates of convergence (both large deviations and Berry–Esseen type bounds) of the AMEs now remains open. Note that if the model

contains the true model, then, in the linear case, it is known that AME2 has a faster rate of convergence than AMLE1 (see Bishwal and Bose (2001)). We conjecture that, in the general case, AME2 would have a faster rate of convergence than the AME1.

(3) It remains open to obtain the asymptotic normality of AMEs under the SIED condition.

(4) Generalization of the results of this paper to multiparameter case is worth investigating.

References

1. Ait-Sahalia, Y. (2002) : Maximum likelihood estimation of discretely sampled diffusions : a closed form approximation approach, *Econometrica* **70** 223-262.
2. Arato, M. (1982) : *Linear Stochastic Systems with Constant Coefficients : A Statistical Approach*, Lecture Notes in Control and Information Sciences **45**, Springer, New York.
3. Arfi, M. (1995) : Nonparametric drift estimation from ergodic samples, *J. Nonparametric Statist.* **5**, 381-389.
4. Arfi, M. and Lecoutre, J.P. (1994) : Drift estimation under strong mixing, *Pub. Inst. Stat. Stat. Univ. Paris* **38**, 111-127.
5. Banon, G. (1978) : Nonparametric identification for diffusion processes, *SIAM J. Control Optim.* **16**, 380-395.
6. Basawa, I.V. and Prakasa Rao, B.L.S. (1980) : *Statistical Inference for Stochastic Processes*, Academic Press, New York-London.
7. Bibby, B.M. and Sørensen, M. (1995a) : Martingale estimation functions for discretely observed diffusion processes, *Bernoulli* **1**, 17-39.
8. Bibby, B.M. and Sørensen, M. (1998) : On estimation for discretely observed diffusions : A review, *Theory of Stochastic Processes* **2**.
9. Bishwal, J.P.N. (2006) : Rates of weak convergence of the approximate minimum contrast estimator for the discretely observed Ornstein–Uhlenbeck process, *Statistics and Probability Letters* **76** (13), 1397-1409.
10. Bishwal, J.P.N. (2007) : A new estimating function for discretely sampled diffusions, *Random Operators and Stochastic Equations* **15**, 65-88.
11. Bishwal, J.P.N. (2008) : *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, **1923**, Springer, Berlin.
12. Bishwal, J.P.N. and Bose, A. (2001) : Rates of convergence of approximate maximum likelihood estimators in the Ornstein–Uhlenbeck process, *Comput. Math. Appl.* **42**, 23-38.
13. Broze, L., Scaillet, O. and Zakoian, J.M. (1998) : Quasi-indirect inference for diffusion processes, *Ecometric Theory* **14**, 161-186.
14. Clement, E. (1993) : Inference statistique des processes de diffusion, Research Report No. 9404, CREST-INSEE, Universite Paris VI.
15. Clement, E. (1995) : Estimation des processes de diffusion par methodes simulees, Research Report No.9507, CREST-INSEE, Universite Paris VI.
16. Clement, E. (1997) : Estimation of diffusion processes by simulated moments methods, *Scand. J. Statist.* **24**(3), 353-369.
17. Coutin, L. (1994) : Estimation des coefficients d'un processes cadlag observe a des temps discrets, *Pub. Inst. Stat. Univ. Paris* **38**, 87-109.
18. Dacunha-Castelle, D. and Florens-Zmirou, D. (1986) : Estimation of the coefficients of a diffusion from discrete observations, *Stochastics*, **19**, 263-284.
19. Dorogovcev, A. Ja. (1976) : The consistency of an estimate of a parameter of a stochastic differential equation, *Theory Prob. Math. Statist.* **10**, 73-82.
20. Duffie, D. and Glynn, P. (1997) : Estimation of continuous time Markov processes sampled at random time intervals, Preprint, Graduate School of Business, Stanford University.
21. Florens-Zmirou, D. (1989) : Approximate discrete-time schemes for statistics of diffusion processes, *Statistics* **20**, 547-557.
22. Genon-Catalot, V. (1990) : Maximum contrast estimation for diffusion processes arom discrete observations, *Statistics* **21**, 99-116.
23. Gikhman, I.I. and Skorohod, A.V. (1972) : *Stochastic Differential Equations*, Springer, Berlin.
24. Gouriéroux, C., Monfort, A. and Renault, E. (1993) : Indirect inference, *Journal of Applied Econometrics* **8**, 585-618.
25. Gouriéroux, C. and Monfort, A. (1994) : Indirect inference for stochastic differential equatios, Discussion paper 9443, CREST-INSEE, University of Paris VI.

26. Harison, V. (1996) : Drift estimation of a certain class of diffusion processes from discrete observations, *Computers Math. Applic.* **31**, 121-133.
27. Ibragimov, I.A. and Has'minskii, R.Z. (1981) : *Statistical Estimation : Asymptotic Theory*, Springer, Berlin.
28. Kasonga, R.A. (1988) : The consistency of a nonlinear least squares estimator from diffusion processes, *Stoch. Proc. Appl.* **30**, 263-275.
29. Kessler, M. (1997) : Estimation of an ergodic diffusion from discrete observations, *Scand. J. Statist.* **24**(2), 211-229.
30. Kessler, M. and Sørensen, M. (1995) : Estimating equations based on eigen function for a discretely observed diffusion process, Research Rep. No. 332, Dept. of Theoretical Statistics, Univ. Åarhus, Denmark.
31. Kloeden, P.E., Platen, H, Schurz, H. and Sørensen, M. (1996) : On the effects of discretization of estimators of drift parameters for diffusion processes, *J. Appl. Probab.* **33**, 1061-1076.
32. Kutoyants, Yu. A. (1984) : *Parameter Estimation for Stochastic Processes* (Translated and edited by B.L.S. Prakasa Rao) Heldermann, Berlin.
33. Kutoyants, Yu. A. (1994) : *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.
34. Kutoyants, Yu. A. (1996) : Some problems of nonparametric estimation by observations of ergodic diffusion process, *Statist. Probab. Lett.*
35. Kutoyants, Yu. A. (2003) : *Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag.
36. Lanksa, V. (1979) : Minimum contrast estimation in diffusion processes, *J. Appl. Prob.* **16**, 65-75.
37. Le Breton, A. (1976) : On continuous and discrete sampling for parameter estimation in diffusion type processes, *Mathematical Programming Study* **5**,124-144.
38. McKeague, I.A. (1984) : Estimation for diffusion processes under misspecified models, *J. Appl. Probab.* **21** (1984), 511-520.
39. Mishra, M.N. and Bishwal, J.P.N. (1995) : Approximate maximum likelihood estimation for diffusion processes from discrete observations, *Stoch. Stoch. Reports* **52**, 1-13.
40. Nguyen, H.T. and Pham, T.D. (1981) : Nonparametric estimation in diffusion model by discrete sampling, *Pub. Inst. Stat. Univ. Paris* **26**, 89-109.
41. Pedersen, A.R. (1995a) : A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations, *Scand. J. Statist.* **22**, 55-71.
42. Pedersen, A.R. (1995b) : Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes, *Bernoulli* **1**, 257-279.
43. Pham, T.D. (1981) : Nonparametric estimation of the drift coefficient in the diffusion equation, *Math. Oper. Statist., Ser. Statistics* **12**, 61-73.
44. Yoshida, N. (1990) : Asymptotic behaviour of M-estimator and related random field for diffusion process, *Ann. Inst. Statist. Math.* **42**, 221-251.
45. Yoshida, N. (1992) : Estimation for diffusion processes from discrete observations, *J. Multivariate Anal.* **41**, 220-242.

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg, 9201 University City Blvd., Charlotte, NC 28223-0001
E-mail address: J.Bishwal@uncc.edu