Eur. J. Math. Anal. 2 (2022) 7
doi: 10.28924/ada/ma.2.7

On the Stratonovich Estimator for the Itô Diffusion

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg, 9201 University City Blvd., Charlotte, NC 28223-0001, USA

Correspondence: J.Bishwal@uncc.edu

Abstract. For the parameter appearing non-linearly in the drift coefficient of homogeneous Itô stochastic differential equation having a stationary ergodic solution, the paper obtains the strong consistency of an approximate maximum likelihood estimator based on Stratonovich type approximation of the continuous Girsanov likelihood, under some regularity conditions, when the corresponding diffusion is observed at equally spaced dense time points over a long time interval in the high frequency regime. Pathwise convergence of stochastic integral approximations and their connection to discrete drift estimators is studied. Often it is shown that discrete drift estimators converge in probability. We obtain convergence of the estimator with probability one. Ornstein-Uhlenbeck process is considered as an example.

1. Introduction and Preliminaries

Parameter estimation in diffusion processes based on discrete observations is being paid a lot of attention now a days in view of its application in many fields such as biology, physics, oceanograpgy and especially in finance, see Kutoyants (2004) and Bishwal (2008, 2021).

Consider the Itô stochastic differential equation

$$\begin{aligned} dX_t &= f(\theta, X_t)dt + dW_t, \ t \geq 0 \\ X_0 &= X^0 \end{aligned} \tag{1.1}$$

where $\{W_t, t \geq 0\}$ is a one dimensional standard Wiener process, $\theta \in \Theta$, Θ is a compact subset of $\mathbb{R}$, f is a known real valued function defined on $\Theta \times \mathbb{R}$, the unknown parameter θ is to be estimated on the basis of observation of the proces $\{X_t, t \geq 0\}$. Let θ_0 be the true value of the parameter which is in the interior of Θ. We assume that the process $\{X_t, t \geq 0\}$ is observed at $0 = t_0 < t_1 < \ldots < t_n = T$ with $\Delta t_i := t_i - t_{i-1} = \frac{T}{n} = h, \quad i = 1, 2, \ldots, n$ and $T = dn^{1/2}$ for some fixed real number $d > 0$. We estimate θ from the observations $\{X_{t_0}, X_{t_1}, \ldots, X_{t_n}\}$. This

Received: 5 Dec 2021.

Key words and phrases. Itô stochastic differential equation; Stratonovich integral; diffusion process; discrete observations; high frequency; approximate maximum likelihood estimators; conditional least squares estimator; strong consistency; Monte Carlo methods.

model was first studied by Dorogovcev (1976) who obtained weak consistency of the conditional least squares estimator (CLSE) under some regularity conditions as $T \to \infty$ and $\frac{T}{n} \to 0$. Kasonga (1988) obtained the strong consistency of the CLSE under some regularity conditions as $n \to \infty$ assuming that $T = dn^{1/2}$ for some fixed real number $d > 0$.

Note that the conditional least squares estimator (CLSE) of θ is defined as

$$\theta_{n,T} := \arg\min_{\theta \in \Theta} Q_{n,T}(\theta)$$

$$\text{where} \quad Q_{n,T}(\theta) = \sum_{i=1}^{n} \frac{\left[X_{t_i} - X_{t_{i-1}} - f(\theta, X_{t_{i-1}})h\right]^2}{\Delta t_i}.$$

Note that the CLSE, the Euler-Maruyama estimator and the IAMLE are the same estimator (see Shoji (1997)). For the Ornstein-Uhlenbeck process, Bishwal and Bose (2001) studied the rates of weak convergence of approximate maximum likelihood estimators, which are of conditional least squares type. For the Ornstein-Uhlenbeck process Bishwal (2010a) studied uniform rate of weak convergence for the minimum contrast estimator, which has close connection to Stratonovich-Milstein scheme. Bishwal (2009a) studied Berry-Esseen inequalities for conditional least squares estimator discretely observed nonlinear diffusions. Bishwal (2009b) studied Stratonovich based approximate M-estimator of discretely sampled nonlinear diffusions. Bishwal(2011a) studied Milstein approximation of posterior density of diffusions. Bishwal (2010b) studied conditional least squares estimation in nonlinear diffusion processes based on Poisson sampling. Bishwal (2011b) obtained some new estimators of integrated volatility using the stochastic Taylor type schemes which could be useful for option pricing in stochastic volatility models. In mathematical finance, almost sure optimal hedging has received recent attention. Gobet and Landon (2014) studied the optimal discretization error in the context of hedging error in a multidimensional Itô model where the convergence is studied in an almost sure sense and the discrete trading dates are stopping times which includes the sampling scheme of Karandikar (1995) who studied pathwise convergence of stochastic integrals. Bishwal (2011c) studied higher order approximation of hedging error in the mean square sense. Almost sure hedging and optimality of discretization error motivates our almost sure consistency in estimation problem.

Florens-Zmirou (1989) studied minimum contrast estimator, based on an Euler-Maruyama type first order approximate discrete time scheme of the SDE (1.1) which is given by

$$Z_{t_i} - Z_{t_{i-1}} = f(\theta, Z_{t_{i-1}})(t_i - t_{i-1}) + W_{t_i} - W_{t_{i-1}}, \ i \geq 1, \ Z_0 = X^0.$$

The log-likelihood function of $\{Z_{t_i}, 0 \leq i \leq n\}$ is given by

$$C \sum_{i=1}^{n} \frac{\left[Z_{t_i} - Z_{t_{i-1}} - f(\theta, Z_{t_{i-1}})h\right]^2}{\Delta t_i}.$$

where C is a constant independent of θ. A contrast for the estimation of θ is derived from the above log-likelihood by substituting $\{Z_{t_i}, 0 \leq i \leq n\}$ with $\{X_{t_i}, 0 \leq i \leq n\}$. The resulting contrast is

$$H_{n,T} = C \sum_{i=1}^{n} \frac{\left[X_{t_i} - X_{t_{i-1}} - f(\theta, X_{t_{i-1}})h\right]^2}{\Delta t_i}.$$

and the resulting minimum contrast estimator, called the Euler estimator, is

$$\check{\theta}_{n,T} := \arg\min_{\theta \in \Theta} H_{n,T}(\theta)$$

Florens-Zmirou (1989) showed L_2 consistency of the estimator as $T \to \infty$ and $\frac{T}{n} \to 0$.

If continuous observation of $\{X_t\}$ on the interval $[0, T]$ were available, then the likelihood function of θ would be

$$L_T(\theta) = \exp\left\{\int_0^T f(\theta, X_t) dX_t - \frac{1}{2}\int_0^T f^2(\theta, X_t) dt\right\}, \tag{1.2}$$

(see Liptser and Shiryayev (1977)). In our case we have discrete data and we have to approximate the likelihood to get the MLE. Taking Itô type approximation of the stochastic integral and rectangle rule approximation of the ordinary integral in (1.2) and obtain the approximate likelihood function

$$L_{n,T}(\theta) = \exp\left\{\sum_{i=1}^{n} f(\theta, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) - \frac{h}{2}\sum_{i=1}^{n} f^2(\theta, X_{t_{i-1}})\right\}. \tag{1.3}$$

An approximate maximum likelihood estimate (AMLE) based on $L_{n,T}$ is defined as

$$\hat{\theta}_{n,T} := \arg\max_{\theta \in \Theta} L_{n,T}(\theta).$$

Weak consistency and other properties of this estimator were studied by Yoshida (1992) as $T \to \infty$ and $\frac{T}{n} \to 0$.

Note that the CLSE, the Euler estimator and the AMLE1 are the same estimator (see Shoji (1997)).

In order to obtain a better estimator, which may have faster rate of convergence, we propose a new algorithm. Note that the Itô and the Stratonovich integrals are connected by

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

(see Ikeda and Watanabe (1989)). We transform the Itô integral in (1.2) to Stratonovich integral and apply Stratonovich type approximation of the stochastic integral and rectangular rule type approximation of the ordinary integrals and obtain the approximate likelihood

$$\begin{aligned} \widetilde{L}_{n,T}(\theta) \quad = \quad & \exp\left\{\frac{1}{2}\sum_{i=1}^{n}(f(\theta, X_{t_{i-1}}) + f(\theta, X_{t_i}))(X_{t_i} - X_{t_{i-1}})\right. \\ & \left. -\frac{h}{2}\sum_{i=1}^{n}(\dot{f}(\theta, X_{t_{i-1}}) + f^2(\theta, X_{t_{i-1}}))\right\}. \end{aligned} \tag{1.4}$$

The Stratonovich approximate maximum likelihood estimator (SAMLE) based on $\widetilde{L}_{n,T}$ is defined as

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

This estimator is known to have faster rate of convergence (in the mean square sense) than the conditional least squares estimator, see Bishwal (2009b).

For Monte Carlo simulations in finance, one would be interested for pathwise convergence of the estimator. In this paper prove the strong consistency of the SAMLE under some regularity conditions given below as $n \to \infty$. We shall use the following notations : $\Delta X_i = X_{t_i} - X_{t_{i-1}}$, $\Delta W_i = W_{t_i} - W_{t_{i-1}}$, C is a generic constant independent of h, n and other variables (perhaps it may depend on θ). *Prime denotes derivative w.r.t. θ and dot denotes derivative w.r.t. x.* Suppose that θ_0 denote the true value of the parameter and $\theta_0 \in \Theta$. We assume the following conditions:

(A1) The parameter space Θ is compact.

(A2) $|f(\theta, x)| \leq K(\theta)(1 + |x|)$,

$$|f(\theta, x) - f(\theta, y)| \leq K(\theta)|x - y|.$$

$|f(\theta, x) - f(\phi, y)| \leq C(x)|\theta - \phi|$ for all $\theta, \phi \in \Theta, x, y \in \mathbb{R}$ where

$$\sup_{\theta\in\Theta} |K(\theta)| = K < \infty, E|C(X^0)|^m = C_m < \infty \text{ for some } m > 16.$$

(A3) The diffusion process X is stationary and ergodic with invariant measure ν, i.e., for any g with $E[g(\cdot)] < \infty$

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

Further $E|X^0|^m < \infty$ for some $m > 16$.

(A4) $E|f(\theta, X_0) - f(\theta, X_0)|^2 = 0$ iff $\theta = \theta_0$.

(A5) f is twice continuously differentiable function in x with

$$E\sup_t |\dot{f}(X_t)|^2 < \infty, \qquad E\sup_t |\ddot{f}(X_t)|^2 < \infty.$$

2. Main Results

We shall use the following theorem to prove the strong consistency of the SAMLE.

Theorem 2.1 *(Frydman (1980). Suppose the random function D_n satisfy the following conditions:*
(C1) With probability one, $D_n(\theta) \to D(\theta)$ uniformly in $\theta \in \Theta$ as $n \to \infty$.
(C2) The limiting nonrandom function D is such that

$$D(\theta_0) \geq D(\theta) \text{ for all } \theta \in \Theta.$$

(C3) $D(\theta) = D(\theta_0)$ *iff* $\theta = \theta_0$.
Then $\theta_n \to \theta_0$ *a.s. as* $n \to \infty$, *where* $\theta_n = \sup_{\theta\in\Theta} D_n(\theta)$.

We need the following lemmas in order to prove our main result.

Lemma 2.1 *Under (A1)- (A5),*

$$\sup_{\theta\in\Theta} \frac{1}{2T}\left\{\sum_{i=1}^{n}\left[v(\theta, X_{t_{i-1}}) + v(\theta, X_{t_i})\right]\Delta W_i - \frac{h}{2}\sum_{i=1}^{n}\left[\dot{v}(\theta, X_{t_{i-1}}) + \dot{v}(\theta, X_{t_i})\right]\right\} \to 0 \text{ a.s.}$$

as $T \to \infty$, $\frac{T}{n} \to 0$.

Proof. Let $v(\theta, x) := f(\theta, x) - f(\theta_0, x)$. The Fourier expansion of $v(\theta, x)$ in $\mathcal{L}(\Theta)$ be given by

$$v(\theta, x) = \sum_{m=1}^{\infty} a_m(x) e^{\pi j m\theta}, \quad j = \sqrt{-1}, \ x \in \mathbb{R}$$

where $a_k(x)$ are the Fourier coefficients. Thus

$$\begin{aligned} & \frac{1}{2T}\left\{\sum_{i=1}^{n}\left[v(\theta, X_{t_{i-1}}) + v(\theta, X_{t_i})\right]\Delta W_i - \frac{h}{2}\sum_{i=1}^{n}\left[\dot{v}(\theta, X_{t_{i-1}}) + \dot{v}(\theta, X_{t_i})\right]\right\} \\ = & \frac{1}{2T}\left\{\sum_{m=1}^{\infty}\sum_{i=1}^{n}\left[a_m(X_{t_{i-1}}) + a_m(X_{t_i})\right] e^{\pi j m\theta}\Delta W_i \right. \\ & \left. -\frac{h}{2}\sum_{m=1}^{\infty}\sum_{i=1}^{n}\left[\dot{a}_m(X_{t_{i-1}}) + \dot{a}_m(X_{t_i})\right] e^{\pi j m\theta}\right\} \end{aligned}$$

where

$$|a_m(x)| \le c_m|x|, \quad \sum_{m=1}^{\infty} m^{1+\gamma} c_m^4 < \infty.$$

Let

$$A_{m,n}(s) := \frac{1}{2}\sum_{i=1}^{n}\left[a_m(X_{t_{i-1}}) + a_m(X_{t_i})\right] I_{(t_{i-1}-t_i]}(s)$$

where $I_{(t_{i-1}-t_i]}$, $i = 1, 2, ..., n$ are indicator functions. Then

$$\frac{1}{2}\sum_{i=1}^{n}\left[a_m(X_{t_{i-1}}) + a_m(X_{t_i})\right]\Delta W_i = \int_0^T A_{m,n}(s) \circ dW_s$$

and

$$\frac{h}{2}\sum_{i=1}^{n}\left[\dot{a}_m(X_{t_{i-1}}) + \dot{a}_m(X_{t_i})\right] = \int_o^T \dot{A}_{m,n} ds.$$

But

$$\int_0^T A_{m,n}(s) \circ dW_s - \frac{1}{2}\int_o^T \dot{A}_{m,n} ds = \int_0^T A_{m,n}(s) dW_s.$$

By exponential inequality for martingales, we have

$$P\left\{\int_0^T A_{m,n}(s)dW_s - \frac{\alpha}{2}\int_o^T A_{m,n}^2 ds > \beta\right\} \leq e^{-\alpha\beta}$$

for any $\alpha, \beta > 0$. Thus

$$P\left\{\frac{1}{T}\int_0^T A_{m,n}(s)dW_s > \frac{\beta}{T} + \frac{\alpha}{2T}\int_o^T A_{m,n}^2 ds\right\} \leq e^{-\alpha\beta}$$

and

$$P\left\{\left|\frac{1}{T}\int_0^T A_{m,n}(s)dW_s\right| > \frac{\beta}{T} + \frac{\alpha h}{8T}\sum_{i=1}^n \left[a_m(X_{t_{i-1}}) + a_m(X_{t_i})\right]^2\right\} \leq 2e^{-\alpha\beta}.$$

Since

$$\frac{h}{2T}\sum_{i=1}^n \left[a_m(X_{t_{i-1}}) + a_m(X_{t_i})\right]^2 \leq c_m^2\frac{h}{T}\sum_{i=1}^n \left[(X_{t_{i-1}})^2 + (X_{t_i})^2\right]$$

and by (A3)

$$\frac{h}{2T}\sum_{i=1}^n \left[(X_{t_{i-1}})^2 + (X_{t_i})^2\right] \to E(X_0^2) > 0 \ a.s.,$$

there exists a random variable V such that

$$\frac{h}{2T}\sum_{i=1}^n \left[(X_{t_{i-1}})^2 + (X_{t_i})^2\right] < V \ a.s.$$

for all $T > 0, n = 1, 2, \ldots$. where $P(V < \infty) = 1$.

Denote

$$Z_{m,n} := \frac{1}{t_n}\int_0^{t_n} A_{m,n}(s)dW_s.$$

Recall that $T = t_n$. Choose

$$\alpha := \frac{m^a}{t_n^\delta}, \quad \beta := \frac{t_n^\gamma}{m^b},$$

where $\delta < \gamma < 1$ and $\frac{1}{2} < b < \frac{1+\gamma}{2}$.

Then

$$P\left(|Z_{m,n}| > \frac{1}{t_n^{1-\gamma}m^b} + \frac{m^a c_m^2 V}{2t_n^\delta}\right) < 2e^{-m^{a-b}t_n^{\gamma-\delta}}.$$

This

$$
\begin{aligned}
&P\left(\sum_{m=1}^{\infty} Z_{m,n}^2 > \sum_{m=1}^{\infty}\left(\frac{1}{t_n^{1-\gamma}m^b}+\frac{m^a c_m^2 V}{2t_n^{\delta}}\right)^2\right)\\
&\leq \sum_{m=1}^{\infty} P\left(Z_{m,n}^2 > \left(\frac{1}{t_n^{1-\gamma}m^b}+\frac{m^a c_m^2 V}{2t_n^{\delta}}\right)^2\right)\\
&= \sum_{m=1}^{\infty} P\left(|Z_{m,n}| > \frac{1}{t_n^{1-\gamma}m^b}+\frac{m^a c_m^2 V}{2t_n^{\delta}}\right)\\
&\leq 2\sum_{m=1}^{\infty} e^{-m^{a-b}t_n^{\gamma-\delta}}\\
&\leq 2e^{-t_n^{\gamma-\delta}}\sum_{m=1}^{\infty} e^{-m^{a-b}}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\sum_{n=1}^{\infty} P\left(\sum_{m=1}^{\infty} Z_{m,n}^2 > \sum_{m=1}^{\infty}\left(\frac{1}{t_n^{1-\gamma}m^b}+\frac{m^a c_m^2 V}{2t_n^{\delta}}\right)^2\right)\\
&\qquad \leq 2\sum_{n=1}^{\infty} e^{-t_n^{1-\gamma}}\sum_{m=1}^{\infty} e^{-m^{a-b}} < \infty
\end{aligned}
$$

since $\gamma-\delta>0$ and $a-b>0$. The above implies

$$
\sum_{n=1}^{\infty} P\left(\sum_{m=1}^{\infty} Z_{m,n}^2 > \frac{2}{t_n^{2(1-\gamma)}}\sum_{m=1}^{\infty} m^{-2b}+\frac{V^2}{t_n^{2\delta}}\sum_m m^{2a}c_m^4\right) < \infty.
$$

By Borel-Cantelli lemma,

$$
\begin{aligned}
&\sum_{m=1}^{\infty}\left(\frac{1}{2t_n}\sum_{i=1}^{n}\left[a_m(X_{t_{i-1}})+a_m(X_{t_i})\right]\Delta W_i-\frac{h}{2t_n}\sum_{i=1}^{n}\left[\dot{v}(\theta,X_{t_{i-1}})+\dot{v}(\theta,X_{t_i})\right]\right)^2\\
&\qquad \longrightarrow 0 \ a.s. \ as \ n\to\infty.
\end{aligned}
$$

This completes the proof of the lemma. □

Lemma 2.2 *Under (A1)– (A5), with probability one,*

$$
\sup_{\theta\in\Theta}\left|\frac{1}{T}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}[f(\theta_0,X_s)-f(\theta_0,X_{t_{i-1}})]v(\theta,X_{t_{i-1}})ds\right|\to 0.
$$

Proof. For $m>0$, we have

$$
\begin{aligned}
&E\left\{\sup_{\theta\in\Theta}\left|\frac{1}{T}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}[f(\theta_0,X_s)-f(\theta_0,X_{t_{i-1}})]v(\theta,X_{t_{i-1}})ds\right|^{2m}\right\}\\
=\ &E\left\{\sup_{\theta\in\Theta}\left|\frac{1}{T}\int_0^T G_n(s)ds\right|^{2m}\right\}.
\end{aligned}
$$

where $G_n(s) = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} [f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})] v(\theta, X_{t_{i-1}})$ if $t_{i-1} \leq s \leq t_i$.

Hölder's inequality implies that

$$\begin{aligned} & E\left\{\sup_{\theta\in\Theta}\left|\frac{1}{T}\int_0^T G_n(s)ds\right|^{2m}\right\} \\ \leq\ & T^{-2m}E\left\{\sup_{\theta\in\Theta} T^{2m-1}\int_0^T |G_n(s)|^{2m}ds\right\} \\ \leq\ & T^{-2m}E\left(\sup_{\theta\in\Theta} T^{2m-1}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} |f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})|^{2m} |v(\theta, X_{t_{i-1}})|^{2m} ds\right) \\ & \leq T^{-1}U_m \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} E(|f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})|^{2m} |C(X_{t_{i-1}})|^{2m} ds) \end{aligned}$$

by condition (A2) where $U_m := \sup_{\theta\in\Theta} |\theta - \theta_0|^{2m} < \infty$.

By Cauchy-Schwarz's inequality the above term is

$$\begin{aligned} & \leq T^{-1}U_m \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} (E|f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})|^{4m})^{1/2} (E(C(X_{t_{i-1}})|^{4m})^{1/2} ds \\ \leq\ & T^{-1}U_m K^{2m}(\theta_0)(E|C(X_0)|^{4m})^{1/2} \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} (E|X_s - X_{t_{i-1}})|^{4m})^{1/2} ds \end{aligned}$$

by condition (A2). Since $E|X_t - X_s|^{2m} \leq M(t-s)^m$, from Gikhman and Skorohod (1975, p.48), the above term

$$\begin{aligned} \leq\ & T^{-1}U_m K^{2m}(\theta_0)(E|C(X_0)|^{4m})^{1/2} M^{1/2} \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} (s - t_{i-1})^m ds \\ =\ & U_m K^{2m}(\theta_0)(E|C(X_0)|^{4m} M)^{1/2} T^{-1} \sum_{i=1}^{n} \frac{(\Delta t_i)^{m+1}}{m+1} \\ \leq\ & \frac{U_m K^{2m}(\theta_0)}{m+1}(E|C(X_0)|^{4m} M)^{1/2} h^m n^{-m/2}, \quad m > 4. \end{aligned}$$

Chebyshev's inequality and the above implies that for any $\epsilon > 0$,

$$\sum_{n=1}^{\infty} P\left\{\sup_{\theta\in\Theta}\left|\frac{1}{T}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} [f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})] v(\theta, X_{t_{i-1}}) ds\right| > \epsilon\right\} < \infty.$$

Hence Borel-Cantelli lemma yields the result. □

Lemma 2.3 *Under (A1)- (A6), with probability one,*

$$\frac{1}{T}\sum_{i=1}^{n} [f(\theta, X_{t_{i-1}}) - f(\theta_0, X_{t_{i-1}})]^2 \Delta t_i \to E|v(\theta, X^0)|^2$$

uniformly in θ *as* $T \to \infty$, $\frac{T}{n} \to 0$.

Proof. By the strong law of large numbers (ergodicity),

$$\frac{1}{T}\int_0^T |v(\theta, X_s)^2 ds \to E|v(\theta, X_0)|^2.$$

a.s. as $T \to \infty$ for each $\theta \in \Theta$. The condition (A2) implies that

$$\frac{1}{T}\int_0^T |v(\theta, X_s)^2 ds \leq \frac{1}{T}|\theta - \theta_0|^2 \int_0^T |C(X_s)|^2 ds$$

$$\leq \sup_{\theta\in\Theta} |\theta - \theta_0|^2 \frac{1}{T}\int_0^T |C(X_s)|^2 ds \leq B$$

almost surely for some random variable B by (A1), (A2) and (A3). It also follows easily by (A1)-(A4) that

$$\left|\frac{1}{T}\int_0^T |v(\theta_1, X_s)^2 ds - \frac{1}{T}\int_0^T |v(\theta_2, X_s)^2 ds\right| \leq J|\theta_1 - \theta_2|$$

almost surely for some random variable J and $\theta_1, \theta_2 \in \Theta$. Thus the family of functions

$$\left\{\frac{1}{T}\int_0^T |v(\cdot, X_s)|^2 ds, \quad T \geq 0\right\}$$

is equicontinuous. Hence by Arzela-Ascoli theorem, the convergence is uniform. Denote

$$g_n^2(\theta) := \frac{h}{2}\sum_{i=1}^n \left[(X_{t_{i-1}})^2 + (X_{t_i})^2\right].$$

Now it is enough to show that

$$\frac{1}{T}\int_0^T |v(\theta, X_s)|^2 ds - \frac{1}{T}g_n^2(\theta) \to 0$$

a.s. uniformly in θ. We have

$$\begin{aligned} & E\left\{\sup_{\theta\in\Theta} |\int_0^T |v(\theta, X_s)^2 ds - g_n^2(\theta)|^{2m}\right\} \\ & E\left\{\sup_{\theta\in\Theta} |\int_0^T |v(\theta, X_s)^2 ds - h\sum_{i=1}^n |v(\theta, X_{t_{i-1}})|^2|^{2m}\right\} \\ = \quad & E\left\{\sup_{\theta\in\Theta} |\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \sum_{i=1}^n (v(\theta, X_s - v(\theta, X_{t_{i-1}}))(v(\theta, X_s + v(\theta, X_{t_{i-1}}))ds|^{2m}\right\}. \end{aligned}$$

Hölder inequality implies the above expectation

$$\leq T^{2m-1} E \sup_{\theta\in\Theta} \sum_{i=1}^{n} \{ \int_{t_{i-1}}^{t_i} |v(\theta, X_s - v(\theta, X_{t_{i-1}})|^{2m} |v(\theta, X_s + v(\theta, X_{t_{i-1}}))|^{2m} \}$$

$$\leq T^{2m-1} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} E[\sup_{\theta\in\Theta} |v(\theta, X_s - v(\theta, X_{t_{i-1}})|^{2m} \sup_{\theta\in\Theta} |v(\theta, X_s + v(\theta, X_{t_{i-1}}))|^{2m}] ds$$

$$\leq T^{2m-1} K^{2m} 2^{2m} U_m \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} E[|X_s - X_{t_{i-1}}|^{2m} (|C(X_s)|^{2m} + |C(X_{t_{i-1}}))|^{2m}] ds$$

$$\leq T^{2m-1} K^{2m} 2^{2m+1} U_m \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} (E|X_s - X_{t_{i-1}})|^{4m})^{1/2} (E|C(X_s)|^{4m} + E|C(X_{t_{i-1}}))|^{4m})^{1/2} ds$$

$$\leq T^{2m-1} K^{2m} 2^{2m+1} U_m M^{1/2} (E|C(X_0)|^{2m}))^{1/2} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} (s - t_{i-1})^m ds|$$

(by stationarity)

$$\leq R_m T^{2m-1} n (T/n)^{m+1}$$

where $U_m := \sup_{\theta\in\Theta} |\theta - \theta_0|^{2m} < \infty$ and $R_m := K^{2m} 2^{2m+2} U_m M^{1/2} (E|C(X_0)|^{4m})^{1/2}$. Hence if $m > 4$,

$$E \left\{ \sup_{\theta\in\Theta} |\frac{1}{T} \int_0^T |v(\theta, X_s)|^2 ds - \frac{1}{T} g_n^2(\theta)|^{2m} \right\} \leq R_m (T/n)^m \leq R_m h^{m/2} n^{-m/2}.$$

Borel-Cantelli argument yields the result. □

Now we are ready to present the main result of the paper:

Theorem 2.2 Under the conditions (A1)-(A5), the SAMLE is strongly consistent, i.e.,

$$\widetilde{\theta}_{n,T} \to \theta_0 \text{ a.s. as } T \to \infty, \frac{T}{n} \to 0.$$

Proof. Let

$$\widetilde{l}_{n,T}(\theta) := \log \widetilde{L}_{n,T}(\theta)$$

and

$$v(\theta, x) := f(\theta, x) - f(\theta_0, x).$$

Note that

$$\begin{aligned}
& \frac{1}{T}\left[\widetilde{l}_{n,T}(\theta)-\widetilde{l}_{n,T}(\theta_0)\right] \\
= \quad & \frac{1}{2T}\sum_{i=1}^{n}[f(\theta,X_{t_{i-1}})+f(\theta,X_{t_i})](X_{t_i}-X_{t_{i-1}}) \\
& -\frac{1}{2T}\sum_{i=1}^{n}[f(\theta_0,X_{t_{i-1}})+f(\theta_0,X_{t_i})](X_{t_i}-X_{t_{i-1}}) \\
& -\frac{1}{2n}\sum_{i=1}^{n}[\dot{f}(\theta,X_{t_{i-1}})-\dot{f}(\theta_0,X_{t_{i-1}})] \\
& -\frac{1}{2n}\sum_{i=1}^{n}[f^2(\theta,X_{t_{i-1}})-f^2(\theta_0,X_{t_{i-1}})] \\
= \quad & \frac{1}{2T}\left\{\sum_{i=1}^{n}\left[v(\theta,X_{t_{i-1}})+v(\theta,X_{t_i})\right]\Delta W_i - h\sum_{i=1}^{n}\dot{v}(\theta,X_{t_{i-1}})\right\} \\
& -\frac{1}{2n}\sum_{i=1}^{n}v^2(\theta,X_{t_{i-1}}) \\
& -\frac{1}{T}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}v(\theta,X_{t_{i-1}})[f(\theta_0,X_t)+f(\theta_0,X_{t_{i-1}})]dt \\
& -\frac{1}{T}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}[v(\theta,X_{t_i})f(\theta_0,X_t)-v(\theta_0,X_{t_{i-1}})f(\theta_0,X_{t_{i-1}})]dt \\
=: \quad & I_1-I_2-I_3-I_4.
\end{aligned}$$

Let

$$D_{n,T}(\theta) := \frac{1}{T}\left[\widetilde{l}_{n,T}(\theta)-\widetilde{l}_{n,T}(\theta_0)\right].$$

Below Lemma 2.1-2.3 show that

$$D_{n,T}(\theta)\to D(\theta) \text{ a.s. as } T\to\infty, \frac{T}{n}\to 0$$

where

$$D(\theta) := -\frac{1}{2}E|f(\theta,X^0)-f(\theta_0,X^0)|^2.$$

Thus condition (C1) of Theorem 2.1 is satisfied. The limiting function $D(\theta)$ satisfies the conditions (C2) and (C3) of Theorem. Hence as a consequence of Theorem 2.1 we obtain the result. □

3. Ornstein-Uhlenbeck Process

Consider the Ornstein-Uhlenbeck process satisfying

$$dX_t = \theta X_t dt + dW_t,\ t \geq 0,\ X_0 = 0,\ \theta < 0.$$

The Euler Estimator (conditional least squares estimator) is given by

$$\breve{\theta}_{n,T} = \frac{\sum_{i=1}^{n} X_{t_{i-1}}(X_{t_i} - X_{t_{i-1}})}{h\sum_{i=1}^{n} X_{t_{i-1}}^2}.$$

Strong consistency of this estimator is obtained in Kasonga (1988). As a consequence of Theorem 2.2, we obtain the strong consistency of three estimators with

$$\widetilde{\theta}_{n,T} = \frac{(X_T^2 - T)/2}{h\sum_{i=1}^{n} X_{t_{i-1}}^2},\ \bar{\theta}_{n,T,3} = \frac{X_T^2/2}{h\sum_{i=1}^{n} X_{t_{i-1}}^2},\ \widehat{\theta}_{n,T,2} = \frac{-T/2}{h\sum_{i=1}^{n} X_{t_{i-1}}^2}.$$

which are SAMLE, YAMLE (Young AMLE) and , AMCE respectively as $T \to \infty$ and $T/n \to 0$. SAMLE is the linear combination of AMCE and YAMLE.

Define the continuous MLE, YMLE and MCE respectively

$$\theta_{T,1} = \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt},\ \theta_{T,2} = \frac{X_T^2/2}{\int_0^T X_t^2 dt},\ \theta_{T,3} = \frac{-T/2}{\int_0^T X_t^2 dt}.$$

Interpreting $\int_0^T X_t dX_t$ to be the Young (1936) integral, it equals $X_T^2/2$. Belfadli *et al.* (2011) (see also El Machkouri *et al.* (2016)) obtained the strong consistency of $\theta_{T,2}$ as $T \to \infty$. The YAMLE $\bar{\theta}_{n,T,2}$ is the Euler discretization of $\theta_{T,2}$. Lanksa (1979) obtained strong consistency of the MCE $\theta_{T,3}$ as $T \to \infty$ whose Euler discretetization is $\widehat{\theta}_{n,T,3}$. Liptser and Shiryayev (1978) obtained strong consistency of the MLE $\theta_{T,1}$ as $T \to \infty$ whose Euler discretetization is $\breve{\theta}_{n,T,1}$.

Concluding Remark It would be interesting to extend the results of the paper to diffusions driven by persistent fractional Brownian motion which are neither Markov processes nor semimartingales, but preserved long memory property of the model.

References

[1] R. Belfadli, K. Es-Sebaiy and Y. Ouknine, Parameter estimation for fractional Ornstein-Uhlenbeck processes: non-ergodic case, Front. Sci. Eng. 1 (2011) 41-56. https://doi.org/10.34874/IMIST.PRSM/fsejournal-v1i1.26873.

[2] J.P.N. Bishwal, Parameter estimation in stochastic differential equations, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. https://doi.org/10.1007/978-3-540-74448-1.

[3] J.P.N. Bishwal, Berry–Esseen inequalities for discretely observed diffusions, Monte Carlo Methods and Applications. 15 (2009) 229-239. https://doi.org/10.1515/MCMA.2009.013.

[4] J.P.N. Bishwal, M-Estimation for discretely sampled diffusions, Theory Stoch. Processes, 15 (31) (2) (2009b) 62-83.

[5] J.P.N. Bishwal, Uniform rate of weak convergence of the minimum contrast estimator in the Ornstein–Uhlenbeck process, Methodol. Comput. Appl. Probab. 12 (2010) 323–334. https://doi.org/10.1007/s11009-008-9099-x.

[6] J.P.N. Bishwal, Conditional least squares estimation in diffusion processes based on Poisson sampling, J. Appl. Probab. Stat. 5 (2) (2010b) 169-180.

[7] J.P.N. Bishwal, Milstein approximation of posterior density of diffusions, Int. J. Pure Appl. Math. 68 (4) (2011a) 403-414.

[8] J.P.N. Bishwal, Some new estimators of integrated volatility, Amer. Open J. Stat. 1 (2) (2011b) 74-80.

[9] J.P.N. Bishwal, Stochastic moment problem and hedging of generalized Black–Scholes options, Appl. Numer. Math. 61 (2011) 1271–1280. https://doi.org/10.1016/j.apnum.2011.08.005.

[10] J.P.N. Bishwal, Parameter estimation in stochastic volatility models, Springer Nature Switzerland AG (forthcoming). (2021).

[11] J.P.N. Bishwal, A. Bose, Rates of convergence of approximate maximum likelihood estimators in the Ornstein-Uhlenbeck process, Computers Math. Appl. 42 (2001) 23–38. https://doi.org/10.1016/S0898-1221(01)00127-4.

[12] M. El Machkouri, K. Es-Sebaiy, Y. Ouknine, Least squares estimator for non-ergodic Ornstein-Uhlebeck processes driven by Gaussian processes, J. Korean Stat. Soc. 45 (2016) 329-341.

[13] D. Florens-zmirou, Approximate discrete-time schemes for statistics of diffusion processes, Statistics. 20 (1989) 547–557. https://doi.org/10.1080/02331888908802205.

[14] R. Frydman, A proof of the consistency of maximum likelihood estimators of nonlinear regression models with autocorrelated errors, Econometrica. 48 (1980) 853-860. https://doi.org/10.2307/1912936.

[15] E. Gobet, N. Landon, Almost sure optimal hedging strategy, Ann. Appl. Probab. 24 (2014). https://doi.org/10.1214/13-AAP959.

[16] N. Ikeda, S. Watanabe, Stochastic differential equations and diffusion processes, Second Edition, North-Holland, Amsterdam (Kodansha Ltd., Tokyo). (1989).

[17] R.L. Karandikar, On pathwise stochastic integration, Stoch. Processes Appl. 57 (1995) 11–18. https://doi.org/10.1016/0304-4149(95)00002-O.

[18] R.A. Kasonga, The consistency of a non-linear least squares estimator from diffusion processes, Stoch. Processes Appl. 30 (1988) 263-275. https://doi.org/10.1016/0304-4149(88)90088-9.

[19] Y.A. Kutoyants, Statistical inference for ergodic diffusion processes, Springer London, London, 2004. https://doi.org/10.1007/978-1-4471-3866-2.

[20] V. Lánska, Minimum contrast estimation in diffusion processes, J. Appl. Probab. 16 (1979) 65–75. https://doi.org/10.2307/3213375.

[21] R.S. Liptser, A.N. Shiryayev, Statistics of Random Processes I: General Theory Springer-Verlag, Berlin. (1977).

[22] R.S. Liptser, A.N. Shiryayev, Statistics of Random Processes II : Applications Springer-Verlag, Berlin. (1978).

[23] I. Shoji, A note on asymptotic properties of estimator derived from the Euler method for diffusion processes at discrete times, Stat. Probab. Lett. 36 (1997) 153-159. https://doi.org/10.1016/S0167-7152(97)00058-8.

[24] N. Yoshida, Estimation for diffusion processes from discrete observation, J. Multivar. Anal. 41 (1992) 220–242. https://doi.org/10.1016/0047-259X(92)90068-Q.

[25] L.C. Young, An inequality of the Hölder type, connected with Stieltjes integration, Acta Math. 67 (1936) 251–282. https://doi.org/10.1007/BF02401743.