Asymptotic normality of discretized maximum likelihood estimator for drift parameter in homogeneous diffusion model

Kostiantyn Ralchenko[a]

[a]*Taras Shevchenko National University of Kyiv,
Department of Probability Theory, Statistics and Actuarial Mathematics,
Volodymyrska 64/13, 01601 Kyiv, Ukraine*

k.ralchenko@gmail.com (K. Ralchenko)

Abstract   We prove the asymptotic normality of the discretized maximum likelihood estimator for the drift parameter in the homogeneous ergodic diffusion model.

1 Introduction

The statistical inference for diffusion models has been thoroughly studied by now; see the books [4, 6–8, 10] and references therein.

In this paper, we consider the homogeneous diffusion process given by the stochastic differential equation

$$dX_t = \theta a(X_t)\, dt + b(X_t)\, dW_t,$$

where W_t is a standard Wiener process, and θ is an unknown parameter.

The standard maximum likelihood estimator for the parameter θ constructed by the observations of X on the interval $[0, T]$ has the form

$$\hat{\theta}_T = \frac{\int_0^T \frac{a(X_t)}{b(X_t)^2} \, dX_t}{\int_0^T \frac{a(X_t)^2}{b(X_t)^2} \, dt};$$

see, for instance, [7, Example 1.37] and [9]. If the equation has a weak solution, the coefficient a is not identically zero, and the functions $\frac{1}{b^2}, \frac{a^2}{b^2}, \frac{a^2}{b^4}$ are locally integrable, then this estimator is strongly consistent [9, Thm. 3.3]. Moreover, if the model is ergodic, then this estimator is asymptotically normal [7, Ex. 1.37]. Note that in the nonergodic case the maximum likelihood estimator $\hat{\theta}_T$ may have different limit distributions; some examples can be found in [7, Sect. 3.5].

If the data are the observations of the trajectory $\{X_t, t \geq 0\}$ at discrete time moments $t_1, t_2, \ldots$, we obtain the discrete-time version of the model. Parameter estimation in such models has been studied since the mid-1980s; see [2, 3, 11]. A review of this problem and many references can be found in [5] and [13]. For recent results, see [6, 9, 12].

In this paper, we are interested in the scheme of observations that is called "rapidly increasing experimental design." The process X is observed at time moments $t_i = i\Delta_n, i = 0, \ldots, n$, such that $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. One of possible approaches to parameter estimation is to consider a discretized version of the continuous-time MLE $\hat{\theta}_T$. The most general results in this direction were obtained by Yoshida [14]. He proved the consistency and asymptotic normality of the discretized MLE in the model, where the process was multidimensional, the drift coefficient depended on θ nonlinearly, and the diffusion coefficient also contained an unknown parameter.

Assume that we observe the process X at discrete time moments $t_k^n = k/n$, $0 \leq k \leq n^{1+\alpha}$, where $0 < \alpha < \frac{1}{2}$. In this scheme, Mishura [9] proposed the following discretized version of the maximum likelihood estimator:

$$\hat{\theta}_n = \frac{\sum_{k=0}^{n^{1+\alpha}} a(X_{\frac{k}{n}})(X_{\frac{k+1}{n}} - X_{\frac{k}{n}})/b(X_{\frac{k}{n}})^2}{n^{-1} \sum_{k=0}^{n^{1+\alpha}} a(X_{\frac{k}{n}})^2/b(X_{\frac{k}{n}})^2}.$$

She proved its strong consistency in the case where the coefficients a and b are bounded. The aim of this paper is to establish the asymptotic normality of this estimator. Additionally, we assume the ergodicity of the model, but the boundedness of the coefficients is not required. In comparison with general results of Yoshida [14], our assumptions are less restrictive. We assume the polynomial growth of the function $1/b$ instead of the condition $\inf_x b(x)^2 > 0$. Also, we do not assume the smoothness of the coefficients and the polynomial growth of their derivatives; any Lipschitz continuous $a(x)$ and $b(x)$ are possible.

The paper is organized as follows. In Section 2, we describe the model and formulate the results. In Section 3, some simulation experiments are considered. The proof of the main theorem is given in Section 4.

2 Model description and main result

Let $(\Omega, \mathfrak{F})$ be a measurable space. Assume that $\theta \in \mathbb{R}$ is fixed but unknown. Consider a probability measure $\mathbf{P}_\theta$ such that $\mathfrak{F}$ is $\mathbf{P}_\theta$-complete.

Let X solve the equation

$$X_t = x_0 + \theta \int_0^t a(X_s)\,ds + \int_0^t b(X_s)\,dW_s, \tag{1}$$

where $x_0 \in \mathbb{R}$, $a, b\colon \mathbb{R} \to \mathbb{R}$ are measurable functions, and $\{W_t, t \geq 0\}$ is a standard Wiener process on $(\Omega, \mathfrak{F}, \mathbf{P}_\theta)$.

Denote $c(x) = \frac{a(x)}{b(x)^2}$, $d(x) = \frac{a(x)^2}{b(x)^2}$, $\varphi_\theta(x) = \exp\{-2\theta \int_0^x c(y)\,dy\}$, and $\Phi_\theta(x) = \int_0^x \varphi_\theta(y)\,dy$.

Assume that the following conditions hold.

(A1) For some $L > 0$ and for any $x, y \in \mathbb{R}$,

$$\big|a(x) - a(y)\big| + \big|b(x) - b(y)\big| \leq L|x - y|.$$

(A2) $\Phi_\theta(+\infty) = -\Phi_\theta(-\infty) = +\infty.$

(A3) $G_\theta := \int_{-\infty}^{+\infty} \frac{dx}{b(x)^2 \varphi_\theta(x)} < \infty.$

It is well known that under assumption (A1) the stochastic differential equation (1) has a unique strong solution. This assumption also yields that the functions $a(x)$ and $b(x)$ satisfy the linear growth condition, that is,

$$\big|a(x)\big| + \big|b(x)\big| \leq M_1(1 + |x|) \tag{2}$$

for some $M > 0$ and for all $x \in \mathbb{R}$.

Assume additionally that

(A4) There exist $K > 0$ and $p \geq 0$ such that

$$\big|b(x)\big|^{-1} \leq K(1 + |x|^p).$$

Then, for some $M_2 > 0$ and for any $x \in \mathbb{R}$,

$$\big|c(x)\big| \leq M_2(1 + |x|^{2p+1}), \qquad \big|d(x)\big| \leq M_2(1 + |x|^{2p+2}). \tag{3}$$

Under assumptions (A2)–(A3), the diffusion process X is positive recurrent; see, for example, [7, Prop. 1.15]. In this case, it has ergodic properties with the invariant density given by

$$\mu_\theta(x) = \frac{1}{G_\theta b(x)^2 \varphi_\theta(x)}, \quad x \in \mathbb{R}.$$

Let ξ_θ denote a random variable with density $\mu_\theta(x)$. Then, for any measurable function h such that $\mathbf{E}_\theta|h(\xi_\theta)| < \infty$,

$$\frac{1}{T} \int_0^T h(X_t)\,dt \to \int_{-\infty}^{+\infty} h(x)\mu_\theta(x)\,dx \equiv \mathbf{E}_\theta h(\xi_\theta) \quad \text{a.s. as } T \to \infty, \tag{4}$$

see [7, Thm. 1.16]. Moreover, according to [1, Sect. II.37], the convergence (4) holds also in L_1, that is,

$$\frac{1}{T}\mathbf{E}_\theta \int_0^T h(X_t)\,dt \to \mathbf{E}_\theta h(\xi_\theta). \tag{5}$$

Assume that the invariant distribution satisfies the condition

(A5) $\mathbf{E}_\theta|\xi_\theta|^r \equiv \int_{-\infty}^{+\infty} |x|^r \mu_\theta(x)\,dx < \infty$ for all $r \geq 0$.

Let $0 < \alpha < 1$. Suppose that we observe the process X at discrete time moments $t_k^n = k/n$, $0 \le k \le n^{1+\alpha}$. Consider the estimator

$$\hat{\theta}_n = \frac{\sum_{k=1}^{n^{1+\alpha}} c(X_{\frac{k-1}{n}}) \Delta X_k^n}{n^{-1} \sum_{k=1}^{n^{1+\alpha}} d(X_{\frac{k-1}{n}})},$$

where $\Delta X_k^n = X_{\frac{k}{n}} - X_{\frac{k-1}{n}}$.

Assume also that

(A6) a is not identically zero.

Then $\mathbf{E}_\theta d(\xi_\theta) > 0$. Note also that by (3) and (A5), $\mathbf{E}_\theta d(\xi_\theta) < \infty$. Now we are ready to formulate the main result.

Theorem 2.1. *Assume that conditions* (A1)–(A6) *hold. Then*

(i) $\hat{\theta}_n \xrightarrow{\mathbf{P}_\theta} \theta$ *as* $n \to \infty$,

(ii) $n^{\alpha/2}(\hat{\theta}_n - \theta) \Rightarrow N(0, 1/\mathbf{E}_\theta d(\xi_\theta))$ *as* $n \to \infty$.

The proof is given in Section 4.

The following result gives sufficient conditions for consistency and asymptotic normality in the case where the parameter θ is positive.

Corollary 2.2. *Let* $\theta > 0$. *Assume that conditions* (A1), (A4) *and* (A6) *are fulfilled and, additionally,*

$$\limsup_{|x| \to \infty} c(x)\, \mathrm{sgn}(x) < 0. \tag{6}$$

Then

(i) $\hat{\theta}_n \xrightarrow{\mathbf{P}_\theta} \theta$ *as* $n \to \infty$,

(ii) $n^{\alpha/2}(\hat{\theta}_n - \theta) \Rightarrow N(0, 1/\mathbf{E}_\theta d(\xi_\theta))$ *as* $n \to \infty$.

Proof. Note that condition (6), together with (A4), implies that assumptions (A2)–(A3) are satisfied and, moreover, all polynomial moments of the invariant density are finite; see [7, p. 3]. Hence, the result follows directly from Theorem 2.1. □

If the coefficients are bounded, then the consistency and asymptotic normality of $\hat{\theta}_n$ can be obtained without assumption (A5).

Corollary 2.3. *Assume that conditions* (A1)–(A3) *are satisfied, the coefficients* $a(x)$ *and* $b(x)$ *are bounded, and* $\inf_{x \in \mathbb{R}} |b(x)| > 0$. *Then*

(i) $\hat{\theta}_n \xrightarrow{\mathbf{P}_\theta} \theta$ *as* $n \to \infty$,

(ii) $n^{\alpha/2}(\hat{\theta}_n - \theta) \Rightarrow N(0, 1/\mathbf{E}_\theta d(\xi_\theta))$ *as* $n \to \infty$.

Sketch of proof. This result can be proved similarly to Theorem 2.1 using the boundedness of $a(x), b(x), c(x), d(x)$ instead of the growth conditions (2), (3), and (A4) to-

gether with the boundedness of moments of the invariant density. In this case, (8) implies the inequality

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m} \le C(m, \theta) n^{-m}$$

for all $m \in \mathbb{N}$ and $t \in [\frac{k-1}{n}, \frac{k}{n}]$, $k = 1, 2, \ldots n^\alpha$. This estimate is used in the proof instead of Lemmas 4.1–4.2. □

Remark 2.4. For $\alpha \in (0, \frac{1}{2})$, Mishura [9] obtained the a.s. convergence in Corollary 2.3(i) without assumptions (A2)–(A3).

3 Some simulation results

In this section, we illustrate quality of the estimator by simulation experiments. We consider the diffusion process (1) with drift parameter $\theta = 2$ and initial value $x_0 = 1$ in three following cases:

(1) $a(x) = 1 - x$, $b(x) = 2 + \sin x$,

(2) $a(x) = -\arctan x$, $b(x) = 1$,

(3) $a(x) = -\frac{x}{1+x^2}$, $b(x) = 1$.

Using the Milstein method, we simulate 100 sample paths of each process and find the estimate $\hat{\theta}_n$ for different values of n and α. The average values of $\hat{\theta}_n$ and the corresponding standard deviations are presented in Tables 1–3.

Table 1. $a(x) = 1 - x$, $b(x) = 2 + \sin x$

		n					
		50	100	500	1000	2000	5000
$\alpha = 0.1$	Mean	3.05812	2.97626	2.73973	2.58453	2.55888	2.53879
	Std. dev.	2.06388	2.00007	1.43273	1.34689	1.26920	1.22077
$\alpha = 0.5$	Mean	2.11065	2.15066	2.08157	2.05626	2.03686	2.03479
	Std. dev.	0.62613	0.56038	0.31621	0.28909	0.22875	0.18187
$\alpha = 0.9$	Mean	2.02509	2.01702	2.02024	2.01308	2.00626	2.00289
	Std. dev.	0.27874	0.19589	0.09995	0.06918	0.04850	0.03028

Table 2. $a(x) = -\arctan x$, $b(x) = 1$

		n					
		50	100	500	1000	2000	5000
$\alpha = 0.1$	Mean	2.69321	2.66637	2.65053	2.66356	2.59903	2.46685
	Std. dev.	2.03142	2.06075	1.82903	1.73034	1.68212	1.50186
$\alpha = 0.5$	Mean	2.12190	2.10459	2.01048	1.99535	2.01712	1.99517
	Std. dev.	0.85304	0.69484	0.48803	0.37807	0.31746	0.25846
$\alpha = 0.9$	Mean	1.95538	1.97446	1.98035	1.99565	2.00266	2.00290
	Std. dev.	0.35057	0.26796	0.12235	0.09050	0.06496	0.04533

Table 3. $a(x) = -\frac{x}{1+x^2}, b(x) = 1$

		n					
		50	100	500	1000	2000	5000
$\alpha = 0.1$	Mean	1.99507	1.99813	1.97122	1.99255	1.98366	1.94811
	Std. dev.	2.44248	2.53060	2.17322	2.13403	2.05527	1.80128
$\alpha = 0.5$	Mean	1.87038	1.87897	1.89022	1.92593	1.94964	1.96624
	Std. dev.	1.01932	0.89315	0.54811	0.49005	0.41787	0.33855
$\alpha = 0.9$	Mean	1.90341	1.92162	2.00240	2.00068	2.00491	1.99347
	Std. dev.	0.47656	0.33693	0.18136	0.13173	0.09595	0.07033

4 Proof of Theorem 2.1

In this section, we prove the main theorem and some auxiliary lemmas. In what follows, $C, C_1, C_2, \ldots$ are positive generic constants that may vary from line to line. If they depend on some arguments, we will write $C(\theta)$, $C(m, \theta)$, and so on.

By (1),

$$
\Delta X_k^n = \theta \int_{\frac{k-1}{n}}^{\frac{k}{n}} a(X_t)\, dt + \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(X_t)\, dW_t
$$

$$
= \theta a(X_{\frac{k-1}{n}}) \frac{1}{n} + \theta \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(a(X_t) - a(X_{\frac{k-1}{n}}) \right) dt + b(X_{\frac{k-1}{n}}) \Delta W_k^n
$$

$$
+ \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(b(X_t) - b(X_{\frac{k-1}{n}}) \right) dW_t.
$$

Therefore,

$$
\hat{\theta}_n = \theta + \sum_{k=1}^{n^{1+\alpha}} \left(c(X_{\frac{k-1}{n}}) \theta \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(a(X_t) - a(X_{\frac{k-1}{n}}) \right) dt + \frac{a(X_{\frac{k-1}{n}})}{b(X_{\frac{k-1}{n}})} \Delta W_k^n \right.
$$

$$
\left. + c(X_{\frac{k-1}{n}}) \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(b(X_t) - b(X_{\frac{k-1}{n}}) \right) dW_t \right) \Big/ \left(\frac{1}{n} \sum_{k=1}^{n^{1+\alpha}} d(X_{\frac{k-1}{n}}) \right).
$$

Then

$$
\hat{\theta}_n - \theta = \frac{n^{-\alpha}(A_n + B_n + E_n)}{D_n}, \tag{7}
$$

where

$$
D_n = n^{-1-\alpha} \sum_{k=1}^{n^{1+\alpha}} d(X_{\frac{k-1}{n}}),
$$

$$
A_n = \sum_{k=1}^{n^{1+\alpha}} c(X_{\frac{k-1}{n}}) \theta \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(a(X_t) - a(X_{\frac{k-1}{n}}) \right) dt,
$$

$$B_n = \sum_{k=1}^{n^{1+\alpha}} \frac{a(X_{\frac{k-1}{n}})}{b(X_{\frac{k-1}{n}})} \Delta W_k^n,$$

$$E_n = \sum_{k=1}^{n^{1+\alpha}} c(X_{\frac{k-1}{n}}) \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(b(X_t) - b(X_{\frac{k-1}{n}}) \right) dW_t.$$

Lemma 4.1. *Let assumptions* (A1)–(A3) *and* (A5) *be fulfilled. Then for every* $m \in \mathbb{N}$, *there exists a constant* $C(m, \theta) > 0$ *such that*

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m} \le C(m, \theta) n^{-m+1+\alpha}.$$

for all $n \in \mathbb{N}$, $1 \le k \le n^{1+\alpha}$, *and* $t \in [\frac{k-1}{n}, \frac{k}{n}]$.

Proof. By (1) and the inequality $(a + b)^{2m} \le 2^{2m-1}(a^{2m} + b^{2m})$,

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m}$$

$$\le 2^{2m-1} \left(\theta^{2m} \mathbf{E}_\theta \left(\int_{\frac{k-1}{n}}^t a(X_s)\, ds \right)^{2m} + \mathbf{E}_\theta \left(\int_{\frac{k-1}{n}}^t b(X_s)\, dW_s \right)^{2m} \right).$$

Using the Burkholder–Davis–Gundy inequality, we obtain

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m}$$

$$\le 2^{2m-1} \left(\theta^{2m} \mathbf{E}_\theta \left(\int_{\frac{k-1}{n}}^t a(X_s)\, ds \right)^{2m} + C(m) \mathbf{E}_\theta \left(\int_{\frac{k-1}{n}}^t b(X_s)^2\, ds \right)^m \right).$$

By Jensen's inequality,

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m} \le 2^{2m-1} \left(\theta^{2m} \left(t - \frac{k-1}{n} \right)^{2m-1} \mathbf{E}_\theta \int_{\frac{k-1}{n}}^t a(X_s)^{2m}\, ds \right.$$

$$\left. + C(m) \left(t - \frac{k-1}{n} \right)^{m-1} \mathbf{E}_\theta \int_{\frac{k-1}{n}}^t b(X_s)^{2m}\, ds \right). \tag{8}$$

Further, we have

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m} \le 2^{2m-1} \left(\theta^{2m} n^{1-2m} \mathbf{E}_\theta \int_0^{n^\alpha} a(X_s)^{2m}\, ds \right.$$

$$\left. + C(m) n^{1-m} \mathbf{E}_\theta \int_0^{n^\alpha} b(X_s)^{2m}\, ds \right).$$

Now it remains to note that by (2) and (5) the integrals $n^{-\alpha} \int_0^{n^\alpha} a(X_s)^{2m}\, ds$ and $n^{-\alpha} \int_0^{n^\alpha} b(X_s)^{2m}\, ds$ have bounded expectations. $\square$

Lemma 4.2. *Under assumption* (A1), *for every* $m \in \mathbb{N}$, *there exists a constant* $C(m, \theta) > 0$ *such that*

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m} \le C(m, \theta) n^{-m} \mathbf{E}_\theta \left(1 + |X_{\frac{k-1}{n}}|^{2m} \right)$$

for all $n \in \mathbb{N}$, $1 \le k \le n^{1+\alpha}$, *and* $t \in [\frac{k-1}{n}, \frac{k}{n}]$.

Proof. By (8),

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m}$$

$$\le C_1(m, \theta) n^{1-m} \left(\mathbf{E}_\theta \int_{\frac{k-1}{n}}^t a(X_s)^{2m}\, ds + \mathbf{E}_\theta \int_{\frac{k-1}{n}}^t b(X_s)^{2m}\, ds \right).$$

Using assumption (A1) and (2), we get

$$\mathbf{E}_\theta \int_{\frac{k-1}{n}}^t a(X_s)^{2m}\, ds \le 2^{2m-1} \mathbf{E}_\theta \int_{\frac{k-1}{n}}^t \left(\left(a(X_s) - a(X_{\frac{k-1}{n}}) \right)^{2m} + a(X_{\frac{k-1}{n}})^{2m} \right) ds$$

$$\le 2^{2m-1} L \mathbf{E}_\theta \int_{\frac{k-1}{n}}^t (X_s - X_{\frac{k-1}{n}})^{2m}\, ds$$

$$+ 2^{2m-1} M \left(t - \frac{k-1}{n} \right) \mathbf{E}_\theta \left(1 + |X_{\frac{k-1}{n}}|^{2m} \right).$$

The same estimate holds for $\mathbf{E}_\theta \int_{\frac{k-1}{n}}^t b(X_s)^{2m}\, ds$. Therefore,

$$\mathbf{E}_\theta (X_t - X_{\frac{k-1}{n}})^{2m} \le C_2(m, \theta) n^{1-m} \int_{\frac{k-1}{n}}^t \mathbf{E}_\theta (X_s - X_{\frac{k-1}{n}})^{2m}\, ds$$

$$+ C_2(m, \theta) n^{-m} \mathbf{E}_\theta \left(1 + |X_{\frac{k-1}{n}}|^{2m} \right),$$

and the result follows from the Gronwall lemma. $\qquad\square$

Lemma 4.3. *Assume that conditions (A1)–(A3) and (A5) are fulfilled. Then for any $m \ge 1$, $1 \le i \le 2m$, and $0 \le j \le 2m$, there exists $C(m, \theta) > 0$ such that*

$$\sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \mathbf{E}_\theta \left(|X_{\frac{k-1}{n}} - X_t|^i |X_t|^j \right) dt \le C(m, \theta) n^{\alpha - \frac{i(m-\alpha)}{2m+2}}.$$

Proof. Applying the Hölder inequality and Lemma 4.1, we get

$$\mathbf{E}_\theta \left(|X_{\frac{k-1}{n}} - X_t|^i |X_t|^j \right) \le \left(\mathbf{E}_\theta |X_{\frac{k-1}{n}} - X_t|^{2m+2} \right)^{\frac{i}{2m+2}} \left(\mathbf{E}_\theta |X_t|^{\frac{j(2m+2)}{2m+2-i}} \right)^{\frac{2m+2-i}{2m+2}}$$

$$\le C_1(m, \theta) n^{-\frac{(m-\alpha)i}{2m+2}} \left(\mathbf{E}_\theta |X_t|^{\frac{j(2m+2)}{2m+2-i}} \right)^{\frac{2m+2-i}{2m+2}}.$$

Then

$$\sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \mathbf{E}_\theta \left(|X_{\frac{k-1}{n}} - X_t|^i |X_t|^j \right) dt$$

$$\le C_1(m, \theta) n^{-\frac{(m-\alpha)i}{2m+2}} \int_0^{n^\alpha} \left(\mathbf{E}_\theta |X_t|^{\frac{j(2m+2)}{2m+2-i}} \right)^{\frac{2m+2-i}{2m+2}} dt.$$

By Jensen's inequality we have

$$\int_0^{n^\alpha} \left(\mathbf{E}_\theta |X_t|^{\frac{j(2m+2)}{2m+2-i}} \right)^{\frac{2m+2-i}{2m+2}} dt \le n^\alpha \left(n^{-\alpha} \int_0^{n^\alpha} \mathbf{E}_\theta |X_t|^{\frac{j(2m+2)}{2m+2-i}} dt \right)^{\frac{2m+2}{2m+2-i}}.$$

By (5) the expression in brackets is bounded. This completes the proof. $\qquad\square$

Lemma 4.4. *Assume that conditions* (A1)–(A3) *and* (A5) *are fulfilled. Then*

(i) *for any* $m = 0, 1, 2, \ldots,$

$$n^{-1-\alpha} \sum_{k=1}^{n^{1+\alpha}} X_{\frac{k-1}{n}}^{2m} \xrightarrow{L_1} \mathbf{E}_\theta \xi_\theta^{2m} \quad as\ n \to \infty;$$

(ii) *if, additionally,* (A4) *holds, then*

$$D_n \xrightarrow{L_1} \mathbf{E}_\theta d(\xi_\theta) \quad as\ n \to \infty.$$

Proof. (i) In the case $m = 0$, the result is trivial. Let $m \geq 1$. By (5) we have

$$n^{-\alpha} \int_0^{n^\alpha} X_t^{2m}\, dt \xrightarrow{L_1} \mathbf{E}_\theta \xi_\theta^{2m} \quad as\ n \to \infty.$$

Hence, it suffices to prove that

$$F_n := \mathbf{E}_\theta \left| n^{-\alpha} \int_0^{n^\alpha} X_t^{2m}\, dt - n^{-1-\alpha} \sum_{k=1}^{n^{1+\alpha}} X_{\frac{k-1}{n}}^{2m} \right|$$

$$= n^{-\alpha} \mathbf{E}_\theta \left| \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(X_t^{2m} - X_{\frac{k-1}{n}}^{2m} \right) dt \right|$$

converges to zero as $n \to \infty$. By the inequality $|x| \leq |x - y| + |y|$,

$$\left| x^{2m} - y^{2m} \right| \leq |x - y| \sum_{i=0}^{2m-1} |x|^i |y|^{2m-1-i} \leq \sum_{i=1}^{2m} C_i |x - y|^i |y|^{2m-i}. \qquad (9)$$

Therefore,

$$F_n \leq \sum_{i=1}^{2m} C_i n^{-\alpha} \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \mathbf{E}_\theta \left(|X_{\frac{k-1}{n}} - X_t|^i |X_t|^{2m-i} \right) dt,$$

and, by Lemma 4.3,

$$F_n \leq C(m, \theta) \sum_{i=1}^{2m} C_i n^{-\frac{i(m-\alpha)}{2m+2}} \to 0 \quad as\ n \to \infty.$$

(ii) For arbitrary x and y,

$$d(x) - d(y) = \frac{a(x)^2}{b(x)^2} - \frac{a(y)^2}{b(y)^2}$$

$$= \left(a(x) - a(y) \right) \left(\frac{a(x)}{b(x)^2} + \frac{a(y)}{b(x)b(y)} \right)$$

$$- \left(b(x) - b(y) \right) \left(\frac{a(x)a(y)}{b(x)^2 b(y)} + \frac{a(x)a(y)}{b(x)b(y)^2} \right).$$

By (A1), (A4), and (2),

$$\left| d(x) - d(y) \right| \le C|x - y|\left(1 + |x|^{2p+1} + \left(1 + |x|^p\right)\left(1 + |y|^{p+1}\right)\right.$$
$$\left. + \left(1 + |x|^{2p+1}\right)\left(1 + |y|^{p+1}\right) + \left(1 + |x|^{p+1}\right)\left(1 + |y|^{2p+1}\right)\right).$$
(10)

The rest of the proof can be done similarly to part (i) using estimate (10) instead of (9). □

Lemma 4.5. *Under the assumptions of Theorem 2.1,*

(i) $n^{-\alpha/2}|A_n| \xrightarrow{\mathbf{P}_\theta} 0$ *as* $n \to \infty$,

(ii) $n^{-\alpha/2}|E_n| \xrightarrow{\mathbf{P}_\theta} 0$ *as* $n \to \infty$.

Proof. (i) By the Cauchy–Schwarz inequality we have

$$\mathbf{E}_\theta|A_n| \le |\theta| \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \mathbf{E}_\theta \left| c(X_{\frac{k-1}{n}})\left(a(X_t) - a(X_{\frac{k-1}{n}})\right) \right| dt$$

$$\le |\theta| \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\mathbf{E}_\theta c(X_{\frac{k-1}{n}})^2\right)^{\frac{1}{2}} \left(\mathbf{E}_\theta \left(a(X_t) - a(X_{\frac{k-1}{n}})\right)^2\right)^{\frac{1}{2}} dt.$$

Using (A1), (3), and Lemma 4.2, we get

$$\mathbf{E}_\theta|A_n| \le C_1|\theta| \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\mathbf{E}_\theta\left(1 + |X_{\frac{k-1}{n}}|^{4p+2}\right)\right)^{\frac{1}{2}} \left(\mathbf{E}_\theta(X_t - X_{\frac{k-1}{n}})^2\right)^{\frac{1}{2}} dt$$

$$\le C_2(\theta)n^{-1/2} \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\mathbf{E}_\theta\left(1 + |X_{\frac{k-1}{n}}|^{4p+4}\right)\right)^{1/2} dt$$

$$= C_2(\theta)n^{-3/2} \sum_{k=1}^{n^{1+\alpha}} \left(\mathbf{E}_\theta\left(1 + |X_{\frac{k-1}{n}}|^{4p+4}\right)\right)^{1/2}$$

$$\le C_2(\theta)n^{\alpha-1/2} \left(n^{-1-\alpha} \sum_{k=1}^{n^{1+\alpha}} \mathbf{E}_\theta\left(1 + |X_{\frac{k-1}{n}}|^{4p+4}\right)\right)^{1/2}.$$

By Lemma 4.4 the expression $n^{-1-\alpha} \sum_{k=1}^{n^{1+\alpha}} \mathbf{E}_\theta(1 + |X_{\frac{k-1}{n}}|^{4p+4})$ is bounded. Therefore,

$$n^{-\alpha/2}\mathbf{E}_\theta|A_n| \le C_3(\theta)n^{\frac{1}{2}(\alpha-1)} \to 0 \quad \text{as } n \to \infty.$$

(ii) We have

$$E_n = \int_0^{n^\alpha} h(t, \omega) \, dW_t,$$

where

$$h(t, \omega) = \sum_{k=1}^{n^{1+\alpha}} c(X_{\frac{k-1}{n}})\left(b(X_t) - b(X_{\frac{k-1}{n}})\right) \mathbb{1}_{[\frac{k-1}{n}, \frac{k}{n})}(t).$$

Then

$$\mathbf{E}_\theta E_n^2 = \mathbf{E}_\theta \int_0^{n^\alpha} h(t,\omega)^2 \, dt = \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \mathbf{E}_\theta \big(c(X_{\frac{k-1}{n}})^2 \big(b(X_t) - b(X_{\frac{k-1}{n}}) \big)^2 \big) \, dt.$$

Similarly to (i), we can estimate $\mathbf{E}_\theta E_n^2 \le C_4(\theta) n^{\alpha-1}$. Therefore, $n^{-\alpha/2}|E_n| \xrightarrow{L_2} 0$ as $n \to \infty$. $\qquad\square$

Lemma 4.6. *Under the assumptions of Theorem 2.1,*

    (i) $n^{-\alpha} B_n \xrightarrow{\mathbf{P}_\theta} 0$ *as $n \to \infty$,*

    (ii) $n^{-\alpha/2} B_n \Rightarrow N(0, \mathbf{E}_\theta d(\xi_\theta))$ *as $n \to \infty$.*

Proof. (i) Let us prove the convergence in L_2. We have

$$B_n = B_n' + B_n'',$$

where

$$B_n' = \int_0^{n^\alpha} \frac{a(X_t)}{b(X_t)} \, dW_t, \qquad B_n'' = \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\frac{a(X_{\frac{k-1}{n}})}{b(X_{\frac{k-1}{n}})} - \frac{a(X_t)}{b(X_t)} \right) dW_t.$$

Then by (5) we have

$$\mathbf{E}_\theta \big(n^{-\alpha/2} B_n' \big)^2 = n^{-\alpha} \mathbf{E}_\theta \int_0^{n^\alpha} \frac{a(X_t)^2}{b(X_t)^2} \, dt = n^{-\alpha} \mathbf{E}_\theta \int_0^{n^\alpha} d(X_t) \, dt \to \mathbf{E}_\theta d(\xi_\theta) \quad (11)$$

as $n \to \infty$. Hence, $n^{-\alpha} B_n' \xrightarrow{L_2} 0$ as $n \to \infty$.

Arguing as in the proof of Lemma 4.5 (ii), we obtain

$$\mathbf{E}_\theta \big(B_n'' \big)^2 = \sum_{k=1}^{n^{1+\alpha}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\frac{a(X_{\frac{k-1}{n}})}{b(X_{\frac{k-1}{n}})} - \frac{a(X_t)}{b(X_t)} \right)^2 dt.$$

Further,

$$\frac{a(x)}{b(x)} - \frac{a(y)}{b(y)} = \frac{a(x)}{b(x)b(y)} \big(b(y) - b(x) \big) + \frac{1}{b(y)} \big(a(x) - a(y) \big).$$

Therefore, by (2) and assumption (A4),

$$\frac{a(x)}{b(x)} - \frac{a(y)}{b(y)} \le \frac{2a(x)^2}{b(x)^2 b(y)^2} \big(b(y) - b(x) \big)^2 + \frac{2}{b(y)^2} \big(a(x) - a(y) \big)^2$$

$$\le C(x-y)^2 \big(1 + |x|^{2p+2} |y|^{2p} + |y|^{2p} \big).$$

Similarly to the proof of Lemma 4.4, we get the convergence

$$n^{-\alpha} \mathbf{E}_\theta \big(B_n'' \big)^2 \to 0 \quad \text{as } n \to \infty. \tag{12}$$

(ii) According to [7, Theorem 1.19], it follows from (11) that $n^{-\alpha/2} B_n' \Rightarrow N(0, \mathbf{E}_\theta d(\xi_\theta))$ as $n \to \infty$. Taking into account the convergence (12), we obtain the result. $\qquad\square$

Now the statement of Theorem 2.1 follows from (7) and Lemmas 4.4–4.6.

Acknowledgments

References

[1] Borodin, A.N., Salminen, P.: Handbook of Brownian Motion: Facts and Formulae, 2nd edn. Birkhäuser, Basel (2002) MR1912205. doi:10.1007/978-3-0348-8163-0

[2] Dacunha-Castelle, D., Florens-Zmirou, D.: Estimation of the coefficients of a diffusion from discrete observations. Stochastics **19**, 263–284 (1986) MR0872464. doi:10.1080/17442508608833428

[3] Florens-Zmirou, D.: Approximate discrete-time schemes for statistics of diffusion processes. Statistics **20**(4), 547–557 (1989) MR1047222. doi:10.1080/02331888908802205

[4] Heyde, C.C.: Quasi-Likelihood and Its Application. A General Approach to Optimal Parameter Estimation. Springer, New York, NY (1997) MR1461808. doi:10.1007/b98823

[5] Iacus, S.M.: Simulation and Inference for Stochastic Differential Equations. With R Examples. Springer, New York, NY (2008) MR2410254. doi:10.1007/978-0-387-75839-8

[6] Kessler, M., Lindner, A., Sørensen, M. (eds.): Statistical Methods for Stochastic Differential Equations. Selected Papers Based on the Presentations at the 7th Séminaire Européen de Statistiques on "Statistics for Stochastic Differential Equations Models," La Manga del Mar Menor, Cartagena, Spain, 7–12 May 2007. CRC Press, Boca Raton, FL (2012) MR2975799

[7] Kutoyants, Y.A.: Statistical Inference for Ergodic Diffusion Processes. Springer, London (2004) MR2144185. doi:10.1007/978-1-4471-3866-2

[8] Liptser, R.S., Shiryayev, A.N.: Statistics of Random Processes. II. Applications. Springer-Verlag, New York, Heidelberg, Berlin (1978) MR0488267

[9] Mishura, Y.: Standard maximum likelihood drift parameter estimator in the homogeneous diffusion model is always strongly consistent. Stat. Probab. Lett. **86**, 24–29 (2014) MR3162713. doi:10.1016/j.spl.2013.12.004

[10] Prakasa Rao, B.L.S.: Asymptotic Theory of Statistical Inference. John Wiley & Sons, New York (1987) MR0874342

[11] Prakasa Rao, B.L.S.: Statistical inference from sampled data for stochastic processes. Contemp. Math. **20**, 249–284 (1988) MR0999016. doi:10.1090/conm/080/999016

[12] Sørensen, M.: Parametric inference for discretely sampled stochastic differential equations. In: Handbook of Financial Time Series. With a Foreword by Robert Engle, pp. 531–553. Springer, Berlin (2009)

[13] Sørensen, H.: Estimation of diffusion parameters for discretely observed diffusion processes. Bernoulli **8**(4), 491–508 (2002) MR1914700

[14] Yoshida, N.: Estimation for diffusion processes from discrete observation. J. Multivar. Anal. **41**(2), 220–242 (1992) MR1172898. doi:10.1016/0047-259X(92)90068-Q