Advances in Difference Equations
a SpringerOpen Journal

Open Access

CrossMark

On drift parameter estimation for mean-reversion type stochastic differential equations with discrete observations

Jingjie Li[1*] and Jiang-Lun Wu[2]

*Correspondence: lxyljj@tjcu.edu.cn
[1] School of Science, Tianjin University of Commerce, Tianjin, 300134, China
Full list of author information is available at the end of the article

Abstract

We are concerned with a parameter estimation for mean-reversion type stochastic differential equations (SDEs) driven by Brownian motion. The equations, involving a small dispersion parameter, are observed at discrete (regularly spaced) time instants. The least square method is utilized to derive an asymptotically consistent estimator. We then discuss the rate of convergence of the least square estimator. The new feature of our study in this paper is that, due to the mean-reversion type drift coefficient in the SDEs, we have to use the Girsanov transformation to simplify the equations, which then gives rise to the corresponding convergence of the least square estimator being with respect to a family of probability measures indexed by the dispersion parameter, while in the literature the existing results have dealt with convergence with respect to a given probability measure.

MSC: 60H10; 62F12; 62M05

Keywords: mean-reversion type SDEs; Girsanov transformation; least square estimator (LSE); discrete observation; consistency of least square estimator; asymptotic distribution of LSE

1 Introduction

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space endowed with a usual filtration $\{\mathcal{F}_t\}_{t \geq 0}$, *i.e.*, $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for $0 \leq s \leq t \leq 1$ and $\mathcal{F}_0$ contains all null set of P. We are interested in the unique solution, denoted by $X = (X_t)_{0 \leq t \leq 1}$, of the following stochastic differential equations (SDEs) of mean-reversion type:

$$dX_t = \big[r + \alpha(X_t, t, \varepsilon)\big] b(X_t, t)\, dt + \varepsilon \sigma(X_t, t)\, dB_t, \quad 0 \leq t \leq 1 \tag{1.1}$$

with initial value $X_0 = x \in \mathbb{R}$, where r is a constant (which is supposed to be unknown), $\varepsilon \in (0, 1]$ is a parameter describing the smallness of dispersion, $\alpha : (x, t, \varepsilon) \in \mathbb{R} \times [0, 1] \times (0, 1] \mapsto \alpha(x, t, \varepsilon) \in \mathbb{R}$ is twice differentiable with respect to x and differentiable with respect to t, both $b : (x, t) \in \mathbb{R} \times [0, 1] \mapsto b(x, t) \in \mathbb{R}$ and $\sigma : (x, t) \in \mathbb{R} \times [0, 1] \mapsto \sigma(x, t) \in \mathbb{R} \setminus \{0\}$ are continuous with respect to t, and $(B_t)_{0 \leq t \leq 1}$ is a one dimensional $\{\mathcal{F}_t\}$-Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{P}, \{\mathcal{F}_t\}_{0 \leq t \leq 1})$.

The above mean-reversion type SDE (with $\varepsilon = 1$) was used widely in modeling price dynamics in mathematical studies of finance and economics. One feature of such a kind of

Springer

modeling is that the function α in the drift coefficient satisfies a typical nonlinear parabolic partial differential equation - the Burgers equation; see, *e.g.*, [1] and references therein.

In our present paper, we are interested in the discrete-time system of (1.1) starting with the initial $X_0 = x$. That is, for any fixed $n \in \mathbb{N}$ and for any given partition $0 = t_0 < t_1 < \cdots < t_{i-1} < t_i < \cdots < t_n = 1$, we define X_{t_i}, $i = 1, 2, \ldots, n$, via the Euler-Maruyama numerical scheme

$$X_{t_k} = x + \sum_{i=1}^{k} \big[r + \alpha(X_{t_{i-1}}, t_{i-1}, \varepsilon) \big] b(X_{t_{i-1}}, t_{i-1}) \Delta t_i + \varepsilon \sum_{i=1}^{k} \sigma(X_{t_{i-1}}, t_{i-1})(B_{t_i} - B_{t_{i-1}})$$

for $1 \leq k \leq n$, where $\Delta t_i = t_i - t_{i-1}$, $i = 1, \ldots, n$.

Before we proceed further, let us give a brief overview of the topic considered in the present paper. Historically, the theory of SDEs has been very well developed since the seminal work of the great Japanese mathematician Kiyosi Itô in the mid 1940s. Since then, SDEs had profound impacts on many areas in science and technology, *cf. e.g.* the excellent textbook [2] (and references therein). Nowadays, the theory of SDEs has played an important role in modeling uncertain and volatile systems arising in many diverse subjects such as economics and finance, biology, chemistry, ecology, physics, and so on. Fundamental issues are to estimate certain parameters (*i.e.*, deterministic quantities) appearing in the random models by certain observations (or by experimental data). Estimating the drift parameter of SDEs is one of such important topics. In the past decades, there were many papers devoted to the drift parameter estimation for SDEs and there are two main methods; see for instance Prakasa Rao [3], Liptser and Shiryaev [4], Kutoyants [5] for the maximum likelihood estimator (MLE) method, and Dorogovcev [6], Le Breton [7], Kasonga [8] for the least square estimator (LSE) method. It turns out that both the MLE and the LSE are asymptotically equivalent and the LSE enjoys the strong consistency property under some regularity conditions. Moreover, Prakasa Rao [9] gave a study of the asymptotic distribution. Further, Shimizu and Yoshida [10] considered a multidimensional diffusion process with jumps whose jump term is driven by a compound Poisson process. They let $\alpha(x, \theta)$ be the drift coefficient and $b(x, \sigma)$ be the diffusion coefficient and study estimation of the parameter $\alpha(x, \theta) = (\theta, \sigma)$. Under certain assumptions, the consistency and asymptotic normality of an estimator were shown. Shimizu [11] considered a similar case and proposed an estimating function with more complicated situation.

On the other hand, based on continuous time observations or discrete-time observations, there are correspondingly two kinds of drift parameter estimates. A parameter estimation for diffusion processes based on continuous-time observation can be found in, *e.g.*, Kutoyants [12], Kutoyants [13], Uchida and Yoshida [14], and Yoshida [15, 16]), in particular, Uchida and Yoshida [14] considered the evaluation problem of statistical models for diffusion processes driven by a small noise. As for a parameter estimation based on discrete-time observations, Sørensen [17] gave an excellent survey of existing estimation techniques for stationary and ergodic diffusion processes observed at discrete points in time. It is more realistic and interesting to consider the parameter estimation for diffusion processes based on discrete observation, since the actual data may be obtained discretely. We begin with our study from this point of view.

Recently, the parameter estimation for mean-reversion type SDEs has received a lot of attention. Long [18] gave an investigation of the parameter estimation for discretely observations on one dimensional Ornstein-Uhlenbeck (O-U) processes driven by small Lévy

noise. One assumed that the drift function $b(x, \theta) = -\theta x$ is linear for both x and θ. Meanwhile the driving Lévy process was $L_t = aB_t + bZ_t$, where a and b are known constants, $\{B_t, t \geq 0\}$ is a standard Brownian motion and Z_t is a α-stable Lévy motion independent of $\{B_t, t \geq 0\}$. In this framework, the author established the consistency and asymptotic normality for the proposed estimators. Long [19] investigated the parameter estimation problem for discrete observations driven by a small Levy noises and further gave a discussion of the case of the drift function $b(x, \theta) = \theta b(x)$. Under some regularity conditions, the author obtained the consistency and rate of convergence of the least squares estimator when the small dispersion parameter $\varepsilon \to 0$ and $n \to \infty$ simultaneously. In a similar framework, Ma [20] extended the results of Long [19] to the case where the driving noise was a general Lévy process. After that, by Hu and Long [21] a least square estimator for Ornstein-Uhlenbeck processes driven by α-stable motions was studied. The main focus of Hu and Long [21] is to study the strong consistency and asymptotic distributions of the least square estimator for generalized O-U processes. After obtaining a least square estimator, the authors proved the strong consistency result and obtained further the rate of convergence of the estimator and an asymptotic distribution. Hu and Long [22] extended the results of [21] to the case that the drift function $b(x, \theta) = \alpha_0 - \theta_0 x$. When $\alpha_0 = 0$, the mean-reverting α-stable motion becomes an O-U process. Under certain conditions, by using a least square method, they showed the consistency property and obtained the asymptotic distribution.

There are many applications of small noise asymptotic to mathematical finance; see, *e.g.*, Kunitomo and Takahashi [23], Long [19], Takahashi [24], Takahashi and Yoshida [25], Uchida and Yoshida [26], and Yoshida [27]. We mention particularly Kunitomo and Takahashi [23] who proposed a new methodology for the valuation problem of financial contingent claims when the underlying asset prices follow a general class of continuous Itô processes. Furthermore, the authors gave two interesting examples of the valuation problems of average options for interest rates.

In the present paper, we consider a fairly general class of stochastic processes solving SDEs of mean-reversion type (1.1). We aim to investigate the least square estimator for the true value of r based on the (discrete) sampling data $(X_{t_i})_{i=1}^n$. Comparing to the existing studies in the literature, a major difficulty in our case is the appearance of the item $\alpha(X_t, t, \varepsilon)$ in the drift coefficient of (1.1). The feature of our study is the utilizing of the Girsanov transformation, which makes it possible to simplify the drift coefficient, therefore the equation, under a family of probability measures indexed by the small dispersion parameter $\varepsilon \in (0, 1]$. As such, our consideration creates a novel point: all involved convergences are with respect to the ε indexed family of probability measures which, to our knowledge, has not been considered in the literature. For the Girsanov transformation for SDEs, the reader is referred to Øksendal [2]. In our case here, we apply the Girsanov transformation to get rid of the term $\alpha(X_t, t, \varepsilon)$, which changes the original probability measure P to a family of (equivalent) probability measures Q_ε. We then derive explicitly the least square estimators which also depend on ε and therefore indexed by ε for the simplified SDE. From this we can prove, under certain conditions, the convergence of the least square estimators under the family Q_ε to a limit which turns out to be the true value of the parameter r. We end up by studying the convergence rate and derive further the asymptotic distribution under a new equivalent probability measure.

The paper is organized as follows. In the next section, we give some preliminaries, starting with the discretization of our equation (1.1) is presented, which will be used in later derivations and proofs. In Section 3, we focus on showing the convergence of the estimators with high frequency and small dispersion simultaneously. Section 4 is devoted to establishing the rate of the relevant convergence and the associated asymptotic distribution. To finish, the last section draws a conclusion.

2 Preliminaries and auxiliary results

In this section, we start with an introduction of the SDEs of mean-reversion type and then we impose some assumptions on our SDE (1.1) to ensure the existence and uniqueness of the solution. Furthermore, we introduce the Girsanov transformation to simplify our equation (1.1). On the other hand, in order to obtain the consistency and asymptotics of our LSE $\hat{r}_{n,\varepsilon}$, an explicit form of $\hat{r}_{n,\varepsilon}$ is defined. In addition, we would like to present some notations and preliminaries which will be needed in the latter sections of the paper. Throughout the paper, we use notation '$\to_Q$' to denote 'convergence in probability Q'; notation '$\to_P$' to denote 'convergence in probability P' and notation '$\Rightarrow$' to denote 'convergence in distribution'.

We assume the coefficients of (1.1) fulfill the following conditions throughout the paper:

(C.1) $\exists$ a constant $L > 0$ such that $|b(x,t) - b(y,t)| \leq L|x - y|$, for $x, y \in \mathbb{R}$, $t \in [0,1]$;

(C.2) $\exists$ a constant $L'' > 0$ such that, for $x, y \in \mathbb{R}$, $t \in [0,1]$ and $\varepsilon \in (0,1]$,

$$\left| \alpha(x,t,\varepsilon)b(x,t) - \alpha(y,t,\varepsilon)b(y,t) \right| \leq L''|x - y|;$$

(C.3) $\exists$ a constant $L' > 0$ such that $|\sigma(x,t) - \sigma(y,t)| \leq L'|x - y|$, for $x, y \in \mathbb{R}$, $t \in [0,1]$;

(C.4) $\exists$ constants $K' > 0$, $m > 0$ such that $\sigma^{-2}(x,t) \leq K'(1 + |x|^m)$, for $x \in \mathbb{R}$, $t \in [0,1]$.

The above conditions guarantee (see, *e.g.*, [1]) the existence of a unique solution of (1.1) for a given initial data $X_0 = x \in \mathbb{R}$. The celebrated Girsanov transformation (also called the transformation of the drift) provides a very useful and efficient approach to solve (1.1). The transformation says the following.

Let $u : \mathbb{R} \times [0,1] \to \mathbb{R}$ satisfy the following Novikov condition:

$$\mathbb{E}\left[\exp\left(\frac{1}{2} \int_0^t \left| u(X_s, s) \right|^2 ds \right) \right] < \infty, \quad \forall t \in [0,1].$$

Then, by the Girsanov theorem (*cf. e.g.* Theorem IV in Section 4.1 of [28]),

$$\exp\left(\int_0^t u(X_s, s)\, dB_s - \frac{1}{2} \int_0^t \left[u(X_s, s) \right]^2 ds \right), \quad t \in [0,1],$$

is an $\{\mathcal{F}_t\}_{t \in [0,1]}$-martingale. Furthermore, for $t \in [0,1]0$, we define

$$Q_t := \exp\left(\int_0^t u(X_s, s)\, dB_s - \frac{1}{2} \int_0^t \left[u(X_s, s) \right]^2 ds \right) \cdot P$$

or equivalently, in terms of the Radon-Nikodym derivative,

$$\frac{dQ_t}{dP} = \exp\left(\int_0^t u(s, X_s)\, dB_s - \frac{1}{2} \int_0^t \left[u(X_s, s) \right]^2 ds \right).$$

Then taking Q_1 in particular, we see that

$$\tilde{B}_t := B_t - \int_0^t u(s, X_s)\, ds, \quad 0 \le t \le 1,$$

is an $\{\mathcal{F}_t\}_{t \in [0,1]}$-Brownian motion under the probability Q_1. Moreover, the solution process X_t of (1.1) solves the following SDE:

$$dX_t = \left\{ \left[\left[r + \alpha(X_t, t, \varepsilon) \right] b(X_t, t) + \varepsilon \sigma(X_t, t) u(t, X_t) \right] \right\} dt + \sigma(X_t, t)\, d\tilde{B}_t, \quad t \in [0,1]. \quad (2.1)$$

In order to get rid of the term $\alpha(X_t, t, \varepsilon)$ for $\varepsilon \in (0,1]$, we can specify u such that

$$\alpha(x, t, \varepsilon) b(x, t) + \varepsilon \sigma(x, t) u(x, t) = 0,$$

and hence

$$\alpha(X_t, t, \varepsilon) b(X_t, t) + \varepsilon \sigma(X_t, t) u(X_t, t) = 0$$

and our equation (2.1) becomes

$$dX_t = r b(X_t, t)\, dt + \varepsilon \sigma(X_t, t)\, d\hat{B}_t^\varepsilon. \quad (2.2)$$

Thus, we set

$$u_\varepsilon(x, t) := \frac{\alpha(x, t, \varepsilon) b(x, t)}{\varepsilon \sigma(x, t)}$$

so that

$$u_\varepsilon(X_t, t) = \frac{\alpha(X_t, t, \varepsilon) b(X_t, t)}{\varepsilon \sigma(X_t, t)}. \quad (2.3)$$

The Novikov condition for this u_ε is then the following:

$$E\left[\exp\left(\frac{1}{2} \int_0^t \left| \frac{\alpha(X_s, s, \varepsilon) b(X_s, s)}{\varepsilon \sigma(X_s, s)} \right|^2 ds \right) \right] < \infty, \quad t \in [0,1], \quad (2.4)$$

which is needed to put on the coefficients of the original equation (1.1). Then we define

$$M_t^\varepsilon = \exp\left(-\int_0^t u_\varepsilon(X_s, s)\, dB_s - \frac{1}{2} \int_0^t u_\varepsilon^2(X_s, s)\, ds \right), \quad t \in [0,1], \quad (2.5)$$

where M_t^ε is an $\{\mathcal{F}_t\}_{t \in [0,1]}$-martingale. For each $\varepsilon \in (0,1]$, let Q_ε be a probability measure on $\mathcal{F}_1$, satisfying

$$dQ_\varepsilon := M_1^\varepsilon\, dP. \quad (2.6)$$

Then we define

$$\hat{B}_t^\varepsilon := \int_0^t u_\varepsilon(X_s, s)\, ds + B_t, \quad (2.7)$$

where $\hat{B}_t^\varepsilon$ is an $\{\mathcal{F}_t\}_{t\in[0,1]}$-Brownian motion with respect to the probability measure Q_ε. Then we arrive at equation (2.2) for X_t, that is,

$$dX_t = rb(X_t, t)\,dt + \varepsilon\sigma(X_t, t)\,d\hat{B}_t^\varepsilon, \quad t \in [0,1],$$

which, from now, we will focus on. We denote the true value of the parameter r by r_0 and the least square estimator of r by $\hat{r}$. As mentioned before we focus on the investigation of the least square estimator for the true value r_0 based on the (discrete) sampling data $(X_{t_i})_{i=1}^n$ obtained by the Euler-Maruyama numerical scheme for the Cauchy problem of equation (2.2) with initial $X_0 = x$. For simplicity, we assume that the process X_t is observed at regularly spaced time points $0 = t_0 < t_1 < \cdots < t_{i-1} < t_i < \cdots < t_n = 1$ with $\{t_i = \frac{i}{n}, i = 0,1,2,\ldots,n\}$, where $n \in \mathbb{N}$ is arbitrarily fixed. That is,

$$X_{t_i} = x + \sum_{i=1}^n rb(X_{t_{i-1}}, t_{i-1})\Delta t_i + \varepsilon\sum_{i=1}^n \sigma(X_{t_{i-1}}, t_{i-1})\big(\hat{B}_{t_i}^\varepsilon - \hat{B}_{t_{i-1}}^\varepsilon\big).$$

Let us start with the use of the least square method to get a consistent estimator. We first discretize (2.2) as follows:

$$X_{t_i} - X_{t_{i-1}} = rb(X_{t_{i-1}}, t_{i-1})\Delta t_i + \varepsilon\sigma(X_{t_{i-1}}, t_{i-1})\Delta\hat{B}_{t_i}^\varepsilon,$$

where $\Delta t_i = t_i - t_{i-1} = \frac{1}{n}$; $\Delta\hat{B}_{t_i}^\varepsilon = \hat{B}_{t_i}^\varepsilon - \hat{B}_{t_{i-1}}^\varepsilon$ is the increment of Brownian motion. Then

$$\frac{X_{t_i} - X_{t_{i-1}} - rb(X_{t_{i-1}}, t_{i-1})\Delta t_i}{\varepsilon\sigma(X_{t_{i-1}}, t_{i-1})} = \Delta\hat{B}_{t_i}^\varepsilon.$$

Since $\Delta\hat{B}_{t_i}^\varepsilon$ is a normal distribution with zero mean on the new probability space $(\Omega, \mathcal{F}, Q_\varepsilon)$, we obtain the variance of $\Delta\hat{B}_{t_i}^\varepsilon$ and denote it by the following contrast function:

$$\rho_{n,\varepsilon}(r) := \sum_{i=1}^n \left| \frac{X_{t_i} - X_{t_{i-1}} - rb(X_{t_{i-1}}, t_{i-1})\Delta t_i}{\varepsilon\sigma(X_{t_{i-1}}, t_{i-1})} \right|^2.$$

In order to get the least square estimator $\hat{r}_{n,\varepsilon}$, we let

$$\frac{\partial\rho_{n,\varepsilon}(r)}{\partial r} = 0,$$

from which we get the solution, denoted by $\hat{r}_{n,\varepsilon}$, which is given as

$$\hat{r}_{n,\varepsilon} := \frac{\sum_{i=1}^n b(X_{t_{i-1}}, t_{i-1})(X_{t_i} - X_{t_{i-1}})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})}{n^{-1}\sum_{i=1}^n b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})}. \tag{2.8}$$

With all these in hand, we now present several auxiliary results which are needed for our later considerations.

Definition 2.1 Let X_t^0 be the solution of the following ordinary differential equation under the true value of the drift parameter:

$$dX_t^0 = r_0 b\big(X_t^0, t\big)\,dt, \qquad X_0^0 = x_0, \tag{2.9}$$

where r_0 be the real value of r.

The following lemma is a reformulation of Theorem 7.3 of [29] (see also [28] or [2]), which is nothing but the Burkholder-Davis-Gundy inequality in our setting.

Lemma 2.1 *Let $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R})$. Define, for $0 \le t \le T$,*

$$x(t) = \int_0^t g(s)\, dB_s$$

and

$$A(t) = \int_0^t |g(s)|^2\, ds.$$

Then, for every $p > 0$, there exist universal positive constants c_p, C_p (depending only on p), such that

$$c_p E |A(t)|^{\frac{p}{2}} \le E \Big(\sup_{0 \le s \le t} |x(s)|^p \Big) \le C_p E |A(t)|^{\frac{p}{2}}$$

for all $t \ge 0$. In particular, one may take $c_p = (p/2)^p$, $C_p = (32/p)^{p/2}$, if $0 < p < 2$; $c_p = 1$, $C_p = 4$, if $p = 2$; $c_p = (2p)^{-p/2}$, $C_p = [p^{p+1}/2(p-1)^{p-1}]^{p/2}$, if $p > 2$.

Let us next show the following two lemmas which will be used in Section 3 and Section 4.

Lemma 2.2 *Under conditions* (C.1), (C.2), (C.3), *we have*

$$\left| X_t - X_t^0 \right| \le \varepsilon e^{L|r_0|t} \sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s)\, d\hat{B}_t^\varepsilon \right|. \tag{2.10}$$

Proof We have

$$X_t^0 = x_0 + r_0 \int_0^t b(X_s^0, s)\, ds. \tag{2.11}$$

From (2.2) we have

$$X_{t_i} - X_{t_{i-1}} = r_0 \int_{t_{i-1}}^{t_i} b(X_s, s)\, ds + \varepsilon \int_{t_{i-1}}^{t_i} \sigma(X_s, s)\, d\hat{B}_s^\varepsilon. \tag{2.12}$$

Together with (2.11) and (2.12), we obtain

$$X_t - X_t^0 = r_0 \int_0^t \big(b(X_s, s) - b(X_s^0, s) \big)\, ds + \varepsilon \int_0^t \sigma(X_s, s)\, d\hat{B}_s^\varepsilon. \tag{2.13}$$

By condition (C.1), we get

$$\left| X_t - X_t^0 \right| \le L|r_0| \int_0^t \left| X_s - X_s^0 \right| ds + \varepsilon \left| \int_0^t \sigma(X_s, s)\, d\hat{B}_s^\varepsilon \right|. \tag{2.14}$$

By the Gronwall inequality, we get

$$\left| X_t - X_t^0 \right| \le \varepsilon e^{L|r_0|t} \sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s)\, d\hat{B}_s^\varepsilon \right|. \qquad \square$$

Lemma 2.3 *Under conditions* (C.1), (C.2), (C.3), *we have*

$$\sup_{0 \leq t \leq 1} \left| X_t - X_t^0 \right| \to_{Q_\varepsilon} 0 \quad as \; \varepsilon \to 0. \tag{2.15}$$

Proof By Lemma 2.2,

$$\sup_{t \in [0,1]} \left| X_t - X_t^0 \right| \leq \varepsilon e^{L|r_0|} \sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s) \, d\hat{B}_t^\varepsilon \right|. \tag{2.16}$$

Let $\eta > 0$, by Lemma 2.1, the Markov inequality, and condition (C.4), we have

$$Q_\varepsilon \left(\varepsilon e^{L|r_0|} \sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| > \eta \right)$$

$$\leq \eta^{-1} e^{L|r_0|} \varepsilon E_{Q_\varepsilon} \left[\sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| \right]$$

$$\leq 4\sqrt{2} \eta^{-1} e^{L|r_0|} \varepsilon E_{Q_\varepsilon} \left[\left(\int_0^\delta \left| \sigma(X_s, s) \right|^2 ds \right)^{\frac{1}{2}} \right]. \tag{2.17}$$

Since $\sigma(x, t)$ satisfies the Lipschitz condition and is continuous with respect to t, $\sigma(x, t)$ meets the following linear growth condition:

$$\left| \sigma(x, t) \right|^2 \leq K \left(1 + |x|^2 \right),$$

where $K > 0$ is a constant. Then we have

$$Q_\varepsilon \left(\varepsilon e^{L|r_0|} \sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| > \eta \right)$$

$$\leq 4\sqrt{2} \eta^{-1} e^{L|r_0|} \varepsilon \left(\int_0^\delta K \left(1 + E_{Q_\varepsilon} |X_s|^2 \right) ds \right)^{\frac{1}{2}}. \tag{2.18}$$

By the Hölder inequality, the Itô isometry, and the Gronwall inequality, we get $E_{Q_\varepsilon} |X_s|^2 \leq C$, where $C > 0$ is a constant. Then we have

$$Q_\varepsilon \left(\varepsilon e^{L|r_0|} \sup_{\delta \in [0,t]} \left| \int_0^\delta \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| > \eta \right) \leq 4\sqrt{2} \eta^{-1} e^{L|r_0|} \varepsilon \left[K(1 + C)\delta \right]^{\frac{1}{2}}. \tag{2.19}$$

The above equation implies that

$$\sup_{0 \leq t \leq 1} \left| X_t - X_t^0 \right| \to_{Q_\varepsilon} 0 \quad \text{as } \varepsilon \to 0. \qquad \square$$

3 Consistency of the least square estimator

This section is devoted to prove the consistency of the least square estimator $\hat{r}_{n,\varepsilon}$. In the following, we first derive an explicit decomposition for $\hat{r}_{n,\varepsilon}$. Based on this, we will show the strong consistency of the LSE $\hat{r}_{n,\varepsilon}$. Note from (2.2), that we have

$$X_{t_i} - X_{t_{i-1}} = r_0 \int_{t_{i-1}}^{t_i} b(X_s, s) \, ds + \varepsilon \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon. \tag{3.1}$$

Combining the above with (2.8) then yields the following decomposition:

$$
\begin{aligned}
\hat{r}_{n,\varepsilon} &= \frac{\sum_{i=1}^{n} b(X_{t_{i-1}}, t_{i-1})(X_{t_i} - X_{t_{i-1}})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})}{n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})} \\
&= \frac{r_0 \sum_{i=1}^{n} b(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})\int_{t_{i-1}}^{t_i} b(X_s, s)\, ds}{n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})} \\
&\quad + \frac{\varepsilon \sum_{i=1}^{n} b(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})\int_{t_{i-1}}^{t_i} \sigma(X_s, s)\, d\hat{B}_s}{n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})} \\
&= r_0 + \frac{r_0 \sum_{i=1}^{n} b(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})\int_{t_{i-1}}^{t_i} (b(X_s, s) - b(X_{t_{i-1}}, t_{i-1}))\, ds}{n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})} \\
&\quad + \frac{\varepsilon \sum_{i=1}^{n} b(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})\int_{t_{i-1}}^{t_i} \sigma(X_s, s)\, d\hat{B}_s}{n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1})} \\
&=: r_0 + \frac{\phi_2(n,\varepsilon)}{\phi_1(n,\varepsilon)} + \frac{\phi_3(n,\varepsilon)}{\phi_1(n,\varepsilon)}.
\end{aligned}
$$

Our first main result is the following theorem.

Theorem 3.1 *We have $\hat{r}_{n,\varepsilon} \to_{Q_\varepsilon} r_0$, as $n \to \infty$ and $\varepsilon \to 0$ with $\varepsilon n^{\frac{1}{2}} \to 0$.*

To prove Theorem 3.1, the following lemmas are needed. We shall study the asymptotic behavior of $\phi_1(n,\varepsilon)$, $\phi_2(n,\varepsilon)$, and $\phi_2(n,\varepsilon)$, respectively.

Lemma 3.1 *Under conditions* (C.1)-(C.4), *we have*

$$
\phi_1(n,\varepsilon) \to_{Q_\varepsilon} \int_0^t \sigma^{-2}(X_t^0, t)b^2(X_t^0, t)\, dt
$$

as both $n \to \infty$ and $\varepsilon \to 0$.

Proof We first have

$$
\begin{aligned}
\phi_1(n,\varepsilon) &= n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}, t_{i-1})\sigma^{-2}(X_{t_{i-1}}, t_{i-1}) \\
&= n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}^0, t_{i-1})\sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) \\
&\quad + n^{-1}\sum_{i=1}^{n} (b^2(X_{t_{i-1}}, t_{i-1}) - b^2(X_{t_{i-1}}^0, t_{i-1}))\sigma^{-2}(X_{t_{i-1}}, t_{i-1}) \\
&\quad + n^{-1}\sum_{i=1}^{n} b^2(X_{t_{i-1}}^0, t_{i-1})(\sigma^{-2}(X_{t_{i-1}}, t_{i-1}) - \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1})) \\
&=: \phi_{1,1}(n,\varepsilon) + \phi_{1,2}(n,\varepsilon) + \phi_{1,3}(n,\varepsilon).
\end{aligned}
$$

For $\phi_{1,1}(n,\varepsilon)$, according to the definition of Riemann integral, we obtain, as $n \to \infty$,

$$
\phi_{1,1}(n,\varepsilon) \to_{Q_\varepsilon} \int_0^1 b^2(X_s^0, s)\sigma^{-2}(X_s^0, s)\, ds.
$$

For $\phi_{1,2}(n,\varepsilon)$, we have the following derivation:

$$
\begin{aligned}
\big|\phi_{1,2}(n,\varepsilon)\big| &\leq n^{-1}\sum_{i=1}^{n}\big|b^2(X_{t_{i-1}},t_{i-1})-b^2\big(X^0_{t_{i-1}},t_{i-1}\big)\big|\sigma^{-2}(X_{t_{i-1}},t_{i-1}) \\
&\leq n^{-1}\sum_{i=1}^{n}K'\big(1+|X_{t_{i-1}}|^m\big)\big|b(X_{t_{i-1}},t_{i-1}) \\
&\quad -b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|\big|b(X_{t_{i-1}},t_{i-1})+b\big(X^0_{t_{i-1}},t_{i-1}\big)\big| \\
&\leq n^{-1}\sum_{i=1}^{n}K'\big(1+|X_{t_{i-1}}|^m\big)\Big(\big|b(X_{t_{i-1}},t_{i-1})-b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|^2 \\
&\quad +2\big|b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|\big|b(X_{t_{i-1}},t_{i-1})-b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|\Big) \\
&\leq n^{-1}K'\sum_{i=1}^{n}\big|b(X_{t_{i-1}},t_{i-1})-b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|^2 \\
&\quad +2n^{-1}K'\sum_{i=1}^{n}\big|b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|\big|b(X_{t_{i-1}},t_{i-1})-b\big(X^0_{t_{i-1}},t_{i-1}\big)\big| \\
&\quad +n^{-1}K'\sum_{i=1}^{n}|X_{t_{i-1}}|^m\big|b(X_{t_{i-1}},t_{i-1})-b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|^2 \\
&\quad +2n^{-1}K'\sum_{i=1}^{n}\big|b\big(X^0_{t_{i-1}},t_{i-1}\big)\big||X_{t_{i-1}}|^m\big|b(X_{t_{i-1}},t_{i-1})-b\big(X^0_{t_{i-1}},t_{i-1}\big)\big| \\
&:=\phi^1_{1,2}(n,\varepsilon)+\phi^2_{1,2}(n,\varepsilon)+\phi^3_{1,2}(n,\varepsilon)+\phi^4_{1,2}(n,\varepsilon).
\end{aligned}
$$

By condition (C.1), we have

$$
\phi^1_{1,2}(n,\varepsilon)\leq n^{-1}L^2K'\sum_{i=1}^{n}\big|X_{t_{i-1}}-X^0_{t_{i-1}}\big|^2\leq L^2K'\Big(\sup_{0\leq t\leq 1}\big|X_t-X^0_t\big|\Big)^2.
$$

By Lemma 2.3, we get $\phi^1_{1,2}(n,\varepsilon)\to_{Q_\varepsilon}0$ as $\varepsilon\to 0$. Next, we have

$$
\begin{aligned}
\phi^2_{1,2}(n,\varepsilon) &\leq 2n^{-1}LK'\sum_{i=1}^{n}\big|b\big(X^0_{t_{i-1}},t_{i-1}\big)\big|\big|X_{t_{i-1}}-X^0_{t_{i-1}}\big| \\
&\leq 2LK'\sup_{0\leq t\leq 1}\big|X_t-X^0_t\big|\int_0^1\big|b\big(X^0_t,t\big)\big|\,dt.
\end{aligned}
$$

By Lemma 2.3, we have $\phi^2_{1,2}(n,\varepsilon)\to_{Q_\varepsilon}0$ as both $n\to\infty$ and $\varepsilon\to 0$. For $\phi^3_{1,2}(n,\varepsilon)$, since

$$
|X_{t_{i-1}}|^m=\big(\big|X_{t_{i-1}}-X^0_{t_{i-1}}\big|+\big|X^0_{t_{i-1}}\big|\big)^m\leq 2^m\big(\big|X_{t_{i-1}}-X^0_{t_{i-1}}\big|^m+\big|X^0_{t_{i-1}}\big|^m\big), \tag{3.2}
$$

where $m\geq 1$. By condition (C.1), we have

$$
\phi^3_{1,2}(n,\varepsilon)\leq n^{-1}K'L^2 2^m\sum_{i=1}^{n}\big|X_{t_{i-1}}-X^0_{t_{i-1}}\big|^{m+2}+n^{-1}K'L^2 2^m\sum_{i=1}^{n}\big|X^0_{t_{i-1}}\big|^m\big|X_{t_{i-1}}-X^0_{t_{i-1}}\big|^2.
$$

By Lemma 2.3, we obtain $\phi_{1,2}^3(n, \varepsilon) \to_{Q_\varepsilon} 0$ as $n \to \infty$ and $\varepsilon \to 0$. And by the same argument, we get $\phi_{1,2}^4(n, \varepsilon) \to_{Q_\varepsilon} 0$ as both $n \to \infty$ and $\varepsilon \to 0$. Thus we get $\phi_{1,2}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as $n \to \infty$. For $\phi_{1,3}(n, \varepsilon)$, by conditions (C.1), (C.4), and (3.2), we have

$$\left| \phi_{1,3}(n, \varepsilon) \right| \leq n^{-1} \sum_{i=1}^{n} \left| \sigma^{-2}(X_{t_{i-1}}, t_{i-1}) - \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) \right| b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$\leq n^{-1} \sum_{i=1}^{n} \sigma^{-2}(X_{t_{i-1}}, t_{i-1}) \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1})$$

$$\times \left| \sigma^2(X_{t_{i-1}}, t_{i-1}) - \sigma^2(X_{t_{i-1}}^0, t_{i-1}) \right| b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$\leq n^{-1} 2K'LK \sum_{i=1}^{n} \left(1 + |X_{t_{i-1}}|^m \right) \left| X_{t_{i-1}} - X_{t_{i-1}}^0 \right| \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$\leq n^{-1} 2K'LK \sum_{i=1}^{n} \left| X_{t_{i-1}} - X_{t_{i-1}}^0 \right| \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$+ n^{-1} 2K'LK2^m \sum_{i=1}^{n} \left| X_{t_{i-1}} - X_{t_{i-1}}^0 \right|^{m+1} \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$+ n^{-1} 2K'LK2^m \sum_{i=1}^{n} \left| X_{t_{i-1}} - X_{t_{i-1}}^0 \right| \left| X_{t_{i-1}}^0 \right|^m \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$\leq 2K'LK \left(\sup_{0 \leq t \leq 1} \left| X_t - X_t^0 \right| + 2^m \sup_{0 \leq t \leq 1} \left| X_t - X_t^0 \right|^{m+1} \right) n^{-1}$$

$$\times \sum_{i=1}^{n} \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) b^2(X_{t_{i-1}}^0, t_{i-1})$$

$$+ 2K'LK2^m \sup_{0 \leq t \leq 1} \left| X_t - X_t^0 \right| n^{-1} \sum_{i=1}^{n} \left| X_{t_{i-1}}^0 \right|^m \sigma^{-2}(X_{t_{i-1}}^0, t_{i-1}) b^2(X_{t_{i-1}}^0, t_{i-1}).$$

We then get $\phi_{1,3}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as both $n \to \infty$ and $\varepsilon \to 0$. □

Next, we turn to the study of the asymptotic behavior of $\phi_2(n, \varepsilon)$.

Lemma 3.2 *For both $n \to \infty$ and $\varepsilon \to 0$,*

$$\phi_2(n, \varepsilon) \to_{Q_\varepsilon} 0.$$

Proof From (2.12), we first show

$$|X_t - X_{t_{i-1}}|$$

$$\leq |r_0| \int_{t_{i-1}}^{t} \left(\left| b(X_s, s) - b(X_{t_{i-1}}, t_{i-1}) \right| + \left| b(X_{t_{i-1}}, t_{i-1}) \right| \right) ds + \varepsilon \left| \int_{t_{i-1}}^{t} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|$$

$$\leq |r_0| L \int_{t_{i-1}}^{t} |X_s - X_{t_{i-1}}| \, ds + n^{-1} |r_0| \left| b(X_{t_{i-1}}, t_{i-1}) \right| + \varepsilon \sup_{t_{i-1} \leq t \leq t_i} \left| \int_{t_{i-1}}^{t} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|.$$

By Gronwall's inequality, we obtain

$$|X_t - X_{t_{i-1}}| \leq e^{r_0 L(t - t_{i-1})} \left[n^{-1} |r_0| \left| b(X_{t_{i-1}}, t_{i-1}) \right| + \varepsilon \sup_{t_{i-1} \leq t \leq t_i} \left| \int_{t_{i-1}}^{t} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| \right].$$

This then yields

$$\sup_{t_{i-1} \le t \le t_i} |X_t - X_{t_{i-1}}| \le e^{r_0 L n^{-1}} \left[n^{-1} |r_0| |b(X_{t_{i-1}}, t_{i-1})| + \varepsilon \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^t \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| \right].$$

By conditions (C.1) and (C.4), we have

$$\phi_2(n, \varepsilon) \le |r_0| \sum_{i=1}^n K'\big(1 + |X_{t_{i-1}}|^m\big) |b(X_{t_{i-1}}, t_{i-1})|$$

$$\times \left| \int_{t_{i-1}}^{t_i} \big(b(X_s, s) - b(X_{t_{i-1}}, t_{i-1})\big) \, ds \right|$$

$$\le K'L|r_0| \sum_{i=1}^n \big(1 + |X_{t_{i-1}}|^m\big) |b(X_{t_{i-1}}, t_{i-1})| \int_{t_{i-1}}^{t_i} |X_s - X_{t_{i-1}}| \, ds$$

$$\le K'L|r_0| \sum_{i=1}^n \big(1 + |X_{t_{i-1}}|^m\big) |b(X_{t_{i-1}}, t_{i-1})| n^{-1} \sup_{t_{i-1} \le t \le t_i} |X_t - X_{t_{i-1}}|$$

$$\le K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n \big(1 + |X_{t_{i-1}}|^m\big) |b(X_{t_{i-1}}, t_{i-1})|^2$$

$$+ K'L|r_0| e^{\frac{|r_0|L}{n}} n^{-1} \varepsilon \sum_{i=1}^n \big(1 + |X_{t_{i-1}}|^m\big) |b(X_{t_{i-1}}, t_{i-1})|$$

$$\times \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^t \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|$$

$$=: \phi_{2,1}(n, \varepsilon) + \phi_{2,2}(n, \varepsilon). \tag{3.3}$$

For $\phi_{2,1}(n, \varepsilon)$, by condition (C.1) and (3.2), we have

$$\phi_{2,1}(n, \varepsilon) \le K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n \big(1 + 2^m |X_{t_{i-1}}^0|^m + 2^m |X_{t_{i-1}} - X_{t_{i-1}}^0|^m\big)$$

$$\times \big(|b(X_{t_{i-1}}^0, t_{i-1})| + L|X_{t_{i-1}} - X_{t_{i-1}}^0|\big)^2$$

$$\le K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n |b(X_{t_{i-1}}^0, t_{i-1})|^2 \big(1 + 2^m |X_{t_{i-1}}^0|^m\big)$$

$$+ K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n |b(X_{t_{i-1}}^0, t_{i-1})|^2 |X_{t_{i-1}} - X_{t_{i-1}}^0|^m$$

$$+ K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n |X_{t_{i-1}} - X_{t_{i-1}}^0|^2 \big(1 + 2^m |X_{t_{i-1}}^0|^m\big)$$

$$+ K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n |X_{t_{i-1}} - X_{t_{i-1}}^0|^{m+2}$$

$$\le K'L|r_0|^2 e^{\frac{|r_0|L}{n}} n^{-2} \sum_{i=1}^n |b(X_{t_{i-1}}^0, t_{i-1})|^2 \big(1 + 2^m |X_{t_{i-1}}^0|^m\big)$$

$$+ K'L|r_0|^2 e^{\frac{|r_0|L}{n}} \sup_{0 \le t \le 1} |X_{t_{i-1}} - X_{t_{i-1}}^0|^m n^{-2} \sum_{i=1}^n |b(X_{t_{i-1}}^0, t_{i-1})|^2$$

$$+ K'L|r_0|^2 e^{\frac{|r_0|L}{n}} \sup_{0 \le t \le 1} |X_{t_{i-1}} - X^0_{t_{i-1}}|^2 n^{-2} \sum_{i=1}^{n} (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$+ K'L|r_0|^2 e^{\frac{|r_0|L}{n}} \sup_{0 \le t \le 1} |X_{t_{i-1}} - X^0_{t_{i-1}}|^{m+2} n^{-1}. \tag{3.4}$$

By Lemma 2.3, it is clear to see that $\phi_{2,1}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as both $n \to \infty$ and $\varepsilon \to 0$.

As for $\phi_{2,2}(n, \varepsilon)$, we have

$$\phi_{2,2} \le K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} (1 + 2^m |X^0_{t_{i-1}}|^m + 2^m |X_{t_{i-1}} - X^0_{t_{i-1}}|^m)$$

$$\times \left(|b(X^0_{t_{i-1}}, t_{i-1})| + L|X_{t_{i-1}} - X^0_{t_{i-1}}| \right) \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$\le K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m) \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| |X_{t_{i-1}} - X^0_{t_{i-1}}|^m \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |X_{t_{i-1}} - X^0_{t_{i-1}}| (1 + 2^m |X^0_{t_{i-1}}|^m) \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |X_{t_{i-1}} - X^0_{t_{i-1}}|^{m+1} \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$\le K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m) \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sup_{0 \le t \le 1} |X_{t_{i-1}} - X^0_{t_{i-1}}|^m \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})|$$

$$\times \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sup_{0 \le t \le 1} |X_{t_{i-1}} - X^0_{t_{i-1}}|^m \sum_{i=1}^{n} (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$\times \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sup_{0 \le t \le 1} |X_{t_{i-1}} - X^0_{t_{i-1}}|^{m+1} \sum_{i=1}^{n} \sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$=: \phi^1_{2,2}(n, \varepsilon) + \phi^2_{2,2}(n, \varepsilon) + \phi^3_{2,2}(n, \varepsilon) + \phi^4_{2,2}(n, \varepsilon). \tag{3.5}$$

For $\phi^1_{2,2}(n, \varepsilon)$, by the Markov inequality, the Hölder inequality, the Gronwall inequality, and Lemma 2.1, for any given $\gamma > 0$, we have

$$Q_\varepsilon (|\phi^1_{2,2}(n, \varepsilon)| > \gamma) \le \gamma^{-1} K'L|r_0|e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$\times E_{Q_\varepsilon} \left[\sup_{t_{i-1} \le t \le t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right| \right]$$

$$\leq \gamma^{-1} K' L |r_0| e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$\times 4\sqrt{2} E_{Q_\varepsilon} \left[\left(\int_{t_{i-1}}^{t_i} |\sigma(X_s, s)|^2 \, ds \right)^{\frac{1}{2}} \right]$$

$$\leq \gamma^{-1} K' L |r_0| e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$\times 4\sqrt{2} \left(\int_{t_{i-1}}^{t_i} E_{Q_\varepsilon} |\sigma(X_s, s)|^2 \, ds \right)^{\frac{1}{2}}$$

$$\leq \gamma^{-1} K' L |r_0| e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$\times 4\sqrt{2} \left(\int_{t_{i-1}}^{t_i} K(1 + E_{Q_\varepsilon} |X_s|^2) \, ds \right)^{\frac{1}{2}}$$

$$\leq \gamma^{-1} K' L |r_0| e^{\frac{|r_0|L}{n}} \varepsilon n^{-1} \sum_{i=1}^{n} |b(X^0_{t_{i-1}}, t_{i-1})| (1 + 2^m |X^0_{t_{i-1}}|^m)$$

$$\times 4\sqrt{2} K^{\frac{1}{2}} (1 + C)^{\frac{1}{2}} n^{-\frac{1}{2}}.$$

It then implies that $\phi^1_{2,2}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as both $n \to \infty$ and $\varepsilon \to 0$. With similar derivations, through using Lemma 2.3, we get furthermore $\phi^2_{2,2}(n, \varepsilon), \phi^3_{2,2}(n, \varepsilon), \phi^4_{2,2}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as both $n \to \infty$ and $\varepsilon \to 0$. □

Finally, let us turn to the asymptotic behavior of $\phi_3(n, \varepsilon)$. We have the followingr result.

Lemma 3.3 *Under conditions* (C.1)-(C.4), *we have* $\phi_3(n, \varepsilon) \to_{Q_\varepsilon} 0$ *as both* $n \to \infty$ *and* $\varepsilon \to 0$ *with* $\varepsilon n^{1/2} \to 0$.

Proof By conditions (C.1) and (C.4), and (3.2), we have

$$|\phi_3(n, \varepsilon)| \leq \varepsilon \sum_{i=1}^{n} K'(1 + |X_{t_{i-1}}|^m) |b(X_{t_{i-1}}, t_{i-1})| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$\leq \varepsilon \sum_{i=1}^{n} K'(1 + 2^m |X^0_{t_{i-1}}|^m + 2^m |X_{t_{i-1}} - X^0_{t_{i-1}}|^m)$$

$$\times \left(|b(X^0_{t_{i-1}}, t_{i-1})| + L|X_{t_{i-1}} - X^0_{t_{i-1}}| \right) \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$\leq \varepsilon K' \sum_{i=1}^{n} (1 + 2^m |X^0_{t_{i-1}}|^m) |b(X^0_{t_{i-1}}, t_{i-1})| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ \varepsilon K' 2^m \sum_{i=1}^{n} |X_{t_{i-1}} - X^0_{t_{i-1}}|^m |b(X^0_{t_{i-1}}, t_{i-1})| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ \varepsilon K' L \sum_{i=1}^{n} (1 + 2^m |X^0_{t_{i-1}}|^m) |X_{t_{i-1}} - X^0_{t_{i-1}}| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$+ \varepsilon K' 2^m L \sum_{i=1}^{n} |X_{t_{i-1}} - X^0_{t_{i-1}}|^{m+1} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}^\varepsilon_s \right|$$

$$\leq \varepsilon K' \sum_{i=1}^{n} \left(1 + 2^m |X_{t_{i-1}}^0|^m\right) |b(X_{t_{i-1}}^0, t_{i-1})| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|$$

$$+ \varepsilon K' 2^m \sup_{0 \leq t \leq 1} |X_{t_{i-1}} - X_{t_{i-1}}^0|^m \sum_{i=1}^{n} |b(X_{t_{i-1}}^0, t_{i-1})| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|$$

$$+ \varepsilon K' L \sup_{0 \leq t \leq 1} |X_{t_{i-1}} - X_{t_{i-1}}^0| \sum_{i=1}^{n} \left(1 + 2^m |X_{t_{i-1}}^0|^m\right) \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|$$

$$+ \varepsilon K' 2^m L \sup_{0 \leq t \leq 1} |X_{t_{i-1}} - X_{t_{i-1}}^0|^{m+1} \sum_{i=1}^{n} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right|$$

$$=: \phi_{3,1}(n, \varepsilon) + \phi_{3,2}(n, \varepsilon) + \phi_{3,3}(n, \varepsilon) + \phi_{3,4}(n, \varepsilon).$$

For $\phi_{3,1}(n, \varepsilon)$, by the Markov inequality and Lemma 2.1, we have, for any given $\gamma > 0$,

$$Q_\varepsilon\left(|\phi_{3,1}(n, \varepsilon)| > \gamma\right) \leq \gamma^{-1} K' \varepsilon \sum_{i=1}^{n} \left(1 + 2^m |X_{t_{i-1}}^0|^m\right) |b(X_{t_{i-1}}^0, t_{i-1})|$$

$$\times E_{Q_\varepsilon}\left[\sup_{t_{i-1} \leq t \leq t_i} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s^\varepsilon \right| \right]$$

$$\leq \gamma^{-1} K' \varepsilon \sum_{i=1}^{n} \left(1 + 2^m |X_{t_{i-1}}^0|^m\right) |b(X_{t_{i-1}}^0, t_{i-1})|$$

$$\times 4\sqrt{2} E_{Q_\varepsilon}\left[\left(\int_{t_{i-1}}^{t_i} |\sigma(X_s, s)|^2 \, ds \right)^{\frac{1}{2}} \right]$$

$$\leq \gamma^{-1} K' \varepsilon \sum_{i=1}^{n} \left(1 + 2^m |X_{t_{i-1}}^0|^m\right) |b(X_{t_{i-1}}^0, t_{i-1})|$$

$$\times 4\sqrt{2} K^{\frac{1}{2}} (1 + C)^{\frac{1}{2}} n^{-\frac{1}{2}}.$$

The above further implies that $\phi_{3,1}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as $n \to \infty$ and $\varepsilon n^{\frac{1}{2}} \to 0$ simultaneously. Here, we remark that the new condition $\varepsilon n^{\frac{1}{2}} \to 0$ has been added to have convergence. In the same manner, we can further obtain $\phi_{3,2}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as $n \to \infty$ and $\varepsilon n^{\frac{1}{2}} \to 0$; $\phi_{3,3}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as $n \to \infty$ and $\varepsilon n^{\frac{1}{2}} \to 0$; $\phi_{3,4}(n, \varepsilon) \to_{Q_\varepsilon} 0$ as $n \to \infty$ and $\varepsilon n^{\frac{1}{2}} \to 0$. $\qquad \square$

Now we are in the position to prove Theorem 3.1.

Proof of Theorem 3.1 Let $n \to \infty$ and $\varepsilon \to 0$ with $\varepsilon n^{\frac{1}{2}} \to 0$. By using Lemma 3.1, Lemma 3.2, and Lemma 3.3, we have therefore

$$\hat{r}_{n,\varepsilon} = r_0 + \frac{\phi_2(n, r)}{\phi_1(n, r)} + \frac{\phi_3(n, r)}{\phi_1(n, r)} \to_{Q_\varepsilon} r_0.$$

This completes the proof. $\qquad \square$

4 Asymptotics of the least square estimator

Our aim of this section is to study the asymptotic behavior of the least square estimator. For the sake of simplicity, we assume that $\alpha(x, t, \varepsilon) = \varepsilon \alpha(x, t)$, so that $Q_\varepsilon = Q$ is independent of ε. The main result of this section is formulated as follows.

Theorem 4.1 *There exist two independent Q-random variables U_1 and U_2 with the standard normal distribution $N(0,1)$ such that*

$$\varepsilon^{-1}(\hat{r}_{n,\varepsilon} - r_0) \quad \Rightarrow_Q \quad \frac{(\int_0^1 |\sigma(X_s^0,s)|^{-4}[(b(X_s^0,s)\sigma(X_s^0,s))_+]^2\,ds)^{\frac{1}{2}} U_1}{\int_0^1 \sigma^{-2}(X_s^0,s)b^2(X_s^0,s)\,ds}$$

$$- \frac{(\int_0^1 |\sigma(X_s^0,s)|^{-4}[(b(X_s^0,s)\sigma(X_s^0,s))_-]^2\,ds)^{\frac{1}{2}} U_2}{\int_0^1 \sigma^{-2}(X_s^0,s)b^2(X_s^0,s)\,ds}$$

as $n \to \infty$, $\varepsilon \to 0$, $n\varepsilon \to \infty$, and $\varepsilon n^{\frac{1}{2}} \to 0$ simultaneously.

To show Theorem 4.1, by Theorem 3.1, we could rewrite $\varepsilon^{-1}(\hat{r}_{n,\varepsilon} - r_0)$ as follows:

$$\varepsilon^{-1}(\hat{r}_{n,\varepsilon} - r_0) = \frac{\varepsilon^{-1}\phi_2(n,\varepsilon)}{\phi_1(n,\varepsilon)} + \frac{\varepsilon^{-1}\phi_3(n,\varepsilon)}{\phi_1(n,\varepsilon)} =: \frac{\Phi_2(n,\varepsilon)}{\phi_1(n,\varepsilon)} + \frac{\Phi_3(n,\varepsilon)}{\phi_1(n,\varepsilon)}.$$

Theorem 4.1 will be proved by verifying the following lemmas, in which we shall discuss the asymptotic behaviors of $\Phi_i(n,\varepsilon)$, $i = 2, 3$, respectively.

Lemma 4.1 *Under conditions* (C.1)-(C.4), *we have $\Phi_2(n,\varepsilon) \to_Q 0$ as $n \to \infty$, $\varepsilon \to 0$, and $n\varepsilon \to \infty$ simultaneously.*

Proof From Lemma 3.3, by (3.3), we have

$$\left|\Phi_2(n,\varepsilon)\right| = \varepsilon^{-1}\left|\phi_2(n,\varepsilon)\right| \leq \varepsilon^{-1}\phi_{2,1}(n,\varepsilon) + \varepsilon^{-1}\phi_{2,2}(n,\varepsilon) =: \Phi_{2,1}(n,\varepsilon) + \Phi_{2,2}(n,\varepsilon).$$

By (3.4), it is easy to see that $\Phi_{2,1}(n,\varepsilon) \to_Q 0$ as $n \to \infty$, $\varepsilon \to 0$, and $n\varepsilon \to \infty$. Notice that we have the new condition $n\varepsilon \to \infty$ for convergence. Similarly, $\Phi_{2,2}(n,\varepsilon) \to_Q 0$ as both $n \to \infty$ and $\varepsilon \to 0$. This then completes the proof of Lemma 4.1. □

Next, let us turn to the consideration of the asymptotic behavior of $\Phi_3(n,\varepsilon)$.

Lemma 4.2 *Under conditions* (C.1)-(C.4), *we have*

$$\Phi_3(n,\varepsilon) \quad \Rightarrow_Q \quad \left(\int_0^1 \left|\sigma\left(X_s^0,s\right)\right|^{-4}\left[\left(b\left(X_s^0,s\right)\sigma\left(X_s^0,s\right)\right)_+\right]^2\,ds\right)^{\frac{1}{2}} U_1$$

$$- \left(\int_0^1 \left|\sigma\left(X_s^0,s\right)\right|^{-4}\left[\left(b\left(X_s^0,s\right)\sigma\left(X_s^0,s\right)\right)_-\right]^2\,ds\right)^{\frac{1}{2}} U_2$$

as both $n \to \infty$ and $\varepsilon \to 0$ with $\varepsilon n^{\frac{1}{2}} \to 0$.

Proof We have

$$\Phi_3(n,\varepsilon) = \sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}},t_{i-1})b(X_{t_{i-1}},t_{i-1})\int_{t_{i-1}}^{t_i} \sigma(X_s,s)\,d\hat{B}_s$$

$$= \sum_{i=1}^n \sigma^{-2}\left(X_{t_{i-1}}^0,t_{i-1}\right)b\left(X_{t_{i-1}}^0,t_{i-1}\right)\int_{t_{i-1}}^{t_i} \sigma(X_s,s)\,d\hat{B}_s$$

$$+ \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) b\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \int_{t_{i-1}}^{t_i} \left(\sigma\left(X_s, s\right) - \sigma\left(X_s^0, s\right)\right) d\hat{B}_s$$

$$+ \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\left[b(X_{t_{i-1}}, t_{i-1}) - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right] \int_{t_{i-1}}^{t_i} \sigma\left(X_s, s\right) d\hat{B}_s$$

$$+ \sum_{i=1}^{n} \left[\sigma^{-2}(X_{t_{i-1}}, t_{i-1}) - \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right] b\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \int_{t_{i-1}}^{t_i} \sigma\left(X_s, s\right) d\hat{B}_s$$

$$+ \sum_{i=1}^{n} \left[\sigma^{-2}(X_{t_{i-1}}, t_{i-1}) - \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right]\left[b(X_{t_{i-1}}, t_{i-1})\right.$$

$$\left. - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right] \int_{t_{i-1}}^{t_i} \sigma\left(X_s, s\right) d\hat{B}_s$$

$$=: \Phi_{3,1}(n, \varepsilon) + \Phi_{3,2}(n, \varepsilon) + \Phi_{3,3}(n, \varepsilon) + \Phi_{3,4}(n, \varepsilon) + \Phi_{3,5}(n, \varepsilon).$$

Let us set $V(s)$ (which is deterministic) as follows:

$$V(s) := \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) b\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \sigma\left(X_s^0, s\right) \mathbb{1}_{(t_{i-1}, t_i]}(s).$$

Let $V_+(s)$ and $V_-(s)$ be the positive and negative parts of $V(s)$, respectively. Then by Theorem 4.1 of Kallenberg [30], there exist two independent Q-Brownian motions $\hat{B}'$ and $\hat{B}''$, which have the same distribution of $\hat{B}$, such that

$$\Phi_{3,1}(n, \varepsilon) = \int_0^1 V(s) \, d\hat{B}_s = \hat{B}' \circ \int_0^1 V_+^2(s) \, ds - \hat{B}'' \circ \int_0^1 V_-^2(s) \, ds.$$

Note that

$$V_+^2 = \sum_{i=1}^{n} \left|\sigma\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right|^{-4} \left(\left(b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right) \sigma\left(X_s^0, s\right)\right)_+^2 \mathbb{1}_{(t_{i-1}, t_i]}(s)$$

and

$$V_-^2 = \sum_{i=1}^{n} \left|\sigma\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right|^{-4} \left(\left(b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right) \sigma\left(X_s^0, s\right)\right)_-^2 \mathbb{1}_{(t_{i-1}, t_i]}(s).$$

Then we have

$$\int_0^1 V_+^2(s) \, ds \rightarrow \int_0^1 \left|\sigma\left(X_s^0, s\right)\right|^{-4} \left(b\left(X_s^0, s\right) \sigma\left(X_s^0, s\right)\right)_+^2 \, ds$$

and

$$\int_0^1 V_-^2(s) \, ds \rightarrow \int_0^1 \left|\sigma\left(X_s^0, s\right)\right|^{-4} \left(b\left(X_s^0, s\right) \sigma\left(X_s^0, s\right)\right)_-^2 \, ds$$

as $n \to \infty$. Then

$$\hat{B}' \circ \int_0^1 V_+^2(s) \, ds \rightarrow \hat{B}' \circ \int_0^1 \left|\sigma\left(X_s^0, s\right)\right|^{-4} \left[\left(b\left(X_s^0, s\right) \sigma\left(X_s^0, s\right)\right)_+\right]^2 \, ds$$

and

$$\hat{B}'' \circ \int_0^1 V_-^2(s)\,ds \to \hat{B}'' \circ \int_0^1 |\sigma(X_s^0,s)|^{-4}\big[(b(X_s^0,s)\sigma(X_s^0,s))_-\big]^2\,ds.$$

So we get, as $n \to \infty$,

$$\Phi_{3,1}(n,\varepsilon) \quad \Rightarrow_Q \quad \left(\int_0^1 |\sigma(X_s^0,s)|^{-4}\big[(b(X_s^0,s)\sigma(X_s^0,s))_+\big]^2\,ds\right)^{\frac{1}{2}} U_1$$

$$-\left(\int_0^1 |\sigma(X_s^0,s)|^{-4}\big[(b(X_s^0,s)\sigma(X_s^0,s))_-\big]^2\,ds\right)^{\frac{1}{2}} U_2.$$

For $\Phi_{3,2}(n,\varepsilon)$, by condition (C.3), the Markov inequality, and Lemma 2.1, Lemma 2.2, for any given $\gamma > 0$, we have

$$Q\big(|\Phi_{3,2}(n,\varepsilon)| > \gamma\big)$$

$$\leq \gamma^{-1} E_Q\left[\sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})|\left|\int_{t_{i-1}}^{t_i}(\sigma(X_s,s)-\sigma(X_s^0,s))\,d\hat{B}_s\right|\right]$$

$$\leq \gamma^{-1} \sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})| E_Q\left[\left|\int_{t_{i-1}}^{t_i}(\sigma(X_s,s)-\sigma(X_s^0,s))\,d\hat{B}_s\right|\right]$$

$$\leq \gamma^{-1} \sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})|4\sqrt{2}E_Q\left[\left(\int_{t_{i-1}}^{t_i}|\sigma(X_s,s)-\sigma(X_s^0,s)|^2\,ds\right)^{\frac{1}{2}}\right]$$

$$\leq \gamma^{-1}L4\sqrt{2}\sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})| E_Q\left[\left(\int_{t_{i-1}}^{t_i}|X_s-X_s^0|^2\,ds\right)^{\frac{1}{2}}\right]$$

$$\leq \gamma^{-1}L4\sqrt{2}n^{-\frac{1}{2}}\sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})| E_Q\left[\sup_{t_{i-1}\leq t\leq t_i}|X_s-X_s^0|\right]$$

$$\leq \gamma^{-1}L4\sqrt{2}n^{-\frac{1}{2}}\sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})|$$

$$\times E_Q\left[\varepsilon e^{L|r_0|t_i}\sup_{t_{i-1}\leq t\leq t_i}\left|\int_0^t \sigma(X_s,s)\,d\hat{B}_s\right|\right]$$

$$\leq \gamma^{-1}L32e^{L|r_0|}n^{-\frac{1}{2}}\varepsilon\sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})|b(X_{t_{i-1}}^0,t_{i-1})|$$

$$\times E_Q\left[\left(\int_0^{t_i}|\sigma(X_s,s)|^2\,ds\right)^{\frac{1}{2}}\right], \tag{4.1}$$

which tends to zero as both $n \to \infty$ and $\varepsilon \to 0$ with $\varepsilon n^{\frac{1}{2}} \to 0$.

For $\Phi_{3,3}(n,\varepsilon)$, we have

$$\Phi_{3,3}(n,\varepsilon) = \sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})\big[b(X_{t_{i-1}},t_{i-1})-b(X_{t_{i-1}}^0,t_{i-1})\big]\int_{t_{i-1}}^{t_i}\sigma(X_s,s)\,d\hat{B}_s$$

$$= \sum_{i=1}^n \sigma^{-2}(X_{t_{i-1}}^0,t_{i-1})\big[b(X_{t_{i-1}},t_{i-1})-b(X_{t_{i-1}}^0,t_{i-1})\big]\int_{t_{i-1}}^{t_i}\sigma(X_s^0,s)\,d\hat{B}_s$$

$$+ \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\left[b(X_{t_{i-1}}, t_{i-1}) - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right]$$

$$\times \int_{t_{i-1}}^{t_i} \left(\sigma(X_s, s) - \sigma\left(X_s^0, s\right)\right) d\hat{B}_s$$

$$=: \Phi_{3,3}^{1}(n, \varepsilon) + \Phi_{3,3}^{2}(n, \varepsilon).$$

For $\Phi_{3,3}^{1}(n, \varepsilon)$, by condition (C.1), we have

$$\Phi_{3,3}^{1}(n, \varepsilon) \leq \left| \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\left[b(X_{t_{i-1}}, t_{i-1}) - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right] \int_{t_{i-1}}^{t_i} \sigma\left(X_s^0, s\right) d\hat{B}_s \right|$$

$$\leq L \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\left|X_{t_{i-1}} - X_{t_{i-1}}^{0}\right| \left| \int_{t_{i-1}}^{t_i} \sigma\left(X_s^0, s\right) d\hat{B}_s \right|$$

$$\leq L \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \varepsilon e^{L|r_0|t_{i-1}} \left| \int_{0}^{t_{i-1}} \sigma(X_s, s) d\hat{B}_s \right| \left| \int_{t_{i-1}}^{t_i} \sigma\left(X_s^0, s\right) d\hat{B}_s \right|.$$

By the Markov inequality and Lemma 2.1, for any given $\gamma > 0$, we get

$$Q\left(\left|\Phi_{3,2}(n, \varepsilon)\right| > \gamma\right)$$

$$\leq \gamma^{-1} L \varepsilon \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) e^{L|r_0|t_{i-1}} E_Q \left| \int_{0}^{t_{i-1}} \sigma(X_s, s) d\hat{B}_s \right| E_Q \left| \int_{t_{i-1}}^{t_i} \sigma\left(X_s^0, s\right) d\hat{B}_s \right|$$

$$\leq 32 \gamma^{-1} L \varepsilon \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) e^{L|r_0|t_{i-1}} E_Q\left[\left(\int_{0}^{t_{i-1}} \left|\sigma(X_s, s)\right|^2 ds \right)^{\frac{1}{2}} \right]$$

$$\times \left[\left(\int_{t_{i-1}}^{t_i} \left|\sigma\left(X_s^0, s\right)\right|^2 ds \right)^{\frac{1}{2}} \right]$$

$$\leq 32 e^{L|r_0|} n^{-\frac{1}{2}} K(1 + C) \gamma^{-1} L \varepsilon \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) t_{i-1}^{\frac{1}{2}}$$

$$= 32 e^{L|r_0|} \varepsilon n^{\frac{1}{2}} K(1 + C) \gamma^{-1} L n^{-1} \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) t_{i-1}^{\frac{1}{2}},$$

which tends to zero as $n \to \infty$ and $\varepsilon n^{\frac{1}{2}} \to 0$.

For $\Phi_{3,3}^{2}(n, \varepsilon)$, by condition (C.1), Lemma 2.3, and the same arguments as used in (4.1), we find

$$\Phi_{3,3}^{2}(n, \varepsilon) \leq \left| \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\left[b(X_{t_{i-1}}, t_{i-1}) - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right] \right.$$

$$\left. \times \int_{t_{i-1}}^{t_i} \left(\sigma(X_s, s) - \sigma\left(X_s^0, s\right)\right) d\hat{B}_s \right|$$

$$\leq L \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\left|X_{t_{i-1}} - X_{t_{i-1}}^{0}\right| \left| \int_{t_{i-1}}^{t_i} \left(\sigma(X_s, s) - \sigma\left(X_s^0, s\right)\right) d\hat{B}_s \right|$$

$$\leq L \sup_{0 \leq t \leq 1} \left|X_t - X_t^0\right| \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \left| \int_{t_{i-1}}^{t_i} \left(\sigma(X_s, s) - \sigma\left(X_s^0, s\right)\right) d\hat{B}_s \right|,$$

which converges to zero as both $n \to \infty$ and $\varepsilon \to 0$ with $\varepsilon n^{\frac{1}{2}} \to 0$. Hence,

$$\Phi_{3,3}(n,\varepsilon) \to_Q 0$$

as $n \to \infty$, $\varepsilon \to 0$, and $\varepsilon n^{\frac{1}{2}} \to 0$.

For $\Phi_{3,4}(n,\varepsilon)$, by conditions (C.1)-(C.4), and (3.2), we have

$$
\begin{aligned}
\left| \Phi_{3,4}(n,\varepsilon) \right| &\leq \sum_{i=1}^{n} \left| \sigma^{-2}(X_{t_{i-1}},t_{i-1}) - \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) \right| \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right| \\
&\leq \sum_{i=1}^{n} \sigma^{-2}(X_{t_{i-1}},t_{i-1}) \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) \left| \sigma^2(X_{t_{i-1}},t_{i-1}) - \sigma^2(X^0_{t_{i-1}},t_{i-1}) \right| \\
&\quad \times \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right| \\
&\leq \sum_{i=1}^{n} K'\left(1 + |X_{t_{i-1}}|^m\right) \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) 2K \left| \sigma(X_{t_{i-1}},t_{i-1}) - \sigma(X^0_{t_{i-1}},t_{i-1}) \right| \\
&\quad \times \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right| \\
&\leq 2KK'L \sum_{i=1}^{n} \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \\
&\quad \times \left| X_{t_{i-1}} - X^0_{t_{i-1}} \right| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right| \\
&\quad + 2KK'L2^m \sum_{i=1}^{n} \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \\
&\quad \times \left| X_{t_{i-1}} - X^0_{t_{i-1}} \right|^{m+1} \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right| \\
&\quad + 2KK'L2^m \sum_{i=1}^{n} \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \\
&\quad \times \left| X^0_{t_{i-1}} \right|^m \left| X_{t_{i-1}} - X^0_{t_{i-1}} \right| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right| \\
&=: \Phi^1_{3,4}(n,\varepsilon) + \Phi^2_{3,4}(n,\varepsilon) + \Phi^3_{3,4}(n,\varepsilon).
\end{aligned}
$$

From the method of the convergence of $\Phi_{3,3}(n,\varepsilon)$, we get $\Phi^1_{3,4}(n,\varepsilon) \to_Q 0$ and $\Phi^3_{3,4}(n, \varepsilon) \to_Q 0$ as $n \to \infty$, $\varepsilon \to 0$, and $\varepsilon n^{\frac{1}{2}} \to 0$.

For $\Phi^2_{3,4}(n,\varepsilon)$, we have

$$
\begin{aligned}
\Phi^2_{3,4}(n,\varepsilon) &\leq \sup_{0 \leq t \leq 1} \left| X_{t_{i-1}} - X^0_{t_{i-1}} \right|^m 2KK'L2^m \sum_{i=1}^{n} \sigma^{-2}(X^0_{t_{i-1}},t_{i-1}) \\
&\quad \times \left| b(X^0_{t_{i-1}},t_{i-1}) \right| \left| X_{t_{i-1}} - X^0_{t_{i-1}} \right| \left| \int_{t_{i-1}}^{t_i} \sigma(X_s,s) \, d\hat{B}_s \right|,
\end{aligned}
$$

which converges to zero in probability, since $m \geq 1$, $\sup_{0 \leq t \leq 1} |X_{t_{i-1}} - X^0_{t_{i-1}}|^m$ converges to zero in probability as $\varepsilon \to 0$. Hence, $\Phi_{3,4}(n,\varepsilon) \to_Q 0$ as $n \to \infty$, $\varepsilon \to 0$, and $\varepsilon n^{\frac{1}{2}} \to 0$.

For $\Phi_{3,5}(n, \varepsilon)$, for some $G > 0$ we have

$$
\begin{aligned}
\left|\Phi_{3,5}(n, \varepsilon)\right| &\leq \sum_{i=1}^{n} \left|\sigma^{-2}(X_{t_{i-1}}, t_{i-1}) - \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right| \\
&\quad \times \left|b(X_{t_{i-1}}, t_{i-1}) - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right| \left|\int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s\right| \\
&\leq \sum_{i=1}^{n} \sigma^{-2}(X_{t_{i-1}}, t_{i-1}) \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \left|\sigma^2(X_{t_{i-1}}, t_{i-1}) - \sigma^2\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right| \\
&\quad \times \left|b(X_{t_{i-1}}, t_{i-1}) - b\left(X_{t_{i-1}}^{0}, t_{i-1}\right)\right| \left|\int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s\right| \\
&\leq G \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \left(1 + |X_{t_{i-1}}|^m\right) \left|X_{t_{i-1}} - X_{t_{i-1}}^{0}\right|^2 \left|\int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s\right| \\
&\leq \sup_{0 \leq t \leq 1} \left|X_{t_{i-1}} - X_{t_{i-1}}^{0}\right| G \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \left(1 + |X_{t_{i-1}}|^m\right) \\
&\quad \times \left|X_{t_{i-1}} - X_{t_{i-1}}^{0}\right| \left|\int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, d\hat{B}_s\right|,
\end{aligned}
$$

which converges to zero in probability Q, since $\sup_{0 \leq t \leq 1} |X_{t_{i-1}} - X_{t_{i-1}}^{0}| \to_Q 0$ as $\varepsilon \to 0$ by Lemma 2.3, and

$$
G \sum_{i=1}^{n} \sigma^{-2}\left(X_{t_{i-1}}^{0}, t_{i-1}\right) \left(1 + |X_{t_{i-1}}|^m\right) \left|X_{t_{i-1}} - X_{t_{i-1}}^{0}\right| \left|\int_{t_{i-1}}^{t_i} \sigma(X_s, s) \, dB_s\right| \to_Q 0
$$

as $n \to \infty$, $\varepsilon \to 0$, and $\varepsilon n^{\frac{1}{2}} \to 0$ by the same arguments for the convergence of $\Phi_{3,4}(n, \varepsilon)$. □

Finally, we are ready to prove Theorem 4.1.

Proof of Theorem 4.1 By using Lemma 3.1, Lemma 4.1, and Lemma 4.2, we have

$$
\begin{aligned}
\varepsilon^{-1}(\hat{r}_{n,\varepsilon} - r_0) &= \frac{\Phi_2(n, \varepsilon)}{\phi_1(n, \varepsilon)} + \frac{\Phi_3(n, \varepsilon)}{\phi_3(n, \varepsilon)} \\
&\Rightarrow_Q \frac{\left(\int_0^1 |\sigma(X_s^0, s)|^{-4} [(b(X_s^0, s)\sigma(X_s^0, s))_+]^2 \, ds\right)^{\frac{1}{2}} U_1}{\int_0^1 \sigma^{-2}(X_s^0, s) b^2(X_s^0, s) \, ds} \\
&\quad - \frac{\left(\int_0^1 |\sigma(X_s^0, s)|^{-4} [(b(X_s^0, s)\sigma(X_s^0, s))_-]^2 \, ds\right)^{\frac{1}{2}} U_2}{\int_0^1 \sigma^{-2}(X_s^0, s) b^2(X_s^0, s) \, ds}
\end{aligned}
$$

as $n \to \infty$, $n\varepsilon \to \infty$, $\varepsilon n^{\frac{1}{2}} \to 0$, and $\varepsilon \to 0$. This completes the proof. □

5 Conclusion

In the paper, we discuss parameter estimation problem for the mean-reversion stochastic differential equations driven by Brownian motion. As mentioned in the introduction, the SDEs of mean-reversion type involve a complex term $\alpha(X_t, t, \varepsilon)$ in the drift coefficient,

which makes a big difficult to apply the existing methods directly. What we have done here is to utilize the celebrated Girsanov transformation (*i.e.*, the drift transformation) to simplify the drift coefficient, which then changes the originally given probability measure P to a family of equivalent probability measures $\{Q_\varepsilon\}_{\varepsilon>0}$. With this in hand, we are able to derive an explicit least square estimator and we can further prove the convergence from the least square estimator to the true value with respect to the family $\{Q_\varepsilon\}_{\varepsilon>0}$ and we can show the asymptotic distribution of the least square estimator. The convergence discussed in this paper is with respect to a family of (equivalent) probability measures, which seems new as all the convergences appearing in the literature are only with respect to a single (*i.e.*, the originally given) probability measure.

Authors' contributions
All authors have made equal contributions. All authors have read and approved the final manuscript.

Author details
[1]School of Science, Tianjin University of Commerce, Tianjin, 300134, China. [2]Department of Mathematics, Swansea University, Swansea, SA2 8PP, UK.

References
1. Wu, J-L, Yang, W: Pricing CDO tranches in an intensity based model with the mean reversion approach. Math. Comput. Model. **52**(5-6), 814-825 (2010)
2. Øksendal, B: Stochastic Differential Equations: An Introduction with Applications. Springer, Berlin (2007)
3. Prakasa Rao, BLS: Statistical Inference for Diffusion Type Processes. Arnold, London; Oxford University Press, New York (1999)
4. Liptser, RS, Shiryaev, AN: Statistics of Random Processes. II: Applications, 2nd edn. Applications of Mathematics. Springer, Berlin (2001)
5. Kutoyants, YA: Statistical Inference for Ergodic Diffusion Processes. Springer, London (2004)
6. Dorogovcev, AJ: The consistency of an estimate of a parameter of a stochastic differential equation. Theory Probab. Math. Stat. **10**, 73-82 (1976)
7. Le Breton, A: On continuous and discrete sampling for parameter estimation in diffusion type processes. Math. Program. Stud. **5**, 124-144 (1976)
8. Kasonga, RA: The consistency of a nonlinear least squares estimator for diffusion processes. Stoch. Process. Appl. **30**, 263-275 (1988)
9. Prakasa Rao, BLS: Asymptotic theory for nonlinear least squares estimator for diffusion processes. Math. Oper.forsch. Stat., Ser. Stat. **14**, 195-209 (1983)
10. Shimizu, Y, Yoshida, N: Estimation of parameters for diffusion processes with jumps from discrete observations. Stat. Inference Stoch. Process. **9**, 227-277 (2006)
11. Shimizu, Y: *M*-Estimation for discretely observed ergodic diffusion processes with infinite jumps. Stat. Inference Stoch. Process. **9**, 179-225 (2006)
12. Kutoyants, YA: Parameter Estimation for Stochastic Process. Heldermann, Berlin (1984)
13. Kutoyants, YA: Identification of Dynamical Systems with Small Noise. Kluwer Academic, Dordrecht (1994)
14. Uchida, M, Yoshida, N: Information criteria for small diffusions via the theory of Malliavin-Watanabe. Stat. Inference Stoch. Process. **7**, 35-67 (2004)
15. Yoshida, N: Conditional expansions and their applications. Stoch. Process. Appl. **107**, 53-81 (2003)
16. Yoshida, N: Asymptotic expansion of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe. Probab. Theory Relat. Fields **92**, 275-311 (1992)
17. Sørensen, H: Parameter inference for diffusion processes observed at discrete points in time: a survey. Int. Stat. Rev. **72**, 337-354 (2004)
18. Long, H: Least squares estimator for discretely observed Ornstein-Uhlenbeck processes with small Lévy noises. Stat. Probab. Lett. **79**, 2076-2085 (2009)
19. Long, H: Parameter estimation for a class of stochastic differential equations driven by small stable noises from discrete observations. Acta Math. Sci. Ser. B **30**(3), 645-663 (2010)
20. Ma, C: A note on 'Least squares estimator for discretely observed Ornstein-Uhlenbeck processes with small Lévy noise'. Stat. Probab. Lett. **80**, 1528-1531 (2010)
21. Hu, Y, Long, H: Least squares estimator for Ornstein-Uhlenbeck processes driven by α-stable Lévy motions. Stoch. Process. Appl. **119**, 2465-2480 (2009)

22. Hu, Y, Long, H: On the singularity of least squares estimator for mean-reverting α-stable motions. Acta Math. Sci. Ser. B **28**(3), 599-608 (2009)
23. Kunitomo, N, Takahashi, A: The asymptotic expansion approach to the valuation of interest rate contingent claims. Math. Finance **11**, 117-151 (2001)
24. Takahashi, A: An asymptotic expansion approach to pricing contingent claims. Asia-Pac. Financ. Mark. **6**, 115-151 (1999)
25. Takahashi, A, Yoshida, N: An asymptotic expansion scheme for optimal investment problems. Stat. Inference Stoch. Process. **7**, 153-188 (2004)
26. Uchida, M, Yoshida, N: Asymptotic expansion for small diffusions applied to option pricing. Stat. Inference Stoch. Process. **7**, 189-223 (2004)
27. Yoshida, N: Asymptotic expansion for statistics related to small diffusions. J. Japan Statist. Soc. **22**, 139-159 (1992)
28. Ikeda, N, Watanabe, S: Stochastic Differential Equations and Diffusion Processes. North-Holland, Amsterdam (1981)
29. Mao, X: Stochastic Differential Equations and Their Applications. Horwood, Chichester (2008)
30. Kallenberg, O: Some time change representations of stable integrals, via predictable transformations of local martingales. Stoch. Process. Appl. **40**, 199-223 (1992)