RESEARCH   Open Access

The law of iterated logarithm for the estimations of diffusion-type processes

Mingzhi Mao[1*] and Gang Huang[1]

*Correspondence:
mingzhi-mao@163.com
[1] School of Mathematics and
Physics, China University of
Geosciences, Wuhan, China

Abstract

This paper mainly discusses the asymptotic behaviours on the lasso-type estimators for diffusion-type processes with a small noise. By constructing the objective function on the estimation, in view of convexity argument, it is proved that the estimator for different values of γ satisfies the iterated logarithm law. The result also presents the exponential convergence principle for the estimator converging to the true value.

MSC: 60J60; 60F05; 60J25; 60F10

Keywords: The law of iterated logarithm; Diffusion-type processes; Approach of argmins; Objective function

1 Introduction

Let $\{X_t^\epsilon\}_{0 \le t \le T}$ be the solution of the following stochastic diffusion-type process:

$$
\begin{aligned}
dX_t^\epsilon &= S_t(\theta, X^\epsilon)\, dt + \epsilon\, dW_t, \\
X_0^\epsilon &= x_0,
\end{aligned}
\tag{1}
$$

where x_0 is a fixed constant, $\theta = (\theta_1, \dots, \theta_p)' \in \Theta \subset \mathbb{R}^p$ is an unknown parameter, $S_t(\theta, x)$ is a known measurable, nonanticipative functions such that (1) has a strong unique solution, W_t is a standard Wiener process, which is usually called random noise, and $\epsilon \in (0, 1]$ the diffusion coefficient.

There is a rich literature on the methods to estimate parameter θ, such as least squares estimation, Bayesian estimation, maximum likelihood estimation, and so on. However, there is no uniform standard to comment on those methods. The limiting properties for various estimators attract the attention of statisticians because of their applicability in mathematical finance, biology and other fields; see [5, 16, 21]. As described by Kutoyants [16], the minimum distance estimation is a relatively new method compared with traditional estimation methods, which deals with less stochastic calculations and robustness. This paper considers a class of minimum distance estimators for the diffusion process (1) and discusses their exponential convergence principle. To estimate the unknown param-

Springer

eter θ, we need to introduce the ordinary differential equation:

$$dX_t^0 = S_t(\theta, X^0)\, dt, \quad 0 \le t \le T, \tag{2}$$

with initial condition $X_0^0 = x_0$.

The minimum distance estimator $\widehat{\theta}^\epsilon$ is given by

$$\begin{aligned}
\widehat{\theta}^\epsilon &= \operatorname*{argmin}_{\theta \in \overline{\Theta}} \| X^\epsilon - X^0 \|^2 \\
&= \operatorname*{argmin}_{\theta \in \overline{\Theta}} \int_0^T (X_t^\epsilon - X_t^0)^2 \alpha(dt),
\end{aligned} \tag{3}$$

where $\overline{\Theta}$ is the closure of Θ, α is some finite measure on $[0, T]$ and $\operatorname{argmin} g = \{x : g(x) = \inf g\}$. Kutoyants discussed the consistency and asymptotic normality of estimator $\widehat{\theta}^\epsilon$ in [17, 18]. Dietz and Kutoyants [7] considered a class of minimum distance estimators for diffusion processes with ergodic properties. For general minimum distance estimators, the reader can refer to Kutoyants [16] and references therein. Recently, Zhao and Zhang [22] studied the minimum distance parameter estimation for stochastic differential equations with small α-stable noises.

However, few works considered the convergence rate of $\widehat{\theta}^\epsilon$ converging to the true value of θ. This induces some stochastic dynamical systems not to be reasonably identified by finite observations of spaced time points. What's more, from the point of view of probability, the law of large numbers, the central limit theorem, and the iterated logarithm law are all contained in the limit theory, which is a whole system. This motivates us to study the convergence rate of $\widehat{\theta}^\epsilon \to \theta^*$ (θ^* is the true value of θ). In order to generalize the model, we use the constrained minimum distance estimator based on L_γ-penalized function contrast:

$$Z_\epsilon(\theta) = \int_0^T (X_t^\epsilon - X_t^0)^2 \alpha(dt) + \lambda_\epsilon \sum_{j=1}^p |\theta_j|^\gamma, \tag{4}$$

where $\gamma > 0$ is a fixed constant and $\lambda_\epsilon > 0$ is a penalty parameter with respect to ϵ. Without loss of generality, we assume that the trend functional of the processes (1) is of integral type:

$$S_t(\theta, X^\epsilon) = V(\theta, t, X^\epsilon) + \int_0^t K(\theta, t, s, X_s^\epsilon)\, ds.$$

Denote the true value of θ by θ^* and the lasso-type estimator of θ by

$$\widehat{\theta}^\epsilon = \operatorname*{argmin}_{\theta \in \overline{\Theta}} Z_\epsilon(\theta). \tag{5}$$

In this paper, we will discuss the limit behaviors of $\dfrac{\widehat{\theta}^\epsilon - \theta^*}{\epsilon \sqrt{2 \log \log(\epsilon^{-1} \vee 3)}}$, i.e., the iterated logarithm law. This shows that the estimator $\widehat{\theta}^\epsilon$ converges to a constant almost everywhere with an exponential convergence rate. We recall that Gregorio and Iacus showed that $\widehat{\theta}^\epsilon$

satisfies the central limit theorem in [11], that is,

$$\epsilon^{-1}(\widehat{\theta}^{\epsilon} - \theta^{*}) \Rightarrow \operatorname*{argmin}_{u} V(u),$$

where $\Rightarrow$ denotes convergence in distribution, $V(u)$ is a fixed random function. Our result also can be considered a supplement of Gregorio and Iacus' work.

2 Preliminaries

This section will present some basic notations and assumptions which will be used in the paper. Define the inner product by $\langle x, y \rangle = \sum_{i=1}^{p} x_i y_i$ in the space $\mathbb{R}^p$. In particular, use $|\cdot|$ for the Euclidean distance, that is, $|y| = \sqrt{y'y} = \sqrt{\sum_{i=1}^{p} \langle x_i, x_i \rangle}$, where y' denotes the transpose of y. Let **B** be some Banach space and write $\|\cdot\|$ for the corresponding norm. If **B** is the space of all continuous bounded functions on $\mathbb{R}^p$, we always define $\|f\| = \sup_{x \in \mathbb{R}^p} |f(x)|$ for any $f \in$ **B**. Let $D(\mathbb{T})$ the space of càdlàg functions (i.e., right continuous with left limits) on $\mathbb{T}$ with the Skorohod topology. For any set $A \subset \mathbb{R}^p$, we define the distance from $x \in \mathbb{R}^p$ to A by $\rho(x, A) = \inf_{y \in A} \rho(x, y)$. If $\{x^{\epsilon}\}$ is a suitable family of points in $\mathbb{R}^p$, then let $\mathbf{C}(\{x^{\epsilon}\})$ denote the cluster set of $\{x^{\epsilon}\}$. That is, $\mathbf{C}(\{x^{\epsilon}\})$ are all possible limit points of the sequence $\{x_n\}$. We sometimes use the notation $\lim_{\epsilon \to 0} x^{\epsilon} = A$ if both $\lim_{\epsilon \to 0} \rho(x^{\epsilon}, A) = 0$ and $\mathbf{C}(\{x^{\epsilon}\}) = A$. Throughout the paper, let $\mathbf{P}_{\theta}$ denote the law of $X_t(\theta)$ under parameter θ. The subscript θ indicates that the process $X_t^{\epsilon}(\theta)$ depends on θ. If it doesn't cause confusion, we always omit θ, i.e., $X_t^{\epsilon} = X_t^{\epsilon}(\theta)$.

Define $V_x(\theta, t, x) = \frac{\partial}{\partial x} V(\theta, t, x)$ and $K_x(\theta, t, s, x) = \frac{\partial}{\partial x} K(\theta, t, s, x)$. Let $Y_t = \{Y_t(\theta), 0 \le t \le T\}$ be the solution of a diffusion-type process

$$dY_t = \left(V_x(\theta, t, X^0(\theta)) Y_t + \int_0^t K_x(\theta, t, s, X_s^0(\theta)) Y_s \, ds \right) dt + dW_t \tag{6}$$

with initial condition $Y_0 = 0$. The process Y_t plays a central role in the study of the asymptotic distribution of the estimators in the theory of diffusion process X_t with small noise. Denote the p-dimensional vector of partial derivatives of $X_t^0(\theta)$ with respect to θ_j $(j = 1, \ldots, p)$ by $\dot{X}_t^0(\theta)$, that is,

$$\dot{X}_t^0(\theta) = \frac{\partial X_t^0(\theta)}{\partial \theta}$$
$$= \left(\frac{\partial}{\partial \theta_1} X_t^0(\theta), \ldots, \frac{\partial}{\partial \theta_p} X_t^0(\theta) \right)'. \tag{7}$$

It is easy to see that $\dot{X}_t^0(\theta)$ satisfies the following differential equation:

$$\frac{d\dot{X}_t^0(\theta)}{dt} = V_x(\theta, t, X_t^0(\theta)) \dot{X}_t^0(\theta) + \dot{V}(\theta, t, X_t^0(\theta))$$
$$+ \int_0^t \left(\dot{K}(\theta, t, s, X_s^0(\theta)) + K_x(\theta, t, s, X_s^0(\theta)) \dot{X}_s^0(\theta) \right) ds \tag{8}$$

and $\dot{X}_0^0(\theta) = 0$, where $\dot{V}(\theta, t, X_t^0(\theta)) = (\frac{\partial}{\partial \theta_1} V(\theta, t, X_t^0(\theta)), \ldots, \frac{\partial}{\partial \theta_p} V(\theta, t, X_t^0(\theta)))'$.

We suppose that the following regular conditions for the trend coefficient $V(\theta, t, x)$ and $K(\theta, t, s, x)$ hold:

(A1) $\epsilon^{-1}\lambda_\epsilon \to \lambda_0 \geq 0$;

(A2) for any $t \in [0, T]$,

$$\sup_{\theta \in \Theta, x \in \mathbb{R}^p} \left| \frac{\partial}{\partial \theta} V(\theta, t, x) \right| < \infty, \qquad \sup_{\theta \in \Theta, x \in \mathbb{R}^p} \left| \frac{\partial}{\partial x} V(\theta, t, x) \right| < \infty$$

and

$$\sup_{s,t \in [0,T]} \sup_{\theta \in \Theta, x \in \mathbb{R}^p} \left| \frac{\partial}{\partial \theta} K(\theta, t, s, x) \right| < \infty, \qquad \sup_{s,t \in [0,T]} \sup_{\theta \in \Theta, x \in \mathbb{R}^p} \left| \frac{\partial}{\partial x} K(\theta, t, s, x) \right| < \infty;$$

(A3) there exist two positive constants $\mathcal{M}_1$ and $\mathcal{M}_2$ such that

$$\sup_{t \in [0,T], \theta \in \Theta} \left| \frac{\partial}{\partial x} V(\theta, t, x) - \frac{\partial}{\partial x} V(\theta, t, y) \right| < \mathcal{M}_1 |x - y|,$$

$$\sup_{t,s \in [0,T], \theta \in \Theta} \left| \frac{\partial}{\partial x} K(\theta, t, s, x) - \frac{\partial}{\partial x} K(\theta, t, s, y) \right| < \mathcal{M}_1 |x - y|$$

and

$$\sup_{t,s \in [0,T]} \sup_{x \in \mathbb{R}^p} \left| \frac{\partial}{\partial x} V(\theta_1, t, s, x) - \frac{\partial}{\partial x} V(\theta_2, t, s, x) \right| \leq \mathcal{M}_2 |\theta_1 - \theta_2|,$$

$$\sup_{t,s \in [0,T]} \sup_{x \in \mathbb{R}^p} \left| \frac{\partial}{\partial x} K(\theta_1, t, s, x) - \frac{\partial}{\partial x} K(\theta_2, t, s, x) \right| \leq \mathcal{M}_2 |\theta_1 - \theta_2|.$$

One can check that under conditions (A2) and (A3), the stochastic differential Eq. (1) satisfies Condition $\mathcal{L}$, that is,

$$\left| V(\theta, t, X_t) - V(\theta, t, Y_t) \right| + \left| K(\theta, t, s, X_t) - K(\theta, t, s, Y_t) \right|$$

$$\leq L_1 \int_0^t |X_s - Y_s| \, dK_s + L_2 |X_t - Y_t|,$$

$$\left| V(\theta, t, X_t) \right| + \left| K(\theta, t, s, X_t) \right| \leq L_1 \int_0^t \left(1 + |X_s|\right) dK_s + L_2 \left(1 + |X_t|\right),$$

where L_1 and L_2 are two positive constants and K_s is a nondecreasing right-continuous function, $0 \leq K_t \leq K_0$, $K_0 > 0$. By virtue of Theorem 4.6 of [19, 20], there exists a unique $D([0, T], \mathbb{R}^p)$-value strong solution for Eq. (1) under conditions (A2) and (A3) (the reader can also see [8] for the theory on existence and uniqueness). In Lemma 4 below, we will show that the deterministic diffusion process $X_t^0(\theta)$ is differentiable with respect to θ at the point θ^* in L_2-norm under conditions (A2) and (A3), i.e.,

$$\int_0^T \left(X_t^0(\theta^* + h) - X_t^0(\theta^*) - h' \frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta = \theta^*} \right)^2 \alpha(dt) = o(|h|).$$

In the book of Kutoyants [16], the reader can see that the assumption is very important for the statistical identification problems. We now introduce the objective function

of $\widehat{\theta^\epsilon} - \theta^*$:

$$V_\epsilon(u) = \frac{1}{\epsilon^2 h(\epsilon)} \Bigg[\int_0^T \big(X_t^\epsilon - X_t^0\big(\theta^* + h(\epsilon)\epsilon u\big)\big)^2 \alpha(dt)$$

$$- \int_0^T \big(X_t^\epsilon - X_t^0\big(\theta^*\big)\big)^2 \alpha(dt) + \lambda_\epsilon \sum_{j=1}^p \big(\big|\theta_j^* + h(\epsilon)\epsilon u_j\big|^\gamma - \big|\theta_j^*\big|^\gamma\big) \Bigg],$$

where $u = (u_1, \dots, u_p)'$. It is easy to see that

$$\frac{1}{\epsilon h(\epsilon)} \big(\widehat{\theta^\epsilon} - \theta^*\big) \in \underset{u \in \mathbb{R}^p}{\operatorname{argmin}}\, V_\epsilon(u). \tag{9}$$

A simple calculation yields

$$V_\epsilon(u) = \frac{1}{\epsilon^2 h(\epsilon)} \int_0^T \big(X_t^0\big(\theta^* + h(\epsilon)\epsilon u\big) - X_t^0\big(\theta^*\big)\big)^2 \alpha(dt)$$

$$+ \frac{\lambda_\epsilon}{\epsilon^2 h(\epsilon)} \sum_{j=1}^p \big\{ \big|\theta_j^* + h(\epsilon)\epsilon u_j\big|^\gamma - \big|\theta_j^*\big|^\gamma \big\}$$

$$- \frac{2}{\epsilon^2 h(\epsilon)} \int_0^T \big(X_t - X_t^0\big(\theta^*\big)\big)\big(X_t^0\big(\theta^* + h(\epsilon)\epsilon u\big) - X_t^0\big(\theta^*\big)\big)\alpha(dt). \tag{10}$$

In Sect. 4, we will show that $V_\epsilon(u)$ can be approached by some stochastic function.

3 Main result

We state our main result as follows:

Theorem 1 *Let* $h(\epsilon) = \sqrt{2\log\log(\epsilon^{-1} \vee 3)}$. *Assume that conditions* (A1)–(A3) *hold. Then, for* $\gamma \geq 1$, *the process* $(\widehat{\theta^\epsilon} - \theta^*)/(\epsilon h(\epsilon))$ *satisfies the iterated logarithm law, that is,*

$$\limsup_{\epsilon \to 0} \rho\bigg(\frac{\widehat{\theta^\epsilon} - \theta^*}{\epsilon h(\epsilon)}, Q^{-1}K \bigg) = 0, \quad a.e.$$

and

$$\mathbf{P}\bigg(\omega : \mathbf{C}\bigg(\bigg\{ \frac{\widehat{\theta^\epsilon} - \theta^*}{\epsilon h(\epsilon)} \bigg\} \bigg) = Q^{-1}K \bigg) = 1,$$

where $\rho(\cdot, \cdot)$ *denotes the Euclidean distance, a.e. stands for almost everywhere, and*

$$Q = \int_0^T \frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \cdot \bigg(\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \bigg)' \alpha(dt)$$

and

$$K = \bigg\{ \int_0^T g \cdot \frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \alpha(dt) : g \in \mathcal{H} \text{ and } I(g) \leq 1 \bigg\}. \tag{11}$$

Here

$$I(g) := \inf_{\phi \in \mathcal{H}: g = Y_t^\phi} \bigg\{ \frac{1}{2} \int_0^T \big|\dot\phi(s)\big|^2 \, ds \bigg\},$$

$$\mathcal{H} = \left\{ \phi : \phi \, is \, an \, absolutely \, continuous \, function \, with \, \phi(0) = 0 \right.$$

$$\left. and \int_0^T |\dot{\phi}(s)|^2 \, ds < \infty \right\} \tag{12}$$

and Y^ϕ is the solution of the integral equation given by

$$Y_t^\phi = \int_0^t \left[V_x(\theta, s, X_s^0) + \int_0^s K_x(\theta, s, u, X_u^0(\theta)) \, du \right] Y_s^\phi \, ds + \phi(t). \tag{13}$$

Remark For $0 < \gamma < 1$, supposing that conditions (A2)–(A3) hold and $\lambda_\epsilon / \epsilon^{1-\gamma} \to \lambda_0 \geq 0$, it can be still proved that $(\widehat{\theta^\epsilon} - \theta^*)/(\epsilon h(\epsilon))$ satisfies the iterated logarithm law. The proof method is the same as that of Theorem 1.

We below present an example, which is an application of the above result.

Example Consider the following diffusion process:

$$dX_t^\epsilon = \theta X_t^\epsilon \, dt + \epsilon \, dW_t, \quad X_0^\epsilon = x_0 \neq 0, 0 \leq t \leq T, \tag{14}$$

where $\theta \in (\kappa_1, \kappa_2) = \Theta$ and W_t is a standard Wiener process, then the limit solution is

$$X_t^0 = x_0 e^{\theta t}, \quad 0 \leq t \leq T.$$

The minimum distance estimator $\widehat{\theta^\epsilon}$ is defined by

$$\widehat{\theta^\epsilon} = \operatorname*{argmin}_{\theta \in \Theta} \int_0^T (X_t^\epsilon - x_0 e^{\theta t})^2 \alpha(dt).$$

It can be checked that the process (14) satisfies the conditions of Theorem 7.5 of [16], so $\widehat{\theta^\epsilon}$ is consistent and asymptotically normal:

$$\frac{\widehat{\theta^\epsilon} - \theta^*}{\epsilon} \Rightarrow \frac{\int_0^T t e^{\theta^* t} Y_t \, dt}{x_0 \int_0^T t^2 e^{2\theta^* t} \alpha(dt)},$$

where Y_t satisfies $dY_t = \theta^* Y_t \, dt + dW_t$, $Y_0 = 0$, $0 \leq t \leq T$. It can be easily proved that $\int_0^T t e^{\theta^* t} Y_t \alpha(dt)$ is a Gaussian process with variance

$$\sigma^2 = \frac{1}{2\theta} \int_0^T \int_0^T st e^{2\theta^*(s+t)} \left[1 - e^{-2\theta^*(t \wedge s)} \right] \alpha(dt) \alpha(ds). \tag{15}$$

By virtue of Theorem 1, a simple calculation can show that the estimator $\widehat{\theta^\epsilon}$ has the following limit behavior:

$$\limsup_{\epsilon \to 0} \frac{\widehat{\theta^\epsilon} - \theta^*}{\epsilon \sqrt{2 \log \log(\epsilon^{-1} \vee 3)}} = \frac{\sigma}{x_0 \int_0^T t^2 e^{2\theta^* t} \alpha(dt)}, \quad a.e.$$

where σ^2 is defined in (15).

4 Proofs

In order to prove Theorem 1, we need the following lemmas. Lemma 2 is about the approach of argmins, which is crucial to study the asymptotic theory for some argmin processes of parametrized convex objective functions (see [10] and [14] for other applications).

Lemma 1 *Suppose that $A^\epsilon(u)$ and $B^\epsilon(u)$, $u \in \mathbb{R}^p$, are two convex bounded functions. Assume that $\lim_{\epsilon \to 0} h(\epsilon) = \infty$ and for any $u \in \mathbb{R}^p$, $\delta > 0$,*

$$\limsup_{\epsilon \to 0} \frac{1}{h^2(\epsilon)} \log \mathbf{P}\big(\big|A^\epsilon(u) - B^\epsilon(u)\big| \geq \delta\big) = -\infty. \tag{16}$$

Then for any compact set $D \subset \mathbb{R}^p$,

$$\limsup_{\epsilon \to 0} \frac{1}{h^2(\epsilon)} \log \mathbf{P}\Big(\sup_{u \in D}\big|A^\epsilon(u) - B^\epsilon(u)\big| \geq \epsilon\Big) = -\infty. \tag{17}$$

Proof The approach stems from Lemma 3 of Kato [13]. For the sake of completeness, we simply state the proof. By the the convexity and boundedness of $A^\epsilon(u)$, there exists a constant $\beta_1 > 0$ satisfying

$$\big|A^\epsilon(u) - A^\epsilon(v)\big| \leq \beta_1 |u - v|, \quad \text{for any } u, v \in D. \tag{18}$$

Similarly, there exists another constant $\beta_2 > 0$ satisfying

$$\big|B^\epsilon(u) - B^\epsilon(v)\big| \leq \beta_2 |u - v|, \quad \text{for any } u, v \in D. \tag{19}$$

Let $\beta_0 = \max\{\beta_1, \beta_2\}$. For any $\epsilon > 0$, there exists a finite set $D_1 \subset D$ such that each point of D lies within the distance $\frac{\epsilon}{3\beta_0}$ of at least one point of D_1. Equation (16) implies that

$$\limsup_{\epsilon \to 0} \frac{1}{h^2(\epsilon)} \log \mathbf{P}\Big(\sup_{v \in D_1}\big|A^\epsilon(v) - B^\epsilon(v)\big| \geq \epsilon\Big) = -\infty. \tag{20}$$

Giving any $u \in K$, let v be a point of D_1 such that $|u - v| \leq \frac{\epsilon}{3\beta_0}$. Then

$$\begin{aligned}
\big|A^\epsilon(u) - B^\epsilon(u)\big| &\leq \big|A^\epsilon(u) - A^\epsilon(v)\big| + \big|A^\epsilon(v) - B^\epsilon(v)\big| + \big|B^\epsilon(v) - B^\epsilon(u)\big| \\
&\leq \beta_1|u - v| + \big|A^\epsilon(v) - B^\epsilon(v)\big| + \beta_2|v - u| \\
&\leq \frac{2\epsilon}{3} + \big|A^\epsilon(v) - B^\epsilon(v)\big|.
\end{aligned}$$

Consequently,

$$\mathbf{P}\big(\big|A^\epsilon(u) - B^\epsilon(u)\big| \geq \epsilon\big) \leq \mathbf{P}\bigg(\big|A^\epsilon(v) - B^\epsilon(v)\big| \geq \frac{\epsilon}{3}\bigg).$$

By virtue of (20), the desired result is obtained.   □

Lemma 2 *Suppose that $A^\epsilon(u)$ and $B^\epsilon(u)$ are two suitable families of convex random functions defined on a compact set $S \in \mathbb{R}^p$, where $\epsilon \in (0,1]$ is the index parameter. Let a^ϵ be the argmin of $A^\epsilon(u)$ and assume that $B^\epsilon(u)$ has a unique argmin b^ϵ. Then for any positive constant δ,*

$$\mathbf{P}\big(\big|a^\epsilon - b^\epsilon\big| > \delta\big) \leq \mathbf{P}\bigg(\widetilde{\triangle}_\epsilon \geq \frac{\eta}{2}\bigg), \tag{21}$$

where $\widetilde{\triangle}_\epsilon = \sup_{u \in \{u : |u-b^\epsilon| \leq \delta\}} |A^\epsilon(u) - B^\epsilon(u)|$ and

$$\eta = \inf_{v \in \mathbb{S}^{p-1}} \big\{ B^\epsilon\big(b^\epsilon + \delta v\big) - B^\epsilon\big(b^\epsilon\big) \big\}. \tag{22}$$

Proof Let $\mathbb{S}^{p-1} = \{x \in \mathbb{R}^p : |x| = 1\}$. For any $v \in \mathbb{S}^{p-1}$, the convexity of $A^\epsilon(u)$ yields

$$\bigg(1 - \frac{\delta}{l}\bigg)A^\epsilon\big(b^\epsilon\big) + \frac{\delta}{l}A^\epsilon\big(b^\epsilon + lv\big) \geq A^\epsilon\big(b^\epsilon + \delta v\big), \quad \forall l > \delta.$$

It is equivalent to

$$\frac{\delta}{l}\big(A^\epsilon\big(b^\epsilon + lv\big) - A^\epsilon\big(b^\epsilon\big)\big) \geq A^\epsilon\big(b^\epsilon + \delta v\big) - A^\epsilon\big(b^\epsilon\big).$$

Let $\triangle_\epsilon(u) = A^\epsilon(u) - B^\epsilon(u)$. We have

$$\begin{aligned} &\frac{\delta}{l}\big(A^\epsilon\big(b^\epsilon + lv\big) - A^\epsilon\big(b^\epsilon\big)\big) \\ &\quad \geq B^\epsilon\big(b^\epsilon + \delta v\big) - B^\epsilon\big(b^\epsilon\big) + \big(\triangle_\epsilon\big(b^\epsilon + \delta v\big) - \triangle_\epsilon\big(b^\epsilon\big)\big) \\ &\quad \geq \eta - 2\widetilde{\triangle}_\epsilon. \end{aligned} \tag{23}$$

Since S is a compact set and b^ϵ is the unique argmin point of B^ϵ, so η is a positive random variable. If $\widetilde{\triangle}_\epsilon < \frac{\eta}{2}$, then $A^\epsilon(b^\epsilon + lv) - A^\epsilon(b^\epsilon) > 0$ for each v. This implies that if $|a^\epsilon - b^\epsilon| > \delta$, then $A^\epsilon(a^\epsilon) - A^\epsilon(b^\epsilon) > 0$. The minimum property of a^ϵ will lead to a contradiction. Thus, for any positive constant δ,

$$\mathbf{P}\big(\big|a^\epsilon - b^\epsilon\big| > \delta\big) \leq \mathbf{P}\bigg(\widetilde{\triangle}_\epsilon \geq \frac{\eta}{2}\bigg).$$

The proof is completed. □

Lemma 3 *Assume that $A^\epsilon(u)$ is a convex random function defined in an open set $S \in \mathbb{R}^p$. Let $B^\epsilon(u) = -u^T U^\epsilon + \frac{1}{2}u^T Q u$, where Q is a symmetric and positive define $p \times p$ matrix and U^ϵ is stochastically bounded. Furthermore, let $1 \leq h(\epsilon) = o(1/\sqrt{\epsilon})$, we assume the following three conditions hold:*

(i) *Random process U^ϵ satisfies the iterated logarithm law, that is, there exists a fixed bounded symmetric set K in $\mathbb{R}^p$ such that*

$$\limsup_{\epsilon \to 0} \rho\bigg(\frac{U^\epsilon}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \wedge 3))}}, K\bigg) = 0, \quad a.e.$$

and

$$\mathbf{P}\left(\omega : \mathbf{C}\left(\left\{\frac{U^\epsilon}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \wedge 3))}}\right\}\right) = K\right) = 1.$$

(ii) *For any $R > 0$ and any $\delta > 0$, there exists an $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0]$,*

$$\mathbf{P}\left(\left|A^\epsilon(u) - B^\epsilon(u)\right| \geq \delta h(\epsilon)\right) \leq e^{-Rh^2(\epsilon)}. \tag{24}$$

Then, a^ϵ, the minimizer of convex process $A^\epsilon(u)$, satisfies the iterated logarithm law, that is,

$$\limsup_{\epsilon \to 0} \rho\left(\frac{a^\epsilon}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \wedge 3))}}, Q^{-1}K\right) = 0, \quad a.e.$$

and

$$\mathbf{P}\left(\omega : \mathbf{C}\left(\left\{\frac{a^\epsilon}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \wedge 3))}}\right\}\right) = Q^{-1}K\right) = 1,$$

where $Q^{-1}K = \{Q^{-1}x : x \in K\}$.

Proof Let $b^\epsilon = Q^{-1}U^\epsilon$. It is easy to see that b^ϵ is the unique minimum point of $B^\epsilon(u)$. The continuous mapping theorem in the iterated logarithm law yields

$$\limsup_{\epsilon \to 0} \rho\left(\frac{b^\epsilon}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \wedge 3))}}, Q^{-1}K\right) = 0, \quad a.e.$$

and

$$\mathbf{P}\left(\omega : \mathbf{C}\left(\left\{\frac{b^\epsilon}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \wedge 3))}}\right\}\right) = Q^{-1}K\right) = 1.$$

Then a simple calculation shows that

$$\begin{aligned}
B^\epsilon(u) - B^\epsilon(b^\epsilon) &= \frac{1}{2}(u - b^\epsilon)' Q(u - b^\epsilon) \\
&\geq c|u - b^\epsilon|^2,
\end{aligned} \tag{25}$$

where $c > 0$ is the smallest eigenvalue of Q.

As in the proof of Lemma 2, note that by the definition of η, we further have

$$\mathbf{P}\left(\left|a^\epsilon - b^\epsilon\right| > \delta\right) \leq \mathbf{P}\left(\widetilde{\triangle}_n \geq \frac{c\delta^2}{2}\right). \tag{26}$$

From (25) and (26), combining condition (ii), we have

$$\mathbf{P}\left(\left|a^\epsilon - b^\epsilon\right| \geq \delta h(\epsilon)\right) \leq e^{-Rh^2(\epsilon)}.$$

This implies that a^ϵ, the minimizer of convex process $A^\epsilon(u)$, satisfies the iterated logarithm law. $\qquad\square$

Lemma 4 *Assume X_t^ϵ is the solution of the following stochastic differential equation:*

$$dX_t^\epsilon = b(X_t^\epsilon)\,dt + \sqrt{\epsilon}\,\sigma(X_t^\epsilon)\,dB_t, \quad X_0^\epsilon = x_0,$$

and that X_t^0 is the solution of the ordinary differential equations

$$dX_t^0 = b(X_t^0)\,dt, \quad X_t^0 = x_0.$$

Assume that $b(\cdot)$ and $\sigma(\cdot)$ are Lipschitz continuous on every compact subset of $\mathbb{R}$ and there exists a positive constant L, for any $x, y \in \mathbb{R}^+$, satisfying

$$xb(x) \le L(1 + |x|^2), \qquad |\sigma(x) - \sigma(y)| \le L|x - y|^2.$$

Then, the process $\dfrac{X_t^\epsilon - X_t^0}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \vee 3)}}$ satisfies the iterated logarithm law, that is,

$$\limsup_{\epsilon \to 0} \rho\left(\frac{X_t^\epsilon - X_t^0}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \vee 3)}}, K\right) = 0, \quad a.e.$$

and

$$\mathbf{P}\left(\omega : \mathbf{C}\left(\left\{\frac{X_t^\epsilon - X_t^0}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \vee 3)}}\right\}\right) = K\right) = 1.$$

Here, $K = \{g : g \in \mathcal{H} \text{ and } I(g) \le 1\}$ and

$$I(g) := \inf_{\phi \in \mathcal{H}: g = Y_t^\phi} \left\{\frac{1}{2}\int_0^T |\dot\phi(s)|^2\,ds\right\},$$

where $\mathcal{H}$ is defined in (12) and Y^ϕ is the solution of the integral equation given by

$$Y_t^\phi = \int_0^t \dot{b}(X_s^0)Y_s^\phi\,ds + \int_0^t \sigma(X_s^0)\dot\phi(s)\,ds.$$

Proof See Theorem 2.2 of [3] and Proposition 3.2 of [15], or Theorem 3.1 of [6]. □

Lemma 5 *Under conditions (A2) and (A3), the deterministic dynamical system $X_t^0(\theta)$ is differentiable with respect to θ at the point θ^* in L_2-norm, that is,*

$$\int_0^T \left(X_t^0(\theta^* + h) - X_t^0(\theta^*) - h'\frac{\partial X_t^0}{\partial\theta}\bigg|_{\theta=\theta^*}\right)^2 \alpha(dt) = o(|h|).$$

Proof From Eq. (2), we have

$$
\begin{aligned}
X_t^0(\theta) &= \int_0^t S_s(\theta, X^0(\theta))\,ds \\
&= \int_0^t \left[V(\theta, s, X^0(\theta)) + \int_0^s K(\theta, s, u, X_u^0(\theta))\,du\right]ds.
\end{aligned}
\tag{27}
$$

So

$$
X_t^0(\theta^* + h) - X_t^0(\theta^*)
$$

$$
= \int_0^t S_s(\theta^* + h, X^0(\theta^* + h)) \, ds - \int_0^t S_s(\theta^*, X^0(\theta^*)) \, ds
$$

$$
= \int_0^t \left[V(\theta^* + h, s, X^0(\theta^* + h)) - V(\theta^*, s, X^0(\theta^* + h)) \right] ds
$$

$$
+ \int_0^t \left[V(\theta^*, s, X^0(\theta^* + h)) - V(\theta^*, s, X^0(\theta^*)) \right] ds
$$

$$
+ \int_0^t \left[\int_0^s K(\theta^* + h, s, u, X_u^0(\theta^* + h)) \, du - \int_0^s K(\theta^*, s, u, X_u^0(\theta^* + h)) \, du \right] ds
$$

$$
+ \int_0^t \left[\int_0^s K(\theta^*, s, u, X_u^0(\theta^* + h)) \, du - \int_0^s K(\theta^*, s, u, X_u^0(\theta^*)) \, du \right] ds.
$$

Applying conditions $(A2)$, $(A3)$ and Gronwall's inequality, we have

$$
\sup_{0 \le s \le T} \int_0^s \left(X_u^0(\theta^* + h) - X_u^0(\theta^*) \right)^2 \alpha(du) \le \mathcal{M} T |h| e^T, \tag{28}
$$

where $\mathcal{M}$ is some positive constant. From (27), we also get that

$$
\frac{\partial X_t^0}{\partial \theta} = \int_0^t \frac{\partial S_s(\theta, X^0)}{\partial \theta} \, ds + \int_0^t \frac{\partial S_s(\theta, X^0)}{\partial x} \cdot \frac{\partial X_s^0}{\partial \theta} \, ds
$$

and

$$
X_t^0(\theta^* + h) - X_t^0(\theta^*) - h' \frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta = \theta^*}
$$

$$
= \int_0^t \left[S_s(\theta^* + h, X^0(\theta^* + h)) - S_s(\theta^* + h, X^0(\theta^*)) \right.
$$

$$
\left. - h' \frac{\partial X_s^0}{\partial \theta} \cdot \frac{\partial S_s(\theta, X^0)}{\partial x} \bigg|_{\theta = \theta^*} \right] ds
$$

$$
+ \int_0^t \left[S_s(\theta^* + h, X^0(\theta^*)) - S_s(\theta^*, X^0(\theta^*)) - h' \frac{\partial S_s(\theta, X^0)}{\partial \theta} \bigg|_{\theta = \theta^*} \right] ds
$$

$$
= \mathbf{I}_1 + \mathbf{I}_2. \tag{29}
$$

Taylor's expansion and $(A3)$ imply that $|\mathbf{I}_2| \le \mathcal{M} |h|^2 T$.

For $\mathbf{I}_1$, let

$$
A_s(\theta) = \int_0^1 \frac{\partial}{\partial x} S_s(\theta + h, X_s^0(\theta) + u(X_s^0(\theta + h) - X_s^0(\theta))) \, du,
$$

then

$$
\int_0^T (\mathbf{I}_1)^2 \alpha(dt)
$$

$$
= \int_0^T \left(\int_0^t \left[A_s(\theta^*)(X_s^0(\theta^* + h) - X_s^0(\theta^*)) - h' \frac{\partial X_s^0}{\partial \theta} \frac{\partial S_s(\theta, X^0)}{\partial x} \bigg|_{\theta = \theta^*} \right] ds \right)^2 \alpha(dt)
$$

$$\leq \int_0^T \left[\int_0^t A_s(\theta^*) \left((X_s^0(\theta^*+h) - X_s^0(\theta^*)) - h'\frac{\partial X_s^0}{\partial \theta} \right) ds \right]^2 \alpha(dt)$$

$$+ \int_0^T \left[\int_0^t \left(A_s(\theta^*) - \frac{\partial S_s(\theta+h, X_s^0)}{\partial x} \bigg|_{\theta=\theta^*} \right) h'\frac{\partial X_s^0}{\partial \theta} ds \right]^2 \alpha(dt)$$

$$+ \int_0^T \left[\int_0^t \left(\frac{\partial S_s(\theta+h, X_s^0)}{\partial x} \bigg|_{\theta=\theta^*} - \frac{\partial S_s(\theta, X_s^0)}{\partial x} \bigg|_{\theta=\theta^*} \right) h'\frac{\partial X_s^0}{\partial \theta} ds \right]^2 \alpha(dt)$$

$$= \mathbf{I}_3 + \mathbf{I}_4 + \mathbf{I}_5.$$

Conditions (*A*2) and (*A*3) yield that function $A_s(\theta)$ is bounded for any $s \in [0, T]$. By the differentiability of function $S_t(\theta, x)$ at x for every fixed θ, we get $\mathbf{I}_4 = o(|h|)$ and $\mathbf{I}_5 = o(|h|)$.

On the other hand, we also have

$$\mathbf{I}_3 \leq \mathcal{M} \int_0^T \left(\int_0^t \left[X_s^0(\theta^*+h) - X_s^0(\theta^*) - h'\frac{\partial X_s^0}{\partial \theta} \bigg|_{\theta=\theta^*} \right] ds \right)^2 \alpha(dt).$$

By applying Gronwall's inequality and combining (28) and (29), we get

$$\int_0^T \left(X_t^0(\theta^*+h) - X_t^0(\theta^*) - h'\frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} \right)^2 \alpha(dt) = o(|h|).$$

The proof is completed. $\qquad \square$

Proof of Theorem 1 Noting that

$$\frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} = \left(\frac{\partial X_t^0}{\partial \theta_1}, \dots, \frac{\partial X_t^0}{\partial \theta_p} \right)^T \bigg|_{\theta=\theta^*}$$

and $h(\epsilon) = \sqrt{2 \log \log(\epsilon^{-1} \vee 3)}$, we define

$$G_\epsilon(u) = h(\epsilon) \left[u' \int_0^T \frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} \cdot \left(\frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} \right)' \alpha(dt) \cdot u \right.$$

$$\left. - 2u' \int_0^T \frac{1}{\epsilon h(\epsilon)} (X_t^\epsilon - X_t^0(\theta^*)) \frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} \alpha(dt) \right]. \tag{30}$$

A simple calculation yields

$$\frac{|V_\epsilon(u) - G_\epsilon(u)|}{h(\epsilon)} = \int_0^T \left[\left(\frac{X_t^0(\theta^*+h(\epsilon)\epsilon u) - X_t^0(\theta^*)}{\epsilon h(\epsilon)} \right)^2 - \left(u'\frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} \right)^2 \right] \alpha(dt)$$

$$+ \frac{\lambda_\epsilon}{\epsilon^2 h^2(\epsilon)} \sum_{j=1}^p \{ |\theta_j^* + h(\epsilon)\epsilon u_j|^\gamma - |\theta_j^*|^\gamma \} - 2\int_0^T \frac{X_t^\epsilon - X_t^0(\theta^*)}{\epsilon h(\epsilon)}$$

$$\cdot \left(\frac{X_t^0(\theta^*+h(\epsilon)\epsilon u) - X_t^0(\theta^*)}{\epsilon h(\epsilon)} - u'\frac{\partial X_t^0}{\partial \theta} \bigg|_{\theta=\theta^*} \right) \alpha(dt)$$

$$:= \mathbf{J}_1 + \mathbf{J}_2 - 2\mathbf{J}_3. \tag{31}$$

For the case of $\gamma > 1$, noting that $\epsilon^{-1}\lambda_\epsilon \to \lambda_0$ as $\epsilon \to 0$, we have

$$\mathbf{J}_2 = \frac{\lambda_\epsilon}{\epsilon h(\epsilon)} \sum_{j=1}^{p} u_j \frac{|\theta_j^* + h(\epsilon)\epsilon u_j|^\gamma - |\theta_j^*|^\gamma}{h(\epsilon)\epsilon u_j} \sim \frac{\lambda_0}{h(\epsilon)} \sum_{j=1}^{p} u_j \, \mathrm{sgn}\big(\theta_j^*\big)\big|\theta_j^*\big|^{\gamma-1} \to 0,$$

where $\mathrm{sgn}(x) = 1$ for $x > 0$; $\mathrm{sgn}(x) = -1$ for $x < 0$ and $\mathrm{sgn}(0) = 0$.

For the case of $\gamma = 1$,

$$\mathbf{J}_2 = \frac{\lambda_\epsilon}{\epsilon^2 h^2(\epsilon)} \sum_{j=1}^{p} \big\{\big|\theta_j^* + \epsilon h(\epsilon)u_j\big| - \big|\theta_j^*\big|\big\}$$

$$\sim \frac{\lambda_0}{h(\epsilon)} \sum_{j=1}^{p} \big(|u_j|\mathcal{I}\big\{\theta_j^* = 0\big\} + u_j \, \mathrm{sgn}\big(\theta_j^*\big)\big|\theta_j^*\big|\mathcal{I}\big\{\theta_j^* \neq 0\big\}\big)$$

$$\to 0,$$

where $\mathcal{I}\{\cdot\}$ denotes the indicator function. From Lemma 5, it implies that $\mathbf{J}_1 \to 0$. Thus for any $\delta > 0$ and sufficiently small positive ϵ, we have

$$\mathbf{P}\big(\big|V_\epsilon(u) - G_\epsilon(u)\big| \geq h(\epsilon)\delta\big) \leq \mathbf{P}\bigg(|\mathbf{J}_3| \geq \frac{\delta}{4}\bigg). \tag{32}$$

In addition, Lemma 5 also implies that for any $\kappa > 0$, there exists a positive constant ϵ_0 such that, when $\epsilon < \epsilon_0$,

$$\left|\frac{X_t^0(\theta^* + h(\epsilon)\epsilon u) - X_t^0(\theta^*)}{\epsilon h(\epsilon)} - u'\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta=\theta^*}\right| \leq \kappa.$$

By the *Cauchy–Schwartz inequality*, we have

$$|\mathbf{J}_3|^2 \leq \int_0^T \left(\frac{X_t^\epsilon - X_t^0(\theta^*)}{\epsilon h(\epsilon)}\right)^2 \alpha(dt)$$

$$\cdot \int_0^T \left(\frac{X_t^0(\theta^* + h(\epsilon)\epsilon u) - X_t^0(\theta^*)}{\epsilon h(\epsilon)} - u'\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta=\theta^*}\right)^2 \alpha(dt)$$

$$\leq T\kappa^2 \int_0^T \left(\frac{X_t^\epsilon - X_t^0(\theta^*)}{\epsilon h(\epsilon)}\right)^2 \alpha(dt).$$

Thus, for sufficiently small $\delta > 0$,

$$\mathbf{P}\bigg(|\mathbf{J}_3| \geq \frac{\delta}{4}\bigg) \leq \mathbf{P}\bigg(\bigg(\int_0^T \bigg(\frac{X_t^\epsilon - X_t^0(\theta^*)}{\epsilon h(\epsilon)}\bigg)^2 dt\bigg)^{\frac{1}{2}} \geq \frac{\delta}{4\sqrt{T}\kappa}\bigg). \tag{33}$$

One can easily check that $(A2)$–$(A3)$ imply Assumption (A) of [6]. By Lemma 4, the stochastic process $(X_t^\epsilon - X_t^0(\theta^*))/\epsilon h(\epsilon)$ satisfies the law of iterated logarithm on $C([0, T]; \mathbb{R})$ with the rate function $I(\cdot)$, that is,

$$\limsup_{\epsilon \to 0} \rho\bigg(\frac{X_t^\epsilon - X_t^0(\theta^*)}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \vee 3)}}, K_1\bigg) = 0, \quad \text{a.e.}$$

and

$$\mathbf{P}\left(\omega : \mathbf{C}\left(\left\{\frac{X_t^\epsilon - X_t^0}{\sqrt{2\epsilon \log\log(\epsilon^{-1} \vee 3)}}\right\}\right) = K_1\right) = 1,$$

where $K_1 = \{g : I(g) \leq \frac{1}{2}\}$ and

$$I(g) := \inf_{\phi \in \mathcal{H}: g = Y_t^\phi}\left\{\frac{1}{2}\int_0^T |\dot\phi(s)|^2 \, ds\right\}.$$

Here $\mathcal{H}$ is defined in (12) and Y^ϕ in (13).

The invariance principle (see Theorem 4.3 of [9]) yields that the stochastic process

$$\int_0^T \frac{1}{\epsilon h(\epsilon)}\left(X_t^\epsilon - X_t^0(\theta^*)\right)\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \alpha(dt)$$

satisfies the law of the iterated logarithm with rate function $I^*(\cdot)$, that is,

$$\limsup_{\epsilon \to 0} \rho\left(\int_0^T \frac{1}{\epsilon h(\epsilon)}\left(X_t^\epsilon - X_t^0(\theta^*)\right)\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \alpha(dt), K\right) = 0, \quad \text{a.e.} \tag{34}$$

and

$$\mathbf{P}\left(\omega : \mathbf{C}\left(\left\{\int_0^T \frac{1}{\epsilon h(\epsilon)}\left(X_t^\epsilon - X_t^0(\theta^*)\right)\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \alpha(dt)\right\}\right) = K\right) = 1, \tag{35}$$

where K is defined in (11) and

$$I^*(x) = \inf_{g \in \mathcal{H}}\left\{I(g) : x = \int_0^T g\frac{\partial X_t^0}{\partial \theta}\bigg|_{\theta = \theta^*} \alpha(dt)\right\}.$$

On the other hand, letting $\kappa \to 0$ in (33), we have

$$\limsup_{\epsilon \to 0} \frac{1}{h^2(\epsilon)} \log \mathbf{P}\left(\left(\int_0^T \left(\frac{X_t^\epsilon - X_t^0(\theta^*)}{\epsilon h(\epsilon)}\right)^2 \alpha(dt)\right)^{\frac{1}{2}} \geq \frac{\delta}{4\sqrt{T}}\right) = -\infty.$$

Hence, for any $\delta > 0$ and $R > 0$, there exists an $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0]$,

$$\mathbf{P}\left(\left|V_\epsilon(u) - G_\epsilon(u)\right| \geq h(\epsilon)\delta\right) \leq e^{Rh^2(\epsilon)}. \tag{36}$$

Combining (30) and (34)–(36), by applying Lemma 3, we see that the process $(\widehat{\theta}^\epsilon - \theta^*)/\epsilon h(\epsilon)$ satisfies the law of the iterated logarithm. The proof is completed. $\qquad\square$

5 Conclusion

In this paper, we discussed the convergence rate on the estimators $\widehat{\theta}^\epsilon$ converging to the true value. A simple example was given to test the feasibility of this result. Due to the complexity and minimization of the convex process, we did not conduct an extensive simulation study to identify these stochastic diffusion processes and illustrate the finite performance of the proposed method. Recently, we got to know that some novel approaches of

modeling such as the accurate discretization method [12], two coupled pendulums methods [4], fractional stochastic modeling [2], and fractional discretization [1] were introduced. Those methods can be helpful when dealing with our simulation at some point. This will become an important research direction for us in the future.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the manuscript and typed, read and approved the final manuscript.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Atangana, A.: Fractional discretization: the African's tortoise walk. Chaos Solitons Fractals **130**, Article ID 109399 (2020)
2. Atangana, A., Bonyah, E.: Fractional stochastic modeling: new approach to capture more heterogeneity. Chaos, Interdiscip. J. Nonlinear Sci. **29**, Article ID 013118 (2019)
3. Baldi, P.: Large deviations and functional iterated logarithm law for diffusion processes. Probab. Theory Relat. Fields **71**, 435–453 (1986)
4. Baleanu, D., Jajarmi, A., Asad, J.H.: Classical and fractional aspects of two coupled pendulums. Rom. Rep. Phys. **71**(1), Article ID 103 (2019)
5. Bressloff, P.C.: Stochastic Processes in Cell Biology. Interdisciplinary Applied Mathematics, vol. 41. Springer, New York (2014)
6. Caramellino, L.: Strassen's law of the iterated logarithm for diffusion processes for small time. Stoch. Process. Appl. **74**(1), 1–19 (1998)
7. Dietz, H.M., Kutoyants, Y.A.: A class of minimum-distance estimators for diffusion processes with ergodic properties. Stat. Risk. Model. **15**(3), 211–228 (1997)
8. Freidlin, M.I., Szücs, J., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Grundlehren der mathematischen Wissenschaften, vol. 260. Springer, New York (2012)
9. Gao, F.G., Wang, S.C.: Asymptotic behaviors for functionals of random dynamical systems. Stoch. Anal. Appl. **34**, 258–277 (2015)
10. Geyer, C.J.: On the asymptotics of convex stochastic optimization. Unpublished manuscript (1996)
11. Gregorio, A.D., Iacus, S.M.: On penalized estimation for dynamical systems with small noise. Electron. J. Stat. **12**, 1614–1630 (2018)
12. Hajipour, M., Jajarmi, A., Baleanu, D.: On the accurate discretization of a highly nonlinear boundary value problem. Numer. Algorithms **79**(3), 679–695 (2018)
13. Kato, K.: Asymptotics for argmin processes: convexity arguments. J. Multivar. Anal. **100**, 1816–1829 (2009)
14. Knight, K., Fu, W.J.: Asymptotics for lasso-type estimators. Ann. Stat. **28**, 1356–1378 (2000)
15. Kouritzin, M.A., Heunis, A.J.: A law of the iterated logarithm for stochastic processes defined by differential equations with a small parameter. Ann. Probab. **22**(2), 659–679 (1994)
16. Kutoyants, Y.: Identification of Dynamical Systems with Small Noise. Kluwer Academic, Dordrecht (1994)
17. Kutoyants, Y., Pilibossian, P.: On minimum L_1-norm estimate of the parameter of the Ornstein–Uhlenbeck process. Stat. Probab. Lett. **20**(2), 117–123 (1994)
18. Kutoyants, Y., Pilibossian, P.: On minimum uniform metric estimate of parameters of diffusion-type processes. Stoch. Process. Appl. **51**(2), 259–267 (1994)
19. Liptser, R.S., Shiryayev, A.N.: Statistics of Random Processes, vol. I. Springer, New York (1977)
20. Liptser, R.S., Shiryayev, A.N.: Statistics of Random Processes, vol. II. Springer, New York (1978)
21. Nkurunziza, S.: Shrinkage strategies in some multiple multi-factor dynamical systems. ESAIM, Probab. Stat. **16**, 139–150 (2012) https://doi.org/10.1051/ps/2010015
22. Zhao, H., Zhang, C.: Minimum distance parameter estimation for SDEs with small α-stable noises. Stat. Probab. Lett. **145**, 301–311 (2019)