Parameter estimation of stochastic differential equation driven by small fractional noise

Shohei Nakajima and Yasutaka Shimizu

Department of Applied Mathematics, Waseda University, Shinjuku City, Tokyo, Japan

ABSTRACT
We study the problem of parametric estimation for continuously observed stochastic processes driven by additive small fractional Brownian motion with the Hurst index $H \in (0, 1)/\{1/2\}$. Under some assumptions on the drift coefficient, we obtain the asymptotic normality and moment convergence of maximum likelihood estimator of the drift parameter when a small dispersion coefficient $\varepsilon \to 0$.

ARTICLE HISTORY
Received 17 November 2021
Accepted 4 July 2022

KEYWORDS
Parameter estimation; stochastic differential equation; fractional Brownian motion; small noise; asymptotic normality

1. Introduction

Let $\{X_t^\varepsilon\}_{t \in [0, T]}$ be a solution to the following stochastic differential equation:

$$X_t^\varepsilon = X_0 + \int_0^t b(X_s^\varepsilon, \theta_0) \, ds + \varepsilon B_t^H, \quad t \in (0, T],$$

where $X_0 \in \mathbb{R}$ is the initial value, $\{B_t^H\}_{t \in [0, T]}$ is a fractional Brownian motion with Hurst index $H \in (0, 1)/\{1/2\}$ and $\theta_0 \in \Theta$ is the parameter which is contained in a bounded and open convex subset $\Theta \subset \mathbb{R}^d$ admitting Sobolev’s inequalities for embedding $W^{1,p}(\Theta) \hookrightarrow C(\Theta)$. Without loss of generality, we assume that $\varepsilon \in (0, 1]$. The main purpose of this paper is the estimation of parameter $\theta_0 \in \Theta$ from a realization $\{X_t^\varepsilon\}_{t \in [0, T]}$ when $\varepsilon \to 0$. In the case where $H = 1/2$, that is, $\{B_t^H\}_{t \in [0, T]}$ is a Brownian motion, estimation problems have been studied by many authors. In particular, the maximum likelihood estimator (MLE) via the likelihood function based on the Girsanov density is the one of the optimal methods for estimation (see [1–3]).

There are several results on the parametric inference for stochastic differential equation driven by fractional Brownian motion. Brouste and Kleptsyna [4] and Kleptsyna and Le Breton [5] studied the parameter estimation problem for continuously observed fractional Ornstein–Uhlenbeck processes. In these papers, the drift function is linear in both $x$ and $\theta$ ($b(x, \theta) = -\theta x$) and the asymptotic normality and moment convergence of MLE are established when the terminal time of observation goes to infinity. In a similar framework, Tudor and Viens [6] discussed the statistical estimation with special drift function $b(x, \theta) = \theta b(x)$.
and they showed the consistency of the MLE. All the drift functions discussed in these papers are linear in $\theta$ and the MLE has an explicit expression. Recently, in the case when the drift function $b(x, \theta)$ is nonlinear in both $x$ and $\theta$, Chiba [7] proposed an M-estimator based on the likelihood function. Unlike previous studies, the estimator proposed in [7] does not have an explicit expression. In order to obtain the asymptotic properties of the estimator, he applied the method investigated by Ibragimov and Has’minskii [8] and established asymptotic properties of the estimator when the Hurst index $H$ is contained in $(1/4, 1/2)$. Their approach is based on the analysis of the likelihood ratio random field, where the large deviation inequality plays an important role to derive the asymptotic properties.

The parametric inference for diffusion processes with small white noise has been well developed (see, e.g. [9–13]). However, parametric estimation problems for the stochastic differential equation driven by small fractional Brownian motion has not been analysed yet. In this paper, we allow the drift function $b(x, \theta)$ to be nonlinear in both $x$ and $\theta$ and the Hurst index $H$ is contained in $(0, 1)/(1/2)$. We aim to deduce asymptotic normality and moment convergence of the MLE of the drift parameter under $\varepsilon \to 0$ in the spirit of Ibragimov and Has’minskii as in [7].

This paper is organized as follows: in Section 2, we make some notations and assumptions to state our main results. In Section 3, we prove main results. Most of the proof is checking out the sufficient conditions of the polynomial type large deviation inequality investigated by Yoshida [14].

2. Main results

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We assume that the parameter space $\Theta \subset \mathbb{R}^d$ to be bounded, open and convex domain admitting Sobolev embedding $W^{1,p}(\Theta) \hookrightarrow C(\bar{\Theta})$ for $p > d$. Here, $C(\bar{\Theta})$ is the set of continuous functions on $\bar{\Theta}$ and $W^{1,p}(\Theta)$ is the set of functions $f$ on $\Theta$ such that $f$ and its derivative in the weak sense are $L^p$ integrable functions. We aim to estimate the unknown parameter $\theta_0 \in \Theta$ in Equation (1) from the completely observed data $\{X_t^\varepsilon\}_{t \in [0,T]}$. Let us define some functions appearing in the likelihood function for Equation (1). We first recall the basic definitions of fractional calculus. Let $f \in L^1(a, b)$ for $a < b$ and $\alpha > 0$. The fractional Riemann–Liouville integrals of $f$ of order $\alpha$ are defined for almost all $x \in (a, b)$ by

$$I_{a+}^\alpha f(x) := \frac{1}{\Gamma(\alpha)} \int_a^x (x - y)^{\alpha - 1} f(y) \, dy$$

and

$$I_{b-}^\alpha f(x) := \frac{1}{\Gamma(\alpha)} \int_x^b (y - x)^{\alpha - 1} f(y) \, dy.$$  

Let $I_{a+}^\alpha (L^p(a, b))$ (resp., $I_{b-}^\alpha (L^p(a, b))$) be the image of $L^p(a, b)$ by the operator $I_{a+}^\alpha (a, b)$ (resp., $I_{b-}^\alpha (L^p(a, b))$) if $f \in I_{a+}^\alpha (L^p(a, b))$ (resp., $I_{b-}^\alpha (L^p(a, b))$) and $0 < \alpha < 1$, then the Weyl derivative are defined by

$$D_{a+}^\alpha f(x) := \frac{1}{\Gamma(1 - \alpha)} \left( \frac{f(x)}{(x - a)^\alpha} + \alpha \int_a^x \frac{f(x) - f(y)}{(x - y)^{\alpha+1}} \, dy \right).$$
where we define an operator $D^\alpha_t f(x) := \frac{1}{\Gamma(1-\alpha)} \left( \frac{f(x)}{(b-x)^\alpha} + \alpha \int_x^b \frac{f(x) - f(y)}{(x-y)^{\alpha+1}} \, dy \right)$. 

We impose some assumptions on coefficient $b$ to derive the likelihood function.

**Assumption 2.1:** The function $b$ in (1) is of $C^{1,4}(\mathbb{R} \times \Theta; \mathbb{R})$-class such that for every $x \in \mathbb{R}$ and $\theta \in \Theta$, the following growth conditions hold:

$$|b(x, \theta)| \leq c(1 + |x|), \quad |\nabla^i b(x, \theta)| \leq c(1 + |x|^N), \quad |\nabla^i \partial_x b(x, \theta)| \leq c(1 + |x|^N),$$

for $0 \leq i \leq 4$ and some constants $c > 0, N \in \mathbb{N}$.

**Assumption 2.2:** There exists $L > 0$ such that for every $x, y \in \mathbb{R}$,

$$\sup_{\theta \in \Theta} |b(x, \theta) - b(y, \theta)| \leq L|x - y|.$$

According to [15,16], the existence and uniqueness of the strong solution to Equation (1) follows under Assumptions 2.1 and 2.2. In addition, for every $0 < \varepsilon < H$, the solution to (1) has $H - \varepsilon$ Hölder continuity. From the Hölder continuity of the solution to (1), we can define the function (see Theorem 13.6 in [17])

$$Q_{H,\theta}^\varepsilon(t) := \begin{cases} \varepsilon d_H^{-1} t^{H - 1 / 2} \mathbb{I}_{0 < H < 1 / 2} \left[ (\cdot)^{1/2-H} b(X^\varepsilon_s, \theta) \right](t) & \text{if } H < 1/2, \\ \varepsilon d_H^{-1} t^{H - 1 / 2} \mathbb{I}_{H > 1 / 2} \left[ (\cdot)^{1/2-H} b(X^\varepsilon_s, \theta) \right](t) & \text{if } H > 1/2, \end{cases}$$

where

$$d_H := \sqrt{\frac{2H(\frac{3}{2} - H) \Gamma(H + \frac{1}{2})}{\Gamma(2 - 2H)}}.$$

For $0 < s < t$, let $k_{H}^{-1}(t,s)$ be a function given by

$$k_{H}^{-1}(t,s) := \begin{cases} \frac{1}{d_H} (s^{1/2-H} \mathbb{I}_{H < 1 / 2} \left[ (\cdot)^{H - 1/2} \right](s)) & \text{if } H < 1/2, \\ \frac{1}{d_H} (s^{1/2-H} \mathbb{I}_{H > 1 / 2} \left[ (\cdot)^{H - 1/2} \right](s)) & \text{if } H > 1/2. \end{cases}$$

We define a semimartingale $\{Z_t\}_{t \geq 0}$ as follows:

$$Z_t := \varepsilon^{-1} \int_0^t k_{H}^{-1}(t,s) \, dX^\varepsilon_s,$$

$$= \int_0^t Q_{H,\theta}^\varepsilon(s) \, ds + W_t,$$

where $\{W_t\}_{0 \leq t \leq T}$ is a Wiener process. Note that we used the Volterra correspondence

$$W_t = \int_0^t k_{H}^{-1}(t,s) \, dB^H_s.$$

Here, we interpret the stochastic integral with respect to a fractional Brownian motion as a Wiener integral. By the Girsanov theorem, the log-likelihood function $\mathbb{L}_{H,\varepsilon}$ for
Equation (1) can be obtained by

\[
\mathbb{L}_{H, \epsilon}(\theta) := \int_{0}^{T} Q_{H, \epsilon}^{\theta}(t) \, dZ_{t} - \frac{1}{2} \int_{0}^{T} Q_{H, \epsilon}^{\theta}(t)^{2} \, dt.
\]

For more details about construction of the likelihood function, see [6]. We define the MLE by

\[
\hat{\theta}_{\epsilon} := \arg\max_{\theta \in \Theta} \mathbb{L}_{H, \epsilon}(\theta).
\]

In order to state our main results about asymptotic properties of \( \hat{\theta}_{\epsilon} \), we make some notations. Let \( \{x_{t}\}_{0 \leq t \leq T} \) be the solution to the differential equation under the true value of the drift parameter:

\[
\begin{align*}
\frac{dx_{t}}{dt} &= b(x_{t}, \theta_{0}), \\
x_{0} &= X_{0}.
\end{align*}
\]

We set the \( d \)-dimensional square matrix \( \Gamma_{H}(\theta_{0}) \) which is an asymptotic covariance matrix of our estimator as

\[
\Gamma_{H}^{ij}(\theta_{0}) := \begin{cases}
\c_{1} \int_{0}^{T} t^{2H-1} \left( \int_{0}^{t} s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_{j}} b(x_{s}, \theta_{0}) \, ds \right) \times \left( \int_{0}^{t} s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_{i}} b(x_{s}, \theta_{0}) \, ds \right) \, dt & \text{if } H < 1/2, \\
\int_{0}^{T} \left( c_{2} t^{1/2-H} \partial_{\theta_{j}} b(x_{t}, \theta_{0}) + c_{3} t^{H-1/2} \partial_{\theta_{j}} b(x_{t}, \theta_{0}) \right) dt & \text{if } H > 1/2,
\end{cases}
\]

where

\[
\begin{align*}
c_{1} &= (d_{H} \Gamma(1/2 - H))^{-2}, \\
c_{2} &= (d_{H} \Gamma(3/2 - H))^{-1} \left\{ 1 + (H - 1/2) \int_{0}^{1} \frac{1 - s^{1/2-H}}{(1 - s)^{H+1/2}} \, ds \right\}, \\
c_{3} &= (H - 1/2)(d_{H} \Gamma(3/2 - H))^{-1}.
\end{align*}
\]

**Assumption 2.3:** The matrix \( \Gamma_{H}(\theta_{0}) \) is positive definite.
In order to guarantee the asymptotic properties of the estimator, we need to impose the identifiability condition. Define

\[ \mathbb{Y}_{H, \varepsilon}(\theta) := \varepsilon^2 (\mathbb{L}_{H, \varepsilon}(\theta) - \mathbb{L}_{H, \varepsilon}(\theta_0)), \]

and let \( \mathbb{Y}_H \) be the expected limit of \( \mathbb{Y}_{H, \varepsilon} \) defined by

\[
\mathbb{Y}_H(\theta) := \begin{cases} 
- \frac{c_1}{2} \int_0^T t^{2H-1} \left( \int_0^t s^{1/2-H} (t-s)^{-1/2-H} b(x_t, \theta) - b(x_t, \theta_0) \right)^2 \, dt & \text{if } H < 1/2, \\
- \frac{1}{2} \int_0^T \left( c_2 t^{1/2-H} (b(x_t, \theta) - b(x_t, \theta_0)) \right)^2 \, dt & \text{if } H = 1/2, \\
+ c_3 t^{H-1/2} \int_0^t \left( (b(x_t, \theta) - b(x_t, \theta_0)) - (b(x_s, \theta) - b(x_s, \theta_0)) \right) s^{1/2-H} \, ds \right)^2 \, dt & \text{if } H > 1/2.
\]

**Assumption 2.4:** There exists a positive constant \( \xi(\theta_0) > 0 \) such that

\[ \mathbb{Y}_H(\theta) \leq -\xi(\theta_0)|\theta - \theta_0|^2, \]

for every \( \theta \in \Theta \).

The following theorem gives the asymptotic properties of the estimator \( \hat{\theta}_\varepsilon \).

**Theorem 2.1:** Suppose that Assumptions 2.1–2.4 are fulfilled. Then the estimator \( \hat{\theta}_\varepsilon \) satisfies that

\[ \varepsilon^{-1} (\hat{\theta}_\varepsilon - \theta_0) \xrightarrow{d} N(0, \Gamma_H(\theta_0)^{-1}), \]

as \( \varepsilon \to 0 \). Moreover, we have

\[ E[f(\varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0))] \to E[f(\xi)], \]

as \( \varepsilon \to 0 \) for every continuous function \( f \) of polynomial growth, where \( \xi \sim N(0, \Gamma_H(\theta_0)^{-1}) \).

**Example 2.1:** We consider a one-dimensional fractional Ornstein–Uhlenbeck process, that is, the drift function in (1) is given by \( b(x, \theta_0) = \theta_0 x \) with \( \theta_0 \in \mathbb{R} \). Then \( x_t \) satisfies the following equation:

\[
\begin{aligned}
\frac{dx_t}{dt} &= \theta_0 x_t, \\
x_0 &= X_0.
\end{aligned}
\]

The explicit solution is given by \( x_t = X_0 e^{\theta_0 t} \). In this case, we can check Assumptions 2.1–2.4 hold true. Indeed, Assumptions 2.1 and 2.2 are obvious. We will show that Assumptions 2.3 and 2.4 hold. In the case \( H < 1/2 \),

\[
\Gamma_H(\theta_0) = c_1 \int_0^T t^{2H-1} \left( \int_0^t s^{1/2-H} (t-s)^{-1/2-H} X_0 e^{\theta_0 s} \, ds \right)^2 \, dt \\
\geq c_1 X_0^2 (e^{\theta_0 T} \wedge 1)^2 \int_0^T t^{2H-1} \left( \int_0^t s^{1/2-H} (t-s)^{-1/2-H} \, ds \right)^2 \, dt \\
= c_1 X_0^2 (e^{\theta_0 T} \wedge 1)^2 \beta(3/2 - H, 1/2 - H)^2 T^{2-2H} > 0.
\]
Thus, Assumption 2.3 holds true. In the same way,

\[-Y_H(\theta) = \frac{c_1}{2} \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H} (t-s)^{-1/2-H} (\theta_0 x_s - \theta x_s) \, ds \right\}^2 \, dt \]

\[= \frac{c_1}{2} (\theta_0 - \theta)^2 X_0^2 \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H} (t-s)^{-1/2-H} e^{\theta_0 s} \, ds \right\}^2 \, dt \]

\[\geq \frac{c_1}{2} (\theta_0 - \theta)^2 X_0^2 (e^{\theta_0 T} \wedge 1)^2 \beta(3/2 - H, 1/2 - H) T^{2-2H}, \]

and Assumption 2.4 holds true. In the case \( H > 1/2 \), we restrict \( \theta_0 > 0 \). Then

\[-Y_H(\theta) = \frac{(\theta - \theta_0)^2}{2} X_0^2 \int_0^T \left\{ c_2 t^{1/2-H} e^{\theta_0 t} + c_3 t^{H-1/2} \int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right\}^2 \, dt \]

\[= \frac{(\theta - \theta_0)^2}{2} X_0^2 \left\{ c_2^2 \int_0^T t^{1-2H} e^{2\theta_0 t} + 2c_2c_3 e^{\theta_0 t} \theta_0 t^{2-2H} \beta(3/2 - H, 3/2 - H) + c_3^2 t^{2H-1} \left( \int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right)^2 \, dt \right\}. \]

By the mean value theorem, we have

\[\int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \geq \theta_0 \int_0^t (t-s)^{1/2-H} s^{1/2-H} e^{\theta_0 s} \, ds \]

\[\geq \theta_0 t^{2-2H} \beta(3/2 - H, 3/2 - H). \]

Therefore,

\[-Y_H(\theta) \geq \frac{(\theta - \theta_0)^2}{2} X_0^2 \left\{ c_2^2 \int_0^T t^{1-2H} e^{2\theta_0 t} + 2c_2c_3 e^{\theta_0 t} \theta_0 t^{2-2H} \beta(3/2 - H, 3/2 - H) + c_3^2 \theta_0^2 t^{3-2H} \beta(3/2 - H, 3/2 - H) \right\} \]

\[\geq \frac{(\theta - \theta_0)^2}{2} X_0^2 \left\{ \frac{c_2^2}{2-2H} T^{2-2H} + \frac{2c_2c_3 \theta_0}{3-2H} T^{3-2H} \beta(3/2 - H, 3/2 - H) + \frac{c_3^2 \theta_0^2}{4-2H} T^{4-2H} \beta(3/2 - H, 3/2 - H) \right\}, \]

and Assumption 2.4 holds. Assumption 2.3 can be confirmed by the same calculation.

**Example 2.2:** Let us consider the drift function

\[b(x, \theta) = \sqrt{\theta + x^2}, \]

with \( \theta \in (m, M), \ 0 < m < M, \ X_0 > 0 \) and \( H \in (0, 1/2) \). Then, we can confirm to \( b \in C^{1,4}(\mathbb{R} \times \Theta) \) and all derivative functions are bounded. Thus, Assumptions 2.1 and 2.2
hold true. We check Assumption 2.4. Note that the function $x_t$ is monotone increasing and satisfies the relation

$$x_t + \sqrt{\theta_0 + x_t^2} = (X_0 + \sqrt{\theta_0 + X_0^2}) e^t.$$  

In particular, for every $t \in [0, T]$, $X_0 \leq x_t < (X_0 + \sqrt{\theta_0 + X_0^2}) e^t$. By the mean value theorem,

$$-\mathbb{E}_H(\theta) = \frac{c_1}{2} \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H} (t-s)^{-1/2-H} (\sqrt{\theta_0 + x_s^2} - \sqrt{\theta + x_t^2}) \, ds \right\}^2 \, dt$$

$$\geq \frac{c_1}{8}(\theta_0 - \theta)^2 \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H} (t-s)^{-1/2-H} \frac{1}{\sqrt{M + x_s^2}} \, ds \right\}^2 \, dt$$

$$\geq \frac{c_1}{8(M + x_0^2)}(\theta_0 - \theta)^2 \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H} (t-s)^{-1/2-H} \, ds \right\}^2 \, dt$$

$$\geq (\theta_0 - \theta)^2 \frac{c_1 \beta (3/2 - H, 1/2 - H) 2^{2-2H}}{8(M + (X_0 + \sqrt{\theta_0 + X_0^2}) e^2 T)},$$

and Assumption 2.4 holds. Assumption 2.3 can be confirmed by the same calculation.

**Example 2.3:** Let $H \in (1/2, 1)$. We consider a simpler drift function than Example 2.2 which is given by

$$b(x, \theta) = \theta \sqrt{1 + x^2},$$

with $\theta \in (m, M)$, $0 < m < M$ and $X_0 > 0$. As in Example 2.2, Assumptions 2.1 and 2.2 can be checked and the function $x_t$ satisfies the relation $x_t + \theta_0 \sqrt{1 + x_t^2} = (X_0 + \theta_0 \sqrt{1 + X_0^2}) e^t$. We check Assumption 2.4 holds true.

$$-\mathbb{E}_H(\theta) = \frac{1}{2} \int_0^T \left( c_2 t^{1/2-H} (\theta \sqrt{1 + x_t^2} - \theta \sqrt{1 + x_t^2}) + c_3 t^{H-1/2} \int_0^t (\theta_0 \sqrt{1 + x_s^2} - \theta \sqrt{1 + x_t^2} - \theta_0 \sqrt{1 + x_t^2} - \theta_0 \sqrt{1 + x_0^2}) (t-s)^{H+1/2} \, ds \right) \, dt$$

$$= (\theta_0 - \theta)^2 \frac{1}{2} \int_0^T (c_2 t^{1/2-H} \sqrt{1 + x_t^2})^2 + 2 c_2 c_3 \sqrt{1 + x_t^2} \int_0^t \sqrt{1 + x_s^2} - \sqrt{1 + x_t^2} (t-s)^{H+1/2} \, ds$$

$$+ \left( c_3 t^{H-1/2} \int_0^t \sqrt{1 + x_s^2} - \sqrt{1 + x_t^2} (t-s)^{H+1/2} \, ds \right) \, dt.$$
We evaluate each of the three terms that appear in the last equality in (3). Using the monotonicity of $x_t$, the first term can be estimated as

$$
\int_0^T t^{1-2H} (1 + x_t^2) \, dt \geq (1 + X_0^2) \int_0^T t^{1-2H} \, dt = \frac{1 + X_0^2}{2 - 2H} T^{2-2H}.
$$

Let us estimate the second term of (3). By the change of the variable formula, we have

$$
\sqrt{1 + x_t^2} - \sqrt{1 + x_s^2} = \int_s^t \frac{x_u}{\sqrt{1 + x_u^2}} \, dx_u
= \theta_0 \int_s^t x_u \sqrt{1 + x_u^2} \frac{1}{\sqrt{1 + x_u^2}} \, du \geq \theta_0 X_0 (t - s).
$$

Thus,

$$
\int_0^T \sqrt{1 + x_t^2} \int_0^t \frac{\sqrt{1 + x_t^2} - \sqrt{1 + x_s^2}}{(t-s)^{H+1/2}} \, ds \, dt
\geq \sqrt{1 + X_0^2} \theta_0 X_0 \int_0^T \int_0^t (t-s)^{1/2-H} s^{1/2-H} \, ds \, dt
= \sqrt{1 + X_0^2} \theta_0 X_0 \beta (3/2 - H, 3/2 - H) T^{3-2H}.
$$

In the same way, we can estimate the third term of (3) as

$$
\int_0^T t^{2H-1} \left( \int_0^t \frac{\sqrt{1 + x_t^2} - \sqrt{1 + x_s^2}}{(t-s)^{H+1/2}} \, ds \right)^2 \, dt
\geq \theta_0^2 X_0^2 \int_0^T t^{2H-1} \left( \int_0^t (t-s)^{1/2-H} s^{1/2-H} \, ds \right)^2 \, dt
= \frac{\theta_0^2 X_0^2 \beta(3/2 - H, 3/2 - H)^2}{4 - 2H} T^{4-2H},
$$

and Assumption 2.4 is valid. Assumption 2.3 can be confirmed by the same calculation.

### 3. Proofs

We first establish a lemma that is used frequently in this paper. Recall that $\{x_t\}_{0 \leq t \leq T}$ be the solution to the following differential equation:

$$
\begin{cases}
\frac{dx_t}{dt} = b(x_t, \theta_0), \\
x_0 = X_0.
\end{cases}
$$

**Lemma 3.1:** For every $p > 0$, there exist constants $c_i > 0$, $i = 1, 2, 3$ such that for every $s, t \in [0, T]$,

$$
E|X_t^e - x_t|^p \leq c_1 e^p,
$$
\[ E|X^\varepsilon_t|^p \leq c_2, \]

and

\[ E|X^\varepsilon_t - X^\varepsilon_s|^p \leq c_3|t - s|^{pH}. \]

**Proof:** By Assumptions 2.2,

\[
|X^\varepsilon_t - x_t| \leq \int_0^t |b(X^\varepsilon_s, \theta_0) - b(x_s, \theta_0)| + \varepsilon |B^H_t| \\
\leq L \int_0^t |X^\varepsilon_s - x_s| \, ds + \varepsilon \sup_{0 \leq t \leq T} |B^H_t|. 
\]

By Gronwall’s inequality, it follows that

\[
|X^\varepsilon_t - x_t| \leq \varepsilon e^{Lt} \sup_{0 \leq t \leq T} |B^H_t|, 
\]

and the first estimate follows. Other estimates hold true by the linear growth condition of the function \( b \).

To show the asymptotic property of estimator \( \hat{\theta}_\varepsilon \), we apply the polynomial type large deviation inequality investigated by Yoshida [14]. Let \( \mathbb{U}_\varepsilon(\theta_0) := \{ u \in \mathbb{R}^d : \theta_0 + \varepsilon u \in \Theta \} \) and define the random field \( \mathbb{Z}_{H,\varepsilon} : \mathbb{U}_\varepsilon(\theta_0) \to \mathbb{R}_+ \) by

\[
\mathbb{Z}_{H,\varepsilon}(u) = \exp \left\{ \mathbb{L}_{H,\varepsilon}(\theta_0 + \varepsilon u) - \mathbb{L}_{H,\varepsilon}(\theta_0) \right\}, \quad u \in \mathbb{U}_\varepsilon(\theta_0).
\]

Applying Taylor’s formula, we have

\[
\log \mathbb{Z}_{H,\varepsilon}(u) = \varepsilon \nabla_\theta \mathbb{L}_{H,\varepsilon}(\theta_0)[u] - \frac{1}{2} u \Gamma_H(\theta_0) u^\top + R_\varepsilon(u),
\]

where

\[
R_\varepsilon(u) = \frac{1}{2} u (\varepsilon^2 \nabla^2_\theta \mathbb{L}_{H,\varepsilon}(\theta_0) - (-\Gamma_H(\theta_0))) u^\top
\]

\[
+ \frac{1}{2} \varepsilon^3 \int_0^1 (1 - s)^2 \nabla^3_\theta \mathbb{L}_{H,\varepsilon}(\theta_0 + s\varepsilon u)[u, u, u] \, ds
\]

and \( \nabla^3_\theta \mathbb{L}_{H,\varepsilon}(\theta)[u, v, w] = \sum_{i, j, k} \partial_i \partial_j \partial_k \mathbb{L}_{H,\varepsilon}(\theta) u_i v_j w_k. \)

**Remark 3.1:** The log-likelihood function \( \mathbb{L}_{H,\varepsilon} \) is differentiable in \( \theta \) under Assumption 2.1, and we have

\[
\nabla_\theta \mathbb{L}_{H,\varepsilon}(\theta) = \int_0^T \nabla_\theta Q^\varepsilon_{H,\theta}(t) \, dZ_t - \int_0^T Q^\varepsilon_{H,\theta}(t) \nabla_\theta Q^\varepsilon_{H,\theta}(t) \, dt,
\]

\[
\nabla^2_\theta \mathbb{L}_{H,\varepsilon}(\theta_0) = \int_0^T \nabla^2_\theta Q^\varepsilon_{H,\theta}(t) \, dZ_t - \int_0^T (\nabla_\theta Q^\varepsilon_{H,\theta}(t))^\top \, dt - \int_0^T Q^\varepsilon_{H,\theta}(t) \nabla^2_\theta Q^\varepsilon_{H,\theta}(t) \, dt.
\]

Throughout this paper, we will use the following notations:
Notation 3.1: For any \( a, b \geq 0 \), the symbol \( a \lesssim b \) means that there exists a universal constant \( C > 0 \) such that \( a \leq Cb \). When \( C \) depends explicitly on a specific quantity, we shall indicate it explicitly through the paper.

The following lemma is one of the sufficient conditions for polynomial type large deviation inequality investigated by Yoshida [14].

Lemma 3.2: For every \( p > 0 \),

\[
\sup_{0 < \varepsilon < 1} E[(\varepsilon^{-d_H} | e^{2\nabla^2_{\theta}} L_{H,\varepsilon}(\theta_0) - (-\Gamma_H(\theta_0))|^p)] < \infty,
\]

where

\[
d_H = \begin{cases} 
1 & \text{if } H < 1/2, \\
1/2 & \text{if } H > 1/2.
\end{cases}
\]

Proof: By Remark 3.1, it follows that

\[
e^{2\nabla^2_{\theta}} L_{H,\varepsilon}(\theta_0) - (-\Gamma_H(\theta_0))
= e^{2} \int_{0}^{T} \nabla^2_{\theta} Q^\varepsilon_{H,\theta_0}(t) \, dW_t - \left( e^{2} \int_{0}^{T} (\nabla_{\theta} Q^\varepsilon_{H,\theta_0}(t))^{\otimes 2} \, dt - \Gamma_H(\theta_0) \right).
\]

At first, we consider the case of \( H < 1/2 \). Note that

\[
Q^\varepsilon_{H,\theta_0}(t) = c^2 \varepsilon^{-1} t^{H-1/2} \int_{0}^{t} s^{1/2-H} (t-s)^{-1/2-H} b(X_{s}^\varepsilon, \theta_0) \, ds.
\]

By Burkholder’s and Minkowski’s inequalities and Lemma 3.1, the stochastic integral part is estimated as

\[
E \left( e^{2} \int_{0}^{T} \partial_{i}\partial_{j} Q^\varepsilon_{H,\theta_0}(t) \, dW_t \right)^p \lesssim \left( e^{4} \int_{0}^{T} \left\| \partial_{i}\partial_{j} Q^\varepsilon_{H,\theta_0}(t) \right\|_{L^p(\Omega)}^2 \, dt \right)^{p/2}
\lesssim e^p \left( \int_{0}^{T} t^{2H-1} \left\| \int_{0}^{t} s^{1/2-H} (t-s)^{-1/2-H} \left\| \partial_{i}\partial_{j} b(X_{s}^\varepsilon, \theta_0) \right\|_{L^p(\Omega)} \, ds \right\|^{2} \, dt \right)^{p/2}
\lesssim e^p \sup_{0 \leq s \leq T} \| 1 + |X_{s}^\varepsilon |^N \|_{L^p(\Omega)}^{p} \left( \int_{0}^{T} t^{1-2H} \, dt \right)^{p/2} \lesssim e^p,
\]

for every \( i, j = 1, \ldots, d \). We shall estimate the second part. For every \( i, j = 1, \ldots, d \),

\[
e^{2} \int_{0}^{T} \partial_{i} Q^\varepsilon_{H,\theta_0}(t) \partial_{j} Q^\varepsilon_{H,\theta_0}(t) \, dt - \Gamma_H(\theta_0)
= c^2 \int_{0}^{T} t^{2H-1} \left( \int_{0}^{t} s^{1/2-H} (t-s)^{-1/2-H} \partial_{i} b(X_{s}^\varepsilon, \theta_0) \, ds \right)
\times \left( \int_{0}^{t} s^{1/2-H} (t-s)^{-1/2-H} \partial_{j} b(X_{s}^\varepsilon, \theta_0) \, ds \right)
\]
By Hölder’s and Minkowski’s inequalities and Lemma 3.1, we can show that

\[
\begin{align*}
&\mathbb{E} \left[ \int_0^T t^{2H-1} \left( \int_0^t s^{1/2-H} (t-s)^{-1/2-H} \partial_\theta (b(x_s, \theta_0) - b(x_s, \theta_0)) \, ds \right) \right. \\
&\left. \times \left( \int_0^t s^{1/2-H} (t-s)^{-1/2-H} \partial_\theta (b(X_s^e, \theta_0) - b(X_s^e, \theta_0)) \, ds \right) \right] \\
&\leq \sup_{0 \leq s \leq T} \| (1 + |X_s|^N + |x_s|^N) |X_s^e - x_s|^p \|_{L^2(\Omega)} \left( \int_0^T t^{1-2H} \, dt \right) \lesssim \varepsilon^p.
\end{align*}
\]

Therefore,

\[
\sup_{0 < \varepsilon < 1} \mathbb{E} [(\varepsilon^{-1} |\varepsilon^2 \nabla_\theta \mathbb{L}_{H, \varepsilon}(\theta_0) - (-\Gamma_H(\theta_0)))^p] < \infty.
\]

The case where \( H > 1/2 \). We have that

\[
Q_{H, \theta_0}^e (t) = c_1 \varepsilon^{-1} t^{1/2-H} b(X_t^e, \theta_0) + c_2 \varepsilon^{-1} t^{H-1/2} \int_0^t \frac{b(X_s^e, \theta_0) - b(X_s^e, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} \, ds.
\]

In a similar way with the case \( H < 1/2 \), the stochastic integral part is evaluated as

\[
\begin{align*}
&\mathbb{E} \left[ \varepsilon^2 \left| \int_0^T \partial_\theta \partial_\theta Q_{H, \theta_0}^e (t) \, dW_t \right| \right]^p \\
&\lesssim \left( \varepsilon^4 \int_0^T \left\| \partial_\theta \partial_\theta Q_{H, \theta_0}^e (t) \right\|_{L^2(\Omega)}^2 \, dt \right)^{p/2} \\
&\lesssim \varepsilon^p \left\{ \int_0^T t^{1-2H} \left\| \partial_\theta \partial_\theta b(X_t^e, \theta_0) \right\|_{L^2(\Omega)}^2 \, dt \\
&+ \int_0^T t^{2H-1} \left\| \int_0^t \frac{\partial_\theta \partial_\theta b(X_s^e, \theta_0) - \partial_\theta \partial_\theta b(X_s^e, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right\|_{L^2(\Omega)}^2 \, dt \right\}^{p/2}
\end{align*}
\]
\[
\lesssim \varepsilon^p \left\{ \left\| 1 + \sup_{0 \leq t \leq T} E[X_t^\varepsilon]^{2N} \right\|_{L^p(\Omega)}^2 \int_0^T t^{1-2H} \, dt \right\}^{p/2} \\
+ \int_0^T t^{2H-1} \left\| \int_0^t \frac{1 + |X_s^\varepsilon|^N + |X_t^\varepsilon|^N}{(t-s)^{H+1/2}} \left| X_t^\varepsilon - X_s^\varepsilon \right|^{1/2-H} \, ds \right\|_{L^p(\Omega)}^2 \, dt \right\}^{p/2} \\
\lesssim \varepsilon^p \left\{ 1 + \int_0^T t^{2H-1} \left( \int_0^t \left\| X_t^\varepsilon - X_s^\varepsilon \right\|_{L^p(\Omega)}^{1/2-H} \, ds \right)^2 \, dt \right\}^{p/2} \lesssim \varepsilon^p,
\]

for every \( i, j = 1, \ldots, d \). We estimate the term \( \varepsilon^2 \int_0^T (\nabla_0 Q_{H, \theta_0}^\varepsilon(t))^2 \, dt - \Gamma_H(\theta_0) \). For every \( i, j = 1, \ldots, d \),

\[
\varepsilon^2 \int_0^T \partial_{\theta_i} Q_{H, \theta_0}^\varepsilon(t) \partial_{\theta_j} Q_{H, \theta_0}^\varepsilon(t) \, dt - \Gamma_H^{ij}(\theta_0)
\]

\[
= c_2 \left( \int_0^T t^{1-2H} \left\{ \partial_{\theta_i} b(X_t^\varepsilon, \theta_0) \partial_{\theta_j} b(X_t^\varepsilon, \theta_0) - \partial_{\theta_i} b(x_t, \theta_0) \partial_{\theta_j} b(x_t, \theta_0) \right\} \, dt \right) \\
+ (c_2 c_3)^{1/2} \left( \int_0^T \partial_{\theta_i} b(X_t^\varepsilon, \theta_0) \int_0^t \partial_{\theta_j} b(X_s^\varepsilon, \theta_0) - \partial_{\theta_j} b(x_s, \theta_0) \right) \frac{1}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \, dt \\
- \int_0^T \partial_{\theta_i} b(x_t, \theta_0) \int_0^t \partial_{\theta_j} b(x_s, \theta_0) \frac{1}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \, dt \\
+ c_3 \left( \int_0^t \left( \int_0^t \partial_{\theta_i} b(X_s^\varepsilon, \theta_0) - \partial_{\theta_i} b(x_s, \theta_0) \right) \frac{1}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right) \right) \\
\times \left( \int_0^t \left( \int_0^t \partial_{\theta_j} b(X_s^\varepsilon, \theta_0) - \partial_{\theta_j} b(x_s, \theta_0) \right) \frac{1}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right) \, dt \\
- \int_0^T \left( \int_0^t \partial_{\theta_i} b(x_t, \theta_0) - \partial_{\theta_i} b(x_s, \theta_0) \right) \frac{1}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right) \\
\times \left( \int_0^t \left( \int_0^t \partial_{\theta_j} b(x_t, \theta_0) - \partial_{\theta_j} b(x_s, \theta_0) \right) \frac{1}{(t-s)^{H+1/2}} s^{1/2-H} \, ds \right) \, dt \right) \right) \right).
\]

Using Lemma 3.1, we have

\[
E \left| \int_0^T t^{1-2H} \left\{ \partial_{\theta_i} b(x_t, \theta_0) \partial_{\theta_j} b(x_t, \theta_0) - \partial_{\theta_i} b(x_t, \theta_0) \partial_{\theta_j} b(x_t, \theta_0) \right\} \, dt \right|^p \lesssim E \left| \int_0^T t^{1-2H} \left( \partial_{\theta_i} b(x_t, \theta_0) \left( \partial_{\theta_j} b(x_t, \theta_0) - \partial_{\theta_j} b(x_t, \theta_0) \right) \\
+ \partial_{\theta_j} b(x_t, \theta_0) \left( \partial_{\theta_i} b(x_t, \theta_0) - \partial_{\theta_i} b(x_t, \theta_0) \right) \right) \right|^p \lesssim \varepsilon^p.
\]
We shall estimate the second term. Note that
\[
\| \partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(X^\epsilon_s, \theta_0) - \partial_\theta b(x_t, \theta_0) + \partial_\theta b(x_s, \theta_0) \|_{L^p(\Omega)} \\
\lesssim (\| \partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(X^\epsilon_s, \theta_0) \|_{L^p(\Omega)} + \| \partial_\theta b(x_t, \theta_0) - \partial_\theta b(x_s, \theta_0) \|_{L^p(\Omega)})^{1/2} \\
\times (\| \partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(x_s, \theta_0) \|_{L^p(\Omega)} + \| \partial_\theta b(X^\epsilon_s, \theta_0) + \partial_\theta b(x_s, \theta_0) \|_{L^p(\Omega)})^{1/2} \\
\lesssim \epsilon^{1/2} |t - s|^{H/2}.
\]

Thus, we obtain that
\[
E \left| \int_0^T \left( \partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(X^\epsilon_s, \theta_0) \right) \int_0^t \frac{\partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(X^\epsilon_s, \theta_0)}{(t-s)^{H+1/2}} \, ds \right|^p \\
= E \left| \int_0^T \left( \partial_\theta b(x_t, \theta_0) \int_0^t \frac{\partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(X^\epsilon_s, \theta_0)}{(t-s)^{H+1/2}} \, ds \right) \right|^p \\
+ \left( \int_0^T (1 + |x_t|^N |x_t|^N) \left\| 1 + |X^\epsilon_t|^N |X^\epsilon_s|^N \left\|_{L^p(\Omega)} \left( \int_0^t \frac{\partial_\theta b(X^\epsilon_t, \theta_0) - \partial_\theta b(X^\epsilon_s, \theta_0) - \partial_\theta b(x_t, \theta_0) + \partial_\theta b(x_s, \theta_0)}{(t-s)^{H+1/2}} \, ds \right) \right)^p \\
\right)
\leq \epsilon^p + \epsilon^{p/2}.
\]

We can estimate the third term in a similar way the second term and we complete the proof.

**Lemma 3.3:** For every \( p \geq 2, \)
\[
\sup_{0 < \epsilon < 1} \epsilon \left[ \sup_{\theta \in \Theta} E \left| \epsilon^2 \nabla_\theta^3 H_{H, \epsilon}(\theta) \right|^p \right] < \infty.
\]

**Proof:** By Sobolev’s inequality, for every \( p > d, \)
\[
\sup_{\theta \in \Theta} \left| \nabla_\theta^3 H_{H, \epsilon}(\theta) \right|^p \lesssim \int_{\Theta} \left( \left| \nabla_\theta^3 H_{H, \epsilon}(\theta) \right|^p + \left| \nabla_\theta^4 H_{H, \epsilon}(\theta) \right|^p \right) d\theta.
\]

By the same argument as in the proof of Lemma 3.2, we can show that
\[
\sup_{0 < \epsilon < 1} \epsilon^{2p} E \left[ \int_{\Theta} \left( \left| \nabla_\theta^3 H_{H, \epsilon}(\theta) \right|^p + \left| \nabla_\theta^4 H_{H, \epsilon}(\theta) \right|^p \right) d\theta \right] < \infty.
\]
Recall that
\[ \partial_\theta \mathbb{L}_{H,\varepsilon} (\theta_0) = \int_0^T Q_{H,\theta_0}^\varepsilon (t) \, dW_t. \]

From the proof of Lemma 3.2, for every \( p \geq 2, \)
\[
E \left| \varepsilon^2 \int_0^T (\nabla_\theta Q_{H,\theta_0}^\varepsilon (t)) \otimes^2 dt - \Gamma_H(\theta_0) \right|^p \to 0,
\]
as \( \varepsilon \to 0. \) Then, we can apply the martingale central limit theorem to the process \( \partial_\theta \mathbb{L}_{H,\varepsilon} (\theta_0) \) and obtain that
\[ \varepsilon \partial_\theta \mathbb{L}_{H,\varepsilon} (\theta_0) \xrightarrow{d} N(0, \Gamma_H(\theta_0)), \quad (4) \]
as \( \varepsilon \to 0. \) Moreover, Lemmas 3.2 and 3.3 and convergence (4) give the local asymptotic normality of \( Z_{H,\varepsilon}(u): \)
\[ Z_{H,\varepsilon}(u) \xrightarrow{d} Z_H(u) := \exp \left( \Delta_H(\theta_0)u - \frac{1}{2} \Gamma_H(\theta_0)[u, u] \right), \]
where \( \Delta_H(\theta_0) \sim N(0, \Gamma_H(\theta_0)) \). In the sequel, we check conditions theorem 3, (c) in [14]. Lemmas 3.2 and 3.3 give \([A1''\prime]\) and \([A 4'\prime]\) in [14] with \( \beta_1 \approx 1/4, \rho_1, \rho_2, \beta, \beta_2 \approx 0. \) The conditions \([B1]\) and \([B2]\) in Yoshida [14] follow from Assumptions 2.3 and 2.4. Moreover, the condition \([A6]\) is true from the following lemma. The proof is similar to the one for Lemmas 3.2 and 3.3.

**Lemma 3.4:** For every \( p \geq 2, \)
\[ \sup_{0 < \varepsilon < 1} E|\varepsilon \partial_\theta \mathbb{L}_{H,\varepsilon} (\theta_0)|^p < \infty \]
and
\[
\sup_{0 < \varepsilon < 1} E \left( \sup_{\theta \in \Theta} \varepsilon^{-1} \left| \mathbb{Y}_{H,\varepsilon} (\theta) - \mathbb{Y}_H(\theta) \right| \right)^p < \infty.
\]
Hence, Theorem 3 of Yoshida [14] yields the inequality
\[
\sup_{0 < \varepsilon < 1} P \left[ \sup_{|u| \geq r} Z_{H,\varepsilon}(u) \geq e^{-r} \right] \lesssim r^{-L}, \quad (5)
\]
hold for any \( r > 0 \) and \( L > 0. \) Since \( u_\varepsilon := \varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0) \) maximizes the random field \( Z_{H,\varepsilon}, \) the sequence \( \{f(u_\varepsilon)\}_\varepsilon \) is uniformly integrable for every continuous function \( f \) such that for
every $x \in \mathbb{R}$, $f(x) \lesssim 1 + |x|^N$ for some $N > 0$. Indeed,

$$\sup_{0 < \varepsilon < 1} P(|u_{\varepsilon}| > r) \lesssim \sup_{0 < \varepsilon < 1} P \left( \sup_{|u| > r} Z_{H,\varepsilon}(u) \geq Z_{H,\varepsilon}(0) \right) \lesssim r^{-L},$$

for every $r > 0$ and $L > 0$. Thus,

$$\sup_{0 < \varepsilon < 1} \mathbb{E}[|f(u_{\varepsilon})|] \lesssim 1 + \int_0^\infty \sup_{0 < \varepsilon < 1} P(|u_{\varepsilon}| > r^{1/N}) \, dr < \infty.$$

Let $B(R) := \{ u \in \mathbb{R}^d ; |u| \leq R \}$. In the sequel, we prove that

$$\log Z_{H,\varepsilon} \overset{d}{\to} \log Z_{H,0} \quad \text{in } C(B(R)), \quad (6)$$

as $\varepsilon \to 0$. If we can show the convergence (6), we obtain the asymptotic normality:

$$\varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0) \overset{d}{\to} N(0, \Gamma_H(\theta_0)^{-1}),$$

as $\varepsilon \to 0$ by Theorem 5 in Yoshida [14]. Due to linearity in $u$ of the weak convergence term $\varepsilon \nabla \log Z_{H,\varepsilon}(\theta_0)[u]$, the convergence of finite-dimensional distribution holds true. It remains to show the tightness of the family $\{ \log Z_{H,\varepsilon}(u) \}_{u \in B(R)}$. By the Kolmogorov tightness criterion, it suffices to show that for every $R > 0$ there exists a constant $p > 0$, $\gamma > d$ and $C > 0$ such that

$$E \left| \log Z_{H,\varepsilon}(u_1) - \log Z_{H,\varepsilon}(u_2) \right|^p \leq C |u_1 - u_2|^{\gamma}, \quad (7)$$

for $u_1, u_2 \in B(R)$. For a number $p > 0$ large enough, the inequality (7) is shown easily by Lemmas 3.2–3.4. Therefore, we complete the proof.

**Remark 3.2:** In this paper, we discussed the MLE constructed by completely observed data $\{ X_t^\varepsilon \}_{t \in [0, T]}$. It is very important in statistical inference to study MLE. However, since the observable data are always discrete in practice, our results cannot be applied directly in any application. Tudor and Viens [6] studied the MLE with the drift function $b(x, \theta) = \theta b(x)$. They discretized the likelihood function when the process are discretely observed and proved the consistency of the estimator. By referring to this, we believe that we can discretize the likelihood function in our model to construct the estimator and guarantee the asymptotic properties.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

The second author was partially supported by JSPS KAKENHI [grant number JP21K03358] and Japan Science and Technology Agency (JST) CREST [grant number JPMJCR14D7], Japan.

**ORCID**

Shohei Nakajima http://orcid.org/0000-0002-0407-7625

Yasutaka Shimizu http://orcid.org/0000-0003-3479-1149
References

[1] Prakasa Rao BLS. Statistical inference for diffusion type processes. Arnold, London, New York (NY): Oxford University Press; 1999.
[2] Liptser RS, Shiryaev AN. Statistics of random processes: II applications. 2nd ed. Berlin, Heidelberg, New York: Springer-Verlag; 2001. (Applications of mathematics).
[3] Kutoyants YA. Statistical inference for ergodic diffusion processes. London, Berlin, Heidelberg: Springer-Verlag; 2004.
[4] Brouste A, Kleptsyna M. Asymptotic properties of MLE for partially observed fractional diffusion system. Stat Inference Stoch Process. 2010;13(1):1–13.
[5] Kleptsyna X, Le Breton M. A statistical analysis of the fractional Ornstein-Uhlenbeck type process. Stat Inference Stoch Process. 2002;5(3):229–248.
[6] Tudor C, Viens F. Statistical aspects of the fractional stochastic calculus. Ann Stat. 2007;35:1183–1212.
[7] Chiba K. An M-estimator for stochastic differential equations driven by fractional Brownian motion with small Hurst parameter. Stat Inference Stoch Process. 2020;23:319–353.
[8] Ibragimov IA, Has’minskii RZ. Statistical estimation: asymptotic theory. New York (NY): Springer; 1981.
[9] Kutoyants YA. Parameter estimation for stochastic processes. Berlin: Heldermann; 1984.
[10] Kutoyants YA. Identification of dynamical systems with small noise. Dordrecht: Kluwer; 1994.
[11] Uchida M, Yoshida N. Information criteria for small diffusions via the theory of Malliavin–Watanabe. Stat Inference Stoch Process. 2004;7:35–67.
[12] Yoshida N. Asymptotic expansion of maximum likelihood estimators for small diffusions via the theory of Malliavin–Watanabe. Probab Theory Relat Fields. 1992;92:275–311.
[13] Yoshida N. Conditional expansions and their applications. Stoch Process Appl. 2003;107:53–81.
[14] Yoshida N. Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. Ann Inst Stat Math. 2011;63(3):431–479.
[15] Nualart D, Ouknine Y. Regularization of differential equations by fractional noise. Stoch Process Appl. 2002;102:103–116.
[16] Nualart D, Răşcanu A. Differential equations driven by fractional Brownian motion. Collect Math. 2002;53:55–81.
[17] Samko SG, Kilbas AA, Marichev OO. Fractional integrals and derivatives. Yverdon: Gordon and Breach; 1993, translated from the 1987 Russian original.