Asymptotic inference for stochastic differential equations driven by fractional Brownian motion

非整数ブラウン運動が駆動する確率微分方程式に対する漸近推測論

February, 2023

Shohei NAKAJIMA
中島　翔平

Asymptotic inference for stochastic differential equations driven by fractional Brownian motion

非整数ブラウン運動が駆動する確率微分方程式に対する漸近推測論

February, 2023

Waseda University Graduate School of Fundamental Science and Engineering

Department of Pure and Applied Mathematics, Research on Stochastic and Statistical Analysis

Shohei NAKAJIMA
中島　翔平

Acknowledgements

I am very grateful to Professor Yasutaka Shimizu who was my supervisor for giving me precious reseach opportunity as a Ph.D student in his laboratory. I learned many things about statistical and mathematical theory. I would also like to thank Professor Tadahisa Funaki and Yoshio Yamada for their guidance of stochastic analysis and partial differential equation respectively while I was enrolled in the master's program.

Contents

1 Introduction

In this thesis, we study asymptotic behavior of maximum likelihood type estimator and leats squares type estimator for parameterized stochastic differential equations driven by a fractional Brownian motion (fBm). Note that the fBm is a centered Gaussian process with $B_0 = 0$ and

$$E[B_t^H B_s^H] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

This covariance formula implies that the sample paths of the fBm have γ-Hölder continuity for any $0 < \gamma < H$. The fBm divided into three different families corresponding to $0 < H < 1/2$, $H = 1/2$ and $1/2 < H < 1$ respectively. When $H = 1/2$, this process is a standard Brownian motion and the increment of the process in disjoint intervals are independent. However, for $H \neq 1/2$ the increments are not independent. When $H > 1/2$, the increments exhibit long-range dependence which is observed in the economics, physics, finance and other fields. When $H < 1/2$, the increments are negatively correlated and it is more important in the field of the mathematical finance (see [9]). Thus, fBm can represent various phenomena and it is important to study the statistical inference for stochastic processes, modeled by stochastic differential equations driven by fBm.

When $H = 1/2$, for which the process is a diffusion process, its parametric inference has been studied by many authors. In particular, the maximum likelihood estimator (MLE) via the likelihood function based on the Girsanov density is one of the optimal estimation methods (see, e.g., Prakasa Rao [36], Liptser and Shiryaev [24], Kutoyants [22] and references therein).

Since fBm is neither a semimartingale nor Markov process, it is not possible to use the tool of the stochastic analysis. In this context, Norros [31] succeeded in analyzing stochastic differential equation driven by fBm within the framework of conventional stochastic analysis by applying a transformation to a certain martingale. This idea has since been widely used in its statistical inference and various studies have been conducted. We refer the reader to Pracasa Rao [37], Mishura [26] and also references therein. Many authors studied the following fractional one-dimensional Orstein–Uhlenbeck process:

$$X_t = X_0 - \theta_0 \int_0^t X_s ds + \sigma B_t^H, \quad t \in (0, T], \tag{1.1}$$

where $X_0 \in \mathbb{R}$ is the initial value, $\{B_t^H\}_{t\in[0,T]}$ is a fBm with Hurst index $H \in (0,1)/\{\frac{1}{2}\}$. In this model, the drift function is linear in θ and the MLE has explicit form. Asymptotic properties of the MLE for unknown parameter θ_0 in (1.1) is well studied when $T \to \infty$.

The main purpose of this thesis is to construct an asymptotic inference theory for parameterized stochastic differential equations driven by fBm with generalized drift functions. In Chapter 3, we consider a stochastic process $\{X_t^\varepsilon\}_{t\in[0,T]}$ which is the solution to the following one-dimensional stochastic differential equations:

$$X_t^\varepsilon = X_0 + \int_0^t b(X_s^\varepsilon, \theta_0)ds + \varepsilon B_t^H, \quad t \in (0,T], \tag{1.2}$$

where $X_0 \in \mathbb{R}$ is the initial value, $\theta_0 \in \Theta$ is the parameter which is contained in a bounded and open convex subset $\Theta \subset \mathbb{R}^d$ and b is the function on $\mathbb{R} \times \Theta$. we consider a maximum likelihood estimator from a realization data $\{X_t^\varepsilon\}_{t\in[0,T]}$ and derive its asymptotic properties under a small dispersion coefficient $\varepsilon \to 0$: small noise asymptotic.

The parametric inference for diffusion processes under small noise asymptotics has been well developed (see, e.g., Kutoyants [20], [21], Uchida and Yoshida [42], Yoshida [43] and [44]). However, parametric estimation problems for the stochastic differential equation driven by fBm has not been analyzed yet. There are some practical advantages in small noise asymptotic. We need technical assumptions such as long-time observation and uniformly-moment conditions in the ergodic setting to obtain limit theorems of estimators. However, it doesn't seem very easy to check those conditions in (1.2) in practice. In the case of small noise asymptotics, we can obtain some limit theorems under relatively milder conditions compared to the ergodic case. In addition, by a suitable scaling, we can deal with a small noise model as a long-term model, approximately. Hence the small noise model is convenient in practice.

In Chapter 4, we consider the following stochastic differential equations which generalize coefficients of noise in (1.2):

$$X_t^\varepsilon = X_0 + \int_0^t b(X_s^\varepsilon, \theta_0)ds + \varepsilon \int_0^t \sigma(X_s^\varepsilon)dB_s^H, \quad t \in (0,T], \tag{1.3}$$

where σ is the function on $\mathbb{R}$ and the integral with respect to B^H is defined as a pathwise Riemann–Stieltjes integral. As in Chapter 3, we consider the M-estimator that naturally follow from the derivation of the MLE from a realization data $\{X_t^\varepsilon\}_{t\in[0,T]}$ and derive its asymptotic properties under a small dispersion coefficient $\varepsilon \to 0$. In addition, we give one of the sufficient conditions which guarantee the absolute continuity between the semimartingale derived by transforming the solution of (1.3) and Wiener measure.

On the other hands, from a practical point of view, the parametric inference for discretely observed data is necessary. In Chapter 5, we consider a stochastic process $\{X_t\}_{t\in[0,T]}$ which is the solution to the following one-dimensional stochastic differential equations:

$$X_t = X_0 + \int_0^t b(X_s, \theta_0)ds + \sigma B_t^H, \quad t \in [0, T], \tag{1.4}$$

where $H \in (1/2, 1)$, θ_0 is a parameter contained a bounded and open convex subset $\Theta \subset \mathbb{R}$, and $\sigma \in \mathbb{R}$ is assumed to be the known diffusion coefficient. we consider the least squares type estimators from discretely observed data $\{X_{kh_n}\}_{k=1}^n$ and derive its asymptotic properties as $n \to \infty,\ h_n \to 0,\ nh_n \to \infty$. The least squares type estimators are relatively tractable estimators, and deriving asymptotic normality is theoretically and practically meaningful. In contrast, Liu *et al.* [25] proved the LAN property for the equation (1.4) when the observation of the process are continuous and the optimal asymptotic variance and converegence rate of estimators have already known. Our results do not achieve that optimal rate of convergence. This indicates that our least squares type estimators are not lilely asymptotic efficient for SDE driven by fBm, unlike the case of Brownian motion. Of course, it can happen the optimality rate of estimators is different between continuous observations and discrete observations. However, Brouste and Iacus [2] and Hu-Nualart-Zhou [14] studied an estimator which is called ergodic type estimator in the case of fractional Orstein–Uhlenbeck process and derived its asymptotic distribution. Their estimators achieve optimal convergence rate suggested by [25] when $H \in (1/2, 3/4)$. Thus we expect that it is possible to improve the convergence rate of our estimator based on discrete samples and more general drift function. In addition, the drift and diffusion parameters and Hurst index should be estimated jointly. [2] and Haress and Hu [12] studied the parametric estimation of fractional Ornstein–Uhlenbeck process based on discrete samples. They proposed the ergodic type estimators and proved joint convergence of all parameters. For more general equations such as (1.3), Kubilius and Skorniakov [19] proposed the estimator of the Hurst index when the function σ is unknown and they proved the consistency and asymptotic normality. However, in this general case, joint convergence of the drift and the diffusion parameters and the Hurst index is still open.

This thesis mainly consists of three already published papers on statistical inference [27], [28] and [29].

2 Preliminaries

2.1 Notations

- Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_t$ be the σ-field generated by the random variables $B_s^H, \ s \in [0, t]$.
- We denote $o_P(1)$ for a sequence of random variables $\{Y_n\}_{n\in\mathbb{N}}$ that converges to zero in probability as $n \to \infty$.
- For a measurable function f on the measure space $(E, \mathcal{B}, \mu)$, we denote $\|Y\|_{L^p(E)} := (\int_E |f|^p d\mu)^{1/p}$.
- $C^{k,l}(\mathbb{R}\times\Theta)$ denotes the space of functions $f : \mathbb{R}\times\Theta \to \mathbb{R}$ such that the function $f(x, \theta)$ is k and l times differentiable with respect to x and θ.
- $C_b^k(\mathbb{R}^n)$ denotes a set of functins which are bounded and $C^k(\mathbb{R}^n)$-class with bounded derivative.
- For any $0 < \lambda < 1$, we define the Hölder space $C^\lambda[0, T]$ which is the set of λ-Hölder continuous functions $g : [0, T] \to \mathbb{R}$ equipped with the norm

  $$\|g\|_{C^\lambda[0,T]} := \|g\|_{\infty,[0,T]} + \|g\|_{\lambda,[0,T]},$$

  with

  $$\|g\|_{\infty,[0,T]} = \sup_{t\in[0,T]} |g(t)|, \qquad \|g\|_{\lambda,[0,T]} = \sup_{s,t\in[0,T]} \frac{|g(t) - g(s)|}{|t - s|^\lambda}.$$
- 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 thesis.

2.2 Fractional calculus

We recall the basic definitions in fractional calculus. See [38] for more details on this subsection. Let $f \in L^1(a, b)$ for $a < b$ and $\alpha > 0$. The fractional Riemann-Liouville integrals of f of order α 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

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

By interchange the order of integration, we can obtain the following semigroup properties

$$I_{a+}^{\alpha} I_{a+}^{\beta} f = I_{a+}^{\alpha+\beta} f,\ I_{b-}^{\alpha} I_{b-}^{\beta} f = I_{b-}^{\alpha+\beta} f,\ \alpha > 0, \beta > 0.$$

For $f \in L^p(a,b),\ g \in L^q(a,b)$ such that $1/p + 1/q \leq 1 + \alpha$, the following integration by parts formula holds

$$\int_a^b I_{a+}^{\alpha} f(x) g(x) dx = \int_a^b f(x) I_{b-}^{\alpha} g(x) dx. \tag{2.1}$$

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+}^{α} (resp. I_{b-}^{α}). For $f \in I_{a+}^{\alpha}(L^p(a,b))$ with $0 < \alpha < 1$, the fractional derivative is defined by

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

Moreover, for $f \in I_{b-}^{\alpha}(L^p(a,b))$,

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

Fractional derivatives also have integration by parts formula

$$\int_a^b D_{a+}^{\alpha} f(x) g(x) dx = \int_a^b f(x) D_{b-}^{\alpha} g(x) dx,$$

for $f \in I_{a+}^{\alpha}(L^p(a,b)),\ g \in I_{b-}^{\alpha}(L^q(a,b)),\ 1/p + 1/q \leq 1 + \alpha$ and the semigroup properties

$$D_{a+}^{\alpha} D_{a+}^{\beta} f = D_{a+}^{\alpha+\beta} f,$$

for $f \in I_{a+}^{\alpha+\beta}(L^1(a,b))$. The following inversion formulas hold

$$I_{a+}^{\alpha} D_{a+}^{\alpha} f = f,$$

for $f \in I_{a+}^{\alpha}(L^p(a,b))$ and

$$D_{a+}^{\alpha} I_{a+}^{\alpha} f = f,$$

for $f \in L^1(a,b)$. The following Proposition guarantees that certain functions are contained in the image of fractional integral which is found in Theorem 13.6 in [38].

Proposition 2.1. *Let $f(x) := (x-a)^{-\mu} g(x)$ where $g \in C^{\lambda}([a,b])$, $a < b$, $\lambda > \alpha$ and $-\alpha < \mu < 1$. Then for $1 \leq p < \infty$ such that $\mu + \alpha < 1/p$ it follows that $f \in I_{a+}^{\alpha}(L^p(a,b))$.*

2.3 Fractional Brownian motion

The fBm with Hurst index $H \in (0,1)$ is a centered Gaussian process and its covariance function is given by

$$E[B_t^H B_s^H] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) =: R_H(s,t).$$

By Kormogorev's continuity criterion and Garsia–Rodemich–Rumsey inequality, we can show that fBm has a version with γ-Hölder continuous trajectories with $0 < \gamma < H$. For $H = 1/2$, fBm reduces to the standard Brownian motion. It is known that fBm is a semimartingale if and only if $H = 1/2$. Therefore, to define stochastic integration, the tool of the stochastic analysis is not applicable and the stochastic integral with respect to fBm needs to define other method.

Let $\mathcal{E}$ be the set of step function on $[0,T]$ and $\mathcal{H}$ be the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the scalar product

$$\langle 1_{[0,t]}, 1_{[0,s]}\rangle_{\mathcal{H}} = R_H(t,s).$$

The mapping $1_{[0,t]} \mapsto B_t^H$ can be extended to a linear isometry between $\mathcal{E}$ and the Gaussian space $\mathcal{H}_1$ spanned by B^H. We denote this isometry by $f \mapsto B^H(f)$ and call it Wiener integral. The covariance of the fBm can be expressed as

$$R_H(t,s) = \int_0^{t\wedge s} k_H(t,u)k_H(s,u)du,$$

where $k_H(t,s)$ is the square integrable kernel defined by

$$k_H(t,s) := \begin{cases} \frac{d_H}{\Gamma(H-1/2)}\Bigg[\left(\frac{t}{s}\right)^{H-1/2}(t-s)^{H-1/2} \\ \qquad\qquad -(H-1/2)s^{1/2-H}\int_s^t u^{H-3/2}(u-s)^{H-1/2}du\Bigg] & \text{if } H < 1/2 \\ \frac{d_H}{\Gamma(H-1/2)} s^{1/2-H}\int_s^t (u-s)^{H-3/2}u^{H-1/2}du & \text{if } H > 1/2, \end{cases}$$

with

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

Consider the linear operator K_H^* from $\mathcal{E}$ to $L^2[0,T]$ defined by

$$(K_H^* f)(s) := \begin{cases} k_H(T,s)f(s) + \int_s^T (f(t)-f(s))\frac{\partial k_H}{\partial t}(t,s)dt & \text{if } H < 1/2 \\ \int_s^T f(t)\frac{\partial k_H}{\partial t}(t,s)dt & \text{if } H > 1/2. \end{cases}$$

Notice that

$$K_H^* 1_{[0,t]} = k_H(t,s) 1_{[0,t]}(s).$$

The operator K_H^* is an isometry between $\mathcal{E}$ and the $L^2[0,T]$ that can be extended to the Hibert space $\mathcal{H}$. Indeed, for $s, t[0,T]$,

$$\begin{aligned}
\langle K_H^* 1_{[0,t]}, K_H^* 1_{[0,s]} \rangle_{L^2[0,T]} &= \langle k_H(t,\cdot) 1_{[0,t]}, k_H(s,\cdot) 1_{[0,s]} \rangle_{L^2[0,T]} \\
&= \int_0^{t \wedge s} k_H(t,u) k_H(s,u) du \\
&= R_H(t,s) = \langle 1_{[0,t]}, 1_{[0,s]} \rangle_{\mathcal{H}}.
\end{aligned}$$

The operator K_H^* can be expressed by the following fractional representation:

$$(K_H^* f)(s) := \begin{cases} d_H s^{1/2-H} D_{T-}^{1/2-H} \left((\cdot)^{H-1/2} f \right)(s) & \text{if } H < 1/2 \\ d_H s^{1/2-H} I_{T-}^{H-1/2} \left((\cdot)^{H-1/2} f \right)(s) & \text{if } H > 1/2. \end{cases}$$

By using the fractional inversion formula, for any $a \in [0,T]$, the inverse operator of K_H^* is given by

$$\left((K_H^*)^{-1} 1_{[0,a]} \right)(s) = \begin{cases} \frac{1}{d_H} s^{1/2-H} I_{a-}^{1/2-H} \left((\cdot)^{H-1/2} \right)(s) 1_{[0,a]}(s) & \text{if } H < 1/2 \\ \frac{1}{d_H} s^{1/2-H} D_{a-}^{H-1/2} \left((\cdot)^{H-1/2} \right)(s) 1_{[0,a]}(s) & \text{if } H > 1/2. \end{cases}$$

Consider the stochastic process $\{W_t\}_{t \in [0,T]}$

$$W_t := B^H \left((K_H^*)^{-1} 1_{[0,t]} \right). \tag{2.2}$$

Then W is a Wiener process. Indeed, for $s, t \in [0,T]$, we have

$$\begin{aligned}
E[W_t W_s] &= E\left[B^H \left((K_H^*)^{-1} 1_{[0,t]} \right) B^H \left((K_H^*)^{-1} 1_{[0,s]} \right) \right] \\
&= \langle B^H \left((K_H^*)^{-1} 1_{[0,t]} \right), B^H \left((K_H^*)^{-1} 1_{[0,s]} \right) \rangle_{\mathcal{H}} \\
&= \langle 1_{[0,t]}, 1_{[0,s]} \rangle = s \wedge t.
\end{aligned}$$

Moreover, the process $B^H(f)$ has an integral representation of the form

$$B_t^H = \int_0^t k_H(t,s) dW_s. \tag{2.3}$$

2.4 Polynomial type large deviation inequality

The statistical inference for unknown parameter $\theta \in \Theta$ will be done besed on the random fields $\mathbb{L}_\varepsilon : \Omega \times \Theta \to \mathbb{R}$. In Chapter 3 and 4, we formulate $\mathbb{L}_\varepsilon$ as a likelihood function and consider the estimator that maximizes it. Let $\mathbb{U}_\varepsilon(\theta_0) := \left\{u \in \mathbb{R}^d : \theta_0 + \varepsilon u \in \Theta\right\}$ and define the random field $\mathbb{Z}_\varepsilon : \mathbb{U}_\varepsilon(\theta_0) \to \mathbb{R}_+$ by

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

Assume that $\mathbb{L}_\varepsilon$ is of C^3 with respect to θ for every $\omega \in \Omega$ and then applying Taylor formula, the random field $\mathbb{Z}_\varepsilon$ is locally asymptotically quadratic at $\theta_0 \in \Theta$ under P_{θ_0}, that is,

$$\mathbb{Z}_\varepsilon(u) = \exp\left(\varepsilon \nabla_\theta \mathbb{L}_\varepsilon(\theta_0) u^* - \frac{1}{2} u \Gamma(\theta_0) u^* + R_\varepsilon(u)\right),$$

where $\Gamma(\theta_0)$ is a d-dimensional deterministic matrix, u^* denotes the tarnspose of u and

$$\begin{aligned} R_\varepsilon(u) = & \frac{1}{2} u \left(\varepsilon^2 \nabla_\theta^2 \mathbb{L}_\varepsilon(\theta_0) - (-\Gamma(\theta_0))\right) u^* \\ & + \frac{1}{2}\varepsilon^3 \int_0^1 (1-s)^2 \nabla_\theta^3 \mathbb{L}_\varepsilon(\theta_0 + s\varepsilon u)[u,u,u] ds. \end{aligned}$$

with $\nabla_\theta^3 \mathbb{L}_\varepsilon(\theta)[x,y,z] = \sum_{i,j,k=1}^d \partial_{\theta_i}\partial_{\theta_j}\partial_{\theta_k} \mathbb{L}_\varepsilon(\theta) x_i y_j z_k$. Moreover, let $\mathbb{Y}_\varepsilon : \Omega \times \Theta \to \mathbb{R}$ be a randon field defined by

$$\mathbb{Y}_\varepsilon(\theta) := \varepsilon^2 \left(\mathbb{L}_\varepsilon(\theta) - \mathbb{L}_\varepsilon(\theta_0)\right),$$

and $\mathbb{Y}_H$ be the expected its limit when $\varepsilon \to 0$.

Assumption 2.1. *For every $p > 0$,*

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

Assumption 2.2. *For every $p > 0$,*

$$\sup_{\varepsilon \in (0,1]} E\left(\varepsilon^{-1/2}\left|-\varepsilon^2 \nabla_\theta^2 \mathbb{L}_\varepsilon(\theta_0) - \Gamma(\theta_0)\right|\right)^p < \infty.$$

Assumption 2.3. *For every $p > 0$,*

$$\sup_{\varepsilon \in (0,1]} E\left(\varepsilon^2 \sup_{\theta \in \Theta}\left|\nabla_\theta^3 \mathbb{L}_\varepsilon(\theta)\right|\right)^p < \infty.$$

Assumption 2.4. *For every $p > 0$,*

$$\sup_{\varepsilon\in(0,1]} E\left(\sup_{\theta\in\Theta}\varepsilon^{-1}\left|\mathbb{Y}_\varepsilon(\theta)-\mathbb{Y}(\theta)\right|\right)^p < \infty.$$

Assumption 2.5. *The matrix $\Gamma(\theta_0)$ is deterministic and positive definite.*

Assumption 2.6. *There exists a positive constant $\xi(\theta_0) > 0$ such that*

$$\mathbb{Y}(\theta) \le -\xi(\theta_0)|\theta-\theta_0|^2,$$

for every $\theta \in \Theta$.

The following theorems are Theorems 3, 5 of Yoshida [45]. Here we give a simplified version of them.

Theorem 2.1. *Under Assumptions 2.1-2.6, there exists a constant $C > 0$ such that*

$$\sup_{0<\varepsilon<1} P\left[\sup_{|u|\ge r}\mathbb{Z}_\varepsilon(u) \ge e^{-r}\right] \le Cr^{-L},$$

for any $r > 0$ and $L > 0$.

Let $B(R) := \{u \in \mathbb{R}^d;\ |u| \le R\}$ for $R > 0$ and $\hat{\theta}_\varepsilon$ be a random variable maximizes $\mathbb{L}_\varepsilon$:

$$\mathbb{L}_\varepsilon(\hat{\theta}_\varepsilon) = \sup_{\theta\in\bar{\Theta}}\mathbb{L}_\varepsilon(\theta),$$

and let $u_\varepsilon := \varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0)$.

Theorem 2.2. *Assume that there exists a random function $\mathbb{Z}$ such that for every $R > 0$,*

$$\mathbb{Z}_\varepsilon \xrightarrow{d} \mathbb{Z},$$

in $C(B(R))$ as $\varepsilon \to 0$ and there exists a measurable mapping $\hat{u}$ that is a unique maximum point of $\mathbb{Z}$. Then $\hat{u}_\varepsilon \xrightarrow{d} \hat{u}$ as $\varepsilon \to 0$.

2.5 Basic inequalities

We summarize the inequalities that will be used frequently throughout this thesis. First, we state Hölder' inequality and Minkowski's inequality. For proofs, see theorem 6.2.7 in [40].

Theorem 2.3 (Hölder's inequality)**.** *Let $p, q \in [1, \infty]$ such that $1/p + 1/q = 1$. Then for every measurable functions f, g on the measure space $(E, \mathcal{B}, \mu)$,*

$$\|fg\|_{L^1(E)} \leq \|f\|_{L^p(E)} \|g\|_{L^q(E)}.$$

Theorem 2.4 (Minkowski's inequality)**.** *Let $(E_i, \mathcal{B}_i, \mu_i)$, $i \in \{1, 2\}$ be σ-finite measure spaces. Then for $1 \leq q \leq p\infty$ and measurable functions f on $(E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2)$,*

$$\left(\int_{E_2} \left(\int_{E_1} |f(x_1, x_2)|^q \mu_1(dx_1) \right)^{\frac{p}{q}} \mu_2(dx_2) \right)^{\frac{1}{p}} \leq \left(\int_{E_1} \left(\int_{E_2} |f(x_1, x_2)|^p \mu_2(dx_2) \right)^{\frac{q}{p}} \mu_1(dx_1) \right)^{\frac{1}{q}}$$

Lemma 2.1 (Gronwall's inequality)**.** *Let f, g be non-negative and local-integrable functions on $[0, \infty)$. Suppose that a non-negative function ϕ on $[0, \infty)$ satisfies*

$$\phi(t) \leq f(t) + \int_0^t g(s)\phi(s) ds.$$

Then the following inequality is valid:

$$\phi(t) \leq f(t) + \int_0^t f(s)g(s) \exp(\int_s^t g(r)dr) ds.$$

The following lemma is often used to evaluate the Hölder norm of some functions. Its proof is found in [8].

Lemma 2.2 (Garsia–Rodemich–Rumsey inequality)**.** *Let $p \geq 1$ and $\alpha > p^{-1}$. Then there exists a constant $C_{\alpha,p} > 0$ such that for any $f \in C[0, T]$ and for all $s, t \in [0, T]$,*

$$|f(t) - f(s)|^p \leq C_{\alpha,p} |t - s|^{\alpha p - 1} \int_0^T \int_0^T \frac{|f(x) - f(y)|^p}{|x - y|^{\alpha p + 1}} dxxy.$$

We prepare fractional version of the Gronwall's inequality (see [34]).

Lemma 2.3. *Let $0 \leq \alpha < 1$, $a, b \geq 0$ and f is non-negative and continuous function on $[0, \infty)$ such that*

$$f(t) \leq a + bt^\alpha \int_0^t (t - s)^{-\alpha} s^{-\alpha} f(s) ds.$$

Then

$$f(t) \leq a d_\alpha \exp(c_\alpha t b^{1/(1-\alpha)}),$$

where c_α and d_α are positive constant depending on α.

In order to establish some Hölder norm estimates of the solution to the SDE, we prepare the following lemma which is found in Exercise 4.24 in [6].

Lemma 2.4. *Let $\alpha \in (0,1)$, $h > 0$, and $Z \in C^\alpha[0,T]$. Assume that*

$$\|Z\|_{\alpha,h} := \sup \frac{|Z_{s,t}|}{|t-s|^\alpha} \le M,$$

where sup is restricted to time $s,t \in [0,T]$ and $|t-s| \le h$. Then

$$\|Z\|_{\alpha,[0,T]} \le M(1 \vee 2(T/h)^{1-\alpha}).$$

We prepare the Burkholder–Davis–Gundy's inequality which is appied to estimate the stochastic integral with respect to Wiener process. For a proof, see Theorem 5.16 in [23].

Theorem 2.5 (Burkholder–Davis–Gundy's inequality)**.** *For every $p > 0$, there exist constants c_p and C_p such that for every continuous martingale $\{M_t\}_{t\in[0,T]}$ satisfying $E[\langle M\rangle_T^{p/2}] < \infty$,*

$$c_p E[\langle M\rangle_T^{p/2}] \le E[\sup_{t\in[0,T]} |M_t|^p] \le C_p E[\langle M\rangle_T^{p/2}],$$

where $\langle M\rangle$ is the quadratic variation of M.

We state Fernique's theorem frequently used in Section 4. Let E be a separable Banach space with norm $\|\cdot\|_E$ and $\mathcal{B}(E)$ be its Borel σ-field. A centered Gaussian measure μ on $(E,\mathcal{E})$ is a probability measure such that $l^*\mu$ is a real Gaussian probability measure on $\mathbb{R}$ with zero mean for every linear functional $l : E \to \mathbb{R}$. For a proof, see [5] or [10].

Theorem 2.6 (Fernique's theorem)**.** *Let μ be an arbitrary centered Gaussian measure on E. Then there exists a constant $c > 0$ such that*

$$\int_E e^{c\|x\|_E^2} \mu(dx) < \infty.$$

Remark 2.1. In a later section, we will use Fernique's theorem on the Hölder space $C^\lambda[0,T]$. However, we need to be a little more careful because the Hölder space is not separable. In order to apply Fernique's theorem, we fix $0 < \beta < \lambda$ and let E be a function space defined as the closure of $C^\beta[0,T]$ in the norm $C^\lambda[0,T]$. Then E is a separable Banach space and we can see that μ can actually realised as a Gaussian measure on E. Therefore, Fernique's theorem is applicable. See [1] and [10] for further details.

Let $f \in C^\lambda[a,b]$ and $g \in C^\mu[a,b]$ with $\lambda + \mu > 1$. Then the Riemann Stieltjes integral $\int_a^b f(s)dg(s)$ exists , see Young [46]. Moreover, the chain rule for the change of variable is valid which found Theorem 4.3.1 in [47].

Lemma 2.5. *Let $f \in C^\lambda[a,b]$ and $F \in C^1(\mathbb{R})$ be functions such that $\partial_x F(f(\cdot)) \in C^\mu[a,b]$ with $\lambda + \mu > 1$. Then*

$$F(f(x)) - F(f(a)) = \int_a^x \partial_x F(f(y))df(y), \; x \in [a,b].$$

Lemma 2.6 (Young's inequality)**.** *Let $f \in C^\lambda[a,b]$ and $g \in C^\mu[a,b]$ with $\lambda + \mu > 1$. Then there exists a constant $C_{\lambda,\mu} > 0$ such that*

$$\left| \int_a^b (f(x) - f(a))\, dg(x) \right| \leq C_{\lambda,\mu} \|f\|_{\lambda,[a,b]} \|g\|_{\mu,[a,b]}.$$

3 Additive noise case

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

$$X_t^\varepsilon = X_0 + \int_0^t b(X_s^\varepsilon, \theta_0)ds + \varepsilon B_t^H, \quad t \in (0,T], \tag{3.1}$$

where $X_0 \in \mathbb{R}$ is the initial value, $\{B_t^H\}_{t\in[0,T]}$ is a fBm with Hurst index $H \in (0,1)/\{\frac{1}{2}\}$ 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 $W^{1,p}(\Theta) \hookrightarrow C(\bar{\Theta})$. Without loss of generality, we assume that $\varepsilon \in (0,1]$. The main purpose in this chapter is the estimation of parameter $\theta_0 \in \Theta$ from a realization $\{X_t^\varepsilon\}_{t\in[0,T]}$ when $\varepsilon \to 0$.

There are several results on the parametric inference for stochastic differential equation driven by fractional Brownian motion. Brouste and Kleptsyna [3], Kleptsyna and Le Breton [17] studied the parameter estimation problem for continuously observed fractional Ornstein–Uhlenbeck processes. In these papers, the drift function is linear in both x and θ ($b(x,\theta) = -\theta x$) and the asymptotic normality and moment convergence of MLE are established when the terminal time of observation goes to infinity. In similar framework, Tudor and Viens [41] 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 θ and the MLE has an explicit expression. Recently, in the case when the drift function $b(x,\theta)$ is nonlinear in both x and θ, Chiba [4] proposed an M-estimator based on the likelihood function. Unlike previous studies, the estimator proposed in [4] 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 [15] and established asymptotic properties of the estimator when the Hurst index H is contained in $(\frac{1}{4}, \frac{1}{2})$. 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. We allow the drift function $b(x,\theta)$ is nonlinear in both x and θ and the Hurst index H is contained in $(0,1)/\{\frac{1}{2}\}$. We aim to deduce asymptotic normality and moment convergence of the MLE of the drift parameter under $\varepsilon \to 0$.

3.1 Construction of the MLE

We aim to estimate the unknown parameter $\theta_0 \in \Theta$ in the equation (3.1) from completely observed data $\{X_t^\varepsilon\}_{t\in[0,T]}$. We impose some assumptions on the parameter space Θ and coefficients b to derive likelihood function.

Assumption 3.1. *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 Θ such that f and its derivative in the weak sense are L^p integrable functions.*

Assumption 3.2. *The function b in (3.1) 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:*

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

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

Assumption 3.3. *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 [33] and [34], the existence and uniqueness of the strong solution to equation (3.1) follows under Assumption 3.2 and 3.3. In addition, for every $0 < \alpha < H$, the solution to (3.1) has α-Hölder continuity. By Proposition 2.1, we can define the function

$$Q^\varepsilon_{H,\theta}(t) := \begin{cases} (\varepsilon d_H)^{-1} t^{H-1/2} I_{0+}^{1/2-H} \left[(\cdot)^{1/2-H} b(X^\varepsilon_\cdot, \theta) \right](t) & \text{if } H < 1/2 \\ (\varepsilon d_H)^{-1} t^{H-1/2} D_{0+}^{H-1/2} \left[(\cdot)^{1/2-H} b(X^\varepsilon_\cdot, \theta) \right](t) & \text{if } H > 1/2. \end{cases}$$

For $0 < s < t$, let $k_H^{-1}(t, s)$ be a functin given by

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

Define a stochastic process $\{W_t\}_{t \in [0,T]}$ by

$$W_t := \int_0^t k_H^{-1}(t, s)\, dB_s^H.$$

Here we interpret the stochastic integral with respect to a fractional Brownian motion as a Wiener integral. Then W is a Wiener process and B^H has the integral representation

$$B_t^H = \int_0^t k_H(t, s)\, dW_s.$$

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

$$\begin{aligned} Z_t :&= \varepsilon^{-1}\int_0^T k_H^{-1}(t,s)dX_s^\varepsilon \\ &= \int_0^t Q_{H,\theta}^\varepsilon(s)ds + W_t. \end{aligned}$$

We can find the following Girsanov theorem in [41].

Proposition 3.1. *Under Assumptions 1-3, the stochastic process $\{X_t^\varepsilon/\varepsilon\}_{t\in[0,T]}$ is a $\mathcal{F}_t$-fractional Brownian motion.*

By Proposition 3.1, the log-likelihood function $\mathbb{L}_{H,\varepsilon}$ for the equation (3.1) can be obtained by

$$\mathbb{L}_{H,\varepsilon}(\theta) := \int_0^T Q_{H,\theta}^\varepsilon(t)dZ_t - \frac{1}{2}\int_0^T Q_{H,\theta}^\varepsilon(t)^2 dt.$$

We define the maximum likelihood estimator by

$$\hat{\theta}_\varepsilon := \arg\max_{\theta\in\bar{\Theta}} \mathbb{L}_{H,\varepsilon}(\theta).$$

3.2 Main results

In order to state our main results about asymptotic properties of $\hat{\theta}^\varepsilon$, 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:

$$\begin{cases} \dfrac{dx_t}{dt} = b(x_t,\theta_0) \\ \;\; x_0 = X_0. \end{cases} \tag{3.2}$$

We set the d–dimensional square matrix $\Gamma_H(\theta_0)$ which is an asymptotic convariance matrix of our estimator as

$$\Gamma_H^{i,j}(\theta_0) := \begin{cases} c_1\int_0^T t^{2H-1}\left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\partial_{\theta_i}b(x_s,\theta_0)ds\right) & \\ \qquad\qquad \times\left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\partial_{\theta_j}b(x_s,\theta_0)ds\right)dt & \text{if } H<1/2 \\ \int_0^T\left(c_2 t^{1/2-H}\partial_{\theta_i}b(x_t,\theta_0) + c_3 t^{H-1/2}\int_0^t \frac{\partial_{\theta_i}b(x_t,\theta_0)-\partial_{\theta_i}b(x_s,\theta_0)}{(t-s)^{H+1/2}}s^{1/2-H}ds\right) & \\ \times\left(c_2 t^{1/2-H}\partial_{\theta_j}b(x_t,\theta_0) + c_3 t^{H-1/2}\int_0^t \frac{\partial_{\theta_j}b(x_t,\theta_0)-\partial_{\theta_j}b(x_s,\theta_0)}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)dt & \text{if } H>1/2, \end{cases}$$

where

$$\begin{aligned}
c_1 &= (d_H\Gamma(1/2-H))^{-2} \\
c_2 &= (d_H\Gamma(3/2-H))^{-1}\left\{1+(H-1/2)\int_0^1 \frac{1-s^{1/2-H}}{(1-s)^{H+1/2}}ds\right\} \\
c_3 &= (H-1/2)\,(d_H\Gamma(3/2-H))^{-1}.
\end{aligned}$$

Assumption 3.4. *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

$$\mathbb{Y}_{H,\varepsilon}(\theta) := \varepsilon^2\left(\mathbb{L}_{H,\varepsilon}(\theta)-\mathbb{L}_{H,\varepsilon}(\theta_0)\right),$$

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

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

Assumption 3.5. *There exists a positive constant $\xi(\theta_0)>0$ such that*

$$\mathbb{Y}_H(\theta) \le -\xi(\theta_0)|\theta-\theta_0|^2,$$

for every $\theta\in\Theta$.

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

Theorem 3.1. *Suppose that the Assumptions 3.1-3.5 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

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

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

3.3 Examples

Example 3.1. We consider a one-dimensional fractional Ornstein–Uhlenbeck process that is, the drift function in (3.1) is given by $b(x, \theta_0) = \theta_0 x$ with $\theta_0 \in \mathbb{R}$. Then x_t satisfies the following equation

$$\begin{cases} \dfrac{dx_t}{dt} = \theta_0 x_t \\ \quad x_0 = X_0. \end{cases}$$

The explicit solution is given by $x_t = X_0 e^{\theta_0 t}$. In this case, we can check Assumptions 1-4 hold true. Indeed, Assumptions 1,2 are obvious. We will show that Assumptions 3.4 and 3.5 hold. In the case $H < 1/2$,

$$\begin{aligned}
\Gamma_H(\theta_0) &= c_1 \int_0^T t^{2H-1} \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} X_0 e^{\theta_0 s} ds \right)^2 dt \\
&\geq c_1 X_0^2 \left(e^{\theta_0 T} \wedge 1\right)^2 \int_0^T t^{2H-1} \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} ds \right)^2 dt \\
&= c_1 X_0^2 \left(e^{\theta_0 T} \wedge 1\right)^2 \beta(3/2-H, 1/2-H)^2 T^{2-2H} > 0.
\end{aligned}$$

Thus the Assumption 3.4 holds true. In the same way,

$$\begin{aligned}
-\mathbb{Y}_H(\theta) &= \frac{c_1}{2} \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H}(t-s)^{-1/2-H} \left(\theta_0 x_s - \theta x_s\right) ds \right\}^2 dt \\
&= \frac{c_1}{2} (\theta_0 - \theta)^2 X_0^2 \int_0^T t^{2H-1} \left\{ \int_0^t s^{1/2-H}(t-s)^{-1/2-H} e^{\theta_0 s} ds \right\}^2 dt \\
&\geq \frac{c_1}{2} (\theta_0 - \theta)^2 X_0^2 \left(e^{\theta_0 T} \wedge 1\right)^2 \beta(3/2-H, 1/2-H)^2 T^{2-2H},
\end{aligned}$$

and the Assumption 3.5 holds true. In the case $H > 1/2$, we restrict $\theta_0 > 0$. Then

$$\begin{aligned}
-\mathbb{Y}_H(\theta) &= \frac{(\theta-\theta_0)^2}{2} X_0^2 \int_0^T \left(c_2 t^{1/2-H} e^{\theta_0 t} + c_3 t^{H-1/2} \int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} ds \right)^2 dt \\
&= \frac{(\theta-\theta_0)^2}{2} X_0^2 \Bigg\{ c_2^2 \int_0^T t^{1-2H} e^{2\theta_0 t} + 2 c_2 c_3 e^{\theta_0 t} \int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} ds \\
&\qquad + c_3^2 t^{2H-1} \left(\int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} ds \right)^2 dt \Bigg\}.
\end{aligned}$$

By the mean value theorem, we have

$$\int_0^t \frac{e^{\theta_0 t} - e^{\theta_0 s}}{(t-s)^{H+1/2}} s^{1/2-H} ds \geq \theta_0 \int_0^t (t-s)^{1/2-H} s^{1/2-H} e^{\theta_0 s} ds$$
$$\geq \theta_0 t^{2-2H} \beta(3/2-H, 3/2-H).$$

Thereofore

$$-\mathbb{Y}_H(\theta) \geq \frac{(\theta-\theta_0)^2}{2} X_0^2 \left\{ c_2^2 \int_0^T t^{1-2H} e^{2\theta_0 t} + 2c_2 c_3 e^{\theta_0 t} \theta_0 t^{2-2H} \beta(3/2-H, 3/2-H) \right.$$
$$\left. + c_3^2 \theta_0^2 t^{3-2H} \beta(3/2-H, 3/2-H)^2 dt \right\}$$
$$\geq \frac{(\theta-\theta_0)^2}{2} X_0^2 \left\{ \frac{c_2^2}{2-2H} T^{2-2H} + \frac{2c_2 c_3 \theta_0}{3-2H} T^{3-2H} \beta(3/2-H, 3/2-H) \right.$$
$$\left. + \frac{c_3^2 \theta_0^2}{4-2H} T^{4-2H} \beta(3/2-H, 3/2-H)^2 \right\},$$

and Assumption 3.5 holds. Assumption 3.4 can be confirmed by the same calculation.

Example 3.2. 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 3.2,3.3 hold true. We check Assumption 3.5. Note that the function x_t is monotone increasing and satisfies the relation

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

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

$$
\begin{aligned}
-\mathbb{Y}_H(\theta) &= \frac{c_1}{2}\int_0^T t^{2H-1}\left\{\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\left(\sqrt{\theta_0+x_s^2}-\sqrt{\theta+x_s^2}\right)ds\right\}^2 dt\\
&\geq \frac{c_1}{8}(\theta_0-\theta)^2\int_0^T t^{2H-1}\left\{\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\frac{1}{\sqrt{M+x_s^2}}ds\right\}^2 dt\\
&\geq \frac{c_1}{8(M+x_T^2)}(\theta_0-\theta)^2\int_0^T t^{2H-1}\left\{\int_0^t s^{1/2-H}(t-s)^{-1/2-H}ds\right\}^2 dt\\
&\geq (\theta_0-\theta)^2\frac{c_1\beta(3/2-H,1/2-H)^2T^{2-2H}}{8\left(M+\left(X_0+\sqrt{\theta_0+X_0^2}\right)^2e^{2T}\right)},
\end{aligned}
$$

and Assumption 3.5 holds. Assumption 3.4 can be confirmed by the same calculation.

Example 3.3. Let $H\in(1/2,1)$. We consider a simpler drift function than Example 3.2 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 3.2, Assumptions 1,2 can be checked and the function x_t satisfies the relation $x_t+\theta_0\sqrt{1+x_t^2}=\left(X_0+\theta_0\sqrt{1+X_0^2}\right)e^t$. We check Assumption 3.5 holds true.

$$
\begin{aligned}
-\mathbb{Y}_H(\theta) &= \frac{1}{2}\int_0^T\Bigg(c_2t^{1/2-H}\left(\theta_0\sqrt{1+x_t^2}-\theta\sqrt{1+x_t^2}\right)\\
&\qquad + c_3t^{H-1/2}\int_0^t\frac{\left(\theta_0\sqrt{1+x_t^2}-\theta\sqrt{1+x_t^2}\right)-\left(\theta_0\sqrt{1+x_s^2}-\theta\sqrt{1+x_s^2}\right)}{(t-s)^{H+1/2}}s^{1/2-H}ds\Bigg)^2 dt\\
&= \frac{(\theta_0-\theta)^2}{2}\int_0^T\left(c_2t^{1/2-H}\sqrt{1+x_t^2}\right)^2+2c_2c_3\sqrt{1+x_t^2}\int_0^t\frac{\sqrt{1+x_t^2}-\sqrt{1+x_s^2}}{(t-s)^{H+1/2}}s^{1/2-H}ds\\
&\qquad + \left(c_3t^{H-1/2}\int_0^t\frac{\sqrt{1+x_t^2}-\sqrt{1+x_s^2}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)^2 dt.
\end{aligned}
\tag{3.3}
$$

We evaluate each of the three terms that appear in the last equality in (3.3). Using the monotonicity of x_t, the first term can be estimated as

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

Let us estimate the second term of (3.3). 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

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

In the same way, we can estimate the third term of (3.3) as

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

and Assumption 3.5 is valid. Assumption 3.4 can be confirmed by the same calculation.

3.4 Proofs

Recall that $\{x_t\}_{0\leq t\leq T}$ be the solution to the following differential equation:

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

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

$$E|X_t^\varepsilon - x_t|^p \leq c_1\varepsilon^p$$
$$E|X_t^\varepsilon|^p \leq c_2,$$

and

$$E|X_t^\varepsilon - X_s^\varepsilon|^p \leq c_3|t-s|^{pH}.$$

Proof. By Assumptions 3.3,

$$
\begin{aligned}
|X_t^\varepsilon - x_t| &\le \int_0^t |b(X_s^\varepsilon,\theta_0) - b(x_s,\theta_0)| + \varepsilon |B_t^H| \\
&\le L \int_0^t |X_s^\varepsilon - x_s| ds + \varepsilon \sup_{0\le t\le T} |B_t^H|.
\end{aligned}
$$

By Gronwall's inequality, it follows that

$$
|X_t^\varepsilon - x_t| \le \varepsilon e^{Lt} \sup_{0\le t\le T} |B_t^H|,
$$

and the first estimate follows. Other estimates hold true by the linear growth condition of the functon b. □

Let $\mathbb{U}_\varepsilon(\theta_0) := \left\{ u \in \mathbb{R}^d : \theta_0 + \varepsilon u \in \Theta \right\}$ 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

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

where

$$
\begin{aligned}
R_\varepsilon(u) = &\frac{1}{2} u \left(\varepsilon^2 \nabla_\theta^2 \mathbb{L}_{H,\varepsilon}(\theta_0) - (-\Gamma_H(\theta_0)) \right) u^* \\
&+ \frac{1}{2}\varepsilon^3 \int_0^1 (1-s)^2 \nabla_\theta^3 \mathbb{L}_{H,\varepsilon}(\theta_0 + s\varepsilon u)[u,u,u] ds.
\end{aligned}
$$

The following lemma gives Assumption 2.2 of Theorem 2.1.

Lemma 3.2. *For every $p > 0$,*

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

where

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

Proof. Note that the function $\mathbb{L}_{H,\varepsilon}$ is twice differentiable in θ and we have

$$
\varepsilon^2 \nabla_\theta^2 \mathbb{L}_{H,\varepsilon}(\theta_0) - (-\Gamma_H(\theta_0)) = \varepsilon^2 \int_0^T \nabla_\theta^2 Q_{H,\theta_0}^\varepsilon(t) dW_t - \left(\varepsilon^2 \int_0^T \left(\nabla_\theta Q_{H,\theta_0}^\varepsilon(t) \right)^{\otimes 2} dt - \Gamma_H(\theta_0) \right).
$$

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

$$Q^{\varepsilon}_{H,\theta_0}(t) = c_1^2 \varepsilon^{-1} t^{H-1/2} \int_0^t s^{1/2-H}(t-s)^{-1/2-H} b(X^{\varepsilon}_s, \theta_0) ds.$$

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

$$\begin{aligned}
&E\left(\varepsilon^2 \left| \int_0^T \partial_{\theta_i}\partial_{\theta_j} Q^{\varepsilon}_{H,\theta_0}(t) dW_t \right| \right)^p \lesssim \left(\varepsilon^4 \int_0^T \left\| \partial_{\theta_i}\partial_{\theta_j} Q^{\varepsilon}_{H,\theta_0}(t) \right\|^2_{L^p(\Omega)} dt \right)^{p/2} \\
&\lesssim \varepsilon^p \left(\int_0^T t^{2H-1} \left| \int_0^t s^{1/2-H}(t-s)^{-1/2-H} \left\| \partial_{\theta_i}\partial_{\theta_j} b(X^{\varepsilon}_s, \theta_0) \right\|_{L^p(\Omega)} ds \right|^2 dt \right)^{p/2} \\
&\lesssim \varepsilon^p \sup_{0 \le s \le T} \left\| 1 + |X^{\varepsilon}_s|^N \right\|^p_{L^p(\Omega)} \left(\int_0^T t^{1-2H} dt \right)^{p/2} \lesssim \varepsilon^p,
\end{aligned}$$

for every $i, j = 1, \cdots, d$. We shall estimate the second part. For every $i, j = 1, \cdots, d$,

$$\begin{aligned}
&\varepsilon^2 \int_0^T \partial_{\theta_i} Q^{\varepsilon}_{H,\theta_0}(t) \partial_{\theta_j} Q^{\varepsilon}_{H,\theta_0}(t) dt - \Gamma^{i,j}_H(\theta_0) \\
&= c_1^2 \int_0^T t^{2H-1} \Bigg\{ \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_i} b(X^{\varepsilon}_s, \theta_0) ds \right) \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_j} b(X^{\varepsilon}_s, \theta_0) ds \right) \\
&\qquad - \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_i} b(x_s, \theta_0) ds \right) \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_j} b(x_s, \theta_0) ds \right) \Bigg\} dt \\
&= c_1^2 \int_0^T t^{2H-1} \Bigg\{ \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_i} \left(b(X^{\varepsilon}_s, \theta_0) - b(x_s, \theta_0) \right) ds \right) \\
&\qquad \times \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_j} b(X^{\varepsilon}_s, \theta_0) ds \right) \\
&\qquad - \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_i} b(x_s, \theta_0) ds \right) \\
&\qquad \times \left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H} \partial_{\theta_j} \left(b(x_s, \theta_0) - b(X^{\varepsilon}_s, \theta_0) \right) ds \right) \Bigg\} dt.
\end{aligned}$$

By Hölder's and Minkowski's inequalities and Lemma 3.1, we can show that

$$
\begin{aligned}
&E\left|\int_0^T t^{2H-1}\left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\partial_{\theta_i}\left(b(X_s^\varepsilon,\theta_0)-b(x_s,\theta_0)\right)ds\right)\right.\\
&\qquad\qquad \left.\times\left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)ds\right)dt\right|^p\\
&\leq \left|\int_0^T t^{2H-1}\left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\left\|\partial_{\theta_i}\left(b(X_s^\varepsilon,\theta_0)-b(x_s,\theta_0)\right)\right\|_{L^{2p}(\Omega)}ds\right)\right.\\
&\qquad\qquad \left.\times\left(\int_0^t s^{1/2-H}(t-s)^{-1/2-H}\left\|\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)\right\|_{L^{2p}(\Omega)}ds\right)dt\right|^p\\
&\lesssim \sup_{0\leq s\leq T}\left\|\left(1+|X_s^\varepsilon|^N+|x_s|^N\right)|X_s^\varepsilon-x_s|\right\|_{L^{2p}(\Omega)}^p\left(\int_0^T t^{1-2H}dt\right)\lesssim\varepsilon^p.
\end{aligned}
$$

Therefore

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

The case where $H>1/2$. We have that

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

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

$$
\begin{aligned}
&E\left(\varepsilon^2 \left|\int_0^T \partial_{\theta_i}\partial_{\theta_j} Q^{\varepsilon}_{H,\theta_0}(t)dW_t\right|\right)^p \lesssim \left(\varepsilon^4 \int_0^T \left\|\partial_{\theta_i}\partial_{\theta_j} Q^{\varepsilon}_{H,\theta_0}(t)\right\|^2_{L^p(\Omega)} dt\right)^{p/2} \\
&\lesssim \varepsilon^p \Bigg\{ \int_0^T t^{1-2H} \|\partial_{\theta_i}\partial_{\theta_j} b(X^{\varepsilon}_t, \theta_0)\|^2_{L^p(\Omega)} dt \\
&\quad + \int_0^T t^{2H-1} \left\| \int_0^t \frac{\partial_{\theta_i}\partial_{\theta_j} b(X^{\varepsilon}_t, \theta_0) - \partial_{\theta_i}\partial_{\theta_j} b(X^{\varepsilon}_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds \right\|^2_{L^p(\Omega)} dt \Bigg\}^{p/2} \\
&\lesssim \varepsilon^p \Bigg\{ \left\| 1 + \sup_{0\le t\le T} E|X^{\varepsilon}_t|^{2N} \right\|^2_{L^p(\Omega)} \int_0^T t^{1-2H} dt \\
&\quad + \int_0^T t^{2H-1} \left\| \int_0^t \frac{\left(1 + |X^{\varepsilon}_t|^N + |X^{\varepsilon}_s|^N\right) |X^{\varepsilon}_t - X^{\varepsilon}_s|}{(t-s)^{H+1/2}} s^{1/2-H} ds \right\|^2_{L^p(\Omega)} dt \Bigg\}^{p/2} \\
&\lesssim \varepsilon^p \left\{ 1 + \int_0^T t^{2H-1} \left(\int_0^t \frac{\|X^{\varepsilon}_t - X^{\varepsilon}_s\|_{L^{2p}(\Omega)}}{(t-s)^{H+1/2}} s^{1/2-H} ds \right)^2 dt \right\}^{p/2} \lesssim \varepsilon^p,
\end{aligned}
$$

for every $i, j = 1, \cdots, d$. We estimate the term $\varepsilon^2 \int_0^T \left(\nabla_\theta Q^{\varepsilon}_{H,\theta_0}(t)\right)^{\otimes 2} dt - \Gamma_H(\theta_0)$. For every $i, j = 1, \cdots, d$

$$
\begin{aligned}
&\varepsilon^2 \int_0^T \partial_{\theta_i} Q^{\varepsilon}_{H,\theta_0}(t) \partial_{\theta_j} Q^{\varepsilon}_{H,\theta_0}(t) dt - \Gamma^{i,j}_H(\theta_0) \\
&= c_2 \left(\int_0^T t^{1-2H} \left\{ \partial_{\theta_i} b(X^{\varepsilon}_t, \theta_0) \partial_{\theta_j} b(X^{\varepsilon}_t, \theta_0) - \partial_{\theta_i} b(x_t, \theta_0) \partial_{\theta_j} b(x_t, \theta_0) \right\} dt \right) \\
&+ (c_2 c_3)^{1/2} \Bigg(\int_0^T \partial_{\theta_i} b(X^{\varepsilon}_t, \theta_0) \int_0^t \frac{\partial_{\theta_j} b(X^{\varepsilon}_t, \theta_0) - \partial_{\theta_j} b(X^{\varepsilon}_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds dt \\
&\qquad - \int_0^T \partial_{\theta_i} b(x_t, \theta_0) \int_0^t \frac{\partial_{\theta_j} b(x_t, \theta_0) - \partial_{\theta_j} b(x_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds dt \Bigg) \\
&+ c_3 \Bigg(\int_0^T \left(\int_0^t \frac{\partial_{\theta_i} b(X^{\varepsilon}_t, \theta_0) - \partial_{\theta_i} b(X^{\varepsilon}_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds \right) \left(\int_0^t \frac{\partial_{\theta_j} b(X^{\varepsilon}_t, \theta_0) - \partial_{\theta_j} b(X^{\varepsilon}_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds \right) dt \\
&\qquad - \int_0^T \left(\int_0^t \frac{\partial_{\theta_i} b(x_t, \theta_0) - \partial_{\theta_i} b(x_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds \right) \left(\int_0^t \frac{\partial_{\theta_j} b(x_t, \theta_0) - \partial_{\theta_j} b(x_s, \theta_0)}{(t-s)^{H+1/2}} s^{1/2-H} ds \right) dt \Bigg).
\end{aligned}
$$

Using Lemma 3.1, we have

$$E\left|\int_0^T t^{1-2H}\left\{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_i}b(x_t,\theta_0)\partial_{\theta_j}b(x_t,\theta_0)\right\}dt\right|^p$$

$$\lesssim E\left|\int_0^T t^{1-2H}\left\{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)\left(\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(x_t,\theta_0)\right)+\partial_{\theta_j}b(x_t,\theta_0)\left(\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_i}b(x_t,\theta_0)\right)\right\}\right|^p$$

$$\lesssim E\left|\int_0^T t^{1-2H}(1+|X_t^\varepsilon|^N+|x_t|^N)^2\,|X_t^\varepsilon-x_t|\right|^p\lesssim\varepsilon^p.$$

We shall estimate the second term. Note that

$$\begin{aligned}
&\left\|\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)-\partial_{\theta_j}b(x_t,\theta_0)+\partial_{\theta_j}b(x_s,\theta_0)\right\|_{L^p(\Omega)}\\
&\lesssim\left(\left\|\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)\right\|_{L^p(\Omega)}+\left\|\partial_{\theta_j}b(x_t,\theta_0)-\partial_{\theta_j}b(x_s,\theta_0)\right\|_{L^p(\Omega)}\right)^{1/2}\\
&\times\left(\left\|\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(x_s,\theta_0)\right\|_{L^p(\Omega)}+\left\|\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)+\partial_{\theta_j}b(x_s,\theta_0)\right\|_{L^p(\Omega)}\right)^{1/2}\\
&\lesssim\varepsilon^{1/2}|t-s|^{H/2}.
\end{aligned}$$

Thus we obtain that

$$\begin{aligned}
&E\left|\int_0^T\left(\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)\int_0^t\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{(t-s)^{H+1/2}}s^{1/2-H}ds\right.\right.\\
&\qquad\left.\left.-\,\partial_{\theta_i}b(x_t,\theta_0)\int_0^t\frac{\partial_{\theta_j}b(x_t,\theta_0)-\partial_{\theta_j}b(x_s,\theta_0)}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)dt\right|\\
&=E\left|\int_0^T\left(\left(\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_i}b(x_t,\theta_0)\right)\int_0^t\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{(t-s)^{H+1/2}}s^{1/2-H}ds\right.\right.\\
&\qquad\left.\left.+\,\partial_{\theta_i}b(x_t,\theta_0)\int_0^t\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)-\partial_{\theta_j}b(x_t,\theta_0)+\partial_{\theta_j}b(x_s,\theta_0)}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)dt\right|^p\\
&\lesssim\left(\int_0^T\left\|(1+|X_t^\varepsilon|^N+|x_t|^N)|X_t^\varepsilon-x_t|\right\|_{L^{2p}(\Omega)}\int_0^t\frac{\left\|\left(1+|X_t^\varepsilon|^N+|X_s^\varepsilon|^N\right)|X_t^\varepsilon-X_s^\varepsilon|\right\|_{L^{2p}(\Omega)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)^p\\
&+\left(\int_0^T(1+|x_t|^N)\int_0^t\frac{\left\|\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)-\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)-\partial_{\theta_j}b(x_t,\theta_0)+\partial_{\theta_j}b(x_s,\theta_0)\right\|_{L^p(\Omega)}}{(t-s)^{H+1/2}}s^{1/2-H}dsdt\right)^p\\
&\lesssim\varepsilon^p+\varepsilon^{p/2}.
\end{aligned}$$

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

Lemma 3.3. *For every $p \geq 2$,*

$$\sup_{0<\varepsilon<1} E\left[\sup_{\theta\in\Theta}\left|\varepsilon^2\nabla_\theta^3\mathbb{L}_{H,\varepsilon}(\theta)\right|^p\right] < \infty.$$

Proof. By Sobolev's inequality, it follows for every $p > d$ that

$$\sup_{\theta\in\Theta}\left|\nabla_\theta^3\mathbb{L}_{H,\varepsilon}(\theta_0)\right|^p \lesssim \int_\Theta \left(\left|\nabla_\theta^3\mathbb{L}_{H,\varepsilon}(\theta)\right|^p + \left|\nabla_\theta^4\mathbb{L}_{H,\varepsilon}(\theta)\right|^p\right) d\theta.$$

By the same argument as in the proof of Lemma 3.2, we can show that

$$\sup_{0<\varepsilon<1} \varepsilon^{2p} E\left[\int_\Theta \left(\left|\nabla_\theta^3\mathbb{L}_{H,\varepsilon}(\theta)\right|^p + \left|\nabla_\theta^4\mathbb{L}_{H,\varepsilon}(\theta)\right|^p\right) d\theta\right] < \infty.$$

□

The proof of the following Lemma 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

$$\sup_{0<\varepsilon<1} E\left(\sup_{\theta\in\Theta}\varepsilon^{-1}\left|\mathbb{Y}_{H,\varepsilon}(\theta) - \mathbb{Y}_H(\theta)\right|\right)^p < \infty.$$

From Lemma 3.4 and the proof of Lemma 3.2, we obtain that

$$\varepsilon\nabla_\theta\mathbb{L}_{H,\varepsilon}(\theta_0) \xrightarrow{d} N(0, \Gamma_H(\theta_0)), \tag{3.4}$$

as $\varepsilon \to 0$ by the martingale central limit theorem. Moreover, Lemmas 3.2, 3.3 and the convergence (3.4) give the local asymptotic normality of $Z_{H,\varepsilon}(u)$:

$$\mathbb{Z}_{H,\epsilon}(u) \xrightarrow{d} \mathbb{Z}_H(u) := \exp\left(\Delta_H(\theta_0)u - \frac{1}{2}\Gamma_H(\theta_0)[u,u]\right),$$

where $\Delta_H(\theta_0) \sim N(0, \Gamma_H(\theta_0))$. By Lemma 3.2, 3.3 and 3.4, we can apply Theorem 2.1 and yields the inequality

$$\sup_{0<\varepsilon<1} P\left[\sup_{|u|\geq r}\mathbb{Z}_{H,\varepsilon}(u) \geq e^{-r}\right] \lesssim r^{-L}, \tag{3.5}$$

hold for any $r > 0$ and $L > 0$. Since $u_\varepsilon := \varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0)$ maximizes the random field $\mathbb{Z}_{H,\varepsilon}$, 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,

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

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

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

In the sequel, we prove that

$$\log \mathbb{Z}_{H,\varepsilon} \xrightarrow{d} \log \mathbb{Z}_H \quad \text{in } C(B(R)), \tag{3.6}$$

as $\varepsilon \to 0$. If we can show the convergence (3.6), we obtain the asymptotic normality:

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

as $\varepsilon \to 0$ by Theorem 2.4. Due to linearity in u of the weak convergence term $\varepsilon \nabla_\theta \mathbb{L}_{H,\varepsilon}(\theta_0)[u]$, the convergence of finite-dimensional distribution holds true. It remains to show the tightness of the family $\{\log \mathbb{Z}_{H,\varepsilon}(u)\}_{u \in B(R)}$. By the Kolmogorov tightness criterion (see [15]), it suffices to show that for every $R > 0$ there exists a constant $p > 0$, $\gamma > d$ and $C > 0$ such that

$$E\left|\log \mathbb{Z}_{H,\varepsilon}(u_1) - \log \mathbb{Z}_{H,\varepsilon}(u_2)\right|^p \leq C|u_1 - u_2|^\gamma, \tag{3.7}$$

for $u_1, u_2 \in B(R)$. For a number $p > 0$ large enough, the inequality (3.7) is shown easily by Lemmas 3.2, 3.3 and 3.4. Therefore, we complete the proof.

4 Multiplicative noise case

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

$$X_t^\varepsilon = X_0 + \int_0^t b(X_s^\varepsilon, \theta_0)\, ds + \varepsilon \int_0^t \sigma(X_s^\varepsilon)\, dB_s^H, \quad t \in [0,T], \tag{4.1}$$

where $X_0 \in \mathbb{R}$ is the initial value, $\theta_0 \in \Theta$ is the parameter which is contained in a bounded and open convex subset $\Theta \subset \mathbb{R}^d$ admitting Sobolev's inequalities for embedding $W^{1,p}(\Theta) \hookrightarrow C(\bar{\Theta})$. The stochastic process $\{B_t^H\}_{t\in[0,T]}$ is a fBm with Hurst index $H \in (1/2, 1)$ and the integral with respect to B^H is defined as a pathwise Riemann-Stieltjes integral. Without loss of generality, we assume that $\varepsilon \in (0,1]$. The main purpose in this section is the estimation of parameter $\theta_0 \in \Theta$ from a realization $\{X_t^\varepsilon\}_{t\in[0,T]}$ when $\varepsilon \to 0$.

There are some methods to define the stochastic integral with respect to the fBm. One of them is the pathwise approach. Since we can expect the solution to (4.1) whose trajectories are the same Hölder continuity to the fBm, the Riemann-Stieltjes integral $\int_0^t \sigma(X_s^\varepsilon)\, dB_s^H$ in (4.1) can be defined in the sense of the Young integral. Under this interpretation, Nualart and Răşcanu [34] showed the existence and uniqueness of the solution to (4.1). They rewrote the stochastic integral as generalized Stieltjes integral and derived to apriori estimate to guarantee global existence of a solution to (4.1). In order to obtain asymptotic properties of the estimator under $\varepsilon \to 0$, we rewrite the stochastic integral as generalized Stieltjes integral in the same way as [34]. Then we can get some estimates and asymptotic behavior of the solution to (4.1). We first prepare these evaluations and then apply the method investigated by Ibragimov and Has'minskii [15]. We aim to deduce asymptotic normality and moment convergence of the maximum likelihood type estimators of the drift parameter under $\varepsilon \to 0$.

4.1 Generalized Riemann–Stieltjes integral

We define the pathwise integral with respect to fractional Brownian motion appearing in equation (4.1). Let f and g are functions such that the limits $f(a+) := \lim_{\delta\downarrow 0} f(a+\delta)$, $g(b-) := \lim_{\delta\downarrow 0} g(b-\delta)$ exist. Let

$$f_{a+}(x) := (f(x) - f(a+))\, 1_{(a,b)}(x)$$
$$g_{b-}(x) := (g(b-) - g(x))\, 1_{(a,b)}(x).$$

Suppose that $f(a+) \in I_{a+}^\alpha(L^p(a,b))$ and $g(b-) \in L^q(a,b)$ for some $p, q \geq 0$, $1/p + 1/q \leq 1$ and $0 < \alpha < 1$. The generalized Stieltjes integral is defined as

$$\int_a^b f(x) dg(x) := \int_a^b D_{a+}^\alpha f_{a+}(x) D_{b-}^{1-\alpha} g_{b-}(x)\, dx + f(a+)\left(g(b-) - g(a+)\right),$$

and if $\alpha p < 1$, it can be rewritten by

$$\int_a^b f(x)dg(x) = \int_a^b D_{a+}^{\alpha} f(x) D_{b-}^{1-\alpha} g_{b-}(x)\, dx.$$

If $f \in C^{\lambda}[a,b]$ and $g \in C^{\mu}[a,b]$ with $\lambda + \mu > 1$, the generalized Riemann–Stieltjes integral $\int_a^b f(x)dg(x)$ coincides with the Riemann–Stieltjes integral. Since the fractional Brownian motion has λ-Hölder continuous trajectiries with $\lambda \in (0, H)$, we can define the integral for $1 - H < \alpha < 1/2$ and $f \in I_{a+}^{\alpha}(L^1(a,b))$,

$$\int_a^b f(x)dB^H(x) := \int_a^b D_{a+}^{\alpha} f(x) D_{b-}^{1-\alpha} B_{b-}^H(x)\, dx,$$

For $1 - \alpha < \beta < H$, we can show that

$$\begin{aligned} \left| \int_a^b f(x)dB^H(x) \right| &\le \sup_{a \le x \le b} \left| D_{b-}^{1-\alpha} B_{b-}^H(x) \right| \int_a^b \left| D_{a+}^{\alpha} f(x) \right| \, dx \\ &\le \frac{\beta(b-a)^{\alpha+\beta-1}}{\Gamma(\alpha)(\alpha+\beta-1)} \|B_\cdot^H\|_{\beta,[a,b]} \int_a^b \left| D_{a+}^{\alpha} f(x) \right| \, dx, \end{aligned} \tag{4.2}$$

where the second inequality follows from the following evaluation:

$$\begin{aligned} \left| D_{b-}^{1-\alpha} B_{b-}^H(x) \right| &= \frac{1}{\Gamma(\alpha)} \left| \frac{B_b^H - B_x^H}{(b-x)^{1-\alpha}} + (1-\alpha) \int_x^b \frac{B_x - B_y}{(y-x)^{2-\alpha}} dy \right| \\ &\le \frac{\|B_\cdot^H\|_{\beta,[a,b]}}{\Gamma(\alpha)} \left| \frac{1}{(b-x)^{1-\alpha-\beta}} + (1-\alpha) \int_x^b \frac{1}{(y-x)^{2-\alpha-\beta}} dy \right| \\ &\le \frac{\beta(b-a)^{\alpha+\beta-1}}{\Gamma(\alpha)(\alpha+\beta-1)} \|B_\cdot^H\|_{\beta,[a,b]}. \end{aligned}$$

4.2 Construction of the maximum likelihood type estimator

We impose some assumptions on coefficients b and σ in order to derive the likelihood function.

Assumption 4.1. *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 $p > d$.*

Assumption 4.2. *The function $b(\cdot, \theta)$ in (4.1) is of $C^1(\mathbb{R})$-class and there exist constants $c, N > 0$ such that, for every $x, y \in \mathbb{R}$ and $\theta \in \Theta$,*

$$\sup_{\theta \in \Theta} |b(x,\theta) - b(y,\theta)| \le c|x-y|, \;\; |\partial_x b(x,\theta)| \le c(1 + |x|^N).$$

Assumption 4.3. *The function $\sigma(\cdot)$ in (4.1) is of $C^1(\mathbb{R})$-class and there exist constants $\gamma, \delta \in [0,1]$ and $c, N > 0$ such that for every $x, y \in \mathbb{R}$ and integer $0 \le i \le 1$,*

$$|\sigma(x)| \le c(1+|x|^\gamma), \ \ |\partial_x^i \sigma(x) - \partial_x^i \sigma(y)| \le c|x-y|, \ \ |\partial_x \sigma(x)| \le c(1+|x|^N).$$

Assumption 4.4. *The function $\frac{b(\cdot,\cdot)}{\sigma(\cdot)}$ is of $C^{2,4}(\mathbb{R}\times\Theta;\mathbb{R})$-class, and there exist constants $c, N > 0$ such that for every $x \in \mathbb{R}$, $\theta \in \Theta$ and integers $0 \le j \le 2$ and $0 \le i \le 4$, it holds that*

$$\left|\partial_x^j \nabla_\theta^i \left(\frac{b(x,\theta)}{\sigma(x)}\right)\right| \le c\left(1+|x|^N\right).$$

Assumption 4.5. *There exists a random variable ξ such that $P(\xi < \infty) = 1$ and*

$$\sup_{t\in[0,T]} \frac{1}{|\sigma(X_t^\varepsilon)|} \le \xi.$$

Since, for every $1-H < \alpha < 1/2 \wedge (2-\gamma)/4$, the solution to (4.1) has $(1-\alpha)$-Hölder continuity under Assumptions 4.2 and 4.3 (see [34]), we can define a function under Assumptions 4.2-4.4,

$$Q_{H,\theta}^\varepsilon(t) := (\varepsilon d_H)^{-1}\, t^{H-1/2} D_{0+}^{H-1/2}\left[(\cdot)^{1/2-H}\frac{b(X_\cdot^\varepsilon,\theta)}{\sigma(X_\cdot^\varepsilon)}\right](t),$$

where

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

In the sequel, we construct a maximum likelihood type estimator under Assumptions 4.2-4.5. Consider the stochastic process

$$Y_t := \int_0^t \sigma^{-1}(X_s^\varepsilon) dX_s = \int_0^t \frac{b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)} ds + \varepsilon B_t^H,$$

We set a semimartingale $\{Z_t\}_{t\ge0}$ as follows:

$$\begin{aligned} Z_t :&= \varepsilon^{-1}\int_0^t k_H^{-1}(t,s)\, dY_s \\ &= \int_0^t Q_{H,\theta}^\varepsilon(s)\, ds + W_t. \end{aligned}$$

Define the probability measure $dP^* := \Lambda_T^{\theta_0} dP$, where

$$\Lambda_T^{\theta_0} := \exp\left(-\int_0^T Q_{H,\theta_0}^{\varepsilon}(s)\, dW_s - \frac{1}{2}\int_0^T Q_{H,\theta_0}^{\varepsilon}(s)^2\, ds \right).$$

Now, let us assume that

$$E\Lambda_T^{\theta_0} = 1, \tag{4.3}$$

one of the sufficient conditions for which is given in Theorem 4.3 later. Then, thanks to the Girsanov theorem, a stochastic process $\{Z_t\}_{t\in[0,T]}$ is an $(\mathcal{F}_t)$-Brownian motion under P^*. Therefore, we see that

$$\int_0^t k_H(t,s)\, dZ_s = \varepsilon^{-1} Y_t,$$

is an $(\mathcal{F}_t)$-fractional Brownian motion.

Let $X^{\varepsilon,0}$ be the solution to the following SDE:

$$X_t^{\varepsilon,0} = X_0 + \varepsilon \int_0^t \sigma(X_s^{0,\varepsilon})\, dB_s^H,$$

and let $P_{\theta_0}^{\varepsilon}$ and P_0^{ε} be probability measures induced by processes $\{X_t^{\varepsilon}\}_{t\in[0,T]}$ and $\{X_t^{\varepsilon,0}\}_{t\in[0,T]}$, respectively. Note that we have the equality P^*-a.s.

$$X_0 + \int_0^t \sigma(X_s^{\varepsilon})\, dY_s = X_0 + \int_0^t b(X_s^{\varepsilon}, \theta_0)\, ds + \varepsilon \int_0^t \sigma(X_s^{\varepsilon})\, dB_s^H.$$

Hence for every $A \in \mathcal{B}(C[0,T])$,

$$P_{\theta_0}^{\varepsilon}(A) = P^*(X^{\varepsilon} \in A) = P(X^{\varepsilon,0} \in A) = \int_{\{X^{\varepsilon}\in A\}} \left(\Lambda_T^{\theta_0}(\omega)^{-1}\right)\, dP^*(\omega),$$

and we find that the Radon-Nikodym derivative of $P_{\theta_0}^{\varepsilon}$ with respect to P_0^{ε} is given by

$$\frac{dP_{\theta_0}^{\varepsilon}}{dP_0^{\varepsilon}}(X^{\varepsilon}) = \exp\left(\int_0^T Q_{H,\theta_0}^{\varepsilon}(s) dW_s + \frac{1}{2}\int_0^T Q_{H,\theta_0}^{\varepsilon}(s)^2\, ds \right). \tag{4.4}$$

We consider the log-likelihood function

$$\mathbb{L}_{H,\varepsilon}(\theta) := \int_0^T Q_{H,\theta}^{\varepsilon}(t) dZ_t - \frac{1}{2}\int_0^T Q_{H,\theta}^{\varepsilon}(t)^2 dt.$$

The maximum likelihood type estimator $\hat{\theta}_\varepsilon$ is defined as

$$\hat{\theta}_\varepsilon := \arg\max_{\theta \in \bar{\Theta}} \mathbb{L}_{H,\varepsilon}(\theta), \tag{4.5}$$

under (4.3).

Remark 4.1. Note that the above argument is under the condition (4.3), which is to be proved so that $\hat{\theta}_\varepsilon$ becomes the "true" MLE. Nevertheless, we consider the estimator (4.5) in the sequel as an M-estimator without care of (4.3) while $\hat{\theta}_\varepsilon$ is well-defined. We need some additional conditions for coefficients b and σ to ensure (4.3), some of those is given in Section 4.5.

4.3 Main results

To describe our main results, we make some notations.

Let $\{x_t\}_{0\le t\le T}$ be the solution to the differential equation under the true value of the drift parameter:

$$\begin{cases} \dfrac{dx_t}{dt} = b(x_t, \theta_0) \\ \ \ x_0 = X_0. \end{cases} \tag{4.6}$$

We set the d-dimensional square matrix $\Gamma_H(\theta_0)$ which is an asymptotic convariance matrix of the estimator as

$$\begin{aligned} \Gamma_H^{i,j}(\theta_0) := \int_0^T & \left(c_1 t^{1/2-H} \frac{\partial_{\theta_i} b(x_t,\theta_0)}{\sigma(x_t)} + c_2 t^{H-1/2} \int_0^t \frac{\frac{\partial_{\theta_i} b(x_t,\theta_0)}{\sigma(x_t)} - \frac{\partial_{\theta_i} b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}} s^{1/2-H} ds \right) \\ & \times \left(c_1 t^{1/2-H} \frac{\partial_{\theta_j} b(x_t,\theta_0)}{\sigma(x_t)} + c_2 t^{H-1/2} \int_0^t \frac{\frac{\partial_{\theta_j} b(x_t,\theta_0)}{\sigma(x_t)} - \frac{\partial_{\theta_j} b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}} s^{1/2-H} ds \right) dt, \end{aligned}$$

where

$$\begin{aligned} c_1 &= (d_H \Gamma(3/2-H))^{-1} \left\{ 1 + (H-1/2) \int_0^1 \frac{s^{1/2-H}-1}{(1-s)^{H+1/2}} ds \right\} \\ c_2 &= (H-1/2) \left(d_H \Gamma(3/2-H) \right)^{-1}. \end{aligned}$$

Assumption 4.6. *The matrix $\Gamma_H(\theta_0)$ is positive definite.*

In order to guarantee the asymptotic properties of the estimator, we need to impose the identifibility condition. Define

$$\mathbb{Y}_{H,\varepsilon}(\theta) := \varepsilon^2 \left(\mathbb{L}_{H,\varepsilon}(\theta) - \mathbb{L}_{H,\varepsilon}(\theta_0) \right),$$

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

$$\mathbb{Y}_H(\theta) := -\frac{1}{2}\int_0^T \Bigg(c_1 t^{1/2-H}\left(\frac{b(x_t,\theta)}{\sigma(x_t)} - \frac{b(x_t,\theta_0)}{\sigma(x_t)}\right) \\ + c_2 t^{H-1/2}\int_0^t \frac{\left(\frac{b(x_t,\theta)}{\sigma(x_t)} - \frac{b(x_t,\theta_0)}{\sigma(x_t)}\right) - \left(\frac{b(x_s,\theta)}{\sigma(x_s)} - \frac{b(x_s,\theta_0)}{\sigma(x_s)}\right)}{(t-s)^{H+1/2}} s^{1/2-H} ds \Bigg)^2 dt.$$

Assumption 4.7. *There exists a positive constant $\xi(\theta_0) > 0$ such that*

$$\mathbb{Y}_H(\theta) \le -\xi(\theta_0)|\theta - \theta_0|^2,$$

for every $\theta \in \Theta$.

The following theorem gives the asymptotic distribution of the estimator $\hat{\theta}_\varepsilon$.

Theorem 4.1. *Suppose the Assumptions 4.1–4.7. Then it holds for $\hat{\theta}_\varepsilon$ in* (4.5) *that*

$$\varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0) \xrightarrow{d} \xi, \quad \varepsilon \to 0,$$

where $\xi \sim N(0, \Gamma_H(\theta_0)^{-1})$. Moreover, we have

$$E\left[f\left(\varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0)\right)\right] \to E[f(\xi)], \quad \varepsilon \to 0,$$

for every continuous function f of polynomial growth.

Remark 4.2. Let us remark that the Theorem 4.1 can be generalized to the multi-dimensional stochastic differential equation considered as in [34]. In fact, the key estimates of the proof of the Theorem 4.1, Lemma 4.1 and 4.2, hold in the multi-dimensional case. However, to avoid complications of the calculation, we have restricted the assertion of Theorem 4.1 to the one-dimensional case.

4.4 Examples

We supply some examples of Thereom 4.1.

Example 4.1. Let us consider the drift and diffusion functions as follows:

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

with $\theta \in (m, M),\ 0 < m < M,\ X_0 > 0$. Then the function $x.$ satisfies the following equation:

$$\begin{cases} \dfrac{dx_t}{dt} = \dfrac{\theta_0 x_t}{1+x_t^2} \\ x_0 = X_0. \end{cases}$$

In this case, we can check Assumptions of Theorem 4.1 and 4.2. It is easy to verify that Assumptions 4.2-4.5 are satisfied. We check Assumption 4.7. Since $X_0 > 0$, the function $x.$ is monotone increasing and we can evaluate for every $t \in [0, T]$,

$$X_0 \le x_t = X_0 + \theta_0 \int_0^t \frac{x_s}{1+x_s^2} ds \le X_0 + \theta_0 \int_0^t x_s ds,$$

and by Gronwall's inequality, we have

$$x_t \le X_0 e^{\theta_0 t}.$$

Recall that $\mathbb{Y}_H(\theta)$ is given by

$$\begin{aligned} -\mathbb{Y}_H(\theta) &:= \frac{1}{2}\int_0^T \Bigg(c_1 t^{1/2-H} \left(\frac{\theta x_t}{\sqrt{1+x_t^2}} - \frac{\theta_0 x_t}{\sqrt{1+x_t^2}} \right) \\ &\qquad + c_2 t^{H-1/2} \int_0^t \frac{\left(\frac{\theta x_t}{\sqrt{1+x_t^2}} - \frac{\theta_0 x_t}{\sqrt{1+x_t^2}} \right) - \left(\frac{\theta x_s}{\sqrt{1+x_s^2}} - \frac{\theta_0 x_s}{\sqrt{1+x_s^2}} \right)}{(t-s)^{H+1/2}} s^{1/2-H} ds \Bigg)^2 dt \\ &= \frac{(\theta-\theta_0)^2}{2} \int_0^T \Bigg(c_1^2 t^{1-2H} \frac{x_t^2}{1+x_t^2} + 2c_1c_2 \frac{x_t}{\sqrt{1+x_t^2}} \int_0^t \frac{\frac{x_t}{\sqrt{1+x_t^2}} - \frac{x_s}{\sqrt{1+x_s^2}}}{(t-s)^{H+1/2}} s^{1/2-H} ds \\ &\qquad + c_2^2 t^{2H-1} \left(\int_0^t \frac{\frac{x_t}{\sqrt{1+x_t^2}} - \frac{x_s}{\sqrt{1+x_s^2}}}{(t-s)^{H+1/2}} s^{1/2-H} ds \right)^2 \Bigg) dt. \end{aligned} \tag{4.7}$$

The first term in the last inequality of (4.7) can be bounded as

$$\int_0^T t^{1-2H} \frac{x_t^2}{1+x_t^2} dt \ge \frac{X_0^2}{1+X_0^2 e^{2\theta_0 T}} \int_0^T t^{1-2H} dt = \frac{X_0^2 T^{2-2H}}{(2-2H)(1+X_0^2 e^{2\theta_0 T})}.$$

We estimate the second term of (4.7). By the change of the variable formula, we have

$$
\begin{aligned}
\frac{x_t}{\sqrt{1+x_t^2}}-\frac{x_s}{\sqrt{1+x_s^2}} &= \int_s^t \left\{\frac{1}{\sqrt{1+x_u^2}}-\frac{x_u^2}{(1+x_u^2)^{3/2}}\right\} dx_u \\
&= \theta_0 \int_s^t \frac{1}{(1+x_u^2)^{3/2}} \frac{x_u}{1+x_u^2} du \\
&\geq \frac{\theta_0 X_0}{\left(1+X_0^2 e^{2\theta_0 T}\right)^{5/2}} (t-s).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\int_0^T \frac{x_t}{\sqrt{1+x_t^2}} \left(\int_0^t \frac{\frac{x_t}{\sqrt{1+x_t^2}}-\frac{x_s}{\sqrt{1+x_s^2}}}{(t-s)^{H+1/2}} s^{1/2-H} ds\right) dt \\
&\geq \frac{\theta_0 X_0}{\left(1+X_0^2 e^{2\theta_0 T}\right)^{5/2}} \int_0^T \frac{x_t}{\sqrt{1+x_t^2}} \left(\int_0^t (t-s)^{1/2-H} s^{1/2-H} ds\right) dt \\
&\geq \frac{\theta_0 X_0^2 \beta\left(3/2-H, 3/2-H\right)}{\left(1+X_0^2 e^{2\theta_0 T}\right)^3} \int_0^T t^{2-2H} dt = \frac{\theta_0 X_0^2 \beta\left(3/2-H, 3/2-H\right) T^{3-2H}}{\left(3-2H\right)\left(1+X_0^2 e^{2\theta_0 T}\right)^3}.
\end{aligned}
$$

In the same way, we can evaluate the third term of (4.7) and Assumption 4.7 holds true. Assumption 4.6 can be confirmed by the same calculation. We check the Assumption 4.8. The boundedness of the function σ itself and its derivative are obvious. We show that the Hölder continuity of the function b/σ. For every $x, y \in \mathbb{R}$, $\theta \in (m, M)$ and $\lambda \in (1-\frac{1}{2H}, 1/2)$,

$$
\begin{aligned}
\theta \left|\frac{x}{\sqrt{1+x^2}}-\frac{y}{\sqrt{1+y^2}}\right| &= \theta \left|\frac{x}{\sqrt{1+x^2}}-\frac{y}{\sqrt{1+y^2}}\right|^{\lambda+(1-\lambda)} \\
&\leq M 2^{1-\lambda} \left|\frac{x}{\sqrt{1+x^2}}-\frac{y}{\sqrt{1+y^2}}\right|^{\lambda},
\end{aligned}
$$

and by the mean value theorem, we conclude that

$$
\theta \left|\frac{x}{\sqrt{1+x^2}}-\frac{y}{\sqrt{1+y^2}}\right| \leq 2M|x-y|^{\lambda}.
$$

Therefore, in this case, the estimator $\hat{\theta}_\varepsilon$ is the MLE and has asymptotic properties described in Theorem 4.1.

Example 4.2. Let $d = 1,\ H \in (3/4, 1)$ and consider the fractional geometric Brownian motion that is, the drift and the diffusion functions are given by

$$b(x, \theta) = \theta x, \quad \sigma(x) = x,$$

with $\theta \in \Theta$. Then we can confirm that Assumptions 4.2-4.5 are satisfied. Moreover, for every $\theta \in \Theta$, the asymptotic variance $\Gamma_H(\theta_0)$ and $\mathbb{Y}_H(\theta)$ have the following form:

$$\begin{aligned}\Gamma_H(\theta_0) &= c_1^2 \theta_0^2 \int_0^T t^{1-2H} dt = \frac{c_1^2 T^{2-2H}}{2-2H} \theta_0^2 \\ \mathbb{Y}_H(\theta) &= -\frac{c_1^2 T^{2-2H}}{2(2-2H)} (\theta - \theta_0)^2.\end{aligned}$$

Thus Assumptions 4.6, 4.7 are also satisfied. On the other hands, since the fuction $\sigma(x) = x$ is not bounded, we can not apply Theorem 4.2. However, in this case, we can confirm that $\hat{\theta}_\varepsilon$ is the true MLE. Indeed, the function $Q^\varepsilon_{H,\theta_0}$ has the form

$$Q^\varepsilon_{H,\theta_0}(t) = c_1 \varepsilon^{-1} t^{1/2-H} \theta_0,$$

and the condition (4.21) holds true. Therefore, $\hat{\theta}_\varepsilon$ is the MLE and has asymptotic properties described in Theorem 4.1.

4.5 Proofs

We first establish a lemma on the evaluation of X^ε.

Lemma 4.1. *Under Assumptions 4.1-4.3, for every $p > 0$, there exist constants $C_i > 0,\ i = 1, 2, 3, 4$ depending on $T > 0,\ \alpha \in (1 - H, 1/2 \wedge (2-\gamma)/4)$ and $\beta \in (1 - \alpha, H)$ such that for $s, t \in [0, T]$,*

$$E|X_t^\varepsilon - x_t|^p \le C_1 \varepsilon^p \tag{4.8}$$

$$E|X_t^\varepsilon|^p \le C_2 (1 + \varepsilon^p) \tag{4.9}$$

$$E\left(\int_0^t \frac{|X_t^\varepsilon - X_s^\varepsilon|}{(t-s)^{\alpha+1}} ds\right)^p \le C_3 (1 + \varepsilon^p), \tag{4.10}$$

and

$$E|X_t^\varepsilon - X_s^\varepsilon|^p \le C_4 |t-s|^{p(H-\beta+1-\alpha)} \le C_4 T^{p(H-\beta+1/2-\alpha)} |t-s|^{\frac{pH}{2}}. \tag{4.11}$$

Proof. By the equations (4.1) and (4.6), we see that

$$|X_t^\varepsilon - x_t| \le \int_0^t |b(X_s^\varepsilon, \theta_0) - b(x_s, \theta_0)| ds + \varepsilon \left| \int_0^t \sigma(X_s^\varepsilon) dB_s^H \right| := \Psi_t + \Phi_t.$$

Note that it follows from the Lipschitz continuity of the function b that

$$|\Psi_t| \lesssim \int_0^t |X_s^\varepsilon - x_s| ds. \tag{4.12}$$

Hence Fubini's theorem yields that

$$\begin{aligned} \int_0^t \frac{|\Psi_t - \Psi_s|}{(t-s)^{1+\alpha}} ds &\lesssim \int_0^t (t-s)^{-1-\alpha} \left(\int_s^t |X_u^\varepsilon - x_u| du \right) ds \\ &= \int_0^t |X_u^\varepsilon - x_u| \left(\int_0^u (t-s)^{-1-\alpha} ds \right) du \\ &\lesssim \int_0^t |X_u^\varepsilon - x_u| (t-u)^{-\alpha} du. \end{aligned} \tag{4.13}$$

As for the stochastic integral term Φ_t, we see from Assumption 4.3 and the inequality (4.2) that

$$\begin{aligned} |\Phi_t| &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \int_0^t |D_{0+}^\alpha \sigma(X_\cdot^\varepsilon)(s)| ds \\ &= \varepsilon \|B^H\|_{\beta,[0.t]} \int_0^t \left| \frac{\sigma(X_s^\varepsilon)}{s^\alpha} + \alpha \int_0^s \frac{\sigma(X_s^\varepsilon) - \sigma(X_u^\varepsilon)}{(s-u)^{1+\alpha}} du \right| ds \\ &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \left\{ t^{1-\alpha} + \int_0^t \left(\frac{|X_s^\varepsilon|^\gamma}{s^\alpha} + \alpha \int_0^s \frac{|X_s^\varepsilon - X_u^\varepsilon|}{(s-u)^{1+\alpha}} du \right) ds \right\} \\ &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \left\{ 1 + \int_0^t \left(\frac{|X_s^\varepsilon - x_s|^\gamma + |x_s|^\gamma}{s^\alpha} + \int_0^s \frac{|X_s^\varepsilon - x_s - X_u^\varepsilon - x_u| + |x_s - x_u|}{(s-u)^{1+\alpha}} du \right) ds \right\}. \end{aligned}$$

We use the mean value theorem to obtain that

$$\begin{aligned} &\varepsilon \|B^H\|_{\beta,[0.t]} \left\{ 1 + \int_0^t \left(\frac{|X_s^\varepsilon - x_s|^\gamma + |x_s|^\gamma}{s^\alpha} + \int_0^s \frac{|X_s^\varepsilon - x_s - X_u^\varepsilon - x_u| + |x_s - x_u|}{(s-u)^{1+\alpha}} du \right) ds \right\} \\ &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \Bigg\{ 1 + \sup_{s\in[0,T]} |x_s|^\gamma t^{1-\alpha} + \sup_{s\in[0,T]} |\dot{x}_s| t^{2-\alpha} \\ &\qquad + \int_0^t \left(\frac{|X_s^\varepsilon - x_s|^\gamma}{s^\alpha} + \int_0^s \frac{|X_s^\varepsilon - x_s - X_u^\varepsilon - x_u|}{(s-u)^{1+\alpha}} du \right) ds \Bigg\}, \end{aligned}$$

and by the Hölder's inequality, we can get

$$|\Phi_t| \lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \left\{ 1 + \int_0^t \left(\frac{|X_s^\varepsilon - x_s|}{s^\alpha} + \int_0^s \frac{|X_s^\varepsilon - x_s - X_u^\varepsilon - x_u|}{(s-u)^{1+\alpha}} du \right) ds \right\}. \quad (4.14)$$

On the other hand, it follows by the additivity of integral that

$$\begin{aligned}
|\Phi_t - \Phi_s| &= \varepsilon \left| \int_s^t \sigma(X_u^\varepsilon) dB_u \right| \\
&= \left| \int_s^t D_{s+}^\alpha \sigma(X_\cdot^\varepsilon)(u) D_{t-}^{1-\alpha} B^H(u) du \right| \\
&\leq \varepsilon \|B^H\|_{\beta,[0.t]} \left| \int_s^t \left(\frac{\sigma(X_u^\varepsilon)}{(u-s)^\alpha} + \alpha \int_s^u \frac{\sigma(X_u^\varepsilon) - \sigma(X_v^\varepsilon)}{(u-v)^{\alpha+1}} dv \right) du \right|.
\end{aligned}$$

Thus, by Fubini's theorem, we obtain that

$$\begin{aligned}
\int_0^t \frac{|\Phi_t - \Phi_s|}{(t-s)^{1+\alpha}} ds &= \varepsilon \int_0^t (t-s)^{-1-\alpha} \left| \int_s^t \sigma(X_u^\varepsilon) dB_u^H \right| ds \\
&\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \int_0^t (t-s)^{-1-\alpha} \left| \int_s^t \left(\frac{\sigma(X_u^\varepsilon)}{(u-s)^\alpha} + \alpha \int_s^u \frac{\sigma(X_u^\varepsilon) - \sigma(X_v^\varepsilon)}{(u-v)^{\alpha+1}} dv \right) du \right| ds \\
&\leq \varepsilon \|B^H\|_{\beta,[0.t]} \left\{ \int_0^t |\sigma(X_u^\varepsilon)| \left(\int_0^u (t-s)^{-1-\alpha} (u-s)^{-\alpha} ds \right) du \right. \\
&\qquad \left. + \int_0^t \int_0^u \frac{|\sigma(X_u^\varepsilon) - \sigma(X_v^\varepsilon)|}{(u-v)^{\alpha+1}} \left(\int_0^v (t-s)^{-\alpha-1} ds \right) dv du \right\}.
\end{aligned}$$

Note that, by the change the variable $s = u - (t-u)y$, we have

$$\begin{aligned}
\int_0^u (t-s)^{-1-\alpha} (u-s)^{-\alpha} ds &= (t-u)^{-2\alpha} \int_0^{\frac{u}{t-u}} (1+y)^{-\alpha-1} y^{-\alpha} dy \\
&\leq (t-u)^{-2\alpha} \int_0^\infty (1+y)^{-\alpha-1} y^{-\alpha} dy = B(2\alpha, 1-\alpha)(t-u)^{-2\alpha}.
\end{aligned}$$

and

$$\int_0^v (t-s)^{-\alpha-1} ds = \alpha^{-1} \left\{ (t-v)^{-\alpha} - t^{-\alpha} \right\} \leq \alpha^{-1} (t-v)^{-\alpha}.$$

It follows that

$$\int_0^t \frac{|\Phi_t-\Phi_s|}{(t-s)^{1+\alpha}}ds \lesssim \varepsilon\|B^H\|_{\beta,[0.t]}\Bigg\{\int_0^t |\sigma(X_u^\varepsilon)|\,(t-u)^{-2\alpha}du + \int_0^t\left(\int_0^u (t-v)^{-\alpha}\frac{|\sigma(X_u^\varepsilon)-\sigma(X_v^\varepsilon)|}{(u-v)^{\alpha+1}}dv\right)du\Bigg\}$$
$$\lesssim \varepsilon\|B^H\|_{\beta,[0.t]}\left\{1+\int_0^t |X_u^\varepsilon|^\gamma\,(t-u)^{-2\alpha}du+\int_0^t\left(\int_0^u (t-v)^{-\alpha}\frac{|X_u^\varepsilon-X_v^\varepsilon|}{(u-v)^{\alpha+1}}dv\right)du\right\}.$$

By using the mean value theorem, we obtain that

$$\begin{aligned}
&\varepsilon\|B^H\|_{\beta,[0.t]}\left\{1+\int_0^t |X_u^\varepsilon|^\gamma\,(t-u)^{-2\alpha}du+\int_0^t\left(\int_0^u (t-v)^{-\alpha}\frac{|X_u^\varepsilon-X_v^\varepsilon|}{(u-v)^{\alpha+1}}dv\right)du\right\}\\
&\lesssim \varepsilon\|B^H\|_{\beta,[0.t]}\Bigg\{1+\sup_{u\in[0,T]}|x_u|^\gamma t^{1-2\alpha}+\sup_{u\in[0,T]}|\dot{x}_u|\int_0^t\left(\int_0^u (t-v)^{-\alpha}(u-v)^{-\alpha}dv\right)du\\
&\quad+\int_0^t |X_u^\varepsilon-x_u|^\gamma\,(t-u)^{-2\alpha}du+\int_0^t\left(\int_0^u (t-v)^{-\alpha}\frac{|X_u^\varepsilon-x_u-X_v^\varepsilon+x_v|}{(u-v)^{\alpha+1}}dv\right)du\Bigg\}\\
&\lesssim \varepsilon\|B^H\|_{\beta,[0.t]}\Bigg\{1+\int_0^t |X_u^\varepsilon-x_u|^\gamma\,(t-u)^{-2\alpha}du\\
&\quad+\int_0^t\left(\int_0^u (t-v)^{-\alpha}\frac{|X_u^\varepsilon-x_u-X_v^\varepsilon+x_v|}{(u-v)^{\alpha+1}}dv\right)du\Bigg\}.
\end{aligned}\tag{4.15}$$

Define

$$\kappa_\gamma := \begin{cases} 2\alpha & \text{if } \gamma=1\\ >1+\frac{2\alpha-1}{\gamma} & \text{if } \frac{1-2\alpha}{1-\alpha}\le\gamma<1\\ \alpha & \text{if } 0\le\gamma<\frac{1-2\alpha}{1-\alpha},\end{cases}$$

and we can show the following inequality

$$\begin{aligned}
\int_0^t \frac{|\Phi_t-\Phi_s|}{(t-s)^{1+\alpha}}ds \lesssim \varepsilon\|B^H\|_{\beta,[0.t]}\Bigg\{&1+\int_0^t |X_u^\varepsilon-x_u|\,(t-u)^{-\kappa_\gamma}du\\
&+\int_0^t (t-u)^{-\alpha}\left(\int_0^u \frac{|X_u^\varepsilon-x_u-X_v^\varepsilon+x_v|}{(u-v)^{\alpha+1}}dv\right)du\Bigg\}.
\end{aligned}\tag{4.16}$$

Indeed, if $\gamma = 1$, the inequality follws immediately from (4.15). If $\frac{1-2\alpha}{1-\alpha} \leq \gamma < 1$, by the Hölder's inequality with $\delta = \kappa_\gamma \gamma > \gamma + 2\alpha - 1$,

$$\begin{aligned}\int_0^t |X_u^\varepsilon - x_u|^\gamma (t-u)^{-2\alpha} du &\leq \left(\int_0^t |X_u^\varepsilon - x_u| (t-u)^{-\frac{\delta}{\gamma}} du \right)^\gamma \left(\int_0^t (t-u)^{-\frac{2\alpha+\delta}{1-\gamma}} du \right)^{1-\gamma} \\ &\lesssim \left(1 + \int_0^t |X_u^\varepsilon - x_u| (t-u)^{-\kappa_\gamma} du \right),\end{aligned}$$

and (4.16) holds true. If $0 \leq \gamma < \frac{1-2\alpha}{1-\alpha}$, (4.16) follows from the same argument as in the case $\frac{1-2\alpha}{1-\alpha} \leq \gamma < 1$. Define the funtion

$$h_t := |X_t^\varepsilon - x_t| + \int_0^t \frac{|X_t^\varepsilon - x_t - (X_s^\varepsilon - x_s)|}{(t-s)^{\alpha+1}} ds.$$

Then by estimates (4.12), (4.13), (4.14) and (4.16), we have

$$\begin{aligned}|h_t| &\leq |\Psi_t| + |\Phi_t| + \int_0^t \frac{|\Psi_t - \Psi_s| + |\Phi_t - \Phi_s|}{(t-s)^{1+\alpha}} ds \\ &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} + \int_0^t |X_s^\varepsilon - x_s| ds + \int_0^t |X_s^\varepsilon - x_s| \, (t-s)^{-\alpha} ds \\ &\quad + \varepsilon \|B^H\|_{\beta,[0.t]} \int_0^t \left(\frac{|X_s^\varepsilon - x_s|}{s^\alpha} + \int_0^s \frac{|X_s^\varepsilon - x_s - X_u^\varepsilon - x_u|}{(s-u)^{1+\alpha}} du \right) ds \\ &\quad + \varepsilon \|B^H\|_{\beta,[0.t]} \left\{ \int_0^t |X_s^\varepsilon - x_s| \, (t-s)^{-\kappa_\gamma} ds + \int_0^t (t-s)^{-\alpha} \left(\int_0^s \frac{|X_s^\varepsilon - x_s - X_u^\varepsilon + x_u|}{(s-u)^{\alpha+1}} du \right) ds \right\} \\ &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} + \left(\varepsilon \|B^H\|_{\beta,[0.t]} + 1 \right) \int_0^t h_s (t-s)^{-\kappa_\gamma} ds + \varepsilon \|B^H\|_{\beta,[0.t]} \int_0^t s^{-\alpha} h_s ds \\ &\lesssim \varepsilon \|B^H\|_{\beta,[0.t]} + \left(\varepsilon \|B^H\|_{\beta,[0.t]} + 1 \right) t^{\kappa_\gamma} \int_0^t h_s (t-s)^{-\kappa_\gamma} s^{-\kappa_\gamma} ds.\end{aligned}$$

By Lemma 2.3, there exists a constant $C > 0$ depending on α and T such that

$$|X_t^\varepsilon - x_t| \leq |h_t| \lesssim \varepsilon \|B^H\|_{\beta,[0.t]} \exp\left\{ C \left(1 + \varepsilon \|B^H\|_{\beta,[0.t]} \right)^{\frac{1}{1-\kappa_\gamma}} \right\}. \tag{4.17}$$

For every $p \geq 1$, it follows that from Lemma 2.2

$$\left(E \|B^H\|_{\beta,[s,t]}^p \right)^{1/p} \lesssim (t-s)^{H-\beta}.$$

Since $\frac{1}{1-\kappa_\gamma} < 2$, we can get

$$E \exp\left\{pC\left(1+\varepsilon\|B^H\|_{\beta,[0.t]}\right)^{\frac{1}{1-\kappa_\gamma}}\right\} < \infty,$$

by Fernique's theorem and we conclude (4.8). The inequality (4.9) and (4.10) follows immediately from (4.17). We will prove the estimate (4.11). Using the additivity of the integral, we obtain that

$$\begin{aligned}
|X_t^\varepsilon - X_s^\varepsilon| &\le \left|\int_s^t b(X_u^\varepsilon,\theta_0)du\right| + \varepsilon\left|\int_s^t \sigma(X_u^\varepsilon)dB_u^H\right| \\
&\lesssim \int_s^t (1+|X_u^\varepsilon|)\,du + \varepsilon\|B^H\|_{\beta,[s,t]}\int_s^t \left|D_{s+}^\alpha \sigma(X_\cdot^\varepsilon)(u)\right| du \\
&\lesssim \sup_{u\in[0,T]}(1+|X_u^\varepsilon|)(t-s) + \varepsilon\|B^H\|_{\beta,[s,t]}\left\{\int_s^t\left(\frac{|\sigma(X_u^\varepsilon)|}{(u-s)^\alpha} + \int_s^u \frac{|\sigma(X_u^\varepsilon)-\sigma(X_v^\varepsilon)|}{(u-v)^{1+\alpha}}dv\right)du\right\}.
\end{aligned}$$

The Assumption 4.2 and the estimate (4.17) lead

$$\begin{aligned}
|X_t^\varepsilon - X_s^\varepsilon| &\lesssim \sup_{u\in[0,T]}(1+|X_u^\varepsilon|)(t-s) + \varepsilon\|B^H\|_{\beta,[s,t]}(t-s)^{1-\alpha}\left(1+\sup_{u\in[0,T]}|X_u^\varepsilon|\right) \\
&\qquad + \varepsilon\|B^H\|_{\beta,[s,t]}\int_s^t\left(\int_0^u \frac{|X_u^\varepsilon - X_v^\varepsilon|}{(u-v)^{1+\alpha}}dv\right)du \\
&\lesssim \left(1+\varepsilon\|B^H\|_{\beta,[0.t]}\exp\left(C\left(1+\varepsilon\|B^H\|_{\beta,[0.t]}\right)^{\frac{1}{1-\kappa_\gamma}}\right)\right)(t-s) \\
&\qquad + \varepsilon\|B^H\|_{\beta,[s,t]}(t-s)^{1-\alpha}\left\{1+\varepsilon\|B^H\|_{\beta,[0.t]}\exp\left(C\left(1+\varepsilon\|B^H\|_{\beta,[0.t]}\right)^{\frac{1}{1-\kappa_\gamma}}\right)\right\}.
\end{aligned}$$

From Lemma 2.2 and Fernique's theorem, we conclude (4.11). Thus the proof is completed. □

Remark 4.3. The function $\mathbb{L}_{H,\varepsilon}$ is differentiable in θ under Assumption 4.4, and we have

$$\begin{aligned}
\nabla_\theta \mathbb{L}_{H,\varepsilon}(\theta) &= \int_0^T \nabla_\theta Q_{H,\theta}^\varepsilon(t)dZ_t - \int_0^T Q_{H,\theta}^\varepsilon(t)\nabla_\theta Q_{H,\theta}^\varepsilon(t)dt, \\
\nabla_\theta^2 \mathbb{L}_{H,\varepsilon}(\theta_0) &= \int_0^T \nabla_\theta^2 Q_{H,\theta}^\varepsilon(t)dZ_t - \int_0^T \left(\nabla_\theta Q_{H,\theta}^\varepsilon(t)\right)^{\otimes 2}dt - \int_0^T Q_{H,\theta}^\varepsilon(t)\nabla_\theta^2 Q_{H,\theta}^\varepsilon(t)dt.
\end{aligned}$$

We prepare some lemmas which give sufficient conditions to obtain the polynomial-type large deviation inequality.

Lemma 4.2. *For every* $p \geq 2$,

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

Proof. By Remark 4.3, it folows that

$$\varepsilon^2\nabla_\theta^2\mathbb{L}_{H,\varepsilon}(\theta_0) - (-\Gamma_H(\theta_0)) = \varepsilon^2\int_0^T \nabla_\theta^2 Q^\varepsilon_{H,\theta_0}(t)dW_t - \left(\varepsilon^2\int_0^T \left(\nabla_\theta Q^\varepsilon_{H,\theta_0}(t)\right)^{\otimes 2}dt - \Gamma_H(\theta_0)\right).$$

Note that the function $Q^\varepsilon_{H,\theta_0}$ has the form

$$\begin{aligned}
Q^\varepsilon_{H,\theta_0}(t) &= (\varepsilon d_H\Gamma(3/2-H))^{-1}\Bigg(t^{1/2-H}\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)} \\
&\qquad + (H-\frac{1}{2})t^{H-1/2}\int_0^t \frac{\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)}t^{1/2-H} - \frac{b(X^\varepsilon_s,\theta_0)}{\sigma(X^\varepsilon_s)}s^{1/2-H}}{(t-s)^{H+1/2}}ds\Bigg) \\
&= c_1\varepsilon^{-1}t^{1/2-H}\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)} + c_2\varepsilon^{-1}t^{H-1/2}\int_0^t \frac{\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)} - \frac{b(X^\varepsilon_s,\theta_0)}{\sigma(X^\varepsilon_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds.
\end{aligned}$$

By Burkholder's and Minkowski's inequalities, we have

$$\begin{aligned}
E\left(\varepsilon^2\left|\int_0^T \partial_{\theta_i}\partial_{\theta_j}Q^\varepsilon_{H,\theta_0}(t)dW_t\right|\right)^p &\lesssim \left(\varepsilon^4\int_0^T \left\|\partial_{\theta_i}\partial_{\theta_j}Q^\varepsilon_{H,\theta_0}(t)\right\|^2_{L^p(\Omega)}dt\right)^{p/2} \\
&\lesssim \varepsilon^p\Bigg\{\int_0^T t^{1-2H}\left\|\frac{\partial_{\theta_i}\partial_{\theta_j}b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)}\right\|^2_{L^p(\Omega)}dt \\
&\qquad + \int_0^T t^{2H-1}\left\|\int_0^t \frac{\frac{\partial_{\theta_i}\partial_{\theta_j}b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)} - \frac{\partial_{\theta_i}\partial_{\theta_j}b(X^\varepsilon_s,\theta_0)}{\sigma(X^\varepsilon_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right\|^2_{L^p(\Omega)}dt\Bigg\}^{p/2}
\end{aligned}$$

From Assumptions 4.4 and the mean value theorem, we obtain that

$$\begin{aligned}
E\left(\varepsilon^2\left|\int_0^T \partial_{\theta_i}\partial_{\theta_j}Q^\varepsilon_{H,\theta_0}(t)dW_t\right|\right)^p &\lesssim \varepsilon^p\Bigg\{\sup_{0\leq t\leq T}\left\|1+|X^\varepsilon_t|^N\right\|^2_{L^p(\Omega)}\int_0^T t^{1-2H}dt \\
&\qquad + \int_0^T t^{2H-1}\left\|\int_0^t \frac{\left(1+|X^\varepsilon_t|^N+|X^\varepsilon_s|^N\right)|X^\varepsilon_t - X^\varepsilon_s|}{(t-s)^{H+1/2}}s^{1/2-H}ds\right\|^2_{L^p(\Omega)}dt\Bigg\}^{p/2} \\
&\lesssim \varepsilon^p\left\{1+\int_0^T t^{2H-1}\left(\int_0^t \frac{\|X^\varepsilon_t - X^\varepsilon_s\|_{L^{2p}(\Omega)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)^2 dt\right\}^{p/2} \lesssim \varepsilon^p,
\end{aligned}$$

for every $i,j=1,\cdots,d$ and some constants $N>0$. We estimate the term $\varepsilon^2\int_0^T\left(\nabla_\theta Q^\varepsilon_{H,\theta_0}(t)\right)^{\otimes 2}dt-\Gamma_H(\theta_0)$. For every $i,j=1,\cdots,d$

$$\begin{aligned}
&\varepsilon^2\int_0^T \partial_{\theta_i}Q^\varepsilon_{H,\theta_0}(t)\partial_{\theta_j}Q^\varepsilon_{H,\theta_0}(t)dt-\Gamma_H^{i,j}(\theta_0)\\
&=c_1^2\left(\int_0^T t^{1-2H}\left\{\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)^2}-\frac{\partial_{\theta_i}b(x_t,\theta_0)\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)^2}\right\}dt\right)\\
&+c_1c_2\Bigg(\int_0^T\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}\int_0^t\frac{\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}}s^{1/2-H}dsdt\\
&\qquad-\int_0^T\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}\int_0^t\frac{\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}-\frac{\partial_{\theta_j}b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}}s^{1/2-H}dsdt\Bigg)\\
&+c_1c_2\Bigg(\int_0^T\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}\int_0^t\frac{\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_i}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}}s^{1/2-H}dsdt\\
&\qquad-\int_0^T\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}\int_0^t\frac{\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}-\frac{\partial_{\theta_i}b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}}s^{1/2-H}dsdt\Bigg)\\
&+c_2^2\Bigg(\int_0^T t^{2H-1}\left(\int_0^t\frac{\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_i}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)\left(\int_0^t\frac{\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)dt\\
&\quad-\int_0^T t^{2H-1}\left(\int_0^t\frac{\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}-\frac{\partial_{\theta_i}b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)\left(\int_0^t\frac{\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}-\frac{\partial_{\theta_j}b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\right)dt\Bigg).
\end{aligned}\tag{4.18}$$

By using Lemma 4.1, the mean value theorem and Assumptions 4.4, we have

$$\begin{aligned}
&E\left|\int_0^T t^{1-2H}\left\{\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)^2}-\frac{\partial_{\theta_i}b(x_t,\theta_0)\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)^2}\right\}dt\right|^p\\
&=E\Bigg|\int_0^T t^{1-2H}\Bigg\{\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}\left(\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}\right)\\
&\qquad\qquad+\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}\left(\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}\right)\Bigg\}dt\Bigg|^p\\
&\lesssim\left(\int_0^T t^{1-2H}\left\|\left(1+|X_t^\varepsilon|^N+|x_t|^N\right)|X_t^\varepsilon-x_t|\right\|_{L^p(\Omega)}dt\right)^p\lesssim\varepsilon^p,
\end{aligned}$$

for some constants $N > 0$. We shall estimate the second term of (4.18). By Lemma 2.5

$$
\begin{aligned}
&\left\| \frac{\partial_{\theta_j} b(X^\varepsilon_t, \theta_0)}{\sigma(X^\varepsilon_t)} - \frac{\partial_{\theta_j} b(X^\varepsilon_s, \theta_0)}{\sigma(X^\varepsilon_s)} - \left(\frac{\partial_{\theta_j} b(x_t, \theta_0)}{\sigma(x_t)} - \frac{\partial_{\theta_j} b(x_s, \theta_0)}{\sigma(x_s)} \right) \right\|_{L^p(\Omega)} \\
&= \left\| \int_s^t \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) dX^\varepsilon_u - \int_s^t \partial_x \left(\frac{\partial_{\theta_j} b(x_u, \theta_0)}{\sigma(x_u)} \right) dx_u \right\|_{L^p(\Omega)} \\
&\leq \left\| \int_s^t \left\{ \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) b(X^\varepsilon_u, \theta_0) - \partial_x \left(\frac{\partial_{\theta_j} b(x_u, \theta_0)}{\sigma(x_u)} \right) b(x_u, \theta_0) \right\} du \right\|_{L^p(\Omega)} \\
&\qquad + \varepsilon \left\| \int_s^t \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) \sigma(X^\varepsilon_u) dB^H_u \right\|_{L^p(\Omega)}.
\end{aligned}
$$

By the mean value theorem and Assumptions 4.2 and 4.4, we have

$$
\begin{aligned}
&\left\| \int_s^t \left\{ \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) b(X^\varepsilon_u, \theta_0) - \partial_x \left(\frac{\partial_{\theta_j} b(x_u, \theta_0)}{\sigma(x_u)} \right) b(x_u, \theta_0) \right\} du \right\|_{L^p(\Omega)} \\
&\lesssim \left\| \int_s^t \left(1 + |x_u|^N + |X^\varepsilon_u|^N \right) |X^\varepsilon_u - x_u| du \right\|_{L^p(\Omega)} \lesssim \varepsilon(t-s).
\end{aligned}
$$

From the mean value theorem and Assumptions 4.3 and 4.4, we can estimate the stochastic integral for $\partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) \sigma(X^\varepsilon_u)$ as with respect to B^H for $\alpha \in (1-H, 1/2)$ and $\beta \in (1-\alpha, H)$

$$
\begin{aligned}
&\left| \int_s^t \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) \sigma(X^\varepsilon_u) dB^H_u \right| \lesssim \|B^H\|_{\beta,[s,t]} \int_s^t \left| D^\alpha_{s+} \left\{ \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_\cdot, \theta_0)}{\sigma(X^\varepsilon_\cdot)} \right) \sigma(X^\varepsilon_\cdot) \right\} (u) \right| du \\
&\lesssim \|B^H\|_{\beta,[s,t]} \int_s^t \left\{ \frac{1 + |X^\varepsilon_u|^N}{(u-s)^\alpha} + \int_s^u \frac{\left| \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) \sigma(X^\varepsilon_u) - \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_v, \theta_0)}{\sigma(X^\varepsilon_v)} \right) \sigma(X^\varepsilon_v) \right|}{(u-v)^{\alpha+1}} dv \right\} du \\
&\lesssim \|B^H\|_{\beta,[s,t]} \left\{ \left(1 + \sup_{s \leq u \leq t} |X^\varepsilon_u|^N \right) \left((t-s)^{1-\alpha} + \int_s^t \left(\int_s^u \frac{|X^\varepsilon_u - X^\varepsilon_v|}{(u-v)^{\alpha+1}} dv \right) du \right) \right\}.
\end{aligned}
$$

By the Garcia–Rodenmich–Rumsey inequality and Lemma 4.1, we get

$$
\varepsilon \left\| \int_s^t \partial_x \left(\frac{\partial_{\theta_j} b(X^\varepsilon_u, \theta_0)}{\sigma(X^\varepsilon_u)} \right) \sigma(X^\varepsilon_u) dB^H_u \right\|_{L^p(\Omega)} \lesssim \varepsilon(t-s)^{1+H-(\alpha+\beta)} \lesssim \varepsilon(t-s)^{1/2}.
$$

Thus by Minkowski, Hölder's inequality, we obtain that

$$
\begin{aligned}
E\Bigg|\int_0^T\Bigg(&\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}\int_0^t\frac{\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\\
&-\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}\int_0^t\frac{\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}-\frac{\partial_{\theta_j}b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\Bigg)dt\Bigg|^p\\
=E\Bigg|\int_0^T\Bigg(&\left(\frac{\partial_{\theta_i}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}\right)\int_0^t\frac{\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\\
&+\frac{\partial_{\theta_i}b(x_t,\theta_0)}{\sigma(x_t)}\int_0^t\frac{\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}-\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}+\frac{\partial_{\theta_j}b(x_s,\theta_0)}{\sigma(x_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\Bigg)dt\Bigg|^p\\
\lesssim\Bigg(&\int_0^T\left\|(1+|X_t^\varepsilon|^N+|x_t|^N)|X_t^\varepsilon-x_t|\right\|_{L^{2p}(\Omega)}\int_0^t\frac{\left\|\left(1+|X_t^\varepsilon|^N+|X_s^\varepsilon|^N\right)|X_t^\varepsilon-X_s^\varepsilon|\right\|_{L^{2p}(\Omega)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\Bigg)^p\\
+\Bigg(&\int_0^T(1+|x_t|^N)\int_0^t\frac{\left\|\frac{\partial_{\theta_j}b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)}-\frac{\partial_{\theta_j}b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}-\frac{\partial_{\theta_j}b(x_t,\theta_0)}{\sigma(x_t)}+\frac{\partial_{\theta_j}b(x_s,\theta_0)}{\sigma(x_s)}\right\|_{L^p(\Omega)}}{(t-s)^{H+1/2}}s^{1/2-H}dsdt\Bigg)^p\lesssim\varepsilon^p.
\end{aligned}
$$

Since the third and fourth terms of (4.18) can be evaluated in the similar way as in the second term of (4.18), the proof is completed. □

The rest of the proof of Theorem 4.1 is similar to the proof of Theorem 3.1 and is therefore omitted.

4.6 A sufficient condition for the absolute continuity

In Section 4.2, we have derived the likelihood ratio (4.4) via the Girsanov theorem under the assumption (4.3), under which our estimator $\hat{\theta}_\varepsilon$ is the true MLE. However, to ensure the condition (4.3), we need a stronger assumption for coefficients b and σ.

Assumption 4.8. *The function σ in (4.1) is of $C_b^1(\mathbb{R})$-class and there exist some constants $\lambda\in(1-\frac{1}{2H},\frac{1}{2})$, $c>0$ such that for every $x,y\in\mathbb{R}$ and $\theta\in\Theta$,*

$$
\left|\frac{b(x,\theta)}{\sigma(x)}-\frac{b(y,\theta)}{\sigma(y)}\right|\le c|x-y|^\lambda.
$$

Theorem 4.2. *Suppose that Assumptions 4.2,4.4 and 4.8 are fulfilled. Then a stochastic process $\{\Lambda_t^{\theta_0}\}_{t\in[0,T]}$ is a martingale.*

To prove Theorem 4.2, we prepare the following result which is found in Lemma 3.2 in [11]. In [11], they impose that the function $\partial_x b$ is bounded. Actually, it is enough to impose the Lipschitz continuity of the function $b(\cdot,\theta)$.

Lemma 4.3. *Under the condition of Theorem 4.2, for every $\alpha \in (1-H, 1/2)$ and $\beta \in (1-\alpha, H)$, there exists a constant $C>0$ such that*

$$\|X^\varepsilon\|_{1-\alpha} \le C\left(1+\|B^H\|_\beta\right)^{1/\beta}.$$

Proof. By Lemma 2.6 and Assumption 4.8, we have

$$\begin{aligned}|X_t^\varepsilon - X_s^\varepsilon - \sigma(X_s^\varepsilon)\left(B_t^H - B_s^H\right)| &\le \int_s^t |b(X_u^\varepsilon,\theta_0)|\,du + \left|\int_s^t \{\sigma(X_u^\varepsilon)-\sigma(X_s^\varepsilon)\} dB_u^H\right| \\ &\lesssim \int_s^t |b(X_u^\varepsilon,\theta_0)|\,du + \|X^\varepsilon\|_{1-\alpha,[s,t]}\|B^H\|_{\beta,[s,t]}(t-s)^{1-\alpha+\beta}.\end{aligned}$$

Note that

$$\sup_{s\le\tau\le\eta\le t}\left|\int_\tau^\eta b(X_u^\varepsilon,\theta_0)du\right| \lesssim (t-s) + |X_s^\varepsilon|(t-s) + \|X^\varepsilon\|_{1-\alpha,[s,t]}(t-s)^{2-\alpha}.$$

Thus there exists $C>0$ depending on T such that

$$\begin{aligned}\|X^\varepsilon\|_{1-\alpha,[s,t]} \le C\Bigg(&1+|X_s^\varepsilon|(t-s)^\alpha + \|B^H\|_{\beta,[0,T]}(t-s)^{\beta-(1-\alpha)} \\ &+ \|X^\varepsilon\|_{1-\alpha,[s,t]}\left(\|B^H\|_{\beta,[0,T]}+1\right)(t-s)^\beta\Bigg).\end{aligned} \tag{4.19}$$

Set $h := \left(2C(\|B^H\|_{\beta,[0,T]}+1)^{-1/\beta}\right)$. Then for $|t-s|<h$,

$$\|X^\varepsilon\|_{1-\alpha,[s,t]} \le 2C\left(1+|X_s^\varepsilon|h^\alpha + \|B^H\|_{\beta,[0,T]}h^{\beta-(1-\alpha)}\right).$$

We aim to estimate $\|X^\varepsilon\|_{\infty,[0,T]}$. By the triangle inequality, we have

$$\begin{aligned}|X^\varepsilon|_{\infty,[s,t]} &\le |X_s^\varepsilon| + \|X^\varepsilon\|_{1-\alpha,[s,t]}(t-s)^{1-\alpha} \\ &\lesssim |X_s^\varepsilon|(1+h) + h^\beta(1+\|B^H\|_{\beta,[0,T]}),\end{aligned}$$

for $|t-s|<h$. Iterating the above estimate for $N=T/h$, we can get

$$\begin{aligned}\|X^\varepsilon\|_{\infty,[0,T]} &\lesssim |x|(1+h)^N + h^\beta(1+\|B^H\|_{\beta,[0,T]})\sum_{k=0}^{N-1}(1+h)^k \\ &\lesssim |x|(1+h)^N + h^\beta(1+\|B^H\|_{\beta,[0,T]})(1+h)^N N \\ &\lesssim (1+h)^{T/h}\left(|x| + h^\beta(1+\|B^H\|_{\beta,[0,T]})T/h\right) \\ &= e^T\left(|x| + h^\beta(1+\|B^H\|_{\beta,[0,T]})T/h\right),\end{aligned} \tag{4.20}$$

where we used the fact $(1+\frac{1}{h})^h \le e$ for the last inequality. Thus we have

$$\begin{aligned}\|X^\varepsilon\|_{1-\alpha,h} &\lesssim 1 + \left(1+\|B^H\|_{\beta,[0,T]}\right) h^{\beta-(1-\alpha)} \\ &\lesssim 1 + \left(1+\|B^H\|_{\beta,[0,T]}\right)^{(1-\alpha)/\beta}.\end{aligned}$$

By using Lemma 2.4,

$$\|X^\varepsilon\|_{1-\alpha} \lesssim \left(1+\left(1+\|B^H\|_{\beta,[0,T]}\right)^{(1-\alpha)/\beta}\right)\left(1\vee h^{-\alpha}\right) \lesssim \left(1+\|B^H\|_{\beta,[0,T]}\right)^{1/\beta},$$

and we complete the proof of Lemma 4.3. □

Proof of Theorem 4.2. It is enough to show that for every $r>0$

$$E\exp\left(r\int_0^T Q^\varepsilon_{H,\theta_0}(t)^2 dt\right) < \infty. \tag{4.21}$$

Recall that the function $Q^\varepsilon_{H,\theta_0}$ has the form

$$\begin{aligned}Q^\varepsilon_{H,\theta_0}(t) &= (\varepsilon d_H \Gamma(3/2-H))^{-1}\left(t^{1/2-H}\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)}\right.\\ &\qquad\left. + (H-\frac{1}{2})t^{H-1/2}\int_0^t \frac{\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)}t^{1/2-H} - \frac{b(X^\varepsilon_s,\theta_0)}{\sigma(X^\varepsilon_s)}s^{1/2-H}}{(t-s)^{H+1/2}}ds\right)\\ &= c_1\varepsilon^{-1}t^{1/2-H}\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)} + c_2\varepsilon^{-1}t^{H-1/2}\int_0^t \frac{\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)} - \frac{b(X^\varepsilon_s,\theta_0)}{\sigma(X^\varepsilon_s)}}{(t-s)^{H+1/2}}s^{1/2-H}ds\\ &=: F(t)+G(t).\end{aligned}$$

By Lemma 4.3, there exists a universal constant $C>0$ such that

$$\begin{aligned}\exp\left(r\int_0^T F(t)^2dt\right) &= \exp\left(rc_1^2\varepsilon^{-2}\int_0^T\left(\frac{b(X^\varepsilon_t,\theta_0)}{\sigma(X^\varepsilon_t)}\right)^2 t^{1-2H}dt\right)\\ &\le \exp\left(Cr\varepsilon^{-2}\int_0^T\left(1+|X_t|^{2\lambda}\right)t^{1-2H}dt\right)\\ &\le \exp\left(Cr(1+\|X^\varepsilon\|^{2\lambda}_{1-\alpha})\int_0^T t^{2\lambda(1-\alpha)+1-2H}dt\right)\\ &\le \exp\left(T^{2\lambda(1-\alpha)+2-2H}Cr\left(1+\|B^H\|_\beta^{2\lambda/\beta}\right)\right),\end{aligned}$$

and by Fernique's theorem, we have

$$E \exp\left(r \int_0^T F(t)^2 dt\right) < \infty.$$

In the same way, we can estimate the term $G(t)$ as

$$\begin{aligned}
\exp\left(r \int_0^T G(t)^2 dt\right) &\leq \exp\left(Cr \int_0^T t^{2H-1} \left(\int_0^t \frac{\frac{b(X_t^\varepsilon,\theta_0)}{\sigma(X_t^\varepsilon)} - \frac{b(X_s^\varepsilon,\theta_0)}{\sigma(X_s^\varepsilon)}}{(t-s)^{H+1/2}} s^{1/2-H} ds\right)^2 dt\right) \\
&\leq \exp\left(Cr \|X^\varepsilon\|_{1-\alpha}^{2\lambda} \int_0^T t^{2H-1} \left(\int_0^t (t-s)^{-1/2-H+\lambda(1-\alpha)} s^{1/2-H} ds\right)^2 dt\right) \\
&= \exp\left(CrT^{2\lambda(1-\alpha)+1} \left(1 + \|B^H\|_\beta^{2\lambda/\beta}\right)\right),
\end{aligned}$$

and we conclude (4.21) by Fernique's theorem. □

5 Discretely observed case

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

$$X_t = X_0 + \int_0^t b(X_s, \theta_0)ds + \sigma B_t^H, \quad t \in [0, T], \tag{5.1}$$

where $X_0 \in \mathbb{R}$ is the initial value, $\{B_t^H\}_{t\in[0,T]}$ is a fBm with Hurst index $H \in (1/2, 1)$ and $\theta_0 \in \Theta$ is the parameter which is contained in a bounded and open convex subset $\Theta \subset \mathbb{R}$. When $H = 1/2$ and processes are observed at discrete points, the least squares estimator (LSE) is asymptotically equivalent to the MLE. For the LSE, the consistency and asymptotic distributions were proved by Kasonaga [16] and Prakasa Rao [35].

Neuenkirch and Tindel [30] studied estimators based on discrete samples for (5.1). They proposed the least squares type estimators and showed its consistency when $1/2 < H < 1$. The purpose of this Chapter is to prove the asymptotic normality of the estimator proposed in [30]. The main tools in the proof of the asymptotic normality are the limit theorem for quadratic variation of fBm and the ergodic theorem to sums of the increments of fBm weighted by a function.

5.1 Main results

We assume that $\{X_t\}_{t\geq 0}$ is observed at points $\{t_k : \ 0 \leq k \leq n\}$ and take equally spaced observations with $t_{k+1} - t_k := h_n$. Define the least squares type procedure,

$$Q_n(\theta) := \frac{1}{nh_n^2} \sum_{k=1}^{n} \left(\left(X_{t_k} - X_{t_{k-1}} - h_n b(X_{t_{k-1}}, \theta) \right)^2 - \sigma^2 h_n^{2H} \right),$$

and the LSE for the true θ_0 is defined as

$$\theta_n := \arg\min_{\theta\in\bar{\Theta}} |Q_n(\theta)|. \tag{5.2}$$

Let us state a one sided dissipative Lipschitz condition and the polynomial growth assumptions for a drift coefficients b, ensuring ergodic properties for process X.

Assumption 5.1. *The function b in (3.1) is of $C^{1,2}(\mathbb{R}\times\Theta)$-class such that, for every $x, y \in \mathbb{R}$ and $\theta \in \Theta$,*

$$\left(b(x,\theta) - b(y,\theta)\right)(x - y) \leq -c|x - y|^2.$$

and the following growth conditions hold true

$$|b(x,\theta)| \le c(1+|x|^N),\ |\partial_x b(x,\theta)| \le c(1+|x|^N),\ |\partial_\theta b(x,\theta)| \le c(1+|x|^N)$$
$$|\partial_\theta \partial_x b(x,\theta)| \le c(1+|x|^N),\ |\partial_\theta^2 b(x,\theta)| \le c(1+|x|^N),$$

for some constants $c > 0,\ N \in \mathbb{N}$.

Assumption 5.2. *There exists a function* $U \in C^{2,2}(\mathbb{R} \times \Theta)$ *such that*

$$\partial_x U(x,\theta) = b(x,\theta),$$

for every $x \in \mathbb{R},\ \theta \in \Theta$.

We impose the condition on the size of the sampling step, which is required to control the contribution of fractional Brownian motion.

Assumption 5.3. $h_n = \kappa n^{-\alpha}$ *with* $0 < \alpha < \min\{\frac{1}{4(1-H)}, 1\}$ *and* $\kappa > 0$.

We set the identibility assumption for consistency of our estimator.

Assumption 5.4. *For every* $\theta_0 \in \Theta$,

$$E\left|b(\bar{X},\theta_0)\right|^2 = E\left|b(\bar{X},\theta)\right|^2,$$

implies that $\theta = \theta_0$ *where* $\bar{X}$ *is the random variable appearing in Proposition 5.1.*

Assumption 5.5.

$$E\left[\left(\partial_\theta b(\bar{X},\theta_0)\right) b(\bar{X},\theta_0)\right] \neq 0.$$

The consistency of LSE (5.2) was given by Neuenkirch and Tindel [30].

Theorem 5.1 (Neuenkirch and Tindel [30]). *Under Assumptions 5.1-5.4, the LSE* θ_n *is strongly consistent with* θ_0:

$$\theta_n \to \theta_0 \quad a.s., \quad n \to \infty.$$

To state the main results, we further make some notations: let

$$\tau_n^H = \begin{cases} \sqrt{n} h_n^{2-2H}, & H \in (1/2, 3/4) \\ \sqrt{\frac{nh_n}{\log(n)}}, & H = 3/4 \\ (nh_n)^{2-2H}, & H \in (3/4, 1), \end{cases}$$

and we set

$$c_H := \frac{1}{2}\sum_{v \in \mathbb{Z}} \left(|v+1|^{2H} + |v-1|^{2H} - 2|v|^{2H}\right)^2 < \infty.$$

Theorem 5.2. *Suppose the same assumptions as in Theorem 5.1, and that* $\frac{1}{2} < \alpha < \min\{\frac{1}{4(1-H)}, 1\}$*, Assumption 5.5 holds true. Then for every* $H \in (1/2, 3/4]$,

$$\tau_n^H(\theta_n - \theta_0) \xrightarrow{d} N\left(0, \frac{\sigma^2 c_H^2}{4}\left(E\left[\left(\partial_\theta b(\bar{X}, \theta_0)\right) b(\bar{X}, \theta_0)\right]\right)^{-2}\right),$$

as $n \to \infty$*. Moreover, if* $H \in (3/4, 1)$,

$$\tau_n^H(\theta_n - \theta_0) \xrightarrow{d} \frac{\sigma\left(E\left[\left(\partial_\theta b(\bar{X}, \theta_0)\right) b(\bar{X}, \theta_0)\right]\right)^{-1}}{2} Z,$$

as $n \to \infty$ *where* Z *is a Rosenblatt random variable that will be defined in Proposition 5.3.*

5.2 Examples

Example 5.1. Let $\theta_0 > 0$ and $x \in \mathbb{R}$. The simplest example of an equation which satisfies the above assumptions is fractional Orstein–Uhlenbeck process:

$$X_t = x - \theta_0 \int_0^t X_s ds + \sigma B_t^H.$$

One of the example of nonlinear SDE is

$$X_t = x - \theta_0 \int_0^t \left(X_s^3 + X_s\right) ds + \sigma B_t^H.$$

5.3 Ergodicity

We shall describe the results of the ergodic theorem in equation (3.1). We will work on the canonical probability space $(\Omega, \mathcal{F}, P)$, where $\Omega = C_0(\mathbb{R})$ is equipped with the topology of the compact convergence, $\mathcal{F}$ is the corresponding Borel σ-algebra, and P is the distribution of the fBm. We define the shift operator $S_t : \Omega \to \Omega$ for each $t \in \mathbb{R}$ and $\omega \in \Omega$ as

$$S_t\omega(\cdot) = \omega(t + \cdot) - \omega(t).$$

The shifted process $(B_s(S_t(\cdot)))_{s\in\mathbb{R}}$ is a 1–dimensional fBm, and, for any integrable random variable $F : \Omega \to \mathbb{R}$ and any $\omega \in \Omega$, we have

$$\lim_{T\to\infty} \int_0^T F(S_t(\omega))dt = E[F].$$

We state the existence and uniqueness of the ergodic limit for X investigated in [7].

Proposition 5.1. *Under Assumption 5.1, for any $\theta \in \Theta$, the equation* (5.1) *has a unique solution $X \in C^{\lambda}(\mathbb{R}_+; \mathbb{R})$ for all $\lambda < H$. In addition, There exists a random variable $\bar{X} : \Omega \to \mathbb{R}$ such that*

$$\lim_{t\to\infty} \left|X_t(\omega) - \bar{X}(S_t\omega)\right| = 0,$$

for all $\omega \in \Omega$ where S_t is the canonical shift operator.

An ergodic theorem for discrete sampling is found in Lemma 3.3 in Neuenkirch and Tindel [30].

Lemma 5.1. *Let $f \in C^{1,1}(\mathbb{R} \times \Theta)$ be such that*

$$|f(x,\theta)| \le C(1+|x|^N),\ |\partial_x f(x,\theta)| \le C(1+|x|^N),\ |\partial_\theta f(x,\theta)| \le C(1+|x|^N),$$

for some $c > 0, N \in \mathbb{N}$. Then

$$\sup_{\theta\in\Theta} \left| \frac{1}{n} \sum_{k=1}^{n} f(X_{t_{k-1}},\theta) - Ef(\bar{X},\theta) \right| \to 0, \quad a.s.$$

In addition, assume that there exists a function $U \in C^{2,1}(\mathbb{R} \times \Theta)$ such that

$$\partial_x U(x,\theta) = f(x,\theta),\ x \in \mathbb{R},\ \theta \in \Theta.$$

Then

$$\sup_{\theta\in\Theta} \left| \frac{1}{nh_n} \sum_{k=1}^{n} f\left(X_{t_{k-1}},\theta\right)(B_{t_k} - B_{t_{k-1}}) + E\left[b(\bar{X},\theta_0) f(\bar{X},\theta)\right] \right| \to 0 \quad a.s.$$

as $n \to \infty$.

We prepare some estimate results for the pth moment of the solution to (3.1). To support these results, we refer to [7] and [30].

Proposition 5.2. *Under Assumption 5.1, for any $\theta \in \Theta$ and $p \ge 1$, there exist constants $c_p, k_p > 0$ such that*

$$E|X_t|^p \le c_p, \qquad E|X_t - X_s|^p \le k_p|t-s|^{pH}$$

for all $s,t \ge 0$.

We need the following convergence results for one–dimensional fBm (cf. Theorem 7.4.1 in [32]).

Proposition 5.3. *For every* $0 < H < 3/4$*,*

$$\frac{1}{\sqrt{n}c_H}\sum_{k=1}^{n}[(B_k^H - B_{k-1}^H)^2 - 1] \xrightarrow{d} N(0,1),$$

while for $H = 3/4$ *it holds*

$$\frac{1}{\sqrt{n\log(n)}c_{3/4}}\sum_{k=1}^{n}[(B_k^H - B_{k-1}^H)^2 - 1] \xrightarrow{d} N(0,1).$$

Finally, for $3/4 < H < 1$

$$\frac{1}{n^{2H-1}}\sum_{k=1}^{n}[(B_k - B_{k-1})^2 - 1]$$

converges in $L^2(\Omega)$ *to some random variable* Z*, which belongs to the Wiener chaos of* B^H *with order 2. The random variable* Z *is called a Rosenblatt random variable.*

5.4 Proofs

Note that

$$\partial_\theta Q_n(\theta) = \frac{-2}{nh_n}\sum_{k=1}^{n}\left(X_{t_k} - X_{t_{k-1}} - h_n b(X_{t_{k-1}},\theta)\right)\partial_\theta b(X_{t_{k-1}},\theta).$$

Since θ_n minimizes $Q_n(\theta)^2$, we obtain that

$$Q_n(\theta_n)\partial_\theta Q_n(\theta_n) = 0.$$

To solve the above equation, we prepare the following lemma.

Lemma 5.2. *Define* $\zeta_n = \tau_n^H(\theta_0 - \theta_n)$*. Then, under assumptions of Theorem 5.2,*

$$\begin{cases} \tau_n^H Q_n(\theta_n) = \tau_n^H \dfrac{1}{nh_n^2}\displaystyle\sum_{k=1}^{n}\left[\sigma^2(B_{t_k}^H - B_{t_{k-1}}^H)^2 - \sigma^2 h_n^{2H}\right] + \dfrac{2\sigma}{nh_n}\displaystyle\sum_{k=1}^{n}(B_{t_k}^H - B_{t_{k-1}}^H)\partial_\theta b(X_{t_{k-1}},\theta_0)\zeta_n \\ \qquad\qquad + \zeta_n o_P(1) + o_P(1) \\ \partial_\theta Q_n(\theta_n) = -\dfrac{2\sigma}{nh_n}\displaystyle\sum_{k=1}^{n}(B_{t_k} - B_{t_{k-1}})\partial_\theta b(X_{t_{k-1}},\theta_n) + o_P(1). \end{cases}$$

Proof. By using (3.1), we have

$$
\begin{aligned}
Q_n(\theta_n) &= \frac{1}{nh_n^2}\sum_{k=1}^{n}\Big(\big(X_{t_k}-X_{t_{k-1}}-h_n b(X_{t_{k-1}},\theta_n)\big)^2-\sigma^2 h_n^{2H}\Big)\\
&= \frac{1}{nh_n^2}\sum_{k=1}^{n}\left\{\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds+\sigma(B^H_{t_k}-B^H_{t_{k-1}})\right)^2-\sigma^2 h_n^{2H}\right\}\\
&= \frac{1}{nh_n^2}\sum_{k=1}^{n}\left\{\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds\right)^2+\sigma^2(B^H_{t_k}-B^H_{t_{k-1}})^2\right.\\
&\qquad\left.+2\sigma(B^H_{t_k}-B^H_{t_{k-1}})\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds-\sigma^2 h_n^{2H}\right\}.
\end{aligned}
$$

For the first term, we can calculate that

$$
\begin{aligned}
&\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds\right)^2\\
&=\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds+\int_{t_{k-1}}^{t_k}\big(b(X_{t_{k-1}},\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds\right)^2\\
&=\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right)^2+\left(\int_{t_{k-1}}^{t_k}\big(b(X_{t_{k-1}},\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds\right)^2\\
&\qquad+2\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right)\left(\int_{t_{k-1}}^{t_k}\big(b(X_{t_{k-1}},\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds\right)
\end{aligned}
$$

Through Minkowski's inequality, mean value theorem, Assumption 5.1, Hölder's inequality and Proposition 5.2 we can estimate that

$$\begin{aligned}\left\|\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right\|_{L^2(\Omega)}^2 &\le \left(\int_{t_{k-1}}^{t_k}\big\|b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big\|_{L^2(\Omega)}\,ds\right)^2\\ &\lesssim \left(\int_{t_{k-1}}^{t_k}\big\|\big(1+|X_s|^N+|X_{t_{k-1}}|^N\big)\big(X_s-X_{t_{k-1}}\big)\big\|_{L^2(\Omega)}\,ds\right)^2\\ &\le \left(\int_{t_{k-1}}^{t_k}\big\|1+|X_s|^N+|X_{t_{k-1}}|^N\big\|_{L^2(\Omega)}\big\|X_s-X_{t_{k-1}}\big\|_{L^2(\Omega)}\,ds\right)^2\\ &\lesssim \left(\int_{t_{k-1}}^{t_k}(s-t_{k-1})^H\,ds\right)^2\lesssim h_n^{2H+2}.\end{aligned} \tag{5.3}$$

Therefore, under the assumption $nh_n^2\to 0$ as $n\to\infty$, we obtain

$$\tau_n^H\frac{1}{nh_n^2}\sum_{k=1}^n\left\|\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right)^2\right\|_{L^1(\Omega)}\lesssim \tau_n^H h_n^{2H}\to 0,$$

as $n\to\infty$. Applying Taylor's formula, we have

$$\begin{aligned}&\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right)\left(\int_{t_{k-1}}^{t_k}\big(b(X_{t_{k-1}},\theta_0)-b(X_{t_{k-1}},\theta_n)\big)\,ds\right)\\ &=h_n\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right)\partial_\theta b(X_{t_{k-1}},\tilde\theta_n)(\theta_0-\theta_n),\end{aligned}$$

where $\tilde\theta_n:=\theta_0+\beta_n(\theta_n-\theta_0),\ 0<\beta_n<1$. Similar calculation of (5.3), we have

$$\begin{aligned}&\frac{1}{nh_n^2}\sum_{k=1}^n E\left|h_n\left(\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right)\partial_\theta b(X_{t_{k-1}},\tilde\theta_n)\right|\\ &\lesssim \frac{1}{nh_n}\sum_{k=1}^n\left\|\int_{t_{k-1}}^{t_k}\big(b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big)\,ds\right\|_{L^2(\Omega)}\big\|1+|X_{t_{k-1}}|^N\big\|_{L^2(\Omega)}\\ &\lesssim \frac{1}{nh_n}\sum_{k=1}^n\int_{t_{k-1}}^{t_k}\big\|b(X_s,\theta_0)-b(X_{t_{k-1}},\theta_0)\big\|_{L^2(\Omega)}\,ds\lesssim h_n^H.\end{aligned}$$

Therefore

$$\tau_n^H \frac{1}{nh_n^2} \sum_{k=1}^{n} \left(\int_{t_{k-1}}^{t_k} \big(b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_0)\big)\, ds \right) \left(\int_{t_{k-1}}^{t_k} \big(b(X_{t_{k-1}},\theta_0) - b(X_{t_{k-1}},\theta_n)\big)\, ds \right) \xrightarrow{L^1} 0,$$

as $n \to \infty$. With Taylor's formula again, we have

$$\tau_n^H \frac{1}{nh_n^2} \sum_{k=1}^{n} \left(\int_{t_{k-1}}^{t_k} \big(b(X_{t_{k-1}},\theta_0) - b(X_{t_{k-1}},\theta_n)\big)\, ds \right)^2 = \frac{\tau_n^H}{n} \sum_{k=1}^{n} \left(\partial_\theta b(X_{t_{k-1}},\tilde{\theta}_n)(\theta_0 - \theta_n) \right)^2.$$

Through Theorem 5.1 and Proposition 5.1, we obtain that

$$\tau_n^H \frac{1}{nh_n^2} \sum_{k=1}^{n} \left(\int_{t_{k-1}}^{t_k} \big(b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_n)\big)\, ds \right)^2 = \zeta_n o_P(1).$$

Let us consider the cross term. At first, we decompose as follows

$$\begin{aligned}
&\tau_n^H \frac{2\sigma}{nh_n^2} \sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}}) \int_{t_{k-1}}^{t_k} \big(b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_n)\big)\, ds \\
&= \tau_n^H \frac{2\sigma}{nh_n^2} \sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}}) \int_{t_{k-1}}^{t_k} \big(b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_0)\big)\, ds \\
&\qquad + \tau_n^H \frac{2\sigma}{nh_n^2} \sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}}) \int_{t_{k-1}}^{t_k} \big(b(X_{t_{k-1}},\theta_0) - b(X_{t_{k-1}},\theta_n)\big)\, ds.
\end{aligned}$$

From Cauchy–Schwartz and Minkowski's inequalities and (5.3), we obtain

$$\begin{aligned}
&E\left| (B_{t_k} - B_{t_{k-1}}) \int_{t_{k-1}}^{t_k} \big(b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_0)\big)\, ds \right| \\
&\le \left\| B_{t_k} - B_{t_{k-1}} \right\|_{L^2(\Omega)} \int_{t_{k-1}}^{t_k} \left\| b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_0) \right\|_{L^2(\Omega)} ds \lesssim h_n^{2H+1}.
\end{aligned}$$

Thus

$$\tau_n^H \frac{2\sigma}{nh_n^2} \sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}}) \int_{t_{k-1}}^{t_k} \big(b(X_s,\theta_0) - b(X_{t_{k-1}},\theta_0)\big)\, ds = o_P(1),$$

and we conclude that

$$\tau_n^H Q_n(\theta_n) = \tau_n^H \frac{1}{nh_n^2} \sum_{k=1}^n \left[\sigma^2 (B_{t_k} - B_{t_{k-1}})^2 - \sigma^2 h_n^{2H}\right] + \frac{2\sigma}{nh_n} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}}) \partial_\theta b(X_{t_{k-1}}, \theta_0) \zeta_n$$
$$+ \zeta_n o_P(1) + o_P(1).$$

We will prove the second equality. Using (3.1), we have

$$\begin{aligned}
\partial_\theta Q_n(\theta_n) &= \frac{-2}{nh_n} \sum_{k=1}^n \left(X_{t_k} - X_{t_{k-1}} - h_n b(X_{t_{k-1}}, \theta_n)\right) \partial_\theta b(X_{t_{k-1}}, \theta_n) \\
&= \frac{-2}{nh_n} \sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} \left(b(X_s, \theta_0) - b(X_{t_{k-1}}, \theta_n)\right) ds + \sigma (B_{t_k} - B_{t_{k-1}}) \right) \partial_\theta b(X_{t_{k-1}}, \theta_n).
\end{aligned}$$

Let us now apply Taylor's formula to obtain that

$$\begin{aligned}
&\frac{-2}{nh_n} \sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} \left(b(X_s, \theta_0) - b(X_{t_{k-1}}, \theta_n)\right) ds \right) \partial_\theta b(X_{t_{k-1}}, \theta_n) \\
&\quad = \frac{-2}{nh_n} \sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} \left(b(X_s, \theta_0) - b(X_{t_{k-1}}, \theta_0)\right) ds \right) \partial_\theta b(X_{t_{k-1}}, \theta_n) \\
&\qquad + \frac{-2}{nh_n} \sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} \left(b(X_{t_{k-1}}, \theta_0) - b(X_{t_{k-1}}, \theta_n)\right) ds \right) \partial_\theta b(X_{t_{k-1}}, \theta_n) \\
&\quad = -\frac{2}{n} \sum_{k=1}^n \partial_\theta b(X_{t_{k-1}}, \theta_0)(\theta_0 - \theta_n) \partial_\theta b(X_{t_{k-1}}, \theta_0) + o_P(1).
\end{aligned}$$

By Lemma 5.1 and Theorem 5.1, we obtain the results of Lemma 5.2. □

Proof of Theorem 5.2. Since the relationship $\tau_n^H Q_n(\theta_n) \partial_\theta Q_n(\theta_n) = 0$ holds, we have

$$\left(\tau_n^H \frac{1}{nh_n^2} \sum_{k=1}^n \left[\sigma^2 (B_{t_k} - B_{t_{k-1}})^2 - \sigma^2 h_n^{2H}\right] + \left(\frac{2\sigma}{nh_n} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}}) \partial_\theta b(X_{t_{k-1}}, \theta_0) + o_P(1) \right) \zeta_n + o_P(1) \right)$$
$$\times \left(-\frac{2\sigma}{nh_n} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}}) \partial_\theta b(X_{t_{k-1}}, \theta_n) + o_P(1) \right) = 0.$$

By using Lemma 5.1, Proposition 5.3, we conclude the statement of Theorem 5.2. □

References

[1] Baldi, P., Ben Arous, G. and Kerkyacharian, G. (1992). Large deviations and the Strassen theorem in Hölder norm, *Stochastic Process Appl.* **42**, (1), 171-180.

[2] Brouste, A., and Iacus, S. M. (2013). Parameter estimation for the discretely observed fractional Ornstein - Uhlenbeck process and the Yuima R package. *Computer Statistics,* **28**(4), 1529-1547.

[3] Brouste, A. and Kleptsyna, M. (2010). Asymptotic properties of MLE for partially observed fractional diffusion system, *Stat. Infer. Stochastic Process.,* **13** (1), 1-13.

[4] Chiba, K. (2020). An M-estimator for stochastic differential equations driven by fractional Brownian motion with small Hurst parameter, *Statistical Inference for Stochastic Processes,* **23**, 319–353.

[5] Da Prato, G. and Zabczyk, J. (1992). *Stochastic Equations in Infinite Dimensions.* Encyclopedia Math. Appl., vol.44, Cambridge University Press. Cambridge.

[6] Fritz, P.K. and Hairer, M. (2014). *A Course on Rough Paths.* Universitext. Springer, Cham.

[7] Garrido–Atienza, M., Kloeden, P. and Neuenkirch, A. (2009). Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion, Appl. Math. Optim, **60** (2), 151–172.

[8] Garsia, A.M., Rodemich, H. and Rumsey Jr. (1971). A real variable lemma and the continuity of paths of some Gaussian processes, Indiana Univ. Math. J., **20**, 565 - 578.

[9] Gatheral, J., Jaisson, T. and Rosenbaum, M. (2017). Volatility is rough. *Quantitative Finance,* **18**, (6), 933-949.

[10] Hairer, M. (2009). An introduction to stochastic PDEs. Available at http://www.hairer.org/notes/SPDEs.pdf.

[11] Hairer, M. and Pillai, N.S. (2011). Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **47**, (2), 601-628.

[12] Haress, E.M. and Hu, Y. (2021). Estimation of all parameters in the fractional Ornstein–Uhlenbeck model under discrete observations. *Statistical Inference for Stochastic Processes,* **24**, 327-351.

[13] Hu, Y., Nualart, D. and Zhou, H. (2019). Drift parameter estimation for nonlinear stochastic differential equations driven by fractional Brownian motion, Stochastics, **91**, 1–25.

[14] Parameter estimation for fractional Orstein–Uhlenbeck processes of general Hurst parameter. *Statistical Inference for Stochastic Processes*, **22**, 111-142.

[15] I.A. Ibragimov and R.Z. Has'minskii, (1981). *Statistical estimation: Asymptotic theory.* New York: Springer.

[16] Kasonga, R.A. (1988). The consistency of a nonlinear least squares estimator for diffusion processes, Stochastic Process. Appl. **30**, 263–275.

[17] Kleptsyna, M. and Le Breton, (2002). A Statistical analysis of the fractional Ornstein Uhlenbeck type process. *Statistical Inference for Stochastic Processes*, **5** (3), 229–248.

[18] Kozachenko, Y., Melnikov, A. and Mishura, Y. (2014). On drift parameter estimation in models with fractional Brownian motion. *Statistics,* **49**, 35-62.

[19] Kubilius, K. and Skorniakov, V. (2016). On some estimators of the Hurst index of the solution of SDE driven by a fractional Brownian motion. *Statistics and Probability Letters,* **109**, 159-167.

[20] Kutoyants, Y.A. (1984). *Parameter Estimation for Stochastic Processes,* Heldermann, Berlin.

[21] Kutoyants, Y.A. (1994). *Identification of Dynamical Systems with Small Noise,* Kluwer, Dordrecht.

[22] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes,* Springer-Verlag, London, Berlin, Heidelberg.

[23] Le Gall, J.F. (2016). *Brownian Motion, Martingales, and Stochastic Calculus.* Springer, Heidelberg.

[24] Liptser, R.S. and Shiryayev, A.N. (2001). *Statistics of Random Processes: II Applications, Second Edition, Applications of Mathematics,* Springer-Verlag, Berlin, Heidelberg, New York.

[25] Liu, Y., Nualart, E. and Tindel, S. (2019). LAN property for stochastic differential equations with additive fractional, Stochastic Process. Appl., **129**, 2880–2902.

[26] Mishura, Y. (2008). *Stochastic Calculus for fractional Brownian motion and related processes*, Lecture Notes in Math. **1929**.

[27] Nakajima, S. and Shimizu,Y. (2022). Asymptotic normality of least squares type estimators to stochastic differential equations driven by fractional Brownian motions. *Statistics and Probability Letters,* **187**, 109476.

[28] Nakajima, S. and Shimizu,Y. (2022). Parameter estimation of stochastic differential equation driven by small fractional noise. *Statistics,* **56**, 919-934.

[29] Nakajima, S. and Shimizu,Y. (2022). Asymptotic inference for stochastic differential equations driven by fractional Brownian motion. *Japanese Journal of Statistics and Data Science*, Available online.

[30] Neuenkirch, A. and Tindel, S. (2014). A least square–type procedure for parameter estimation in stochastic differential equations with additive fractional noise, Stat. Inference Stoch. Process. **17** (1), 99–120.

[31] Norros, I., Valkeila, E. and Virtamo, J. (1999). An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions, *Bernoulli,* **5**, 571–587.

[32] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality,* Cambridge University Press, Cambridge.

[33] Nualart, D. and Ouknine, Y. (2002). Regularization of differential equations by fractional noise. *Stochastic Process. Appl.* **102**, 103–116.

[34] Nualart, D. and Răşcanu, A. (2002). Differential equations driven by fractional Brownian motion. *Collect. Math.* **53**, 55-81.

[35] Prakasa Rao, B.L.S. (1983). Asymptotic theory for nonlinear least squares estimator for diffusion processes, Math. Operations forschung Statist Ser. Statist. **14**, 195–209.

[36] Prakasa Rao, B.L.S. (1999). *Statistical Inference for Diffusion Type Processes,* Arnold, London, Oxford University Press, New York.

[37] Prakasa Rao, B.L.S. (2010). *Statistical Inference for Fractional Diffusion Processes*. Wiley Series in Probability and Statistics. Wiley, Chichester.

[38] Samko, S.G., Kilbas, A.A. and Marichev, O.O. (1993). *Fractional integrals and derivatives,* translated from the 1987 Russian original, Gordon and Breach, Yverdon.

[39] Shimizu, Y. (2017). Threshold Estimation for Stochastic Processes with Small Noise. *Scandinavian Journal of Statistics*, **44** (4), 951-988.

[40] Stroock, D.W. (2011). Essential of integration theory for analysis, Graduate Texts in Mathematics, **262**, Springer, xii+243 pp.

[41] Tudor, C. and Viens, F. (2007). Statistical aspects of the fractional stochastic calculus, *Ann. Stat.* **35**, 1183–1212.

[42] Uchida, M. Yoshida, N. (2004). Information criteria for small diffusions via the theory of Malliavin-Watanabe, *Stat. Inference Stoch. Process.* **7**, 35–67.

[43] Yoshida, N. (1992). Asymptotic expansion of maximum likelihood estimators for small diffusions via the theory of Malliavin–Watanabe, *Probab. Theory Relat. Fields*, **92**, 275–311.

[44] Yoshida, N. (2003). Conditional expansions and their applications, *Stochastic Process. Appl.* **107**, 53–81.

[45] Yoshida, N. (2011). Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Ann Inst Stat Math* **63** (3), 431–479.

[46] Young, LC. (1936). An inequality of Hölder type connected with Stieltjes integration. Acta Math, **67**, 251-282.

[47] Zähle, M. (1998). Integration with respect to fractal functions and stochastic calculus. *Probab. Theorey Related Fields* **111**, 333-374.

List of research achievements for application of Doctor of Science, Waseda University

Full Name： 中島　翔平　　seal or signature

Date Submitted(yyyy/mm/dd): 2022/11/1

種類別 (By Type)	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
	○Asymptotic inference for stochastic differential equations driven by fractional Brownian motion, Japanese Journal of Statistics and Data Science, available online, 2022 October, S. Nakajima and Y. Shimizu.
	○Parameter estimation of stochastic differential equation driven by small fractional noise, Statistics, 56 (4), 2022, July, 919-934, S. Nakajima and Y. Shimizu
	○Asymptotic normality of least squares type estimators to stochastic differential equations driven by fractional Brownian motions, Statistics and Probability Letters, 187, 109476, 2022, April, S. Nakajima and Y. Shimizu
	Existence of weak solutions to SPDEs with fractional Laplacian and non-Lipschitz coefficients , Stochastics and Partial Differential Equations: Analysis and Computations, 10, 255-277, 2021, June, S. Nakajima.