Statistical Inference for Stochastic Processes (2021) 24:327–351
https://doi.org/10.1007/s11203-020-09235-z

Estimation of all parameters in the fractional Ornstein–Uhlenbeck model under discrete observations

El Mehdi Haress[1] · Yaozhong Hu[2]

Received: 15 May 2020 / Accepted: 24 December 2020 / Published online: 13 January 2021
© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract
Let the Ornstein–Uhlenbeck process $(X_t)_{t\geq 0}$ driven by a fractional Brownian motion B^H described by $dX_t = -\theta X_t dt + \sigma dB_t^H$ be observed at discrete time instants $t_k = kh$, $k = 0, 1, 2, \ldots, 2n+2$. We propose an ergodic type statistical estimator $\hat{\theta}_n$, $\hat{H}_n$ and $\hat{\sigma}_n$ to estimate all the parameters θ, H and σ in the above Ornstein–Uhlenbeck model simultaneously. We prove the strong consistence and the rate of convergence of the estimator. The step size h can be arbitrarily fixed and will not be forced to go zero, which is usually a reality. The tools to use are the generalized moment approach (via ergodic theorem) and the Malliavin calculus.

Keywords Fractional Brownian motion · Fractional Ornstein–Uhlenbeck · Parameter estimation · Malliavin calculus · Ergodicity · Stationary processes · Newton method · Central limit theorem

Mathematics Subject Classification 62M09 · 60G22 · 60H10 · 60H30

1 Introduction

The Ornstein–Uhlenbeck process $(X_t)_{t\geq 0}$ is described by the following Langevin equation:

$$dX_t = -\theta X_t dt + \sigma \mathrm{d}B_t^H, \tag{1.1}$$

where $\theta > 0$ so that the process is ergodic and where for simplicity of the presentation we assume $X_0 = 0$. Other initial value can be treated exactly in the same way. We assume that the process $(X_t)_{t\geq 0}$ is observed at discrete time instants $t_k = kh$ and we want to use the

Supported by NSERC discovery grant and a startup fund of University of Alberta.

✉ Yaozhong Hu
yaozhong@ualberta.ca

El Mehdi Haress
el-mehdi.haress@student.ecp.fr

1 University of Paris-Saclay, Gif-sur-Yvette, France

2 Department of Mathematical and Statistical Sciences, University of Alberta at Edmonton, Edmonton T6G 2G1, Canada

observations $\{X_h, X_{2h}, \ldots, X_{2n+2h}\}$ to estimate the parameters θ, H and σ that appear in the above Langevin equation simultaneously.

Before we continue let us briefly recall some recent relevant works obtained in literature. Most of the works deal with the estimator of the drift parameter θ. In fact, when the Ornstein–Uhlenbeck process $(X_t)_{t\geq 0}$ can be observed continuously and when the parameters σ and H are assumed to be known, we have the following results :

1. The maximum likelihood estimator for θ defined by θ_T^{mle} is studied Tudor and Viens (2007) (see also the references therein for earlier references), and is proved to be strongly consistent. The asymptotic behavior of the bias and the mean square of θ_T^{mle} is also given. In this paper, a strongly consistent estimator of σ is also proposed.
2. A least squares estimator defined by $\tilde{\theta}_T = \frac{-\int_0^T X_t dX_t}{\int_0^T X_t^2 dt}$ was studied in Chen et al. (2017), Hu and Nualart (2010) and Hu et al. (2019). It is proved that $\tilde{\theta}_T \to \theta$ almost surely as $T \to \infty$. It is also proved that when $H \leq 3/4$, $\sqrt{T}(\tilde{\theta}_T - \theta)$ converges in law to a mean zero normal random variable. The variance of this normal variable is also obtained. When $H \geq 3/4$, the rate of convergence is also known Hu et al. (2019).

Usually in reality the process can only be observed at discrete times $\{t_k = kh, k = 1, 2, \ldots, n\}$ for some fixed observation time lag $h > 0$. In this very interesting case, there are very limited works. Let us only mention two (Hu and Song 2013; Panloup et al. 2019). To the best of knowledge there is only one work (Brouste and Iacus 2013) that estimates all the parameters θ, H and σ at the same time, but the observations are assumed to be made continuously.

The diffusion coefficient σ represents the "volatility" and it is commonly believed that it should be computed (hence estimated) by the $1/H$ variations (see Hu et al. 2019 and references therein). To use the $1/H$ variations one has to assume the process can be observed continuously (or we have high frequency data). Namely, it is a common belief that σ can only be estimated when one has high frequency data.

In this work, we assume that the process can only be observed at discrete times $\{t_k = kh, k = 1, 2, \ldots, n\}$ for some arbitrarily fixed observation time lag $h > 0$ (without the requirement that $h \to 0$). We want to estimate θ, H and σ simultaneously. The idea we use is the ergodic theorem, namely, we find the explicit form of the limit distribution of $\frac{1}{n}\sum_{k=1}^n f(X_{kh})$ and use it to estimate our parameters. People may naturally think that if we appropriately choose three different f, then we may obtain three different equations to obtain all the three parameters θ, H and σ.

However, this is impossible since as long as we proceed this way, we shall find out that whatever we choose f, we cannot get independent equations. Motivated by a recent work [4], we may try to add the limit distribution of $\frac{1}{n}\sum_{k=1}^n g(X_{kh}, X_{(k+1)h})$ to find all the parameters. However, this is still impossible because regardless how we choose f and g we obtain only two independent equations. This is because regardless how we choose f and g the limits depends only on the covariance of the limiting Gaussians (see Y_0 and Y_h ulteriorly). Finally, we propose to use the following quantities to estimate all the three parameters θ, H and σ:

$$\frac{\sum_{k=1}^n X_{kh}^2}{n}, \quad \frac{\sum_{k=1}^n X_{kh}X_{kh+h}}{n}, \quad \frac{\sum_{k=1}^n X_{kh}X_{kh+2h}}{n}. \tag{1.2}$$

We shall study the strong consistence and joint limiting law of our estimators. The above three series converge to $\mathbb{E}(Y_0^2)$, $\mathbb{E}(Y_0 Y_h)$, and $\mathbb{E}(Y_0 Y_{2h})$ respectively. It should be emphasized that it seems that we cannot use the joint distribution of Y_0, Y_h alone to estimate all the three parameters θ, H and σ, we need to the joint distribution of Y_0, Y_h, Y_{2h}.

The paper is organized as follows. In Sect. 2, we recall some known results. The construction and the strong consistency of the estimators are provided in Sect. 3. Central limit theorems are obtained in Sect. 4. To make the paper more readable, we delay some proofs in Append A. To use our estimators we need the determinant of some functions to be nondegenerate. This is given in Appendix B. Some numerical simulations to validate our estimators are illustrated in Appendix C.

2 Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. The expectation on this space is denoted by $\mathbb{E}$. The fractional Brownian motion $(B_t^H, t \in \mathbb{R})$ with Hurst parameter $H \in (0, 1)$ is a zero mean Gaussian process with the following covariance structure:

$$\mathbb{E}(B_t^H B_s^H) = R_H(t, s) = \frac{1}{2}(\mid t \mid^{2H} + \mid s \mid^{2H} - \mid t - s \mid^{2H}), \quad \forall\, t, s \in \mathbb{R}. \tag{2.1}$$

On stochastic analysis of this fractional Brownian motion, such as stochastic integral $\int_a^b f(t) dB_t^H$, chaos expansion, and stochastic differential equation $dX_t = b(X_t)dt + \sigma(X_t)dB_t^H$ we refer to Biagini et al. (2008).

For any $s, t \in \mathbb{R}$, we define

$$\langle I_{[0,t]}, I_{[0,s]} \rangle_{\mathcal{H}} = R_H(s, t), \tag{2.2}$$

where $I_{[a,b]}$ denotes the indicate function on $[a, b]$ and we use $I_{[b,a]} = -I_{[a,b]}$ for any $a < b$. We can first extend this scalar product to general elementary functions $f(\cdot) = \sum_{i=1}^n a_i I_{[0,s_i]}(\cdot)$ by (bi-)linearity and then to general function by a limiting argument. We can then obtain the reproducing kernel Hilbert space, denoted by $\mathcal{H}$, associated with this Gaussian process B_t^H (see e.g. Hu and Nualart 2010 for more details).

Let $\mathcal{S}$ be the space of smooth and cylindrical random variables of the form

$$F = f(B^H(\phi_1), \ldots, B^H(\phi_n)), \quad \phi_1, \ldots, \phi_n \in C_0^\infty([0, T]),$$

where $f \in C_b^\infty(\mathbf{R}^n)$ and $B^H(\phi) = \int_0^\infty \phi(t) dB_t^H$. For such a variable F, we define its Malliavin derivative as the $\mathcal{H}$ valued random element:

$$DF = \sum_{k=1}^n \frac{\partial f}{\partial x_i}(B^H(\phi_1), \ldots, B^H(\phi_n))\phi_i.$$

We shall use the following result in Sect. 4 to obtain the central limit theorem. We refer to Hu (2017) and many other references for a proof.

Proposition 2.1 *Let $\{F_n, n \geq 1\}$ be a sequence of random variables in the space of p-th Wiener Chaos, $p \geq 2$,such that $\lim_{n\to\infty} \mathbb{E}(F_n^2) = \sigma^2$. Then the following statements are equivalent:*

(i) F_n converges in law to $N(0, \sigma^2)$ as n tends to infinity.
(ii) $\|DF_n\|_{\mathcal{H}}^2$ converges in L^2 to a constant as n tends to infinity.

3 Estimators of *θ*,*H* and *σ*

If $X_0 = 0$, then the solution X_t to (1.1) can be expressed as

$$X_t = \sigma \int_0^t e^{-\theta(t-s)} \mathrm{d}B_s^H. \tag{3.1}$$

The associated stationary solution, the solution of (1.1) with the initial value

$$Y_0 = \int_{-\infty}^0 e^{\theta s} \mathrm{d}B_s^H, \tag{3.2}$$

can be expressed as

$$Y_t = \int_{-\infty}^t e^{-\theta(t-s)} \mathrm{d}B_s^H = e^{-\theta t} Y_0 + X_t. \tag{3.3}$$

Y_t is stationary, namely, the Y_t has the same distribution as that of Y_0 which is also the limiting normal distribution of X_t (when $t \to \infty$). Let's consider the following three quantities :

$$\begin{cases} \eta_n = \frac{1}{n} \sum_{k=1}^n X_{kh}^2, \\ \eta_{h,n} = \frac{1}{n} \sum_{k=1}^n X_{kh} X_{kh+h}, \\ \eta_{2h,n} = \frac{1}{n} \sum_{k=1}^n X_{kh} X_{kh+2h}. \end{cases} \tag{3.4}$$

As in Kubilius et al. (2017, Section 1.3.2.2), we have the following ergodic result:

$$\lim_{n\to\infty} \eta_n = \mathbb{E}(Y_0^2) = \sigma^2 H \Gamma(2H) \theta^{-2H}. \tag{3.5}$$

Now we want to have a similar result for $\eta_{h,n}$. First, let's study the ergodicity of the processes $\{Y_{t+h} - Y_t\}_{t\geq 0}$. According to Magdziarz and Weron (2011), a centered Gaussian wide-sense stationary process M_t is ergodic if $\mathbb{E}(M_t M_0) \to 0$ as t tends to infinity. We shall apply this result to $M_t = Y_{t+h} - Y_t$, $t \geq 0$. Obviously, it is a centered Gaussian stationary process and

$$\mathbb{E}((Y_{t+h} - Y_t)(Y_h - Y_0)) = \mathbb{E}(Y_{t+h} Y_h) - \mathbb{E}(Y_{t+h} Y_0) - \mathbb{E}(Y_t Y_h) + \mathbb{E}(Y_t Y_0).$$

In Cheridito et al. (2003, Theorem 2.3), it is proved that $\mathbb{E}(Y_t Y_0) \to 0$ as t goes to infinity. Thus, it is easy to see that $\mathbb{E}((Y_{t+h} - Y_t)(Y_h - Y_0)) \to 0$. Hence, we see that the process $\{Y_{t+h} - Y_t\}_{t\geq 0}$ is ergodic. This implies

$$\frac{\sum_{k=1}^n [Y_{(k+1)h} - Y_{kh}]^2}{n} \to_{n\to\infty} \mathbb{E}([Y_h - Y_0]^2).$$

This combined with (3.5) yields the following Lemma.

Theorem 3.1 *Let η_n, $\eta_{h,n}$ and $\eta_{2h,n}$ be defined by (3.4). Then as $n \to \infty$ we have almost surely*

$$\lim_{n\to\infty} \eta_n = \mathbb{E}(Y_0^2) = \sigma^2 H \Gamma(2H) \theta^{-2H}; \tag{3.6}$$

$$\lim_{n\to\infty} \eta_{h,n} = \mathbb{E}(Y_0 Y_h) = \sigma^2 \frac{\Gamma(2H+1)\sin(\pi H)}{2\pi} \int_{-\infty}^{\infty} e^{ixh} \frac{|x|^{1-2H}}{\theta^2 + x^2} dx; \tag{3.7}$$

$$\lim_{n\to\infty} \eta_{2h,n} = \mathbb{E}(Y_0 Y_{2h}) = \sigma^2 \frac{\Gamma(2H+1)\sin(\pi H)}{2\pi} \int_{-\infty}^{\infty} e^{2ixh} \frac{|x|^{1-2H}}{\theta^2 + x^2} dx. \tag{3.8}$$

The explicit expressions of $\mathbb{E}(Y_0 Y_h)$ and $\mathbb{E}(Y_0 Y_{2h})$ are borrowed from Cheridito et al. (2003, Remark 2.4).

From the above theorem we propose the following construction for the estimators of the parameters θ, H and σ.

First let us define

$$\begin{cases} f_1(\theta, H, \sigma) := \sigma^2 H \Gamma(2H) \theta^{-2H} \,; \\ f_2(\theta, H, \sigma) := \frac{1}{\pi} \sigma^2 \Gamma(2H+1) \sin(\pi H) \int_0^\infty \cos(hx) \frac{x^{1-2H}}{\theta^2 + x^2} dx \,; \\ f_3(\theta, H, \sigma) := \frac{1}{\pi} \sigma^2 \Gamma(2H+1) \sin(\pi H) \int_0^\infty \cos(2hx) \frac{x^{1-2H}}{\theta^2 + x^2} dx . \end{cases} \tag{3.9}$$

It is elementary to verify (we fix $h > 0$) that $f_1(\theta, H, \sigma)$, $f_2(\theta, H, \sigma)$, $f_3(\theta, H, \sigma)$ are continuously differentiable functions of $\theta > 0, \sigma > 0$ and $H \in (0, 1)$. Let $f(\theta, H, \sigma) = (f_1(\theta, H, \sigma), f_2(\theta, H, \sigma), f_3(\theta, H, \sigma))^T$ be a vector function defined on $\theta > 0, \sigma > 0$ and $H \in (0, 1)$. Then we set

$$\begin{cases} f_1(\theta, H, \sigma) = \eta_n = \frac{1}{n} \sum_{k=1}^n X_{kh}^2 \,; \\ f_2(\theta, H, \sigma) = \eta_{h,n} = \frac{1}{n} \sum_{k=1}^n X_{kh} X_{kh+h} \,; \\ f_3(\theta, H, \sigma) = \eta_{2h,n} = \frac{1}{n} \sum_{k=1}^n X_{kh} X_{kh+2h} , \end{cases} \tag{3.10}$$

as a system of three equations for the three unknowns (θ, H, σ). The Jacobian of f, denoted by $J(\theta, H, \sigma)$, is an elementary function whose explicit form can be obtained in a straightforward way. However, this explicit expression is extremely complicated and involves the complicated integrations as well. It is hard to find the range of the parameters analytically so that the determinant of the Jacobian $J(\theta, H, \sigma)$ is not singular (nonzero). In Appendix B, we shall give a more detailed account for the determinant of the Jacobian $J(\theta, H, \sigma)$ and in particular we shall demonstrate

$$\det(J(\theta, H, \sigma)) \neq 0, \quad \forall\, (\theta, H, \sigma) \in \mathbb{D}_h, \tag{3.11}$$

where

$$\mathbb{D}_h = \{(\theta, H, \sigma) : \, 2/h < \theta < 10/h, \quad H \in (0.3, 1/2) \cup (1/2, 3/4), \sigma > 0\} . \tag{3.12}$$

Our approach there is a numerical one. We can try to plot more values to enlarge the domain $\mathbb{D}_h$. However, we shall not pursue along this direction. By the inverse function theorem, we see that for any point $(\theta_0, H_0, \sigma_0)$ in $\mathbb{D}_h$, there is a neighbourhood U of $(\theta_0, H_0, \sigma_0)$ and a neighbourhood V of $f(\theta_0, H_0, \sigma_0)$ such that the function f has a continuously differentiable inverse f^{-1} from V to U. From Theorem 3.1 we know that if the true parameter is $(\theta_0, H_0, \sigma_0)$, then $\nu_n = (\eta_n, \eta_{h,n}, \eta_{2h,n})$ converges almost surely to $f(\theta_0, H_0, \sigma_0)$ as $n \to \infty$. This means that there is a $N = N(\omega)$ such that when $n \geq N$, $\nu_n = (\eta_n, \eta_{h,n}, \eta_{2h,n}) \in V$. In other words, when n is sufficiently large, the Eq. (3.10) has a (unique) solution in the neighbourhood of $(\theta_0, H_0, \sigma_0)$.

Theorem 3.2 *If $(\theta, H, \sigma) \in \mathbb{D}_h$, then when n is sufficiently large the Eq.* (3.10) *has a solution in $\mathbb{D}_h$ and in a neighbourhood of (θ, H, σ) the solution is unique denoted by $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$. Moreover, $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$ converge almost surely to (θ, H, σ) as n tends to infinity.*

We shall use $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$ to estimate the parameters (θ, H, σ). We call $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$ the ergodic (or generalized moment) estimator of (θ, H, σ).

It seems hard to explicitly obtain the explicit solution of the system of Eq. (3.10). However, it is a classical problem. There are copious numeric approaches to find the approximate solution. We shall give some validation of our estimators numerically in “Appendix C”.

Theorem 3.2 states that the ergodic estimator exists uniquely in a neighbourhood of the true parameter (θ, H, σ). However, does the Eq. (3.10) have more than one solution on the domain $\mathbb{D}_h$? The global inverse function theorem is much more sophisticated. There are several extension of the Hadamard–Caccioppoli theorem (e.g. Mustafa et al. 2007). However, it seems that these works can hardly be applied to our situation. It seems impossible to use the determinant alone to determine if a mapping has a global inverse or not. For example, the function $(f(x, y), g(x, y)) = (e^x \cos y, e^x \sin y)$ has a strictly positive determinant on $\mathbb{R}^2$. This function is a surjection from $\mathbb{R}^2$ onto $\mathbb{R}^2\backslash\{0\}$, but it is not an injection. For this reason we are not going to obtain rigorous results on the uniqueness of the solution to (3.10) on the whole domain $\mathbb{D}_h$ in the present paper. However, we propose the following two points in statistical practice to determine the estimator $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$.

(1) Dividing the second and third equations by the first one in (3.9) and noticing $\Gamma(2H+1) = 2H\Gamma(2H)$ we have

$$\begin{cases} \frac{f_2}{f_1} = \frac{2\sin(\pi H)\theta^{2H}}{\pi} \int_0^\infty \cos(hx) \frac{x^{1-2H}}{\theta^2+x^2} dx, \\ \frac{f_3}{f_1} = \frac{2\sin(\pi H)\theta^{2H}}{\pi} \int_0^\infty \cos(2hx) \frac{x^{1-2H}}{\theta^2+x^2} dx, \end{cases} \tag{3.13}$$

where we recall that f_1, f_2, f_3 are given by the right hand side (3.10), which are determined from the real observations of the process. Denote

$$\mathcal{I}_h(\theta, H) := \frac{2\sin(\pi H)\theta^{2H}}{\pi} \int_0^\infty \cos(hx) \frac{x^{1-2H}}{\theta^2 + x^2} dx. \tag{3.14}$$

We obtain a system of equations for (θ, H):

$$\begin{cases} \mathcal{I}_h(\theta, H) = \frac{f_2}{f_1}, \\ \mathcal{I}_{2h}(\theta, H) = \frac{f_3}{f_1}. \end{cases} \tag{3.15}$$

When the real data are observed and when one knows a priori the domain (say the projection of $\mathbb{D}_h$ onto the (θ, H) plane) of the parameter (θ, H), one can plot the function

$$g(\theta, H) := \left\| \mathcal{I}_h(\theta, H) - \frac{f_2}{f_1} \right\|^2 + \left\| \mathcal{I}_{2h}(\theta, H) - \frac{f_3}{f_1} \right\|^2$$

on that domain to see if it reaches its minimum 0 only at one point (θ, H). We carry out a simulation of the process with $\theta = 6$, $\sigma = 2$ and $H = 0.7$ for the f_1, f_2, f_3 and we plot the function g for $h = 0.1$ as Fig. 1 (for $n = 2^{10}$). A quick computation shows that g reaches its minimum 0 only for one point $(\theta = 7.833, H = 0.7133)$.

(2) In the case that one finds several solutions to (3.10), a second way to select which one as the ergodic estimator may follow the following principle. Choose appropriately some positive integers N_1, N_2, N_3 and let

$$\tilde{\mathbb{Z}} = \{(p, q, m),\ p = 1, \dots, N_1,\ q = 1, \dots, N_2,\ m = 1, \dots, N_3\}.$$

For each $(p, q, m) \in \tilde{\mathbb{Z}}$ compute $\eta_{p,q,m} = \frac{1}{n}\sum_{k=1}^n X_{kh}^p X_{kh+mh}^q$ and we know that this quantity will convergence to $\mathbb{E}(Y_0^p Y_{mh}^q)$ as $n \to \infty$. Thus, we may choose the one which minimizes $\sum_{(p,q,m)\in\tilde{\mathbb{Z}}} \left(\eta_{p,q,m} - \mathbb{E}(Y_0^p Y_{mh}^q)\right)^2$.

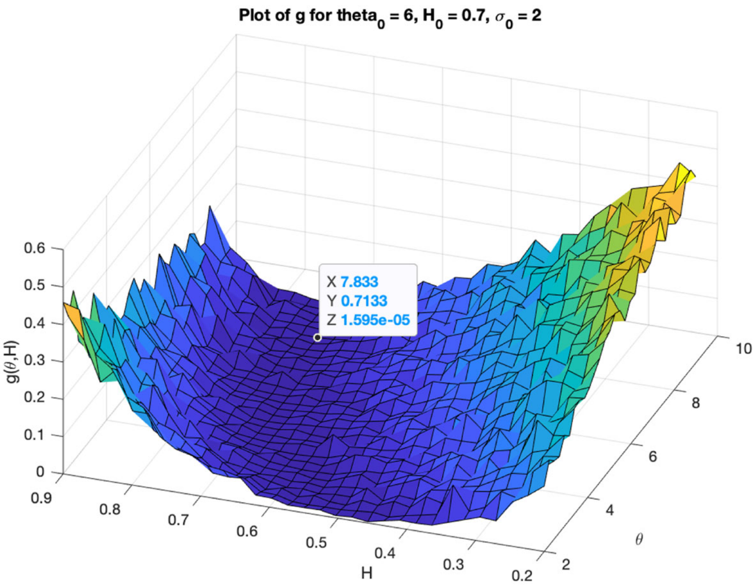

Fig. 1 Plot of the function g for $\theta = 6$, $\sigma = 2$, $H = 0.7$ and $n = 2^{10}$

4 Central limit theorem

In this section, we shall prove central limit theorem associated with our ergodic estimator $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$. We shall prove that $\sqrt{n}(\tilde{\theta}_n - \theta, \tilde{H}_n - H, \tilde{\sigma}_n - \sigma)$ converges in law to a mean zero normal vector.

Let's first consider the random variable F_n defined by

$$F_n = \begin{pmatrix} \sqrt{n}(\eta_n - \mathbb{E}(\eta_n)) \\ \sqrt{n}(\eta_{h,n} - \mathbb{E}(\eta_{h,n})) \\ \sqrt{n}(\eta_{2h,n} - \mathbb{E}(\eta_{2h,n})) \end{pmatrix}. \tag{4.1}$$

Our first goal is to show that F_n converges in law to a multivariate normal distribution using Proposition 2.1. So we consider a linear combination:

$$G_n = \alpha\sqrt{n}(\eta_n - \mathbb{E}(\eta_n)) + \beta\sqrt{n}(\eta_{h,n} - \mathbb{E}(\eta_{h,n})) + \gamma\sqrt{n}(\eta_{2h,n} - \mathbb{E}(\eta_{2h,n})), \tag{4.2}$$

and show that it converges to a normal distribution.

We will use the following Feynman diagram formula (Hu 2017), where interested readers can find a proof.

Proposition 4.1 *Let X_1, X_2, X_3, X_4 be jointly Gaussian random variables with mean zero. Then*

$$\mathbb{E}(X_1X_2X_3X_4) = \mathbb{E}(X_1X_2)\mathbb{E}(X_3X_4) + \mathbb{E}(X_1X_3)\mathbb{E}(X_2X_4) + \mathbb{E}(X_1X_4)\mathbb{E}(X_2X_3).$$

An immediate consequence of this result is

Proposition 4.2 *Let X_1, X_2, X_3, X_4 be jointly Gaussian random variables with mean zero. Then*

$$\begin{cases} \mathbb{E}\left[(X_1X_2 - \mathbb{E}(X_1X_2))(X_3X_4 - \mathbb{E}(X_3X_4))\right] \\ \quad = \mathbb{E}(X_1X_3)\mathbb{E}(X_2X_4) + \mathbb{E}(X_1X_4)\mathbb{E}(X_2X_3)\,; & (4.3) \\ \mathbb{E}\left[(X_1^2 - \mathbb{E}(X_1^2))(X_2X_3 - \mathbb{E}(X_2X_3))\right] = 2\mathbb{E}(X_1X_2)\mathbb{E}(X_1X_3)\,; & (4.4) \\ \mathbb{E}\left[(X_1^2 - \mathbb{E}(X_1^2))(X_2^2 - \mathbb{E}(X_2^2))\right] = 2\left[\mathbb{E}(X_1X_2)\right]^2. & (4.5) \end{cases}$$

Theorem 4.3 *Let $H \in (0, 1/2) \cup (1/2, 3/4)$. Let X_t be the Ornstein–Uhlenbeck process defined by Eq. (1.1) and let η_n, $\eta_{h,n}$, $\eta_{2h,n}$ be defined by (3.4). Then*

$$\begin{pmatrix} \sqrt{n}(\eta_n - \mathbb{E}(\eta_n)) \\ \sqrt{n}(\eta_{h,n} - \mathbb{E}(\eta_{h,n})) \\ \sqrt{n}(\eta_{2h,n} - \mathbb{E}(\eta_{2h,n})) \end{pmatrix} \to N(0, \Sigma), \tag{4.6}$$

where $\Sigma = (\Sigma(i,j))_{1\le i,j\le 3}$ is a positive semidefinite symmetric matrix whose elements are given by

$$\Sigma(1,1) = 2\left[\mathbb{E}(Y_0^2)\right]^2 + 4\sum_{m=1}^{\infty}\left[\mathbb{E}(Y_0Y_{mh})\right]^2\ ; \tag{4.7}$$

$$\Sigma(2,2) = \left[\mathbb{E}(Y_0^2)\right]^2 + \left[\mathbb{E}(Y_0Y_h)\right]^2 + 2\sum_{m=1}^{\infty}\left[\mathbb{E}(Y_0Y_{mh})\right]^2 + 2\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{(m-1)h})\mathbb{E}(Y_0Y_{(m+1)h})\,; \tag{4.8}$$

$$\Sigma(3,3) = \left[\mathbb{E}(Y_0^2)\right]^2 + \left[\mathbb{E}(Y_0Y_{2h})\right]^2 + 2\sum_{m=1}^{\infty}\left[\mathbb{E}(Y_0Y_{mh})\right]^2 + 2\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{|m-2|h})\mathbb{E}(Y_0Y_{(m+2)h})\,; \tag{4.9}$$

$$\Sigma(1,2) = \Sigma(2,1) = 4\sum_{m=0}^{\infty}\mathbb{E}(Y_0Y_{mh})\mathbb{E}(Y_0Y_{(m+1)h})\,; \tag{4.10}$$

$$\begin{aligned}\Sigma(2,3) = \Sigma(3,2) &= \mathbb{E}(Y_0^2)\mathbb{E}(Y_0Y_h) + \sum_{m=1}^{\infty}\mathbb{E}(Y_0^2)\left[\mathbb{E}(Y_0Y_{(m+1)h}) + \mathbb{E}(Y_0Y_{(m-1)h})\right] \\ &+ \mathbb{E}(Y_0Y_h)\mathbb{E}(Y_0Y_{2h}) + \sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{(m+2)h})\mathbb{E}(Y_0Y_{(m-1)h}) \\ &+ \sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{|m-2|h})\mathbb{E}(Y_0Y_{(m+1)h})\end{aligned} \tag{4.11}$$

$$\begin{aligned}\Sigma(1,3) = \Sigma(3,1) &= \mathbb{E}(Y_0^2)\mathbb{E}(Y_0Y_{2h}) + \sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{mh})\mathbb{E}(Y_0Y_{(m+2)h}) \\ &+ \sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{mh})\mathbb{E}(Y_0Y_{|m-2|h}).\end{aligned} \tag{4.12}$$

Remark 4.4 (1) It is easy from the following proof to see that all entries $\Sigma(i, j)$ of the covariance matrix Σ are finite.

(2) In an earlier work of Hu and Song it is said (Hu and Song 2013, equation (19.19)) that the variance Σ (corresponding to our $\Sigma(1, 1)$ in our notation) is independent of the time lag h. But there was an error on the bound of A_n on Hu and Song (2013, page 434, line 14). So, A_n there does not go to zero. Its limit is re-calculated in this work.

Proof We write

$$\mathbb{E}(G_n^2) = (\alpha, \beta, \gamma)\Sigma_n(\alpha, \beta, \gamma)^T, \quad \Sigma_n = (\Sigma_n(i, j))_{1\le i,j\le 3},$$

where Σ_n is a symmetric 3×3 matrix given by

$$\begin{cases} \Sigma_n(1, 1) = n\mathbb{E}\left[(\eta_n - \mathbb{E}(\eta_n))^2\right]; \\ \Sigma_n(1, 2) = \Sigma_n(2, 1) = n\mathbb{E}\left[(\eta_n - \mathbb{E}(\eta_n))(\eta_{h,n} - \mathbb{E}(\eta_{h,n}))\right]; \\ \Sigma_n(1, 3) = \Sigma_n(3, 1) = n\mathbb{E}\left[(\eta_n - \mathbb{E}(\eta_n))(\eta_{2h,n} - \mathbb{E}(\eta_{2h,n}))\right]; \\ \Sigma_n(2, 2) = n\mathbb{E}\left[(\eta_{h,n} - \mathbb{E}(\eta_{h,n}))^2\right]; \\ \Sigma_n(2, 3) = \Sigma_n(3, 2) = n\mathbb{E}\left[(\eta_{h,n} - \mathbb{E}(\eta_{h,n}))(\eta_{2h,n} - \mathbb{E}(\eta_{2h,n}))\right]; \\ \Sigma_n(3, 3) = n\mathbb{E}\left[(\eta_{2h,n} - \mathbb{E}(\eta_{2h,n}))^2\right]. \end{cases}$$

First, we compute the limit of $\Sigma_n(1, 1)$. From the Definition (3.4) of η_n and Proposition 4.2, we have

$$\begin{aligned} \Sigma_n(1, 1) &= \frac{1}{n}\sum_{k,k'=1}^{n} \mathbb{E}\left[(X_{kh}^2 - \mathbb{E}\left[(X_{kh})^2\right])(X_{k'h}^2 - \mathbb{E}\left[(X_{k'h})^2\right])\right] \\ &= \frac{2}{n}\sum_{k,k'=1}^{n} [\mathbb{E}(X_{kh}X_{k'h})]^2. \end{aligned}$$

By Lemma A.2 with $a = b = c = d = 0$, we see that

$$\Sigma_n(1, 1) \to \Sigma(1, 1) = 2\left[\mathbb{E}(Y_0^2)\right]^2 + 4\sum_{m=1}^{\infty} [\mathbb{E}(Y_0 Y_{mh})]^2. \tag{4.13}$$

This proves (4.7). For $\Sigma_n(2, 2)$ we have

$$\begin{aligned} \Sigma_n(2, 2) &= \frac{1}{n}\sum_{k,k'=1}^{n} \mathbb{E}(X_{kh}X_{(k'+1)h})\mathbb{E}(X_{(k+1)h}X_{k'h}) \\ &\quad + \frac{1}{n}\sum_{k,k'=1}^{n} \mathbb{E}(X_{kh}X_{k'h})\mathbb{E}(X_{(k+1)h}X_{(k'+1)h}) \\ &= I_{1,n} + I_{2,n}. \end{aligned} \tag{4.14}$$

By Lemma A.2 with $a = d = 0$ and $b = c = 1$, we see that that

$$I_{1,n} \to [\mathbb{E}(Y_0 Y_h)]^2 + 2\sum_{m=1}^{\infty} \mathbb{E}(Y_0 Y_{(m-1)h})\mathbb{E}(Y_0 Y_{(m+1)h}). \tag{4.15}$$

By Lemma A.2 with $a = b = 0$ and $c = d = 1$, we have

$$I_{2,n} \to \left[\mathbb{E}(Y_0^2)\right]^2 + 2\sum_{m=1}^{\infty} [\mathbb{E}(Y_0 Y_{mh})]^2. \tag{4.16}$$

This proves (4.8). As for $\Sigma_n(3,3)$ we have

$$\begin{aligned}\Sigma_n(3,3) &= \frac{1}{n}\sum_{k,k'=1}^{n}\mathbb{E}(X_{kh}X_{(k'+2)h})\mathbb{E}(X_{(k+2)h}X_{k'h}) \\ &\quad+\frac{1}{n}\sum_{k,k'=1}^{n}\mathbb{E}(X_{kh}X_{k'h})\mathbb{E}(X_{(k+2)h}X_{(k'+2)h}) \\ &\to [\mathbb{E}\,(Y_0Y_{2h})]^2+2\sum_{m=1}^{\infty}\mathbb{E}\left(Y_0Y_{(m+2)}\right)\mathbb{E}\left(Y_0Y_{|m-2|h}\right) \\ &\quad+\left[\mathbb{E}\left(Y_0^2\right)\right]^2+2\sum_{m=1}^{\infty}[\mathbb{E}\,(Y_0Y_{mh})]^2\,. \end{aligned} \tag{4.17}$$

This proves (4.9).

Now let consider the limit of $\Sigma_n(1,2)$. From the Definition (3.4) and from Proposition 4.2 it follows

$$\begin{aligned}\Sigma_n(1,2) &= \frac{2}{n}\sum_{k,k'=1}^{n}\mathbb{E}(X_{kh}X_{k'h})\mathbb{E}(X_{kh}X_{(k'+1)h}) \\ &\to \mathbb{E}(Y_0^2)\mathbb{E}(Y_0Y_1)+\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_m)\mathbb{E}(Y_0Y_{m+1})+\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_m)\mathbb{E}(Y_0Y_{m-1}) \\ &= 4\sum_{m=0}^{\infty}\mathbb{E}(Y_0Y_{mh})\mathbb{E}(Y_0Y_{(m+1)h}). \end{aligned} \tag{4.18}$$

This proves (4.10). As for $\Sigma_n(2,3)$ we have similarly

$$\begin{aligned}\Sigma_n(2,3) &= \frac{1}{n}\sum_{k,k'=1}^{n}\mathbb{E}(X_{kh}X_{k'h})\mathbb{E}(X_{(k+1)h}X_{(k'+2)h}) \\ &\quad+\frac{1}{n}\sum_{k,k'=1}^{n}\mathbb{E}(X_{kh}X_{(k'+2)h})\mathbb{E}(X_{(k+1)h}X_{k'h}) \\ &\to \mathbb{E}(Y_0^2)\mathbb{E}(Y_0Y_h)+\sum_{m=1}^{\infty}\mathbb{E}(Y_0^2)\left[\mathbb{E}(Y_0Y_{(m+1)h})+\mathbb{E}(Y_0Y_{(m-1)h})\right] \\ &\quad+\mathbb{E}(Y_0Y_h)\mathbb{E}(Y_0Y_{2h})+\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{(m+2)h})\mathbb{E}(Y_0Y_{(m-1)h}) \\ &\quad+\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{|m-2|h})\mathbb{E}(Y_0Y_{(m+1)h}). \end{aligned} \tag{4.19}$$

This is (4.11). Lastly, to get (4.12) we use

$$\begin{aligned}\Sigma_n(1,3) &= \frac{2}{n}\sum_{k,k'=1}^{n}\mathbb{E}(X_{kh}X_{k'h})\mathbb{E}(X_{kh}X_{(k'+2)h}) \\ &\to \mathbb{E}(Y_0^2)\mathbb{E}(Y_0Y_{2h})+\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{mh})\mathbb{E}(Y_0Y_{(m+2)h}) \end{aligned}$$

$$+\sum_{m=1}^{\infty}\mathbb{E}(Y_0Y_{mh})\mathbb{E}(Y_0Y_{|m-2|h}). \tag{4.20}$$

Combining (4.13)–(4.20) yields

$$\lim_{n\to\infty}\mathbb{E}(G_n^2)=(\alpha,\beta,\gamma)\Sigma(\alpha,\beta,\gamma)^T. \tag{4.21}$$

Using Lemma A.3, we know that $J_n := \langle DG_n, DG_n\rangle_{\mathcal{H}}$ converges to a constant. Then by Proposition 2.1, we know G_n converges in law to a normal random variable.

Since G_n converges to a normal for any real vales α, β, and γ, we know by the Cramér-Wold theorem that F_n converges to a mean zero Gaussian random vector, proving theorem. □

Now using the delta method and the above theorem we immediately have the following theorem.

Theorem 4.5 *Let $(\theta, H, \sigma) \in \mathbb{D}_h$. Let X_t be the Ornstein–Uhlenbeck process defined by Eq. (1.1) and let $(\tilde{\theta}_n, \tilde{H}_n, \tilde{\sigma}_n)$ be the ergodic estimator defined by (3.10). Then as $n\to\infty$, we have*

$$\begin{pmatrix}\sqrt{n}(\tilde{\theta}_n-\theta)\\ \sqrt{n}(\tilde{H}_n-H)\\ \sqrt{n}(\tilde{\sigma}_n-\sigma)\end{pmatrix}\xrightarrow{d}N(0,\tilde{\Sigma})$$

,

where J denotes the Jacobian matrix of f, studied in Appendix B, Σ is defined in 4.3 and

$$\tilde{\Sigma}=[J(\theta,H,\sigma)]^{-1}\,\Sigma\left[J^T(\theta,H,\sigma)\right]^{-1}. \tag{4.22}$$

Acknowledgements We thank the referees for the constructive comments.

Appendix A: Detailed computations

First, we need the following lemma.

Lemma A.1 *Let X_t be the Ornstein–Uhlenbeck process defined by (1.1). Then*

$$|\mathbb{E}(X_tX_s)|\le C(1\wedge|t-s|^{2H-2})\le(1+|t-s|)^{2H-2}. \tag{A.1}$$

The above inequality also holds true for Y_t.

Proof From Cheridito et al. (2003, Theorem 2.3), we have that

$$\mathbb{E}(Y_sY_t)\le C_{H,\theta}|t-s|^{2H-2}\quad\text{for }|t-s|\text{sufficiently large}. \tag{A.2}$$

But $X_t = Y_t - e^{-\theta t}Y_0$. This combined with (A.2) proves (A.1). □

Lemma A.2 *Let X_t be defined by (1.1) and let a, b, c, d be integers. When $H\in(0,\frac{1}{2})\cup(\frac{1}{2},\frac{3}{4})$ we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k,k'=1}^{n}\mathbb{E}\left(X_{kh+ah}X_{k'h+bh}\right)\mathbb{E}\left(X_{kh+ch}X_{k'h+dh}\right)$$

$$= \mathbb{E}\left(Y_0 Y_{|b-a|}\right) \mathbb{E}\left(Y_0 Y_{|d-c|}\right) + \sum_{m=1}^{\infty} \mathbb{E}\left(Y_0 Y_{|m+b-a|}\right) \mathbb{E}\left(Y_0 Y_{|m+d-c|}\right)$$

$$+ \sum_{m=1}^{\infty} \mathbb{E}\left(Y_0 Y_{|m+a-b|}\right) \mathbb{E}\left(Y_0 Y_{|m+c-d|}\right). \tag{A.3}$$

Proof To simplify notations we shall use X_k, Y_k to represent X_{kh}, Y_{kh} etc. From the relation (3.3) it is easy to see that

$$\begin{aligned} \mathbb{E}(X_{k+a} X_{k'+b}) &= \mathbb{E}(Y_{k+a} Y_{k'+b}) - e^{-\theta(k'+b)h} \mathbb{E}(Y_0 Y_{k+a}) \\ &\quad - e^{-\theta(k+a)h} \mathbb{E}(Y_0 Y_{k'+b}) + e^{-\theta(k+k'+a+b)h} \mathbb{E}(Y_0^2) \\ &=: \sum_{i=1}^{4} I_{i,k,k'}, \end{aligned} \tag{A.4}$$

where $I_{i,k,k'} = I_{i,a,b,k,k'}$, $i = 1, \ldots, 4$, denote the above i-th term.

Let us consider $\frac{1}{n} \sum_{k,k'=1}^{n} I_{i,k,k'}^2$ for $i = 2, 3, 4$. First, we consider $i = 2$. By Cheridito et al. (2003, Theorem 2.3), we know that $\mathbb{E}(Y_0 Y_k)$ converges to 0 when $k \to \infty$. Thus by the Toeplitz theorem, we have

$$\begin{aligned} \frac{1}{n} \sum_{k,k'=1}^{n} I_{2,k,k'}^2 &= \frac{1}{n} \sum_{k,k'=1}^{n} e^{-2\theta(k'+b)h} \left[\mathbb{E}(Y_0 Y_{k+a})\right]^2 \\ &\le C \frac{1}{n} \sum_{k}^{n} \left[\mathbb{E}(Y_0 Y_{k+a})\right]^2 \to 0. \end{aligned} \tag{A.5}$$

Exactly in the same way we have

$$\frac{1}{n} \sum_{k,k'=1}^{n} I_{3,k,k'}^2 \to 0. \tag{A.6}$$

When $i = 4$, we have easily

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k,k'=1}^{n} I_{4,k,k'}^2 = \frac{1}{n} \sum_{k,k'=1}^{n} e^{-2\theta(k+k'+a+b)h} \left[\mathbb{E}(Y_0^2)\right]^2 \to 0. \tag{A.7}$$

Now we have

$$\begin{aligned} &\frac{1}{n} \sum_{k,k'=1}^{n} \mathbb{E}\left(X_{kh+ah} X_{k'h+bh}\right) \mathbb{E}\left(X_{kh+ch} X_{k'h+dh}\right) \\ &= \frac{1}{n} \sum_{i,j=1}^{4} \sum_{k,k'=1}^{n} I_{i,a,b,k,k'} I_{j,c,d,k,k'} \\ &= \frac{1}{n} \sum_{k,k'=1}^{n} I_{1,a,b,k,k'} I_{1,c,d,k,k'} + \frac{1}{n} \sum_{i\neq 1, \text{or } j\neq 1} \sum_{k,k'=1}^{n} I_{i,a,b,k,k'} I_{j,c,d,k,k'} \\ &= \mathcal{I}_{1,1,n} + \sum_{i\neq 1, \text{or } j\neq 1} \mathcal{I}_{i,j,n}. \end{aligned}$$

First, let us consider $\mathcal{I}_{1,1,n}$. By the stationarity of Y_n, we have

$$
\begin{aligned}
\mathcal{I}_{1,1,n} &= \frac{1}{n}\sum_{k,k'=1}^{n} \mathbb{E}(Y_{k+a}Y_{k'+b})\mathbb{E}(Y_{k+c}Y_{k'+d}) \\
&= \frac{1}{n}\sum_{k,k'=1}^{n} \mathbb{E}(Y_0 Y_{|k'-k+b-a|})\mathbb{E}(Y_0 Y_{|k'-k+d-c|}) \\
&= \mathbb{E}(Y_0 Y_{|b-a|})\mathbb{E}(Y_0 Y_{|d-c|}) + \frac{1}{n}\sum_{m=1}^{n-1}(n-m)\mathbb{E}(Y_0 Y_{|m+b-a|})\mathbb{E}(Y_0 Y_{|m+d-c|}) \\
&\quad + \frac{1}{n}\sum_{m=1}^{n-1}(n-m)\mathbb{E}(Y_0 Y_{|-m+b-a|})\mathbb{E}(Y_0 Y_{|-m+d-c|}) \\
&= \mathbb{E}(Y_0 Y_{|b-a|})\mathbb{E}(Y_0 Y_{|d-c|}) + \sum_{m=1}^{n-1}\mathbb{E}(Y_0 Y_{|m+b-a|})\mathbb{E}(Y_0 Y_{|m+d-c|}) \\
&\quad + \sum_{m=1}^{n-1}\mathbb{E}(Y_0 Y_{|m+a-b|})\mathbb{E}(Y_0 Y_{|m+a-b|}) + \frac{1}{n}\sum_{m=1}^{n-1} m\mathbb{E}(Y_0 Y_{|m+b-a|})\mathbb{E}(Y_0 Y_{|m+d-c|}) \\
&\quad + \frac{1}{n}\sum_{m=1}^{n-1} m\mathbb{E}(Y_0 Y_{|m+a-b|})\mathbb{E}(Y_0 Y_{|m+c-d|}).
\end{aligned} \tag{A.8}
$$

By Lemma A.1 for Y_t or an expression of $\mathbb{E}(Y_0 Y_m)$ given in Cheridito et al. (2003, Theorem 2.3):

$$
\mathbb{E}(Y_0 Y_m) = \frac{1}{2}\sigma^2 \sum_{n=1}^{N} \theta^{-2n}(\Pi_{k=0}^{2n-1}(2H-k))m^{2H-2n} + O(m^{2H-2N-2}).
$$

This means $\mathbb{E}(Y_0 Y_m) = O(m^{2H-2})$ as $m \to \infty$, which in turn means that $\left|\mathbb{E}(Y_0 Y_{|m+\rho_1|})\mathbb{E}(Y_0 Y_{|m+\rho_2|})\right| = O(m^{4H-4})$ for any arbitrarily given integers ρ_1 and ρ_2. Hence, when $H < \frac{3}{4}$, $\sum_{m=0}^{n-1}\mathbb{E}(Y_0 Y_{|m+\rho_1|})\mathbb{E}(Y_0 Y_{|m+\rho_2|})$ converges as n tends to infinity. This shows that the second and third terms in (A.8) are convergent.

Notice that for $H < \frac{3}{4}$, $m\mathbb{E}(Y_0 Y_m)^2 = O(m^{4H-3}) \to 0$ as $m \to \infty$. By Toeplitz theorem we have

$$
\frac{1}{n}\sum_{m=0}^{n-1} m\left|\mathbb{E}(Y_0 Y_{|m+\rho_1|})\mathbb{E}(Y_0 Y_{|m+\rho_2|})\right| \to 0 \quad \text{as } n \to \infty.
$$

Thus, the fourth and fifth terms in (A.8) converges to 0. This implies that $\mathcal{I}_{1.1.n}$ converges to the right hand side of (A.3).

When one of the i or j is not equal to 1, we have by the Hölder inequality

$$
\mathcal{I}_{i,j,n} \le \left(\frac{1}{n}\sum_{k,k'=1}^{n} I^2_{i,a,b,k,k'}\right)^{1/2}\left(\frac{1}{n}\sum_{k,k'=1}^{n} I^2_{j,c,d,k,k'}\right)^{1/2}
$$

which will go to 0 since $\frac{1}{n}\sum_{k,k'=1}^{n} I^2_{i,a,b,k,k'}$, $n = 1, 2, \ldots$ is bounded when $i = 1$ and converges to zero when $i \neq 1$ by (A.5)–(A.7). □

Let G_n be defined by (4.2) in Sect. 4. Its Malliavin derivative is given by

$$DG_n = \frac{1}{\sqrt{n}}2\alpha\sum_{k=1}^{n} X_k DX_k + \frac{1}{\sqrt{n}}\beta\sum_{k=1}^{n}(X_k DX_{k+1} + X_{k+1}DX_k)$$
$$+\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\gamma(X_k DX_{k+2} + X_{k+2}DX_k). \tag{A.9}$$

Lemma A.3 *Define the sequence of random variables $J_n := \langle DG_n, DG_n\rangle_{\mathcal{H}}$. Then*

$$\lim_{n\to\infty}\mathbb{E}\,[J_n - \mathbb{E}(J_n)]^2 = 0. \tag{A.10}$$

Proof It is easy to see that J_n is a linear combination of terms of the following forms (with the coefficients being a quadratic forms of α, β, γ):

$$\tilde{J}_n := \frac{1}{n}\sum_{k',k=1}^{n}\langle DX_{k_1}, DX_{k_1'}\rangle_{\mathcal{H}} X_{k_2}X_{k_2'}$$
$$= \frac{1}{n}\sum_{k',k=1}^{n}\mathbb{E}(X_{k_1}X_{k_1'})X_{k_2}X_{k_2'}, \tag{A.11}$$

where k_1, k_2 may take $k, k+1, k+2$, and k_1', k_2' may take $k', k'+1, k'+2$. For example, one term is to take $k_1 = k_2 = k$ and $k_1' = k'+1, k_2' = k'$ which corresponds to the product:

$$\left\langle \frac{1}{\sqrt{n}}2\alpha\sum_{k=1}^{n}X_k DX_k, \frac{1}{\sqrt{n}}\beta\sum_{k=1}^{n}(X_k DX_{k+1})\right\rangle$$
$$= \frac{2\alpha\beta}{n}\sum_{k',k=1}^{n}\mathbb{E}(X_k X_{k'+1})X_k X_{k'} =: 2\alpha\beta\tilde{J}_{0,n}. \tag{A.12}$$

We will first give a detail argument to explain why

$$\mathbb{E}\left[\tilde{J}_{0,n} - \mathbb{E}(\tilde{J}_{0,n})\right]^2 \to 0$$

and then we outline the procedure that similar claims hold true for any terms in (A.11). Note that $\mathbb{E}(\tilde{J}_{0,n})$ will not converge to 0.

From the Proposition 4.2 it follows

$$\mathbb{E}\left[\tilde{J}_{0,n} - \mathbb{E}(\tilde{J}_{0,n})\right]^2 = \frac{1}{n^2}\sum_{k,k',j,j'=1}^{n}\mathbb{E}(X_k X_{k'+1})\mathbb{E}(X_j X_{j'+1})\mathbb{E}(X_k X_j)\mathbb{E}(X_{k'}X_{j'})$$
$$+\frac{1}{n^2}\sum_{k,k',j,j'=1}^{n}\mathbb{E}(X_k X_{k'+1})\mathbb{E}(X_j X_{j'+1})\mathbb{E}(X_k X_{j'})\mathbb{E}(X_{k'}X_j)$$
$$=: I_{1,n} + I_{2,n}.$$

Using (A.1) we have

$$I_{1,n} \le \frac{1}{n^2}\sum_{k,k',j,j'=1}^{n}(1+|k'-k|)^{2H-2}(1+|j'-j|)^{2H-2}$$
$$(1+|j-k|)^{2H-2}(1+|k'-j'|)^{2H-2};$$

$$I_{2,n} \le \frac{1}{n^2} \sum_{k,k',j,j'=1}^{n} (1+|k'-k|)^{2H-2}(1+|j'-j|)^{2H-2} (1+|j'-k|)^{2H-2}(1+|k'-j|)^{2H-2}.$$

Now it is elementary to see that $I_{1,n} \to 0$ and $I_{2,n} \to 0$ when $n \to \infty$.

Now we deal with the general term

$$\tilde{J}_{1,n} := \frac{1}{n} \sum_{k',k=1}^{n} \mathbb{E}(X_{k_1} X_{k_1'}) X_{k_2} X_{k_2'}$$

in (A.11), where k_1, k_2 may take $k, k+1, k+2$, and k_1', k_2' may take $k', k'+1, k'+2$. We use Proposition 4.2 to obtain

$$\begin{aligned}
\mathbb{E}\left[\tilde{J}_{1,n} - \mathbb{E}(\tilde{J}_{1,n})\right]^2 &= \frac{1}{n^2} \sum_{k,k',j,j'=1}^{n} \mathbb{E}(X_{k_1} X_{k_1'})\mathbb{E}(X_{j_1} X_{j_1'})\mathbb{E}(X_{k_2} X_{j_2})\mathbb{E}(X_{k_2'} X_{j_2'}) \\
&\quad + \frac{1}{n^2} \sum_{k,k',j,j'=1}^{n} \mathbb{E}(X_{k_1} X_{k_1'})\mathbb{E}(X_{j_1} X_{j_1'})\mathbb{E}(X_{k_2} X_{j_2'})\mathbb{E}(X_{k_2'} X_{j_2}) \\
&=: \tilde{I}_{1,n} + \tilde{I}_{2,n},
\end{aligned}$$

where k_1, k_2 may take $k, k+1, k+2$, and k_1', k_2' may take $k', k'+1, k'+2$, j_1, j_2 may take $j, j+1, j+2$, and j_1', j_2' may take $j', j'+1, j'+2$. Using (A.1) we have

$$\begin{aligned}
\tilde{I}_{1,n} &\le \frac{1}{n^2} \sum_{k,k',j,j'=1}^{n} (1+|k'-k|)^{2H-2}(1+|j'-j|)^{2H-2} (1+|j-k|)^{2H-2}(1+|k'-j'|)^{2H-2}; \\
\tilde{I}_{2,n} &\le \frac{1}{n^2} \sum_{k,k',j,j'=1}^{n} (1+|k'-k|)^{2H-2}(1+|j'-j|)^{2H-2} (1+|j'-k|)^{2H-2}(1+|k'-j|)^{2H-2}.
\end{aligned}$$

Now it is elementary to see that $I_{1,n} \to 0$ and $I_{2,n} \to 0$ when $n \to \infty$. □

Appendix B: Determinant of the Jacobian of *f*

The goal of this section is to compute the determinant of the Jacobian of

$$f(\theta, H, \sigma) = \begin{cases} \frac{1}{\pi}\sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty \frac{x^{1-2H}}{\theta^2+x^2}dx\,; \\ \frac{1}{\pi}\sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(hx)\frac{x^{1-2H}}{\theta^2+x^2}dx\,; \\ \frac{1}{\pi}\sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(2hx)\frac{x^{1-2H}}{\theta^2+x^2}dx, \end{cases} \tag{B.1}$$

(we use the integral form of the first component of f to simplify the computation of the determinant).

The Jacobian matrix of f is equivalent (their determinants are up to a sign) to $J = (C_1, C_2, C_3)$, where the column vectors are given by

$$C_1 = \begin{pmatrix} 2\sigma\Gamma(2H+1)\sin(\pi H)\int_0^\infty \frac{x^{1-2H}}{\theta^2+x^2}dx \\ 2\sigma\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ 2\sigma\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(2hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \end{pmatrix};$$

$$C_2 = \begin{pmatrix} -2\theta\sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty \frac{x^{1-2H}}{(\theta^2+x^2)^2}dx \\ -2\theta\sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(hx)\frac{x^{1-2H}}{(\theta^2+x^2)^2}dx \\ -2\theta\sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(2hx)\frac{x^{1-2H}}{(\theta^2+x^2)^2}dx \end{pmatrix};$$

and $C_3 = C_{3,1} + C_{3,2} + C_{3,3}$, where

$$C_{3,1} = \begin{pmatrix} \sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty -2\log(x)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ \sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty -2\log(x)\cos(hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ \sigma^2\Gamma(2H+1)\sin(\pi H)\int_0^\infty -2\log(x)\cos(2hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \end{pmatrix};$$

$$C_{3,2} = \begin{pmatrix} \sigma^2\pi\Gamma(2H+1)\cos(\pi H)\int_0^\infty \frac{x^{1-2H}}{\theta^2+x^2}dx \\ \sigma^2\pi\Gamma(2H+1)\cos(\pi H)\int_0^\infty \cos(hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ \sigma^2\pi\Gamma(2H+1)\cos(\pi H)\int_0^\infty \cos(2hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \end{pmatrix};$$

and

$$C_{3,3} = \begin{pmatrix} \sigma^2\partial_H\Gamma(2H+1)\sin(\pi H)\int_0^\infty \frac{x^{1-2H}}{\theta^2+x^2}dx \\ \sigma^2\partial_H\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ \sigma^2\partial_H\Gamma(2H+1)\sin(\pi H)\int_0^\infty \cos(2hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \end{pmatrix}.$$

By the linearity of the determinant, we have

$$\det(J) = \det(C_1, C_2, C_{3,1}) + \det(C_1, C_2, C_{3,2}) + \det(C_1, C_2, C_{3,3})$$

It is easy to see that $\det(C_1, C_2, C_{3,2}) = \det(C_1, C_2, C_{3,3}) = 0$ (C_1 is a proportional to $C_{3,2}$ and to $C_{3,3}$). Therefore

$$\det(J) = \det(C_1, C_2, C_{3,1}). \tag{B.2}$$

Notice that

$$\det(C_1, C_2, C_{3,1}) = -4\theta\sigma^5\Gamma^3(2H+1)\sin^3(\pi H)\det(M), \tag{B.3}$$

where

$$M = \begin{pmatrix} \int_0^\infty \frac{x^{1-2H}}{(\theta^2+x^2)}dx & \int_0^\infty \frac{x^{1-2H}}{(\theta^2+x^2)^2}dx & \int_0^\infty -2\log(x)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ \int_0^\infty \cos(hx)\frac{x^{1-2H}}{(\theta^2+x^2)}dx & \int_0^\infty \cos(hx)\frac{x^{1-2H}}{(\theta^2+x^2)^2}dx & \int_0^\infty -2\log(x)\cos(hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \\ \int_0^\infty \cos(2hx)\frac{x^{1-2H}}{(\theta^2+x^2)}dx & \int_0^\infty \cos(2hx)\frac{x^{1-2H}}{(\theta^2+x^2)^2}dx & \int_0^\infty -2\log(x)\cos(2hx)\frac{x^{1-2H}}{\theta^2+x^2}dx \end{pmatrix}$$

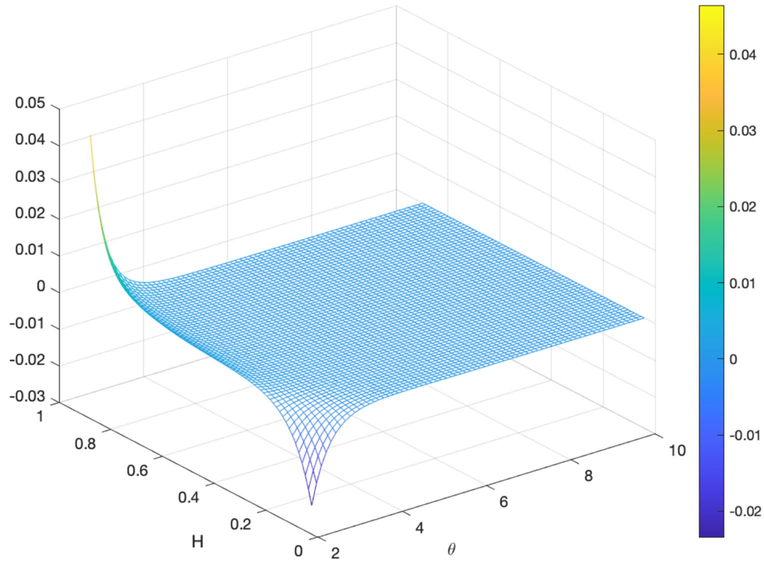

Fig. 2 Determinant of M for $H \in (0, 1)$ and $\theta \in (2, 10)$

Since $\theta > 0$, $\sigma > 0$, $\sin(\pi H) > 0$ and $\Gamma(2H + 1) > 0$ (for $H \in (0, 1)$), $\det(J) = 0$ if and only if $\det(M) = 0$.

The determinant $\det(J)$ or the determinant $\det(M)$ depends also on h. To remove this dependence, we write $M = (M_{ij})_{1\le i,j\le 3}$, where

$$M_{11} = \int_0^\infty h^{2H} \frac{x^{1-2H}}{(h^2\theta^2 + x^2)} dx, \qquad M_{12} = \int_0^\infty h^{2H+2} \frac{x^{1-2H}}{(h^2\theta^2 + x^2)^2} dx$$

$$M_{13} = \int_0^\infty -2h^{2H} \log(\frac{x}{h}) \frac{x^{1-2H}}{h^2\theta^2 + x^2} dx, \qquad M_{21} = \int_0^\infty h^{2H} \cos(x) \frac{x^{1-2H}}{(h^2\theta^2 + x^2)} dx$$

$$M_{22} = \int_0^\infty h^{2H+2} \cos(x) \frac{x^{1-2H}}{(h^2\theta^2 + x^2)^2} dx,$$

$$M_{23} = \int_0^\infty -2h^{2H} \log(\frac{x}{h}) \cos(x) \frac{x^{1-2H}}{h^2\theta^2 + x^2} dx$$

$$M_{31} = \int_0^\infty h^{2H} \cos(2x) \frac{x^{1-2H}}{(h^2\theta^2 + x^2)} dx,$$

$$M_{32} = \int_0^\infty h^{2H+2} \cos(2hx) \frac{x^{1-2H}}{(h^2\theta^2 + x^2)^2} dx$$

$$M_{33} = \int_0^\infty -2h^{2H} \log(\frac{x}{h}) \cos(2x) \frac{x^{1-2H}}{h^2\theta^2 + x^2} dx$$

Since $\log(\frac{x}{h}) = \log(x) - \log(h)$, the determinant of M is equal to h^{6H+2} multiply the determinant of the following matrix:

$$N = \begin{pmatrix} \int_0^\infty \frac{x^{1-2H}}{(h^2\theta^2+x^2)} dx & \int_0^\infty \frac{x^{1-2H}}{(h^2\theta^2+x^2)^2} dx & \int_0^\infty -2\log(x) \frac{x^{1-2H}}{h^2\theta^2+x^2} dx \\ \int_0^\infty \cos(x) \frac{x^{1-2H}}{(h^2\theta^2+x^2)} dx & \int_0^\infty \cos(x) \frac{x^{1-2H}}{(h^2\theta^2+x^2)^2} dx & \int_0^\infty -2\log(x)\cos(x) \frac{x^{1-2H}}{h^2\theta^2+x^2} dx \\ \int_0^\infty \cos(2x) \frac{x^{1-2H}}{(h^2\theta^2+x^2)} dx & \int_0^\infty \cos(2x) \frac{x^{1-2H}}{(h^2\theta^2+x^2)^2} dx & \int_0^\infty -2\log(x)\cos(2x) \frac{x^{1-2H}}{h^2\theta^2+x^2} dx \end{pmatrix}$$

Namely, the determinant $\det(J)$ is a negative number multiplied by the determinant $\det(N)$. Denote $\theta' = h\theta$. The determinant of N a function of two variables only: θ' and H. The plot in Fig. 2 shows that $\det(N)$ is positive for $H \in (0.03, 1)$ and $\theta' \in (2, 10)$. Combining this with (B.2) and (B.3), we see that on

$$\mathbb{D}_h = \{H > 0.03, 2 < \theta h < 10, \sigma > 0\} \tag{B.4}$$

$\det(J)$ is strictly negative hence is not singular.

Appendix C: Numerical results

For all the experiments, we take $h = 1$.

C.1. Strong consistency of the estimators

In this subsection, we illustrate the almost-sure convergence by plotting different trajectories of the estimators. We observe that when $\log_2(n) \geq 14$, the estimators become very close to the true parameter.

However, since our estimators are random (they depend on the sample $\{X_{kh}\}_{k=1}^{n}$), what's important to see in these figures is the deviations from the true parameter we are estimating. Even if three trajectories are not enough to make statements about the variance, the figures predict that the variance of $\tilde{\theta}_n$ is very high compared to the other estimators (see Figs. 3, 4) and that, for H close to 0 (see Fig. 5), the deviations of $\tilde{H}_n$ increase.

C.2. Mean and standard deviation/Asymptotic behavior of the estimators

It is important to check the mean and deviation of our estimators. For example, a large variance implies a large deviation and therefore a "weak" estimator. That is why we plotted the mean and variance of our estimators for $n = 2^{12}$ over 100 samples.

As we observe, the standard deviation (s.d) of $\tilde{\theta}_n$ is larger than the s.d of $\tilde{\sigma}_n$ which is larger than the s.d of $\tilde{H}_n$ (see Tables 1, 2). Notice also that the s.d of $\tilde{H}_n$ increases as H decreases.

In Hu and Song (2013), the variance of the θ estimator is proportional to θ^2. In our case, it is difficult to compute the variances of our estimators (they depend on the matrix Σ (see Theorem 4.3) and the Jacobian of the function f (see Eq. (3.9)), however we should probably expect something similar which could explain the gap in the variances since the values of θ are usually bigger that the values taken by σ or H.

Having access to 100 estimates of each parameter, we are also able to plot the distributions of our estimators to show that they effectively have a Gaussian nature (4.5) (Figs. 6, 7, 8).

Remark C.1 In practice, one may already know the value of one parameter, σ for example. In this case, it is important to point out that the estimators perform a lot better. For example, in Fig. 9, we plot the density of θ_n and H_n for $\sigma = 1$, $H = 0.6$, $\theta = 6$ and for $\log_2(n)$. Observe how the variance of the estimators is a lot smaller and the shape of the density is smoother.

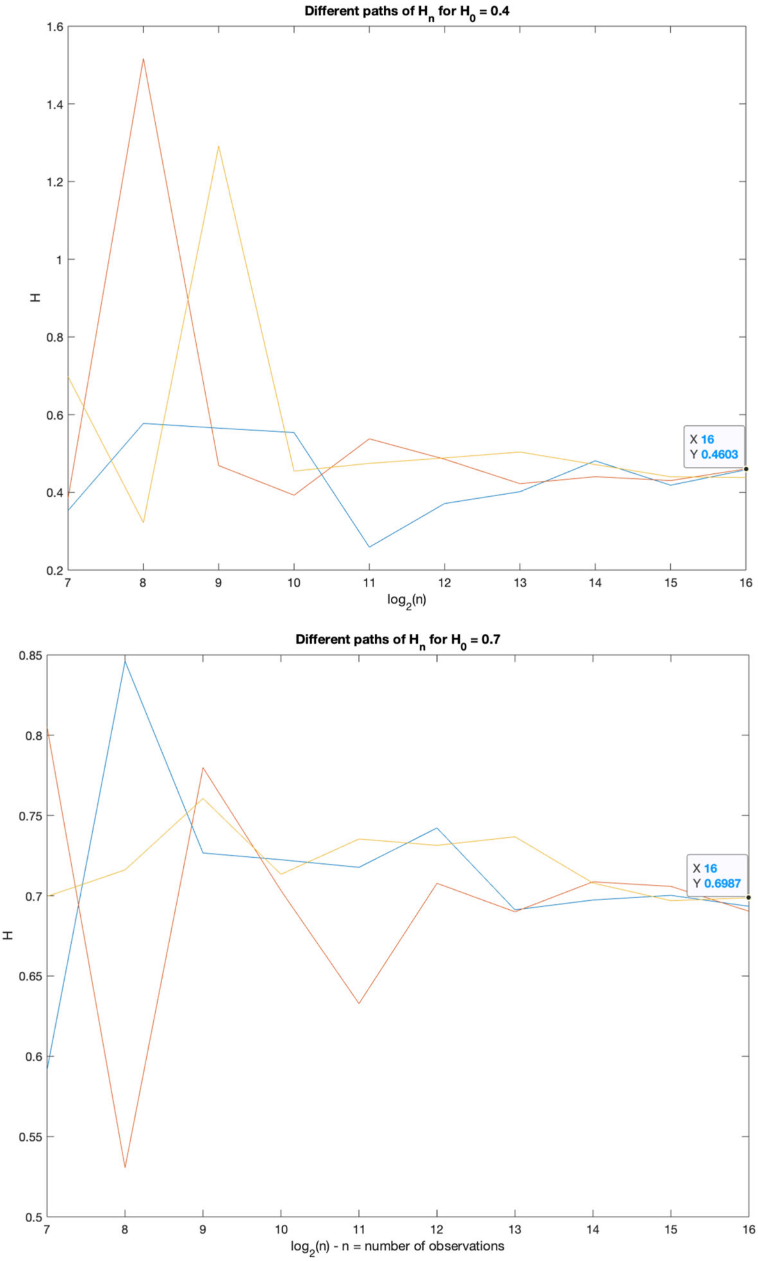

Fig. 3 Convergence of $\widetilde{H}_n$ for $H = 0.7$ and $H = 0.4$ ($\theta = 6, \sigma = 2$)

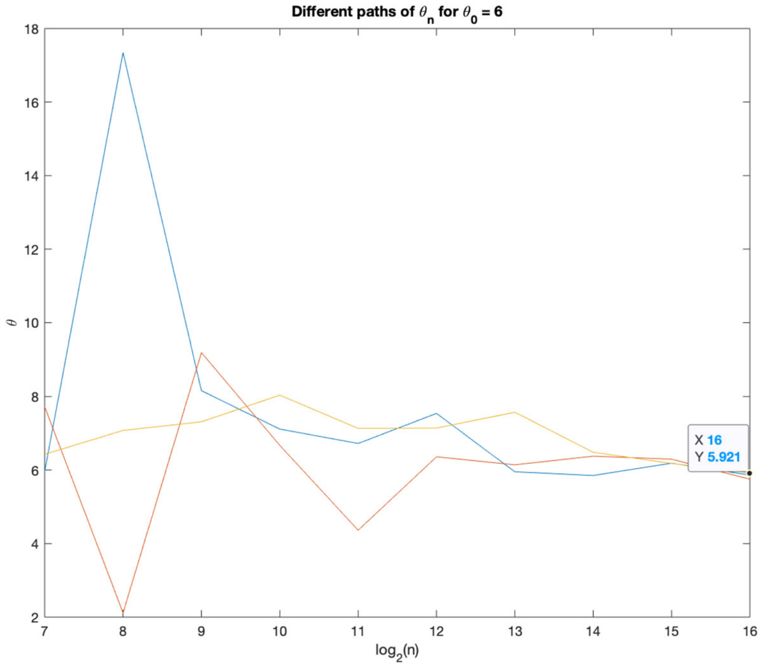

Fig. 4 Convergence of $\widetilde{\theta_n}$ for $\theta = 6$, $H = 0.7$, $\sigma = 2$

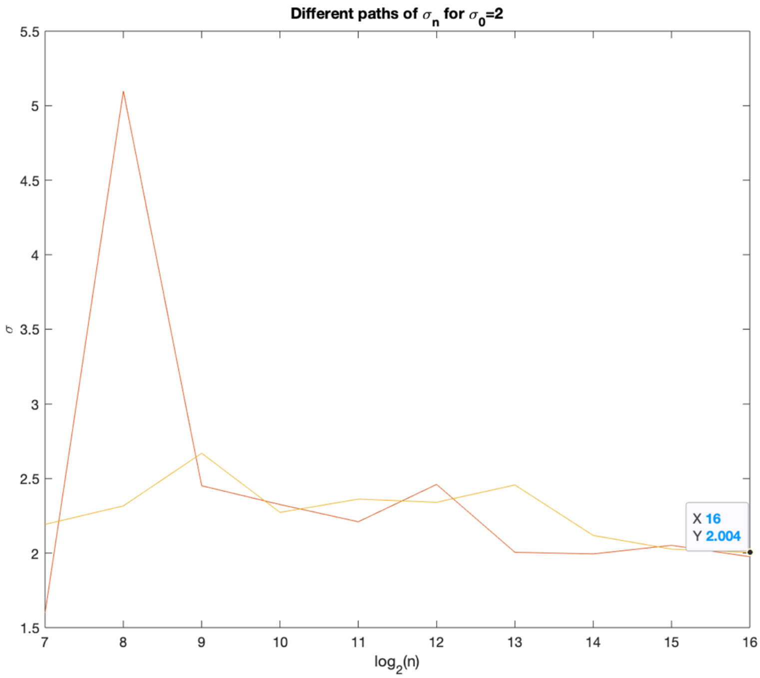

Fig. 5 Convergence of $\widetilde{\sigma_n}$ for $\theta = 6$, $H = 0.7$, $\sigma = 2$

Table 1 $H = 0.7, \theta = 6$ and $\sigma = 2$

	Mean	Standard deviation
$\widetilde{H_n}$	0.704	0.0221
$\widetilde{\theta_n}$	6.2983	0.8288
$\widetilde{\sigma_n}$	2.0921	0.2117

Table 2 $H = 0.4, \theta = 6$ and $\sigma = 2$

	Mean	Standard deviation
$\widetilde{H_n}$	0.4392	0.0531
$\widetilde{\theta_n}$	6.832	1.3227
$\widetilde{\sigma_n}$	2.4785	0.3833

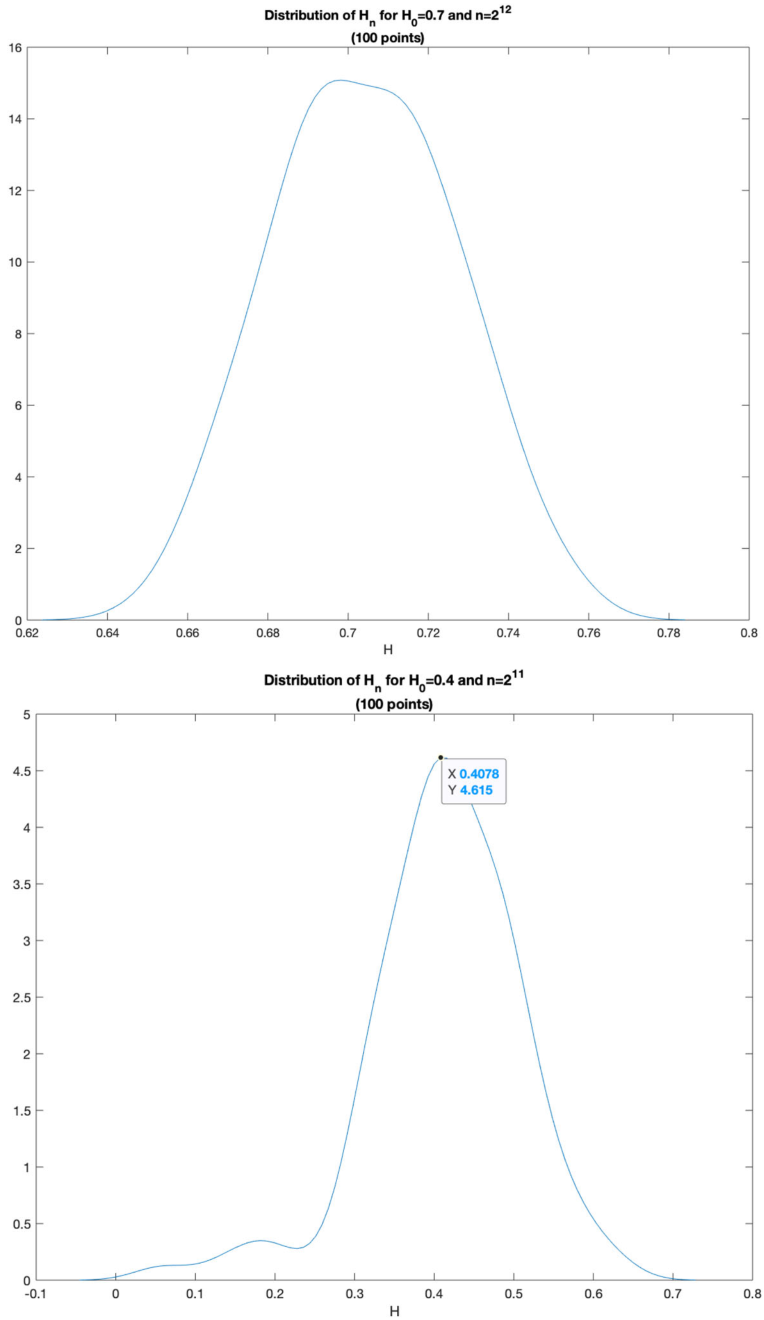

Fig. 6 Distribution of $\widetilde{H_n}$ for $H = 0.7$ and $H = 0.4$ while $\theta = 6$, $\sigma = 2$

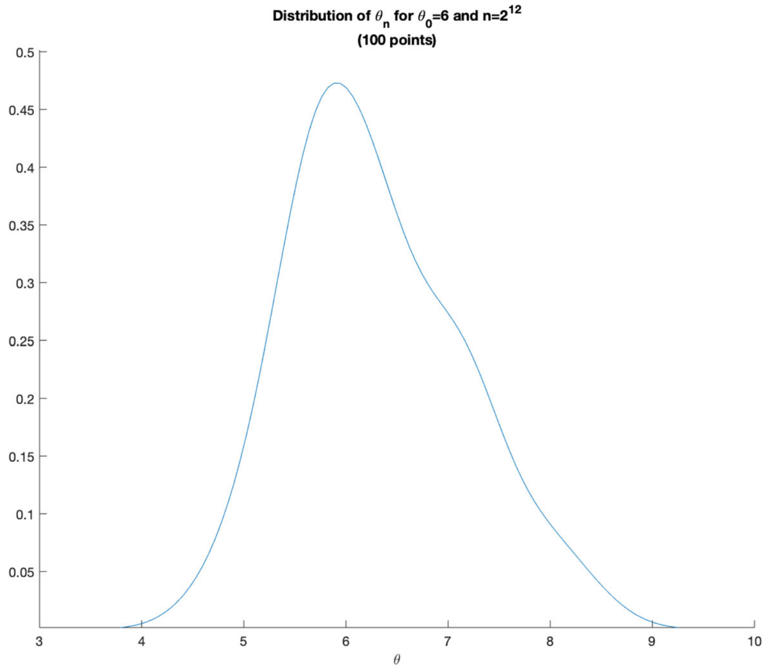

Fig. 7 Distribution of $\widetilde{\theta_n}$ for $\theta = 6$, $H = 0.7$, $\sigma = 2$

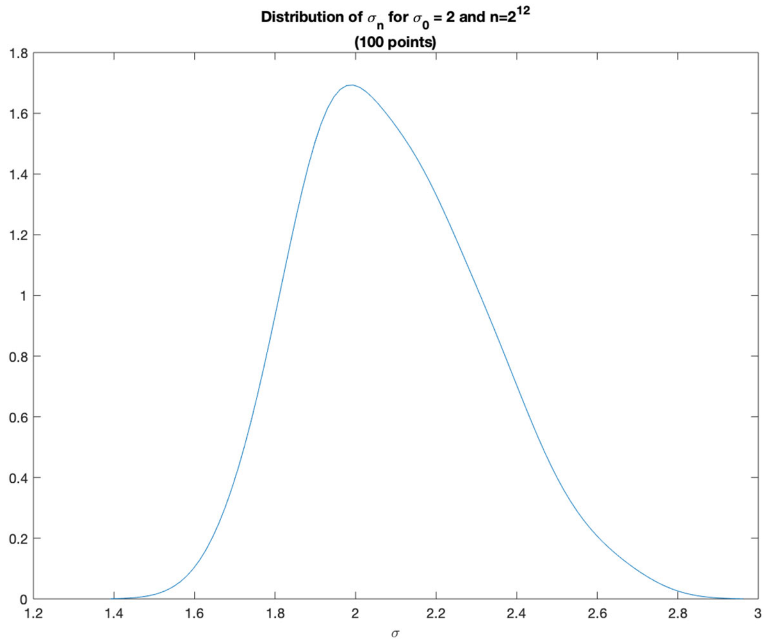

Fig. 8 Distribution of $\widetilde{\sigma_n}$ for $\theta = 6$, $H = 0.7$, $\sigma = 2$

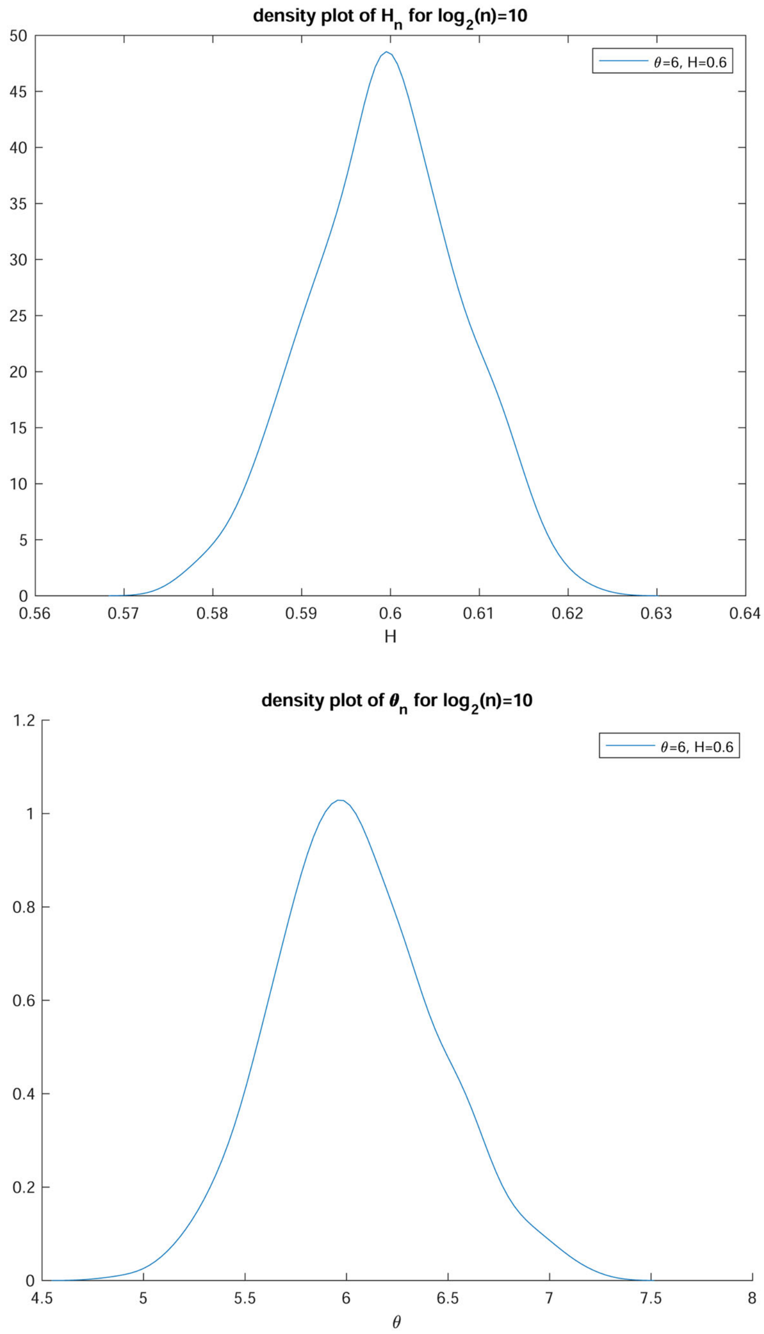

Fig. 9 Density plots of θ_n and H_n when σ is known $(= 1)$

References

Biagini F, Hu Y, Ø ksendal B, Zhang T (2008) Stochastic calculus for fractional Brownian motion and applications. Probability and its applications. Springer, New York

Brouste A, Iacus SM (2013) Parameter estimation for the discretely observed fractional Ornstein–Uhlenbeck process and the Yuima R package. Comput Stat 28(4):1529–1547

Chen Y, Hu Y, Wang Z (2017) Parameter estimation of complex fractional Ornstein–Uhlenbeck processes with fractional noise. ALEA Lat Am J Probab Math Stat 14(1):613–629

Cheng Y, Hu Y, Long H (2020) Generalized moment estimation for Ornstein-Uhlenbeck processes driven by α-stable lévy motions from discrete time observations. Stat Inference Stoch Process 23(1):53–81

Cheridito P, Kawaguchi H, Maejima M (2003) Fractional Ornstein-Uhlenbeck processes. Electron J Probab 8(3):1–14

Hu Y (2017) Analysis on Gaussian spaces. World Scientific Publishing Co., Pte. Ltd, Hackensack

Hu Y, Nualart D (2010) Parameter estimation for fractional Ornstein–Uhlenbeck processes. Stat Probab Lett 80(11–12):1030–1038

Hu Y, Song J (2013) Parameter estimation for fractional Ornstein–Uhlenbeck processes with discrete observations. In: Viens F, Feng J, Hu Y, Nualart E (eds) Malliavin calculus and stochastic analysis. Springer proceedings in mathematics & statistics, vol 34. Springer, Boston, pp 427–442

Hu Y, Nualart D, Zhou H (2019) Parameter estimation for fractional Ornstein–Uhlenbeck processes of general Hurst parameter. Stat Inference Stoch Process 22(1):111–142

Kubilius K, Mishura IS, Ralchenko K (2017) Parameter estimation in fractional diffusion models, vol 8. Springer, Berlin

Magdziarz M, Weron A (2011) Ergodic properties of anomalous diffusion processes. Ann Phys 326(9):2431–2443

Mustafa OG, Rogovchenko-Yuri V (2007) Estimates for domains of local invertibility of diffeomorphisms. Proc Am Math Soc 135(1):69–75

Panloup F, Tindel S, Varvenne M (2019) A general drift estimation procedure for stochastic differential equations with additive fractional noise. arXiv:1903.10769

Tudor CA, Viens FG et al (2007) Statistical aspects of the fractional stochastic calculus. Ann Stat 35(3):1183–1212

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