Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 129 (2019) 3018–3054

stochastic processes and their applications

www.elsevier.com/locate/spa

Optimal rates for parameter estimation of stationary Gaussian processes

Khalifa Es-Sebaiy[a], Frederi G. Viens[b,*]

[a] Department of Mathematics, Faculty of Science, Kuwait University, Kuwait

[b] Department Statistics and Probability, Michigan State University, 619 Red Cedar Rd., East Lansing, MI 48824, USA

Received 13 April 2017; received in revised form 11 July 2018; accepted 22 August 2018

Available online 21 September 2018

Abstract

We study rates of convergence in central limit theorems for partial sums of polynomial functionals of general stationary and asymptotically stationary Gaussian sequences, using tools from analysis on Wiener space. In the quadratic case, thanks to newly developed optimal tools, we derive sharp results, i.e. upper and lower bounds of the same order, where the convergence rates are given explicitly in the Wasserstein distance via an analysis of the functionals' absolute third moments. These results are tailored to the question of parameter estimation, which introduces a need to control variance convergence rates. We apply our result to study drift parameter estimation problems for some stochastic differential equations driven by fractional Brownian motion with fixed-time-step observations.

MSC: 60F05; 60G15; 62F12; 62M09.

Keywords: Central limit theorem; Berry–Esséen; Stationary Gaussian processes; Nourdin–Peccati analysis; Parameter estimation; Fractional Brownian motion

1. Introduction

While statistical inference for Itô-type diffusions has a long history, statistical estimation for equations driven by fractional Brownian motion (fBm) is much more recent, partly because the development of stochastic calculus with respect to fBm, which provides tools to study such

* Corresponding author.

E-mail addresses: khalifa.essebaiy@ku.edu.kw (K. Es-Sebaiy), viens@msu.edu (F.G. Viens).

https://doi.org/10.1016/j.spa.2018.08.010

models, is itself a recent and ongoing endeavor, and partly because these tools can themselves be unwieldy in comparison with the convenience and power of martingale methods and the Markov property which accompany Itô models. Our purpose in this article is to show how the analysis on Wiener space, particularly via tools recently developed to study the convergence-in-law properties in Wiener chaos, can be brought to bear on parameter estimation questions for fBm-driven models, and more generally for arbitrary stationary Gaussian models.

There are several approaches to estimating drift parameters in fBm-driven models, which have been developed over the course of the past 15 years. The approaches we mention below are related to the methods in this article.

- The MLE approach in [18,30]. In general the techniques used to construct maximum likelihood estimators (MLE) for drift parameters are based on Girsanov transforms for fBm and depend on the properties of the deterministic fractional operators (determined by the Hurst parameter) related to the fBm. Typically, the MLE is not easily computable. In particular, it relies on being able to compute stochastic integrals with respect to fBm. This is difficult or hopeless for most models since approximating pathwise integrals w.r.t. fBm, when they exist, is challenging, while Skorohod-type integrals cannot be computed based on the data except in special cases. The work in [30] is the only one in which a strongly consistent discretization of the MLE was based on long-horizon asymptotics without also requiring an in-fill (small time step) condition, though it did not establish any asymptotic distribution.
- A least-squares (LS) approach was proposed in [15]. The study of the asymptotic properties of the estimator is based on certain criteria formulated in terms of the Malliavin calculus (see [23]). It should be noted that in [15], the full LS estimator relies on an unobservable Skorohod integral, and the authors proposed a modified version of this estimator which can be computed based on in-fill asymptotics; however, this modified estimator bears no immediate relation to an LS one (see [14] for examples of what constitutes a discretization of an LS estimator for fBm models, and for a comparison with MLE methods, which coincide with LS methods if and only if $H = 1/2$). In the ergodic case, the statistical inference for several fractional Ornstein–Uhlenbeck (fOU) models via LS methods was recently developed in the papers [15,1,2,14,16,7,21,8]. The case of non-ergodic fOU process of the first kind and of the second kind can be found in [3,12] and [13] respectively.

We bring new techniques to statistical inference for stochastic differential equations (SDEs) related to stationary Gaussian processes. Some of these ideas can be summarized as follows:

- Since the theory of inference for these fBm-driven SDEs is still near its inception, and most authors are concerned with linear problems, whose solutions are Gaussian, this Gaussian property should be exploited to its fullest extent, given the best tools currently available.
  - Therefore we choose to consider polynomial variations of these processes, which then necessarily live in a finite sum of Wiener chaoses, whose properties are now well understood thanks to new Malliavin-calculus advances which were initiated by Nourdin and Peccati in 2008; in particular, we rely a general observation and their so-called optimal 4th moment theorem, in [24].
  - As a consequence, we are able to compute upper bounds in the total variation (TV) norm for the rate of normal convergence of our estimator. In particular, for the quadratic case, we prove a Berry–Esséen theorem (speed on the order of $1/\sqrt{n}$) for

this TV norm which we show is sharp in some cases by finding a lower bound with the same speed. No authors as far as we know have ever provided such quantitative estimates of the speed of asymptotic normality for any drift estimators for any fBm-driven model, let alone shown that they are sharp.

- Rather than starting from the continuous-time setting of SDEs, and then attempt to discretize resulting LS estimators, as was done in many of the aforementioned works including our own [14], we work from discretely observed data from the continuous-time SDEs, and design estimators based on such Gaussian sequences. In fact, we show that one can develop estimators valid for any Gaussian sequence, with suitable conditions on the sequence's auto-correlation function, and then apply them to fBm-driven SDEs of interest. In this way, we are able to provide estimators for many other models, while the models studied in [16,1,2,14] become particular cases in our approach.
- Since our method relies on conditions which need only be checked intrinsically on the auto-correlation function, it can apply equally well to in-fill situations and increasing-horizon situations.
  - It turns out that, as an artefact of trying to discretize estimators based on continuous paths, prior works were never able to avoid an in-fill assumption on the data (and sometimes even required both in-fill and increasing-horizon assumptions). In this paper, we illustrate our methods by showing that in-fill assumptions are never needed for the examples we cover.
  - Essentially, as explained in more detail further below, if a Gaussian stochastic process has a memory correlation length which is bounded above by that of a fBm with Hurst parameter $H < 5/8$, then our polynomial variations estimator based on discrete data (fixed time step) is asymptotically normal as the number of observations n increases, with a TV speed as good as $1/\sqrt{n}$, as mentioned above. In the case of quadratic variations, we can say more: the TV speed remains $1/\sqrt{n}$ up to $H < 2/3$, and then slows down as H ranges from $2/3$ to $3/4$ (in fact, the power of n in this speed interpolates linearly from $-1/2$ to 0: for $H = 3/4$, one obtains a logarithmic speed), and these rates are optimal. In the general case, the rates are not known to be optimal, but our results do also imply power rates for H between $5/8$ and $3/4$.
- Finally, we provide a systematic study of how to go from stationary observations, to observations coming from a Gaussian process which may not be stationary, by implementing a fully quantitative strategy to control the contribution of the non-stationarity to the TV convergence speeds. In the examples we cover, which are those of recent interest in the literature, the non-stationarity term vanishes exponentially fast, which is more than enough for our generic condition to hold, but slower power convergences would yield the same results, for summable powers.

Our article is structured as follows. Section 2 provides some basic elements of analysis on Wiener space which are helpful for some of the arguments we use. Section 3 provides the general theory of polynomial variation for general Gaussian sequences, covering the stationary case (Section 3.4, with examples in Section 3.5), non-stationary cases (Section 4), which include optimality in the quadratic case even under non-stationarity (Section 4.2) and a strategy of how to access a specific parameter other than the polynomial's variance (Section 4.3). Finally, three sets of examples based on fractional Ornstein–Uhlenbeck constructions are given in Sections 5 and 6. Some of the technical results used in various proofs, including the proof of the basic Berry–Esséen theorem in the stationary case, are in the Appendix.

The authors thank I. Nourdin for helpful discussions which resulted in the material herein at the ends of Sections 4.1 and 4.2, and in Corollary 17.

2. Elements of analysis on Wiener space

Here we summarize a few essential facts from the analysis on Wiener space and the Malliavin calculus. Though these facts and notation are essential underpinnings of the tools and results of this paper, most of our results and arguments can be understood without knowledge of the elements in this section. The interested reader can find more details in [27, Chapter 1] and [23, Chapter 2].

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a standard Wiener space, its standard Wiener process W, where for a deterministic function $h \in L^2(\mathbf{R}_+) =: \mathcal{H}$, the Wiener integral $\int_{\mathbf{R}_+} h(s)\, dW(s)$ is also denoted by $W(h)$. The inner product $\int_{\mathbf{R}_+} f(s)\, g(s)\, ds$ will be denoted by $\langle f, g\rangle_{\mathcal{H}}$. For every $q \geqslant 1$, let $\mathcal{H}_q$ be the qth Wiener chaos of W, that is, the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_q(W(h)), h \in \mathcal{H}, \|h\|_{\mathcal{H}} = 1\}$ where H_q is the qth Hermite polynomial. The mapping $I_q(h^{\otimes q}) := q! H_q(W(h))$ provides a linear isometry between the symmetric tensor product $\mathcal{H}^{\odot q}$ (equipped with the modified norm $\|.\|_{\mathcal{H}^{\odot q}} = \frac{1}{\sqrt{q!}}\|.\|_{\mathcal{H}^{\otimes q}}$) and $\mathcal{H}_q$. It also turns out that $I_q(h^{\otimes q})$ is the multiple Wiener integral of $h^{\otimes q}$ w.r.t. W. For every $f, g \in \mathcal{H}^{\odot q}$ the following product formula holds

$$E\left(I_q(f) I_q(g)\right) = q! \langle f, g\rangle_{\mathcal{H}^{\otimes q}}.$$

For $h \in \mathcal{H}^{\otimes q}$, the multiple Wiener integrals $I_q(h)$, which exhaust the set $\mathcal{H}_q$, satisfy a hypercontractivity property (equivalence in $\mathcal{H}_q$ of all L^p norms for all $p \geqslant 2$), which implies that for any $F \in \oplus_{l=1}^{q} \mathcal{H}_l$, we have

$$\left(E\left[|F|^p\right]\right)^{1/p} \leqslant c_{p,q} \left(E\left[|F|^2\right]\right)^{1/2} \text{ for any } p \geqslant 2. \quad (1)$$

Though we will not insist on their use in the main body of the paper, leaving associated technicalities to the proof of one of our main theorems in the Appendix, the Malliavin derivative operator D on $L^2(\Omega)$ plays a fundamental role in evaluating distances between random variables therein. For any function $\Phi \in C^1(\mathbf{R})$ with bounded derivative, and any $h \in \mathcal{H}$, we define the Malliavin derivative of the random variable $X := \Phi(W(h))$ to be consistent with the following chain rule:

$$DX : X \mapsto D_r X := \Phi'(W(h))\, h(r) \in L^2(\Omega \times \mathbf{R}_+).$$

A similar chain rule holds for multivariate Φ. One then extends D to the so-called Gross–Sobolev subset $\mathbf{D}^{1,2} \subsetneqq L^2(\Omega)$ by closing D inside $L^2(\Omega)$ under the norm defined by

$$\|X\|_{1,2}^2 = E\left[X^2\right] + E\left[\int_{\mathbf{R}_+} |D_r X|^2 dr\right].$$

Now recall that, if X, Y are two real-valued random variables, then the total variation distance between the law of X and the law of Y is given by

$$d_{TV}(X, Y) = \sup_{A \in \mathcal{B}(\mathbb{R})} |P[X \in A] - P[Y \in A]|.$$

If X, Y are two real-valued integrable random variables, then the Wasserstein distance between the law of X and the law of Y is given by

$$d_W(X,Y) = \sup_{f\in Lip(1)} |Ef(X) - Ef(Y)|$$

where $Lip(1)$ indicates the collection of all Lipschitz functions with Lipschitz constant $\leqslant 1$. Let N denote the standard normal law. All Wiener chaos random variable are in the domain $\mathbf{D}^{1,2}$ of D, and are orthogonal in $L^2(\Omega)$. The so-called Wiener chaos expansion is the fact that any $X \in \mathbf{D}^{1,2}$ can be written as $X = EX + \sum_{q=1}^{\infty} X_q$ where $X_q \in \mathcal{H}_q$. We define a linear operator L which is diagonalizable under the $\mathcal{H}_q$'s by saying that $\mathcal{H}_q$ is the eigenspace of L with eigenvalue $-q$, i.e. for any $X \in \mathcal{H}_q$, $LX = -qX$. The kernel of L is the constants. The operator $-L^{-1}$ is the negative pseudo-inverse of L, so that for any $X \in \mathcal{H}_q$, $-L^{-1}X = q^{-1}X$. Since the variables we will be dealing with in this article are finite sums of elements of $\mathcal{H}_q$, the operator $-L^{-1}$ is easy to manipulate thereon.

Two key estimates linking total variation distance and the Malliavin calculus are the following.

- Let $X \in \mathbf{D}^{1,2}$ with $E[X] = 0$. Then (see [24, Proposition 2.4]),

$$d_{TV}(X,N) \leqslant 2E\left|1 - \left\langle DX, -DL^{-1}X\right\rangle_{\mathcal{H}}\right|. \tag{2}$$

- Let $(X_n)_{n\geqslant 1}$ be a sequence of random variables in a fixed Wiener chaos of order $q \geqslant 2$ such that $Var[X_n] = 1$, and assume X_n converges to a normal law in distribution, which is equivalent to $\lim_n E\left[X_n^4\right] = 3$ (this equivalence, proved originally in [28], is known as the *fourth moment theorem*). Then we have the following optimal estimate for $d_{TV}(X_n, N)$, known as the optimal 4th moment theorem, proved in [24]: there exist two constants $c, C > 0$ depending only on the sequence $(X_n)_{n\geqslant 1}$ but not on n, such that

$$c\max\left\{E\left[X_n^4\right]-3, \left|E\left[X_n^3\right]\right|\right\} \leqslant d_{TV}(X_n,N) \leqslant C\max\left\{E\left[X_n^4\right]-3, \left|E\left[X_n^3\right]\right|\right\}.$$

We recall that the fourth cumulant of a standardized random variable X is $\kappa_4(X) := E\left[X^4\right] - 3$.

3. Parameter estimation for stationary Gaussian processes

3.1. Berry–Esséen bound of finite sum of multiple integrals

We work with a centered stationary Gaussian process (sequence) $Z = (Z_k)_{k\in\mathbb{Z}}$ with covariance

$$r_Z(k) := E(Z_0 Z_k) \text{ for every } k \in \mathbb{Z} \text{ such that } r_Z(0) > 0.$$

For any centered Gaussian sequence Z indexed by $\mathbb{Z}$, stationary or not, it is always possible to represent the entirely family of Z_n's jointly as Wiener integrals using a corresponding family of functions $\varepsilon_n \in \mathcal{H}$ as

$$Z_n/\sqrt{r_Z(0)} = I_1(\varepsilon_n)$$

in the notation of Section 2. In all that follows, we will use this representation.

Theorem 1. *Let* $F = \sum_{k=2}^{q} I_k(g_k)$ *with* $q \geqslant 2$ *is a positive integer and* $g_k \in \mathcal{H}^{\odot k}$, $k = 2, \ldots, q$. *Denote* $N \sim \mathcal{N}(0, 1)$. *Then there exists a constant* C_q *depending only on* q *such that*

$$d_{TV}\left(\frac{F}{\sqrt{E\left[F^2\right]}}, N\right) \leqslant \frac{C_q}{E\left[F^2\right]} \max_{2\leqslant k\leqslant q} \|g_k\|_{\mathcal{H}^{\otimes k}} \max_{\substack{1\leqslant s\leqslant k\\ 2\leqslant k\leqslant q}} \|g_k \otimes_s g_k\|_{\mathcal{H}^{\otimes 2k-2s}}^{\frac{1}{2}}. \tag{3}$$

In addition, if $g_k = \frac{d_k}{\sqrt{n}} \sum_{i=0}^{n-1} \varepsilon_i^{\otimes k}$ *where* ε_i *are given above and* d_k *are constants, and* $\sigma > 0$, *then*

$$d_{TV}\left(\frac{F}{\sigma}, N\right) \leqslant \frac{C_q}{\sigma^2}\sqrt{Var(I_2(g_2))\sqrt{\kappa_4(I_2(g_2))}} + 2E\left|1 - \frac{E[F^2]}{\sigma^2}\right|, \tag{4}$$

where C_q *is a constant which depends only on* q *and* d_k, $k = 2, \ldots, q$. *Moreover,*

$$\kappa_4(I_2(g_2)) = \mathcal{O}\left(\frac{\left(\sum_{|j|<n} |r_Z(j)|^{4/3}\right)^3}{n}\right). \tag{5}$$

Proof. See Appendix. ■

3.2. *Notation and basic question*

Fix a polynomial function f_q of degree q where $q \geqslant 2$ is an even integer such that $f_q(x) = \widetilde{f}_q(x^2)$, where $\widetilde{f}_q$ is a polynomial. Since polynomials of x^2 can be expressed in the sub-basis of even-rank Hermite polynomials, f_q possesses the following decomposition

$$f_q(x) := \sum_{k=0}^{q/2} d_{f_q,2k} H_{2k}\left(\frac{x}{\sqrt{r_Z(0)}}\right) \tag{6}$$

where for every $k = 1, \ldots, \frac{q}{2}$, $d_{f_q,2k} \in \mathbb{R}$ with $d_{f_q,q} \neq 0$. For simplicity, we will assume that f_q is normalized so that the coefficient of highest order $d_{f_q,q} = r_Z^{\frac{q}{2}}(0)$. From Section 2, we can write for every $i \geqslant 0$

$$f_q(Z_i) = \sum_{k=0}^{q/2} d_{f_q,2k} H_{2k}\left(\frac{Z_i}{\sqrt{r_Z(0)}}\right) = \sum_{k=0}^{q/2} d_{f_q,2k} I_{2k}\left(\varepsilon_i^{\otimes 2k}\right) \tag{7}$$

with $Z_i/\sqrt{r_Z(0)} = I_1(\varepsilon_i)$. Define the following partial sum

$$Q_{f_q,n}(Z) := \frac{1}{n}\sum_{i=0}^{n-1} f_q(Z_i) \qquad \text{and} \qquad \lambda_{f_q}(Z) := E\left[f_q(Z_0)\right].$$

The polynomial variation $Q_{f_q,n}(Z)$, as an empirical mean, should converge to $\lambda_{f_q}(Z)$. Our aim in this section is to estimate the parameter $\lambda_{f_q}(Z)$ by computing the speed of convergence of $Q_{f_q,n}(Z)$ to it. The "quadratic" case $q = 2$ is of special importance. In this case, the quadratic function f_2 will typically be taken as

$$f_2(x) = x^2 = r_Z(0) + r_Z(0)\, H_2\left(\frac{x}{\sqrt{r_Z(0)}}\right), \tag{8}$$

and may also be taken as $H_2(x) = x^2 - 1$ when convenient. We will see in Theorem 1 that certain functionals related to the quadratic case control the estimator's asymptotics no matter what q is. We will also provide an optimal treatment in the case $q = 2$ itself in Section 4.2.

3.3. Consistency

Theorem 2. *Suppose that there exist $\varepsilon > 0$ and $C > 0$ such that for every $n \geqslant 0$*

$$\sum_{k=0}^{n-1} r_Z(j)^2 \leqslant Cn^{1-\varepsilon}. \tag{9}$$

Then $Q_{f_q,n}(Z)$ is a consistent estimator of $\lambda_{f_q}(Z)$, i.e. almost surely, $\lim_{n\to\infty} Q_{f_q,n}(Z) = \lambda_{f_q}(Z)$.

Proof. It follows from (7) that

$$E\left[\left(Q_{f_q,n}(Z) - \lambda_{f_q}(Z)\right)^2\right] = E\left[\left(\frac{1}{n}\sum_{j=0}^{n-1} f_q(Z_j) - Ef_q(Z_j)\right)^2\right]$$
$$= \sum_{k=1}^{q/2} d_{f_q,2k}^2 \frac{(2k)!}{n^2} \sum_{i,j=0}^{n-1} \left(\frac{E(Z_i Z_j)}{r_Z(0)}\right)^{2k}$$
$$= \sum_{k=1}^{q/2} d_{f_q,2k}^2 \frac{(2k)!}{n^2} \sum_{i,j=0}^{n-1} \left(\frac{r_Z(i-j)}{r_Z(0)}\right)^{2k} = \sum_{k=1}^{q/2} d_{f_q,2k}^2 \frac{(2k)!}{n}$$
$$\times \left(1 + \frac{2}{n}\sum_{j=1}^{n-1}(n-j)\left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k}\right)$$
$$= \sum_{k=1}^{q/2} d_{f_q,2k}^2 \frac{(2k)!}{n} \left(1 + 2\sum_{j=1}^{n-1}\left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k} - \frac{2}{n}\sum_{j=1}^{n-1} j\left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k}\right). \tag{10}$$

Now, using (10), (9), (1) and Lemma 33 in the Appendix, the claim follows. ■

Remark 3. If Z is ergodic, the convergence in Theorem 2 is immediate.

3.4. Asymptotic distribution

Consider the following renormalized partial sum

$$U_{f_q,n}(Z) = \sqrt{n}\left(Q_{f_q,n}(Z) - \lambda_{f_q}(Z)\right). \tag{11}$$

Then, by (7) we can write

$$U_{f_q,n}(Z) = \sum_{k=1}^{q/2} I_{2k}(g_{2k,n}) \tag{12}$$

where

$$g_{2k,n} := d_{f_q,2k} \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varepsilon_i^{\otimes 2k}.$$

The following condition will play an important role in our analysis:

$$u_{f_2}(Z) := 2 \sum_{j \in \mathbb{Z}} r_Z(j)^2 < \infty. \tag{13}$$

Under this condition, we can pursue the analysis further: since $r_Z(j)^2$ converges to 0 as $|j| \to \infty$, hence for any k, $r_Z(j)^{2k}$ is dominated by $r_Z(j)^2$ for large $|j|$, and the last term in (10) can be estimated as follows. We first fix an $\varepsilon \in (0,1)$ and write for any $n \geqslant 2$

$$\frac{1}{n}\sum_{j=1}^{n-1} j \cdot \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k} = \frac{1}{n}\sum_{j=1}^{[\varepsilon n]-1} j \cdot \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k} + \frac{1}{n}\sum_{j=[\varepsilon n]}^{n-1} j \cdot \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k}$$

$$\leqslant \frac{\varepsilon n}{n}\sum_{j=1}^{[\varepsilon n]-1} \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k} + \frac{n}{n}\sum_{j=[\varepsilon n]}^{n} \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k} \leqslant \frac{\varepsilon}{2 r_Z(0)^2} u_{f_2}(Z) + \sum_{j=[\varepsilon n]}^{\infty} \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2}.$$

By Condition (13), with ε fixed, one can choose n so large that $\sum_{j=[\varepsilon n]}^{\infty} r_Z(j)^2 < \varepsilon$. Thus the last term in (10) can be made arbitrarily small. This immediately implies the following useful result.

Lemma 4. *Under Condition (13), for every even $q \geqslant 2$,*

$$u_{f_q}(Z) := \lim_{n\to\infty} E\left[U_{f_q,n}^2(Z)\right] = \sum_{k=1}^{q/2} d_{f_q,2k}^2 (2k)! \sum_{j\in\mathbb{Z}} \left(\frac{r_Z(j)}{r_Z(0)}\right)^{2k} < \infty. \tag{14}$$

The following notation will be convenient.

$$F_{f_q,n}(Z) := \frac{U_{f_q,n}(Z)}{\sqrt{E\left[U_{f_q,n}^2(Z)\right]}}. \tag{15}$$

In the next corollary, we look at two examples, one under Condition (13) and one when it fails but normality still holds. In the former case, we replace the normalization term $\sqrt{E\left[U_{f_q,n}^2(Z)\right]}$ which is an unobservable sequence because it depends on the parameter-dependent sequence r_Z, by the constant $\sqrt{u_{f_q}(Z)}$. While this constant also depends on the parameter $\lambda_{f_q}(Z)$, it allows one to measure the total variation distance of the data-based estimator $U_{f_q,n}(Z)$ itself to the fixed law $\mathcal{N}\left(0, u_{f_q}(Z)\right)$, consistent with common methodological practice. This change of normalization results in an additional term to reflect the speed of convergence of $\sqrt{E\left[U_{f_q,n}^2(Z)\right]}$ to $\sqrt{u_{f_q}(Z)}$.

Corollary 5. *(1) If $r_Z(k) \sim ck^{-\frac{1}{2}}$, then*

$$E\left[U_{f_q,n}^2(Z)\right] \sim 4c^2 d_{q,2}^2(Z) \log(n). \tag{16}$$

and the upper bound on $d_{TV}\left(F_{f_q,n}(Z), N\right)$ from Theorem 1 holds with

$$\kappa_4(F_{f_2,n}(Z)) = \mathcal{O}\left(\log^{-2}(n)\right). \tag{17}$$

(2) Under Condition (13), i.e. if $\sum_{j\in\mathbb{Z}}|r_Z(j)|^2<\infty$*, we have*

$$\begin{aligned} & d_{TV}\left(U_{f_q,n}(Z),\mathcal{N}\left(0,u_{f_q}(Z)\right)\right) \\ & \leqslant C_q(Z)\sqrt[4]{\frac{\kappa_4(U_{f_2,n}(Z))}{\left|u_{f_q}(Z)\right|^2}}+2\left|1-\frac{E\left[U^2_{f_q,n}(Z)\right]}{u_{f_q}(Z)}\right|. \end{aligned} \tag{18}$$

(3) Under the additional assumption that r_Z *is asymptotically of constant sign and monotone, the expressions in (18) converge to* 0.

Proof. The estimate (16) is a direct consequence of (10). Also, by (16) and (5) we obtain (17). The result of point (1) is established.

Next, we prove the estimate (18). We first note that $d_{TV}\left(U_{f_q,n}(Z),\sqrt{u_{f_q}(Z)}N\right)$ is identical to $d_{TV}\left(U_{f_q,n}(Z)/\sqrt{u_{f_q}(Z)},N\right)$. By the expression in (10), the ratio $E\left[U^2_{f_q,n}(Z)\right]/u_{f_q}(Z)$ is in (0, 1). Therefore the estimate (18) is an elementary consequence of (4).

To prove the corollary's final claim in point (3), we first note that by Lemma 4, the term $\left|1-E\left[U^2_{f_q,n}(Z)\right]/u_{f_q}(Z)\right|$ tends to 0. Thus we only need to show that under Condition (13) and the additional monotonicity assumption on r_Z, $\kappa_4(U_{f_2,n}(Z))$ also tends to 0. By the conclusion of Theorem 1 and the finiteness of $u_{f_2}(Z)$, we have

$$\kappa_4(U_{f_2,n}(Z))=\mathcal{O}\left(n^{-1}\left(\sum_{|j|<n}|r_Z(j)|^{4/3}\right)^3\right)=:K_4(n). \tag{19}$$

Next, we borrow from [22, Proposition 1] that, under the additional assumptions on r_Z in the last statement of the theorem, and using the finiteness of $u_{f_2}(Z)$,

$$K_4(n)=\mathcal{O}\left(n^{-1/3}\left(n^{-1/2}\left(\sum_{|j|<n}|r_Z(j)|^{3/2}\right)^2\right)^{4/3}\right).$$

From Jensen's inequality and Lemma 4 we get

$$n^{-1/2}\left(\sum_{|j|<n}|r_Z(j)|^{3/2}\right)^2\leqslant 4n^{3/2}\left(\frac{1}{2n}\sum_{|j|<n}|r_Z(j)|^2\right)^{3/2}\leqslant\sqrt{2}u_{f_2}(Z)<\infty.$$

Thus by (19), $\kappa_4(U_{f_2,n}(Z))=\mathcal{O}\left(n^{-1/3}\right)$, which finishes the proof of the corollary. ■

Corollary 6. *Under the notation of Corollary 5, assume that for some* $H<5/8$*,*

$$|r_Z(k)|\leqslant ck^{2H-2}.$$

Then Condition (13) holds and for some constant C depending only on q and $r_Z(0)$,

$$d_{TV}\left(U_{f_q,n}(Z),\mathcal{N}\left(0,u_{f_q}(Z)\right)\right)\leqslant Cn^{-1/4}.$$

Remark 7. In most cases where Condition (13) fails, the series' divergence occurs so fast that $U_{f_q,n}(Z)$'s asymptotics are not normal. Under certain special circumstances, namely a slowly modulated $(2H-2)$-self-similarity assumption on r_Z, classical tools such as in [11] can be used to show that $U_{f_q,n}(Z)$ tends to a so-called (scaled) Rosenblatt law $G_\infty^{(H)}$. A now classical result of Davydov and Martinova [10] was revived in recent years in [6,22] to estimate total-variation distances to $G_\infty^{(H)}$. This can be achieved in our context as well, though for the sake of conciseness, we omit this study, only stating two basic results here, whose proofs would proceed as in [22] and [6] respectively.

1. Assume that for some $H \in (3/4, 1)$ and some $\beta > 0$, asymptotically
$$(1+o(1))\log^{-\beta}(|k|)\,|k|^{2H-2} \leqslant |r_Z(k)|/r_Z(0) \leqslant (1+o(1))\log^{\beta}(|k|)\,|k|^{2H-2},$$
then for some constant C depending on r and H,
$$d_{TV}\left(\frac{U_{f_2,n}(Z)}{2\sum_{|k|>n} r_Z(k)^2}, \frac{1}{2}\sqrt{\frac{4H-3}{2\Gamma(2-2H)\cos\left(\frac{(2-2H)\pi}{2}\right)}}G_\infty^{(H)}\right) \leqslant \frac{C}{\sqrt{\log n}}.$$
2. If $\beta = 0$, then f_2 can be replaced above by f_q for any even q, and $3/4$ can be replaced by $1-1/(2q)$, and $\sqrt{\log n}$ by $n^{H-1+1/(2q)}$.

The law of $G_\infty^{(H)}$ can be represented under a standard white noise measure W on $\mathbb{C}$ as
$$G_\infty^{(H)} = \int\int_{\mathbb{R}^2} e^{i(x+y)}\frac{e^{i(x+y)}-1}{i(x+y)}|xy|^{1/2-H}W(dx)W(dy).$$

3.5. Examples

3.5.1. Hermite variation

Let $q \geqslant 2$ be an even integer. Then the qth Hermite polynomial H_q can be written as in (6). Indeed, it follows from the fact that for q even, $H_q(x) = \sum_{k=0}^{\frac{q}{2}} \frac{q!(-1)^k}{k!(q-2k)!2^k}x^{q-2k}$. Then we can write,
$$H_q(x) = EH_q(Z_0) + \sum_{k=1}^{q/2} d_{H_q,2k}(Z)\,H_{2k}\left(\frac{x}{\sqrt{r_Z(0)}}\right)$$
where for any $k \in \{1,\dots,\frac{q}{2}-1\}$
$$\begin{aligned} d_{H_q,q-2k}(Z) &= (-1)^k\left(r_Z^{\frac{q}{2}}(0) - r_Z^{\frac{q}{2}-1}(0)\right)a_{q-2}^{q}a_{q-4}^{q-2}\dots a_{q-2k}^{q-2k+2}\\ &+\quad (-1)^{k-1}\left(r_Z^{\frac{q}{2}}(0) - r_Z^{\frac{q}{2}-2}(0)\right)a_{q-4}^{q}a_{q-6}^{q-4}\dots a_{q-2k}^{q-2k+2}\\ &+\quad \dots\\ &+\quad (-1)^1\left(r_Z^{\frac{q}{2}}(0) - r_Z^{\frac{q}{2}-k}(0)\right)a_{q-2k}^{q}\end{aligned}$$
and $d_{H_q,q}(Z) = r_Z^{\frac{q}{2}}(0)$, where for every p even, the constants
$$a_{p-2k}^{p} = \frac{p!(-1)^k}{k!(p-2k)!2^k} \qquad k = 0,\dots,p/2$$
are the ones which satisfy
$$H_p(x) = \sum_{k=0}^{p/2} a_{p-2k}^{p}x^{p-2k}.$$

Consequently, the Hermite variation

$$Q_{H_q,n}(Z) := \frac{1}{n}\sum_{k=0}^{n-1} H_q(Z_k)$$

satisfies the results given in Sections 3.3 and 3.4. Moreover the parameter $\lambda_{H_q}(Z)$ has the following explicit expression

$$\begin{aligned}\lambda_{H_q}(Z) &= E\left[H_q(Z_0)\right] = \sum_{k=0}^{\frac{q}{2}} \frac{q!(-1)^k}{k!(q-2k)!2^k} E\left(Z_0^{q-2k}\right)\\ &= \frac{q!}{2^{q/2}}\sum_{k=0}^{\frac{q}{2}} \frac{(-1)^k}{k!(\frac{q}{2}-k)!}\left[E\left(Z_0^2\right)\right]^{\frac{q}{2}-k}\\ &= \frac{q!}{(\frac{q}{2})!2^{q/2}}\left(E\left(Z_0^2\right)-1\right)^{q/2}. \qquad (20)\end{aligned}$$

Thus the results of the previous sections provide explicit means for computing total variation speeds of convergence in a generalized method of moments based on Hermite polynomials for estimating the variance parameter $E\left(Z_0^2\right) = r_Z(0)$. Because of the simple form of (20) as a function of $r_Z(0)$, for any sequence satisfying Condition (13), one can immediately test the hypothesis of whether $r_Z(0)$ equals a specific value σ^2, using Corollary 5 to account precisely for the error term due to non-infinite sample size. Since the corollary provides the error in total variation distance, this error is uniform over σ^2 by definition.

3.5.2. Power variation

Let $q \geqslant 2$ be an even integer. Let $c_{q,2k} = \frac{1}{(2k)!}\int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} x^q H_{2k}(x)\,dx$ be the coefficients of the monomial $\phi_q(x) := x^q$ expanded in the basis of Hermite polynomials :

$$x^q = \sum_{k=0}^{q/2} c_{q,2k} H_{2k}(x).$$

It is known (see [29, Formula 18.18.20]) that

$$c_{q,2k} = \frac{q!}{2^{\frac{q}{2}-k}\left(\frac{q}{2}-k\right)!\,(2k)\,!}.$$

Thus, by relying directly on the results we just saw in the case of Hermite polynomials, the polynomial function ϕ_q can be written as in (6). As a consequence, the power variation

$$Q_{\phi_q,n}(Z) = \frac{1}{n}\sum_{i=0}^{n-1}(Z_i)^q$$

satisfies the results given in Sections 3.3 and 3.4. In this case, the parameter $\lambda_{\phi_q}(Z)$ has the following explicit expression

$$\lambda_{\phi_q}(Z) = E\left[(Z_0)^q\right] = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left[E\left(Z_0^2\right)\right]^{q/2}. \qquad (21)$$

4. Parameter estimation for non-stationary Gaussian processes

In practice, it is often the case that the data comes from a sequence which has visibly not yet reached a stationary regime. This is a typical situation for the solution of a stochastic system which initiates from a point mass rather than the system's stationary distribution; we will see examples of this in Sections 5 and 6. The rate at which stationarity is reached heavily affects other rates of convergence, including the total variation speeds in the central limit theorem. To illustrate this phenomenon more broadly than in the two aforementioned sections, in this section we consider a general class of models which can be written as the sum of a stationary model and a non-stationary nuisance term which vanishes asymptotically.

4.1. General case

Let $q \geqslant 2$ be an even integer. For a polynomial f_q of the form (6), and a random sequence X, recall the polynomial variation notation introduced in Section 3.2:

$$Q_{f_q,n}(X) := \frac{1}{n}\sum_{i=0}^{n-1} f_q(X_i).$$

Let $(Z_k)_{k\in\mathbb{Z}}$ be a centered stationary Gaussian process and let $(Y_k)_{k\in\mathbb{Z}}$ be a process such that the following condition holds: there exists a constant $\gamma > 0$ such that for every $p \geqslant 1$ and for all $n \in \mathbb{N}$,

$$\left\| Q_{f_q,n}(Z+Y) - Q_{f_q,n}(Z)\right\|_{L^p(\Omega)} = \mathcal{O}\left(n^{-\gamma}\right). \tag{22}$$

Combining (22), Lemma 33 in the Appendix, and Theorem 2 we get the following result.

Theorem 8. *Assume that the conditions (22) and (9) hold. Then*

$$Q_{f_q,n}(Z+Y) \longrightarrow \lambda_{f_q}(Z)$$

almost surely as $n \to \infty$.

In Corollary 5, we handled a discrepancy at the level of deterministic normalizing constants, while retaining statements with the total variation distance. In this section, our discrepancy comes at a slightly higher price because it is stochastic. We use instead the Wasserstein distance d_W, in order to rely on the following elementary lemma whose proof is obvious.

Lemma 9. *Let Y and Z be random variables defined on the same probability space. Then*

$$d_W\left(Y+Z, N\right) \leqslant d_W\left(Z, N\right) + \|Y\|_{L^1(\Omega)}.$$

Another lemma, proved for instance in [23, Theorem 5.1.3], relates the Wasserstein distance to a connection between Stein's method and the Malliavin calculus.

Lemma 10. *If F has mean 0, variance 1, and a square-integrable Malliavin derivative, then*

$$d_W\left(F, \mathcal{N}(0,1)\right) \leqslant \sqrt{\frac{2}{\pi}} E\left[\left|1 - \left\langle DF, -DL^{-1}F\right\rangle\right|\right].$$

By combining these two lemmas (the second one applies because variables with finite chaos expansions are infinitely Malliavin-differentiable with finite moments of all orders) and the proof of Theorem 1, by (22) we immediately obtain the following upper bounds.

Theorem 11. *Under hypothesis (22) and the assumptions of Theorem 1, for some constant C depending on the relation in (22),*

$$d_W\left(\frac{U_{f_q,n}(Z+Y)}{\sqrt{E\left[U^2_{f_q,n}(Z)\right]}},N\right)\leqslant C\frac{n^{\frac{1}{2}-\gamma}}{\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}+\frac{C_q\,(Z)\sqrt{2}}{\sqrt{\pi}}\sqrt[4]{\kappa_4(F_{f_2,n}(Z))}.$$

In addition, under Condition (13), i.e. if $\sum_{j\in\mathbb{Z}}|r_Z(j)|^2<\infty$,

$$d_W\left(\frac{U_{f_q,n}(Z+Y)}{\sqrt{u_{f_q}(Z)}},N\right)\leqslant C\frac{n^{\frac{1}{2}-\gamma}}{\sqrt{u_{f_q}(Z)}}+\frac{C_q\,(Z)\sqrt{2}}{\sqrt{\pi}}\sqrt[4]{\frac{\kappa_4(U_{f_2,n}(Z))}{u_{f_q}(Z)^2}}+\sqrt{\frac{8}{\pi}}\left|1-\frac{E\left[U^2_{f_q,n}(Z)\right]}{u_{f_q}(Z)}\right|. \tag{23}$$

One might hope to use the aforementioned result [10] of Davydov and Martynova to obtain bounds similar to those in the theorem above in the total variation distance instead of the Wasserstein metric. Unfortunately, [10] states that one can bound the total variation distance between two variables in the same chaos by the square root of their standard deviation, rather than the standard deviation itself. In other words, when working in a specific chaos, a version of Lemma 9 exists for the TV distance, except that the term $\|Y\|_{L^1(\Omega)}$ must be replaced by $\|Y\|^{1/2}_{L^2(\Omega)}$. This creates an inefficiency whereby, in the conclusion of Theorem 11, the term $n^{1/2-\gamma}$ would have to be replaced by the larger $n^{1/4-\gamma/2}$, which may switch the dominant term in the theorem's upper bound from $\left[\kappa_4(U_{f_2,n}(Z))\right]^{1/4}$ to $n^{1/4-\gamma/2}$. Another problem in applying this strategy is that one must work in a fixed chaos, which limits its application to Hermite variations. A further issue stems from the recent realization that the proof in [10] is sufficiently elliptic that one is not entirely sure that the result is strictly correct. A new proof was obtained in [26] by which the aforementioned $\|Y\|^{1/2}_{L^2(\Omega)}$ must be replaced by $\|Y\|^{1/4}_{L^2(\Omega)}$, and a recent paper [5] appears to reach nearly the original statement in [10] modulo a logarithmic correction; also see [31]. We do not discuss the possible extension to the TV distance further here; the situation is more promising in the quadratic case, which we take up in the next subsection.

4.2. Quadratic case

In this subsection we assume that $q=2$. In this special case, consistent with the notation in (8), without loss of generality up to deterministic shifting and scaling, the only relevant polynomial of interest is $f_2(x)=x^2$. Thus the question introduced in Section 3.2 is to estimate the variance $r_Z(0)=E\left[(Z_0)^2\right]$ where Z is our stationary Gaussian process. Using the notation introduced in that section, we thus have the following expression for our normalized partial sum

$$U_{f_2,n}(Z)=\frac{r_Z(0)}{\sqrt{n}}\sum_{k=0}^{n-1}H_2\left(\frac{Z_k}{\sqrt{r_Z(0)}}\right)=I_2\left(\frac{r_Z(0)}{\sqrt{n}}\sum_{k=0}^{n-1}\varepsilon_k^{\otimes 2}\right),$$

where again ε_k is defined by $Z_k/\sqrt{r_Z(0)}=I_1(\varepsilon_k)$. Using the notation in Section 3.4, the standardized version of $U_{f_2,n}(Z)$ is thus

$$F_{f_2,n}(Z)=\frac{U_{f_2,n}(Z)}{\sqrt{E\left[U^2_{f_2,n}(Z)\right]}}.$$

Define the third cumulant $\kappa_3\left(F_{f_2,n}(Z)\right) := E\left[F_{f_2,n}(Z)^3\right]$. We will apply the sharp asymptotics established in [24] (see bullet points in Section 2), by which a sequence of variance-one random variables F_n in a fixed Wiener chaos which converges in law to the normal has total variation distance to the normal commensurate with the maximum of its third and fourth cumulant. We will also apply an explicit version of this theorem, due to [22], tailored to quadratic variations of stationary Gaussian processes. For positive-valued sequences a and b, we will use the commensurability notation

$$a_n \asymp b_n \iff 0 < c := \inf_n \frac{a_n}{b_n} \leqslant \sup_n \frac{a_n}{b_n} =: C < \infty$$

where the extrema may be over all positive integers, or all integers exceeding a value n_0. Our first result is the following.

Proposition 12. *(1) With $f_2(x) = x^2$, assume that $\kappa_4(F_{f_2,n}(Z)) \longrightarrow 0$. Then*

$$d_{TV}\left(F_{f_2,n}(Z), N\right) \asymp \max\left\{\kappa_4\left(F_{f_2,n}(Z)\right), \left|\kappa_3\left(F_{f_2,n}(Z)\right)\right|\right\}. \tag{24}$$

(2) If r_Z is asymptotically of constant sign and monotone, then $\kappa_4(F_{f_2,n}(Z)) \longrightarrow 0$ if and only if $\kappa_3(F_{f_2,n}(Z)) \longrightarrow 0$, and in this case,

$$d_{TV}\left(F_{f_2,n}(Z), N\right) \asymp \left|\kappa_3\left(F_{f_2,n}(Z)\right)\right| = \left|E\left((F_{f_2,n}(Z))^3\right)\right|, \tag{25}$$

and moreover,

$$\left|E\left((F_{f_2,n}(Z))^3\right)\right| \asymp \frac{\left(\sum_{|k|<n} |r_Z(k)|^{3/2}\right)^2}{\left(\sum_{|k|<n} |r_Z(k)|^2\right)^{3/2}\sqrt{n}}.$$

Proof. The result (24) in Point (1) is a direct consequence of the main result in [24] (see also [4]). The statements in point (2) come directly from [22, Theorem 3]. ■

The methods used to prove Corollary 5 and Theorem 11 immediately lead from the upper bound statements in Proposition 12 to the following corollary.

Corollary 13. *If the hypothesis (22) holds, under the assumptions in part (2) of Proposition 12, for some constant $C > 0$,*

$$d_W\left(\frac{U_{f_2,n}(Z+Y)}{\sqrt{E\left[U^2_{f_2,n}(Z)\right]}}, N\right) \leqslant C\left(\frac{n^{\frac{1}{2}-\gamma}}{\sqrt{E\left[U^2_{f_2,n}(Z)\right]}} + \left|E\left((F_{f_2,n}(Z))^3\right)\right|\right).$$

In addition, if Condition (13) holds, i.e. $\sum_{j\in\mathbb{Z}}|r_Z(j)|^2 < \infty$, then

$$d_W\left(\frac{U_{f_2,n}(Z+Y)}{\sqrt{u_{f_2}(Z)}}, N\right) \leqslant C\left(\frac{n^{\frac{1}{2}-\gamma}}{\sqrt{u_{f_2}(Z)}} + \left|E\left((F_{f_2,n}(Z))^3\right)\right|\right) + C\frac{\sum_{|j|>n}|r_Z(j)|^2}{u_{f_2}(Z)}. \tag{26}$$

Unfortunately, these techniques say nothing about how to obtain lower bounds when one adds discrepancies corresponding to the speed of convergence of the series $\sum_j |r_Z(j)|^2$, and to a non-stationary term. We now investigate some slight strengthening of Conditions (13) and (22) which allow for such lower-bound statements, starting with some elementary considerations.

From (25) and the remainder of Point (2) in Proposition 12, there exists a constant $c_1(Z)$ depending only on the law of Z such that

$$c_1(Z)\frac{\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2}{\left(\sum_{|k|<n}|r_Z(k)|^2\right)^{3/2}\sqrt{n}} \leqslant d_{TV}\left(F_{f_2,n}(Z),N\right) \leqslant C_1(Z)\frac{\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2}{\left(\sum_{|k|<n}|r_Z(k)|^2\right)^{3/2}\sqrt{n}}. \tag{27}$$

Now assume merely that (13) holds: $\sum_j |r_Z(j)|^2$ converges. Thus, for some constant $c_2'(Z)$ depending only on the law of Z,

$$d_{TV}\left(F_{f_2,n}(Z),N\right) \geqslant \frac{c_2'(Z)}{\sqrt{n}}. \tag{28}$$

In cases where $\sum |r_Z(k)|^{3/2}$ diverges, we evidently get a larger lower bound than (28), which would make the rest of the analysis easier. To keep track of multiplicative constants as best we can, we define

$$L(Z) := \lim_{n\to\infty}\frac{\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2}{\left(\sum_{|k|<n}|r_Z(k)|^2\right)^{3/2}}, \tag{29}$$

which exists and is positive under condition (13), with the understanding that when $L(Z)$ is $+\infty$, one may replace it by an arbitrarily large constant for n large enough. See Corollary 16.

Thus in (28), we may take $c_2'(Z) = c_1(Z)L(Z)$ where $c_1(Z)$ is the lower bound constant from (25), i.e. as defined in (27) . Finally, we relate (28) to the Wasserstein distance through the following lemma, proved in the Appendix.

Lemma 14. *Lower bound statements in Proposition 12 hold for d_W with an additional factor 2.*

Thus, under Condition (13), by the previous development and Lemma 14, with

$$c_2(Z) := 2c_1(Z)L(Z), \tag{30}$$

we finally get

$$d_W\left(F_{f_2,n}(Z),N\right) \geqslant \frac{c_2(Z)}{\sqrt{n}}, \tag{31}$$

and we are ready to state and prove our lower bound theorem in the quadratic case under a sharpening of condition (22) and a quantitative version of Condition (13).

Theorem 15. *Assume the following two conditions.*

- *Let $(Y_k)_{k\in\mathbb{Z}}$ be a process such that for all $n\in\mathbb{N}$, for some finite constant $c_3>0$,*

$$\left\|Q_{f_2,n}(Z+Y)-Q_{f_2,n}(Z)\right\|_{L^1(\Omega)} \leqslant \frac{c_3\sqrt{u_{f_2}(Z)}}{n}. \tag{32}$$

- *Condition (13) holds and for some finite constant $c_4 > 0$*

$$\left| u_{f_2}(Z) - E\left[U_{f_2,n}^2(Z)\right]\right| \leqslant \frac{2\,c_4\,u_{f_2}(Z)}{\sqrt{n}}. \tag{33}$$

With the positive constants $c_1(Z)$, $C_1(Z)$, and $L(Z)$ defined via (27), (29), and $c_2 = 2c_1(Z)L(Z)$(30), which exist by Proposition 12, if $c_3 + c_4 < c_2$, with $\varepsilon > 0$ such that $(c_3+c_4)(1+\varepsilon) < c_2$, then there exists n_0 large enough that for all $n > n_0$,

$$\begin{aligned}\frac{2c_1(Z)L(Z) - (c_3+c_4)(1+\varepsilon)}{\sqrt{n}} &\leqslant d_W\left(\frac{U_{f_2,n}(Y+Z)}{\sqrt{u_{f_2}(Z)}}, N\right)\\ &\leqslant \frac{C_1(Z)L(Z) + c_3 + c_4}{\sqrt{n}}.\end{aligned}$$

Proof. By using Lemma 9, the lower bound (31) implies

$$\frac{c_2}{\sqrt{n}} \leqslant d_W\left(\frac{U_{f_2,n}(Y+Z)}{\sqrt{E\left[U_{f_2,n}(Z)^2\right]}}, N\right) + \frac{1}{\sqrt{E\left[U_{f_2,n}(Z)^2\right]}} E\left[\left|U_{f_2,n}(Y+Z) - U_{f_2,n}(Z)\right|\right].$$

Then by assumption (32),

$$\frac{c_2}{\sqrt{n}} \leqslant d_W\left(\frac{U_{f_2,n}(Y+Z)}{\sqrt{E\left[U_{f_2,n}(Z)^2\right]}}, N\right) + \frac{\sqrt{u_{f_2}(Z)}}{\sqrt{E\left[U_{f_2,n}(Z)^2\right]}}\frac{c_3}{\sqrt{n}}.$$

Now using the trivial consequence of Lemma 9 by which, for any random variable Z and constants a, b, $d_W(aZ, N) \leqslant d_W(bZ, N) + |a-b|\,\|Z\|_{L^1(\Omega)}$, we get

$$\begin{aligned}\frac{c_2}{\sqrt{n}} \leqslant\; & d_W\left(\frac{U_{f_2,n}(Y+Z)}{\sqrt{u_{f_2}(Z)}}, N\right)\\ &+ \left|u_{f_2}(Z)^{-1/2} - E\left[U_{f_2,n}(Z)^2\right]^{-1/2}\right| E\left[\left|U_{f_2,n}(Z+Y)\right|\right]\\ &+ \frac{\sqrt{u_{f_2}(Z)}}{\sqrt{E\left[U_{f_2,n}(Z)^2\right]}}\frac{c_3}{\sqrt{n}}.\end{aligned} \tag{34}$$

Regarding the middle term in the right-hand side above, we claim the following: for any $\varepsilon > 0$ and for n large enough,

$$\left|u_{f_2}(Z)^{-1/2} - E\left[U_{f_2,n}(Z)^2\right]^{-1/2}\right| E\left[\left|U_{f_2,n}(Z+Y)\right|\right] \leqslant c_4 n^{-1/2}(1+\varepsilon). \tag{35}$$

Let us prove (35). To lighten the notation, we drop the subscripts. By assumption (32), we have

$$\begin{aligned}&\left|u(Z)^{-1/2} - E\left[U(Z)^2\right]^{-1/2}\right| E\left[|U(Z+Y)|\right]\\ &\leqslant \sqrt{n}\left|u(Z)^{-1/2} - E\left[U(Z)^2\right]^{-1/2}\right| \left(E\left[|Q(Z) - \lambda(Z)|\right] + E\left[|Q(Z+Y) - Q(Z)|\right]\right)\\ &\leqslant \left|u(Z)^{-1/2} - E\left[U(Z)^2\right]^{-1/2}\right| \left(\sqrt{E\left[U(Z)^2\right]} + c_3 n^{-1/2}\sqrt{u(Z)}\right).\end{aligned}$$

Since $E\left[U(Z)^2\right] \to u(Z)$, after some simple algebra, for any fixed $\varepsilon > 0$ and n large enough,

$$\left|u(Z)^{-1/2} - E\left[U(Z)^2\right]^{-1/2}\right| E\left[|U(Z+Y)|\right] \leqslant \left|u(Z)^{-1/2} - E\left[U(Z)^2\right]^{-1/2}\right| \times (1+\varepsilon)\sqrt{u(Z)} \leqslant \frac{1+\varepsilon}{2u(Z)}\left|u(Z) - E\left[U(Z)^2\right]\right|.$$

Thus (35) follows immediately from assumption (33). Combining (35) with (34), and again using $E\left[U_{f_2,n}(Z)^2\right] \to u_{f_2}(Z)$, we finally obtain that for any $\varepsilon > 0$ and for n large enough

$$\frac{c_2}{\sqrt{n}} \leqslant d_W\left(\frac{U_{f_2,n}(Y+Z)}{\sqrt{u_{f_2}(Z)}}, N\right) + \frac{(c_3+c_4)(1+\varepsilon)}{\sqrt{n}}.$$

Since, $c_4 + c_4 < c_2$, $\varepsilon > 0$ exists such that $c_2 - (1+\varepsilon)(c_3+c_4) > 0$, which finishes the lower bound of the theorem. The upper bound is easier to prove, and follows from the same estimates as for the lower bound. Details are omitted. ■

The following corollary, which takes advantage of the slower convergence of $d_W\left(F_{f_2,n}(Z), N\right)$ when $L(Z) = +\infty$, is an immediate consequence of the proof of the theorem. It shows that optimality holds for all normal convergence rates in the Wasserstein metric.

Corollary 16. *Under the hypotheses of Theorem 15, if $L(Z) = +\infty$, i.e. if $\sum_{k=0}^{\infty}|r_Z(k)|^{3/2} = +\infty$, then for all c_3 and c_4,*

$$\frac{c_1(Z)}{\left(2^{-1}u_{f_2}(Z)\right)^{3/2}} \frac{\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2}{\sqrt{n}} \leqslant d_W\left(\frac{U_{f_2,n}(Y+Z)}{\sqrt{u_{f_2}(Z)}}, N\right) \leqslant \frac{C_1(Z)}{\left(2^{-1}u_{f_2}(Z)\right)^{3/2}} \frac{\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2}{\sqrt{n}}.$$

The above conclusion also holds, with appropriately modified constants to replace c_1 and C_1, under weaker versions of the assumptions (32) and (33) where the term $n^{-1/2}$ is multiplied by $\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2$, under the same resulting restriction on c_3 and c_4 as in Theorem 15.

In order to state an optimal result for the TV distance nonetheless, we base the next corollary on the validity of [10], with an understanding that logarithmic corrections could be required based on the results in [5] (see the discussion at the end of Section 4.1). We record this result here, whose extension along the lines of Corollary 16 is left to the reader.

Corollary 17. *Assume Z is stationary, condition (13) holds, and for some finite constant $c_4 > 0$*

$$\left|u_{f_2}(Z) - E\left[U_{f_2,n}^2(Z)\right]\right| \leqslant \frac{2\,c_4\,u_{f_2}(Z)}{n}. \tag{36}$$

This assumption holds for instance if $r_Z(k) \leqslant ck^{-\alpha}$ for some $\alpha \geqslant 1$. With the positive constants $c_1(Z)$, $C_1(Z)$, $L(Z)$ and c_2 as in Theorem 15, if there exists $\varepsilon > 0$ such that $c_4(1+\varepsilon) < c_2$, *then there exists n_0 large enough that for all $n > n_0$,*

$$\frac{2c_1(Z)\,L(Z) - c_4(1+\varepsilon)}{\sqrt{n}} \leqslant d_{TV}\left(\frac{U_{f_2,n}(Z)}{\sqrt{u_{f_2}(Z)}}, N\right) \leqslant \frac{C_1(Z)\,L(Z) + c_4}{\sqrt{n}}.$$

Remark 18. The non-central limit theorem in Remark 7 part (1) also holds if Z is replaced by $Z+Y$ under assumption (22) if $\gamma > 1/2$; and similarly for part (2) if $\gamma > H - (q-1)/2q$. These results' proofs, which are omitted, follow the results in Remark 7 and from the tools in this section and those in [22] and [6].

4.3. *Towards a Berry–Esséen theorem for parameter estimators*

In the previous two sections, we saw how to prove asymptotically normality for the empirical sums of the form $U_{f_q,n}(Z)$ (or $U_{f_q,n}(Y+Z)$ where Y is a non-stationary correction process), with convergence speed theorems in total variation and Wasserstein distances. These apply to parameter estimation if the quantity one is after is the expected value $\lambda_{f_q}(Z) := E\left[f_q(Z_0)\right]$. In this section we evaluate the same question if the parameter one seeks is implicit in $\lambda_{f_q}(Z)$.

Thus assume that one is looking for the unknown parameter $\theta > 0$ and that there is a homeomorphism g such that

$$\lambda_{f_q}(Z) = g^{-1}(\theta) := \theta^*.$$

As stated, so far, for a degree-q polynomial f_q of the form (6) we have studied the "estimator"

$$\widehat{\theta}_n = Q_{f_q,n}(Z) = \frac{1}{n}\sum_{i=0}^{n-1} f_q(Z_i).$$

We have proved the following in Section 3 (see for instance Theorems 2 and 1, Corollaries 5 and 6): $\widehat{\theta}_n \longrightarrow \theta^*$ almost surely and

$$d_W\left(\sqrt{n}E\left[U^2_{f_q,n}(Z)\right]^{-1/2}\left(\widehat{\theta}_n - \theta^*\right), \mathcal{N}(0,1)\right) \leqslant \varphi(n)$$

where

$$U_{f_q,n}(Z) = \sqrt{n}\left(Q_{f_q,n}(Z) - \lambda_{f_q}(Z)\right) = \widehat{\theta}_n - \theta^*.$$

and where $\varphi(n)$ tends to 0 as $n \to \infty$ at various speeds which can be determined thanks to the precise statements in Corollary 5, for instance $\varphi(n) = 1/\sqrt{n}$ in Corollary 6, which is the classical Berry–Esséen speed. By using the relation between θ and λ, we naturally define the estimator of θ by

$$\check{\theta}_n := g\left(\widehat{\theta}_n\right).$$

This is a consistent estimator by Theorem 2 since g is continuous by assumption: $\check{\theta}_n \longrightarrow \theta$ almost surely. Now assume g is a diffeomorphism. By the mean-value theorem we can write

$$\left(\check{\theta}_n - \theta\right) = g'(\xi_n)\left(\widehat{\theta}_n - \theta^*\right),$$

where ξ_n is a random variable which belongs to $[|\widehat{\theta}_n, \theta^*|]$. We can write

$$d_W\left(\frac{\sqrt{n}}{g'(\theta^*)\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\check{\theta}_n-\theta\right), \mathcal{N}(0,1)\right)$$
$$\leqslant d_W\left(\frac{\sqrt{n}}{g'(\theta^*)\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\check{\theta}_n-\theta\right), \frac{\sqrt{n}}{\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\widehat{\theta}_n-\theta^*\right)\right)$$
$$+ d_W\left(\frac{\sqrt{n}}{\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\widehat{\theta}_n-\theta^*\right), \mathcal{N}(0,1)\right).$$

The last term above is controlled by $\varphi(n)$ as mentioned. Now assume that g is twice continuously differentiable. Then by the mean-value theorem again, for ζ_n some random variable which belongs to $[|\xi_n, \theta^*|] \subset [|\widehat{\theta}_n, \theta^*|]$, the other term above is controlled as follows

$$d_W\left(\frac{\sqrt{n}}{g'(\theta^*)\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\check{\theta}_n-\theta\right), \frac{\sqrt{n}}{\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\widehat{\theta}_n-\theta^*\right)\right)$$
$$\cdot|g'(\theta^*)|\sqrt{E\left[U^2_{f_q,n}(Z)\right]}$$
$$\leqslant E\left|\sqrt{n}\left(\widehat{\theta}_n-\theta^*\right)\left(g'(\xi_n)-g'(\theta^*)\right)\right| = E\left|\sqrt{n}\left(\widehat{\theta}_n-\theta^*\right)g''(\zeta_n)\left(\xi_n-\theta^*\right)\right|$$
$$\leqslant E\left|\sqrt{n}\left(\widehat{\theta}_n-\theta^*\right)^2 g''(\zeta_n)\right|$$
$$\leqslant \sqrt{n}\left[E\left(\left(\widehat{\theta}_n-\theta^*\right)^{2p}\right)\right]^{1/p}\left[E\left(g''(\zeta_n)^{p'}\right)\right]^{1/p'},$$

where p and p' are conjugate reals greater than 1, i.e. $1/p + 1/p' = 1$. Moreover, by (1),

$$\sqrt{n}\left[E\left(\left(\widehat{\theta}_n-\theta^*\right)^{2p}\right)\right]^{1/p} \leqslant c_p\sqrt{n}E\left(\left(\widehat{\theta}_n-\theta^*\right)^2\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Therefore the only question left to transfer the quantitative results of Section 3 to $\check{\theta}_n$ is whether one can prove, for instance, that $g''(\zeta_n)$ has a bounded moment of order greater than 1. We will see several examples in Section 5 where this is easy to check. More generally, we advocate checking this on a case-by-case basis when the function g can be identified. In the meantime, we summarize this discussion with the following general principle, which follows from the above discussion.

Theorem 19. *Consider the setup from Corollary 5, in which $\widehat{\theta}_n = Q_{f_q,n}(Z) := \frac{1}{n}\sum_{i=0}^{n-1} f_q(Z_i)$ and $\theta^* = E\left[f_q(Z_0)\right]$, with $\varphi(n)$ an upper bound for the expression in (18) which converges to* 0. *Assume that there exist a twice-differentiable invertible function g and a value θ such that*

$$g^{-1}(\theta) := \theta^*.$$

If $g''\left(\widehat{\theta}_n\right)$ has a moment of order greater than 1 which is bounded in n, the expression

$$\check{\theta}_n := g\left(\widehat{\theta}_n\right)$$

is a strongly consistent and asymptotically normal estimator of θ and

$$d_W\left(\frac{\sqrt{n}}{g'(\theta^*)\sqrt{E\left[U^2_{f_q,n}(Z)\right]}}\left(\check{\theta}_n-\theta\right),\mathcal{N}(0,1)\right)\leqslant C\frac{1}{\sqrt{n}}+\varphi(n),$$

where $\varphi(n)$ is the speed of convergence in Corollary 5.

5. Applications to Ornstein–Uhlenbeck processes: the scalar case

5.1. Fractional Ornstein–Uhlenbeck process: general case

Consider an Ornstein–Uhlenbeck process $X=\{X_t,t\geqslant 0\}$ driven by a fractional Brownian motion $B^H=\left\{B^H_t,t\geqslant 0\right\}$ of Hurst index $H\in(0,1)$. That is, X is the solution of the following linear stochastic differential equation

$$X_0=0;\quad dX_t=-\theta X_t dt+dB^H_t,\quad t\geqslant 0, \tag{37}$$

whereas $\theta>0$ is considered as unknown parameter. The solution X of (37) has the following explicit expression:

$$X_t=\int_0^t e^{-\theta(t-s)}dB^H_s. \tag{38}$$

Thus, we can write

$$X_t=Z^\theta_t-e^{-\theta t}Z^\theta_0 \tag{39}$$

where

$$Z^\theta_t=\int_{-\infty}^t e^{-\theta(t-s)}dB^H_s. \tag{40}$$

Moreover, it is known that Z^θ is an ergodic stationary Gaussian process. It is the stationary solution of Eq. (37). We are thus in the setup of Section 4 with $Z=Z^\theta$ and $Y=-e^{-\theta t}Z^\theta_0$. Consequently, to apply the results of that section, we only need to check that Condition (22) holds. It does, according to the following result.

Lemma 20. *Let X and Z^θ be the processes given in (37) and (40) respectively. Then for every $p\geqslant 1$ and for all $n\in\mathbb{N}$,*

$$\left\|Q_{f_q,n}(X)-Q_{f_q,n}(Z^\theta)\right\|_{L^p(\Omega)}=\mathcal{O}\left(n^{-1}\right).$$

Proof. By (39) and (6) we have

$$\left\|Q_{f_q,n}(X)-Q_{f_q,n}(Z^\theta)\right\|_{L^p(\Omega)}\leqslant\frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{q/2}d_{f_q,2k}\times\left\|H_{2k}\left(\frac{X_i}{\sqrt{r_Z(0)}}\right)-H_{2k}\left(\frac{Z^\theta_i}{\sqrt{r_Z(0)}}\right)\right\|_{L^p(\Omega)}.$$

Combining this and the fact that

$$H_{2k}\left(\frac{X_i}{\sqrt{r_Z(0)}}\right) - H_{2k}\left(\frac{Z_i^\theta}{\sqrt{r_Z(0)}}\right) = \sum_{l=0}^{k} \frac{(2k)!(-1)^l}{l!(2k-2l)!2^l} \times \sum_{j=1}^{2k-2l} \frac{(-1)^j \binom{2k-2l}{j} e^{-\theta i j}}{r_Z^{k-l}(0)} (Z_0^\theta)^j (Z_i^\theta)^{2k-2l-j}.$$

we deduce that there exists a constant $c(\theta, f_q)$ depending on f_q and θ such that

$$\left\| Q_{f_q,n}(X) - Q_{f_q,n}(Z^\theta) \right\|_{L^p(\Omega)} \leqslant c(\theta, f_q) \frac{1}{n} \sum_{i=0}^{n-1} e^{-i\theta}.$$

Thus the lemma is obtained. ■

As a consequence, by using Z^θ ergodic, Lemma 20 and Theorem 8, we conclude that

$$Q_{f_q,n}(X) \longrightarrow \lambda_{f_q}(Z^\theta)$$

almost surely as $n \to \infty$. Moreover, by the Gaussian property of Z^θ and Lemma 35 in the Appendix, we can write

$$\lambda_{f_q}(Z^\theta) := \mu_{f_q}(\theta)$$

where μ_{f_q} is a univariate function of θ determined by the polynomial f_q. Hence, in the case when the function μ_{f_q} is invertible, we obtain the following estimator for θ

$$\check{\theta}_{f_q,n} := \mu_{f_q}^{-1}\left[Q_{f_q,n}(X)\right]. \tag{41}$$

These considerations allow us to state and prove the following strong consistency and asymptotic normality of $\check{\theta}_{f_q,n}$. For asymptotic normality, we assume that μ_{f_q} is a diffeomorphism. Examples of this situation are given in Section 5.2.

Proposition 21. *Assume $H \in (0, 1)$ and μ_{f_q} is a homomorphism. Let $\widehat{\theta}_{f_q,n}$ be the estimator given in (41). Then, as $n \longrightarrow \infty$, almost surely, $\check{\theta}_{f_q,n} \longrightarrow \theta$.*

Proposition 22. *Denote $N \sim \mathcal{N}(0, 1)$. Then*

- *If $H \in (0, \frac{5}{8})$, for any q,*

$$d_W\left(\frac{U_{f_q,n}(X)}{\sqrt{u_{f_q}(Z^\theta)}}, N\right) \leqslant \frac{C}{n^{1/4}}.$$

- *If $H \in (\frac{5}{8}, \frac{3}{4})$, for any q,*

$$d_W\left(\frac{U_{f_q,n}(X)}{\sqrt{u_{f_q}(Z^\theta)}}, N\right) \leqslant \frac{C}{n^{(4H-3)/2}}.$$

- *In particular, in both cases, assuming μ_{f_q} is a diffeomorphism,*

$$\sqrt{n}\left(\check{\theta}_{f_q,n} - \theta\right) \xrightarrow{law} \mathcal{N}\left(0, \frac{u_{f_q}(Z^\theta)}{(\mu'_{f_q}(\theta))^2}\right).$$

- *If* $H = \frac{3}{4}$,

$$d_W\left(\frac{U_{f_q,n}(X)}{\sqrt{E\left[U^2_{f_q,n}(Z^\theta)\right]}}, N\right) \leqslant C\log^{-\frac{1}{4}}(n).$$

In particular,

$$\sqrt{\frac{n}{\log(n)}}\left(\check{\theta}_{f_q,n} - \theta\right) \xrightarrow{law} \mathcal{N}\left(0, \frac{9d^2_{q,2}(Z)}{16\theta^4(\mu'_{f_q}(\theta))^2}\right).$$

Proof. In this proof, C represents a constant which may change from line to line. It was proved in [22] (also see [4]) that, with r_Z the covariance function of Z^θ,

$$\kappa_4(U_{f_2,n}(Z^\theta)) \leqslant C\left(\sum_{|k|<n} |r_Z(k)|^{4/3}\right)^3$$

while

$$E\left[U^2_{f_q,n}(Z^\theta)\right] \leqslant C\sum_{|k|<n} |r_Z(k)|^2.$$

Then by Lemma 35, for all $H < 3/4$, we easily get $E\left[U^2_{f_q,n}(Z^\theta)\right] \leqslant u_{f_q}(Z^\theta) < \infty$ and in particular

$$\left|1 - \frac{E\left[U^2_{f_q,n}(Z^\theta)\right]}{u_{f_q}(Z^\theta)}\right| \leqslant C\sum_{|k|>n} |r_Z(k)|^2 \leqslant Cn^{4H-3}. \tag{42}$$

Also by Lemma 35, for $H < 5/8$,

$$\kappa_4(U_{f_2,n}(Z^\theta)) \leqslant Cn^{-1}$$

while for $H > 5/8$,

$$\kappa_4(U_{f_2,n}(Z^\theta)) \leqslant Cn^{2(4H-3)}.$$

Then by Theorem 11 and by Lemma 20 which shows that $\gamma = 1$, we get

$$d_W\left(\frac{U_{f_q,n}(Z+Y)}{\sqrt{u_{f_q}(Z)}}, N\right) \leqslant Cn^{-1/2} + \sqrt[4]{\kappa_4(U_{f_2,n}(Z^\theta))} + Cn^{4H-3}$$

depending on whether H is larger or smaller than $5/8$ we get the announced result, since $n^{-1/4}$ and $n^{(4H-3)/2}$ coming from $\sqrt[4]{\kappa_4(U_{f_2,n}(Z^\theta))}$ dominate the error terms n^{4H-3} and $n^{-1/2}$.

Now, by assumption, μ_{f_q} has a continuously differentiable derivative. Thus, by the mean value theorem, there exists a random variable $\xi_{f_q,n}$ between θ and $\widehat{\theta}_{f_q,n}$ such that

$$\sqrt{n}\left(\mu_{f_q}(\widehat{\theta}_{q,n}) - \mu_{f_q}(\theta)\right) = \mu'_{f_q}(\xi_{f_q,n})\sqrt{n}\left(\widehat{\theta}_{q,n} - \theta\right).$$

By the asymptotic normality of $\mu_{f_q}(\widehat{\theta}_{q,n}) - \mu_{f_q}(\theta)$ and the a.-s. convergence of $\mu'_{f_q}(\xi_{f_q,n})$ to $\mu'_{f_q}(\theta)$, the theorem's statement when $H < 3/4$ follows. The case of $H = 3/4$ is treated similarly. ■

5.2. Examples and a Berry–Esséen theorem for drift estimators

In the two following examples, the function μ_{f_q} is an explicit diffeomorphism except at $\theta = 0$.

- Assume that $f_q = H_q$. Using (20) and Lemma 35, we have
$$\mu_{H_q}(\theta) = \lambda_{H_q}(Z^\theta) = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left(H\Gamma(2H)\theta^{-2H} - 1\right)^{q/2}.$$
In this case, the function μ is a diffeomorphism with bounded derivatives when the range is restricted to $\mathbf{R}_+$. Since, by the previous strong consistency proposition, $Q_{f_q,n}(X)$ ends up in $\mathbf{R}_+$ almost surely, the estimator $\check{\theta}_{f_q,n}$ is asymptotically equivalent to the one in which the function $g = \mu_{f_q}^{-1}$ is restricted to $\mathbf{R}_+$. This observation will be helpful below when applying the results of Section 4.3.
- Assume that $f_q = \phi_q$ with $\phi_q(x) = x^q$. From (21) and Lemma 35 we obtain
$$\mu_{\phi_q}(\theta) = \lambda_{\phi_q}(Z^\theta) = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left[H\Gamma(2H)\theta^{-2H}\right]^{q/2}.$$
The singularity of μ at $\theta = 0$ poses some technical problems when one tries to translate the consistency of $Q_{f_q,n}(X)$ into that of $\check{\theta}_{f_q,n}$ thanks to Section 4.3, which we investigate below.

We now show how the principle described in Section 4.3 can be used to estimate the speed of convergence for the estimator $\check{\theta}_{f_q,n}$ itself. To work in a specific situation, we look at the above two examples, assuming $q = 2$.

5.2.1. Berry–Esséen theorem for a Hermite-variations-based estimator for θ

In the notation of Section 4.3, using the convention of replacing $Q_{H_2,n}(Z)$ by $\left|Q_{H_2,n}(Z)\right|$, in the case of the Hermite polynomial H_2 we have

$$\mu_{H_2}(\theta) = \lambda_{H_2}(Z^\theta) = H\Gamma(2H)|\theta|^{-2H} - 1 = g^{-1}(\theta)$$

and thus

$$g(x) = g_{H_2}(x) := (H\Gamma(2H))^{-1/(2H)}(1 + |x|)^{-1/(2H)}, \tag{43}$$

and $g''(x)$ is proportional to $(1 + |x|)^{-1/(2H)-2}$. This function is bounded on $\mathbf{R}_+$. Hence, according to Theorem 19, using the speed of convergence from Proposition 12, we obtain the following.

Proposition 23. *For the stationary fractional Ornstein–Uhlenbeck Z^θ in (40), with $Q_{H_2,n}(Z^\theta) = \frac{1}{n}\sum_{k=1}^{n} Z^\theta(k)^2 - 1$ and g as in (43), we get*

$$d_W\left(\frac{\sqrt{n}}{g'(\theta^*)\sqrt{E\left[U_{H_2,n}^2(Z)\right]}}\left(g\left(Q_{H_2,n}\left(Z^\theta\right)\right) - \theta\right), \mathcal{N}(0,1)\right) \leqslant C\frac{1}{\sqrt{n}} + \varphi(n)$$

where $g(x) = (H\Gamma(2H))^{-1/(2H)}(1+|x|)^{-1/(2H)}$ *and* $\theta^* = g^{-1}(\theta)$ *and*

$$\varphi(n) \asymp \frac{\left(\sum_{|k|<n} |k|^{3H-3}\right)^2}{\left(\sum_{|k|<n} |k|^{4H-4}\right)^{3/2} \sqrt{n}}.$$

In particular, for $H < 2/3$, $\varphi(n) \asymp 1/\sqrt{n}$.

Remark 24. Improvements to the above proposition which include the use of the asymptotic variance $u_{f_q}(Z^\theta)$ and the nonstationary fractional OU process X in (38) also hold. These are omitted here for the sake of conciseness; the reader will find these topics covered in Section 5.3below.

5.2.2. Comments and strategy for θ *estimators with singular variance function*

For the power-2 function, in the notation of Section 4.3 we have

$$\mu_{\phi_2}(\theta) = \lambda_{\phi_2}(Z^\theta) = H\Gamma(2H)\theta^{-2H} = g^{-1}(\theta)$$

and thus $g''(x) = c_H x^{-1/(2H)-2}$ which has a singularity at 0 and thus is not bounded. A result can be obtained immediately from the above proposition since $g\left(Q_{\phi_2,n}\left(Z^\theta\right) - 1\right) = g\left(Q_{H_2,n}\left(Z^\theta\right)\right)$ is the estimator studied in that proposition. However, for illustrative purposes, we finish this section by outlining a method for dealing with the singularity, since this works for any g such that g'' is asymptotically decreasing like a negative power, and any process that has an infinite Karhunen–Loève expansion.

According to Theorem 19, we are asking whether for some $p' > 1$,

$$\sup_n E\left[|\widehat{\theta}_n|^{-p'/(2H)-2p'}\right] < \infty.$$

This condition is not entirely trivial, and can fail in some simple pathologically degenerate cases such as if Z is constant, since then $\widehat{\theta}_n = n^{-1}\sum_{k=1}^n Z(k)^2 = Z(0)^2$ is a chi-squared variable with one degree of freedom, which has no moments of negative order less than $-1/2$. This pathology does not occur for the fOU process, though the argument is slightly involved, since the limit law of the renormalized $\widehat{\theta}_n$ is normal, which does not have higher negative moments either. We decompose

$$E\left[g''\left(\widehat{\theta}_n\right)\right] = E\left[g''\left(\widehat{\theta}_n\right)\mathbf{1}_{\widehat{\theta}_n<1/\sqrt{n}}\right] + E\left[g''\left(\widehat{\theta}_n\right)\mathbf{1}_{\widehat{\theta}_n\geqslant 1/\sqrt{n}}\right].$$

By the asymptotic normality of $\left(\widehat{\theta}_n - \theta^*\right)\sqrt{n}$, we get

$$\begin{aligned}
&E\left[g''\left(\widehat{\theta}_n\right)\mathbf{1}_{|\widehat{\theta}_n|\geqslant 1/\sqrt{n}}\right] \sim \int_{-\sqrt{n}\theta^*+1}^{\infty} \left|\frac{z}{\sqrt{n}} + \theta^*\right|^{-p'(2+1/(2H))} e^{-z^2/2}dz \\
&= \int_{-\sqrt{n}\theta^*+1}^{-\sqrt{n}\theta^*/2} \left|\frac{z}{\sqrt{n}} + \theta^*\right|^{-p'(2+1/(2H))} e^{-z^2/2}dz \\
&\quad + \int_{-\sqrt{n}\theta^*/2}^{\infty} \left|\frac{z}{\sqrt{n}} + \theta^*\right|^{-p'(2+1/(2H))} e^{-z^2/2}dz \\
&\leqslant cst\; e^{-n\theta^*/8} n^{p'(2+1/(2H))} + \left(\theta^*/2\right)^{-p'(2+1/(2H))}
\end{aligned}$$

which is bounded for all n.

For the second piece, the normal approximation would not yield a finite bound, thus we must return to the original expression of $\widehat{\theta}_n$ as a 2nd chaos variable. It is known (see [20, page 522]) that $Z^{\theta}(k)$ has a Karhunen–Loève expansion $\sum_{m=0}^{\infty}\sqrt{\lambda_m}e_m(k)\,W_m$ (where the W_m are i.i.d. standard normal, and the e_m are orthonormal in $L^2([0,n])$) such that $\lambda_m \sim cm^{2H-2}$. Thus, the expansion of Z^{θ} contains infinitely many independent terms. One also knows (see [23, Section 2.7.4]) that $\hat{\theta}_n$, like any variable in the second chaos, can be expanded as $\sum_{m=0}^{\infty}\mu_m W_m^2$ where the μ_m are summable. One can check that the infinity of distinct terms in the expansion of Z^y implies that for any fixed n, the expansion of $\hat{\theta}_n$ also contains infinitely many terms, and that the coefficients are positive. Therefore, for any fixed m_0, there exists a positive constant c_{m_0} such that $\hat{\theta}_n \geqslant c_{m_0}\sum_{m=0}^{m_0}W_m^2 =: S_{m_0}$, which is a random variable with χ^2 distribution with m_0 degrees of freedom. Hence the density of S at the origin is of the order $z^{(m_0-1)/2}$, which means it has a negative moment of order $-p'(2+1/(2H))$ as soon as $m_0 > p'(2+1/(2H)) - 1$. From this it follows that $E\left[g''\left(\widehat{\theta}_n\right)\mathbf{1}_{\widehat{\theta}_n<1/\sqrt{n}}\right]$ is bounded. Thus g and $\hat{\theta}$ comply with the conditions of Theorem 19. In fact, since p' can be taken arbitrarily close to 1, we only need to be able to choose $m_0 > 1 + 1/(2H)$. For instance, if $H > 1/2$, this means that for Theorem 19 to work with $q = 2$, one only needs to show that the second-chaos series decomposition of $\widehat{\theta}_n$ contains 2 independent terms. This covers all Gaussian processes except for the trivial case of the constant process.

5.3. Optimal Berry–Esséen theorem in the quadratic case

As in the previous sections, the convergence speed for general q has no reason to be optimal. We illustrate this by studying the case $q = 2$, where we can improve the rate convergence thanks to the optimal rates obtained in Section 4.2, and even obtain optimal two-sided bounds when $H < 5/8$. Note that the results in this section deal with the fully realistic scenario where observations come from the non-stationary process X in (38) and there is no reference to normalizing constants other than finite asymptotic variances.

5.3.1. Setting up the rates of convergence

First assume that $H \leqslant 3/4$. We find that $\kappa_3(F_{f_2,n}(Z^{\theta})) \longrightarrow 0$ and more precisely (see [22]) that

$$\left|E\left((F_{f_2,n}(Z^{\theta}))^3\right)\right| \asymp \frac{\left(\sum_{|k|<n}|r_Z(k)|^{3/2}\right)^2}{\left(\sum_{|k|<n}|r_Z(k)|^2\right)^{3/2}\sqrt{n}}$$
$$\leqslant C \times \begin{cases} n^{-\frac{1}{2}}, & \text{if } 0 < H < \frac{2}{3} \\ \log^2(n)n^{-\frac{1}{2}}, & \text{if } H = \frac{2}{3} \\ n^{6H-\frac{9}{2}}, & \text{if } \frac{2}{3} < H < \frac{3}{4} \\ \log^{-3/2}(n), & \text{if } H = \frac{3}{4}. \end{cases} \tag{44}$$

By using Corollary 13, we see that we must compare the rates therein to the rates obtained in (44). By (42) the rate which controls the convergence of the variances is n^{4H-3}. This can be dominated by $1/\sqrt{n}$ if and only if $n < 5/8$. For $H \in [2/3, 3/4)$, n^{4H-3} dominates the rates in (44). The rate which controls the non-stationarity term is always of order $1/\sqrt{n}$ because of Lemma 20, which is always the lowest-order term. Hence the improved rates in (44) only come into play when $H < 5/8$ when normalizing by the asymptotic variance. In other words, we have the following two estimates, where the second one avoids the use of non-empirical statistics.

Proposition 25. *Denote $N \sim \mathcal{N}(0,1)$. If $H \in (0, \frac{3}{4}]$,*

$$d_W\left(\frac{U_{f_2,n}(X)}{\sqrt{E\left[U^2_{f_2,n}(Z^\theta)\right]}}, N\right) \leqslant C \times \begin{cases} n^{-\frac{1}{2}}, & \text{if } 0 < H < \frac{2}{3} \\ \log^2(n) n^{-\frac{1}{2}}, & \text{if } H = \frac{2}{3} \\ n^{6H-\frac{9}{2}}, & \text{if } \frac{2}{3} < H < \frac{3}{4} \\ \log^{-3/2}(n), & \text{if } H = \frac{3}{4} \end{cases}$$

and

$$d_W\left(U_{f_2,n}(X), \sqrt{u_{f_2}(Z^\theta)}N\right) \leqslant C \times \begin{cases} n^{-\frac{1}{2}}, & \text{if } 0 < H < \frac{5}{8} \\ n^{4H-3}, & \text{if } \frac{5}{8} \leqslant H < \frac{3}{4}. \end{cases}$$

Next we obtain optimal rates of convergence in the Wasserstein distance when $H < 5/8$.

5.3.2. Applying the optimal theorem

To apply Theorem 15 we must check that conditions (32) and (33) are met. We just saw that this is the case when $H < 5/8$. However, we must also check that the corresponding constants c_3 and c_4 are sufficiently small. Since $H < 5/8$, by (42), the constant c_4 can be made arbitrarily small for n large enough. It remains to show that c_3 can be chosen small. A direct application of Lemma 20 is insufficient for this purpose. Therefore, we must modify our estimator slightly, by discarding some of the first terms. We thus fix an integer $i_0 > 0$ and define

$$\tilde{Q}_{f_2,n}(X) := \frac{1}{n}\sum_{i=i_0}^{i_0+n-1} f_2(X_i). \tag{45}$$

It is easy to check that this is still a consistent and asymptotically normal estimator of $r_Z(0)$. By the proof of Lemma 20, we see that

$$\left\|\tilde{Q}_{f_2,n}(X) - \tilde{Q}_{f_2,n}\left(Z^\theta\right)\right\|_{L^p(\Omega)} \leqslant c(\theta, f_2)\frac{1}{n}\sum_{i=i_0}^{i_0+n-1} e^{-i\theta} \leqslant c(\theta, f_2)\frac{1}{n}e^{-i_0\theta}. \tag{46}$$

Since $\theta > 0$, we can make the last expression above as small as we want by choosing i_0 sufficiently large. Thus Theorem 15 applies, and we have the following optimal Berry–Esséen theorem for the variance estimator $\tilde{Q}_{f_2,n}(X)$, which, as we saw in Section 5.1, gives access to estimators for θ.

Proposition 26. *If $H \in \left(0, \frac{5}{8}\right)$, then there exists an integer $i_0 > 0$ such that the quadratic variation $\tilde{Q}_{f_2,n}$ defined in (45) satisfies*

$$d_W\left(\sqrt{n}\left[\tilde{Q}_{f_2,n}(X) - r_Z(0)\right], \mathcal{N}(0, u_{f_2}(Z^\theta))\right) \asymp \frac{1}{\sqrt{n}}.$$

Moreover, with g as in (43), we have

$$d_W\left(\sqrt{n}\left(g\left(\tilde{Q}_{H_2,n}(X)\right) - \theta\right), \mathcal{N}(0, g'(\theta^*)^2 u_{H_2}(Z^\theta))\right) \asymp \frac{1}{\sqrt{n}}.$$

Proof. The first result follows from the considerations immediately above. The second follows from Theorem 19 exactly as did the result in Proposition 23; we omit the details. ■

6. Application to Ornstein–Uhlenbeck processes : multi-parameter examples

In the previous section, we provided a full study of univariate parameter estimation for a fractional Ornstein–Uhlenbeck process, including all details of how to apply our general theory. In this final section of our article, we give two more examples of applications of our methods. For the sake of conciseness, we focus on the results, providing only a minimal amount of computations and proofs, since these are all modeled on the arguments in Section 5.

6.1. OU driven by fractional Ornstein–Uhlenbeck process

In this section we assume that $X = \{X_t, t \geqslant 0\}$ is an Ornstein–Uhlenbeck process driven by a fractional Ornstein–Uhlenbeck process $V = \{V_t, t \geqslant 0\}$. This is given by the following linear stochastic differential equations

$$\begin{cases} X_0 = 0; \quad dX_t = -\theta X_t dt + dV_t, \quad t \geqslant 0 \\ V_0 = 0; \quad dV_t = -\rho V_t dt + dB_t^H, \quad t \geqslant 0, \end{cases} \tag{47}$$

where $B^H = \left\{B_t^H, t \geqslant 0\right\}$ is a fractional Brownian motion of Hurst index $H \in (0, 1)$, whereas $\theta > 0$ and $\rho > 0$ are two unknown parameters such that $\theta \neq \rho$.

Using the notation (40), the explicit solution to this linear system, noted for instance in [14], implies the following decomposition of X_t:

$$X_t = \frac{\rho}{\rho - \theta} X_t^\rho + \frac{\theta}{\theta - \rho} X_t^\theta, \tag{48}$$

where

$$X_t^\theta = Z_t^\theta - e^{-\theta t} Z_0^\theta \tag{49}$$

with

$$Z_t^\theta = \int_{-\infty}^t e^{-\theta(t-s)} dB_s^H. \tag{50}$$

On the other hand, we can also write the system (47) as follows

$$dX_t = -(\theta + \rho) X_t dt - \rho\theta \Sigma_t dt + dB_t^H, \tag{51}$$

where for $0 \leqslant t \leqslant T$

$$\Sigma_t = \int_0^t X_s ds = \frac{V_t - X_t}{\theta} = \frac{X_t^\theta - X_t^\rho}{\rho - \theta}. \tag{52}$$

Moreover, the process $\left(Z_t^{\theta}, Z_t^{\theta'}\right)$ is an ergodic stationary Gaussian process. As consequence

$$X_t = \frac{\rho}{\rho-\theta} Z_t^{\rho} + \frac{\theta}{\theta-\rho} Z_t^{\theta} - \left(\frac{\rho e^{-\rho t}}{\rho-\theta} Z_0^{\rho} + \frac{\theta e^{-\theta t}}{\theta-\rho} Z_0^{\theta}\right)$$
$$:= \quad Z_t^{\theta,\rho} - \left(\frac{\rho e^{-\rho t}}{\rho-\theta} Z_0^{\rho} + \frac{\theta e^{-\theta t}}{\theta-\rho} Z_0^{\theta}\right) \tag{53}$$

and

$$\Sigma_t = \frac{Z_t^{\theta} - Z_t^{\rho}}{\rho-\theta} - \frac{e^{-\theta t} Z_0^{\theta} - e^{-\rho t} Z_0^{\rho}}{\rho-\theta} := \Sigma_t^{\theta,\rho} - \frac{e^{-\theta t} Z_0^{\theta} - e^{-\rho t} Z_0^{\rho}}{\rho-\theta}. \tag{54}$$

Moreover, $Z^{\theta,\rho}$ and $\Sigma^{\theta,\rho}$ are ergodic stationary Gaussian processes.

Now, assume that the processes X and Σ are observed equidistantly in time with the step size $\Delta_n = 1$. We will construct estimators for (θ, ρ). By using the ergodicity of $Z^{\theta,\rho}$ and $\Sigma^{\theta,\rho}$, Lemma 20 and Theorem 8, we conclude that

$$\left(Q_{f_q,n}(X), Q_{f_q,n}(\Sigma)\right) \longrightarrow \left(\lambda_{f_q}(Z^{\theta,\rho}), \lambda_{f_q}(\Sigma^{\theta,\rho})\right)$$

almost surely as $n \to \infty$.

Moreover, by the Gaussian property of $Z^{\theta,\rho}$ and $\Sigma^{\theta,\rho}$, and the expressions $\eta_X(\theta,\rho)$ and $\eta_\Sigma(\theta,\rho)$ for the variances of $Z^{\theta,\rho}$ and $\Sigma^{\theta,\rho}$ which are given respectively in (66) and (67) after Lemma 36 in the Appendix, we can write

$$\left(\lambda_{f_q}(Z^{\theta,\rho}), \lambda_{f_q}(\Sigma^{\theta,\rho})\right) = \delta_{f_q}(\theta,\rho)$$

where δ_{f_q} is a function which can be expressed via $\eta_X(\theta,\rho)$ and $\eta_\Sigma(\theta,\rho)$. Hence, in the case when the function δ_{f_q} is invertible, we obtain the following estimator for θ

$$(\widehat{\theta}_{f_q,n}, \widehat{\rho}_{f_q,n}) := \delta_{f_q}^{-1}\left[\left(Q_{f_q,n}(X), Q_{f_q,n}(\Sigma)\right)\right]. \tag{55}$$

Proposition 27. *Assume $H \in (0,1)$ and δ_{f_q} is a homomorphism. Let $(\widehat{\theta}_{f_q,n}, \widehat{\rho}_{f_q,n})$ be the estimator given in (55). Then, as $n \longrightarrow \infty$*

$$(\widehat{\theta}_{f_q,n}, \widehat{\rho}_{f_q,n}) \longrightarrow (\theta,\rho) \tag{56}$$

almost surely.

Examples. In the two following examples, the function δ_{f_q} is invertible and explicit, based on the expressions for $\eta_X(\theta,\rho)$ and $\eta_\Sigma(\theta,\rho)$ given respectively in (66) and (67) in the Appendix.

- Suppose that $f_q = H_q$. Using (20), (66) and (67), we have

$$\delta_{H_q}(\theta,\rho) = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left((\eta_X(\theta,\rho)-1)^{q/2}, (\eta_\Sigma(\theta,\rho)-1)^{q/2}\right).$$

- Suppose that $f_q = \phi_q$ with $\phi_q(x) = x^q$. From (21), (66) and (67) we obtain

$$\delta_{\phi_q}(\theta,\rho) = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left((\eta_X(\theta,\rho))^{q/2}, (\eta_\Sigma(\theta,\rho))^{q/2}\right).$$

Theorem 28. *Let $H \in \left(0, \frac{3}{4}\right)$. Define*

$$\Gamma_{f_q}(\theta,\rho) = \begin{pmatrix} u_{f_q}(Z^{\theta,\rho}) & u_{f_q}(Z^{\theta,\rho}, \Sigma^{\theta,\rho}) \\ u_{f_q}(Z^{\theta,\rho}, \Sigma^{\theta,\rho}) & u_{f_q}(\Sigma^{\theta,\rho}) \end{pmatrix} \tag{57}$$

where

$$u_{f_q}(Z^{\theta,\rho}, \Sigma^{\theta,\rho}) = \sum_{k=0}^{q/2} d^2_{f_q,2k}(2k)! \sum_{j\in\mathbb{Z}^*} \left(\frac{E\left(Z_0^{\theta,\rho}\Sigma_j^{\theta,\rho}\right)}{\sqrt{r_{Z^{\theta,\rho}}(0) r_{\Sigma^{\theta,\rho}}(0)}} \right)^{2k}.$$

Then

$$d_W\left(\left(U_{f_q,n}(X), U_{f_q,n}(\Sigma)\right); \mathcal{N}\left(0, \Gamma_{f_q}(\theta,\rho)\right)\right) \leqslant \frac{C}{n^{1/4}}. \tag{58}$$

Hence, for any $H \in \left(0, \frac{3}{4}\right)$,

$$\sqrt{n}\left(\widehat{\theta}_{f_q,n} - \theta, \widehat{\rho}_{f_q,n} - \rho\right) \xrightarrow{\mathcal{L}} \mathcal{N} \times \left(0,\; J_{\delta^{-1}_{f_q}}(\eta_X(\theta,\rho), \eta_\Sigma(\theta,\rho))\, \Gamma_{f_q}(\theta,\rho)\, J^T_{\delta^{-1}_{f_q}}(\eta_X(\theta,\rho), \eta_\Sigma(\theta,\rho))\right) \tag{59}$$

where $J_{\delta^{-1}_{f_q}}$ *is the Jacobian matrix of* $\delta^{-1}_{f_q}$.

Proof. Combining (53), (54), Lemma 36 and Theorem 11, we obtain (58). Applying Taylor's formula we can write

$$\sqrt{n}\left(\widehat{\theta}_{q,n} - \theta, \widehat{\rho}_{q,n} - \rho\right) = J^T_{\delta^{-1}_{f_q}}\left(\lambda_{f_q}(Z^{\theta,\rho}), \lambda_{f_q}(\Sigma^{\theta,\rho})\right)\left(U_{f_q,n}(X), U_{f_q,n}(\Sigma)\right) + d_n,$$

where d_n converges in distribution to zero, because

$$\|d_n\| \leqslant C\sqrt{n}\left\|\left(Q_{f_q,n}(X) - \lambda_{f_q}(Z^{\theta,\rho}), Q_{f_q,n}(\Sigma) - \lambda_{f_q}(\Sigma^{\theta,\rho})\right)\right\|^2 \longrightarrow 0$$

almost surely as $n \to \infty$ by using (58). Thus the 2-d random vector in the left-hand side of (59) is the sum of a term converging in law to 0 and another converging almost surely to 0; thus it converges in law to 0, establishing (59). ∎

Example. Here we assume that $f_q = \phi_q$ and $q = 2$, and we can recompute the expression for the function $\delta_{\phi_2} : (0,+\infty)^2 \mapsto (0,+\infty)^2$ as

$$\begin{aligned}\delta_{\phi_2}(x,y) &= (\eta_X(x,y), \eta_\Sigma(x,y)) \\ &= H\Gamma(2H) \times \begin{cases} \frac{1}{y^2 - x^2}\left(y^{2-2H} - x^{2-2H}, x^{-2H} - y^{-2H}\right) & \text{if } x \neq y \\ \left((1-H)x^{-2H}, Hx^{-2H-2}\right) & \text{if } y = x. \end{cases}\end{aligned}$$

Since for every $(x,y) \in (0,+\infty)^2$ with $x \neq y$ the Jacobian of δ_{ϕ_2} computes as

$$J_{\delta_{\phi_2}}(x,y) = \Gamma(2H+1) \times \begin{pmatrix} \frac{(1-H)\,x^{1-2H}\left(x^2 - y^2\right) - x\left(x^{2-2H} - y^{2-2H}\right)}{\left(x^2 - y^2\right)^2} & \frac{(1-H)\,y^{1-2H}\left(y^2 - x^2\right) - y\left(y^{2-2H} - x^{2-2H}\right)}{\left(x^2 - y^2\right)^2} \\ \frac{Hx^{-2H-1}\left(x^2 - y^2\right) + x\left(x^{-2H} - y^{-2H}\right)}{\left(x^2 - y^2\right)^2} & \frac{Hy^{-2H-1}\left(y^2 - x^2\right) + y\left(y^{-2H} - x^{-2H}\right)}{\left(x^2 - y^2\right)^2} \end{pmatrix},$$

which is non-zero in $(0,+\infty)^2$. So δ_{ϕ_2} is a diffeomorphism in $(0,+\infty)^2$ and its inverse $\delta_{\phi_2}^{-1}$ has a Jacobian

$$J_{\delta_{\phi_2}^{-1}}(a,b) = \frac{\Gamma(2H+1)}{\det J_{F_2}(x,y)} \times \begin{pmatrix} \frac{Hy^{-2H-1}\left(y^2-x^2\right)+y\left(y^{-2H}-x^{-2H}\right)}{\left(x^2-y^2\right)^2} & -\frac{(1-H)\,y^{1-2H}\left(y^2-x^2\right)-y\left(y^{2-2H}-x^{2-2H}\right)}{\left(x^2-y^2\right)^2} \\ -\frac{Hx^{-2H-1}\left(x^2-y^2\right)+x\left(x^{-2H}-y^{-2H}\right)}{\left(x^2-y^2\right)^2} & \frac{(1-H)\,x^{1-2H}\left(x^2-y^2\right)-x\left(x^{2-2H}-y^{2-2H}\right)}{\left(x^2-y^2\right)^2} \end{pmatrix};$$

where $(x,y)=\delta_{\phi_2}^{-1}(a,b)$. Thus the asymptotic covariance matrix in (59) is explicit. Moreover, similarly to the results obtained in Section 5, we can prove the following, all details being omitted.

Proposition 29. *Let $(\alpha,\beta)\in\mathbf{R}^2$. Under the assumptions and notation of Theorem 28,*

- *if $H\in(0,\frac{5}{8})$,*

$$d_W\left(\alpha U_{\phi_2,n}(X)+\beta U_{\phi_2,n}(\Sigma);\mathcal{N}\left(0,(\alpha,\beta)\,\Gamma_{\phi_2}(\theta,\rho)(\alpha,\beta)^{Tr}\right)\right)\asymp\frac{1}{\sqrt{n}},$$

- *if $H\in(\frac{5}{8},\frac{3}{4})$,*

$$d_W\left(\alpha U_{\phi_2,n}(X)+\beta U_{\phi_2,n}(\Sigma);\mathcal{N}\left(0,(\alpha,\beta)\,\Gamma_{\phi_2}(\theta,\rho)(\alpha,\beta)^{Tr}\right)\right)\leqslant\frac{C}{n^{4H-3}}.$$

6.2. *Fractional Ornstein–Uhlenbeck process of the second kind*

The last example we consider is the so-called fractional Ornstein–Uhlenbeck process of the second kind, defined via the stochastic differential equation

$$S_0=0,\ \text{and}\ dS_t=-\alpha S_t dt+dY_t^{(1)},\quad t\geqslant 0, \tag{60}$$

where $Y_t^{(1)}=\int_0^t e^{-s}dB_{a_s}^H$ with $a_s=He^{\frac{s}{H}}$ and $B^H=\left\{B_t^H,t\geqslant 0\right\}$ is a fractional Brownian motion with Hurst parameter $H\in(\frac{1}{2},1)$, and where $\alpha>0$ is the unknown real parameter which we would like to estimate. Eq. (60) admits an explicit solution

$$S_t=e^{-\alpha t}\int_0^t e^{\alpha s}dY_s^{(1)}=H^{(1-\alpha)H}e^{-\alpha t}\int_{a_0}^{a_t}r^{(\alpha-1)H}dB_r^H.$$

Hence we can also write

$$S_t=S_t^\alpha-e^{-\alpha t}S_0^\alpha$$

where

$$S_t^\alpha=H^{(1-\alpha)H}e^{-\alpha t}\int_0^{a_t}r^{(\alpha-1)H}dB_r^H.$$

Using a similar argument to that in Lemma 20 we have for every $p\geqslant 1$ and for all $n\in\mathbb{N}$,

$$\left\|Q_{f_q,n}(S)-Q_{f_q,n}(S^\alpha)\right\|_{L^p(\Omega)}=\mathcal{O}\left(n^{-1}\right). \tag{61}$$

As consequence, by using S^α ergodic and (61) we conclude that, almost surely as $n\to\infty$,

$$Q_{f_q,n}(S)\longrightarrow\lambda_{f_q}(S^\alpha).$$

Moreover, by the Gaussian property of U^α and (69) we can write

$$\lambda_{f_q}(S^\alpha) := \nu_{f_q}(\alpha),$$

where ν_{f_q} is a function. Hence, in the case when the function ν_{f_q} is a homeomorphism, we obtain the following strongly consistent estimator for α

$$\widehat{\alpha}_{f_q,n} := \nu_{f_q}^{-1}\left[Q_{f_q,n}(S)\right]. \tag{62}$$

Proposition 30. *Assume $H \in \left(\frac{1}{2}, 1\right)$ and ν_{f_q} is a homeomorphism. Let $\widehat{\alpha}_{f_q,n}$ be the estimator given in (62). Then, almost surely as $n \longrightarrow \infty$*

$$\widehat{\alpha}_{f_q,n} \longrightarrow \alpha.$$

Examples. In the two following examples, the function ν_{f_q} is homeomorphic and explicit.

- Suppose that $f_q = H_q$. Using (20) and (69), we have
$$\nu_{H_q}(\alpha) = \lambda_{H_q}(S^\alpha) = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left(\frac{(2H-1)H^{2H}}{\alpha}\beta(1-H+\alpha H, 2H-1) - 1\right)^{q/2}.$$
- Suppose that $f_q = \phi_q$ with $\phi_q(x) = x^q$. From (21) and (69) we obtain $\nu_{\phi_q}(\alpha) = \lambda_{\phi_q}(S^\alpha) = \frac{q!}{(\frac{q}{2})!2^{q/2}}\left[\frac{(2H-1)H^{2H}}{\alpha}\beta(1-H+\alpha H, 2H-1)\right]^{q/2}$.
- The reader will check that in both cases above, the function $\alpha \mapsto v(\alpha)$ is monotone (decreasing) and convex from $\mathbf{R}_+$ to $\mathbf{R}_+$, and that the moment condition of Theorem 19 on $\left(v^{-1}\right)''$ is satisfied.

Now, we study the asymptotic distribution of $\widehat{\alpha}_{f_q,n}$. By (70) which is established in Lemma 37 in the Appendix, we have for every $H \in (\frac{1}{2}, 1)$, $\sum_{j\in\mathbb{Z}}|r_{S^\alpha}(j)|^2 < \infty$ and $\kappa_4(U_{f_q,n}(S^\alpha)) = \mathcal{O}(\frac{1}{n})$. Thus, applying (23) we deduce the following result.

Proposition 31. *Suppose that $H \in (\frac{1}{2}, 1)$ and $\alpha > 0$. Then*

$$d_W\left(u_{f_q}(S^\alpha)^{-1/2}U_{f_q,n}(S), N\right) \leqslant Cn^{-\frac{1}{4}}.$$

In particular,

$$\sqrt{n}\left(\widehat{\alpha}_{f_q,n} - \alpha\right) \overset{law}{\longrightarrow} \mathcal{N}\left(0, u_{f_q}(S^\alpha)\left(\left(\nu_{f_q}^{-1}\right)'(\alpha)\right)^{-2}\right).$$

Quadratic case. In this case we can improve the rate convergence of $\widehat{\alpha}_{f_2,n}$. By using Theorem 15, the estimates $\kappa_4(U_{f_2,n}(S^\alpha)) = \mathcal{O}(\frac{1}{n})$ and $\left|E\left((F_{f_2,n}(S^\alpha))^3\right)\right| = \mathcal{O}(\frac{1}{\sqrt{n}})$, and invoking the properties of u_{f_2} described in the examples (bullet points) above to invoke Theorem 19, we get the following.

Proposition 32. *Let $H \in \left(\frac{1}{2}, 1\right)$. Then*

$$d_W\left(u_{f_2}(S^\alpha)^{-1/2}U_{f_2,n}(S), N\right) \asymp \frac{1}{\sqrt{n}},$$

and

$$d_W\left(\sqrt{n}\left(\widehat{\alpha}_{f_2,n} - \alpha\right), \mathcal{N}(0, u_{f_2}(S^\alpha))\left(\left(\nu_{f_2}^{-1}\right)'(\alpha)\right)^{-2}\right) \leqslant \frac{C}{\sqrt{n}}.$$

Acknowledgments

F.V.'s research was partially supported by the U.S. National Science Foundation under award numbers DMS 1407762 and DMS 1734183.

Appendix

The following result is a well-known direct consequence of the Borel–Cantelli Lemma (see e.g. [19]).

Lemma 33. *Let $\gamma > 0$ and let $(Z_n)_{n\in\mathbb{N}}$ be a sequence of random variables. If for every $p \geqslant 1$ there exists a constant $c_p > 0$ such that for all $n \in \mathbb{N}$,*

$$\|Z_n\|_{L^p(\Omega)} \leqslant c_p \cdot n^{-\gamma},$$

then for all $\varepsilon > 0$ there exists a random variable η_ε such that

$$|Z_n| \leqslant \eta_\varepsilon \cdot n^{-\gamma+\varepsilon} \quad \textit{almost surely}$$

for all $n \in \mathbb{N}$. Moreover, $\mathbb{E}|\eta_\varepsilon|^p < \infty$ for all $p \geqslant 1$.

Proof of Theorem 1. Let us first prove (3). By (2) we have

$$d_{TV}\left(\frac{F}{\sqrt{E\left[F^2\right]}}, N\right) \leqslant \frac{2}{E\left[F^2\right]} E\left|E\left[F^2\right] - \left\langle DF, -DL^{-1}F\right\rangle_{\mathcal{H}}\right|.$$

On the other hand, exploiting the fact that

$$E\left[E\left[(I_k(g_k))^2\right] - \langle DI_k(g_k), -DL^{-1}I_k(g_k)\rangle_{\mathcal{H}}\right] = 0, \quad \text{and}$$
$$E\left(I_k(g_k)I_l(g_l)\right) = 0 \ \text{ if } k \neq l,$$

we obtain

$$E\left|E\left[F^2\right] - \left\langle DF, -DL^{-1}F\right\rangle_{\mathcal{H}}\right| \leqslant \sum_{k=2}^{q} \sqrt{Var\left(\frac{1}{k}\|DI_k(g_k)\|^2_{\mathcal{H}}\right)} + \sum_{2\leqslant k\neq l\leqslant q} \frac{1}{l} E\left|\langle DI_k(g_k), DI_l(g_l)\rangle_{\mathcal{H}}\right|.$$

Moreover, by [25, Lemma 3.1] we have

$$Var\left(\frac{1}{k}\|DI_k(g_k)\|^2_{\mathcal{H}}\right) = \frac{1}{k^2}\sum_{j=1}^{k-1} j^2 j!^2 \binom{k}{j}^4 (2k-2j)! \|g_k \otimes_j g_k\|^2_{\mathcal{H}^{\otimes 2k-2j}},$$

and for $k < l$

$$E\left[\left(\frac{1}{l}\langle DI_k(g_k), DI_l(g_l)\rangle_{\mathcal{H}}\right)^2\right] \leqslant (k)!\binom{l-1}{k-1}^2 (l-k)! E\left[(I_k(g_k))^2\right] \|g_l \otimes_{l-k} g_l\|_{\mathcal{H}^{\otimes 2k}}$$
$$+ \frac{k^2}{2}\sum_{j=1}^{k-1} (l-1)!^2 \binom{k-1}{j-1}^2 \binom{l-1}{j-1}^2 (k+l-2j)! \left(\|g_k \otimes_{k-j} g_k\|^2_{\mathcal{H}^{\otimes 2j}} + \|g_l \otimes_{l-j} g_l\|^2_{\mathcal{H}^{\otimes 2j}}\right).$$

Therefore

$$E\left|E\left[F^2\right]-\langle DF,-DL^{-1}F\rangle_{\mathcal{H}}\right| \leqslant C_q\sqrt{\max_{2\leqslant k\leqslant q}\|g_k\|^2_{\mathcal{H}^{\otimes k}}\sqrt{\max_{\substack{1\leqslant s\leqslant k\\ 2\leqslant k\leqslant q}}\|g_k\otimes_s g_k\|^2_{\mathcal{H}^{\otimes 2k-2s}}}+\max_{\substack{1\leqslant s\leqslant k\\ 2\leqslant k\leqslant q}}\|g_k\otimes_s g_k\|^2_{\mathcal{H}^{\otimes 2k-2s}}}.$$

Moreover, since $\|g_k\otimes_s g_k\|_{\mathcal{H}^{\otimes 2k-2s}}\leqslant\|g_k\|^2_{\mathcal{H}^{\otimes k}}$ for all $1\leqslant s\leqslant k$, we get

$$E\left|E\left[F^2\right]-\langle DF,-DL^{-1}F\rangle_{\mathcal{H}}\right|\leqslant C_q\max_{2\leqslant k\leqslant q}\|g_k\|_{\mathcal{H}^{\otimes k}}\max_{\substack{1\leqslant s\leqslant k\\ 2\leqslant k\leqslant q}}\|g_k\otimes_s g_k\|^{\frac12}_{\mathcal{H}^{\otimes 2k-2s}}, \tag{63}$$

which implies (3).

Now, let us prove (4). Since, for any $k\in\mathbb{Z}$, $|r_Z(k)|\leqslant r_Z(0)$, we have

$$\begin{aligned}\|g_k\|^2_{\mathcal{H}^{\otimes k}}&=\frac{d_k^2}{n}\sum_{i,j=0}^{n-1}\left(\frac{r_Z(i-j)}{r_Z(0)}\right)^k\leqslant\frac{d_k^2}{n}\sum_{i,j=0}^{n-1}\left(\frac{r_Z(i-j)}{r_Z(0)}\right)^2=\frac{d_k^2}{d_2^2}\|g_2\|^2_{\mathcal{H}^{\otimes 2}}\\&=\frac{d_k^2}{d_2^2}Var(I_2(g_2)).\end{aligned}$$

On the other hand, for every $1\leqslant s\leqslant k-1$ with $k\in\{2,\dots,q\}$,

$$\begin{aligned}&\|g_k\otimes_s g_k\|^2_{\mathcal{H}^{\otimes 2k-2s}}\\&\leqslant d_k^4n^{-2}\sum_{k_1,k_2,k_3,k_4=0}^{n-1}\left(\frac{r_Z(k_1-k_2)}{r_Z(0)}\right)^s\left(\frac{r_Z(k_3-k_4)}{r_Z(0)}\right)^s\left(\frac{r_Z(k_1-k_3)}{r_Z(0)}\right)^{2k-s}\\&\quad\times\left(\frac{r_Z(k_2-k_4)}{r_Z(0)}\right)^{2k-s}\\&\leqslant d_k^4n^{-2}r_Z(0)^{-4}\sum_{k_1,k_2,k_3,k_4=0}^{n-1}r_Z(k_1-k_2)r_Z(k_3-k_4)r_Z(k_1-k_3)r_Z(k_2-k_4)\\&=\frac{d_k^4}{d_2^4}\kappa_4(I_2(g_2)).\end{aligned}$$

Therefore, using (63),

$$E\left|E\left[F^2\right]-\langle DF,-DL^{-1}F\rangle_{\mathcal{H}}\right|\leqslant C_q\sqrt{Var(I_2(g_2))\sqrt{\kappa_4(I_2(g_2))}}.$$

Furthermore, by applying again (2) we can write

$$\begin{aligned}d_{TV}\left(\frac{F}{\sigma},N\right)&\leqslant\frac{2}{\sigma^2}E\left|\sigma^2-\langle DF,-DL^{-1}F\rangle_{\mathcal{H}}\right|\\&\leqslant\frac{2}{\sigma^2}E\left|E\left[F^2\right]-\langle DF,-DL^{-1}F\rangle_{\mathcal{H}}\right|+2E\left|1-\frac{E[F^2]}{\sigma^2}\right|.\end{aligned}$$

Thus the estimate (4) is obtained. The upper bound of the estimate (5) is proved in [4, Proposition 6.4]. ■

Proof of Lemma 14. An inspection of the proof of the main lower bound result in [24] shows that their lower bound on $d_{TV}(F_n, N)$ is in fact a lower bound on

$$\frac{1}{2}\max\{|E(\cos F_n) - E(\cos N)|\,;\,|E(\sin F_n) - E(\sin N)|\}.$$

Since sin and cos are 1-Lipschitz functions, by definition of d_W, this expression is also a lower bound on $\frac{1}{2}d_W(F_n, N)$. This proves the lemma. ■

Lemma 34. *Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1]$, $m, m' > 0$ and $-\infty \leqslant a < b \leqslant c < d < \infty$. Then*

$$E\left(\int_a^b e^{ms} dB^H(s) \int_c^d e^{m't} dB^H(t)\right) = H(2H-1)\int_a^b ds e^{ms}\int_c^d dt e^{m't}(t-s)^{2H-2}.$$

Proof. We use the same argument as in the proof of [9, Lemma 2.1]. ■

Lemma 35. *Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, $m, m' > 0$ and let Z^θ be the process defined in (40). Then,*

$$r_{Z^\theta}(0) = H\Gamma(2H)\theta^{-2H} \qquad \text{and} \qquad r_{Z^\theta}(t) \sim \frac{H(2H-1)}{\theta^2}|t|^{2H-2}$$

for large $|t|$.

Proof. see [9, Theorem 2.3] or Lemma 36. ■

Lemma 36. *Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, $m, m' > 0$ and let Z^m be the process defined in (50). Then,*

$$E\left[Z_0^m Z_0^{m'}\right] = \frac{H\Gamma(2H)}{m+m'}\left(m^{1-2H} + (m')^{1-2H}\right) \tag{64}$$

and for large $|t|$

$$E\left[Z_0^m Z_t^{m'}\right] \sim \frac{H(2H-1)}{mm'}|t|^{2H-2}. \tag{65}$$

This implies that for $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$

$$\eta_X(\theta,\rho) := E\left[\left(Z_0^{\theta,\rho}\right)^2\right] = \frac{H\Gamma(2H)}{\rho^2-\theta^2}[\rho^{2-2H} - \theta^{2-2H}], \tag{66}$$

$$\eta_\Sigma(\theta,\rho) := E\left[\left(\Sigma_0^{\theta,\rho}\right)^2\right] = \frac{H\Gamma(2H)}{\rho^2-\theta^2}[\theta^{-2H} - \rho^{-2H}], \tag{67}$$

and for every $t > 0$

$$E\left[\left(Z_t^{\theta,\rho}\Sigma_t^{\theta,\rho}\right)\right] = E\left[\left(Z_0^{\theta,\rho}\Sigma_0^{\theta,\rho}\right)\right] = 0. \tag{68}$$

Proof. By using [9, Proposition A.1], we can write

$$\begin{aligned} E\left[Z_0^m Z_0^{m'}\right] &= mm'\int_{-\infty}^0\int_{-\infty}^0 e^{mu}e^{m'v}E\left(B_u^H B_v^H\right)dudv \\ &= \frac{mm'}{2}\int_0^\infty\int_0^\infty e^{mu}e^{m'v}\left(u^{2H} + v^{2H} - |v-u|^{2H}\right)dudv \\ &= \frac{\Gamma(2H+1)}{2(m+m')}\left(m^{1-2H} + (m')^{1-2H}\right). \end{aligned}$$

Thus the estimate (64) is proved. Now, let $0 < \varepsilon < 1$

$$
\begin{aligned}
E\left(Z_0^m Z_t^{m'}\right) &= e^{-m't} E\left(\int_{-\infty}^0 e^{mu} dB_u^H \int_{-\infty}^t e^{m'v} dB_v^H\right) \\
&= e^{-m't} E\left(\int_{-\infty}^0 e^{mu} dB_u^H \int_{-\infty}^{\varepsilon t} e^{m'v} dB_v^H\right) \\
&\quad + e^{-m't} E\left(\int_{-\infty}^0 e^{mu} dB_u^H \int_{\varepsilon t}^t e^{m'v} dB_v^H\right) \\
&:= A + B
\end{aligned}
$$

where, using [9, Proposition A.1] it is easy to see that $|A| = O\left(e^{-m't}\right)$. On the other hand, by Lemma 34 and integration by parts and linear changes of variables

$$
\begin{aligned}
B &= H(2H-1)e^{-m't} \int_{-\infty}^0 du\, e^{mu} \int_{\varepsilon t}^t dv\, e^{m'v} (v-u)^{2H-2} \\
&= \frac{H(2H-1)}{m+m'} \left(\int_t^\infty e^{-m(z-t)} z^{2H-2} dz + \int_{\varepsilon t}^t e^{-m'(t-z)} z^{2H-2} dz \right. \\
&\quad \left. + e^{-m't(1-\varepsilon)} \int_{\varepsilon t}^\infty e^{-m(z-\varepsilon t)} z^{2H-2} dz \right) \\
&= \frac{H(2H-1)}{(m+m')} \left(\frac{t^{2H-2}}{m} + \frac{2H-2}{m} \int_t^\infty e^{-m(z-t)} z^{2H-3} dz + \frac{t^{2H-2}}{m'} \right. \\
&\quad - \frac{(\varepsilon t)^{2H-2}}{m'} e^{-m'(1-\varepsilon)t} - \frac{2H-2}{m'} \int_{\varepsilon t}^t e^{-m'(t-z)} z^{2H-3} dz \\
&\quad \left. + e^{-m't(1-\varepsilon)} \int_{\varepsilon t}^\infty e^{-m(z-\varepsilon t)} z^{2H-2} dz \right) \\
&= \frac{H(2H-1)}{mm'} t^{2H-2} + o\left(t^{2H-2}\right),
\end{aligned}
$$

the last inequality coming from the fact that

$$
\int_t^\infty e^{-m(z-t)} z^{2H-3} dz \leqslant t^{-1} \int_0^\infty e^{-my} dy \to 0, \quad \text{as } t \to \infty,
$$

$$
\begin{aligned}
t^{2-2H} \int_{\varepsilon t}^t e^{-m'(t-z)} z^{2H-3} dz &\leqslant \varepsilon^{2H-3} t^{-1} \int_{\varepsilon t}^t e^{-m'(t-z)} dz \\
&= \varepsilon^{2H-3} t^{-1} \int_0^{(1-\varepsilon)t} e^{-m'y} dy \to 0, \quad \text{as } t \to \infty,
\end{aligned}
$$

and

$$
t^{2-2H} e^{-m't(1-\varepsilon)} \to 0, \quad \text{as } t \to \infty.
$$

So, we conclude that the estimate (65) is obtained. ■

Lemma 37. *Let $H \in (\frac{1}{2}, 1)$. Then,*

$$
E\left[\left(S_0^\alpha\right)^2\right] = \frac{(2H-1)H^{2H}}{\alpha} \beta(1 - H + \alpha H, 2H - 1). \tag{69}
$$

and for large $|t|$

$$
r_{S^\alpha}(t) = E\left[S_0^\alpha S_t^\alpha\right] = O\left(e^{-\min\{\alpha, \frac{1-H}{H}\}t}\right). \tag{70}
$$

Proof. We prove the first point (69). We have

$$\begin{aligned} E\left[\left(S_0^\alpha\right)^2\right] &= 2H(2H-1)H^{2(1-\alpha)H}\int_0^{a_0} dy y^{(\alpha-1)H}\int_0^y dx\, x^{(\alpha-1)H}(y-x)^{2H-2} \\ &= 2H(2H-1)H^{2(1-\alpha)H}\int_0^{a_0} dy y^{2\alpha H-1}\int_0^1 dz\, z^{(\alpha-1)H}(1-z)^{2H-2} \\ &= \frac{(2H-1)H^{2H}}{\alpha}\beta(1-H+\alpha H, 2H-1). \end{aligned}$$

Thus (69) is obtained. For the point (70) see [17]. ■

References

[1] E. Azmoodeh, G.I. Morlanes, Drift parameter estimation for fractional Ornstein–Uhlenbeck process of the second kind, Statistics (2013), http://dx.doi.org/10.1080/02331888.2013.863888.

[2] E. Azmoodeh, L. Viitasaari, Parameter estimation based on discrete observations of fractional Ornstein–Uhlenbeck process of the second kind, Stat. Inference Stoch. Process. 18 (3) (2015) 205–227.

[3] R. Belfadli, K. Es-Sebaiy, Y. Ouknine, Parameter estimation for fractional Ornstein–Uhlenbeck processes: Non-ergodic case, Front. Sci. Eng. Int. J. Hassan II Acad. Sci. Technol. 1 (1) (2011) 1–16.

[4] H. Biermé, A. Bonami, I. Nourdin, G. Peccati, Optimal Berry-Esséen rates on the Wiener space: the barrier of third and fourth cumulants, ALEA 9 (2) (2012) 473–500.

[5] V. Bogachev, E. Kosov, G. Zelenov, Fractional smoothness of distributions of polynomials and a fractional analog of the Hardy-Landau-Littlewood inequality, Trans. Amer. Math. Soc. 370 (6) (2018) 4401–4432.

[6] J.C. Breton, I. Nourdin, Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion, Electron. Commun. Probab. 13 (2008) 482–493.

[7] A. Brouste, S.M. Iacus, Parameter estimation for the discretely observed fractional Ornstein–Uhlenbeck process and the Yuima R package, Comput. Statist. 28 (4) (2012) 1529–1547.

[8] P. Cénac, K. Es-Sebaiy, Almost sure central limit theorems for random ratios and applications to LSE for fractional Ornstein–Uhlenbeck processes, Probab. Math. Statist. 35 (2) (2015) 285–300.

[9] P. Cheridito, H. Kawaguchi, M. Maejima, Fractional Ornstein–Uhlenbeck processes, Electron. J. Probab. 8 (2003) 1–14.

[10] Y.A. Davydov, G.V. Martynova, Limit behavior of multiple stochastic integral, in: Statistics and Control of Random Process, Preila, Nauka, Moscow, 1987, pp. 55–57 (in Russian).

[11] R.L. Dobrushin, P. Major, Non-central limit theorems for nonlinear functionals of Gaussian fields, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 50 (1979) 27–52.

[12] M. El Machkouri, K. Es-Sebaiy, Y. Ouknine, Least squares estimator for non-ergodic Ornstein–Uhlenbeck processes driven by Gaussian processes, J. Korean Stat. Soc. 45 (2016) 329–341.

[13] B. El Onsy, K. Es-Sebaiy, C. Tudor, Statistical analysis of the non-ergodic fractional Ornstein–Uhlenbeck process of the second kind, Commun. Stoch. Anal. 11 (2) (2017) 119–136.

[14] B. El Onsy, K. Es-Sebaiy, F. Viens, Parameter Estimation for Ornstein–Uhlenbeck driven by fractional Ornstein–Uhlenbeck processes, Stochastics 89 (2) (2017) 431–468.

[15] Y. Hu, D. Nualart, Parameter estimation for fractional Ornstein–Uhlenbeck processes, Statist. Probab. Lett. 80 (2010) 1030–1038.

[16] Y. Hu, J. Song, Parameter estimation for fractional Ornstein–Uhlenbeck processes with discrete observations, in: F. Viens, et al. (Eds.), Malliavin Calculus and Stochastic Analysis: A Festschrift in Honor of David Nualart, Springer, 2013, pp. 427–442.

[17] T. Kaarakka, P. Salminen, On fractional Ornstein–Uhlenbeck Processes, Commun. Stoch. Anal. 5 (1) (2011) 121–133.

[18] M. Kleptsyna, A. Le Breton, Statistical analysis of the fractional Ornstein- Uhlenbeck type process, Stat. Inference Stoch. Process. 5 (2002) 229–241.

[19] P. Kloeden, A. Neuenkirch, The pathwise convergence of approximation schemes for stochastic differential equations, LMS J. Comput. Math. 10 (2007) 235–253.

[20] H. Luschgy, G. Pagès, Functional quantization of Gaussian processes, J. Funct. Anal. 196 (2) (2002) 486–531.

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

[22] L. Neufcourt, F. Viens, A third-moment theorem and precise asymptotics for variations of stationary Gaussian sequences, ALEA Lat. Am. J. Probab. Math. Stat. 13 (1) (2016) 239–264.

[23] I. Nourdin, G. Peccati, Normal Approximations with Malliavin Calculus : From Stein's Method to Universality, in: Cambridge Tracts in Mathematics, vol. 192, Cambridge University Press, Cambridge, 2012.

[24] I. Nourdin, G. Peccati, The optimal fourth moment theorem, Proc. Amer. Math. Soc. 143 (2015) 3123–3133.

[25] I. Nourdin, G. Peccati, M. Podolskij, Quantitative Breuer-Major Theorems, Stochastic Process. Appl. 121 (4) (2013) 793–812.

[26] I. Nourdin, G. Poly, Convergence in total variation on Wiener chaos, Stochastic Process. Appl. 123 (2013) 651–674.

[27] D. Nualart, The Malliavin Calculus and Related Topics, Springer-Verlag, Berlin, 2006.

[28] D. Nualart, G. Peccati, Central limit theorems for sequences of multiple stochastic integrals, Ann. Probab. 33 (1) (2005) 177–193.

[29] Orthogonal Polynomials, Classical Orthogonal Polynomials, Sums, in: Digital Library of Mathematical Functions, National Institute of Standards and Technology. Available at https://dlmf.nist.gov/18.18#E20.

[30] C. Tudor, F. Viens, Statistical aspects of the fractional stochastic calculus, Ann. Statist. 35 (3) (2007) 1183–1212.

[31] R. Zintout, The total variation distance between two double Wiener-Itô integrals, Statist. Probab. Lett. 83 (10) (2013) 2160–2167.