Variations and estimators for the selfsimilarity order through Malliavin calculus

Ciprian A. Tudor, Frederi Viens

Variations and estimators for the selfsimilarity order through Malliavin calculus

Ciprian A. Tudor[1] Frederi G. Viens[2]

[1]SAMOS-MATISSE, Centre d'Economie de La Sorbonne,
Université de Paris 1 Panthéon-Sorbonne ,
90, rue de Tolbiac, 75634, Paris, France.
tudor@univ-paris1.fr

[2] Dept. Statistics and Dept. Mathematics, Purdue University,
150 N. University St., West Lafayette, IN 47907-2067, USA.
viens@purdue.edu

September 25, 2007

Abstract

Using multiple stochastic integrals, we analyze the asymptotic behavior of quadratic variations for Gaussian and non-Gaussian selfsimilar processes. We apply our results to the study of statistical estimators for the selfsimilarity index.

2000 AMS Classification Numbers: 60F05, 60H05, 60G18.

Key words: multiple stochastic integrals, Malliavin calculus, Hermite process, fractional Brownian motion, non-central limit theorems, quadratic variations.

1 Introduction

A selfsimilar process is a stochastic process such that any part of its trajectory is invariant under time scaling. Selfsimilar processes are of considerable interest in practice in modeling various phenomena, including internet traffic (see e.g. [26]), hydrology (see e.g. [11]), or economics (see e.g. [10], [25]). In various applications, empirical data also shows strong correlation of observations, indicating the presence, in addition to self-similarity, of long-range dependence. We refer to the monographs [5] or [19] for various properties and fields of applications of such processes.

In this work we will focus our attention on a special class of self-similar processes which also exhibit long range dependence, which appear as limits in the so-called Non Central Limit Theorem, and belong to the class of so-called Hermite processes. We study the behavior of the quadratic variations for such processes extending recent results by [13], [14], [12] and we apply the results to the study of estimators for the selfsimilarity index both in the Gaussian and non-Gaussian case. The Hermite processes were introduced by Taqqu (see [21], [22]) and by Dobrushin and Major (see

[4]). The Hermite process of order $k \geq 1$ can be written for $t \geq 0$ as

$$Z_H^k(t) = c(H,k) \int_{\mathbb{R}^k} \int_0^t \left(\prod_{j=1}^k (s - y_i)_+^{-\left(\frac{1}{2} + \frac{1-H}{k}\right)} \right) ds dW(y_1) \dots dW(y_k), \tag{1}$$

where $x_+ = \max(x, 0)$, the self-similarity (Hurst) index H belongs to the interval $(\frac{1}{2}, 1)$ and the above integral is a multiple Wiener-Itô stochastic integral with respect to a Brownian motion $(W(y))_{y \in \mathbb{R}}$ (see [15]). Among its basic properties, we recall that

- it exhibits long-range dependence (the long-range covariances decays at the rate of the non-summable power function n^{2H-2})
- it is H-self-similar in the sense that for any $c > 0$, $\left(Z_H^k(ct)\right)_{t \geq 0} \overset{(d)}{=} \left(c^H Z_H^k(t)\right)_{t \geq 0}$, where $\overset{(d)}{=}$ means equivalence of all finite dimensional distributions
- it has stationary increments, that is, the distribution of $\left(Z_H^k(t+h) - Z_H^k(h)\right)_{t \geq 0}$ does not depend on $h > 0$.
- the covariance function is

  $$\mathbf{E}\left[Z_H^k(t) Z_H^k(s)\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \quad s, t \in [0, T];$$

  consequently, for every $s, t \in [0, T]$ the expected squared increment of the Hermite process is

  $$\mathbf{E}\left[\left(Z_H^k(t) - Z_H^k(s)\right)^2\right] = |t-s|^{2H}, \tag{2}$$

  from which it follows by Kolmogorov's continuity criterion that this process is almost-surely Holdër continuous of any order $\delta < H$.
- if $k \geq 2$, then Z_H^k is non-Gaussian; the k-th Hermite process lives in the so-called k-th Wiener chaos of B, since it is a k-th order Wiener integral; when $k = 2$, it is called the Rosenblatt process; if $k = 1$, the process is Gaussian, and one recognizes the covariance structure, as well as the moving-average representation w.r.t. a Wiener process, of the fractional Brownian motion (fBm), which is the best-known Hermite process.

The stochastic analysis of fBm has been intensively develop in recent years and its applications are many. Other Hermite processes are less studied, but are still of interest because of their long range dependence, self-similarity and stationarity of increments. The great popularity of fBm in modeling is due to these properties, and that one prefers fBm rather than higher order Hermite process because it is a Gaussian process, and its calculus is much easier. But in concrete situations when empirical data attests to the presence of self-similarity and long memory without the Gaussian property, one can use a Hermite process living in a higher chaos. The Hurst parameter H characterizes all the important properties of a Hermite process, as seen above. Therefore, estimating H properly is of the utmost importance. Several statistics have been introduced to this end, such as wavelets, k-variations, variograms, maximum likelihood estimators, or spectral methods. Information on these various approaches can be found in the book of Beran [1].

In this paper we will use the k-variations statistics to estimate H. Let us recall the context. Suppose that a process $(X_t)_{t \in [0,1]}$ is observed at discrete times $\{0, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\}$ and let a be

a "filter" of length $l \geq 0$ and $p \geq 1$ a fixed power; that is, a is an $l+1$ dimensional vector $a = (a_0, a_1, \ldots, a_l)$ such that $\sum_{q=0}^{l} a_q q^r = 0$ for $0 \leq r \leq p-1$ and $\sum_{q=0}^{l} a_q q^p \neq 0$. Then the k-variation statistic associated to the filter a is defined as

$$V_N(k,a) = \frac{1}{N-l} \sum_{i=l}^{N-1} \left[\frac{V_a\left(\frac{i}{N}\right)^k}{EV_a\left(\frac{i}{N}\right)^k} - 1 \right]$$

where

$$V_a\left(\frac{i}{N}\right) = \sum_{q=0}^{l} a_q X_{\frac{i-q}{N}}.$$

When X is fBm, these statistics are used to derive strongly consistent estimators for the Hurst index, and their associated normal convergence results. A detailed study can be found in [6], [9] or more recently in [3]. The behavior of $V_N(k,a)$ is used to derive similar behaviors for the corresponding estimators. The basic result is that, if $p > H + \frac{1}{4}$, then the renormalized k-variation* $V_N(k,a)$ converges to a standard normal distribution. The easiest and most natural case is that of the filter $a = \{1, -1\}$, in which case $p = 1$; one then has the restriction $H < \frac{3}{4}$. The techniques used to prove such convergence in the fBm case in the above references are strongly related to the Gaussian property of the observations; they appear not to extend to non-Gaussian situations.

Our purpose here is to develop new techniques that can be applied to both the fBm case and other non-Gaussian self-similar processes. Since this is the first attempt in such a direction, we keep things as simple as possible: we treat the case of the filter $a = \{1, -1\}$ with a k-variation order $= 2$ (quadratic variation), but clearly the method can be generalized; we further specialize to the simplest non-Gaussian Hermite process, i.e. the one of order 2, the Rosenblatt process.

We will apply the Malliavin calculus, Wiener-Itô chaos expansions, and recent results on the convergence of multiple stochastic integrals proved in [17], [8], [16] or [18]. The key point is the following: if the observed process X lives in some finite Wiener chaos, then the statistics V_N can be decomposed, using product formulas and Wiener chaos calculus, into a finite sum of multiple integrals. Then one can attempt to apply the criteria in [16] to study the convergence in law of such sequences and to derive results on the estimators for the Hurst parameter of the observed process. The criteria in [16] are necessary and sufficient conditions for convergence to the Gaussian law; in some instances, these criteria fail (e.g. the fBm case with $H > 3/4$), in which case a proof of non-normal convergence "by hand", working directly with the chaoses, will be employed. It is the basic Wiener chaos calculus that makes this possible. Non-central limit theorem techniques which have been used in the past for such convergences do not seem to work in this situation.

We now summarize the main results in this paper in some detail. As stated above, we use quadratic variation with $a = \{1, -1\}$. We consider the following two processes, observed at the discrete times $\{i/N\}_{i=0}^{N}$: the fBm process $X = B$, and the Rosenblatt process $X = Z$. In either case, the standardized quadratic variation, and the Hurst parameter estimator, are given by

$$V_N = V_N\left(2, \{-1, 1\}\right) := \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\left| X\left(i/N\right) - X\left(\left(i-1\right)/N\right) \right|^2}{N^{-2H}} - 1 \right), \tag{3}$$

$$\hat{H}_N = \hat{H}_N(2, \{-1, 1\}) := \frac{1}{2} - \frac{1}{2 \log N} \log \sum_{i=1}^{N} \left(X(\frac{i}{N}) - X(\frac{i-1}{N}) \right)^2.$$

*Note that, unfortunately in this paragraph, to be consistent with the literature, the letter k is being used for the order of the variation.

The H-dependent constants $c_{j,H}$ (et. al.) refered to below are defined explicitly in lines (7), (9), (12), (32), (46), (48), and (43). Here and throughout, $L^2(\Omega)$ denotes the set of square-integrable random variables measurable w.r.t. the sigma-field generated by W. This sigma-field is the same as that generated by B or by Z. The term "Rosenblatt random variable" denotes a r.v. whose distribution is the same as that of $Z(1)$.

In this paper we prove that, as $N \to \infty$,

1. if $X = B$ and $H \in (0, 3/4)$, then $\sqrt{N/c_{1,H}} V_N$ converges in $L^2(\Omega)$ to a standard normal random variable; so does $\sqrt{N}\log(N)\frac{2}{\sqrt{c_{1,H}}}\left(\hat{H}_N - H\right)$;

2. if $X = B$ and $H \in (3/4, 1)$, then $\sqrt{N^{4-4H}/c_{2,H}} V_N$ converges in $L^2(\Omega)$ to a standard Rosenblatt random variable with parameter $H_0 = 2H - 1$; so does $N^{1-H}\log(N)\frac{2}{\sqrt{c_{2,H}}}\left(\hat{H}_N - H\right)$;

3. if $X = B$ and $H = 3/4$, then $\sqrt{N/\left(c'_{1,H}\log N\right)} V_N$ converges in $L^2(\Omega)$ to a standard normal random variable; so does $\sqrt{N\log N}\frac{2}{\sqrt{c'_{1,H}}}(\hat{H}_N(2,a) - H)$;

4. if $X = Z$ and $H \in (1/2, 1)$, then $N^{1-H} V_N(2,a)/\left(4d(H)\right)$ converges in $L^2(\Omega)$ to the Rosenblatt random variable $Z(1)$; so does $\frac{N^{1-H}}{2d(H)}\log(N)\ (\hat{H}_N(2,a) - H)$;

5. if $X = Z$ and $H \in (1/2, 3/4)$, then $\frac{\sqrt{N}}{\sqrt{e_{1,H}}}\left[V_N(2,a) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1)\right]$ converges in distribution to a standard normal distribution; so does $\frac{\sqrt{N}}{\sqrt{e_{1,H}}}\left[-2\log(N)\ (\hat{H}_N(2,a) - H) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1)\right]$;

6. if $X = Z$ and $H = 3/4$, then $\frac{\sqrt{N}}{\sqrt{e_{3,H}}}\left[V_N(2,a) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1)\right]$ converges in distribution to a standard normal distribution; so does $\frac{\sqrt{N}}{\sqrt{e_{3,H}}}\left[-2\log(N)\ (\hat{H}_N(2,a) - H) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1)\right]$.
   Note that $Z(1)$ is the **actual observed value** of the Rosenblatt process at time 1.

In the fBm case of points 1, 2, and 3 above, the indicated convergences for the standardized V_N's, at least in distribution, have been known for some time, in works such as [22] or [7]. In addition to these results, we are able to treat the limit case $H = \frac{3}{4}$ because our technology works directly, and our convergences are stronger, in $L^2(\Omega)$. Another motivation to write here the proofs comes from the fact that certain terms from the Gaussian case will appear also in the non-Gaussian case. We also tackle the vectorial convergence of these quantities. The application to parameter estimation was spelled out recently in the case of $H < 3/4$ in [3]. Our proof, however, is more elementary and self-contained, and indicates what is left of this parameter estimation when $H = 3/4$ and $H > 3/4$; in the latter case, the standardized estimator's asymptotic normality is no longer available, one has to deal instead with asymptotic Rosenblatt distribution.

In the Rosenblatt process case of points 4, 5, and 6, all results are new, for a subject that has received too narrow a treatment in the literature, presumably because standard techniques inherited from the Non Central Limit Theorem (and based sometimes on the Fourier transform formula for the driving Gaussian process) are difficult to apply (see [4], [22], [2]). Our Wiener chaos calculus approach allows us to show that the standardized quadratic variation and corresponding estimator both converge to a Rosenblatt random variable in $L^2(\Omega)$. Here our method has a crucial advantage: we are able to determine which Rosenblatt random variable it converges to; it is none other than

the observed value $Z(1)$. This is where the fact we are able to prove $L^2(\Omega)$ convergence, not just convergence in distribution, is crucial. Indeed, when $H \leq 3/4$, substracting an appropriately normalized version of this observed value from the quadratic variation and its associated estimator, we prove that asymptotic normality does hold in this case. This unexpected result has important consequences for the statistics of the Rosenblatt process, since it permits the use of standard artillery in parameter estimation and testing.

Our asymptotic normality result for the Rosenblatt process was specifically made possible by showing that V_N can be decomposed into two terms: a term in the fourth Wiener chaos and a term in the second Wiener chaos; and while the second-Wiener chaos term always converges to the Rosenblatt r.v. $Z(1)$, the fourth chaos term converges to a Gaussian r.v. for $H \leq \frac{3}{4}$. We conjecture that this asymptotic normality should also occur for Hermite processes of higher order $k \geq 3$, and that the threshold $H = 3/4$ is universal.

Beyond our basic interest concerning parameter estimation problems, we would also like to situate our paper in the context of some recent and interesting works on the asymptotic behavior of p-variations (or weighted variations) for Gaussian processes, namely the papers [13], [14], [12] and [20]. These recent papers study the behavior of sequences of the type

$$\sum_{i=1}^{N} h\left(X\left(i/N\right)\right)\left(\frac{\left|X\left(i/N\right)-X\left(\left(i-1\right)/N\right)\right|^2}{N^{-2H}}-1\right)$$

where X is a Gaussian process (fractional Brownian motion in [12], [13] and [14] and the solution of the heat equation in [20]) and h is a regular deterministic function. In the fractional Brownian motion case, the behavior of such sums varies according to the values of the Hurst parameter, the limit being sometimes a Gaussian random variable and sometimes a deterministic integral. We believe our work is the first to tackle a Non-Gaussian case, that is, when the process X above is a Rosenblatt process. Although we restrict ourselves to the case when $h = 1$ we still observe the appearance of interesting limits, depending on the Hurst index: while in general the limit of the suitably normalized sequence is a Rosenblatt random variable (with the same Hurst parameter H as the data, which poses a problem for statistical applications), the adjusted variations (that is to say, the sequences obtained by precisely subtracting the portion responsible for the non-Gaussian convergence) do converge to a Gaussian limit for H between $\frac{1}{2}$ and $\frac{3}{4}$.

This article is structured as follows. Section 2 presents preliminaries on fractional stochastic analysis. Section 3 presents detailed calculations and proofs of our results in the Gaussian case; many of these calculations will be useful in the non-Gaussian case. Section 4 contains proofs of our results for the non-Gaussian, Rosenblatt process. Section 5 establishes our parameter estimation results.

2 Preliminaries

Here we describe the elements from stochastic analysis that we will need in the paper. Consider $(B_t^H)_{t\in[0,1]}$ a fractional Brownian motion with Hurst parameter $H \in (0,1)$ and denote by $\mathcal{H}$ its canonical Hilbert space. If $H = \frac{1}{2}$ then $B^{\frac{1}{2}}$ is the standard Wiener process and in this case $\mathcal{H} = L^2([0,1])$. Otherwise $\mathcal{H}$ is the unique Hilbert space of functions on $[0,1]$ extending the rule

$$\left\langle \mathbf{1}_{[0,s]}; \mathbf{1}_{[0,t]} \right\rangle_{\mathcal{H}} = R_H(s,t) := 2^{-1}\left(s^H + t^H - |t-s|^H\right).$$

Denote by I_n the multiple stochastic integral with respect to B^H. I_n is actually an isometry between the Hilbert space $\mathcal{H}^{\otimes n}$ equipped with the modified norm $\frac{1}{\sqrt{n!}}\|\cdot\|_{\mathcal{H}^{\otimes n}}$ and the Wiener chaos of order n which is defined as the closed linear span of the random variables $H_n(B^H(\varphi))$ where $\varphi \in \mathcal{H}$ and H_n is the Hermite polynomial of degree n. We will sometimes use the representation of $B = B^H$ with respect to a standard Brownian motion W: there exists a Wiener process W and a deterministic kernel $K(t,s)$ for $0 \leq s \leq t$ such that $B(t) = I_1^W(K(t,\cdot))$ where I_1^W is the Wiener integral with respect to W (see [15]).

We recall that any square integrable random variable which is measurable with respect to the σ-algebra generated by B^H can be expanded into an orthogonal sum of multiple stochastic integrals

$$F = \sum_{n\geq 0} I_n(f_n)$$

where $f_n \in \mathcal{H}^{\otimes n}$ are symmetric functions and $I_0(f_0) = \mathbf{E}[F]$.

We will need the general formula for calculating products of Wiener chaos integrals of any orders p, q for any symmetric integrands $f \in \mathcal{H}^{\otimes p}$ and $g \in \mathcal{H}^{\otimes q}$, is

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q} r! C_p^r C_q^r I_{p+q-2r}(f \otimes_r g) \tag{4}$$

as given for instance in D. Nualart's book [15, Proposition 1.1.3]; the contraction $f \otimes_r g$ is an element of $\mathcal{H}^{\otimes(p+q-2r)}$ defined by

$$f \otimes_r g = \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \ldots e_{i_r}\rangle_{\mathcal{H}^r} \langle g, e_{i_1} \otimes \ldots e_{i_r}\rangle_{\mathcal{H}^r}$$

where $(e_i)_{i\geq 1}$ is an orthogonal system in $\mathcal{H}$.

We now introduce the Malliavin derivative for random variables in a finite chaos. If $f \in \mathcal{H}^{\otimes}$ is symmetric then we will use the following rule to differentiate in the Malliavin sense

$$D_t I_n(f) = n I_{n-1}(f_n(\cdot, t)), \qquad t \in [0,1].$$

It is possible to characterize the convergence in law of a sequence of multiple integrals to a Gaussian random variable. We will use the following result (see Theorem 4 in [16], see also [17]).

Theorem 1 *Let $F_k = I_n(f_k)$ be a sequence of square integrable random variables in the n th Wiener chaos such that $\mathbf{E}(F_k^2) \to 1$ as $k \to \infty$. Then the following are equivalent:*

i) The sequence $(F_k)_{k\geq 0}$ converges to the normal law $N(0,1)$.

ii) One has $\mathbf{E}(F_k^4) \to 3$ as $k \to \infty$.

iii) For all $1 \leq l \leq n-1$ it holds $\lim_{k\to\infty} \|f_k \otimes_l f_k\|^{\mathcal{H}^{\otimes 2(n-l)}} \to 0$.

iv) $\|DF_k\|_{\mathcal{H}}^2 \to n$ in $L^2(\Omega)$ as $k \to \infty$, where D is the Malliavin derivative is with respect B.

Criterion (iv) is due to [16]; we will refer to it as the Nualart–Ortiz-Latorre criterion. A multidimensional version of the above theorem has been proved in [18]; it will be used in Section 3 to study the vectorial convergence of our estimators.

3 The Gaussian data case

3.1 Strategy: try the Nualart–Ortiz-Latorre characterization

With $a = \{1, -1\}$ and $k = 2$, with $B =$ standard fractional Brownian motion (fBm), we begin by using our tools to reproduce the results found in J.-F. Cœurjolly [3], seeking to show a central limit theorem for the standardized quadratic variation. With the notation $A_i = \mathbf{1}_{((i-1)/N, i/N]}$ and $B(A_i) = B(i/N) - B((i-1)/N)$ so that $Var[B(A_i)] = N^{-2H}$, this variation, given in (3), can be expressed as

$$V_N = N^{2H-1} \sum_{i=1}^{N} \left(|B(A_i)|^2 - N^{-2H} \right). \tag{5}$$

To show it converges to the standard normal law $\mathcal{N}(0,1)$ after an appropriate scaling, instead of employing Cœurjolly's Gaussian method, we propose a Wiener chaos approach which can generalize to higher order cases than fBm. The same tools will be used for non-Gaussian observation in Section 4, and some of the calculations there will be lifted from the current Gaussian section, which we therefore present here in some detail. In fact, our tools allow some new results even in the Gaussian case.

The strategy is to attempt to use the characterization of $\mathcal{N}(0,1)$ by Nualart and Ortiz-Latorre [16, Theorem 4] as presented in the previous section: if $\{F_N = I_n(f_N) : N \in \mathbf{N}\}$ is a sequence of nth Wiener chaos random variables, and $\mathbf{E}\left[|F_N|^2\right] = \|f_N\|^2_{\mathcal{H}^{\otimes 2}}$ converges to 1 as N tends to infinity, then $\lim_{N\to\infty} F_N = \mathcal{N}(0,1)$ in distribution if and only if

$$\lim_{N\to\infty} \mathbf{E}\left[\left(\|DF_N\|^2_{\mathcal{H}} - n \right)^2 \right] = 0. \tag{6}$$

In the next two subsections, we prove the following.

Theorem 2 *Let $H \in (0,1)$ and B be a fractional Brownian motion with parameter H. Consider the standardized quadratic variation V_N given by (5).*

If $H \in (1/2, 3/4)$, let

$$c_{1,H} := 2 + \sum_{k=1}^{\infty} \left(2k^{2H} - (k-1)^{2H} - (k+1)^{2H} \right)^2; \tag{7}$$

then

$$F_N := \sqrt{N/c_{1,H}} V_N \tag{8}$$

converges in $L^2(\Omega)$ to a standard normal random variable.

If $H \in (3/4, 1)$, let

$$c_{2,H} := 2H^2(2H-1)/(4H-3); \tag{9}$$

then

$$\bar{F}_N := \sqrt{N^{4-4H}/c_{2,H}} V_N \tag{10}$$

converges in $L^2(\Omega)$ to a standard Rosenblatt random variable with parameter $H_0 = 2H - 1$; this random variable is equal to

$$\frac{(4H-3)^{1/2}}{4H(2H-1)^{1/2}} \iint_{[0,1]^2} \left(\int_{r\vee s}^{1} \frac{\partial K^H}{\partial u}(u,s) \frac{\partial K^H}{\partial u}(u,r)\, du \right) dW(r)\, dW(s) \tag{11}$$

where W is the standard Brownian motion used in the representation $B(t) = I_1(K(t,\cdot))$.

If $H = 3/4$, let

$$c'_{1,H} := (2H(2H-1))^2 = 9/16 \tag{12}$$

and define

$$\tilde{F}_N := \sqrt{\frac{N}{c'_{1,H}\log N}} V_N. \tag{13}$$

Then $\tilde{F}_N$ converges in distribution to a standard normal distribution.

The Nualart-Ortiz-Latorre characterization is useful only when a Gaussian limit exists; otherwise, which is the case when $H > 3/4$, a different argument will need to be used.

3.2 Expectation evaluation

Our first task is to evaluate $\mathbf{E}\left[V_N^2\right]$. The product formula (4), in our present, nearly trivial, case, yields

$$\begin{aligned} V_N &= N^{2H-1}\sum_{i=1}^{N}\left(|I_1(A_i)|^2 - N^{-2H}\right) \\ &= N^{2H-1}\sum_{i=1}^{N}\left(I_2(A_i\otimes A_i) + I_0(A_i\otimes_2 A_i) - N^{-2H}\right) \\ &= N^{2H-1}\sum_{i=1}^{N} I_2(A_i\otimes A_i) = N^{2H-1} I_2\left(\sum_{i=1}^{N} A_i\otimes A_i\right). \end{aligned}$$

Now using the formula for the second moment of a second chaos r.v. (use (4) and/or see [15, page 8]),

$$\begin{aligned} \mathbf{E}\left[|V_N|^2\right] &= N^{4H-2}2!\left\|\sum_{i=1}^{N} A_i\otimes A_i\right\|_{\mathcal{H}}^2 \\ &= N^{4H-2}2!\sum_{i=1}^{N}\sum_{j=1}^{N}\langle A_i\otimes A_i, A_j\otimes A_j\rangle_{\mathcal{H}} \\ &= 2N^{4H-2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\langle A_i, A_j\rangle_{\mathcal{H}}\right|^2. \end{aligned} \tag{14}$$

To calculate this quantity, we notice that $\langle A_i, A_j\rangle_{\mathcal{H}} = \mathbf{E}\left[B(A_i)B(A_j)\right]$ can be calculated explicitly using the covariance function R_H

$$\langle A_i, A_j\rangle_{\mathcal{H}} = 2^{-1}\left(2\left|\frac{i-j}{N}\right|^{2H} - \left|\frac{i-j-1}{N}\right|^{2H} - \left|\frac{i-j+1}{N}\right|^{2H}\right).$$

This expression is close to $H(2H-1)N^{-2}\left|(i-j)/N\right|^{2H-2}$, but we must take care whether the series $\sum_k k^{4H-4}$ converges or diverges. Let us consider first the case of convergence.

Case 1: $H < 3/4$. In this case, isolating the diagonal term, and writing the remaining term as twice the sum over $i > j$, we can write

$$\begin{aligned}\mathbf{E}\left[|V_N|^2\right] &= N^{4H-2}2^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(2\left|\frac{i-j}{N}\right|^{2H} - \left|\frac{i-j-1}{N}\right|^{2H} - \left|\frac{i-j+1}{N}\right|^{2H}\right)^2 \\ &= N^{4H-2}2N^{-4H+1} \\ &+ N^{4H-2}\sum_{i=1}^{N}\sum_{k=1}^{N-1}\left(2\left|\frac{k}{N}\right|^{2H} - \left|\frac{k-1}{N}\right|^{2H} - \left|\frac{k+1}{N}\right|^{2H}\right)^2 \\ &= 2/N + N^{-2}\sum_{k=1}^{N-1}(N-k)\left(2k^{2H} - (k-1)^{2H} - (k+1)^{2H}\right)^2.\end{aligned}$$

Thanks to Lemma 16 in the Appendix, the above series converges, and we have proved the limit

$$\lim_{N\to\infty}\mathbf{E}\left[\left|\sqrt{N}V_N\right|^2\right] = c_{1,H}.$$

In this case, we therefore set

$$F_N := \sqrt{\frac{N}{c_{1,H}}}V_N \tag{15}$$

and, in the next subsection, we will try to apply the criterion (6). We also record, for latter use, the result

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\langle A_i, A_j\rangle_{\mathcal{H}}\right|^2 = 2^{-1}c_{1,H}N^{1-4H} + o\left(N^{1-4H}\right) \tag{16}$$

Case 2: $H > 3/4$. In this case, we will instead compare the series in $\mathbf{E}\left[|V_N|^2\right]$ to an integral. Our first task is to do away with the tridiagonal term. Therefore let us first consider the sum in (14) for $|i-j| \leq 1$:

$$\sum_{i,j=1,\cdots,N;|i-j|\leq 1}\left|\langle A_i, A_j\rangle_{\mathcal{H}}\right|^2 = N\left|2^{-1}\left(\left(-2+2-2^{2H}\right)N^{-2H}\right)\right|^2 = 2^{4H-2}N^{1-4H}. \tag{17}$$

The corresponding term in (14) is thus equal to $2^{4H-1}N^{-1}$. On the other hand, we immediately get by Lemma 17 in the Appendix that $N^2\sum_{i,j=1,\cdots,N;|i-j|\geq 1}\left|\langle A_i, A_j\rangle_{\mathcal{H}}\right|^2$ compares to a Riemann sum in such a way that it converges to $H^2(2H-1)/(4H-3)$. This also shows that the non-tridiagonal term, which is of order N^{-2}, dominates the tridiagonal term (17), because it is of order $N^{1-4H} \ll N^{-2}$ (since $H > 3/4$). Thus we have the following result, which we record for further use below:

$$\lim_{N\to\infty}N^2\sum_{i,j=1,\cdots,N}\left|\langle A_i, A_j\rangle_{\mathcal{H}}\right|^2 = \frac{H^2(2H-1)}{4H-3}, \tag{18}$$

In conclusion when $H > 3/4$, according to (14), this result translates as

$$\lim_{N\to\infty}\mathbf{E}\left[\left|N^{2-2H}V_N\right|^2\right] = c_{2,H}.$$

In this case, we will therefore set

$$F_N := N^{2-2H} V_N / \sqrt{c_{2,H}} \tag{19}$$

and try to apply the criterion (6).

Case 3: $H = 3/4$. This case must be treated separately. In this case in the usual way, we have

$$\mathbf{E}\left[(V_N)^2\right] = \frac{2}{N} + \frac{1}{N}\sum_{k=0}^{N-1}\left(2k^{2H} - (k-1)^{2H} - (k+1)^{2H}\right)^2 - \frac{1}{N^2}\sum_{k=0}^{N-1} k\left(2k^{2H} - (k-1)^{2H} - (k+1)^{2H}\right)^2$$

and since $2k^{2H} - (k-1)^{2H} - (k+1)^{2H}$ behaves as $(3/4)\, k^{-1/2}$ we get $\mathbf{E}\left[(V_N)^2\right] \sim c'_{1,H} (\log N)/N$. Thus, $\lim_{N\to\infty} \mathbf{E}\left[\left|\tilde{F}_N\right|^2\right] = 1$ where $c'_{1,H} = 3/16$ and

$$\tilde{F}_N := \left(\frac{N}{c'_{1,H} \log N}\right)^{\frac{1}{2}} V_N. \tag{20}$$

3.3 Derivative calculations

We now attempt to show that $\|DF_N\|^2_{\mathcal{H}}$ converges in $L^2(\Omega)$ to $n = 2$, where F_N is given by (8), (10) or (13) depending on whether H is bigger or smaller than $3/4$ or equal to $3/4$.

We first calculate the Malliavin derivative of V_N, and investigate the expectation $\mathbf{E}\left[\|DF_N\|^2_{\mathcal{H}}\right]$ to check that, at least, it converges to 2. We use the rule $D_r I_2(f) = 2I_1(f(\cdot, r))$ when f is symmetric. We have

$$\begin{aligned} D_r V_N &= N^{2H-1} D_r I_2\left(\sum_{i=1}^N A_i \otimes A_i\right) \\ &= 2N^{2H-1}\sum_{i=1}^N I_1(A_i \otimes A_i(\cdot, r)) = 2N^{2H-1}\sum_{i=1}^N A_i(r)\, I_1(A_i). \end{aligned}$$

Hence

$$\|DV_N\|^2_{\mathcal{H}} = 4N^{4H-2}\sum_{i,j=1}^N I_1(A_i)\, I_1(A_j)\, \langle A_i; A_j\rangle_{\mathcal{H}} \tag{21}$$

and therefore

$$\mathbf{E}\left[\|DV_N\|^2_{\mathcal{H}}\right] = 4N^{4H-2}\sum_{i,j=1}^N \left|\langle A_i; A_j\rangle_{\mathcal{H}}\right|^2$$

We note immediately from (14) that $\mathbf{E}\left[\|DV_N\|^2_{\mathcal{H}}\right] = 2\mathbf{E}\left[V_N^2\right]$. From the results of the previous section, whether in the case $H < 3/4$, $H = 3/4$ or $H \geq 3/4$, we then automatically get that

$$\lim_{N\to\infty} \mathbf{E}\left[\|DF_N\|^2_{\mathcal{H}}\right] = 2.$$

Thus it is now sufficient to show that $\|DF_N\|_{\mathcal{H}}^2 - \mathbf{E}\left[\|DF_N\|_{\mathcal{H}}^2\right]$ converges to 0 in $L^2(\Omega)$. For this calculation, we will use the following convenient Malliavin-derivative tool, which can be found in [24, Chapter 9].

Lemma 3 *Let Y be a centered random variable in the space $\mathbf{D}^{1,2}$ relative to B_H. Then for any $m = 1, 2, 3, \cdots$, there exists a constant c_m depending only on m such that*

$$\mathbf{E}\left[|Y|^m\right] \leq c_m \mathbf{E}\left[\|DY\|_{\mathcal{H}}^m\right]$$

To simplify the calculation, consider first

$$Y = \sum_{i,j=1}^{N} I_1(A_i) I_1(A_j) \langle A_i; A_j \rangle_{\mathcal{H}}.$$

We calculate

$$\begin{aligned} D_r Y &= \sum_{i,j=1}^{N} \left(I_1(A_i) A_j(r) + I_1(A_j) A_i(r)\right) \langle A_i; A_j \rangle_{\mathcal{H}} \\ &= 2 \sum_{i=1}^{N} \sum_{j=1}^{i} B(A_i) A_j(r) \langle A_i; A_j \rangle_{\mathcal{H}}. \end{aligned}$$

Thus

$$\|DY\|_{\mathcal{H}}^2 = 4 \sum_{i,i'=1}^{N} B(A_i) B(A_{i'}) \sum_{j=1}^{i} \sum_{j'=1}^{i'} \langle A_i; A_j \rangle_{\mathcal{H}} \left\langle A_{i'}; A_{j'} \right\rangle_{\mathcal{H}} \left\langle A_j; A_{j'} \right\rangle_{\mathcal{H}}.$$

We can calculate the expectation of this expression immediately:

$$\mathbf{E}\left[\|DY\|_{\mathcal{H}}^2\right] = 4 \sum_{i,i'=1}^{N} \sum_{j=1}^{i} \sum_{j'=1}^{i'} \langle A_i; A_{i'} \rangle_{\mathcal{H}} \langle A_i; A_j \rangle_{\mathcal{H}} \left\langle A_{i'}; A_{j'} \right\rangle_{\mathcal{H}} \left\langle A_j; A_{j'} \right\rangle_{\mathcal{H}} \tag{22}$$

Case 1: $H < 3/4$. In this case, using the scaling $F_N = N^{1/2} V_N / \sqrt{c_{1,H}}$ and the expression (21) which yields $\left\| D \|DV_N\|_{\mathcal{H}}^2 \right\|_{\mathcal{H}}^2 = \left(4N^{4H-2}\right)^2 \|DY\|_{\mathcal{H}}^2$, (22) becomes

$$\begin{aligned} \mathbf{E}\left[\left\| D \|DF_N\|_{\mathcal{H}}^2 \right\|_{\mathcal{H}}^2\right] &= (c_{1,H})^{-2} N^2 \left(4N^{4H-2}\right)^2 \|DY\|_{\mathcal{H}}^2 \\ &= \frac{16}{c_{1,H}^2} N^{8H-2} \|DY\|_{\mathcal{H}}^2 \\ &= \frac{64}{c_{1,H}^2} N^{8H-2} \sum_{i,i'=1}^{N} \sum_{j=1}^{i} \sum_{j'=1}^{i'} \langle A_i; A_{i'} \rangle_{\mathcal{H}} \langle A_i; A_j \rangle_{\mathcal{H}} \left\langle A_{i'}; A_{j'} \right\rangle_{\mathcal{H}} \left\langle A_j; A_{j'} \right\rangle_{\mathcal{H}} \end{aligned} \tag{23}$$

By Lemma 18 in the appendix, the conclusion is that $\mathbf{E}\left[\left\|D\left\|DF_N\right\|_{\mathcal{H}}^2\right\|_{\mathcal{H}}^2\right]$ is asymptotically equivalent to a constant multiple of N^{8H-6}. By Lemma 3, for some constant c we get that

$$\mathbf{E}\left[\left(\left\|DF_N\right\|_{\mathcal{H}}^2-2\right)^2\right]\leq cN^{8H-6}\to 0,$$

which, together with the calculations in Section 3.2, is the Nualart–Ortiz-Latorre necessary and sufficient condition (6) for F_N to converges in law to a standard normal, i.e.

$$\lim_{N\to\infty}\sqrt{N}V_N\overset{\mathcal{D}}{=}\mathcal{N}\left(0,c_{1,H}\right).$$

Convergence in $L^2\left(\Omega\right)$ can also be established, using techniques similar to those in Case 2; the details are left to the reader.

Case 2: $H>3/4$. In this case, using the scaling $\bar{F}_N=N^{2-2H}V_N/\sqrt{c_{2,H}}$ and the expression $\left\|D\left\|DV_N\right\|_{\mathcal{H}}^2\right\|_{\mathcal{H}}^2=\left(4N^{4H-2}\right)^2\left\|DY\right\|_{\mathcal{H}}^2$ from (21), we have in (22):

$$\mathbf{E}\left[\left\|D\left\|D\bar{F}_N\right\|_{\mathcal{H}}^2\right\|_{\mathcal{H}}^2\right]=\frac{64N^4}{c_{2,H}^2}\sum_{i,i'=1}^{N}\sum_{j=1}^{i}\sum_{j'=1}^{i'}\left\langle A_i;A_{i'}\right\rangle_{\mathcal{H}}\left\langle A_i;A_j\right\rangle_{\mathcal{H}}\left\langle A_{i'};A_{j'}\right\rangle_{\mathcal{H}}\left\langle A_j;A_{j'}\right\rangle_{\mathcal{H}}.$$

A calculation of the same nature as in Lemma 18 proves that the above is asymptotically equivalent to

$$\frac{16\left|H\left(2H-1\right)\right|^4}{c_{2,H}^2}\int_{[0,1]^4}\left|\left(x-x'\right)\left(x-y\right)\left(x'-y'\right)\left(y-y'\right)\right|^{2H-2}dxdx'dydy'.$$

This is a positive constant depending on H only. Therefore, the Nualart–Ortiz-Latorre condition is not met, which proves that $N^{2-2H}V_N$ does not converge to a Gaussian distribution.

In order to find the limit of $N^{2-2H}V_N$, we calculate the second Malliavin derivative of $\bar{F}_N$, and investigate its limit. It turns out that the limit of the normalized V_N is a Rosenblatt random variable. In order to prove this, we will use the representation of the Rosenblatt process with respect to standard Brownian motion, and because of this, we now consider the Malliavin derivative $\tilde{D}$ with respect to the standard Brownian motion W underlying the fBm process B_H. In other words, such a W exists such that

$$\bar{F}_N=Nc_{2,H}^{-1/2}\sum_{i=1}^{N}\tilde{I}_2\left(\tilde{A}_i\otimes\tilde{A}_i\right)$$

where $\tilde{I}_2$ is the double Wiener integral w.r.t. W and, with $B_H\left(t\right)=\int_0^t K^H\left(t,s\right)dW\left(s\right)$ (see for instance [15, Chapter 5] for a formula for K^H), $\tilde{A}_i$ is defined by

$$\tilde{A}_i\left(s\right)=\mathbf{1}_{[0,\frac{i+1}{N}]}\left(s\right)K^H\left(\frac{i+1}{N},s\right)-\mathbf{1}_{[0,\frac{i}{N}]}\left(s\right)K^H\left(\frac{i}{N},s\right).\tag{24}$$

We thus have

$$\tilde{D}_{r,s}^2\bar{F}_N=2Nc_{2,H}^{-1/2}\sum_{i=1}^{N}\tilde{A}_i\left(r\right)\tilde{A}_i\left(s\right).\tag{25}$$

We claim that, up to a constant, the function $(r,s) \mapsto \tilde{D}^2_{r,s}\bar{F}_N$ converges in $L^2\left([0,1]^2\right)$ to $2c_{2,H}^{-1/2}L$ where L is the function

$$(r,s) \mapsto L(r,s) := \int_{r\vee s}^1 \frac{\partial K^H}{\partial u}(u,s)\frac{\partial K^H}{\partial u}(u,r)\,du. \tag{26}$$

This fact is proved in Lemma 19 in the Appendix.

Now, since $\bar{F}_N$ is a second-chaos variable, we can write $\bar{F}_N = \tilde{I}_2(f_N)$ for some symmetric function f_N in $L^2\left([0,1]^2\right)$. By basic properties of the derivative, we have $f_N = 2^{-1}\tilde{D}^2\bar{F}_N$. Now define the following Rosenblatt random variable with parameter $H_0 = 2H-1$:

$$Y = d(H_0)\,\tilde{I}_2(L) = d(H_0)\iint_{[0,1]^2}\left(\int_{r\vee s}^1 \frac{\partial K^H}{\partial u}(u,s)\frac{\partial K^H}{\partial u}(u,r)\,du\right)dW(r)\,dW(s),$$

where

$$\begin{aligned} d(H_0) &= (H_0+1)^{-1}\left(2(2H_0-1)/H_0\right)^{1/2} \\ &= (4H-3)^{1/2}(2H-1)^{-1/2}/\left(\sqrt{2}H\right) \\ &= c_{2,H}^{-1/2} \end{aligned}$$

Therefore we have, by the isometry property,

$$\mathbf{E}\left[\left|\bar{F}_N - c_{2,H}^{-1/2}\tilde{I}_2(L)\right|^2\right] = \left\|2^{-1}\tilde{D}^2\bar{F}_N - c_{2,H}^{-1/2}L\right\|^2_{L^2([0,1]^2)},$$

which, by the previous claim, proves that $\bar{F}_N$ converges to the Rosenblatt random variable $Y = c_{2,H}^{-1/2}\tilde{I}_2(L)$ in $L^2(\Omega)$.

Case 3: $H = 3/4$. In this final, limit case, we get

$$\begin{aligned} \|DV_N\|^2_{\mathcal{H}} &= 4N^{4H-2}\sum_{i,j=1}^N I_1(A_i)I_1(A_j)\langle A_i,A_j\rangle_{\mathcal{H}} \\ &= 4N^{4H-2}\sum_{i,j=1}^N \left[I_2(A_i\otimes A_j) + \langle A_i,A_j\rangle_{\mathcal{H}}\right]\langle A_i,A_j\rangle_{\mathcal{H}} \end{aligned}$$

and we obtain, by the evaluation of mean square of V_N, the limit as $N\to\infty$

$$\mathbf{E}\left[\|D\tilde{F}_N\|^2_{\mathcal{H}}\right] = \frac{N}{\log N}\left(c'_{1,H}\right)^{-1}4N^{4H-2}\sum_{i,j=1}^N \langle A_i,A_j\rangle^2_{\mathcal{H}} \to 2. \tag{27}$$

Therefore it is sufficient to check that the sum

$$\frac{N}{\log N}N^{4H-2}\sum_{i,j=1}^N\sum_{i,j=1}^N I_2(A_i\otimes A_j)$$

converges to zero in $L^2(\Omega)$ as $N \to \infty$. Indeed, its L^2 norm is given by

$$\frac{N^{8H-2}}{(\log N)^2} \sum_{i,j,k,l} \langle A_i \otimes A_j, A_k \otimes A_l \rangle_{\mathcal{H}\otimes\mathcal{H}} \langle A_i, A_j \rangle_{\mathcal{H}} \langle A_k, A_l \rangle_{\mathcal{H}}$$
$$= \frac{N^{8H-2}}{(\log N)^2} \sum_{i,j,k,l} \langle A_i, A_k \rangle_{\mathcal{H}} \langle A_j, A_l \rangle_{\mathcal{H}} \langle A_i, A_j \rangle_{\mathcal{H}} \langle A_k, A_l \rangle_{\mathcal{H}}.$$

The sum has already been evaluated. The last expression's behavior is of order $\frac{N^{8H-6}}{(\log N)^2}$ and so its goes to zero as $N \to \infty$. As a consequence $\|D_N \tilde{F}_N\|^2_{\mathcal{H}}$ converges to 2 in $L^2(\Omega)$. This finishes the proof of Theorem 2. ■

Remark 4 *For every $H \in (0,1)$ it holds that $V_N(2,a)$ converges almost surely to zero. Indeed, we already showed in this Section 3 the $L^2(\Omega)$ convergence to zero as $N \to \infty$; to obtain almost sure convergence we only need to use an argument in [3] (proof of Proposition 1) for empirical means of discrete stationary processes.*

3.4 Multidimensional convergence of the 2- variations

This section is devoted to the study of the vectorial convergence of the 2-variations statistics. We will restrict ourselves to the case $H \leq \frac{3}{4}$ in which the limit of the components are Gaussian random variables. We make this choice in order to benefit from some recent results in [18] that characterize the convergence in law of a vector of multiple stochastic integrals to a Gaussian vector. Our strategy is based on the following result (Proposition 2 in [18]).

Proposition 5 *Let $d \geq 2$, and fix $n \geq 2$ as well as a collection of kernels*

$$\left\{ \left(f_1^{(k)}, ..., f_d^{(k)} \right) : k \geq 1 \right\}$$

such that $f_j^{(k)} \in \mathcal{H}^n$ is symmetric and for every $k \geq 1$ and every $j = 1, ..., d$, and

$$\lim_{k\to\infty} j! \left\| f_j^{(k)} \right\|^2_{H^{\otimes n}} = C_{jj}, \quad \forall j = 1, ..., d \tag{28}$$
$$\lim_{k\to\infty} \mathbf{E}\left[I_n^X \left(f_i^{(k)} \right) I_n^X \left(f_j^{(k)} \right) \right] = C_{ij}, \quad \forall 1 \leq i < j \leq d.$$

Then, the following conditions are equivalent:

(i) as k goes to infinity, the vector $\left(I_n^X \left(f_1^{(k)} \right), ..., I_n^X \left(f_d^{(k)} \right) \right)$ converges in distribution to a d-dimensional Gaussian vector $N_d(0, \mathbf{C}_d) = (N_1, ..., N_d)$ with covariance matrix $\mathbf{C}_d$;

(ii) for every $j = 1, ..., d$, $I_n^X \left(f_j^{(k)} \right)$ converges in distribution to N_j, that is, to the law of a centered Gaussian random variable with variance C_{jj};

Now, define the following filters constructed from the filter $a = \{1, -1\}$:

$$a^1 = \{1, -1\}, a^2 = \{1, 0, -1\}, a^3 = \{1, 0, 0, -1\}, \quad \ldots a^M = (1, 0, 0, \ldots - 1\}$$

where M is an integer at each step p, the vector a^p has $p-1$ zeros. Note that for every $p=1,\dots,M$, the filter a^p is a $p+1$ dimensional vector.

Consider the statistics based on the above filters ($1 \leq p \leq M$)

$$\begin{aligned} V_N(2,a^p) &= \frac{1}{N-p}\sum_{i=p}^{N}\left[\frac{\left(B(\frac{i}{N})-B(\frac{i-p}{N})\right)^2}{\mathbf{E}\left(B(\frac{i}{N})-B(\frac{i-p}{N})\right)^2}-1\right] \\ &= \frac{1}{N-p}\sum_{i=p}^{N}\left[\left(I_1(A_{i,p})\right)^2\left(\frac{p}{N}\right)^{-2H}-1\right] \\ &= \frac{1}{N-p}\left(\frac{p}{N}\right)^{-2H}\sum_{i=p}^{N} I_2\left(A_{i,p}\otimes A_{i,p}\right) \end{aligned}$$

where we denoted by

$$A_{i,p} = 1_{[\frac{i-p}{N},\frac{i}{N}]}, \qquad 1\leq p\leq M, p\leq i\leq N.$$

We have the following vectorial limit theorem.

Theorem 6 *Let B be a fBm with $H\in(0,3/4)$ and let $M\geq 1$. For $1\leq p,q\leq M$ define*

$$c_{p,q,H} := \frac{1}{(pq)^{2H}}\sum_{k\geq 1}\left(|k|^{2H}+|k-p+q|^{2H}-|k-p|^{2H}-|k+q|^{2H}\right)^2, \text{ and } c_{p,H}:=c_{p,p,H}$$

and

$$\bar{F}_N(a^p) := \sqrt{N}c_{p,H}^{-1}V_N(2,a_p). \tag{29}$$

Then the vector $(\bar{F}_N(a^1),\dots,F_N(a^M))$ converges as $N\to\infty$, to a Gaussian vector with covariance matrix $C=C_{i,j}$ where $C_{p,q}=\frac{c_{p,q,H}}{\sqrt{c_{p,H}c_{q,H}}}$.

If $H=\frac{3}{4}$, define

$$d_{p,q,H} := \frac{1}{(pq)^{2H}}\frac{3}{16}, \text{ and } d_{p,H}:=d_{p,p,H},$$

and

$$\tilde{F}_N(a^p) = \sqrt{\frac{N}{\log N}}d_{p,H}^{-1/2}V_N(2,a^p).$$

Then the vector $(F_N(a^1),\dots,F_N(a^M))$ converges as $N\to\infty$, to a Gaussian vector with covariance matrix $D=D_{i,j}$ where $D_{p,q}=\frac{d_{p,q,H}}{\sqrt{d_{p,H}d_{q,H}}}$.

Proof. Let us estimate the covariance of two such statistics

$$\begin{aligned} &\mathbf{E}\left[V_N(2,a^p)V_N(2,a^q)\right] \\ &= \frac{N^{4H}}{(N-p)(N-q)}\frac{1}{(pq)^{2H}}2\sum_{i=p}^{N}\sum_{j=q}^{N}\langle A_{i,p}\otimes A_{i,p}, A_{j,q}\otimes A_{j,q}\rangle_{\mathcal{H}\otimes\mathcal{H}} \\ &= \frac{N^{4H}}{(N-p)(N-q)}\frac{2}{(pq)^{2H}}\sum_{i=p}^{N}\sum_{j=q}^{N}\langle A_{i,p}\otimes A_{j,q}\rangle_{\mathcal{H}}^2. \end{aligned}$$

The next step is to compute the scalar product

$$\begin{aligned}\langle A_{i,p}\otimes A_{j,q}\rangle_{\mathcal{H}} &= \langle 1_{[\frac{i-p}{N},\frac{i}{N}]}, 1_{[\frac{j-q}{N},\frac{j}{N}]}\rangle_{\mathcal{H}} \\ &= \frac{1}{2}\left[\left|\frac{i-j}{N}\right|^{2H}+\left|\frac{i-j-p+q}{N}\right|^{2H}-\left|\frac{i-j-p}{N}\right|^{2H}-\left|\frac{i-j+q}{N}\right|^{2H}\right].\end{aligned}$$

Assume that $p\geq q$. We need to estimate the sum

$$\begin{aligned}&\sum_{i=p}^{N}\sum_{j=q}^{N}\left[\left|\frac{i-j}{N}\right|^{2H}+\left|\frac{i-j-p+q}{N}\right|^{2H}-\left|\frac{i-j-p}{N}\right|^{2H}-\left|\frac{i-j+q}{N}\right|^{2H}\right] \\ &= \frac{1}{N^{4H}}\sum_{j=q}^{p-1}\sum_{i=p}^{N}\left(|i-j|^{2H}+|i-j-p+q|^{2H}-|i-j-p|^{2H}-|i-j+q|^{2H}\right)^2 \\ &+ \frac{1}{N^{4H}}\sum_{j=p}^{N}\sum_{i=p}^{N}\left(|i-j|^{2H}+|i-j-p+q|^{2H}-|i-j-p|^{2H}-|i-j+q|^{2H}\right)^2 \\ &= \frac{2}{N^{4H}}\sum_{j=p}^{N}\sum_{k=1}^{N-j}\left(|k|^{2H}+|k-p+q|^{2H}-|k-p|^{2H}-|k+q|^{2H}\right)^2 \\ &= \frac{2}{N^{4H}}\sum_{k=1}^{N-p}(N-k-p)\left(|k|^{2H}+|k-p+q|^{2H}-|k-p|^{2H}-|k+q|^{2H}\right)^2 \\ &= \frac{2}{N^{4H}}\sum_{k=1}^{N-p}(N-k-p)k^{4H}g(\frac{1}{k})^2\end{aligned}$$

where we denoted by

$$g(x)=1+(1-(p-q)x)^{2H}-(1-px)^{2H}-(1+qx)^{2H}.$$

By the asymptotic behavior of the function g around zero, we obtain for large k

$$g(\frac{1}{k})\sim 2H(2H-1)pq\frac{1}{k^2}.$$

We distinguish again the cases $H<\frac{3}{4}$ and $k=\frac{3}{4}$ and we conclude that

$$\mathbf{E}\left[V_N(2,a^p)V_N(2,a^q)\right]\sim_{N\to\infty} c_{p,q,H}\frac{1}{N},\qquad \text{for } H<\frac{3}{4}$$

and

$$\mathbf{E}\left[V_N(2,a^p)V_N(2,a^q)\right]\sim_{N\to\infty} d_{p,q,H}\frac{\log N}{N}\qquad \text{for } H=\frac{3}{4}$$

where the constant $c_{p,q,H}$ and $b_{p,q,H}$ have been defined in the statement of the theorem. The conclusion then follows from Proposition 2 in [18]. ■

4 The non-Gaussian data case

Here we enter the non-Gaussian realm. Our observed process is now a Rosenblatt process $(Z(t))_{t\in[0,1]}$ with self-similarity index $H \in (\frac{1}{2}, 1)$.

Recall that this process is self-similar with stationary increments, and lives in the second Wiener chaos. Our goal, as before, is to estimate its H from discrete observations of its sample paths. As far as we know, this direction has seen little or no attention in the literature, and the classical techniques (e.g, the ones from [21], [22] or [4]) do not work well for it. Therefore, the use of Malliavin calculus and multiple stochastic integrals is of interest.

In a natural extension of the Rosenblatt random variable exhibited in the last section, the Rosenblatt process can be represented as (see [23])

$$Z(t) = d(H) \int_0^t \int_0^t \left[\int_{y_1 \vee y_2}^t \frac{\partial K^{H'}}{\partial u}(u, y_1) \frac{\partial K^{H'}}{\partial u}(u, y_2) du \right] dW(y_1) dW(y_2) \tag{30}$$

where $(W_t, t \in [0,1])$ is some standard Brownian motion, K is the standard kernel of fractional Brownian motion (see any reference on fBm, such as [15, Chapter 5]),

$$H' = \frac{H+1}{2} \tag{31}$$

and

$$d(H) = \frac{1}{H+1} \left(\frac{H}{2(2H-1)} \right)^{-\frac{1}{2}}. \tag{32}$$

We will denote the kernel of the Rosenblatt process with respect to W by

$$L_t(y_1, y_2) := d(H) \left[\int_{y_1 \vee y_2}^t \frac{\partial K^{H'}}{\partial u}(u, y_1) \frac{\partial K^{H'}}{\partial u}(u, y_2) du \right] 1_{[0,t]^2}(y_1, y_2). \tag{33}$$

¿From this point on, we will use the same technical strategy as we introduced at the end of the subsection on derivative calculations, which is to write all stochastic integrals directly with respect to W, and all Malliavin derivatives with respect to W as well. Abusing and simplifying the notation (dropping any tildes), D now denotes the Malliavin derivative and I_n denotes the nth Wiener integral, both with respect to W. In other words, in particular, for every t

$$Z(t) = I_2\left(L_t(\cdot)\right).$$

Consider now the filter $a = \{-1, 1\}$ and the 2-variations given by

$$\begin{aligned} V_N(2,a) &= \frac{1}{N-1} \sum_{i=1}^N \frac{\left(Z(\frac{i}{N}) - Z(\frac{i-1}{N})\right)^2}{\mathbf{E}\left(Z(\frac{i}{N}) - Z(\frac{i-1}{N})\right)^2} - 1 \\ &= N^{2H-1} \sum_{i=1}^N \left[\left(Z(\frac{i}{N}) - Z(\frac{i-1}{N}) \right)^2 - N^{-2H} \right]. \end{aligned}$$

Recall the product formula for multiple Wiener-Itô integrals (4) to obtain

$$I_2(f) I_2(g) = I_4(f \otimes f) + 4 I_2(f \otimes_1 f) + 2 I_0(f \otimes_2 f);$$

overriding the notation A_i in the previous section, we set

$$A_i := L_{\frac{i}{N}} - L_{\frac{i-1}{N}}; \tag{34}$$

we can thus write

$$\begin{aligned}&\left(Z(\frac{i}{N}) - Z(\frac{i-1}{N})\right)^2 = (I_2(A_i))^2\\ &= I_4(A_i \otimes A_i) + 4I_2(A_i \otimes_1 A_i) + 2I_0(A_i \otimes_2 A_i)\end{aligned}$$

and this implies that the 2-variation is decomposed into a 4th chaos term and a 2nd chaos term:

$$\begin{aligned}V_N(2,a) &= N^{2H-1}\sum_{i=1}^{N}\left(I_4(A_i \otimes A_i) + 4I_2(A_i \otimes_1 A_i)\right)\\ &:= T_4 + T_2.\end{aligned}$$

A detailed study of the two terms above will shed light on some interesting facts: $H \leq \frac{3}{4}$ the term T_4 continue to exihibit "normal" behavior (renormalized, it converges in law to a Gaussian distribution), while the term T_2, which turns out to be dominant, never converges to a Gaussian law. One can say that the second Wiener chaos portion is "ill-behaved"; however, once it is subtracted, one obtains a sequence converging to $\mathcal{N}(0,1)$, which has an impact for statistical applications.

4.1 Expectation evaluations

Let us first estimate the L^2 norm of the term denoted by T_4. By the isometry formula for multiple stochastic integrals

$$\mathbf{E}(T_4^2) = 4!N^{4H-2}\sum_{i,j=1}^{N}\langle A_i \otimes A_i, A_j \otimes A_j\rangle_{L^2[0,1]^4} = 4!N^{4H-2}\sum_{i,j=1}^{N}\left|\langle A_i, A_j\rangle_{L^2[0,1]^2}\right|^2.$$

The scalar product $\langle A_i, A_j\rangle^2_{L^2[0,1]^2}$ computes as

$$\begin{aligned}\left|\langle A_i, A_j\rangle_{L^2[0,1]^2}\right|^2 &= \int_0^1\int_0^1 A_i(y_1,y_2)A_j(y_1,y_2)dy_1dy_2\\ &= d(H)^2\int_0^1\int_0^1 dy_1dy_2\\ &\left(\int_{y_1\vee y_2}^{\frac{i}{N}}\frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial u}(u,y_2)du - \int_{y_1\vee y_2}^{\frac{i-1}{N}}\frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial u}(u,y_2)du\right)\\ &\left(\int_{y_1\vee y_2}^{\frac{j}{N}}\frac{\partial K^{H'}}{\partial v}(v,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_2)dv - \int_{y_1\vee y_2}^{\frac{j-1}{N}}\frac{\partial K^{H'}}{\partial v}(v,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_2)dv\right)\\ &= d(H)^2\int_{\frac{i-1}{N}}^{\frac{i}{N}}\int_{\frac{j-1}{N}}^{\frac{j}{N}}dudv\left[\int_0^{u\wedge v}\frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_1)dy_1\right]^2.\end{aligned}$$

Note the following fact, with $a(H) = H'(2H'-1) = H(H+1)/2$:

$$\int_0^{u\wedge v}\frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_1)dy_1 = a(H')|u-v|^{2H'-2}; \tag{35}$$

this will be used repeatedly, and in this case it gives

$$\left|\langle A_i, A_j\rangle_{L^2[0,1]^2}\right|^2 = d(H)^2 a(H)^2 \int_{I_i}\int_{I_j} |u-v|^{2H-2} du dv$$

where we used the notation $I_i = \left(\frac{i-1}{N}, \frac{i}{N}\right]$.

We finally obtain

$$\left|\langle A_i, A_j\rangle_{L^2[0,1]^2}\right|^2 = \frac{d(H)^2 a(H)^2}{H(2H-1)}\frac{1}{2}\left[2\left|\frac{i-j}{N}\right|^{2H} - \left|\frac{i-j+1}{N}\right|^{2H} - \left|\frac{i-j-1}{N}\right|^{2H}\right] \tag{36}$$

and this expression has already been studied in the previous section. With the constants $c_{1,H}$, $c_{2,H}$, and $c'_{1,H}$ given respectively in (7), (9), and (12), using Lemmas 16, 17, and the analogous result for $H = 3/4$, we get, asymptotically for large N,

$$\mathbf{E}(T_4^2) \sim \frac{1}{N} 4! \frac{d(H)^4 a(H)^4}{(H(2H-1))^2}\frac{1}{2} c_{1,H}, \quad H < \frac{3}{4}, \tag{37}$$

$$\mathbf{E}(T_4^2) \sim N^{4H-4} 4! \frac{d(H)^4 a(H)^4}{(H(2H-1))^2} c_{2,H}, \quad H > \frac{3}{4}. \tag{38}$$

and

$$\mathbf{E}(T_4^2) \sim 4! \frac{d(H)^4 a(H)^4}{(H(2H-1))^2} c'_{1,H} \frac{\log N}{N}, \quad H = \frac{3}{4}. \tag{39}$$

Let us evaluate now the mean square of the second term

$$T_2 := N^{2H-1} 4 \sum_{i=1}^{N} I_2(A_i \otimes_1 A_i);$$

the contraction $A_i \otimes_1 A_i$ is given by

$$\begin{aligned}
&(A_i \otimes_1 A_i)(y_1, y_2) = \int_0^1 \int_0^1 A_i(x, y_1) A_i(x, y_2) dx \\
&= d(H)^2 \int_0^1 dx \\
&\left(\int_{x\wedge y_1}^{\frac{i}{N}} \frac{\partial K^{H'}}{\partial u}(u,x)\frac{\partial K^{H'}}{\partial u}(u,y_1)du - \int_{x\wedge y_1}^{\frac{i-1}{N}} \frac{\partial K^{H'}}{\partial u}(u,x)\frac{\partial K^{H'}}{\partial u}(u,y_1)du\right) \\
&\left(\int_{x\wedge y_2}^{\frac{i}{N}} \frac{\partial K^{H'}}{\partial v}(v,x)\frac{\partial K^{H'}}{\partial v}(v,y_2)dv - \int_{x\wedge y_2}^{\frac{i-1}{N}} \frac{\partial K^{H'}}{\partial v}(v,x)\frac{\partial K^{H'}}{\partial v}(v,y_2)dv\right) \\
&= d(H)^2 \int_{I_i}\int_{I_i} du dv \left(\int_0^{u\wedge v} dx \frac{\partial K^{H'}}{\partial u}(u,x)\frac{\partial K^{H'}}{\partial v}(v,x)\right) \frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_2)
\end{aligned}$$

and we obtain, using (35)

$$(A_i \otimes_1 A_i)(y_1, y_2) = a(H) d(H)^2 \int_{I_i}\int_{I_i} |u-v|^{2H'-2} \frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_2) du dv. \tag{40}$$

Then

$$\mathbf{E}(T_2^2) = N^{4H-2}16 \cdot 2! \sum_{i,j=1}^{N} \langle A_i \otimes_1 A_i, A_j \otimes_1 A_j \rangle_{L^2[0,1]^2}$$

where

$$\begin{aligned}
&\langle A_i \otimes_1 A_i, A_j \otimes_1 A_j \rangle_{L^2[0,1]^2} \\
&= \int_0^1 \int_0^1 dy_1 dy_2 (A_i \otimes_1 A_i)(y_1, y_2)(A_j \otimes_1 A_j)(y_1, y_2) \\
&= a(H)^2 d(H)^4 \int_0^1 \int_0^1 dy_1 dy_2 \int_{I_i} \int_{I_i} \int_{I_j} \int_{I_j} du' dv' du dv \\
&|u-v|^{2H'-2}|u'-v'|^{2H'-2} \frac{\partial K^{H'}}{\partial u}(u, y_1) \frac{\partial K^{H'}}{\partial v}(v, y_2) \frac{\partial K^{H'}}{\partial u'}(u', y_1) \frac{\partial K^{H'}}{\partial v'}(v', y_2) \\
&= a(H)^4 d(H)^4 \int_{I_i} \int_{I_i} \int_{I_j} \int_{I_j} |u-v|^{2H'-2}|u'-v'|^{2H'-2}|u-u'|^{2H'-2}|v-v'|^{2H'-2} du' dv' dv du
\end{aligned}$$

and so

$$\begin{aligned}
\mathbf{E}(T_2^2) = N^{4H-2}32a(H)^4 d(H)^4 \sum_{i,j=1}^{N} \int_{I_i} \int_{I_i} \int_{I_j} \int_{I_j} \\
\times |u-v|^{2H'-2}|u'-v'|^{2H'-2}|u-u'|^{2H'-2}|v-v'|^{2H'-2} du' dv' dv du
\end{aligned} \tag{41}$$

By Lemma 20 in the Appendix, we conclude that

$$\begin{aligned}
\lim_{N\to\infty} \mathbf{E}(T_2^2) N^{2-2H} &= 64a(H)^4 d(H)^4 \left(\frac{1}{H} - \frac{1}{H+1}\right) \left| \int_{[0,1]^2} |u-v|^{2H'-2} du dv \right|^2 \qquad (42) \\
&= 64a(H)^2 d(H)^4 \left(\frac{1}{H} - \frac{1}{H+1}\right) \\
&= 16d\,(H)^2 := c_{3,H}. \qquad (43)
\end{aligned}$$

Taking into account the estimations (37) and (42), we see that $\mathbf{E}\left[T_4^2\right]$ is of order N^{-1}, which is smaller than $\mathbf{E}\left[T_2^2\right]$ which is of order N^{H-1}; therefore the mean-square behavior of V_N is given by that of the term T_2 only, which means we obtain for every $H > 1/2$

$$\lim_{N\to\infty} \mathbf{E}\left[\left(N^{1-H} V_N(2,a) \frac{1}{\sqrt{c_{3,H}}}\right)^2\right] = 1. \tag{44}$$

4.2 Normality of the 4th chaos term T_4 when $H \leq 3/4$

We proved above that $\lim_{N\to\infty} \mathbf{E}(G_N^2) = 1$ for $H < 3/4$ where

$$G_N := \sqrt{N} N^{2H-1} e_{1,H}^{-1/2} I_4 \left(\sum_{i=1}^{N} A_i \otimes A_i\right) \tag{45}$$

and

$$e_{1,H} = \frac{4! d(H)^4 a(H)^4 c_{1,H}}{2(H(2H-1))^2}. \tag{46}$$

Similarly, for $H = \frac{3}{4}$, we showed that $\lim_{N\to\infty} \mathbf{E}(\tilde{G}_N)^2 = 1$ where

$$\tilde{G}_N := \sqrt{\frac{N}{\log N}} N^{2H-1} e_{3,H}^{-1} I_4 \left(\sum_{i=1}^{N} A_i \otimes A_i \right) \tag{47}$$

and

$$e_{3,H} := 4! \frac{d(H)^4 a(H)^4}{(H(2H-1))^2} c'_{1,H} \tag{48}$$

Using the criterion of Nualart and Ortiz-Latorre, we prove the following asymptotic normality for G_N and $\tilde{G}_N$.

Theorem 7 *If $H \in (1/2, 3/4)$, then G_N given by (45) converges in distribution as*

$$\lim_{N\to\infty} G_N = \mathcal{N}(0,1). \tag{49}$$

If $H = 3/4$ then $\tilde{G}_N$ given by (47) converges in distribution as

$$\lim_{N\to\infty} \tilde{G}_N = \mathcal{N}(0,1). \tag{50}$$

Proof. *Step 0: setup and expectation evaluation.* Using the derivation rule for multiple stochastic integrals, the Malliavin derivative of G_N is

$$D_r G_N = \sqrt{N} N^{2H-1} e_{1,H}^{-1/2} 4 \sum_{i=1}^{N} I_3(A_i \otimes A_i)(\cdot, r)$$

and its norm is

$$\|DG_N\|^2_{L^2[0,1]} = N^{4H-1} 16 e_{1,H}^{-1} \sum_{i,j=1}^{N} \int_0^1 dr I_3(A_i \otimes A_i)(\cdot, r) I_3(A_j \otimes A_j)(\cdot, r).$$

The product formula (4) gives

$$\begin{aligned}
\|DG_N\|^2_{L^2[0,1]} = N^{4H-1} 16 e_{1,H}^{-1} \sum_{i,j=1}^{N} \int_0^1 dr \, [& I_6 \left((A_i \otimes A_i)(\cdot, r) \otimes (A_j \otimes A_j)(\cdot, r) \right) \\
& + 9 I_4 \left((A_i \otimes A_i)(\cdot, r) \otimes_1 (A_j \otimes A_j)(\cdot, r) \right) \\
& + 9 I_2 \left((A_i \otimes A_i)(\cdot, r) \otimes_2 (A_j \otimes A_j)(\cdot, r) \right) \\
& + 3! I_0 \left((A_i \otimes A_i)(\cdot, r) \otimes_3 (A_j \otimes A_j)(\cdot, r) \right)] \\
& := J_6 + J_4 + J_2 + J_0.
\end{aligned}$$

First note that, for the non-random term J_0 that gives the expected value of the above, we have

$$\begin{aligned}
J_0 &= 16 e_{1,H} N^{4H-1} 3! \sum_{i,j=1}^{N} \int_{[0,1]^4} A_i(y_1, y_2) A_i(y_3, y_4) A_j(y_1, y_2) A_j(y_3, y_4) dy_1 dy_2 dy_3 dy_4 \\
&= 96 N^{4H-1} e_{1,H}^{-1} \sum_{i,j=1}^{N} \left| \langle A_i, A_j \rangle_{L^2[0,1]^2} \right|^2.
\end{aligned}$$

This sum has already been treated: we know from (37) that $J_0/4$ converges to 1, i.e. that $\lim_{N\to\infty} \mathbf{E}[\|DG_N\|^2_{L^2[0,1]}] = 4$. This mean, by the Nualart–Ortiz-Latorre criterion, that we only need to show that all other terms J_6, J_4, J_2 converge to zero in $L^2(\Omega)$ as $N \to \infty$.

Step 1: order-6 chaos term. We consider first the term J_6:

$$\begin{aligned} J_6 &= N^{4H-1} 16 e_{1,H}^{-1} \sum_{i,j=1}^{N} \int_0^1 dr I_6\left((A_i \otimes_1 A_i) \otimes (A_j \otimes_1 A_j(\cdot, r))\right) \\ &= N^{4H-1} 16 e_{1,H}^{-1} \sum_{i,j=1}^{N} I_6\left((A_i \otimes A_j) \otimes (A_i \otimes_1 A_j(\cdot, r))\right). \end{aligned}$$

We study the mean square of this term. We have

$$\begin{aligned} &\mathbf{E}\left[\left(\sum_{i,j=1}^{N} I_6\left((A_i \otimes A_j) \otimes (A_i \otimes_1 A_j(\cdot, r))\right)\right)^2\right] \\ &= 6! \sum_{i,j,k,l} \langle (A_i \otimes A_j) \otimes (A_i \otimes_1 A_j(\cdot, r)), (A_k \otimes A_l) \otimes (A_k \otimes_1 A_l(\cdot, r))\rangle \\ &= 6! \sum_{i,j,k,l} \langle A_i, A_k\rangle_{L^2[0,1]^2} \langle A_j, A_l\rangle_{L^2[0,1]^2} \langle A_i \otimes_1 A_j, A_k \otimes_1 A_l\rangle_{L^2[0,1]^2} \\ &= 6! \sum_{i,j,k,l} \langle A_i, A_k\rangle_{L^2[0,1]^2} \langle A_j, A_l\rangle_{L^2[0,1]^2} \int_0^1 \int_0^1 dy_1 dy_2 \\ &\times \int_{I_i} \int_{I_j} dv du a(H) |u - v|^{2H'-2} \frac{\partial K^{H'}}{\partial u}(u, y_1) \frac{\partial K^{H'}}{\partial v}(v, y_2) \\ &\times \int_{I_k} \int_{I_l} dv' du' a(H) |u - v|^{2H'-2} \frac{\partial K^{H'}}{\partial u'}(u', y_1) \frac{\partial K^{H'}}{\partial v'}(v', y_2) \end{aligned}$$

We get

$$\begin{aligned} \mathbf{E}\left[J_6^2\right] &= cst \cdot N^{8H-2} \sum_{i,j,k,l} \langle A_i, A_k\rangle_{L^2[0,1]^2} \langle A_j, A_l\rangle_{L^2[0,1]^2} \\ &\times \int_{I_i} du \int_{I_j} dv \int_{I_k} du' \int_{I_l} dv' |u - v|^{2H'-2} |u - u'|^{2H'-2} |v - v'|^{2H'-2} |u' - v'|^{2H'-2} \\ &\times \left[2\left|\frac{i-k}{N}\right|^{2H} - \left|\frac{i-k+1}{N}\right|^{2H} - \left|\frac{i-k-1}{N}\right|^{2H}\right] \\ &\times \left[2\left|\frac{j-l}{N}\right|^{2H} - \left|\frac{j-l+1}{N}\right|^{2H} - \left|\frac{j-l-1}{N}\right|^{2H}\right]. \end{aligned}$$

Lemma 21 in the Appendix then proves that

$$\mathbf{E}\left[J_6^2\right] \sim cst \cdot N^{8H-6} \tag{51}$$

and this goes to zero as N goes to ∞ because $H < \frac{3}{4}$.

Step 2: chaos terms of order 4 and 2. To treat the term

$$J_4 = cst \cdot N^{4H-1} \sum_{i,j=1}^{N} \int_0^1 dr I_4 \left((A_i \otimes A_i)(\cdot, r) \otimes_1 (A_j \otimes A_j)(\cdot, r) \right),$$

since $I_4(g) = I_4(\tilde{g})$ where $\tilde{g}$ denotes the symmetrization of the function g, we can write

$$\begin{aligned} J_4 = cst \cdot N^{4H-1} \sum_{i,j=1}^{N} \langle A_i, A_j \rangle_{L^2(0,1]^2} (A_i \otimes A_j) \\ + cst \cdot N^{4H-1} \sum_{i,j=1}^{N} (A_i \otimes_1 A_j) \otimes (A_i \otimes_1 A_j). \end{aligned}$$

Both terms above have been treated in previous computations. One can show that

$$\mathbf{E}\left[J_4^2\right] \sim cst \cdot N^{8H-6} \tag{52}$$

which tends to zero for $H < \frac{3}{4}$. A similar behavior can be obtained for the last term J_2

$$\mathbf{E}\left[J_2^2\right] \sim cst.N^{8H-6} \tag{53}$$

Step 3: conclusion. Putting (51), (52), (53) and (27) together, and recalling the convergence result for $\mathbf{E}\left[T_4^2\right]$ proved in the previous subsection, we can apply the Nualart–Ortiz-Latorre criterion, and use the same method as in the case $H < \frac{3}{4}$ for $H = 3/4$, to conclude the theorem's proof. ■

Remark 8 *The convergences in Theorem 7 also hold in* $L^2(\Omega)$.

It is also possible to prove a multidimensional version of the above limit theorem. For every $p = 1, \ldots, M$ let us consider the 4th chaos term appearing in the decomposition of $V_N(2, a^p)$ where a^p is the filter of length $p+1$ constructed from the initial filter $a = \{1, -1\}$

$$T_4(a^p) = N^{2H-1} \sum_{i=p}^{N} I_4 \left(A_{i,p} \otimes A_{i,p} \right)$$

with

$$A_{i,p} = L_{\frac{i}{N}} - L_{\frac{i-p}{N}}, \quad 1 \le p \le i, 1 \le i \le M.$$

Theorem 9 *Assume* $H < \frac{3}{4}$. *Define, for* $1 \le p \le M$,

$$e_{p,q,H} := \frac{d(H)^2 a(H)^2}{H(2H-1)} \frac{1}{2} \frac{1}{(pq)^{2H}} \sum_{k \ge 1} \left(|k|^{2H} + |k-p+q|^{2H} - |k-p|^{2H} - |k+q|^{2H} \right)^2, \; e_{p,H} := e_{p,p,H},$$

and

$$G_N(a^p) := \sqrt{N} e_{p,H}^{-1/2} T_4(a^p) = \sqrt{N} e_{p,H}^{-1/2} \sum_{i=p}^{N} I_4 \left(A_{i,p} \otimes A_{i,p} \right).$$

Then the vector

$$(G_N(a^1, \dots, G_N(a^M))$$

converges to a centered M dimensional Gaussian vector with covariance matrix $E = E_{i,j}$ where $E_{p,q} = \frac{e_{p,q,H}}{\sqrt{e_{p,H} e_{q,H}}}$.

If $H = \frac{3}{4}$, define

$$f_{p,q,H} := \frac{d(H)^2 a(H)^2}{H(2H-1)} \frac{1}{2} \frac{1}{(pq)^{2H}} \left(2H(2H-1)pq\right)^2, \; f_{p,H} := f_{p,p,H},$$

and

$$\tilde{G}_N = \sqrt{\frac{N}{\log N}} f_{p,H}^{-1/2} T_4(a^p).$$

Then the vector

$$(\tilde{G}_N(a^1, \dots, \tilde{G}_N(a^M))$$

converges to a centered M-dimensional Gaussian vector with covariance matrix $F = F_{i,j}$ where $F_{p,q} = \frac{f_{p,q,H}}{\sqrt{f_{p,H} f_{q,H}}}$.

Proof. This is a consequence of the result in [18] and the estimation of the covariances. For example, if $H < \frac{3}{4}$ we have

$$\lim_N \mathbf{E}\left[G_N(a^p) G_N(a^q)\right] = e_{p,q,H}$$

and if $H = \frac{3}{4}$ then

$$\lim_N \mathbf{E}\left[\tilde{G}_N(a^p) \tilde{G}_N(a^q)\right] = f_{p,q,H}.$$

■

4.3 Anormality of the second chaos term T_2, and limit of the 2-variation

This paragraph studies the asymptotic behavior of the term denoted by T_2 which appears in the decomposition of $V_N(2, a)$. Recall that this is the dominant term, given by

$$T_2 = 4N^{2H-1} I_2 \left(\sum_{i=1}^{N} (A_i \otimes_1 A_i) \right)$$

and, with $c_{3,H}$ given in (43), we showed that

$$\lim_{N \to \infty} \mathbf{E}\left[\left(N^{1-H} T_2 c_{3,H}^{-1/2} \right)^2 \right] = 1.$$

The following proposition, which can be proved using straightforward calculation which we omit, shows by the Nualart–Ortiz-Latorre criterion that T_2 cannot converge to a normal distribution.

Proposition 10 *With $T_N := N^{1-H} T_2 c_{3,H}^{-1/2}$, then, in $L^2(\Omega)$, $\lim_{N\to\infty} \|DT_N\|^2_{L^2[0,1]} = 2 + c$ where c is a strictly positive constant.*

However, it is straightforward to find the limit of T_2, and thus of V_N, in $L^2(\Omega)$ in this case. We have the following result.

Theorem 11 *For all $H \in (1/2,1)$, the normalized 2-variation $N^{1-H}V_N(2,a)/(4d(H))$ converges in $L^2(\Omega)$ to the Rosenblatt random variable $Z(1)$. Note that this is the actual observed value of the Rosenblatt process at time 1.*

Proof. The proof proceeds exactly as in the proof of the case $H \in (3/4,1)$ in Theorem 2. Since we already proved that $N^{1-H}T_4$ converges to 0 in $L^2(\Omega)$, it is sufficient to prove that $N^{1-H}T_2/(4d(H)) - Z(1)$ converges to 0 in $L^2(\Omega)$. Since T_2 is a second-chaos random variable, i.e. is of the form $I_2(f_N)$ where f_N is a symmetric function in $L^2\left([0,1]^2\right)$, it is sufficient to prove that f_N converges to L_1 in $L^2\left([0,1]^2\right)$, where L_1 is given by (33). We have already calculated f_N fairly explicitly in (40): we have that

$$f_N(y_1,y_2) = 4N^{2H-1}a(H)\,d(H)^2 \sum_{i=1}^{N} \iint_{I_i\times I_i} dudv\,|u-v|^{2H'-2}\,\frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial v}(v,y_2)$$

where the summand is zero if $i/N < y_1 \vee y_2$. The detailed approximation arguments go exactly as in the proof of Theorem 2. We give the main lines here only. On the interval $I_i \times I_i$, we may replace the evaluation of $\frac{\partial K^{H'}}{\partial u}$ and $\frac{\partial K^{H'}}{\partial v}$ at u and v by setting $u = v = i/N$. We then get that $f_N(y_1,y_2)$ is asymptotically equivalent to

$$4N^{2H-1}a(H)\,d(H)^2 \sum_{i=1}^{N} \mathbf{1}_{i/N<y_1\vee y_2}\frac{\partial K^{H'}}{\partial u}(i/N,y_1)\frac{\partial K^{H'}}{\partial v}(i/N,y_2)\iint_{I_i\times I_i} dudv\,|u-v|^{2H'-2}$$
$$= 4N^{H-1}d(H)^2\,\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{i/N<y_1\vee y_2}\frac{\partial K^{H'}}{\partial u}(i/N,y_1)\frac{\partial K^{H'}}{\partial v}(i/N,y_2).$$

The above series normalized by N^{-1} is equivalent to a Riemann integral proportional to $L_1(y_1,y_2)$; this justifies that $N^{1-H}f_N$ converges in $L^2\left([0,1]^2\right)$ to $4d(H)\,L_1$, which finishes the proof of the theorem. ■

Remark 12 *As in Remark 4, one can show that the 2-variations $V_N(2,a)$ converge to zero almost surely as N goes to infinity.*

4.4 Normality of the adjusted variations

According to Theorem 11 which we just proved, in the Rosenblatt case, the standardization of the random variable $V_N(2,a)$ does not converge to the normal law. But this statistics, that can be written as $V_N = T_4 + T_2$ has a small *normal part*, which is given by the asymptotics of the term T_4, as we can see from Theorem 7. Therefore, $V_N - T_2$ will converge (under suitable scaling) to the Gaussian distribution. Of course, the term T_2 which represents an iterated stochastic integral is not practical because it cannot be observed. But, replacing it with its limit $Z(1)$ (this *IS* observed) one can defined an adjusted version of the statistics V_N that converges, after standardization, to the normal law.

Theorem 13 *Let $(Z(t), t \in [0,1])$ be a Rosenblatt process with self-similarity index $H \in (\frac{1}{2},1)$ and let previous notations for constants prevail. Then, in distribution,*

if $H < \frac{3}{4}$,

$$\lim_{N\to\infty} \frac{\sqrt{N}}{\sqrt{e_{1,H}}}\left[V_N(2,a) - \frac{\sqrt{c_{3,H}}}{N^{1-H}}Z(1)\right] = \mathcal{N}(0,1),$$

if $H = \frac{3}{4}$,

$$\lim_{N\to\infty} \frac{\sqrt{N/\log N}}{\sqrt{e_{3,H}}} \left[V_N(2,a) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1) \right] = \mathcal{N}(0,1)$$

Proof. *Step 0: setup.* We have

$$\begin{aligned} \frac{\sqrt{N}}{\sqrt{e_{3,H}}} \left[V_N(2,a) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1) \right] &= \frac{\sqrt{N}}{\sqrt{e_{3,H}}} \left[V_N(2,a) - T_2 + T_2 - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1) \right] \\ &= \frac{\sqrt{N}}{\sqrt{e_{3,H}}} T_4 + \frac{\sqrt{N}}{\sqrt{e_{3,H}}} \frac{\sqrt{c_{3,H}}}{N^{1-H}} \left[\frac{N^{1-H}}{\sqrt{c_{3,H}}} T_2 - Z(1) \right]. \end{aligned}$$

We have already proved that $\sqrt{N/e_{3,H}} T_4$ converges in law as $N \to \infty$ to the standard normal law $N(0,1)$. It then suffices to prove that

$$\lim_{N\to\infty} N^{H-\frac{1}{2}} \left[\frac{N^{1-H}}{\sqrt{c_{3,H}}} T_2 - Z(1) \right] = 0 \text{ in } L^2(\Omega). \tag{54}$$

In fact $\sqrt{c_{3,H}} = 4d(H)$. Let us estimate the difference

$$\frac{N^{1-H}}{4d(H)} T_2 - Z(1).$$

As in the proof of Theorem 11 we can write $T_2 = I_2(f_N)$ where

$$f_N(y_1,y_2) = 4N^{2H-1} a(H)\, d(H)^2 \sum_{i=1}^{N} \iint_{I_i\times I_i} du dv\, |u-v|^{2H'-2} \frac{\partial K^{H'}}{\partial u}(u,y_1) \frac{\partial K^{H'}}{\partial v}(v,y_2).$$

Then

$$\mathbf{E}\left[\left| \frac{N^{1-H}}{4d(H)} T_2 - Z(1) \right|^2 \right] = 2! \left\| \frac{N^{1-H}}{4d(H)} f_N - L_1 \right\|^2_{L^2([0,1]^2)}$$

where L_1 is the kernel of the Rosenblatt process at time 1 (see (33)).

For every y_1, y_2,

$$\begin{aligned} &\frac{N^{1-H}}{4d(H)} f_N(y_1,y_2) - L_1(y_1,y_2) \\ &= N^H a(H) d(H) \sum_{i=1}^{N} \iint_{I_i\times I_i} du dv\, |u-v|^{2H'-2} \\ &\times \left[\frac{\partial K^{H'}}{\partial u}(u,y_1) \frac{\partial K^{H'}}{\partial v}(v,y_2) - \partial_1 K^{H'}(\frac{i}{N},y_1) \partial_1 K^{H'}(\frac{i}{N},y_2) \right] \\ &+ N^H a(H) d(H) \sum_{i=1}^{N} \left(\iint_{I_i\times I_i} du dv\, |u-v|^{2H'-2} \right) \partial_1 K^{H'}(\frac{i}{N},y_1) \partial_1 K^{H'}(\frac{i}{N},y_2) - L_1(y_1,y_2) \\ &=: A(y_1,y_2) + B(y_1,y_2). \end{aligned}$$

Step 1: estimating the term A. We will show that

$$\mathbf{E}\left|N^{H-\frac{1}{2}}\left\|A\right\|_{L^2([0,1]^2)}\right|^2 \to 0 \text{ as } N\to\infty. \tag{55}$$

By repeating some previous calculations, we obtain

$$\begin{aligned}
&\mathbf{E}\left|N^{H-\frac{1}{2}}\left\|A\right\|_{L^2([0,1]^2)}\right|^2\\
&\leq N^{2H-1}c(H)N^{2H}\sum_{i,j=1}^{N}\int_{I_i}\int_{I_i}\int_{I_j}\int_{I_j}du'dv'dudv|u-v|^{2H'-2}|u'-v'|^{2H'-2}\\
&\times\left[|u-u'|^{2H'-2}|v-v'|^{2H'-2}-|u-\frac{j}{n}|^{2H'-2}|v-\frac{j}{n}|^{2H'-2}\right.\\
&\left.-|u'-\frac{i}{n}|^{2H'-2}|v'-\frac{i}{n}|^{2H'-2}+|\frac{i}{N}-\frac{j}{N}|^{2H-2}\right].
\end{aligned}$$

The change of variables $\bar{u}=\left(u-\frac{i}{N}\right)N$ (and similar operations for the other integrals) imply

$$\begin{aligned}
&\mathbf{E}\left|N^{H-\frac{1}{2}}\left\|A\right\|_{L^2([0,1]^2)}\right|^2\\
&\leq N^{2H-1}c(H)N^{-2H}\sum_{i,j=1}^{N}\int_{[0,1]^4}dudvdu'dv'|u-v|^{2H'-2}|u'-v'|^{2H'-2}\\
&\times\left[|u-u'+i-j|^{2H'-2}|v-v'+i-j|^{2H'-2}-|u+i-j-1|^{2H'-2}|v+i-j-1|^{2H'-2}\right.\\
&\left.-|u'+j-i-1|^{2H'-2}|v'+j-i-1|^{2H'-2}+|i-j|^{2H'-2}\right].
\end{aligned}$$

We need to study the non-diagonal part which is again the dominant part. This is equal to (modulo constants $c(H)$)

$$\begin{aligned}
&\frac{1}{N}\sum_{k=1}^{N}(N-k)\int_{[0,1]^4}dudvdu'dv'|u-v|^{2H'-2}|u'-v'|^{2H'-2}\\
&\times\left[|u-u'+k|^{2H'-2}|v-v'+k|^{2H'-2}-|u+k-1|^{2H'-2}|v+k-1|^{2H'-2}\right.\\
&\left.-|u'-k-1|^{2H'-2}|v'-k-1|^{2H'-2}+|k|^{2H'-2}\right]\\
&=\frac{1}{N}\sum_{k=1}^{N}(N-k)k^{2H-2}\int_{[0,1]^4}dudvdu'dv'|u-v|^{2H'-2}|u'-v'|^{2H'-2}F_{u,v,u',v'}(k)
\end{aligned}$$

A study of the asymptotic behavior of $F_{u,v,u',v'}(k)$ show that it converges to the constant $\frac{1}{k}$ for every u,v,u',v'. Therefore $\frac{1}{N}\sum_{k=1}^{N}(N-k)k^{2H-2}F_{u,v,u',v'}(k)$ will be of the order of N^{2H-2} and an application of the dominated convergence theorem gives (55).

Step 2: estimating the term B. Recall that

$$\begin{aligned}
B&:=N^Ha(H)d(H)\left(\iint_{I_i\times I_i}dudv\,|u-v|^{2H'-2}\right)\partial_1K^{H'}(\frac{i}{N},y_1)\partial_1K^{H'}(\frac{i}{N},y_2)-L_1(y_1,y_2)\\
&=\frac{1}{N}d(H)\sum_{i=1}^{N}1_{y_1\vee y_2\leq\frac{i}{N}}\partial_1K^{H'}(\frac{i}{N},y_1)\partial_1K^{H'}(\frac{i}{N},y_2)-d(H)\int_{y_1\vee y_2}^{1}\frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial u}(u,y_2)du.
\end{aligned}$$

We recognize here the difference between a Riemann sum and its associated integral, implying that the above term goes to zero for fixed y_1 and y_2. But we need more, that is, to control the speed of this convergence. We will show

$$\lim_{N\to\infty} \mathbf{E}\left|N^{H-\frac{1}{2}}B\right|^2 = 0. \tag{56}$$

One can write

$$B = d(H)\sum_{i=0}^{n-1}\int_{I_{i+1}}\left[1_{y_1\vee y_2\le\frac{i}{N}}\partial_1 K^{H'}(\frac{i}{N},y_1)\partial_1 K^{H'}(\frac{i}{N},y_2) - \frac{\partial K^{H'}}{\partial u}(u,y_1)\frac{\partial K^{H'}}{\partial u}(u,y_2)du\right]$$

and then

$$\begin{aligned}
&\mathbf{E}\left|N^{H-\frac{1}{2}}B\right|^2\\
&= \sum_{i,j=1}^{N}\int_{I_i}\int_{I_j} du'du\\
&\left[\left|\frac{i}{N}-\frac{j}{N}\right|^{2H'-2} - \left|\frac{i}{N}-u'\right|^{2H'-2} - \left|u-\frac{j}{N}\right|^{2H'-2} + |u-v|^{2H'-2}\right]\\
&= c(H)N^{2H-1}\frac{1}{N^{2H'-2}}\frac{1}{N^2}\sum_{i,j}\int_{[0,1]^2} dudu'\\
&\left[|i-j|^{2H'-2} - |u+i-j|^{2H'-2} - |u'+j-i|^{2H'-2} + |u-u'+i-j|^{2H'-2}\right].
\end{aligned}$$

The behavior for large N of the above expression will be given by

$$N^{H-2}\sum_k (N-k)\left[k^{2H'-2} - |u+k|^{2H'-2} - |u'-k|^{2H'-2} + |u-u'+k|^{2H'-2}\right]$$

and since $\sum_k (N-k)\left[k^{2H'-2} - |u+k|^{2H'-2} - |u'-k|^{2H'-2} + |u-u'+k|^{2H'-2}\right]$ is equivalent to $N^{2H'-1} = N^H$ it holds that $\mathbf{E}\left|N^{H-\frac{1}{2}}B\right|^2$ behaves as N^{2H-2} and this clearly implies (56). ■

5 The estimators for the self-similarity index

In this part we construct estimators for the self-similarity exponent of a Hermite process based on the discrete observations of the driving process at times $0, \frac{1}{N}, \ldots, 1$. It is known that the asymptotic behavior of the statistics $V_N(2,a)$ is related to the asymptotic properties of a class of estimators for the Hurst parameter H. This is mentioned for instance in [3].

We recall the setup of how this works. Suppose that the observed process X is a Hermite process; it may be Gaussian (fractional Brownian motion) or non-Gaussian (Rosenblatt process or even a higher order Hermite process). With $a = \{-1, +1\}$, the 2-variation is denoted by

$$S_N(2,a) = \frac{1}{N}\sum_{i=1}^{N}\left(X(\frac{i}{N}) - X(\frac{i-1}{N})\right)^2; \tag{57}$$

Recall that $\mathbf{E}\left[S_N(2,a)\right] = N^{-2H}$.By estimating $\mathbf{E}\left[S_N(2,a)\right]$ by $S_N(2,a)$ we can construct the estimator

$$\hat{H}_N(2,a) = -\frac{\log S_N(2,a)}{2\log N}. \tag{58}$$

To prove that this is a strongly consistent estimator for H, we begin by writing

$$1 + V_N(2,a) = S_N(2,a)N^{2H}$$

where V_N is the original quantity defined in (3), and thus

$$\begin{aligned}\log\left(1 + V_N(2,a)\right) &= \log S_N(2,a) + 2H\log N \\ &= -2(\hat{H}_N(2,a) - H)\log N.\end{aligned}$$

Moreover, by Remarks 4 and 12, $V_N(2,a)$ converges almost surely to 0, and thus $\log\left(1 + V_N(2,a)\right) = V_N(2,a)(1 + o(1))$ where $o\left(1\right)$ converges to 0 almost surely as $N \to \infty$. Hence we obtain

$$V_N(2,a) = 2(H - \hat{H}_N(2,a))\left(\log N\right)(1 + o(1)). \tag{59}$$

Relation (59) means that V_N's behavior immediately give the behavior of $\hat{H}_N - H$.

Specifically, we can now state our convergence results. First, the Gaussian case.

Theorem 14 *Suppose that $H > \frac{1}{2}$ and assume that the observed process is a fBm with Hurst parameter H. Then strong consistency holds for $\hat{H}_N$, i.e. almost surely,*

$$\lim_{N\to\infty} \hat{H}_N(2,a) = H \tag{60}$$

and

- *if $H \in (\frac{1}{2}, \frac{3}{4})$, then, in distribution as $N \to \infty$,*

$$\sqrt{N}\log(N)\frac{2}{\sqrt{c_{1,H}}}\left(\hat{H}_N(2,a) - H\right) \to \mathcal{N}(0,1)$$

- *if $H \in (\frac{3}{4}, 1)$, then, in distribution as $N \to \infty$,*

$$N^{1-H}\log(N)\frac{2}{\sqrt{c_{2,H}}}\left(\hat{H}_N(2,a) - H\right) \to Z$$

  where Z is the law of a standard Rosenblatt random variable (see 11).

- *if $H = \frac{3}{4}$, then, in distribution as $N \to \infty$,*

$$\sqrt{N\log N}\frac{2}{\sqrt{c_{\prime,1,H}}}(\hat{H}_N(2,a) - H) \to \mathcal{N}(0,1).$$

Proof. This follows from the relation (59) and Theorem 2. ∎

In the Rosenblatt data case, the renormalized error $\hat{H}_N - H$ does not converge to the normal law. But one can obtain from Theorem 13 an adjusted version of this error that converges to the normal distribution.

Theorem 15 *Suppose that $H > \frac{1}{2}$ and the observed process is a Rosenblatt process with self-similarity index H. Then, strong consistency holds for $\hat{H}_N$, i.e. almost surely,*

$$\lim_{N\to\infty} \hat{H}_N(2,a) = H. \tag{61}$$

In addition, we have the convergence in $L^2(\Omega)$,

$$\lim_{N\to\infty} \frac{N^{1-H}}{2d(H)} \log(N)\ (\hat{H}_N(2,a) - H) = Z(1), \tag{62}$$

where $Z(1)$ is the observed process at time 1. Moreover,

- *If $H < \frac{3}{4}$, then, in distribution as $N \to \infty$,*

$$\frac{\sqrt{N}}{\sqrt{e_{1,H}}} \left[-2\log(N)\ (\hat{H}_N(2,a) - H) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1) \right] \to N(0,1)$$

- *and if $H = \frac{3}{4}$, then, in distribution as $N \to \infty$,*

$$\frac{\sqrt{N}}{\sqrt{e_{3,H}\log N}} \left[-2\log(N)\ (\hat{H}_N(2,a) - H) - \frac{\sqrt{c_{3,H}}}{N^{1-H}} Z(1) \right] \to N(0,1)$$

Proof. This follows from Theorem 13, Theorem 11 and relation (59). ∎

6 Appendix

Lemma 16 *When $H \in (0,3/4)$, $\sum_{k=1}^{\infty} \left(2k^{2H} - (k-1)^{2H} - (k+1)^{2H}\right)^2$ is finite*

Proof. Since $2k^{2H} - (k-1)^{2H} - (k+1)^{2H} = k^{2H} f\left(\frac{1}{k}\right)$, with $f(x) := 2 - (1-x)^{2H} - (1+x)^{2H}$ being asymptotically equivalent to $H(2H-1)x^2$ for small x, the general term of the series is equivalent to $(2H)^2 (2H-1)^2 k^{4H-4}$. Therefore the series converges to a constant depending only on H. ∎

Lemma 17 *When $H \in (3/4,1))$, $N^2 \sum_{i,j=1,\cdots,N;|i-j|\geq 2} \left(2\left|\frac{i-j}{N}\right|^{2H} - \left|\frac{i-j-1}{N}\right|^{2H} - \left|\frac{i-j+1}{N}\right|^{2H}\right)^2$ converges to $H^2(2H-1)/(H-3/4)$ as $N \to \infty$.*

Proof. Let us write $x = |i-j|/N$, $\alpha = 2H$, and $h = 1/N$. Thence using a mean value theorem of order 3, we have

$$2x^\alpha - (x-h)^\alpha - (x+h)^\alpha = -h^2 \alpha(\alpha-1) x^{\alpha-2} + ch^3 \xi^{\alpha-3}$$

for some $\xi \in (x-h, x+h)$ and some constant c. Under the restriction $x \geq 2h$, we have $x/2 \leq x-h$, which implies that the above correction term $ch^3 |\xi|^{\alpha-3} \leq c' h^3 x^{\alpha-3}$ for some other constant c'. Now

we can write the series of interest as

$$\sum_{i,j=1,\cdots,N;|i-j|\geq 2}\left(2\left|\frac{i-j}{N}\right|^{2H}-\left|\frac{i-j-1}{N}\right|^{2H}-\left|\frac{i-j+1}{N}\right|^{2H}\right)^2 \tag{63}$$

$$\leq \sum_{i,j=1,\cdots,N;|i-j|\geq 2} N^{-4}\left|2H\left(2H-1\right)\right|^2\left|\frac{i-j}{N}\right|^{4H-4} \tag{64}$$

$$+2\sum_{i,j=1,\cdots,N;|i-j|\geq 2} c''N^{-5}\left|\frac{i-j}{N}\right|^{4H-5} \tag{65}$$

$$+\sum_{i,j=1,\cdots,N;|i-j|\geq 2} c''N^{-6}\left|\frac{i-j}{N}\right|^{4H-6} \tag{66}$$

where c'' is another constant. Replacing the + signs in lines (65) and (66) by − signs, we obtain the opposite inequality in line (64). We will show that the terms in lines (65) and (66) are of a lower order in N than the term in line (64). This will imply that $\sum_{i,j=1,\cdots,N;|i-j|\geq 2}\left|\langle A_i,A_j\rangle_{\mathcal{H}}\right|^2$ is asymptotically equivalent to the right-hand side of line (64).

Using a limit of a Riemann sum, we have

$$\lim_{N\to\infty} N^{-2}\sum_{i,j=1,\cdots,N;|i-j|\geq 2}\left|\left(i-j\right)/N\right|^{4H-4}=\iint_{[0,1]^2}\left|x-y\right|^{4H-4}dxdy=\frac{1}{\left(2H-1\right)\left(4H-3\right)}.$$

Therefore the term on the right-hand side of line (64) is asymptotically equivalent to the expression $N^{-2}H^2\left(2H-1\right)/\left(4H-3\right)$. On the other hand, for lines (65) and (66), the series cannot be compared to Riemann sum. Rather, they converge (indeed, $4H-5<-1$). We have

$$\sum_{i,j=1,\cdots,N;|i-j|\geq 2}\left|i-j\right|^{4H-5}=2N\sum_{k=2}^{N-1}k^{4H-5}\leq 2Nc_H;$$

$$\sum_{i,j=1,\cdots,N;|i-j|\geq 2}\left|i-j\right|^{4H-6}=2N\sum_{k=2}^{N-1}k^{4H-6}\leq 2Nc'_H.$$

Therefore both terms in lines (65) and (66) are smaller than a constant times N^{1-4H}, which in our case is negligible compared to N^{-2}. In conclusion, we have proved that N^2 times the series (63) converges to $\frac{\left|2H(2H-1)\right|^2}{(2H-1)(4H-3)}=\frac{H^2(2H-1)}{H-3/4}$, which concludes the proof. ■

Lemma 18 *With $H\in(0,3/4)$,*

$$\langle A_i,A_j\rangle_{\mathcal{H}}=2^{-1}\left(2\left|\frac{i-j}{N}\right|^{2H}-\left|\frac{i-j-1}{N}\right|^{2H}-\left|\frac{i-j+1}{N}\right|^{2H}\right),$$

as $N\to\infty$, we have

$$N^{8H-2}\sum_{i,i'=1}^{N}\sum_{j=1}^{i}\sum_{j'=1}^{i'}\left\langle A_i;A_{i'}\right\rangle_{\mathcal{H}}\left\langle A_i;A_j\right\rangle_{\mathcal{H}}\left\langle A_{i'};A_{j'}\right\rangle_{\mathcal{H}}\left\langle A_j;A_{j'}\right\rangle_{\mathcal{H}}.=o\left(N^{-4}\right).$$

Proof. We can deal with the diagonal terms first. With $i = i'$ and $j = j'$, and thanks to the asymptotics in (16), the corresponding contribution is of order

$$N^{8H-2} \sum_{i,j=1}^{N} N^{-4H} \left| \langle A_i; A_j \rangle_{\mathcal{H}} \right|^2 \asymp N^{4H-2-1-4H} = N^{-1}.$$

It is trivial to check that the terms with $i = i'$ and $j = j' \pm 1$, as well as the terms with $i = i' \pm 1$ and $j = j' \pm 1$ yield again the order N^{-1}. By changing the roles of the indices, we also treat all terms of the type $|i - i'| \leq 2$ and $|j - i| \leq 2$.

Now for the hyperplane terms with $i = i'$ and $|j - j'| \geq 2$, $|j - i| \geq 2$, $|j' - i| \geq 2$, we can use the relations of the form

$$\langle A_i; A_j \rangle_{\mathcal{H}} \leq 2^{2-2H} H (2H - 1) N^{-2} \left| (i - j) / N \right|^{2H-2},$$

holding also for the pairs (i, j') and (j, j'), to obtain that the corresponding contribution in (23) is of the order

$$\begin{aligned}
&N^{8H-2} \sum_{i=1}^{N} \sum_{|j-j'|\geq 2; |j-i|\geq 2; |j'-i|\geq 2} N^{-2H} N^{-6} \left| (i - j) / N \right|^{2H-2} \left| \left(i - j'\right) / N \right|^{2H-2} \left| \left(j - j'\right) / N \right|^{2H-2} \\
&= N^{6H-5} \sum_{i=1}^{N} \sum_{|j-j'|\geq 2; |j-i|\geq 2; |j'-i|\geq 2} N^{-3} \left| (i - j) / N \right|^{2H-2} \left| \left(i - j'\right) / N \right|^{2H-2} \left| \left(j - j'\right) / N \right|^{2H-2} \\
&\asymp N^{6H-5} \ll N^{-1/2}
\end{aligned}$$

where we used the fact that the last summation above converges as a Riemann sum to the finite integral $\int_{[0,1]^3} \left| (x - y) (x - z) (y - z) \right|^{2H-2} dxdydz$, and then the fact that $H < 3/4$. The hyperplane term with $i = i' \pm 1$ and $|j - j'| \geq 2$, $|j - i| \geq 2$, $|j' - i| \geq 2$, the calculation is identical.

Lastly, and similarly to the case just treated, when all indices are distant by at least 2 units, we can again use the upper bound $N^{-2} \left| (i - j) / N \right|^{2H-2}$ for $\langle A_i; A_j \rangle_{\mathcal{H}}$ and all other three pairs, obtaining a contribution of the form

$$\begin{aligned}
&N^{8H-2} \sum_{|i-i'|\geq 2; |j-j'|\geq 2; |j-i|\geq 2; |j'-i|\geq 2} N^{-8} \left| \frac{i - i'}{N} \right|^{2H-2} \left| \frac{i - j}{N} \right|^{2H-2} \left| \frac{i - j'}{N} \right|^{2H-2} \left| \frac{j - j'}{N} \right|^{2H-2} \\
&\asymp N^{8H-6} \int_{[0,1]^4} \left| \left(x - x'\right) (x - y) \left(x' - z\right) (y - z) \right|^{2H-2} dx'dxdydz;
\end{aligned}$$

since $H < 3/4$, we have $8H - 6 < 0$, and the above goes to 0 as well, albeit slower than the other terms. ∎

Lemma 19 *With $H \in (3/4, 1)$, and*

$$\tilde{A}_i (s) = \mathbf{1}_{[0, \frac{i+1}{N}]} (s) K^H \left(\frac{i+1}{N}, s \right) - \mathbf{1}_{[0, \frac{i}{N}]} (s) K^H \left(\frac{i}{N}, s \right),$$

we have that $N \sum_{i=1}^{N} \tilde{A}_i (r) \tilde{A}_i (s)$ converges in $L^2 \left([0,1]^2 \right)$ to the function

$$(r, s) \mapsto L (r, s) := \int_{r \vee s}^{1} \frac{\partial K^H}{\partial u} (u, s) \frac{\partial K^H}{\partial u} (u, r) du.$$

Proof. $\tilde{A}_i(s)$ can be rewritten as

$$\begin{aligned}\tilde{A}_i(s) &= \mathbf{1}_{[0,\frac{i}{N}]}(s)\left(K^H\left(\frac{i+1}{N},s\right)-K^H\left(\frac{i}{N},s\right)\right)+\mathbf{1}_{[\frac{i}{N},\frac{i+1}{N}]}(s)\,K^H\left(\frac{i+1}{N},s\right)\\ &= N^{-1}\mathbf{1}_{[0,\frac{i}{N}]}(s)\,\frac{\partial K^H}{\partial u}(\xi_i,s)+\mathbf{1}_{[\frac{i}{N},\frac{i+1}{N}]}(s)\,K^H\left(\frac{i+1}{N},s\right)\\ &=: B_i(s)+C_i(s)\end{aligned}$$

where $\xi_i=\xi_i(s)$ depends on s but is nonetheless in the interval $[i/N,(i+1)/N]$. Using this formula in (25) yields square-type terms with $B_i(s)B_i(r)$ and $C_i(s)C_i(r)$, and a cross-product term. This last term is treated like the term involving $C_i(s)C_i(r)$, and we leave it to the reader. Now, using the fact that $K(t,s)\le c(t/s)^{H-1/2}(t-s)^{H-1/2}$ we write

$$\begin{aligned}&\iint_{[0,1]^2}drds\left|2Nc_{2,H}^{-1/2}\sum_{i=1}^N C_i(s)C_i(r)\right|^2\\ &\le 4N^2c_{2,H}^{-1}\iint_{[0,1]^2}drds\sum_{i=1}^N\sum_{j=1}^N\mathbf{1}_{[\frac{i}{N},\frac{i+1}{N}]}(s)\,\mathbf{1}_{[\frac{j}{N},\frac{j+1}{N}]}(r)\left(\frac{i+1}{Ns}\right)^{H-1/2}\left(\frac{j+1}{Nr}\right)^{H-1/2}N^{2-4H}\\ &\le 4N^{2-4H}c_{2,H}^{-1}\iint_{[0,1]^2}dtdu\sum_{i=1}^N\sum_{j=1}^N N^{-2}\left(1+\frac{1}{i}\right)^{H-1/2}\left(1+\frac{1}{j}\right)^{H-1/2}\\ &\le 8N^{2-4H}c_{2,H}^{-1}.\end{aligned}$$

Since $H>1/2$, this proves that the portion of $\tilde{D}^2\bar{F}_N$ corresponding to C_i tends to 0 in $L^2\left([0,1]^2\right)$. For the dominant term, we calculate

$$\begin{aligned}&\left|2Nc_{2,H}^{-1/2}\sum_{i=1}^N B_i(r)B_i(s)-2c_{2,H}^{-1/2}L(r,s)\right|\\ &=2c_{2,H}^{-1/2}\left|\sum_{i=1}^N\mathbf{1}_{[0,\frac{i}{N}]}(r\vee s)\,\frac{\partial K^H}{\partial u}(\xi_i(r),r)\,\frac{\partial K^H}{\partial u}(\xi_i(s),s)-\int_{r\vee s}^1\frac{\partial K^H}{\partial u}(u,s)\,\frac{\partial K^H}{\partial u}(u,r)\,du\right|.\end{aligned}$$

This converges to 0 pointwise as a limit of Riemann sums; since $H>1/2$, the kernel's derivative $\frac{\partial K^H}{\partial u}(u,s)$ is of constant sign and integrable on the unit square, so that by dominated convergence, the above convergence to 0 occurs also in $L^2\left([0,1]^2\right)$. We have proved our claim, and the multiplicative constant is $2c_{2,H}^{-1/2}$. ∎

Lemma 20 *For all $H>1/2$, with $I_i=\left(\frac{i-1}{N},\frac{i}{N}\right]$,*

$$\begin{aligned}&\lim_{N\to\infty}N^{2H}\sum_{i,j=1}^N\int_{I_i}\int_{I_i}\int_{I_j}\int_{I_j}|u-v|^{2H'-2}|u'-v'|^{2H'-2}|u-u'|^{2H'-2}|v-v'|^{2H'-2}du'dv'dvdu\\ &=2a(H)^{-2}\left(\frac{1}{H}-\frac{1}{H+1}\right)\end{aligned}$$

Proof. We make the change of variables

$$\bar{u} = (u - \frac{i-1}{N})N$$

with $d\bar{u} = Ndu$ and we proceed similarly for the other variables u', v, v'. We will obtain

$$\begin{aligned}
&\mathbf{E}(T_2^2) \\
&= N^{4H-2} 32 a(H)^4 d(H)^4 \frac{1}{N^4} \frac{1}{N^{8H'-8}} \sum_{i,j=1}^{N} \int_0^1 \int_0^1 \int_0^1 \int_0^1 dudvdu'dv' \\
&\times |u-v|^{2H'-2} |u'-v'|^{2H'-2} |u-u'+i-j|^{2H'-2} |v-v'+i-j|^{2H'-2}.
\end{aligned}$$

Since $8H' - 8 = 4H - 4$,

$$\begin{aligned}
&\mathbf{E}(T_2^2) \\
&= \frac{1}{N^2} 32 a(H)^4 d(H)^4 \sum_{i,j=1}^{N} \int_0^1 \int_0^1 \int_0^1 \int_0^1 dudvdu'dv' \\
&\times |u-v|^{2H'-2} |u'-v'|^{2H'-2} |u-u'+i-j|^{2H'-2} |v-v'+i-j|^{2H'-2}.
\end{aligned}$$

This sum can be divided into two parts: a diagonal part containing the terms $i = j$ and a non-diagonal part containing the terms $i \neq j$. As in the calculations contained in the previous sections, one can see that the non-diagonal part is dominant. Therefore the behavior of $\mathbf{E}(T_2^2)$ will be given by

$$\begin{aligned}
T_2' &= \frac{1}{N^2} 32 a(H)^4 d(H)^4 2 \sum_{i>j} \int_0^1 \int_0^1 \int_0^1 \int_0^1 dudvdu'dv' \\
&\times |u-v|^{2H'-2} |u'-v'|^{2H'-2} |u-u'+i-j|^{2H'-2} |v-v'+i-j|^{2H'-2} \\
&= \frac{1}{N^2} 32 a(H)^4 d(H)^4 2 \sum_{i=1}^{N} \sum_{k=1}^{N-i} \int_0^1 \int_0^1 \int_0^1 \int_0^1 dudvdu'dv' \\
&\times |u-v|^{2H'-2} |u'-v'|^{2H'-2} |u-u'+k|^{2H'-2} |v-v'+k|^{2H'-2} \\
&= \frac{1}{N^2} 32 a(H)^4 d(H)^4 2 \sum_{k=1}^{N} (N-k) \int_0^1 \int_0^1 \int_0^1 \int_0^1 dudvdu'dv' \\
&\times |u-v|^{2H'-2} |u'-v'|^{2H'-2} |u-u'+k|^{2H'-2} |v-v'+k|^{2H'-2}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{1}{N^2} \sum_{k=1}^{N} (N-k) |u-u'+k|^{2H'-2} |v-v'+k|^{2H'-2} \\
&= N^{2H'-2} \frac{1}{N} \sum_{k=1}^{N} (1 - \frac{k}{N}) |\frac{u-u'}{N} + \frac{k}{N}|^{2H'-2} |\frac{v-v'}{N} + \frac{k}{N}|^{2H'-2}.
\end{aligned}$$

Because the terms of the form $(u-u')/N$ are negligible in front of k/N for all but the smallest k's, the above expression is asymptotically equivalent to the Riemann sum approximation of the Riemann integral

$$\int_0^1 (1-x)\, x^{4H'-4} = 1/(2H-1) - 1/(2H)$$

where we used $2H'-2 = H-1$. ■

Lemma 21 . *For $H \in (1/2, 3/4)$, we have $\mathbf{E}\left[J_6^2\right] \sim cst \cdot N^{8H-6}$.*

Proof. With the notation as in Step 1 of the proof of Theorem 7, making the change of variables $\bar{u} = (u - \frac{i-1}{N})N$ and similarly for the other integrands, we obtain

$$\begin{aligned}
&\mathbf{E}\left[J_6^2\right] \\
&= cst \cdot N^{8H-2} \frac{1}{N^{8H'-8}} \frac{1}{N^4} \frac{1}{N^{4H}} \sum_{i,j,k,l} \int_{[0,1]^4} dudvdu'dv' \\
&\times |u-v+i-j|^{2H'-2} |u-u'+i-k|^{2H'-2} |u'-v+j-k|^{2H'-2} |v-v'+k-l|^{2H'-2} \\
&\times \left(2|i-k|^{2H} - |i-k+1|^{2H} - |i-k-1|^{2H}\right) \left(2|j-l|^{2H} - |j-l+1|^{2H} - |j-l-1|^{2H}\right) \\
&= cst \cdot \frac{1}{N^2} \sum_{i,j,k,l} \int_{[0,1]^4} dudvdu'dv' \\
&\times |u-v+i-j|^{2H'-2} |u-u'+i-k|^{2H'-2} |u'-v+j-k|^{2H'-2} |v-v'+k-l|^{2H'-2} \\
&\times \left(2|i-k|^{2H} - |i-k+1|^{2H} - |i-k-1|^{2H}\right) \left(2|j-l|^{2H} - |j-l+1|^{2H} - |j-l-1|^{2H}\right)
\end{aligned}$$

Again we use the fact that the dominant part in the above expression is the one when all indices are distant by at least two units. In this case, up to a constant, we have the upper bound $|i-k|^{2H-2}$ for the quantity $\left(2|i-k|^{2H} - |i-k+1|^{2H} - |i-k-1|^{2H}\right)$. By using Riemann sums, we can write

$$\mathbf{E}\left[J_6^2\right] = cst \cdot \frac{1}{N^2} \left(\frac{1}{N^4} \sum_{i,j,k,l} f(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}, \frac{l}{N}) \right) N^{8H'-8} N^{4H-4}$$

where f is a Riemann-integrable function on $[0,1]^4$ and the Riemann sum converges to the finite integral of f therein. The lemma follows. ■

References

[1] J. Beran (1994): *Statistics for Long-Memory Processes.* Chapman and Hall.

[2] P. Breuer and P. Major (1983): Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Analysis*, **13** (3), 425-441.

[3] J.F. Coeurjolly (2001): Estimating the parameters of a fractional Brownian motion by discrte variations of its sample paths. *Statistical Inference for Stochastic Processes*, **4**, 199-227.

[4] R.L. Dobrushin and P. Major (1979): Non-central limit theorems for non-linear functionals of Gaussian fields. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **50**, 27-52.

[5] P. Embrechts and M. Maejima (2002): *Selfsimilar processes.* Princeton University Press, Princeton, New York.

[6] X. Guyon and J. León (1989): Convergence en loi des H-variations d'un processus gaussien stationnaire sur $\mathbf{R}$. *Annales IHP*, **25**, 265-282.

[7] S.B. Hariz (2002): Limit theorems for the nonlinear functionals of stationary Gaussian processes. *J. of Multivariate Analysis*, **80**, 191-216.

[8] Y.Z. Hu and D. Nualart (2005): Renormalized self-intersection local time for fractional Brownian motion. *The Annals of Probability*, **33** (3), 948-983.

[9] G. Lang and J. Istas (1997): Quadratic variations and estimators of the Hölder index of a Gaussian process. *Annales IHP*, **33**, 407-436.

[10] B. Mandelbrot (1963): The variation of certain speculative prices. *Journal of Business* **XXXVI**, 392-417.

[11] A.I. McLeod and K.W. Kipel (1978): Preservation of the rescaled adjusted range: a reaassement of the Hurst exponent. *Water Resourc. Res.*, **14** (3), 491-508.

[12] J.R. León and C. Ludena (2007): Limits for weighted p-variations and likewise functionals of fractional diffusions with drift. *Stochastic Proc. Appl.*, **117**, 271-296.

[13] I. Nourdin (2007): Asymptotic behavior of certain weighted quadratic variations and cubic varitions of fractional Brownian motion. Preprint, accepted for publication in *The Annals of Probability.*

[14] I. Nourdin and D. Nualart (2007): *Convergence in law for certain weighted quadratic variations of fractional Brownian motion.* Preprint.

[15] D. Nualart (2006): *Malliavin calculus and related topics, 2nd ed.* Springer.

[16] D. Nualart and S. Ortiz-Latorre (2006): Central limit theorems for multiple stochastic integrals and Malliavin calculus. Preprint, to appear in *Stochastic Processes and their Applications.*

[17] D. Nualart and G. Peccati (2005): Central limit theorems for sequences of multiple stochastic integrals. *The Annals of Probability*, **33** (1), 173-193.

[18] G. Peccati and C.A. Tudor (2004): Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités*, **XXXIV**, 247-262.

[19] G. Samorodnitsky and M. Taqqu (1994): *Stable Non-Gaussian random variables.* Chapman and Hall, London.

[20] J. Swanson (2007): Variations of the solution to a stochastic heat equation. *Annals of Probability*, to appear.

[21] M. Taqqu (1975): Weak convergence to the fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **31**, 287-302.

[22] M. Taqqu (1979): Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **50**, 53-83.

[23] C.A. Tudor (2007): Analysis of the Rosenblatt process. *ESAIM-PS*, to appear.

[24] A.S. Ustunel (1995): *An Introduction to Analysis on Wiener Space,* Springer-Verlag.

[25] W. Willinger, M. Taqqu and V. Teverovsky (1999): Long range dependence and stock returns. *Finance and Stochastics*, **3**, 1-13.

[26] W. Willinger, M. Taqqu, W.E. Leland and D.V. Wilson (1995): Selfsimilarity in high speed packet traffic: analysis and modelisation of ethernet traffic measurements. *Statist. Sci.* **10**, 67-85.