Variations of the fractional Brownian motion via 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 24, 2008

Abstract

Using recent criteria for the convergence of sequences of multiple stochastic integrals based on the Malliavin calculus, we analyze the asymptotic behavior of quadratic variations for the fractional Brownian motion (fBm) and we apply our results to the design of a strongly consistent statistical estimators for the fBm's self-similarity parameter H.

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

Key words: multiple stochastic integral, fractional Brownian motion, Malliavin calculus, quadratic variation, Hurst parameter, selfsimilarity, statistical estimation.

1 Introduction

1.1 Context and motivation

The motivation for this work is to examine the variations of fractional Brownian motion (fBm) using tools from stochastic analysis. The stochastic analysis of fBm has been intensively developped in recent years and its applications are many. The Hurst parameter H characterizes all the important properties of the fBm and 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 $a = (a_0, a_1, \dots, a_l)$ 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

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

where for $i \in \{l, \cdots, N\}$,

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

When X is fBm, these statistics are used to derive strongly consistent estimators for the Hurst parameter, and their associated normal convergence results. A detailed study can be found in [5], [7] 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 for fBm 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 selfsimilar 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), for fBm itself, but the method can be generalized. We will apply the Malliavin calculus, Wiener-Itô chaos expansions, and recent results on the convergence of multiple stochastic integrals proved in [10], [9] or [11]. 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 [9] 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 [9] 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.

This article is structured as follows. Section 2 presents preliminaries on fractional stochastic analysis. Section 3 presents strategies and proof outlines; some calculations, recorded as lemmas, are proved in the Appendix. Section 4 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 Brownian motion (Wiener process) W and in this case $\mathcal{H} = L^2([0,1])$. Otherwise $\mathcal{H}$ is the Hilbert space on $[0,1]$ extending (by linearity and closure under the inner product $\langle \mathbf{1}_{[0,s]}; \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}}$) the rule

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

As there should be no risk of confusion, we will drop the superscript and always denote $B = B^H$. Denote by I_n the multiple stochastic integral with respect to B. This I_n is actually an isometry between the Hilbert space $\mathcal{H}^{\odot n}$(symmetric tensor product) equipped with the scaled 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 $H_n(B(\varphi))$ where $\varphi \in \mathcal{H}$ with $\|\varphi\|_{\mathcal{H}} = 1$ and H_n is the Hermite polynomial of degree n. We will sometimes use the representation of B with respect to a standard Brownian motion W: there exists a Wiener process W and a deterministic kernel $K^H(t,s)$ for $0 \le s \le t$ such that $B(t) = I_1^W\left(K^H(t,\cdot)\right)$ where I_1^W is the Wiener integral with respect to W (see [8]).

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

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

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

We now introduce the Malliavin derivative for random variables in a finite chaos. If $f \in \mathcal{H}^{\odot n}$ 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 distribution of a sequence of multiple integrals to the standard normal law. We will use the following result (see Theorem 4 in [9], see also [10]).

Theorem 1 *Let $(F_k, k \geq 1)$, $F_k = I_n(f_k)$ (with $f_k \in \mathcal{H}^{\odot n}$ for every $k \geq 1$) 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 in distribution to the normal law $\mathcal{N}(0,1)$.

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 that $\lim_{k\to\infty} \|f_k \otimes_l f_k\|_{\mathcal{H}^{\otimes 2(n-l)}} = 0$.

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

Criterion (iv) is due to [9]; we will refer to it as the Nualart–Ortiz-Latorre criterion. A multidimensional version of the above theorem has been proved in [11]; it will be used in Section 3.4 to study the vectorial convergence of our variations.

3 Variations of fractional Brownian motion

3.1 Strategy: try the Nualart–Ortiz-Latorre characterization

With $a = \{1, -1\}$ and $k = 2$, with B = fractional Brownian motion (fBm), we will use the tools in the previous section to reproduce the results found in [13], [6] , and summarized in [3], seeking to show a central limit theorem for the standardized quadratic variation. When a central limit theorem does not hold, our tools will also provide the correct limiting distribution.

With the notation $A_i = \mathbf{1}_{((i-1)/N, i/N]}$ $(i = 1, \ldots, N)$ and $B(A_i) = B(i/N) - B((i-1)/N)$ so that $Var[B(A_i)] = N^{-2H}$, this variation can be expressed as

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

To show it converges to the standard normal law $\mathcal{N}(0,1)$ after an appropriate scaling, instead of employing a Gaussian method, we propose a Wiener chaos approach which can generalize to higher order cases than fBm. In fact, our tools allow some new results even in the Gaussian case, e.g. when $H = 3/4$. The strategy is to attempt to use the characterization by Nualart and Ortiz-Latorre [9, Theorem 4] as presented in the previous section. In the next two subsections, we prove the following.

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

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

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

then

$$F_N := \sqrt{N/c_{1,H}}\, V_N \tag{3}$$

converges in distribution to a the standard normal law.

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

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

then

$$\bar{F}_N := \sqrt{N^{4-4H}/c_{2,H}} V_N \tag{5}$$

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

$$\frac{(4H-3)^{1/2}}{4H(2H-1)^{1/2}} \iint_{[0,1]^2} \left(\int_{r \vee s}^1 \frac{\partial K^H}{\partial u}(u,s) \frac{\partial K^H}{\partial u}(u,r)\, du \right) dW(r)\, dW(s) \tag{6}$$

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

- *If $H = 3/4$, let*

$$c'_{1,H} := (2H(2H-1))^2 = 9/16 \tag{7}$$

and define

$$\tilde{F}_N := \sqrt{\frac{N}{c'_{1,H} \log N}} V_N. \tag{8}$$

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

The central limit theorem part can be actually stated for $H \in (0, \frac{3}{4}]$ but we prefer to stay in the context of long-memory. 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

The product formula of multiple integrals (see [8]) in our present case, yields $V_N = N^{2H-1} I_2 \left(\sum_{i=1}^N A_i \otimes_1 A_i \right)$ and

$$\mathbf{E}\left[|V_N|^2 \right] = 2N^{4H-2} \sum_{i=1}^N \sum_{j=1}^N \left| \langle A_i, A_j \rangle_{\mathcal{H}} \right|^2, \tag{9}$$

where, as we said, $A_i = \mathbf{1}_{((i-1)/N, i/N]}$ for $i = 1, \ldots, N$. To calculate this quantity, we notice that $\langle A_i, A_j \rangle_{\mathcal{H}} = \mathbf{E}[B(A_i) B(A_j)]$ can be calculated explicitly via R_H, as $2^{-1}(2\left|\frac{i-j}{N}\right|^{2H} - \left|\frac{i-j-1}{N}\right|^{2H} - \left|\frac{i-j+1}{N}\right|^{2H})$.This expression is close to $H(2H-1) N^{-2} |(i-j)/N|^{2H-2}$, 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

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

Lemma 7 in [15] implies the above sequence converges, and thus $\lim_{N\to\infty} \mathbf{E}\left[\left| \sqrt{N} V_N \right|^2 \right] = c_{1,H}$.

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; in the sum defining this quantity, the tridiagonal term corresponding to $|i-j| \leq 1$ can be ignored as we did in the previous step when $i = j$. By Lemma 8 in [15] we have that $N^2 \sum_{i,j=1,\cdots,N; |i-j| \geq 1} 2 \left| \langle A_i, A_j \rangle_{\mathcal{H}} \right|^2$ compares to a Riemann sum in such a way that it converges to $H^2 (2H-1)/(4H-3)$. This non-tridiagonal term, which is of order N^{-2}, dominates the tridiagonal term which is of order $N^{1-4H} \ll N^{-2}$ (since $H > 3/4$). In conclusion when $H > 3/4$, according to (9),

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

Case 3: $H = 3/4$. In this case, 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 $\mathbf{E}\left[(V_N)^2\right] \simeq c'_{1,H}\left(\log N\right)/N$. Thus, $\lim_{N\to\infty} \mathbf{E}\left[\left|\tilde{F}_N\right|^2\right] = 1$ where $c'_{1,H} = 9/16$ and $\tilde{F}_N := \left(\frac{N}{c'_{1,H}\log N}\right)^{\frac{1}{2}} V_N$.

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 (3), (5) or (8) depending on whether H is bigger or smaller than 3/4 or equal to 3/4. We will see that this only works for $H \leq 3/4$. Using the rule $D_r I_2(f) = 2I_1(f(\cdot, r))$ when f is symmetric, we have $D_r V_N = 2N^{2H-1}\sum_{i=1}^{N} A_i(r)\, I_1(A_i)$. Hence

$$\|DV_N\|^2_{\mathcal{H}} = 4N^{4H-2}\sum_{i,j=1}^{N} I_1(A_i)\, I_1(A_j)\,\langle A_i; A_j\rangle_{\mathcal{H}} \tag{10}$$

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 (9) that $\mathbf{E}\left[\|DV_N\|^2_{\mathcal{H}}\right] = 2\mathbf{E}\left[V_N^2\right]$, and fom the results of the previous section, $\lim_{N\to\infty}\mathbf{E}\left[\|DF_N\|^2_{\mathcal{H}}\right] = 2$ holds in all cases.

Thus it is now sufficient to show that $\|DF_N\|^2_{\mathcal{H}} - \mathbf{E}\left[\|DF_N\|^2_{\mathcal{H}}\right]$ converges to 0 in $L^2(\Omega)$.

A simple use of the product formula for multiple integrals gives

$$\|DV_N\|^2_{\mathcal{H}} - \mathbf{E}\left[\|DV_N\|^2_{\mathcal{H}}\right] = 4N^{4H-2}\sum_{i,j=1}^{N}\langle A_i; A_j\rangle_{\mathcal{H}} I_2(A_i \otimes A_j)$$

and thus

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

Case 1: $1/2 < H < 3/4$. In this case, by Lemma 7 in the appendix, the conclusion is that

$$E\left[\|DF_N\|^2_{\mathcal{H}} - \mathbf{E}\left[\|DF_N\|^2_{\mathcal{H}}\right]\right]^2$$

is asymptotically equivalent to a constant multiple of N^{8H-6}. This, together with the calculations in Section 3.2, is the Nualart–Ortiz-Latorre necessary and sufficient condition for F_N to converges in law to a standard normal.

Case 2: $H = 3/4$. This case is treated similarly to the previous one. We record the following for later use. As $N \to \infty$

$$\mathbf{E}\left[\|D\tilde{F}_N\|^2_{\mathcal{H}}\right] = \frac{N}{\log N}\left(c'_{1,H}\right)^{-1} 4N^{4H-2}\sum_{i,j=1}^{N}\langle A_i, A_j\rangle^2_{\mathcal{H}} \to 2. \tag{11}$$

Case 3: $H > 3/4$. In this case, using the scaling $\bar{F}_N = N^{2-2H} V_N / \sqrt{c_{2,H}}$ one checks that $\left[D\bar{F}_N \right\|_{\mathcal{H}}^2 - \mathbf{E}\left[D\bar{F}_N \right\|_{\mathcal{H}}^2$ does not converge to 0. Therefore, the Nualart–Ortiz-Latorre characterization cannot be applied, and if there is a limit, it cannot be Gaussian. In fact we will prove directly that the limit is not Gaussian. In order to find the limit of $\bar{F}_N$, let us return to the definition of this quantity in terms of the Wiener process W such that $B(t) = \int_0^t K^H(t,s)\, dW(s)$. We then note that $\bar{F}_N$ can be written as $\bar{F}_N = \tilde{I}_2(f_N)$ where $\tilde{I}_2$ is the double Wiener integral operator with respect to W, and $f_N = N c_{2,H}^{-1/2} \sum_{i=1}^N \tilde{A}_i \otimes \tilde{A}_i$ where

$$\tilde{A}_i(s) = \mathbf{1}_{[0,\frac{i+1}{N}]}(s) K^H\left(\frac{i+1}{N}, s\right) - \mathbf{1}_{[0,\frac{i}{N}]}(s) K^H\left(\frac{i}{N}, s\right). \tag{12}$$

Lemma 8 in the Appendix shows that f_N converges in $L^2\left([0,1]^2\right)$ to $c_{2,H}^{-1/2} L_1^{2H-1}$ where L_1^{2H-1} is the function

$$(r,s) \mapsto L_1^{2H-1}(r,s) := \int_{r\vee s}^1 \frac{\partial K^H}{\partial u}(u,s) \frac{\partial K^H}{\partial u}(u,r)\, du. \tag{13}$$

Now define the random variable $Y := d(H_0)\, \tilde{I}_2\left(L_1^{2H-1}\right)$ where

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

This Y is a standard Rosenblatt random variable with parameter $H_0 = 2H-1$, as can be seen for instance in [14]. By the isometry property for stochastic integrals,

$$\mathbf{E}\left[\left|\bar{F}_N - Y\right|^2\right] = \left\| f_N - c_{2,H}^{-1/2} L_1^{2H-1} \right\|_{L^2([0,1]^2)}^2,$$

which, by the convergence of Lemma 8, proves that $\bar{F}_N$ converges to the Rosenblatt random variable $Y = c_{2,H}^{-1/2} \tilde{I}_2\left(L_1^{2H-1}\right)$ in $L^2(\Omega)$.

This finishes the proof of Theorem 2. ■

Remark 3 *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 convergence in probability 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 $H \le \frac{3}{4}$ in which the limit of the components are Gaussian random variables. We make this choice in order to benefit from some recent results in [11] 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 [11]).

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

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

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

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

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

where we denoted by

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

We have the following vectorial limit theorem.

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

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

with $c'_{p,q,H} = \frac{(|p-q|^{2H}-p^{2H}-q^{2H})^2}{2(pq)^{2H}}$ *and*

$$\bar{F}_N(a^p) := \sqrt{N}c_{p,H}^{-1}V_N(2,a_p). \tag{14}$$

Then the vector $(\bar{F}_N(a^1),\ldots,F_N(a^M))$ 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 $(F_N(a^1),\ldots,F_N(a^M))$ 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

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

The next step is to compute the scalar product

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

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

$$\sum_{i=p}^{N}\sum_{j=q}^{N}\left[\left|\frac{i-j}{N}\right|^{2H}+\left|\frac{i-j-p+q}{N}\right|^{2H}-\left|\frac{i-j-p}{N}\right|^{2H}-\left|\frac{i-j+q}{N}\right|^{2H}\right]^2$$
$$=\frac{1}{N^{4H}}\sum_{j=q}^{p-1}\sum_{i=p}^{N}\left(|i-j|^{2H}+|i-j-p+q|^{2H}-|i-j-p|^{2H}-|i-j+q|^{2H}\right)^2$$
$$+\frac{1}{N^{4H}}\sum_{j=p}^{N}\sum_{i=p}^{N}\left(|i-j|^{2H}+|i-j-p+q|^{2H}-|i-j-p|^{2H}-|i-j+q|^{2H}\right)^2+c'_{p,q,H}$$
$$=\frac{2}{N^{4H}}\sum_{j=p}^{N}\sum_{k=1}^{N-j}\left(|k|^{2H}+|k-p+q|^{2H}-|k-p|^{2H}-|k+q|^{2H}\right)^2+c'_{p,q,H}$$
$$=\frac{2}{N^{4H}}\sum_{k=1}^{N-p}(N-k-p)\left(|k|^{2H}+|k-p+q|^{2H}-|k-p|^{2H}-|k+q|^{2H}\right)^2$$
$$=\frac{2}{N^{4H}}\sum_{k=1}^{N-p}(N-k-p)k^{4H}g(\frac{1}{k})^2+c'_{p,q,H}$$

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 [11]. ■

Remark 5 *Is it also possible to give an analogue of Theorem 4 in the case $H\in\left(\frac{3}{4},1\right)$. In this situation we don't need to use Proposition 2 in [11] because we have L^2 convergence. Define, for $p=1,\dots,M$ the sequence $\bar{F}_N(a^p)=b_{p,H}N^{2-2H}V_N(2,a^p)$ where $b_{p,H}$ is a suitable normalizing constant such that $E\left(\bar{F}_N(a^p)\right)^2$ converges to 1 as $N\to\infty$. Then the vector $(\bar{F}_N(a^1),\dots,F_N(a^M))$ converges as $N\to\infty$, in $L^2(\Omega)$ to the vector $(Z_1^{2H-1}(1),\dots,Z_M^{2H-1}(1))$ with $Z_p^{2H-1}(1)$ $(p=1,\dots,M)$ Rosenblatt random variables with selfsimilarity index $2H-1$ as defined in Theorem 2. In this case the L^2 convergence of each component will imply the L^2 convergence of the vector.*

4 The estimators for the selfsimilarity parameter

In this part we construct estimators for the selfsimilarity exponent of a Hermite process based on the discrete observations of the driving process at times $0,\frac{1}{N},\dots,1$. 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

$$S_N(2,a) = \frac{1}{N} \sum_{i=1}^{N} \left(X(\frac{i}{N}) - X(\frac{i-1}{N}) \right)^2 ; \tag{15}$$

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

$$\hat{H}_N(2,a) = -\frac{\log S_N(2,a)}{2 \log N}. \tag{16}$$

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 2-variation, and thus

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

Moreover, by Remark 3, $V_N(2,a)$ converges almost surely to 0, and thus $\log(1 + V_N(2,a)) = V_N(2,a)(1 + o(1))$ where $o(1)$ converges to 0 almost surely as $N \to \infty$. Hence we obtain

$$V_N(2,a) = 2(H - \hat{H}_N(2,a)) (\log N) (1 + o(1)). \tag{17}$$

Relation (17) 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 6 *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,*

$$\lim_{N\to\infty} \hat{H}_N(2,a) = H \tag{18}$$

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 6).

- *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 (17) and Theorem 2. ■

5 Appendix

Lemma 7 *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

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

Proof. As a general rule that we will exemplify below, we have the following: if $i = i'$ or $i = i' \pm 1$ the term $\langle A_i; A_{i'} \rangle_{\mathcal{H}}$ will give a contribution of order $\frac{1}{N^{2H}}$ while if $|i - i'| \geq 2$ the same term will have a contribution less that $cst. \frac{|i-i'|^{2H-2}}{N^{2H-2}} N^{-2}$. Using this rule, although several cases appear, the main term will be obtained when all indices all distant by at least two units.

We can deal with the diagonal terms first. With $i = i'$ and $j = j'$, the corresponding contribution is of order

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

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

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

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

$$\begin{aligned}
&\sum_{i=1}^{N} \sum_{|j-j'| \geq 2; |j-i| \geq 2; |j'-i| \geq 2} N^{-2H} N^{-6} |(i-j)/N|^{2H-2} |(i-j')/N|^{2H-2} |(j-j')/N|^{2H-2} \\
&= N^{-3-2H} \sum_{i=1}^{N} \sum_{|j-j'| \geq 2; |j-i| \geq 2; |j'-i| \geq 2} N^{-3} |(i-j)/N|^{2H-2} |(i-j')/N|^{2H-2} |(j-j')/N|^{2H-2} \\
&\asymp N^{-3-2H} = \mathcal{O}(N^{-4})
\end{aligned}$$

where we used the fact that the last summation above converges as a Riemann sum to the finite integral $\int_{[0,1]^3} |(x-y)(x-z)(y-z)|^{2H-2} dxdydz$, and then the fact that $H < 3/4$. On the hyperplanes term of the form $i = i' \pm 1$ and $|j - j'| \geq 2$, $|j - i| \geq 2$, $|j' - i| \geq 2$, or $|i - i'| \geq 2$, $|i - j| \geq 2$, and $|j - j'| \geq 2$, 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 $N^{-2} |(i-j)/N|^{2H-2}$ for $\langle A_i; A_j \rangle_{\mathcal{H}}$ and all other three pairs, obtaining a contribution of the form

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

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

Lemma 8 *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 $L_N(r,s) := N\sum_{i=1}^N \tilde{A}_i(r)\,\tilde{A}_i(s)$ converges in $L^2\left([0,1]^2\right)$ to the function

$$(r,s) \mapsto L_1^{2H-1}(r,s) := \int_{r\vee s}^1 \frac{\partial K^H}{\partial u}(u,s)\,\frac{\partial K^H}{\partial u}(u,r)\,du.$$

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 $[i/N, (i+1)/N]$. The product $\tilde{A}_i(r)\,\tilde{A}_i(s)$ 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 $K(t,s) \leq c\,(t/s)^{H-1/2}(t-s)^{H-1/2}$ we write

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

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. At this point we can conclude that the sequence $L_N(y_1,y_2)$ converges (in probability for instance) to $L_1^{2H-1}(y_1,y_2)$ for every $y_1, y_2 \in [0,1]$. Our desired result will follow if we proved that the sequence $(L_N)_{N\geq 1}$ is a Cauchy sequence in $L^2([0,1]^2)$. It holds that

$$\begin{aligned}
&\|L_N - L_M\|^2_{L^2([0,1]^2)}\\
&= N^2 \sum_{i,j=0}^{N-1} \left[E\left(B^H(\frac{i+1}{N}) - B^H(\frac{i}{N}) \right) \left(B^H(\frac{j+1}{N}) - B^H(\frac{j}{N}) \right) \right]^2\\
&+ M^2 \sum_{i,j=0}^{M-1} \left[E\left(B^H(\frac{i+1}{M}) - B^H(\frac{i}{M}) \right) \left(B^H(\frac{j+1}{M}) - B^H(\frac{j}{M}) \right) \right]^2\\
&- 2MN \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left[E\left(B^H(\frac{i+1}{N}) - B^H(\frac{i}{N}) \right) \left(B^H(\frac{j+1}{M}) - B^H(\frac{j}{M}) \right) \right]^2
\end{aligned}$$

and we already seen that $N^2 \sum_{i,j=0}^{N-1} \left[E\left(B^H(\frac{i+1}{N}) - B^H(\frac{i}{N}) \right) \left(B^H(\frac{j+1}{N}) - B^H(\frac{j}{N}) \right) \right]^2$ converges to the constant $H^2(2H-1)/(H-3/4)$.

We regard now the sum

$$\begin{aligned}
&MN \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left[E\left(B^H(\frac{i+1}{N}) - B^H(\frac{i}{N}) \right) \left(B^H(\frac{j+1}{M}) - B^H(\frac{j}{M}) \right) \right]^2\\
&= MN \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left[\left| \frac{i+1}{N} - \frac{j+1}{M} \right|^{2H} + \left| \frac{i}{N} - \frac{j}{M} \right|^{2H} - \left| \frac{i+1}{N} - \frac{j}{M} \right|^{2H} - \left| \frac{i}{N} - \frac{j+1}{M} \right|^{2H} \right]^2.
\end{aligned}$$

For any two -variables function g such that $\frac{\partial g}{\partial x \partial y}(x,y)$ exists and belongs to $L^2([0,1]^2)$ it can be easily shown (by a Riemann sum argument) that

$$MN \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left[g(\frac{i+1}{N}, \frac{j+1}{M}) + g(\frac{i}{N}, \frac{j}{M}) - g(\frac{i+1}{N}, \frac{j}{M}) - g(\frac{i}{N}, \frac{j+1}{M}) \right]^2$$

can be written as

$$\frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left(\frac{\partial g}{\partial x \partial y}(a_i, b_j) \right)^2$$

with a_i located between $\frac{i}{N}$ and $\frac{i+1}{N}$ and b_j located between $\frac{j}{M}$ and $\frac{j+1}{M}$ and consequently it converges to $(2H(2H-1))^2 \int_0^1 \int_0^1 |x-y|^{4H-4} dx dy = H^2(2H-1)/(H-3/4)$. ■

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

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

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

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

[9] 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.*

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

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

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

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

[14] C.A. Tudor (2008): Analysis of the Rosenblatt process. *ESAIM-PS*, **12**, 230-257.

[15] C.A. Tudor and F. Viens (2008): Variations and estimators for the selfsimilarity index via Malliavin calculus. Preprint, under revision for *The Annals of Probability.*