

Identification of multifractional Brownian motion

JEAN-FRANÇOIS COEURJOLLY

Labsad, BSHM, UPMF, Avenue Centrale, B.P. 47, 38040 Grenoble Cedex 9, France.

E-mail: jean-francois.coeurjolly@upmf-grenoble.fr.

We develop a method for estimating the Hurst function of a multifractional Brownian motion, which is an extension of the fractional Brownian motion in the sense that the path regularity can now vary with time. This method is based on a local estimation of the second-order moment of a unique discretized filtered path. The effectiveness of our procedure is investigated in a short simulation study.

Keywords: filtering; fractional Brownian motion; functional estimation; multifractional Brownian motion

1. Introduction

Since the pioneering work of Mandelbrot and Van Ness (1968), self-similar processes and, in particular, fractional Brownian motion have been widely used to model data that exhibit long-range dependence and scaling phenomena. However, in certain situations occurring either in the field of turbulence (Frisch 1999) or in biomechanics (Collins and De Luca 1994), a more flexible model is necessary in order to control the dependence structure locally and to allow the path regularity to vary with time.

With such a perspective, a stochastic model leading to an important extension of fractional Brownian motion has recently been developed. This model, called the multifractional Brownian motion, was obtained in two different ways. Both involve replacing the Hurst parameter H by a function of time within the two main stochastic integral representations of fractional Brownian motion. The first representation is a mean average approach and was proposed by Peltier and Lévy Véhel (1995). This leads to the process denoted by $(W_1(t))_{t \geq 0}$. The second one is a spectral approach introduced by Benassi *et al.* (1998). This process is denoted by $(W_2(t))_{t \geq 0}$. These processes are defined as follows:

$$W_1(t) = C\{\pi K(2H(t))\}^{1/2} \int_{\mathbb{R}} f_i(s) dB_1(s), \quad (1)$$

with

$$f_t(s) = \frac{1}{\Gamma(H(t) + 1/2)} \left\{ |t - s|^{H(t)-1/2} \mathbb{1}_{]1-\infty, t]}(s) - |s|^{H(t)-1/2} \mathbb{1}_{]1-\infty, 0]}(s) \right\},$$

$$W_2(t) = C \{K(2H(t))/2\}^{1/2} \int_{\mathbb{R}} \frac{(\exp(it\lambda) - 1)}{|\lambda|^{H(t)+1/2}} dB_2(\lambda), \tag{2}$$

where C is a positive constant, B_1 and B_2 are two Brownian motions, and K is the function defined on $]0, 2[$ by $K(\alpha) = \Gamma(\alpha + 1)\sin(\alpha\pi/2)/\pi$. The processes W_1 and W_2 are well defined (i.e. square-integrable) if the function $H(\cdot)$ is Hölderian of order $0 < \eta \leq 1$ on $[0, 1]$ (the set of such functions is denoted by $\mathcal{C}^\eta([0, 1])$). Under these conditions, Cohen (1999) proved the equality in distribution of both processes normalized in such a way that

$$\mathbb{E}(W_1(t)^2) = \mathbb{E}(W_2(t)^2) = C^2 |t|^{2H(t)}.$$

From now on, a multifractional Brownian motion with Hurst function $H(\cdot)$ and scaling parameter C , defined by (1) or (2), is denoted by $(W(t))_{t \geq 0}$. As a Gaussian model, this process can be defined as the only centred Gaussian process, zero at origin and with covariance function defined for $H \in \mathcal{C}^\eta([0, 1])$ and $s, t \in [0, 1]$ by

$$\mathbb{E}(W(t)W(s)) = \frac{C^2}{2} g(H(t), H(s)) \{ |t|^{H(t)+H(s)} + |s|^{H(t)+H(s)} - |t - s|^{H(t)+H(s)} \}, \tag{3}$$

where g is given by

$$g(H(t), H(s)) = K(H(t) + H(s))^{-1} \{K(2H(t))K(2H(s))\}^{1/2}. \tag{4}$$

The covariance function can easily be obtained using the representation theorem for $|u|^\alpha$ (see, for example, von Bahr and Esseen 1965):

$$|u|^\alpha = K(\alpha) \int_{\mathbb{R}} \frac{1 - \cos(\lambda u)}{|\lambda|^{\alpha+1}} d\lambda, \forall u \in \mathbb{R}, 0 < \alpha < 2.$$

Multifractional Brownian motion leads to a more flexible model since it satisfies our fractional Brownian motion extension conditions.

The main objective of this paper is to develop and study an estimation procedure for multifractional Brownian motion. This problem was partially examined by Benassi *et al.* (1998), where an estimator of a continuously differentiable Hurst function is derived and its consistency is proved. We seek to extend and complete their work by considering Hölderian Hurst functions (of arbitrary order $\eta > 0$) and by establishing limit theorems associated with the functional estimators. These results constitute our main contribution and allow us to construct confidence intervals, confidence bands and parametric tests.

Let us formulate the estimation problem. The identification of such a model is a difficult task since the increment process of a multifractional Brownian motion is no longer stationary, no longer a self-similar process, and its path regularity explicitly varies with time. However, several nice properties of fractional Brownian motion still hold locally for multifractional Brownian motion. Indeed, assuming that $H \in \mathcal{C}^\eta([0, 1])$ and is such that $\sup_t H(t) < \min(1, \eta)$, Benassi *et al.* (1998) proved that

$$\lim_{\epsilon \rightarrow 0^+} \left(\frac{W(t + \epsilon u) - W(t)}{\epsilon^{2H(t)}} \right)_{u \in \mathbb{R}^+} \stackrel{d}{=} C(B_{H(t)}(u))_{u \in \mathbb{R}^+}, \tag{5}$$

where $B_{H(t)}(u)$ denotes a standard fractional Brownian motion with parameter $H(t)$ defined on \mathbb{R}^+ . To estimate the Hurst function $H(\cdot)$ of a multifractional Brownian motion, this result suggests the local adaptation of global methods used to identify a fractional Brownian motion. The method we propose is a local version of the quadratic variations method studied by Istas and Lang (1997), Kent and Wood (1997) and Coeurjolly (2001). It involves first filtering the observations of a self-similar (or locally self-similar at 0) stationary Gaussian process to weaken the dependence of the observations, and then estimating the empirical second-order moment of the filtered series. For the fractional Brownian motion case, this method exhibits nice properties: it produces estimators having rate of convergence that achieve Cramér–Rao bounds (for fractional Brownian motion parameters) (Coeurjolly and Istas 2001) and it is computationally fast, numerically stable and behaves efficiently with respect to the maximum likelihood for small sample sizes (Coeurjolly 2000, p. 35).

The rest of this paper is organized as follows. Section 2 introduces some notation and defines the local H_2 -variations statistic. We prove some convergence results and apply them to the identification problem in Section 3. We derive estimators of Hölderian Hurst functions, and prove their consistency and asymptotical normality. When C is assumed to be known the estimators derived in Coeurjolly (2000) had higher convergence rate, but such an assumption seems quite unrealistic. Our method is a local one and as such depends on a neighbourhood size whose choice is discussed in Section 4, where a procedure to estimate the optimal neighbourhood is proposed and studied. A simulation study is conducted in Section 5 to explore the qualities of the estimators. Finally, proofs of different results are presented in Section 6.

2. Local H_2 -variations of a multifractional Brownian motion

In this section we introduce some notation and derive convergence results for the local second-order moment of the discrete variations of a multifractional Brownian motion. Later on, only a path $(W_t)_{t \in [0,1]}$ of a multifractional Brownian motion with Hurst function $H(\cdot)$ and scaling coefficient C is considered, and our statistical model corresponds to its discretized version $(\mathbf{W}) = (W(i/N))_{i=1, \dots, N}$. The Hurst function $H(\cdot)$ is assumed to be a Hölderian function defined on $[0, 1]$, of order $0 < \eta \leq 1$, and such that $\sup_t H(t) < \min(1, \eta)$. Denote by a a filter of length $\ell + 1$ and of order $p \geq 1$, that is, a vector with real components such that

$$\sum_{q=0}^{\ell} a_q q^i = 0, \quad \text{for } i = 0, \dots, p - 1,$$

$$\sum_{q=0}^{\ell} a_q q^p \neq 0.$$

Let (\mathbf{V}^a) be the series obtained by filtering (\mathbf{W}) with a , that is,

$$V^a\left(\frac{j}{N}\right) = \sum_{q=0}^{\ell} a_q W\left(\frac{j-q}{N}\right), \quad \text{for } j = \ell + 1, \dots, N - 1.$$

For example, when $a = (1, -1)$, (\mathbf{V}^a) represents the increments (\mathbf{W}) , and when $a = (1, -2, 1)$, (\mathbf{V}^a) represents the second-order differences of (\mathbf{W}) . As Lemma 1 reveals, filtering the discretized path of a multifractional Brownian motion allows the series to be made locally stationary and the dependence structure between observations to be destroyed locally. Now, denote by $V_{N,\varepsilon}(t, a)$ the following random variable:

$$V_{N,\varepsilon}(t, a) = \frac{1}{v_N(t)} \sum_{j \in \mathcal{V}_{N,\varepsilon}(t)} \left\{ \frac{V^a(j/N)^2}{\mathbb{E}(V^a(j/N)^2)} - 1 \right\}, \tag{6}$$

where $\mathcal{V}_{N,\varepsilon}(t)$ denotes a neighbourhood of t , defined, for a parameter $\varepsilon > 0$, by $\mathcal{V}_{N,\varepsilon}(t) = \{j = \ell + 1, \dots, N, |j/N - t| \leq \varepsilon\}$, and where $v_N(t) = \#\mathcal{V}_{N,\varepsilon}(t)$. From now on, we define ε as a function of N , say ε_N , such that $\varepsilon = \varepsilon_N \rightarrow 0$ and $N\varepsilon_N \rightarrow +\infty$ as $N \rightarrow +\infty$. The neighbourhood $\mathcal{V}_{N,\varepsilon}(t)$ is sure to contain asymptotically an infinite number of points and to be of length asymptotically zero. More precisely, we suppose the specific form below for ε_N :

$$\varepsilon = \varepsilon_N = \kappa N^{-\alpha} \log(N)^\beta, \quad \text{with } \kappa > 0, 0 < \alpha < 1, \beta \in \mathbb{R}. \tag{7}$$

Remark. The statistics $V_{N,\varepsilon}(t, a)$ can be seen as the local H_2 -variations of a certain Gaussian process (H_2 being the second Hermite polynomial defined by $H_2(t) = t^2 - 1$).

We can now state convergence results for the local H_2 -variations of a discretized path of a multifractional Brownian motion (almost surely and in distribution for the topology of Skorohod).

Proposition 1. (i) Let $t \in [0, 1]$, let a be a filter of order $p \geq 1$, and let ε_N be of the form (7). Then, as $N \rightarrow +\infty$, we have almost surely

$$V_{N,\varepsilon}(t, a) \rightarrow 0. \tag{8}$$

(ii) Let a be a filter of order $p > \bar{H} + 1/4$, where $\bar{H} = \sup_t H(t)$, and let ε_N be of the form (7). Then, as $N \rightarrow +\infty$, the following convergence in distribution on $]0, 1[$ holds:

$$\sqrt{2N\varepsilon_N} V_{N,\varepsilon}(\cdot, a) \longrightarrow \mathbb{G}, \tag{9}$$

where $\mathbb{G} = \{\mathbb{G}(t), t \in]0, 1[\}$ is a centred Gaussian process with covariance function given for $s, t \in]0, 1[$ by

$$\text{cov}(\mathbb{G}(s), \mathbb{G}(t)) = 2 \sum_{k \in \mathbb{Z}} \frac{\pi_{H(s)/2+H(t)/2}^a(k)^2}{\pi_{H(s)}^a(0)\pi_{H(t)}^a(0)}. \tag{10}$$

Moreover, the function $\pi^a(k)$, defined for $H \in]0, 1[$, is given by

$$\pi_H^a(k) = -\frac{1}{2} \sum_{q,q'=0}^{\ell} a_q a_{q'} |q - q' + k|^{2H}. \tag{11}$$

These two results are local versions of the ones obtained for the fractional Brownian motion case (Coeurjolly 2001, Proposition 1). Note that a filter of order at least 2 ensures asymptotic normality for all the values of the function $H(\cdot)$. For a filter of order 1 (i.e. the filter $a = (1, -1)$), this convergence is available if and only if $0 < \bar{H} < 3/4$.

We now turn our attention to the identification of a multifractional Brownian motion and, in particular, the estimation of the Hurst function using a method of moments.

3. Applications to the Hurst function estimation

Let us introduce, for $m \geq 1$, the filter defined, for $i = 0, \dots, m\ell$, by

$$a_i^m = \begin{cases} a_j, & \text{if } i = jm, \\ 0, & \text{otherwise,} \end{cases}$$

which is nothing more than the filter a dilated m times. Define

$$S_{N,\varepsilon}(t, a^m) = \frac{1}{v_N(t)} \sum_{j \in \mathcal{V}_{N,\varepsilon}(t)} V^{a^m} \left(\frac{j}{N} \right)^2, \quad \text{for } t \in [0, 1]. \tag{12}$$

The interest of the sequence $(a^m)_{m \geq 1}$ relies on the fact that $\pi_H^{a^m}(0) = m^{2H} \pi_H^a(0)$. By virtue of Lemma 1, we have

$$\begin{aligned} \mathbb{E}(S_{N,\varepsilon}(t, a^m)) &= \frac{C \times \pi_{H(t)}^{a^m}(0)}{N^{2H(t)}} + \mathcal{O}(\varepsilon_N^\eta \log(N)) \\ &= m^{2H(t)} \times \frac{C \times \pi_{H(t)}^a(0)}{N^{2H(t)}} + \mathcal{O}(\varepsilon_N^\eta \log(N)), \end{aligned}$$

which can be restated as

$$\log \mathbb{E}(S_{N,\varepsilon}(t, a^m)) \sim 2H(t) \log(m) + \log(C \times g_{N,a}(H(t))), \quad \text{as } N \rightarrow +\infty.$$

Let $M \geq 2$ be an integer. The above relation suggests estimating $H(t)$ by a simple local linear regression of $L_{N,\varepsilon}(t, a, M) = (\log(S_{N,\varepsilon}(t, a^m)))_{m=1, \dots, M}$ on $(\log(m))_{m=1, \dots, M}$. We obtain a class of estimators, defined for $t \in]0, 1[$ by

$$\hat{H}_{N,\varepsilon}(t, a, M) = \frac{A^t}{2\|A\|^2} L_{N,\varepsilon}(t, a, M), \tag{13}$$

where A is the vector defined for $m = 1, \dots, M$ by $A_m = \log(m) - M^{-1} \sum_{m=1}^M \log(m)$. This class is a local version of the one obtained for estimating the Hurst parameter of a non-standard fractional Brownian motion (Coeurjolly 2001). Note that the functional estimator of $H(\cdot)$ is clearly independent of the value of C .

Proposition 2. *Let a be a filter of order $p > \bar{H} + 1/4$, where $\bar{H} = \sup_t H(t)$, $M \geq 2$ an integer, and assume ε_N is of the form (7).*

(i) *Then, as $N \rightarrow +\infty$, we have, for all $t \in [0, 1]$,*

$$\text{bias}(\hat{H}_{N,\varepsilon}(t, a, M)) = \mathcal{O}(\varepsilon_N^\eta \log(N)), \quad \text{var}(\hat{H}_{N,\varepsilon}(t, a, M)) = \mathcal{O}\left(\frac{1}{N\varepsilon_N}\right)$$

and, almost surely, $\hat{H}_{N,\varepsilon}(t, a, M) \rightarrow H(t)$. (14)

(ii) *Assume ε_N is of the form (7) with $\alpha \geq 1/(2\eta + 1)$ and $\beta < 0$. On $]0, 1[$, the following convergence in distribution holds:*

$$\sqrt{2N\varepsilon_N}(\hat{H}_{N,\varepsilon}(\cdot, a, M) - H(\cdot)) \rightarrow \mathbb{G}', \tag{15}$$

where $\mathbb{G}' = \{\mathbb{G}'(t), t \in]0, 1[\}$ is a centred Gaussian process with covariance function given for $t, t' \in]0, 1[$ by:

$$\text{cov}(\mathbb{G}'(t), \mathbb{G}'(t')) = \frac{1}{4\|A\|^4} A' \Sigma \left(\frac{H(t)}{2} + \frac{H(t')}{2}, H(t), H(t') \right) A, \tag{16}$$

with $\Sigma(H_1, H_2, H_3)$ the $M \times M$ matrix whose (m, n) th entry is

$$(\Sigma(H_1, H_2, H_3))_{m,n} = 2 \sum_{j \in \mathbb{Z}} \frac{\pi_{H_1}^{a^m, a^n}(j)^2}{\pi_{H_2}^{a^m}(0)\pi_{H_3}^{a^n}(0)}, \quad m, n = 1, \dots, M.$$

where

$$\pi_H^{a^m, a^n}(j) = \sum_{q=0}^{m\ell} \sum_{q'=0}^{n\ell} a_q a_{q'} |mq - nq' + j|^{2H}. \tag{17}$$

Remarks.

- Benassi *et al.* (1998) proved the consistency of $\hat{H}_{N,\varepsilon}$ (for the particular filter $(1, -2, 1)$) under the condition that $\varepsilon_N = \mathcal{O}(N^{-\alpha} \log(N)^\beta)$ with $0 < \alpha < 1/2$.
- The condition $\alpha \geq 1/(2\eta + 1)$ and $\beta < 0$ in (ii) ensures that $\varepsilon_N^{2\eta} \log(N)^2 = o(N\varepsilon_N)$.
- To estimate $H(t)$ we could have performed a weighted linear regression of $L_{N,\varepsilon}(t, a, M)$ on $(\log(m))_{m=1, \dots, M}$. In this case, it is easy to derive a result similar to Proposition 2 by locally adapting the corresponding result in Coeurjolly (2000, Proposition 2.5). We have not done this here, because we believe that the gain is too small with respect to the computational cost involved in the estimation of the covariance matrix Σ .

4. Optimal neighbourhood

We now analyse the asymptotic behaviour of the mean integrated squared error (MISE) of $\hat{H}_{N,\varepsilon}(\cdot) - H(\cdot)$. Such a criterion is widely used in functional estimation. Thanks to Proposition 2, it is clear that the MISE (depending on ε_N) has the following behaviour:

$$\text{MISE}(\varepsilon_N) = \mathbb{E} \left(\int_0^1 \{ \hat{H}_{N,\varepsilon}(t) - H(t) \}^2 dt \right) = \mathcal{O} \left(\varepsilon_N^{2\eta} \log(N)^2 \right) + \mathcal{O} \left(\frac{1}{N\varepsilon_N} \right).$$

Considering a discretized version of the MISE, say

$$R_N(\varepsilon) = \frac{1}{N} \sum_{i=1}^N \left(\hat{H}_{N,\varepsilon} \left(\frac{i}{N} \right) - H \left(\frac{i}{N} \right) \right)^2,$$

it is immediately evident that the asymptotic behaviour of $\mathbb{E}(R_N(\varepsilon))$ is the same as that of the MISE. Let $\varepsilon_N^* = \text{argmin}_{\varepsilon_N} E(R_N(\varepsilon_N))$. From (7), $\varepsilon_n^* = K_n^{*-\alpha} \log(n)^{\beta^*}$ and one can easily see that

$$\alpha^* = \frac{1}{2\eta + 1}, \beta^* = -\frac{2}{2\eta + 1}, \text{ and } \text{MISE}(\varepsilon_N^*) = \mathcal{O} \left(N^{-2\eta/(2\eta+1)} \log(N)^{2/(2\eta+1)} \right).$$

Define also

$$R'_N(\varepsilon) = \frac{1}{N} \sum_{i=1}^N \hat{H}_{N,\varepsilon} \left(\frac{i}{N} \right)^2 - \frac{2}{N} \sum_{i=1}^N \hat{H}_{N,\varepsilon} \left(\frac{i}{N} \right) H \left(\frac{i}{N} \right).$$

Since the function $H(\cdot)$ is independent of ε_N , we have

$$\text{argmin}_{\varepsilon_N} \mathbb{E}(R'_N(\varepsilon_N)) = \text{argmin}_{\varepsilon_N} \mathbb{E}(R_N(\varepsilon_N)) = \mathcal{O} \left(N^{-2\eta/(2\eta+1)} \log(N)^{2/(2\eta+1)} \right).$$

The above asymptotic result suggests a natural procedure for estimating the optimal neighbourhood. We propose estimating $R'_N(\varepsilon)$ by

$$\hat{R}'_N(\varepsilon) = \frac{1}{N} \sum_{i=1}^N \hat{H}_{N,\varepsilon} \left(\frac{i}{N} \right)^2 - \frac{2}{N} \sum_{i=1}^N \hat{H}_{N,\varepsilon} \left(\frac{i}{N} \right) \tilde{H}_{N,\varepsilon^2} \left(\frac{i}{N} \right)$$

where $\tilde{H}_{N,\varepsilon^2}(i/N)$ is defined by

$$\tilde{H}_{N,\varepsilon^2}(i/N) = \frac{1}{\#\mathcal{V}_{N,\varepsilon^2}(i/N)} \sum_{j \in \mathcal{V}_{N,\varepsilon^2}(i/N)} \hat{H}_{N,\varepsilon^2}(j/N) \tag{18}$$

and represents the average of $\hat{H}_{N,\varepsilon^2}(\cdot)$, in a neighbourhood of i/N of size (of the order of) $N\varepsilon_N^2$, which is the functional estimation of $H(\cdot)$ calculated with a neighbourhood of the same size. Now write

$$\hat{\varepsilon}_N^* = \text{argmin}_{\varepsilon_N \in E} \hat{R}'_N(\varepsilon_N),$$

where the set E is defined by

$$E = \left\{ (\varepsilon_N)_{N \geq 1}, \varepsilon_N = \kappa N^{-\alpha} \log(N)^\beta, \text{ with } \frac{1}{2\eta + 1} \leq \alpha < 1, \text{ and } \beta < -\frac{1}{6} \right\}. \tag{19}$$

The set E is assumed to be discrete and to contain at most n^ρ elements, for some $\rho > 0$.

The following result proves the almost sure convergence of $\hat{H}_{N,\hat{\varepsilon}_N^*}$ towards $H(t)$ and proves its optimality in the sense that the average of the empirical risk calculated with $\hat{\varepsilon}_N^*$ is equivalent to $\mathbb{E}(R_N(\hat{\varepsilon}_N^*))$.

Proposition 3. (i) For all $t \in [0, 1]$, we have almost surely, as $N \rightarrow +\infty$,

$$\hat{H}_{N,\hat{\varepsilon}_N^*} \rightarrow H(t). \tag{20}$$

(ii) As $N \rightarrow +\infty$,

$$\frac{\mathbb{E}(R_N(\hat{\varepsilon}_N^*))}{\mathbb{E}(R_N(\varepsilon_N^*))} \rightarrow 1, \quad \text{i.e. } \mathbb{E}(R_N(\hat{\varepsilon}_N^*)) = \mathcal{O}\left(N^{-2\eta/(2\eta+1)} \log(N)^{2/(2\eta+1)}\right). \tag{21}$$

5. A simulation study

To generate a sample path of a standard multifractional Brownian motion discretized at times $i/N, i = 1, \dots, N$, with covariance matrix $C_{H(\cdot)}$, one can simply extract the square root of $C_{H(\cdot)}$. Then, one generates $Z \rightsquigarrow \mathcal{N}(0, I_N)$ and estimates $\mathbf{W} := C_{H(\cdot)}^{1/2} Z$. This method (which is exact in theory) is sufficiently fast for reasonable sample size N (which is the case here since we chose $N = 2500$). For larger N , it becomes expensive in time and memory, and is numerically unstable. We consider two types of Hurst function: a linear function and a logistic one:

$$H_1(t) = 0.1 + (0.9 - 0.1)t, \tag{22}$$

$$H_2(t) = 0.3 + 0.3/(1 + \exp(-100(t - 0.7))). \tag{23}$$

We generate $R = 50$ series of length $N = 2500$. A Daubechies filter of order 4, $a = Db4$ (with two zero moments) is used to define $\hat{H}_{N,\varepsilon}$. We fix the number of filters to $M = 5$. Let us concentrate first on the optimal neighbourhood. For the sake of simplicity, we choose ε such that $N\varepsilon$ is an integer. Define $h = N\varepsilon, h^* = N\varepsilon^*$ and $\hat{h}^* = N\hat{\varepsilon}^*$. Table 1 summarizes the different estimates of $h^*, \hat{h}^*, \hat{E}(R_N(\varepsilon^*))$ and $\hat{E}(\hat{R}'_N(\hat{\varepsilon}^*))$.

Then we applied the previous procedure for each of the $R = 50$ paths to the estimation of the Hurst function. Figure 1 displays the empirical distribution of the functional estimator together with the true function and the estimated confidence bands (at a confidence level $1 - \alpha = 0.95$).

6. Proofs

6.1. Local quadratic variations

Before analysing the asymptotic behaviour of $V_{N,\varepsilon}(t, a)$, we need the following lemma concerning the correlation structure of (\mathbf{V}^a) .

Lemma 1. Let a be a filter of order $p \geq 1$, let $t, t' \in [0, 1]$, let $j \in \mathcal{V}_{N,\varepsilon}(t), j' \in \mathcal{V}_{N,\varepsilon}(t')$ and let $\varepsilon = \varepsilon_N$ be of the form (7).

Table 1. Theoretical and estimated optimal empirical windows and associated risks ($\hat{\mathbb{E}}$ denotes the empirical mean based on the 50 paths generated)

	h^*	$\hat{\mathbb{E}}(\hat{h}^*)$	$\hat{E}(R_N(\varepsilon^*))$	$\hat{E}(R_N(\hat{\varepsilon}^*))$
$H(t)$ linear	203	239.76	3.88×10^{-3}	3.631×10^{-3}
$H(t)$ logistic	192	226	4.17×10^{-3}	3.91×10^{-3}

(i) We have, as $N \rightarrow +\infty$,

$$\mathbb{E}\left(V^a\left(\frac{j}{N}\right)V^a\left(\frac{j'}{N}\right)\right) = \frac{C^2}{N^{H(t)+H(t')}}\pi_{H(t)/2+H(t')/2}^a(j'-j) \times \{1 + \mathcal{O}(\varepsilon_N^\eta \log(N))\}, \quad (24)$$

where

$$\pi_H^a(k) = -\frac{1}{2} \sum_{q,q'=0}^{\ell} a_q a_{q'} |q - q' + k|^{2H}.$$

(ii) Define $Z(j) = \frac{V^a(j/n)}{\mathbb{E}(V^a(j/n)^2)^{1/2}}$. We have

$$\mathbb{E}(Z(j)Z(j')) = \frac{\pi_{H(t)/2+H(t')/2}^a(j'-j)}{\{\pi_{H(t)}^a(0)\pi_{H(t')}^a(0)\}^{1/2}} \times \{1 + \mathcal{O}(\varepsilon_N^\eta \log(N))\}. \quad (25)$$

(iii) Moreover, as $k \rightarrow +\infty$, we have $\pi_H^a(k) = \mathcal{O}(|k|^{2H-2p})$, $\forall H \in]0, 1[$.

Proof. (i) To compute the covariance function, the stochastic representation of a multifractional Brownian motion is used. For the sake of simplicity, let us denote $C'(t) = C \times K(2H(t))^{1/2}/\sqrt{2}$. From (2) and the change of variables $\lambda = Nu$, we obtain

$$\begin{aligned} &\mathbb{E}\left(V^a\left(\frac{j}{N}\right)V^a\left(\frac{j'}{N}\right)\right) \\ &= \int_{\mathbb{R}} \sum_{q,q'} a_q a_{q'} (e^{i(j-q)u} - 1)(e^{-i(j'-q')u} - 1) \frac{C'((j-q)/N)C'((j'-q')/N)}{|Nu|^{H((j-q)/N)+H((j'-q')/N)}} N du \\ &= A + B, \end{aligned}$$

where

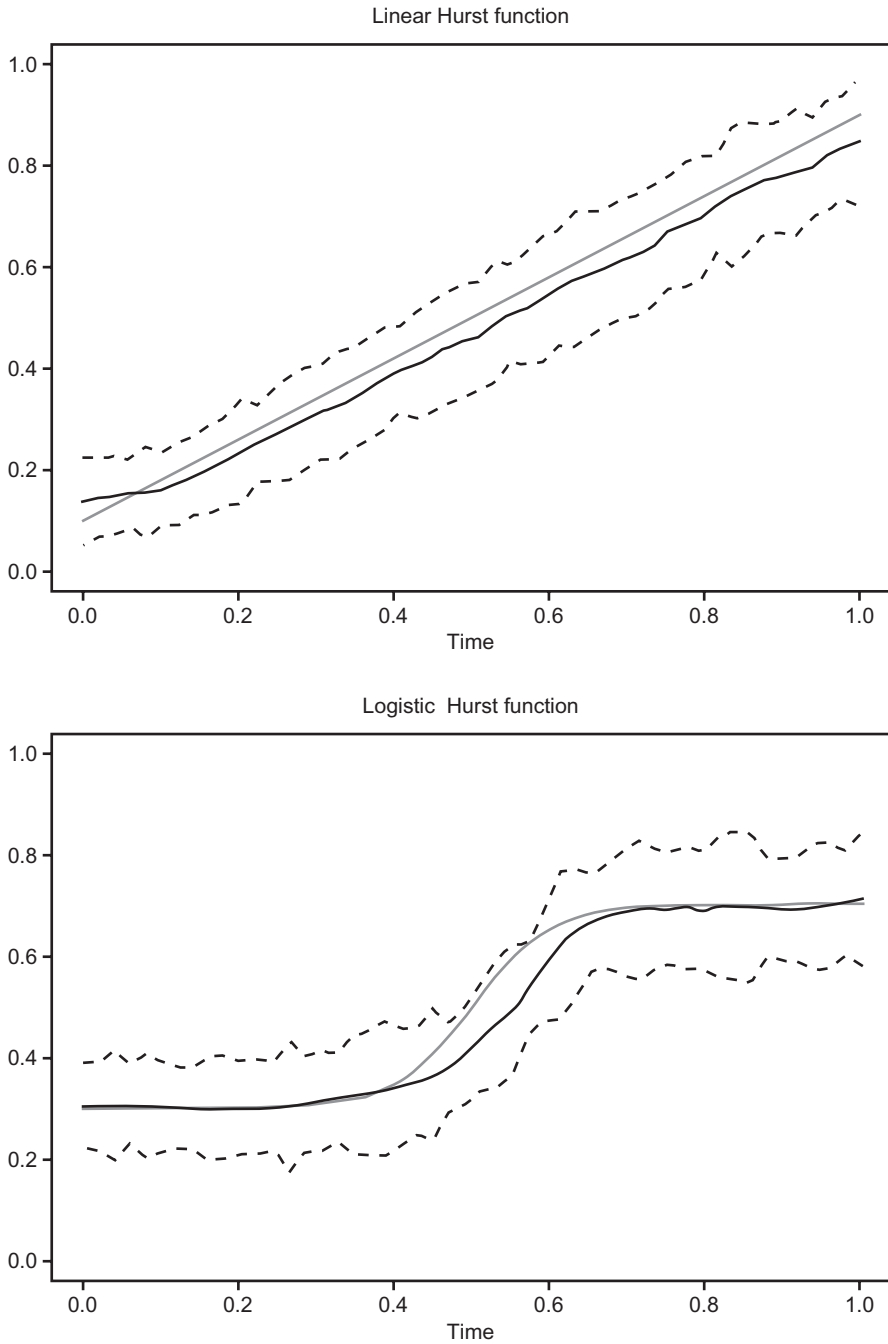


Figure 1. Empirical distributions for the functional estimators of two Hurst functions defined by (22) and (23), and theoretical discretized confidence bands to the level $1 - \alpha = 0.95$.

$$\begin{aligned}
 A &= \int_{\mathbb{R}} \sum_{q,q'=0}^{\ell} a_q a_{q'} (e^{i(j-q)u} - 1) (e^{-i(j'-q')u} - 1) \frac{C'(t)C'(t')}{N^{H(t)+H(t')}} \frac{du}{|u|^{H(t)+H(t')+1}}, \\
 B &= \int_{\mathbb{R}} \sum_{q,q'=0}^{\ell} a_q a_{q'} (e^{i(j-q)u} - 1) (e^{-i(j'-q')u} - 1) \frac{|N|^{-H(t)-H(t')}}{|u|^{H(t)+H(t')+1}} \\
 &\quad \times \left\{ \frac{C'((j-q)/N)C'((j'-q')/N)}{|Nu|^{H((j-q)/N)+H((j'-q')/N)-H(t)-H(t')}} - C'(t)C'(t') \right\} du.
 \end{aligned}$$

Since the filter is of order at least 1,

$$\sum_{q,q'=0}^{\ell} a_q a_{q'} (e^{i(j-q)u} - 1) (e^{-i(j'-q')u} - 1) = \sum_{q,q'=0}^{\ell} e^{i(j-j'+q'-q)u} a_q a_{q'},$$

which allows us to rewrite A as

$$\begin{aligned}
 A &= N^{-H(t)-H(t')} C'(t)C'(t') \sum_{q,q'=0}^{\ell} a_q a_{q'} \int_{\mathbb{R}} \frac{e^{i(j-j'+q'-q)u}}{|u|^{H(t)+H(t')+1}} du \\
 &= -\frac{1}{2} N^{-H(t)-H(t')} C^2 \sum_{q,q'=0}^{\ell} a_q a_{q'} |q - q' + j' - j|^{H(t)+H(t')} \\
 &= \frac{C^2}{N^{H(t)+H(t')}} \pi_{H(t)/2+H(t')/2}^{\alpha} (j' - j).
 \end{aligned}$$

Since $H \in C^{\eta}([0, 1])$,

$$|Nu|^{H((j-q)/N)-H(t)+H((j'-q')/N)-H(t')} = 1 + \mathcal{O}(\varepsilon_N^{\eta} \log(|Nu|)), \tag{26}$$

$$C'\left(\frac{j-q}{N}\right) = C'(t) + \mathcal{O}(\varepsilon_N^{\eta}) \quad \text{and} \quad C'\left(\frac{j'-q'}{N}\right) = C'(t') + \mathcal{O}(\varepsilon_N^{\eta}). \tag{27}$$

Equations (26) and (27) enable the following upper bound to be obtained for B :

$$\begin{aligned}
 B &\leq \frac{C^2}{N^{H(t)+H(t')}} \pi_{H(t)/2+H(t')/2}^{\alpha} (j' - j) \mathcal{O}(\varepsilon_N^{\eta} \log(N)) \\
 &\quad + C'(t)C'(t') \sum_{q,q'=0}^{\ell} a_q a_{q'} \left(\int_{\mathbb{R}} \frac{e^{i(j-j'+q'-q)u}}{|u|^{H(t)+H(t')+1}} \log(|u|) du \right) \times \mathcal{O}(\varepsilon_N^{\eta}).
 \end{aligned}$$

In a neighbourhood of 0,

$$\sum_{q,q'} a_q a_{q'} \frac{e^{i(j-j'+q'-q)u}}{|u|^{H(t)+H(t')+1}} \log(|u|) = o(|u|^{2p-H(t)-H(t')} \log(|u|)).$$

Thus we can conclude that

$$\mathbb{E}\left(V^a\left(\frac{j}{N}\right)V^a\left(\frac{j'}{N}\right)\right) = \frac{C^2}{N^{H(t)+H(t')}}\pi_{H(t)/2+H(t')/2}^a(j'-j)\{1 + \mathcal{O}(\varepsilon_N^\eta \log(N))\}.$$

(ii) The proof is trivial.

(iii) We refer the reader to Coeurjolly (2001, Lemma 1) for the proof of the asymptotic expansion of $\pi_H^a(k)$. □

Proof of Proposition 1(i). Let us define

$$Z(j) = \frac{V^a(j/N)}{\mathbb{E}(V^a(j/N)^2)^{1/2}}$$

and let H_2 the second Hermite polynomial defined by $H_2(u) = u^2 - 1$. We obtain

$$\mathbb{E}(V_{N,\varepsilon}(t, a)^2) = \frac{1}{v_N(t)^2} \sum_{j,j' \in \mathcal{V}_{N,\varepsilon}(t)} \mathbb{E}(H_2(Z(j))H_2(Z(j'))),$$

From Lemma 1(ii),

$$\begin{aligned} \mathbb{E}(H_2(Z(j))H_2(Z(j'))) &= 2\mathbb{E}(Z(j)Z(j'))^2 \\ &= 2\frac{\pi_{H(t)}^a(j'-j)^2}{\pi_{H(t)}^a(0)^2} \{1 + \mathcal{O}(\varepsilon_N^\eta \log(N))\}, \quad \text{as } N \rightarrow +\infty. \end{aligned}$$

Thus, as $N \rightarrow +\infty$,

$$\begin{aligned} \mathbb{E}(V_{N,\varepsilon}(t, a)^2) &\sim \frac{2}{v_N(t)^2} \sum_{j,j'} \frac{\pi_{H(t)}^a(j'-j)^2}{\pi_{H(t)}^a(0)^2} \\ &\sim \frac{2}{v_N(t)^2} \sum_{|j| \leq v_N(t)} (v_N(t) - |j|) \frac{\pi_{H(t)}^a(j)^2}{\pi_{H(t)}^a(0)^2} \\ &= \mathcal{O}\left(\frac{1}{v_N(t)} \sum_{|j| \leq v_N(t)} \frac{\pi_{H(t)}^a(j)^2}{\pi_{H(t)}^a(0)^2}\right). \end{aligned}$$

Lemma 1 gives the upper bound $\pi_{H(t)}^a(j) \leq \mathcal{O}(|j|^{4H(t)-4p})$. Thus,

$$\mathbb{E}(V_{N,\varepsilon}(t, a)^2) = \begin{cases} \mathcal{O}\left(\frac{1}{v_N(t)}\right), & \text{if } p > H(t) + 1/4 \text{ (i.e. if } p \geq 2 \text{ or } p = 1 \text{ and } H(t) < 3/4), \\ \mathcal{O}\left(\frac{\log(v_N(t))}{v_N(t)}\right), & \text{if } p = 1 \text{ and } H(t) = 3/4, \\ \mathcal{O}\left(\frac{1}{v_N(t)^{4-4H(t)}}\right), & \text{if } p = 1 \text{ and } H(t) > 3/4. \end{cases} \tag{28}$$

Therefore, for all $p \geq 1$, and for all $H(t) \in]0, 1[$, there exists $\alpha > 1$ such that $\mathbb{E}(V_{N,\varepsilon}(t, a)^2) = \mathcal{O}(v_N(t)^{-\alpha})$. We adapt a result of Doob (1953, p. 492) giving a condition for which the empirical mean of a stationary (centred discretized) process tends almost surely to 0. Let $\bar{V}_{v_N(t)} = V_{N,\varepsilon}(t, a)$.

Let $\alpha' \in \mathbb{R}$ and $m \in \mathbb{N}$ be such that $\alpha\alpha' > 1$ and $v_N(t) > m^{\alpha'}$. We have $\mathbb{E}(\bar{V}_{v_N(t)}^2) = \mathcal{O}(m^{-\alpha\alpha'})$. If $(w_m)_{m \geq 1}$ denotes the sequence of integers defined by $w_m = [m^{\alpha'}] + 1$, then, for all $\varepsilon > 0$,

$$\mathbb{P}(|\bar{V}_{w_m}| > \varepsilon) \leq \frac{K^2}{\varepsilon^2} \frac{1}{m^{\alpha\alpha'}}.$$

From the Borel–Cantelli lemma, we have, as $m \rightarrow +\infty$, $\bar{V}_{w_m} \rightarrow 0$ almost surely. Furthermore,

$$\begin{aligned} \mathbb{E}\left(\max_{w_m \leq v_N(t) \leq w_{m+1}} \left| \bar{V}_{v_N(t)} - \frac{w_m}{v_N(t)} \bar{V}_{w_m} \right|^2\right) &\leq \frac{1}{|w_m|^2} \mathbb{E}\left(\left(\sum_{j=w_m}^{w_{m+1}} H_2(Z(j))\right)^2\right) \\ &= \frac{2}{|w_m|^2} \sum_{|j| \leq w_{m+1} - w_m} (w_{m+1} - w_m - |j|) \frac{\pi_{H(t)}^a(j' - j)^2}{\pi_{H(t)}^a(0)^2} \\ &\leq \frac{K}{w_m^2} \frac{(w_{m+1} - w_m)^2}{w_m^2} \leq \frac{K'}{m^2}. \end{aligned}$$

Therefore, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\max_{w_m \leq v_N(t) \leq w_{m+1}} \left| \bar{V}_{v_N(t)} - \frac{w_m}{v_N(t)} \bar{V}_{w_m} \right| > \varepsilon\right) \leq \frac{K'^2}{\varepsilon^2} \frac{1}{m^2},$$

and by Borel–Cantelli lemma

$$\bar{V}_{v_N(t)} - \frac{w_m}{v_N(t)} \bar{V}_{w_m} \rightarrow 0, \quad \text{almost surely as } m \rightarrow +\infty.$$

Thus, we have the almost sure convergence of $\bar{V}_{v_N(t)}$ since $\bar{V}_{w_m} \rightarrow 0$ almost surely and since $v_N(t) > m^{\alpha'}$.

Finally, from previous computations, note that if $p > H(t) + 1/4$, then

$$\mathbb{E}(V_{N,\varepsilon}(t, a)^2) \sim \frac{2}{v_N(t)} \sum_{j \in \mathbb{Z}} \frac{\pi_{H(t)}^a(j' - j)^2}{\pi_{H(t)}^a(0)^2}, \quad \text{as } N \rightarrow +\infty.$$

□

Lemma 2. (i) If $p > H(t) + 1/4$, the following convergence in distribution holds:

$$\sqrt{v_N(t)} V_{N,\varepsilon}(t, a) \xrightarrow{d} \mathcal{N}(0, \sigma^2(H(t), a)), \tag{29}$$

with

$$\sigma^2(H(t), a) = 2 \sum_{j \in \mathbb{Z}} \frac{\pi_{H(t)}^a(j)^2}{\pi_{H(t)}^a(0)^2}. \tag{30}$$

(ii) Let $d > 1$ and let $t_1, \dots, t_d \in [0, 1]$. Then if $p > \bar{H} + 1/4$, with $\bar{H} = \sup_t H(t)$, we have

$$\left(\sqrt{v_N(t_1)} V_{N,\varepsilon}(t_1, a), \dots, \sqrt{v_N(t_d)} V_{N,\varepsilon}(t_d, a) \right)^T \xrightarrow{d} (\mathbb{G}(t_1), \dots, \mathbb{G}(t_d))^T, \tag{31}$$

where $(\mathbb{G}(t_1), \dots, \mathbb{G}(t_d))^T$ is a centred Gaussian vector, such that, for all $i, j \in \{1, \dots, d\}$,

$$\text{cov}(\mathbb{G}(t_i), \mathbb{G}(t_j)) = 2 \sum_{k \in \mathbb{Z}} \frac{\pi_{H(t_i)/2+H(t_j)/2}^a(k)^2}{\pi_{H(t_i)}^a(0)\pi_{H(t_j)}^a(0)}.$$

Proof. (i) Recall that $Z(j)$ denotes the random variable $V^a(j/N)/\mathbb{E}(V^a(j/N)^2)^{1/2}$. From Theorem 1 of Breuer and Major (1983, p. 429) adapted to non-stationary Gaussian vectors, the necessary condition to obtain an asymptotic normality result for $V_{N,\varepsilon}(t, a)$ is the squared summability of $\mathbb{E}((j')Z(j+j'))$, for all $j' \in \mathbb{Z}$. But

$$\sum_{j \in \mathbb{Z}} \mathbb{E}(Z(j')Z(j'+j))^2 \sim \sum_{j \in \mathbb{Z}} \mathcal{O}(|j|^{4H(t)-4p}), \quad \text{as } N \rightarrow +\infty.$$

The result is obtained using the fact that $p > H(t) + 1/4$.

(ii) We treat the case $d = 2$, since the case $d > 2$ can easily be derived. Define for $\lambda, \mu \in \mathbb{R}$ and $t_1, t_2 \in]0, 1[$, the random variable $T_{N,\varepsilon}(\lambda, \mu) = \lambda V_{N,\varepsilon}(t_1, a) + \mu V_{N,\varepsilon}(t_2, a)$. Note that

$$\begin{aligned} T_{N,\varepsilon}(\lambda, \mu) &= \frac{1}{v_N(t_1)} \sum_{j_1 \in \mathcal{V}_{N,\varepsilon}(t_1)} \lambda H_2(Z(j_1)) + \frac{1}{v_N(t_2)} \sum_{j_2 \in \mathcal{V}_{N,\varepsilon}(t_2)} \mu H_2(Z(j_2)) \\ &\sim \frac{1}{v_N} \left(\sum_{j_1 \in \mathcal{V}_{N,\varepsilon}(t_1)} \lambda v(t_1) H_2(Z(j_1)) + \sum_{j_2 \in \mathcal{V}_{N,\varepsilon}(t_2)} \mu v(t_2) H_2(Z(j_2)) \right) \end{aligned}$$

where $v(t) = \lim_{N \rightarrow +\infty} v_N(t_1)/v_N$. Now rewrite $T_{N,\varepsilon}(\lambda, \mu)$ as a simple sum:

$$T_{N,\varepsilon}(\lambda, \mu) = \frac{1}{v_N} \sum_{j=1}^{v_N(t_1)+v_N(t_2)} g_j(Z(j^*)),$$

where

$$j^* = \begin{cases} (\mathcal{V}_{N,\varepsilon}(t_1))_j, & \text{if } 1 \leq j \leq v_N(t_1), \\ (\mathcal{V}_{N,\varepsilon}(t_2))_{j-v_N(t_1)}, & \text{if } v_N(t_1) < j \leq v_N(t_1) + v_N(t_2), \end{cases}$$

and

$$g_j(\cdot) = \begin{cases} \lambda v(t_1) H_2(\cdot), & \text{if } 1 \leq j \leq v_N(t_1) \\ \mu v(t_2) H_2(\cdot) & \text{if } v_N(t_1) < j \leq v_N(t_1) + v_N(t_2). \end{cases}$$

The function g_j clearly has Hermite rank 2. Moreover, for $j_1^*, j_2^* \in \{1, \dots, v_N(t_1) + v_N(t_2)\}$ ($j_1^* \leq j_2^*$), it follows from Lemma 1(ii) that

$$\mathbb{E}(Z(j_1^*)Z(j_2^*)) \xrightarrow{N \rightarrow +\infty} \begin{cases} \frac{\pi_{H(t_1)}^a(j_2^* - j_1^*)}{\pi_{H(t_1)}^a(0)} = \mathcal{O}(|j_2^* - j_1^*|^{2H(t_1)-2p}), & \text{if } j_1^*, j_2^* \leq v_N(t_1), \\ \frac{\pi_{H(t_2)}^a(j_2^* - j_1^*)}{\pi_{H(t_2)}^a(0)} = \mathcal{O}(|j_2^* - j_1^*|^{H(t_1)+H(t_2)-2p}), & \text{if } j_1^*, j_2^* > v_N(t_1), \\ \frac{\pi_{H(t_1)/2+H(t_2)/2}^a(j_2^* - j_1^*)}{\{\pi_{H(t_1)}^a(0)\pi_{H(t_2)}^a(0)\}^{1/2}} = \mathcal{O}(|j_2^* - j_1^*|^{H(t_1)+H(t_2)-2p}), & \text{otherwise.} \end{cases}$$

Thus, for all $j' \in \{1, \dots, v_N(t_1) + v_N(t_2)\}$, and since $p > \bar{H} + 1/4$, we obtain

$$\sum_{j=1}^{v_N(t_1)+v_N(t_2)} \mathbb{E}(Z(j^*)Z(j'^*))^2 = \mathcal{O}(1).$$

From Theorem 1 of Breuer and Major (1983, p. 429) adapted to non-stationary Gaussian vectors, there exists $\sigma^2(t_1, t_2)$ such that, for all $\lambda, \mu \in \mathbb{R}$, $\sqrt{N}\varepsilon T_{N,\varepsilon}(\lambda, \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(t_1, t_2))$. As a conclusion, the vector $(\sqrt{v_N(t_1)}V_{N,\varepsilon}(t_1, a), \sqrt{v_N(t_2)}V_{N,\varepsilon}(t_2))^T$ is asymptotically Gaussian. Finally, from previous computations,

$$\text{cov}\left(\sqrt{v_N(t_1)}V_{N,\varepsilon}(t_1, a), \sqrt{v_N(t_2)}V_{N,\varepsilon}(t_2, a)\right) \rightarrow 2 \sum_{j \in \mathbb{Z}} \frac{\pi_{H(t_1)/2+H(t_2)/2}^a(j)^2}{\pi_{H(t_1)}^a(0)\pi_{H(t_2)}^a(0)}. \tag{32}$$

□

To obtain the convergence in distribution of $V_{N,\varepsilon}(\cdot, a)$ for the topology of Skorohod, we need the following inequality, ensuring a tightness criterion.

Lemma 3. *Let a be a filter of order $p > \bar{H} + 1/4$ where $\bar{H} = \sup_t H(t)$, let r be an odd integer greater than 4 and let $t, t' \in I_{N,\varepsilon_N} = [\ell/N + \varepsilon_N, (N - 1)/N - \varepsilon_N]$. For $t \in I_{N,\varepsilon_N}$, let $v_N(t) = 2N\varepsilon_N$. Then*

$$\mathbb{E}\left((2N\varepsilon_N)^{r/2}(V_{N,\varepsilon}(t, a) - V_{N,\varepsilon}(t', a))^r\right) = \mathcal{O}(|t - t'|^{rm}). \tag{33}$$

Proof. Let $v_N = 2N\varepsilon_N$. Let $r \geq 4$ be an integer and let $t^* = [N(t' - t)]$. Then

$$\begin{aligned} \mathbb{E}\left(v_N^{r/2}(V_{N,\varepsilon}(t, a) - V_{N,\varepsilon}(t', a))^r\right) &= \frac{1}{v_N^{r/2}} \mathbb{E}\left\{ \sum_{j \in \mathcal{V}_{N,\varepsilon}(t)} H_2(Z(j)) - H_2(Z(j + t^*)) \right\}^r \\ &= \frac{1}{v_N^{r/2}} \sum_{j_1, \dots, j_r} \sum_{q=0}^r (-1)^q C_r^q \mathbb{E}(H_2(Z(j_1 + t^*)) \dots H_2(Z(j_q + t^*)) H_2(Z(j_{q+1})) \dots H_2(Z(j_r))). \end{aligned}$$

From the diagram formula (see, for example, Taqqu 1975),

$$\mathbb{E}(H_2(Z(j_1 + t^*)) \dots H_2(Z(j_q + t^*))H_2(Z(j_{q+1})) \dots H_2(Z(j_r))) = T_1 + T_2,$$

where T_1 (T_2) represents the terms obtained by the product of covariances (terms obtained by the product of covariances to the power 2).

Up to permutations on indices, each term of T_1 can be rewritten as

$$T_{j_1, \dots, j_r}^1 = \mathbb{E}(Z(j_1 + t^*)Z(j_2 + t^*)) \dots \mathbb{E}(Z(j_{q-1} + t^*)Z(j_q + t^*)) \dots \mathbb{E}(Z(j_{r-1})Z(j_r)). \tag{34}$$

From Lemma 1, there exists $K > 0$, $N_1 \in \mathbb{N}^*$, such that, for all $N \geq N_1$,

$$\begin{aligned} & \frac{1}{v_N^{r/2}} \sum_{j_1, \dots, j_r} T_{j_1, \dots, j_r}^1 \\ & \leq \frac{1}{v_N^{r/2}} \sum_{j_1, \dots, j_r} \pi_{H(t')}^a(j_2 - j_1) \dots \pi_{H(t')}^a(j_q - j_{q-1}) \pi_{H(t)/2 + H(t')/2}^a(j_{q+1} - j_q) \dots \\ & \quad \dots \pi_{H(t)}^a(j_r - j_{r-1}) \pi_{H(t)/2 + H(t')/2}^a(j_1 - j_r) \\ & \leq \frac{K}{v_N^{r/2}} \sum_{j_2, \dots, j_r} \left\{ \sum_{j_1} \pi_{H(t)/2 + H(t')/2}^a(j_1 - j_r) \pi_{H(t')}^a(j_2 - j_1) \right\} \dots \pi_{H(t)}^a(j_r - j_{r-1}). \end{aligned} \tag{35}$$

Let A_1, A_2 and A_3 be the covariance matrices related to the operators $\pi_{H(t)}^a$, $\pi_{H(t')}^a$ and $\pi_{H(t)/2 + H(t')/2}^a$, and let Q_α be the set of squared matrices with terms satisfying

$$|(A)_{j,k}| \leq c(1 + |k - j|)^{-\alpha}, \quad \alpha > 0, c > 0.$$

It is clear that A_1, A_2 and $A_3 \in Q_{2p-2\bar{H}}$, where $\bar{H} = \sup_t H(t)$. Moreover, Jaffard (1990) has proved that Q_α is an algebra, for $\alpha > 1$. Thus,

$$A_i A_j \in Q_{2p-2\bar{H}}, \quad \forall i, j = 1, 2, 3.$$

Iterating this argument leads to the existence of a matrix $B \in Q_{2p-2\bar{H}}$ such that

$$\frac{1}{v_N^{r/2}} \sum_{j_1, \dots, j_r} T_{j_1, \dots, j_r}^1 \leq \frac{C_1}{v_N^{r/2}} \sum_{j_r} (B)_{j_r j_r} \leq \frac{C_2}{v_N^{r/2-1}}.$$

Consequently,

$$\frac{1}{v_N^{r/2}} \sum_{j_1, \dots, j_r} \sum_{q=0}^r (-1)^q C_r^q \times T_1 \rightarrow 0, \quad \text{as } (N, \varepsilon) \rightarrow (+\infty, 0). \tag{36}$$

Turning now to the diagram formula and Lemma 1(ii), there exists $K > 0$, $N_2 \in \mathbb{N}^*$ such that, for all $N \geq N_2$,

$$\begin{aligned} & \frac{1}{v_N^{r/2}} \sum_{q=0}^r (-1)^q C_r^q \sum_{j_1, \dots, j_r} T_2 \leq K \sum_{q=0}^r (-1)^q C_r^q \\ & \times \sum_{\substack{i=0, \dots, \min(q, r-q) \\ q-i \text{ pair}}} \left\{ A_{\max(q, r-q)}^{|r-2q|} S\left(\frac{H(t)}{2} + \frac{H(t')}{2}\right)^i N_{r-q-i} S(H(t))^{(r-q-i)/2} N_{q-i} S(H(t'))^{(q-i)/2} \right\}. \end{aligned} \tag{37}$$

where

$$S(H) = \sum_{j \in \mathbb{Z}} \pi_H^a(j)^2 \quad \text{and} \quad N_\alpha = (\alpha - 1) \times (\alpha - 3) \times \dots \times 3 \times 1.$$

Using (36) and (37), we verify that

$$\mathbb{E}\left(v_N^{r/2} (V_{N,\varepsilon}(t, a) - V_{N,\varepsilon}(t', a))^r\right) = \mathcal{O}\left\{\left(S(H(t)) - 2S\left(\frac{H(t)}{2} + \frac{H(t')}{2}\right) + S(H(t'))\right)^{r/2}\right\}. \tag{38}$$

Let

$$U(t, t') = S(H(t)) - 2S\left(\frac{H(t)}{2} + \frac{H(t')}{2}\right) + S(H(t')).$$

Using (11), we obtain:

$$\begin{aligned} U(t, t') &= \frac{1}{4} \sum_{j \in \mathbb{Z}} \sum_{q_1, \dots, q_4} a_{q_1} \dots a_{q_4} (|q_1 - q_2 + j| |j|)^{H(t)+H(t')} \\ &\quad \times \left\{ (|q_1 - q_2 + j| |q_3 - q_4 + j|)^{H(t)-H(t')} - 1 + (|q_1 - q_2 + j| |q_3 - q_4 + j|)^{H(t')-H(t)} - 1 \right\} \\ &= \frac{1}{4} \sum_{j \in \mathbb{Z}} \sum_{q_1, \dots, q_4} a_{q_1} \dots a_{q_4} (|q_1 - q_2 + j| |q_3 - q_4 + j|)^{H(t)+H(t')} \\ &\quad \times \left\{ |t - t'|^{2\eta} \log(|q_1 - q_2 + j| |q_3 - q_4 + j|)^2 (1 + o(1)) \right\} \\ &= |t - t'|^{2\eta} \sum_{j \in \mathbb{Z}} \mathcal{O}(|j|^{2H(t)+2H(t')-4p} \log(1 + |j|)). \end{aligned}$$

The series converges if $p > \bar{H} + 1/4$. Consequently, $U(t, t') = \mathcal{O}(|t - t'|^{2\eta})$. □

Proof of Proposition 1(ii). Let $r = 2(1 + \lceil 1/\eta \rceil)$. It follows from Lemma 3 that

$$\mathbb{E}\left((2N\varepsilon_N)^{r/2} (V_{N,\varepsilon}(t, a) - V_{N,\varepsilon}(t', a))^r\right) = \mathcal{O}(|t - t'|^{r\eta}).$$

By Lemma 2 and since $r\eta > 1$, we obtain the convergence in distribution, for the topology of Skorohod, of $V_{N,\varepsilon}(\cdot, a)$ towards the Gaussian process \mathbb{G} with covariance function defined by (10). □

6.2. Identification of multifractional Brownian motion

For ease of presentation, let

$$\xi_{N,\varepsilon}(t, a) = L_{N,\varepsilon}(t, a) - X_M \alpha(t) = \log \left(N^{2H(t)} \frac{S_{N,\varepsilon}(t, a)}{C^2 \pi_{H(t)}^a(0)} \right).$$

Proof of Proposition 2(i). Note that

$$\hat{H}_{N,\varepsilon}(t, a, M) - H(t) = \frac{A^t}{2\|A\|^2} (\xi_{N,\varepsilon}(t, a^m))_{m=1,\dots,M}, \tag{39}$$

and that, almost surely

$$N^{2H(t)} \frac{S_{N,\varepsilon}(t, a^m)}{C^2 \pi_{H(t)}^{a^m}(0)} - 1 = V_{N,\varepsilon}(t, a^m) + \mathcal{O}(\varepsilon_N^\eta \log(N)). \tag{40}$$

From Proposition 1(i), we have that almost surely

$$N^{2H(t)} \frac{S_{N,\varepsilon}(t, a^m)}{C^2 \pi_{H(t)}^{a^m}(0)} \rightarrow 1;$$

therefore $\xi_{N,\varepsilon}(t, a^m) \xrightarrow{\text{a.s.}} 0$, which implies the almost sure convergence of $\hat{H}_{N,\varepsilon}(t, a, M)$ towards $H(t)$. Observe, moreover, that $\mathbb{E}(\xi_{N,\varepsilon}(t, a^m)) = \mathcal{O}(\varepsilon_N^\eta \log(N))$, so $\mathbb{E}(\hat{H}_{N,\varepsilon}(t, a, M) - H(t)) = \mathcal{O}(\varepsilon_N^\eta \log(N))$, then that $\text{var}(\xi_{N,\varepsilon}(t, a^m)) = \mathcal{O}(v_N(t)^{-1})$, and so $\text{var}(\hat{H}_{N,\varepsilon}(t, a, M)) = \mathcal{O}((N\varepsilon_N)^{-1})$. \square

Before proving the convergence in distribution, we examine the finite-dimensional convergence of our estimators.

Lemma 4. *Let a be a filter of order $p > \bar{H} + 1/4$, $M \geq 2$ an integer and assume that ε_N is of the form (7) with $\alpha \geq 1/(2\eta + 1)$ and $\beta < 0$. Let $d \geq 1$ and let $t_1, \dots, t_d \in [0, 1]$. Then, writing $B_{N,\hat{H}}(t) = \hat{H}_{N,\varepsilon}(t, a, M) - H(t)$, we have*

$$\left(\sqrt{v_N(t_1, a, M)} B_{N,\hat{H}}(t_1), \dots, \sqrt{v_N(t_d, a, M)} B_{N,\hat{H}}(t_d) \right)^\top \xrightarrow{d} (\mathbb{G}'(t_1), \dots, \mathbb{G}'(t_d))^\top, \tag{41}$$

where $(\mathbb{G}'(t_1), \dots, \mathbb{G}'(t_d))^\top$ is a centred Gaussian vector such that, for all $i, j \in \{1, \dots, d\}$,

$$\text{cov}(\mathbb{G}'(t_i), \mathbb{G}'(t_j)) = \frac{1}{4\|A\|^4} A^\top \Sigma \left(\frac{H(t_i)}{2} + \frac{H(t_j)}{2}, H(t_i), H(t_j) \right) A,$$

with $\Sigma(H_1, H_2, H_3)$ the $M \times M$ matrix whose (m, n) th entry is

$$(\Sigma(H_1, H_2, H_3))_{m,n} = 2 \sum_{j \in \mathbb{Z}} \frac{\pi_{H_1}^{a^m, a^n}(j)^2}{\pi_{H_2}^{a^m}(0) \pi_{H_3}^{a^n}(0)}, \quad m, n = 1, \dots, M,$$

with

$$\pi_H^{a^m, a^n}(j) = \sum_{q=0}^{m\ell} \sum_{q'=0}^{n\ell} a_q a'_{q'} |mq - nq' + j|^{2H},$$

and where A is the vector defined for $m = 1, \dots, M$ by $A_m = \log(m) - M^{-1} \sum_{m=1}^M \log(m)$.

Proof. Let us concentrate on the case $d = 1$. From (40), we have almost surely

$$\xi_{N,\varepsilon}(t, a^m) = \log \left(N^{2H(t)} \frac{S_{N,\varepsilon}(t, a^m)}{C^2 \pi_{H(t)}^{a^m}(0)} \right) = V_{N,\varepsilon}(t, a^m)(1 + o(1)) + \mathcal{O}(\varepsilon_N^\eta \log(N)). \tag{42}$$

Since $\alpha \geq 1/(2\eta + 1)$ and $\beta < 0$, it follows from (42) that $\sqrt{v_N(t)} \xi_{N,\varepsilon}(t, a^m)$ tends in distribution to the same limit as the random variable $\sqrt{v_N(t, a^m)} V_{N,\varepsilon}(t, a^m)$ and from Coeurjolly (2001, Proposition 3) and Lemma 2(i) we obtain

$$(\sqrt{v_N(t)} (\xi_{N,\varepsilon}(t, a^m))_{m=1, \dots, M})^T \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma(H(t), H(t), H(t))), \tag{43}$$

where $\Sigma(H(t), H(t), H(t))$ is the $M \times M$ matrix whose (m, n) th entry is

$$\begin{aligned} (\Sigma(H(t), H(t), H(t)))_{m,n} &= \lim_{N \rightarrow +\infty} \sqrt{v_N(t)} \sqrt{v_N(t)} \mathbb{E}(V_{N,\varepsilon}(t, a^m) V_{N,\varepsilon}(t, a^n)) \\ &= 2 \sum_{j \in \mathbb{Z}} \frac{\pi_{H(t)}^{a^m, a^n}(j)^2}{\pi_{H(t)}^{a^m}(0) \pi_{H(t)}^{a^n}(0)}, \quad m, n = 1, \dots, M, \end{aligned}$$

with

$$\pi_H^{a^m, a^n}(j) = \sum_{q=0}^{m\ell} \sum_{q'=0}^{n\ell} a_q a'_{q'} |mq - nq' + j|^{2H}.$$

The results (39) and (43) ensure that $\sqrt{v_N(t)} (\hat{H}_{N,\varepsilon}(t, a) - H(t))$ is asymptotically Gaussian.

The case $d > 1$ is easily deduced using Lemma 2(ii), which implies the finite-dimensional Gaussian convergence of $\xi_{N,\varepsilon}(\cdot, a)$. We end with the following computation for $t, t' \in [0, 1]$:

$$\begin{aligned} &\text{cov}(\sqrt{v_N(t)} \{ \hat{H}_{N,\varepsilon}(t, a, M) - H(t) \}, \sqrt{v_N(t')} \{ \hat{H}_{N,\varepsilon}(t', a, M) - H(t') \}) \\ &\sim \frac{A^T}{4 \|A\|^4} \mathbb{E}(\sqrt{v_N(t)} \{ V_{N,\varepsilon}(t, a^1), \dots, V_{N,\varepsilon}(t, a^M) \}^T \sqrt{v_N(t')} \{ V_{N,\varepsilon}(t', a^1), \dots, V_{N,\varepsilon}(t', a^M) \}) A \\ &\rightarrow \frac{1}{4 \|A\|^4} A^T \Sigma \left(\frac{H(t)}{2} + \frac{H(t')}{2}, H(t), H(t') \right) A, \quad \text{as } N \rightarrow +\infty. \end{aligned}$$

Proof of Proposition 2(ii). Let $r \geq 4$ be an integer, and let $t, t' \in]0, 1[$. By (42), we have almost surely

$$(\xi_{N,\varepsilon}(t, a^m) - \xi_{N,\varepsilon}(t', a^m)) = (V_{N,\varepsilon}(t, a^m) - V_{N,\varepsilon}(t', a^m))(1 + o(1)) + \mathcal{O}(\varepsilon_N^\eta \log(N)). \tag{44}$$

Thus, if ε_N is such that $\alpha \geq 1/(2\eta + 1)$ and $\beta < 0$, we obtain

$$\mathbb{E}((2N\varepsilon_N)^{r/2} \{ \xi_{N,\varepsilon}(t, a^m) - \xi_{N,\varepsilon}(t', a^m) \}^r) = \mathcal{O}(\mathbb{E}((2N\varepsilon_N)^{r/2} \{ V_{N,\varepsilon}(t, a^m) - V_{N,\varepsilon}(t', a^m) \}^r)).$$

Choosing r large enough, we obtain, using Lemma 3, the convergence in distribution on the range $]0, 1[$ of $\sqrt{2N\varepsilon_N}\xi_{N,\varepsilon}(\cdot, a^m)$, and then of $\sqrt{2N\varepsilon_N}(\hat{H}_{N,\varepsilon}(\cdot, a, M) - H(\cdot))$ using (39). \square

Proof of Proposition 3. (i) In fact it is sufficient to prove that, for all $t \in [0, 1]$, almost surely

$$\sup_{\varepsilon_N \in E} |\hat{H}_{N,\varepsilon}(t)| \rightarrow 0, \quad \text{as } N \rightarrow +\infty.$$

Let $\lambda > 0$ and $k \geq 1$ be an integer. We have, from (19) and from Chebyshev’s inequality,

$$\begin{aligned} \mathbb{P}\left(\sup_{\varepsilon_N \in E} \left| \hat{H}_{N,\varepsilon}(t) - H(t) \right| \geq \lambda\right) &\leq n^\rho \sup_{\varepsilon_N \in E} \mathbb{P}\left(\left| \hat{H}_{N,\varepsilon}(t) - H(t) \right| \geq \lambda\right) \\ &\leq n^\rho \frac{1}{\rho^{2k}} \sup_{\varepsilon_N \in E} \mathbb{E}\left(\left(\hat{H}_{N,\varepsilon}(t) - H(t)\right)^{2k}\right). \end{aligned} \tag{45}$$

Using (44) and the fact that $\alpha \geq 1/(2\eta + 1)$ and $\beta < 0$, we have

$$\mathbb{E}\left(\left(\hat{H}_{N,\varepsilon}(t) - H(t)\right)^{2k}\right) = \mathcal{O}\left(\frac{1}{(N\varepsilon_N)^k}\right).$$

Choosing k sufficiently large leads to the summability of $\sum_n \mathbb{P}(\sup_{\varepsilon_N \in E} |\hat{H}_{N,\varepsilon}(t) - H(t)| \geq \lambda)$ and to the result, using the Borel–Cantelli lemma.

(ii) Since $\hat{\varepsilon}_N^* = \operatorname{argmin}_{\varepsilon_N \in E} \hat{R}'_N(\varepsilon_N)$ and since $\varepsilon_N^* \in E$, we have almost surely $\hat{R}'_N(\hat{\varepsilon}_N^*) \leq \hat{R}'_N(\varepsilon_N^*)$ and so $\mathbb{E}(\hat{R}'_N(\hat{\varepsilon}_N^*)) \leq \mathbb{E}(\hat{R}'_N(\varepsilon_N^*))$. Now

$$\begin{aligned} \frac{|\mathbb{E}(R_N(\hat{\varepsilon}_N^*)) - \mathbb{E}(R_N(\varepsilon_N^*))|}{\mathbb{E}(R_N(\varepsilon_N^*))} &= \frac{\mathbb{E}(R_N(\hat{\varepsilon}_N^*)) - \mathbb{E}(R_N(\varepsilon_N^*))}{\mathbb{E}(R_N(\varepsilon_N^*))} = \frac{\mathbb{E}(R'_N(\hat{\varepsilon}_N^*)) - \mathbb{E}(R'_N(\varepsilon_N^*))}{\mathbb{E}(R_N(\varepsilon_N^*))} \\ &\leq \frac{\mathbb{E}(R'_N(\hat{\varepsilon}_N^*)) - \mathbb{E}(R'_N(\varepsilon_N^*)) + \mathbb{E}(\hat{R}'_N(\varepsilon_N^*)) - \mathbb{E}(\hat{R}'_N(\hat{\varepsilon}_N^*))}{\mathbb{E}(R_N(\varepsilon_N^*))}, \end{aligned}$$

since $\mathbb{E}(\hat{R}'_N(\varepsilon_N^*)) - \mathbb{E}(\hat{R}'_N(\hat{\varepsilon}_N^*)) \geq 0$. Finally, we have

$$\frac{|\mathbb{E}(R_N(\hat{\varepsilon}_N^*)) - \mathbb{E}(R_N(\varepsilon_N^*))|}{\mathbb{E}(R_N(\varepsilon_N^*))} \leq 2 \sup_{\varepsilon_N \in E} \frac{|\mathbb{E}(R'_N(\varepsilon_N) - \hat{R}'_N(\varepsilon_N))|}{\mathbb{E}(R_N(\varepsilon_N^*))}. \tag{46}$$

For $\varepsilon_N \in E$, there exists $N_0 \in \mathbb{N}$ such that, for all $N \geq N_0$,

$$\begin{aligned} \mathbb{E}(R'_N(\varepsilon_N) - \hat{R}'_N(\varepsilon_N)) &= \frac{2}{N} \sum_{i=0}^{N-1} \mathbb{E}\left(\left(\tilde{H}_{N,\varepsilon^2}\left(\frac{i}{N}\right) - H\left(\frac{i}{N}\right)\right) \hat{H}_{N,\varepsilon}\left(\frac{i}{N}\right)\right) \\ &\leq \frac{4}{N} \sum_{i=0}^{N-1} \mathbb{E}\left(\tilde{H}_{N,\varepsilon^2}\left(\frac{i}{N}\right) - H\left(\frac{i}{N}\right)\right). \end{aligned} \tag{47}$$

Moreover,

$$\begin{aligned} \mathbb{E}\left(\tilde{H}_{N,\varepsilon^2}\left(\frac{i}{N}\right) - H\left(\frac{i}{N}\right)\right) &= \frac{1}{\#\mathcal{V}_{N,\varepsilon^2}(i/N)} \sum_{j \in \mathcal{V}_{N,\varepsilon^2}(i/N)} \mathbb{E}\left(\hat{H}_{N,\varepsilon^2}\left(\frac{j}{N}\right) - H\left(\frac{j}{N}\right)\right) \\ &\quad + \frac{1}{\#\mathcal{V}_{N,\varepsilon^2}(i/N)} \sum_{j \in \mathcal{V}_{N,\varepsilon^2}(i/N)} \left(H\left(\frac{j}{N}\right) - H\left(\frac{i}{N}\right)\right) \\ &= \mathcal{O}\left(\varepsilon_N^{2\eta} \log(N)\right) + \mathcal{O}\left(\varepsilon_N^{2\eta}\right) = \mathcal{O}\left(\varepsilon_N^{2\eta} \log(N)\right). \end{aligned} \tag{48}$$

Combining (47) and (48), we obtain

$$\frac{\mathbb{E}\left(R'_N(\varepsilon_N) - \hat{R}'_N(\varepsilon_N)\right)}{\mathbb{E}\left(R_N(\varepsilon_N^*)\right)} = \mathcal{O}\left(N^{2\eta/(2\eta+1)-2\alpha\eta} \log(N)^{1+2\beta\eta-2/(2\eta+1)}\right).$$

Since $0 < \eta \leq 1$, the proof is achieved using the definition of the set E (see (19), and (46)). □

Acknowledgement

This work was financially supported by a grant from IMAG project AMOA. The author thanks Professor Anestis Antoniadis for helpful comments and the editor and a referee for their suggestions and guidance.

References

- Benassi, A., Cohen, S. and Istas, J. (1998) Identifying the multifractional function of a Gaussian process. *Statist. Probab. Lett.*, **39**, 337–345.
- Breuer, P. and Major, P. (1983) Central limit theorems for non-linear functionals of Gaussian fields. *J. Multivariate Anal.*, **13**, 425–441.
- Coeurjolly, J.-F. (2000) Inférence statistique pour les mouvements browniens fractionnaires et multifractionnaires. Doctoral thesis, Université de Grenoble.
- Coeurjolly, J.-F. (2001) Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statist. Inference Stochastic Process.*, **4**, 199–227.
- Coeurjolly, J.-F. and Istas, J. (2001). Cramér–Rao bounds for fractional Brownian motions. *Statist. Probab. Lett.*, **53**, 435–448.
- Cohen, S. (1999) From self-similarity to local self-similarity: the estimation problem. In M. Dekking, J. Lévy Véhel, E. Lutton, and C. Tricot (eds), *Fractals: Theory and Applications in Engineering*, pp. 3–16. London: Springer-Verlag.
- Collins, J.J. and De Luca, C.J. (1994) Upright, correlated random walks: A statistical-biomechanics approach to the human postural control system. *Chaos*, **5**(1), 57–63.
- Doob, J.L. (1953) *Stochastic Processes*. New York: Wiley.
- Frisch, U. (1997) *Turbulence. The Legacy of A.N. Kolmogorov*. Cambridge: Cambridge University Press.

- Istas, J. and Lang, G. (1997) Quadratic variations and estimation of the Hölder index of a Gaussian process. *Ann. Inst. H. Poincaré Probab. Statist.*, **33**, 407–436.
- Jaffard, S. (1990) Propriétés des matrices ‘bien localisées’ près de leur diagonale et quelques applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **7**, 461–476.
- Kent, J.T. and Wood, A.T.A. (1997) Estimating the fractal dimension of a locally self-similar Gaussian process using increments. *J. Roy. Statist. Soc. Ser. B*, **59**, 679–700.
- Mandelbrot, B. and Van Ness, J. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, **10**, 422–437.
- Peltier, R.-F. and Lévy Véhel, J. (1995) Multifractional Brownian motion: definition and preliminary results. INRIA Research Report No. 2645.
- Taqq, M.S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **40**, 203–238.
- Von Bahr, B. and Esseen, C.G. (1965) Inequalities for the r th moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.*, **36**, 299–303.

Received April 2003 and revised November 2004