Uniformly and strongly consistent estimation for the Hurst function of a Linear Multifractional Stable Motion

Antoine Ayache
and
Julien Hamonier

Université Lille 1 - Laboratoire Paul Painlevé
and
Université Lille 2 - Laboratoire de Biomathématiques

Organization of the talk

Fractional Brownian Motion

Fractional Brownian Motion (FBM) of Hurst parameter $H \in (0,1)$ denoted by $\{B_H(t) \,:\, t \in [0,1]\}$ and defined as

$$B_H(t) = \int_{\mathbb{R}} \Big((t-s)_+^{H-1/2} - (-s)_+^{H-1/2} \Big) \mathrm{Z}_2\,(ds)\,, \qquad (1.1)$$

is a quite classical random model for real-life fractal signals. Observe that:

- for all $(x,\kappa) \in \mathbb{R}^2$, one has $(x)_+^{\kappa} = x^{\kappa}$ when $x > 0$ and $(x)_+^{\kappa} = 0$ else;
- $\mathrm{Z}_2\,(ds)$ denotes an independently scattered Gaussian random measure on $\mathbb{R}$, with Lebesgue measure as its control measure. That is $\int_{\mathbb{R}} (\,\cdot\,) \mathrm{Z}_2\,(ds)$ is a usual Wiener integral.

FBM is an H-self-similar centered Gaussian process with stationary increments and a covariance function given for all $(t_1, t_2) \in [0,1]^2$ by

$$\mathbb{E}\big(B_H(t_1)B_H(t_2)\big) = 2^{-1}c(H)\Big(t_1^{2H} + t_2^{2H} - |t_1 - t_2|^{2H}\Big), \qquad (1.2)$$

where $c(H) = \mathbb{E}\big(B_H(1)\big)^2$.

Although this model offers the advantage of simplicity, it lacks flexibility and thus does not always fit with reality. An important limitation is that local fractal properties of FBM sample paths are not really allowed to evolve over time: roughness remains almost the same all along sample paths:

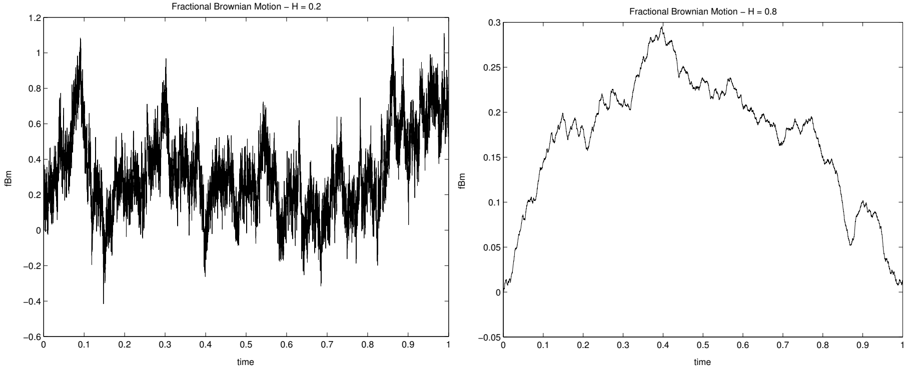

Figure : Simulation of an FBM sample path with $H = 0.2$ (left) and with $H = 0.8$ (right)

This limitation is mainly due to the constancy over time of H the Hurst parameter governing FBM.

Multifractional Brownian Motion

In order to overcome this drawback, various multifractional stochastic processes have been introduced since the 90's and studied by many authors: Angulo, Anh, Ayache, Balança, Bardet, Benassi, Bertrand, Bianchi, Biermé, Boufoussi, Clausel, Coeurjolly, Cohen, Dozzi, Falconer, Guerbaz, Le Guével, Hamonier, Herbin, Istas, Jaffard, Lacaux, Leonenko, Lévy Véhel, Lifshits, Meerschaert, Pantanella, Peltier, Peng, Pianese, Ruiz-Medina, Roux, Surgailis, Stoev, Taqqu, Vedel, Wu, Xiao, ...

→ Roughly speaking, the main idea behind this new class of processes is that Hurst parameter H becomes a function $H(t)$ depending on the time variable t.

→ The paradigmatic example of such processes is the centered Gaussian Multifractional Brownian Motion (MBM) $\{B_{H(t)}(t) : t \in [0,1]\}$, having a covariance function given for all $(t_1, t_2) \in [0,1]^2$ by

$$\begin{aligned} &\mathbb{E}\big(B_{H(t_1)}(t_1)B_{H(t_2)}(t_2)\big) \qquad (1.3)\\ &= c\big(H(t_1), H(t_2)\big)\Big(t_1^{H(t_1)+H(t_2)} + t_2^{H(t_1)+H(t_2)} - |t_1 - t_2|^{H(t_1)+H(t_2)}\Big). \end{aligned}$$

⇒ Although the factor $c\big(H(t_1), H(t_2)\big)$ can usually be neglected, MBM as well as the other multifractional processes have complex dependence structures.

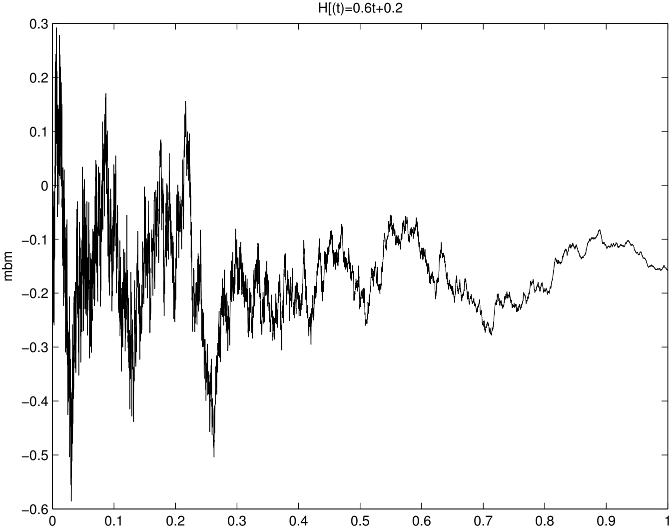

Figure : Simulation of an MBM sample path with $H(t) = 0.6t + 0.2$ for all $t \in [0, 1]$

Organization of the talk

Let us now present in the case of MBM the classical strategy for statistical estimation of $H(t_0)$, the value of the Hurst function at an arbitrary fixed time $t_0 \in [0,1]$. This strategy was introduced in (Benassi, Cohen and Istas 1998).
$\rightarrow$ The observations consist in a sample $\big\{B_{H(k/N)}(k/N)\,:\,k\in\{0,\ldots,N\}\big\}$ of an MBM sample path, where N is an integer large enough. Usually it is assumed that

$$H(\cdot) \text{ is a } \rho_H\text{-Hölder function such that } 1 \geq \rho_H > \sup_{t\in[0,1]} H(t). \qquad (*)$$

$\rightarrow$ The estimator of $H(t_0)$ is built through L-th order discrete variations of the $B_{H(k/N)}(k/N)$'s, where the integer $L \geq 2$ is arbitrary and fixed.
$\rightarrow$ Let us define those variations. For each $q \in \{0,\ldots,L\}$, one sets $a_q = (-1)^{L-q}\binom{L}{q} = (-1)^{L-q}\frac{L!}{q!\,(L-q)!}$; observe that for all $m \in \{0,\ldots,L-1\}$, $\sum_{q=0}^{L} q^m a_q = 0$ and $\sum_{q=0}^{L} q^L a_q \neq 0$. For any $k \in \{0,\ldots,N-L\}$, the L-th order discrete variation of MBM at k/N is denoted by $d_{N,k}^{\mathrm{MBM}}$ and defined as

$$d_{N,k}^{\mathrm{MBM}} = \sum_{q=0}^{L} a_q B_{H((k+q)/N)}((k+q)/N) \simeq \sum_{q=0}^{L} a_q B_{H(k/N)}((k+q)/N). \qquad (2.4)$$

$\rightarrow$ From now on $d_{N,k}^{\mathrm{MBM}}$ is identified with $\sum_{q=0}^{L} a_q B_{H(k/N)}((k+q)/N)$. In doing so, the approximation error is, almost surely uniformly in k, of the same order as $N^{-\rho_H}$, which can be considered to be negligible thanks to Assumption $(*)$.

Thus, one has for all $(k_1, k_2) \in \{0, \ldots, N-L\}^2$,

$$\left|\mathrm{cov}\left(d_{N,k_1}^{\mathrm{MBM}}, d_{N,k_2}^{\mathrm{MBM}}\right)\right|$$

$$\asymp N^{-H(k_1/N)-H(k_2/N)} \left|\sum_{0\le q_1,q_2\le L} a_{q_1} a_{q_2} \left|k_1 - k_2 + q_1 - q_2\right|^{H(k_1/N)+H(k_2/N)}\right|$$

(applying Taylor formula)

$$\asymp N^{-H(k_1/N)-H(k_2/N)} \left(1 + \left|k_1 - k_2\right|\right)^{H(k_1/N)+H(k_2/N)-2L}; \qquad (2.5)$$

in other words

$$N^{H(k_1/N)+H(k_2/N)} \left(1 + \left|k_1 - k_2\right|\right)^{-H(k_1/N)-H(k_2/N)+2L} \left|\mathrm{cov}\left(d_{N,k_1}^{\mathrm{MBM}}, d_{N,k_2}^{\mathrm{MBM}}\right)\right|$$

is bounded from above and from below by positive and finite constants non depending on N, k_1 and k_2. Notice that (2.5) implies that $\|d_{N,k}^{\mathrm{MBM}}\|_2$ the standard deviation of the centered Gaussian random variable $d_{N,k}^{\mathrm{MBM}}$ satisfies

$$\|d_{N,k}^{\mathrm{MBM}}\|_2 \asymp N^{-H(k/N)}. \qquad (2.6)$$

In view of (2.6), it turns out that the quantity $H(t_0)$ is mainly connected with the $d_{N,k}^{\mathrm{MBM}}$'s located "near to" t_0.

In order to clearly define the notion of "near to". For any non-degenerate compact interval $I \subseteq [0,1]$, let $\nu_N(I)$ be the set

$$\nu_N(I) = \big\{k \in \{0,\ldots,N-L\} : k/N \in I\big\}; \tag{2.7}$$

it is rather denoted by $\nu_N(t_0,\gamma)$ when $I = [0,1] \cap [t_0 - N^{-\gamma}, t_0 + N^{-\gamma}]$, where $\gamma \in (0,1)$ is a parameter. Observe that the set $\nu_N(t_0,\gamma)$ is non-empty as soon as $N \geq (L+1)^{1/(1-\gamma)}$ and that $\|d_{N,k}^{\mathrm{MBM}}\|_2 \asymp N^{-H(t_0)}$ if $k \in \nu_N(t_0,\gamma)$.

Theorem 2.1 ((Benassi, Cohen and Istas 1998) and (Coeurjolly 2005 & 2006))

For any $N \geq (L+1)^{1/(1-\gamma)}$, let $V_N(t_0,\gamma)$ be the empirical mean defined as

$$V_N(t_0,\gamma) = |\nu_N(t_0,\gamma)|^{-1} \sum_{k \in \nu_N(t_0,\gamma)} \left| d_{N,k}^{MBM} \right|^2, \tag{2.8}$$

where $|\nu_N(t_0,\gamma)|$ is the cardinality of $\nu_N(t_0,\gamma)$. Then

$$\widehat{H}_N(t_0,\gamma) = 2^{-1} \log_2 \left(\frac{V_N(t_0,\gamma)}{V_{2N}(t_0,\gamma)} \right), \tag{2.9}$$

is an almost surely convergent estimator of $H(t_0)$.

Scketch of the Proof: It is enough to show that "a strong law of large numbers" holds for the empirical mean $V_N(t_0,\gamma)$, more precisely:

$$\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)} \xrightarrow[n\to+\infty]{a.s.} 1. \tag{2.10}$$

(2.10) will result from Borel-Cantelli Lemma. One has for any $\eta > 0$,

$$\mathbb{P}\Big(\Big|\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)} - 1\Big| > \eta\Big) = \mathbb{P}\Big(\Big|V_N(t_0,\gamma) - \mathbb{E}\big(V_N(t_0,\gamma)\big)\Big| > \eta\,\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big). \tag{2.11}$$

Next applying Bienaymé-Tchebychev inequality one gets

$$\mathbb{P}\Big(\Big|\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)} - 1\Big| > \eta\Big) \le \eta^{-2}\,\frac{\mathrm{Var}\big(V_N(t_0,\gamma)\big)}{\big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\big)^2}\,. \tag{2.12}$$

Let us now try to show that

$$\sum_{N\ge(L+1)^{1/(1-\gamma)}}^{+\infty} \frac{\mathrm{Var}\big(V_N(t_0,\gamma)\big)}{\big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\big)^2} < +\infty. \tag{2.13}$$

One has

$$\big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\big)^2 = |\nu_N(t_0,\gamma)|^{-2}\Big(\sum_{k\in\nu_N(t_0,\gamma)} \|d_{N,k}^{\mathrm{MBM}}\|_2^2\Big)^2 \asymp N^{-4H(t_0)}. \qquad (2.14)$$

On the other hand, for any centered 2-D Gaussian vector (Z_1, Z_2),

$$\operatorname{cov}\big(Z_1^2, Z_2^2\big) = 2\big(\operatorname{cov}(Z_1, Z_2)\big)^2; \qquad (2.15)$$

therefore, one gets

$$\begin{aligned}
\operatorname{Var}\big(V_N(t_0,\gamma)\big) &= |\nu_N(t_0,\gamma)|^{-2} \sum_{k_1,k_2\in\nu_N(t_0,\gamma)} \operatorname{cov}\big((d_{N,k_1}^{\mathrm{MBM}})^2, (d_{N,k_2}^{\mathrm{MBM}})^2\big) \\
&= 2\,|\nu_N(t_0,\gamma)|^{-2} \sum_{k_1,k_2\in\nu_N(t_0,\gamma)} \big(\operatorname{cov}(d_{N,k_1}^{\mathrm{MBM}}, d_{N,k_2}^{\mathrm{MBM}})\big)^2 \\
&\asymp |\nu_N(t_0,\gamma)|^{-2}\, N^{-4H(t_0)} \sum_{k_1,k_2\in\nu_N(t_0,\gamma)} \Big(1+\big|k_1-k_2\big|\Big)^{4(H(t_0)-L)} \\
&\asymp |\nu_N(t_0,\gamma)|^{-1}\, N^{-4H(t_0)}, \qquad (2.16)
\end{aligned}$$

since for all fixed $k_1 \in \nu_N(t_0,\gamma)$,

$\sum_{k_2\in\nu_N(t_0,\gamma)} \Big(1+\big|k_1-k_2\big|\Big)^{4(H(t_0)-L)} \le \sum_{n\in\mathbb{Z}} \big(1+|n|\big)^{4(H(t_0)-L)} < +\infty.$

Thus, it follows that

$$\frac{\mathrm{Var}\big(V_N(t_0,\gamma)\big)}{\big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\big)^2} \asymp |\nu_N(t_0,\gamma)|^{-1} \asymp N^{\gamma-1}. \tag{2.17}$$

Unfortunately, the fact that $\gamma \in (0,1)$ implies that

$$\sum_{N \geq (L+1)^{1/(1-\gamma)}}^{+\infty} \frac{\mathrm{Var}\big(V_N(t_0,\gamma)\big)}{\big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\big)^2} = +\infty. \tag{2.18}$$

More effort is necessary for obtaining the theorem! Rather than using

$$\mathbb{P}\left(\left|\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)} - 1\right| > \eta\right) \leq \eta^{-2} \frac{\mathrm{Var}\big(V_N(t_0,\gamma)\big)}{\big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\big)^2}, \tag{2.19}$$

one needs to use

$$\mathbb{P}\left(\left|\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)} - 1\right| > \eta\right) = \mathbb{P}\left(\left|\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)} - 1\right|^4 > \eta^4\right)$$

$$\leq \eta^{-4} \frac{\mathbb{E}\left(\left|V_N(t_0,\gamma) - \mathbb{E}\big(V_N(t_0,\gamma)\big)\right|^4\right)}{\left(\mathbb{E}\big(V_N(t_0,\gamma)\big)\right)^4} \quad \text{(Markov inequality)} \tag{2.20}$$

An appropriate upper bound for $\mathbb{E}\Big(\Big|V_N(t_0,\gamma)-\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big|^4\Big)$ can be derived from the following lemma:

Lemma 2.1

There exists a constant $c>0$, such that for each positive integer m, and for any centered non-degenerate Gaussian vector $(Z_1,\ldots,Z_m)$, one has

$$\mathbb{E}\Big(\Big|\sum_{k=1}^{m}\big(Z_k^2-\mathbb{E}(Z_k^2)\big)\Big|^4\Big)\le c\Big(\mathrm{Var}\Big(\sum_{k=1}^{m}Z_k^2\Big)\Big)^2. \qquad (2.21)$$

Thus, using (2.21) and (2.16), one gets

$$\begin{aligned}\mathbb{E}\Big(\Big|V_N(t_0,\gamma)-\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big|^4\Big) &\le c\Big(\mathrm{Var}\big(V_N(t_0,\gamma)\big)\Big)^2\\ &\le c_1|\nu_N(t_0,\gamma)|^{-2}\,N^{-8H(t_0)}. \qquad (2.22)\end{aligned}$$

Next, (2.22) and (2.14) imply that

$$\frac{\mathbb{E}\Big(\Big|V_N(t_0,\gamma)-\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big|^4\Big)}{\Big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big)^4} \le c_2|\nu_N(t_0,\gamma)|^{-2} \asymp N^{2(\gamma-1)}. \tag{2.23}$$

Hence, when $\gamma \in (0,1/2)$, one has, for all $\eta > 0$,

$$\sum_{N\ge(L+1)^{1/(1-\gamma)}}^{+\infty} \mathbb{P}\Big(\Big|\frac{V_N(t_0,\gamma)}{\mathbb{E}\big(V_N(t_0,\gamma)\big)}-1\Big|>\eta\Big)$$

$$\le c_3\eta^{-4} \sum_{N\ge(L+1)^{1/(1-\gamma)}}^{+\infty} \frac{\mathbb{E}\Big(\Big|V_N(t_0,\gamma)-\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big|^4\Big)}{\Big(\mathbb{E}\big(V_N(t_0,\gamma)\big)\Big)^4} < +\infty,$$

which ends the proof of the theorem.

Organization of the talk

Linear Multifractional Stable Motion

Linear Multifractional Stable Motion (LMSM) denoted by $\{Y(t) : t \in [0,1]\}$ is a quite natural extension of MBM to the setting of heavy-tailed stable distributions. It was introduced in (Stoev and Taqqu 2004) and it is defined as

$$Y(t) = \int_{\mathbb{R}} \Big((t-s)_+^{H(t)-1/\alpha} - (-s)_+^{H(t)-1/\alpha} \Big) Z_\alpha(ds), \tag{3.24}$$

where $Z_\alpha(ds)$ is an independently scattered symmetric α-stable ($\mathcal{S}\alpha\mathcal{S}$) random measure on $\mathbb{R}$, with Lebesgue measure as its control measure (see the book Samorodnitsky and Taqqu 1994). We assume that $\alpha \in (1,2)$.

$\rightarrow$ The stochastic integral $\mathbb{I}(f) = \int_{\mathbb{R}} f(s) Z_\alpha(ds)$ is defined for any $f \in L^\alpha(\mathbb{R})$.

Recall that $\mathbb{I}(f)$ is a real-valued $\mathcal{S}\alpha\mathcal{S}$ random variable i.e. $\mathbb{E}(e^{i\xi\,\mathbb{I}(f)}) = e^{-\sigma^\alpha\,|\xi|^\alpha}$, for all $\xi \in \mathbb{R}$. The scale parameter σ is denoted by $\|\mathbb{I}(f)\|_\alpha$ and given by

$$\|\mathbb{I}(f)\|_\alpha = \Big(\int_{\mathbb{R}} |f(s)|^\alpha \, ds \Big)^{1/\alpha} = \|f\|_{L^\alpha(\mathbb{R})}. \tag{3.25}$$

Also, recall that for any $\gamma > 0$, one has $\mathbb{E}(|\mathbb{I}(f)|^\gamma) < +\infty$ iff $\gamma < \alpha$, moreover

$$\mathbb{E}(|\mathbb{I}(f)|^\gamma) = c(\gamma) \|\mathbb{I}(f)\|_\alpha^\gamma, \tag{3.26}$$

where the constant $c(\gamma)$ only depends on γ.

$\rightarrow$ Let us emphasize that the independently scattered property of $\mathrm{Z}_\alpha(ds)$ will play a crucial role in the sequel; it means that: for each positive integer n and all functions $f_1, \ldots, f_n$ belonging to $L^\alpha(\mathbb{R})$, the coordinates of the $\mathcal{S}\alpha\mathcal{S}$ random vector $(\mathbb{I}(f_1), \ldots, \mathbb{I}(f_n))$ are independent random variables as soon as the supports of $f_1, \ldots, f_n$ are disjoint up to Lebesgue-negligible sets.

A natural question: is it possible to extend Theorem 2.1, on the estimation of $H(t_0)$, to the setting of the LMSM $\{Y(t) : t \in [0,1]\}$?
More precisely, one assumes that $\beta \in (0, 1/4]$, then one sets

$$V_N^\beta(t_0,\gamma) = |\nu_N(t_0,\gamma)|^{-1} \sum_{k \in \nu_N(t_0,\gamma)} \left|d_{N,k}\right|^\beta, \tag{3.27}$$

and

$$\widehat{H}_N^\beta(t_0,\gamma) = \beta^{-1} \log_2 \left(\frac{V_N^\beta(t_0,\gamma)}{V_{2N}^\beta(t_0,\gamma)} \right), \tag{3.28}$$

where $d_{N,k} = \sum_{q=0}^{L} a_q Y((k+q)/N)$ is the L-th order discrete variation of the LMSM at k/N.
Is it true that $\widehat{H}_N^\beta(t_0,\gamma)$ converges almost surely to $H(t_0)$ when N goes to $+\infty$?

From now on, in addition to the ρ_H-Hölder condition $(*)$ already imposed to $H(\cdot)$, one assumes that the latter function is with values in a compact interval $[\underline{H}, \overline{H}] \subset (1/\alpha, \rho_H)$. Then, for each $k \in \{0, \dots, N-L\}$,

$$d_{N,k} \simeq N^{-(H(k/N)-1/\alpha)} \int_{-\infty}^{N^{-1}(k+L)} \Phi_\alpha\big(Ns-k, H(k/N)\big) \mathrm{Z}_\alpha(ds)\,, \tag{3.29}$$

where $\Phi_\alpha(u,v) = \sum_{l=0}^{L} a_l (l-u)_+^{v-1/\alpha}$, for all $(u,v) \in \mathbb{R} \times (1/\alpha, 1)$. Notice that, similarly to the Gaussian case $\alpha = 2$, the approximation error in (3.29), is, almost surely uniformly in k, of the same order as $N^{-\rho_H}$, which can be considered to be negligible thanks to Assumption $(*)$.

→ The continuous function Φ_α has the following useful localization properties:

$$\mathrm{supp}\big(\Phi_\alpha(\cdot, v)\big) =]-\infty, L], \text{ for all fixed } v \in (1/\alpha, 1) \tag{3.30}$$

and

$$\sup\Big\{ \big(1+|u|\big)^{L+1/\alpha-v} \big|\Phi_\alpha(u,v)\big| \,:\, (u,v) \in]-\infty, L] \times (1/\alpha, 1) \Big\} < +\infty. \tag{3.31}$$

Let us now try to rewrite the proof of Theorem 2.1 in the case of LMSM. First, notice that one has $\mathbb{E}\big(\big(V_N^\beta(t_0,\gamma)\big)^4\big) < +\infty$ since $\beta \in (0,1/4]$. Let us try to make use of Borel-Cantelli Lemma in order to show that

$$\frac{V_N^\beta(t_0,\gamma)}{\mathbb{E}\big(V_N^\beta(t_0,\gamma)\big)} \xrightarrow[n\to+\infty]{a.s.} 1. \tag{3.32}$$

Similarly to the Gaussian case, the inequality

$$\mathbb{P}\Bigg(\Bigg|\frac{V_N^\beta(t_0,\gamma)}{\mathbb{E}\big(V_N^\beta(t_0,\gamma)\big)} - 1\Bigg| > \eta\Bigg) \le \eta^{-4}\frac{\mathbb{E}\Big(\Big|V_N^\beta(t_0,\gamma) - \mathbb{E}\big(V_N^\beta(t_0,\gamma)\big)\Big|^4\Big)}{\Big(\mathbb{E}\big(V_N^\beta(t_0,\gamma)\big)\Big)^4}, \tag{3.33}$$

holds for any $\eta > 0$. The integral representation of $d_{N,k}$, allows to show that

$$\|d_{N,k}\|_\alpha \asymp N^{-H(k/N)}, \tag{3.34}$$

which in turns allows to obtain that

$$\mathbb{E}\big(V_N^\beta(t_0,\gamma)\big) \asymp N^{-\beta H(t_0)}. \tag{3.35}$$

Yet, in contrast with the Gaussian case, it is not clear how one can do in order to obtain a convenient upper bound for

$$\mathbb{E}\Big(\Big|V_N^\beta(t_0,\gamma)-\mathbb{E}\big(V_N^\beta(t_0,\gamma)\big)\Big|^4\Big).$$

Indeed, this would require to look for an appropriate extension of Lemma 2.1 to the setting of heavy-tailed $\mathcal{S}\alpha\mathcal{S}$ distributions, as for instance to provide a positive answer to the following question.

A question: is it true that there exists a constant $c>0$, such that for each positive integer m and for any non-degenerate $\mathcal{S}\alpha\mathcal{S}$ random vector $(S_1,\ldots,S_m)$, one has

$$\mathbb{E}\Big(\Big|\sum_{k=1}^m\big(S_k^\beta-\mathbb{E}(S_k^\beta)\big)\Big|^4\Big)\le c\Big(\mathrm{Var}\Big(\sum_{k=1}^m S_k^\beta\Big)\Big)^2\ \ ? \tag{3.36}$$

We have not found such an extension of Lemma 2.1 in the literature. Let us now explain how to overcome this difficulty.

An important decomposition of $d_{N,k}$

Assume that $\delta \in (0,1)$ is fixed and that $e_N = e_N(\delta)$ is the positive integer

$$e_N = \left[N^{\delta} \right], \tag{3.37}$$

where $[\,\cdot\,]$ denotes the integer part function. One has

$$d_{N,k} \simeq N^{-(H(k/N)-1/\alpha)} \int_{-\infty}^{N^{-1}(k+L)} \Phi_\alpha\big(Ns-k, H(k/N)\big) \mathrm{Z}_\alpha(ds) = d_{N,k}^{1,\delta} + d_{N,k}^{2,\delta}, \tag{3.38}$$

where

$$d_{N,k}^{1,\delta} = N^{-(H(k/N)-1/\alpha)} \int_{N^{-1}(k-e_N+L)}^{N^{-1}(k+L)} \Phi_\alpha\big(Ns-k, H(k/N)\big) \mathrm{Z}_\alpha(ds) \tag{3.39}$$

and

$$d_{N,k}^{2,\delta} = N^{-(H(k/N)-1/\alpha)} \int_{-\infty}^{N^{-1}(k-e_N+L)} \Phi_\alpha\big(Ns-k, H(k/N)\big) \mathrm{Z}_\alpha(ds)\,. \tag{3.40}$$

This decomposition of $d_{N,k}$ is important for two reasons

→ First reason: Roughly speaking, one has $d_{N,k}^{1,\delta} \simeq d_{N,k}$.

Heuristic explanation: Using the integral representations of $d_{N,k}^{1,\delta}$ and $d_{N,k}^{2,\delta}$, more particularly the localization properties of the function Φ_α in them, one can show that

$$\|d_{N,k}^{1,\delta}\|_\alpha \asymp N^{-H(k/N)} \asymp \|d_{N,k}\|_\alpha \tag{3.41}$$

and

$$\|d_{N,k}^{2,\delta}\|_\alpha = \mathcal{O}\big(N^{-(1-\delta)H(k/N)-\delta L}\big) = o(\|d_{N,k}^{1,\delta}\|_\alpha). \tag{3.42}$$

→ Second reason: Let $I \subseteq [0,1]$ be an arbitrary non-degenerate compact interval, and, as previously, let $\nu_N(I) = \big\{k \in \{0,\ldots,N-L\} : k/N \in I\big\}$. For each fixed $r \in \{0,\ldots,e_N-1\}$, denote by $\mathcal{J}_{N,r}$ the set

$$\mathcal{J}_{N,r} = \big\{k \in \nu_N(I) : \exists\, q \in \mathbb{Z}_+ \text{ s.t. } k = qe_N + r\big\}. \tag{3.43}$$

Then $\{\widetilde{d}_{N,k}^{1,\delta} : k \in \mathcal{J}_{N,r}\}$ is a finite sequence of independent $\mathcal{S}\alpha\mathcal{S}$ random variables. Also, observe that

$$\nu_N(I) = \bigcup_{r=0}^{e_N-1} \mathcal{J}_{N,r} \quad \text{(disjoint union)}. \tag{3.44}$$

Thus, setting

$$V_N^{\beta,1,\delta}(I) = |\nu_N(I)|^{-1} \sum_{k\in\nu_N(I)} |d_{N,k}^{1,\delta}|^{\beta}, \tag{3.45}$$

one has

$$V_N^{\beta,1,\delta}(I) \simeq V_N^{\beta}(I) = |\nu_N(I)|^{-1} \sum_{k\in\nu_N(I)} |d_{N,k}|^{\beta}. \tag{3.46}$$

It seems to be less difficult to obtain "a strong law of large numbers" for the empirical mean $V_N^{\beta,1,\delta}(I)$, than for the empirical mean $V_N^{\beta}(I)$. Indeed, in view of the independence of the $d_{N,k}^{1,\delta}$'s, $k \in \mathcal{J}_{N,r}$, it is much less difficult to find, for all $\eta > 0$, an appropriate upper bound for the probability

$$p_N^1(\eta) = \mathbb{P}\left(\Big| \frac{V_N^{\beta,1,\delta}(I)}{\mathbb{E}\big(V_N^{\beta,1,\delta}(I)\big)} - 1 \Big| > \eta \right), \tag{3.47}$$

than for the probability

$$p_N(\eta) = \mathbb{P}\left(\Big| \frac{V_N^{\beta}(I)}{\mathbb{E}\big(V_N^{\beta}(I)\big)} - 1 \Big| > \eta \right). \tag{3.48}$$

More precisely, for finding an appropriate upper bound for the probability $p_N^1(\eta)$, the main thing to do is to conveniently bound from above the quantity

$$\begin{aligned} A_N &= \mathbb{E}\Big(\Big|V_N^{\beta,1,\delta}(I) - \mathbb{E}\big(V_N^{\beta,1,\delta}(I)\big)\Big|^4\Big) \\ &= |\nu_N(I)|^{-4}\, \mathbb{E}\Big(\Big(\sum_{k\in\nu_N(I)} \Delta_{N,k}\Big)^4\Big), \end{aligned} \tag{3.49}$$

where the $\Delta_{N,k}$'s are the centered random variables defined as

$$\Delta_{N,k} = |d_{N,k}^{1,\delta}|^\beta - \mathbb{E}\big(|d_{N,k}^{1,\delta}|^\beta\big). \tag{3.50}$$

The fact that

$$\nu_N(I) = \bigcup_{r=0}^{e_N-1} \mathcal{J}_{N,r} \quad \text{(disjoint union)}$$

and the convexity property of the function $x \mapsto x^4$ imply that

$$\mathbb{E}\Big(\Big(\sum_{k\in\nu_N(I)} \Delta_{N,k}\Big)^4\Big) \le e_N^3 \sum_{r=0}^{e_N-1} \mathbb{E}\Big(\Big(\sum_{k\in\mathcal{J}_{N,r}} \Delta_{N,k}\Big)^4\Big). \tag{3.51}$$

Next, using the fact that for each fixed $r \in \{0, \dots, e_N - 1\}$, $\{\Delta_{N,k} : k \in \mathcal{J}_{N,r}\}$ is a finite sequence of independent centered random variables, one gets,

$$\mathbb{E}\Big(\Big(\sum_{k\in\mathcal{J}_{N,r}} \Delta_{N,k}\Big)^4\Big) \tag{3.52}$$
$$= \sum_{k\in\mathcal{J}_{N,r}} \mathbb{E}\Big(\big(\Delta_{N,k}\big)^4\Big) + \sum_{(k',k'')\in\mathcal{J}_{N,r}^2,\, k'\neq k''} \mathbb{E}\Big(\big(\Delta_{N,k'}\big)^2\Big)\,\mathbb{E}\Big(\big(\Delta_{N,k''}\big)^2\Big).$$

Moreover, one can derive from the equality $\Delta_{N,k} = |d_{N,k}^{1,\delta}|^\beta - \mathbb{E}\big(|d_{N,k}^{1,\delta}|^\beta\big)$ and the convexity property of the functions $x \mapsto x^4$ and $x \mapsto x^2$, that

$$\mathbb{E}\Big(\big(\Delta_{N,k}\big)^4\Big) \leq 8\,\mathbb{E}\big(|d_{N,k}^{1,\delta}|^{4\beta}\big) + 8\Big(\mathbb{E}\big(|d_{N,k}^{1,\delta}|^{\beta}\big)\Big)^4 \leq c^2 N^{-4\beta H(k/N)} \tag{3.53}$$

and

$$\mathbb{E}\Big(\big(\Delta_{N,k}\big)^2\Big) \leq 2\,\mathbb{E}\big(|d_{N,k}^{1,\delta}|^{2\beta}\big) + 2\Big(\mathbb{E}\big(|d_{N,k}^{1,\delta}|^{\beta}\big)\Big)^2 \leq cN^{-2\beta H(k/N)}, \tag{3.54}$$

where c is a constant non depending on N and k.

In conclusion: Our computations allow to get a convenient upper bound for A_N, which is the main ingredient of the proofs of our two main results that we are now going to state.

Theorem 3.1

Assume that $\beta \in (0, 1/4]$ and L are arbitrary and such that

$$L > \frac{2}{\beta} + 1. \tag{3.55}$$

Let $I \subseteq [0,1]$ be an arbitrary fixed compact interval with a positive Lebesgue measure $\lambda(I)$. For each integer $N \geq (L+1)\lambda(I)^{-1}$, we set

$$\widehat{H}_N^{\beta}(I) = \beta^{-1} \log_2 \left(\frac{V_N^{\beta}(I)}{V_{2N}^{\beta}(I)} \right). \tag{3.56}$$

Then, there exists an almost surely finite random variable $C > 0$, such that one has almost surely for all $N \geq (L+1)\lambda(I)^{-1}$,

$$\left| \widehat{H}_N^{\beta}(I) - \min_{t \in I} H(t) \right| \leq C \, \frac{\log \log N}{\log N}. \tag{3.57}$$

Construction of an estimator for the function $H(\cdot)$

→ One splits, for any large integer N, the interval $[0,1]$, into a finite sequence $(\mathcal{I}_{N,n})_{0\le n\le M_N}$ of M_N+1 compact subintervals with the same length θ_N, except the last one $\mathcal{I}_{N,M_N}$ having a length between θ_N and $2\theta_N$.

→ Then, one denotes by $\left\{\widetilde{H}^{\beta}_{N,\theta_N}(t)\,:\,t\in[0,1]\right\}$ the stochastic process with piecewise linear sample paths obtained as a linear interpolation between the M_N+2 random points given by the coordinates

$$\big(0,\widehat{H}^{\beta}_N(\mathcal{I}_{N,0})\big);\ldots;\big((M_N-1)\theta_N,\widehat{H}^{\beta}_N(\mathcal{I}_{N,M_N-1})\big);\big(M_N\theta_N,\widehat{H}^{\beta}_N(\mathcal{I}_{N,M_N})\big);\big(1,\widehat{H}^{\beta}_N(\mathcal{I}_{N,M_N})\big);$$

where for all $n\in\{0,\ldots,M_n\}$, $\widehat{H}^{\beta}_N\big(\mathcal{I}_{N,n}\big)$ is the estimator of $\min_{t\in\mathcal{I}_{N,n}} H(t)$ provided by the previous theorem.

Notice that Bardet and Surgailis (2013) gave, in a general Gaussian multifractional frame, uniformly and strongly consistent estimators for Hurst functions.

Theorem 3.2

Assume that $\beta \in (0, 1/4]$ *and* L *are arbitrary and such that* $L > \frac{2}{\beta} + 1$*. Also, assume that*

$$\lim_{N \to +\infty} N^{\frac{\beta(L-1)-2}{4\beta(L-1)+2}} \theta_N = +\infty. \tag{3.58}$$

Then, there exists an almost surely finite random variable $C > 0$*, such that one has almost surely for all integer* N *big enough,*

$$\left\| H - \widetilde{H}^{\beta}_{N,\theta_N} \right\|_\infty = \sup_{t \in [0,1]} \left| H(t) - \widetilde{H}^{\beta}_{N,\theta_N}(t) \right| \tag{3.59}$$

$$\leq C \left(\theta_N^{-1} N^{-\frac{\beta(L-1)-2}{4\beta(L-1)+2}} + N^{-\beta(\rho_H - \sup_{t\in[0,1]} H(t))} (\log N)^2 + \theta_N^{\rho_H} \right).$$

Condition (3.58) means that the length θ_N of the intervals $\mathcal{I}_{N,n}$ must not go very fast to zero, namely, its rate of convergence has to be at least slower than $N^{-1/4}$. On the other hand, even if this rate is extremely slow, $\widetilde{H}^{\beta}_{N,\theta_N}(\cdot)$ remains an almost surely uniformly convergent estimator of $H(\cdot)$; yet, it is absurd to choose θ_N in this way! A reasonable choice for θ_N, would be $\theta_N = a\, N^{-\frac{\zeta(\beta(L-1)-2)}{4\beta(L-1)+2}}$, where $a > 0$ and $\zeta \in (0,1)$ are two parameters to adjust according to situation.