Estimating self-similarity through complex variations

Jacques Istas

Laboratoire Jean Kuntzmann,
Université de Grenoble et CNRS,
F-38041 GRENOBLE cedex 9
e-mail: Jacques.Istas@imag.fr

Abstract: We estimate the self-similarity index of a H-sssi process through complex variations. The advantage of the complex variations is that they do not require existence of moments and can therefore be used for infinite variance processes.

Contents

1. Introduction

A real-valued process X is H-self-similar if, for all $a > 0$

$$\{X(at),\ t \in \mathbb{R}\} \quad \overset{(d)}{=} \quad a^H \{X(t),\ t \in \mathbb{R}\},$$

where $\stackrel{(d)}{=}$ stands for equality of the finite dimensional distributions. Process X has stationary increments if, for all $s \in \mathbb{R}$,

$$\{X(t+s) - X(s),\ t \in \mathbb{R}\} \stackrel{(d)}{=} \{X(t) - X(0), t \in \mathbb{R}\}.$$

A process that is H-self-similar with stationary increments will be called H-sssi. The aim of this paper is to estimate H from a discrete sample of X over the time interval $[0, 1]$. More precisely, one observes a H-sssi process X at a discrete sampling k/n, $k = 0, \ldots, n$. Let $\mathbf{a} = (a_0, a_1, \ldots, a_K)$ be a finite sequence with $L + 1$ vanishing moments

$$\sum_{k=0}^{K} a_k k^{\ell} = 0 \quad \ell = 0, \ldots, L,$$

$$\sum_{k=0}^{K} a_k k^{L+1} \neq 0,$$

with convention $0^0 = 1$. The increments of X with respect to the sequence $\mathbf{a}$ are defined by

$$\Delta_{p,n} X = \sum_{k=0}^{K} a_k X\left(\frac{p+k}{n}\right).$$

Classical variations ($L = 0$) are associated with $a_0 = 1$, $a_1 = -1$

$$\Delta_{p,n} X = X(p/n) - X((p+1)/n).$$

An example with $L = 1$ is given by $a_0 = -1, a_1 = 2, a_2 = -1$

$$\Delta_{p,n} X = 2X((p+1)/n) - X(p/n) - X((p+2)/n).$$

A usual statistical tool is the ϕ-variation

$$V_n(\phi, X) = \frac{1}{n - K + 1} \sum_{p=0}^{n-K} \phi(|\Delta_{p,n} X|), \tag{1.1}$$

where ϕ is a given function. When X is a fractional Brownian motion, generalized quadratic variations ($\phi(x) = x^2$, $L \geq 1$) provides a consistent estimate of H, asymptotically normal with rate $\sqrt{n}$ [10]. These generalized quadratic variations can be used for Rosenblatt or more generally Hermite processes [4, 5, 17], that are processes with finite variance, but the $\sqrt{n}$-rate of convergence is not always obtained. A similar tool is the wavelet's one. Roughly speaking, the increments of the process X are replaced by wavelet coefficients [1, 11]. For H-sssi processes with infinite variance, for instance α-stable processes, the use of (generalized) quadratic variations is hopeless. One could try with functions $\phi(x) = x^p$, $0 < p < \alpha$. Indeed, the p-variations still lead to estimate for fractional Brow-

nian motions [12]. But, for stable processes, this requires an a priori accurate estimate of the stability index α. One can try with log-variations $\phi(x) = \log|x|$ as in [7, 9]. Indeed, this estimate requires only the existence of logarithmic moments. Unfortunately, the rate of convergence of the self-similarity index estimate is very slow (logarithmic [9, sec. 6.2]). [14] propose an estimate based on weighted sums of logarithm of wavelet coefficients, but on a time interval $[0, T]$ with $T \to +\infty$.

At present, in the statistical literature, there is no self-similarity index consistent estimate with a reasonable rate of convergence. One only knows that the self-similarity index is identifiable from the observation of one sample path over a bounded interval. Let us come back to the ϕ-variations, with $\phi(x) = x^p$. These power functions are used since, for all $x, y > 0$, $\phi(x/y) = \phi(x)/\phi(y)$. But, for any $p \in \mathbb{C}$, we still have $(x/y)^p = x^p/y^p$. Among these complex powers, the purely imaginary powers ($p = iM$, $M \in \mathbb{R}$) are particularly interesting, since $|x^{iM}| = 1$. Indeed, for any positive random variable U, the expectation of U^i still exists. In this paper, we therefore work with complex variations

$$V_n(X) \quad = \quad \frac{1}{n - K + 1} \sum_{p=0}^{n-K} |\Delta_{p,n} X|^{iM}.$$

These complex variations provide a self-similarity index estimate without assumptions on the existence of moments of X. Under suitable conditions, we prove the consistency and obtain the rate of convergence.

We then consider family of examples. H-sssi second-order processes exist for $0 < H \leq 1$. Stable H-sssi processes exist for $0 < H \leq \max(1, 1/\alpha)$. Therefore, we consider examples that cover the range of admissible parameters H: H-sssi processes with independent increments, fractional Brownian motions, well-balanced linear fractional stable motions and Takenaka's processes. In the case of fractional Brownian motion, we provide a central limit theorem and a numerical comparison between quadratic variations, log variations and complex variations. Finally, we prove that the distribution of $X(1)$ is identifiable, even if we do not build a tractable estimate.

2. Main result

2.1. Settings and assumptions

For $r > 0$, $M \in \mathbb{R}$, set

$$r^{iM} \quad = \quad e^{iM \log(r)}.$$

For $z \in \mathbb{C}^\star$, let

$$z \quad = \quad |z| e^{i \arg(z)}, \quad 0 \leq \arg(z) < 2\pi.$$

We will make some of the following assumptions.

1. $\mathbb{P}(\Delta_{0,1}X = 0) = 0$.
2. $\mathbb{E}|\Delta_{0,1}X|^{iM} \neq 0$.
3. (a)

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{|k| \leq n} |cov(|\Delta_{k,1}X|^{iM}, |\Delta_{0,1}X|^{iM})| = 0.$$

(b) There exists $0 < \gamma < 1/2$ such that

$$\limsup_{n \to +\infty} \frac{1}{n^{2\gamma}} \sum_{|k| \leq n} |cov(|\Delta_{k,1}X|^{iM}, |\Delta_{0,1}X|^{iM})| = \Sigma^2.$$

Trivial cases, like $X(t) = Ut$ $(U \sim \mathcal{N}(0,1))$ and $L \geq 1$, are forbidden by the first condition. For the second condition, the function $M \mapsto \mathbb{E}|\Delta_{0,1}X|^{iM}$ is continuous and equal to 1 at $M = 0$. There always exists a neighborhood of 0 such that $\mathbb{E}|\Delta_{0,1}X|^{iM} \neq 0$. Moreover, we prove that $\mathbb{E}|\Delta_{0,1}X|^{iM}$ is never vanishing for Gaussian and stable variables. We show in the section "Examples" how to check the third condition on several H-sssi processes. Note that the limit case $\gamma = 0$ corresponds to the case when the series $\sum_{k \in \mathbb{Z}} |cov(|\Delta_{k,1}X|^{iM}, |\Delta_{0,1}X|^{iM})|$ is convergent.

Let us define the estimate of the self-similarity index H

$$\widehat{H}_n = \frac{1}{M \log 2} \arg(V_{n/2}(X)/V_n(X)).$$

Let us make an heuristic explanation of this estimator $\widehat{H}_n$. As proved later, the expectation of $V_n(X)$ is equal to $n^{-iHM}\mathbb{E}|\Delta_{0,1}X|^{iM}$. If $V_n(X)$ converges, as $n \to +\infty$, to its expectation, the ratio $V_{n/2}/V_n$ will converge to 2^{iHM}. Since the modulus of $V_{n/2}/V_n$ may be not equal to 1, we need to take the argument of $V_{n/2}/V_n$ to build the estimator $\widehat{H}_n$.

We have the following, where $O_\mathbb{P}$ is defined by

- $X_n = O_\mathbb{P}(1)$ iff $\forall \varepsilon > 0$, $\exists C > 0$ such that $\sup_n \mathbb{P}(|X_n| \geq C) \leq \varepsilon$.
- $Y_n = O_\mathbb{P}(\delta_n)$ means $Y_n = \delta_n X_n$ with $X_n = O_\mathbb{P}(1)$.

Theorem 2.1. *Assume that $0 < H < 2\pi/(M \log(2))$. Then, assuming 3)a)*

-

$$\lim_{n \to +\infty} \widehat{H}_n = H \ (\mathbb{P}). \tag{2.1}$$

- *Assuming moreover 3)b)*

$$\widehat{H}_n - H = O_\mathbb{P}(n^{\gamma - 1/2}). \tag{2.2}$$

2.2. Proof

Set

$$W_n(X) = n^{iHM} V_n(X).$$

Since process X is H-sssi

$$\mathbb{E} W_n(X) = \mathbb{E}|\Delta_{0,1}X|^{iM}.$$

Then

$$
\begin{aligned}
\mathbb{E}|W_n(X)|^2 &= \frac{1}{(n-K+1)^2} \sum_{p,p'=0}^{n-K} \mathbb{E}|\Delta_{p,1}X|^{iM}|\Delta_{p',1}X|^{-iM} \\
&= \frac{1}{n-K+1} \sum_{|k|\leq n-K} \left(1 - \frac{|k|}{n-K+1}\right) \mathbb{E}|\Delta_{k,1}X|^{iM}|\Delta_{0,1}X|^{-iM}.
\end{aligned}
$$

It follows

$$
\begin{aligned}
\mathbb{E}|W_n(X) - \mathbb{E}|\Delta_{0,1}X|^{iM}|^2 &\leq \frac{1}{n-K+1} \sum_{|k|\leq n-K} \left(1 - \frac{|k|}{n-K+1}\right) \\
&\quad \times |cov(|\Delta_{k,1}X|^{iM}, |\Delta_{0,1}X|^{iM})| \qquad (2.3)
\end{aligned}
$$

and (2.1) is proved. Using (2.3), assumption 3)b) and the dominated convergence theorem then yield

$$\limsup_{n\to+\infty} n^{1-2\gamma}\mathbb{E}|W_n(X) - \mathbb{E}|\Delta_{0,1}X|^{iM}|^2 \leq \Sigma^2.$$

It follows that $W_n(X) - \mathbb{E}|\Delta_{0,1}X|^{iM} = O_{\mathbb{P}}(n^{\gamma-1/2})$. Since $z \mapsto \arg(z)$ is a C^1 function around $\mathbb{E}|\Delta_{0,1}X|^{iM}$, (2.2) is proved.

3. Examples

3.1. H-sssi processes with independent increments

For the Brownian motion $(H = 1/2)$, Theorem 2.1 applies obviously with any L and any $M < 4\pi/\log(2)$.

For the Lévy α-stable processes $(0 < \alpha < 2)$, where $H = 1/\alpha$, Theorem 2.1 applies obviously for any L and $H \in (0, 2\pi/(M\log(2)))$. This is a drawback of Theorem 2.1: one needs an a priori upper bound of H. Nevertheless, note that in many examples, the index H belongs to $(0,1)$, and any $M < 2\pi/\log(2)$ is appropriate.

3.2. Fractional Brownian motions

3.2.1. Theoretical results

Fractional Brownian motion is the unique, up to a constant, Gaussian H-sssi process, with $H \in (0,1]$. The case $H = 1/2$ has just been treated. Our assumptions exclude the trivial case $H = 1$. Let us consider the other cases. We prove that Theorem 2.1 holds with $L \geq 1$ and any $M < 2\pi/\log(2)$. Moreover, we prove a central limit theorem.

Let us first prove that Assumption 2) is satisfied for any M.

Lemma 3.1. *Let X be a centered Gaussian variable with variance $\sigma^2 \neq 0$. Then, for all $M \in \mathbb{R}$, $\mathbb{E}|X|^{iM} \neq 0$.*

Proof.

$$\mathbb{E}|X|^{iM} = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \int_0^{+\infty} x^{iM} e^{-x^2/(2\sigma^2)} dx.$$

The change of variable $z = x^2/(2\sigma^2)$ leads to

$$\begin{aligned}
\mathbb{E}|X|^{iM} &= \sigma^{iM} \frac{2^{iM/2}}{\sqrt{\pi}} \int_0^{+\infty} z^{(iM-1)/2} e^{-z} dz \\
&= \sigma^{iM} \frac{2^{iM/2}}{\sqrt{\pi}} \Gamma\left(\frac{1+iM}{2}\right),
\end{aligned}$$

where Γ is the Gamma function. Since $\Gamma\left(\frac{1+iM}{2}\right) \neq 0$, Lemma 3.1 is proved. $\square$

Proposition 3.1. *Assume $M < 2\pi/\log(2)$ and $L \geq 1$. Then $\sqrt{n}(\widehat{H}_n - H)$ converges in distribution, as $n \to +\infty$, to a centered Gaussian variable.*

We start with the following Lemma.

Lemma 3.2. *Let $(U, V) \sim \mathcal{N}_2\left((0,0); \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)\right)$, $|\rho| \leq 1$. There exists $C > 0$, such that, $\forall |\rho| \leq 1$*

$$|cov(|U|^{iM}, |V|^{iM})| \leq C\rho^2. \tag{3.1}$$

Proof. Let $\Sigma = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$. Then

$$\begin{aligned}
det(\Sigma) &= 1 - \rho^2, \\
\Sigma^{-1} &= \frac{1}{1-\rho^2} \left(\begin{array}{cc} 1 & -\rho \\ -\rho & 1 \end{array}\right)
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{E}(|U|^{iM}|V|^{-iM}) &= \frac{1}{2\pi\sqrt{(1-\rho^2)}} \int_{\mathbb{R}^2} |x|^{iM}|y|^{-iM} \\
&\quad \times \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 + y^2 - 2\rho xy)\right\} dxdy,
\end{aligned}$$

and

$$cov(|U|^{iM}, |V|^{iM}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} |x|^{iM} |y|^{-iM} \exp(-(x^2+y^2)/2) A_\rho(x,y) dx dy,$$

where

$$A_\rho(x,y) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{\rho^2}{1-\rho^2}(x^2+y^2)\right) \exp\left(\frac{\rho xy}{1-\rho^2}\right) - 1.$$

Since

$$\int_{\mathbb{R}} |x|^{iM} x e^{-x^2/2} dx = 0,$$

one gets

$$cov(|U|^{iM}, |V|^{iM}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} |x|^{iM} |y|^{-iM} \exp\left(-\frac{x^2+y^2}{2}\right)$$
$$\times B_\rho(x,y) dx dy, \tag{3.2}$$

where

$$B_\rho(x,y) = A_\rho(x,y) - \rho xy.$$

One then checks that

$$\lim_{\rho \to 0} \frac{B_\rho(x,y)}{\rho^2} = 1/2 + x^2 y^2/2 - (x^2+y^2)/2.$$

$B_0(x,y) = 0$, $\left(\frac{\partial B_\rho(x,y)}{\partial \rho}\right)_{\rho=0} = 0$ and

$$\frac{\partial^2 B_\rho(x,y)}{\partial \rho^2} = P_\rho(x,y) \exp\left(-\frac{\rho^2}{1-\rho^2}(x^2+y^2)\right) \exp\left(\frac{\rho xy}{1-\rho^2}\right),$$

where $P_\rho(x,y)$ is a fourth degree polynomial that depends continuously on ρ. By a Taylor expansion of order 2 of the function $\rho \mapsto B_\rho(x,y)$, there exists $\tilde{\rho} \in (0, \rho)$ such that

$$B_\rho(x,y) = \rho^2 P_{\tilde{\rho}}(x,y) \exp\left(-\frac{\tilde{\rho}^2}{1-\tilde{\rho}^2}(x^2+y^2)\right) \exp\left(\frac{\tilde{\rho} xy}{1-\tilde{\rho}^2}\right).$$

On the compact set $|\rho| \le 1/2$, the polynomials $P_\rho(x,y)$ can be bounded by a fourth degree polynomial $P(x,y)$. More precisely, for all $x, y \in \mathbb{R}$, $|P_\rho(x,y)| \le |P(|x|, |y|)|$. Then

$$\exp\left(-\frac{\tilde{\rho}^2}{1-\tilde{\rho}^2}(x^2+y^2)\right) \exp\left(\frac{\tilde{\rho} xy}{1-\tilde{\rho}^2}\right) = \exp\left(-\frac{\tilde{\rho}^2}{1-\tilde{\rho}}(x-y)^2\right)$$
$$\times \exp\left(\frac{\tilde{\rho}}{2(1+\tilde{\rho})} xy\right)$$
$$\le \exp(|xy|/3).$$

This leads to

$$|B_\rho(x,y)| \quad \leq \quad \rho^2 |P(|x|,|y|)| \exp(|xy|/3).$$

One then checks that

$$\int_{\mathbb{R}^2} \exp\left(-(x^2+y^2)/2\right) |P(|x|,|y|)| \exp(|xy|/3) dx dy \quad < \quad \infty.$$

The dominated convergence theorem is applied to (3.2) as $\rho \to 0$. Since $|cov(|U|^{iM},|V|^{iM})|$ is bounded by 1, (3.1) is proved for any $|\rho| \leq 1/2$. $\square$

Let us now prove Proposition 3.1. We next compute the covariance between $\Delta_{k,1}X$ and $\Delta_{0,1}X$

$$cov(\Delta_{k,1}X, \Delta_{0,1}X) \quad = \quad \frac{1}{2} \sum_{p,p'=0}^{K} a_p a_{p'} |k+p-p'|^{2H}.$$

A Taylor expansion up to order 3 leads to

$$|cov(\Delta_{k,1}X, \Delta_{0,1}X)| \quad \leq \quad Ck^{2H-3}.$$

Lemma 3.2 ensure the convergences of the series $\sum_{p\in\mathbb{Z}} |cov(|\Delta_{k,1}X|^{iM}, |\Delta_{0,1}X|^{iM})|$ for any $H \in (0,1)$.

$W_n(X)$ has the same distribution as $\frac{1}{n-K+1} \sum_{k=0}^{n-K} |\Delta_{0,1}X|^{iM}$. Let us recall the Breuer-Major's theorem [2, 13]. The Hermite rank d of a function f is the order of the first non-trivial coefficient in the Hermite expansion of f. Let $(X_k), k \in \mathbb{Z}$ be a centered stationary Gaussian sequence. Let $\rho(k) = \mathbb{E}(X_0 X_k)$. Assume $d \geq 1$ and $\sum_{k\in\mathbb{Z}} |\rho(k)|^d < \infty$. Then $\frac{1}{\sqrt{n}} \sum_1^n f(X_k)$ converges in distribution, as $n \to +\infty$, to a centered Gaussian variable. Since

$$\int_{\mathbb{R}} |x|^{iM} x e^{-x^2/2} dx \quad = \quad 0,$$

the Hermite rank of any linear combination $\alpha Re(|x|^{iM}) + \beta Im(|x|^{iM})$ is at least 2. The Breuer-Major's theorem can be applied to any real function $f(x) = \alpha Re(|x|^{iM}) + \beta Im(|x|^{iM})$. This proves the convergence of the couple $(Re(|x|^{iM}), Im(|x|^{iM}))$. Therefore, $\sqrt{n}(W_n(X) - \mathbb{E}|\Delta_{0,1}X|^{iM})$ converges in distribution, as $n \to +\infty$, to a centered Gaussian variable. For the convergence of the couple $(W_{n/2}, W_n)$, one needs a bivariate Breuer-Major's theorem. The proof of this bivariate Breuer-Major's theorem is similar to the proof of the classical one. We sketch the proof, following [13, ch. 6 & 7]. Let f and g be two real functions which Hermite ranks are at least $d \geq 1$. Let (f_q) and (g_q) be the coefficients of the Hermite expansion

$$f \quad = \quad \sum_{q \geq d} f_q H_q,$$

$$g \quad = \quad \sum_{q \geq d} g_q H_q,$$

where (H_q) are the Hermite polynomials. The classical Breuer-Major's theorem ensures that $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} H_q(X_k)$ converges in distribution, as $n \to +\infty$ to a centered Gaussian variable $\sqrt{q! \sum_{k \in \mathbb{Z}} \rho(k)^q} N_q$, where N_q is a centered standard Gaussian variable. Since the Hermite polynomials are orthogonal, condition (6.2.11) of Theorem 6.2.3 of [13] is straightforwardly satisfied. It follows, for any given $m \geq d$ that

$$\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} H_q(X_k), d \leq q \leq m \right) \tag{3.3}$$

converges in distribution, as $n \to +\infty$, to a centered Gaussian vector

$$\left(\sqrt{q! \sum_{k \in \mathbb{Z}} \rho(k)^q} N_q, d \leq q \leq m \right)$$

where the centered standard Gaussian variables (N_q) are independent. We then mimic the proof of Theorem 7.2.4 of [13]. Condition (d) is satisfied, one can get $m = +\infty$ in (3.3). With other words, we have proved that the couple

$$\left(\frac{1}{\sqrt{n}} \sum_{1}^{n} f(X_k), \frac{1}{\sqrt{n}} \sum_{1}^{n} g(X_k) \right)$$

converges in distribution, as $n \to +\infty$, to the centered Gaussian vector

$$\left(\sum_{q \geq d} f_q \sqrt{q! \sum_{k \in \mathbb{Z}} \rho(k)^q} N_q, \sum_{q \geq d} g_q \sqrt{q! \sum_{k \in \mathbb{Z}} \rho(k)^q} N_q \right)$$

The proof of the convergence of the couple $(W_{n/2}, W_n)$ is similar. One proves the convergence of the vector

$$\left(\frac{1}{\sqrt{n-K+1}} \sum_{k=0}^{n-K} H_q\left(\frac{\Delta_{k,n} X}{\mathrm{var}(\Delta_{k,n} X)} \right), \frac{1}{\sqrt{n/2-K+1}} \sum_{k=0}^{n/2-K} H_q\left(\frac{\Delta_{k,n/2} X}{\mathrm{var}(\Delta_{k,n/2} X)} \right) \right)$$

first for a given q, with $d \leq q$, using the fourth moment theorem [13, Th.5.2.7], then for any given m, with $d \leq q \leq m$, using the orthogonality of the Hermite polynomials, then for any $d \leq q$. Since the estimator $\hat{H}_n$ is a smooth function of the couple $(W_{n/2}, W_n)$, the central limit theorem of Proposition 3.1 is proved.

3.2.2. Numerical comparisons

Let us now numerically compare the estimate $\widehat{H}_n$ of Theorem 2.1 with quadratic and log variations estimates. More precisely, we use $\phi(x) = x^2$ and $\phi(x) = \log(x)$ in (1.1). We compute the mean estimation $\overline{\widehat{H}}_M$ and the mean square error

$$\overline{\widehat{H}}_{SP} = \frac{1}{SP} \sum_{m=1}^{SP} \left(\hat{H}_m - H \right)^2,$$

TABLE 1
Quadratic variations (on the top), log-variations (in the middle) and complex variations (on the bottom). Mean estimation is indicated, mean square errors are in brackets. Simulations are done with the R package FieldSim [3]

	$N = 7$	$N = 8$	$N = 9$
$H = 0.5$	0.485 (0.030)	0.500 (0.012)	0.495 (0.007)
$H = 0.7$	0.690 (0.017)	0.690 (0.010)	0.695 (0.005)
$H = 0.8$	0.795 (0.018)	0.795 (0.009)	0.800 (0.004)

	$N = 7$	$N = 8$	$N = 9$
$H = 0.5$	0.490 (0.0007)	0.490 (0.0003)	0.490 (0.0002)
$H = 0.7$	0.725 (0.0013)	0.725 (0.0007)	0.720 (0.0005)
$H = 0.8$	0.859 (0.0039)	0.855 (0.0030)	0.850 (0.0024)

	$N = 7$	$N = 8$	$N = 9$
$H = 0.5$	0.500 (0.050)	0.500 (0.022)	0.490 (0.012)
$H = 0.7$	0.680 (0.040)	0.690 (0.023)	0.710 (0.010)
$H = 0.8$	0.825 (0.037)	0.805 (0.020)	0.800 (0.008)

for $SP = 200$ sample paths, for different values of H and $n = 2^N$. At first sighat, quadratic variations leads to better results, that is consistent with [6]. Logarithmic variations leads to worst results and complex variations are in between.

3.3. Well-balanced linear fractional stable motions

Let M be a symmetric α-stable random measure ($0 < \alpha < 2$) with Lebesgue control measure. We refer to [15] for basic facts on integration with respect to a stable random measure. The well-balanced linear fractional stable motions are defined by [15, p.140]

$$X(t) \quad = \quad \int_{\mathbb{R}} (|t - s|^{H-1/\alpha} - |s|^{H-1/\alpha}) M(ds),$$

where $0 < H < 1$ and $H \neq 1/\alpha$. Process X is H-sssi. We prove, in the following serie of Lemmas 3.3, 3.4, 3.5 and 3.6, that, for all $L \geq 0$

$$\lim_{n \to +\infty} \widehat{H}_n \quad = \quad H \ (\mathbb{P}),$$

and that there exists $C > 0$ such that

$$\widehat{H}_n - H \quad = \quad O_{\mathbb{P}}(n^{-\min(1/2, 1/4(\alpha L - \alpha(H-1)))}, \text{ if } \alpha L - \alpha(H - 1) \neq 2$$
$$\widehat{H}_n - H \quad = \quad O_{\mathbb{P}}\left(\sqrt{\frac{\log(n)}{n}} \right) \text{ if } \alpha L - \alpha(H - 1) = 2.$$

Therefore, Theorem 2.1 holds with $L > 2/\alpha + H - 1$ and any $M < 2\pi/\log(2)$. With other words, to obtain the optimal rate of convergence, one needs to know an a priori bound for the stability index α.

Lemma 3.3. *Let $p \in \mathbb{C}$, with $0 < \mathcal{R}(p) < 1$. For all $a \in \mathbb{R}$*

$$\int_{\mathbb{R}} \frac{1 - e^{iax}}{|x|^{1+p}} dx = C_p |a|^p,$$

$$C_p = \frac{\pi}{\Gamma(p+1) \sin(\pi p/2)}.$$

Proof. Following formula 3.823 of [8], for $a > 0$ and $0 < \mathcal{R}(p) < 2$

$$\int_0^{+\infty} \frac{\sin^2(ax)}{x^{1+p}} dx = -\Gamma(-p) \cos(p\pi/2) 2^{p-1} a^p.$$

Since $2\sin^2(x) = 1 - \cos(2x)$, for all $a \in \mathbb{R}$, $0 < \mathcal{R}(p) < 2$

$$\int_{\mathbb{R}} \frac{1 - \cos(ax)}{|x|^{1+p}} dx = -2\Gamma(-p) \cos(p\pi/2) |a|^p.$$

One then has, with $0 < \mathcal{R}(p) < 1$

$$\int_{\mathbb{R}} \frac{\sin(ax)}{|x|^{1+p}} dx = 0.$$

Therefore, for all $a \in \mathbb{R}$, $0 < \mathcal{R}(p) < 1$

$$\int_{\mathbb{R}} \frac{1 - e^{iax}}{|x|^{1+p}} dx = -2\Gamma(-p) \cos(p\pi/2) |a|^p.$$

Using $-p\Gamma(-p) = \Gamma(1-p)$ and $\Gamma(1-p)\Gamma(p) = \pi/\sin(p\pi)$, Lemma 3.3 is proved. $\square$

Let us now prove that Assumption 2), for a standard stable variable, is satisfied for any M.

Lemma 3.4. *Let X be a standard α-stable variable with characteristic function*

$$\mathbb{E}(\exp(i\lambda X)) = \exp(-|\lambda|^\alpha).$$

Then, for all $M \in \mathbb{R}$, $\mathbb{E}|X|^{iM} \neq 0$.

Proof. From Lemma 3.3, with $0 < \mathcal{R}(p) < \min(\alpha, 1)$

$$C_p \mathbb{E}|X|^p = 2 \int_0^{+\infty} \frac{1 - e^{-x^\alpha}}{x^{1+p}} dx.$$

Let us perform the change of variable $x^\alpha = y^2$

$$C_p \mathbb{E}|X|^p = \frac{4}{\alpha} \int_0^{+\infty} \frac{1 - e^{-y^2}}{y^{1+2p/\alpha}} dy. \tag{3.4}$$

Let Z be a centered standard Gaussian variable. The same can be written for Z

$$C_p \mathbb{E}|Z|^p = 2 \int_0^{+\infty} \frac{1 - e^{-y^2}}{y^{1+p}} dy. \tag{3.5}$$

Following Lemma 3.1

$$\mathbb{E}|Z|^p = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{1+p}{2}\right). \qquad (3.6)$$

(3.4), (3.5) and (3.6) lead to

$$\mathbb{E}|X|^p = \frac{2^{(2+p)/2}}{\alpha\sqrt{\pi}}\Gamma\left(\frac{1}{2}+\frac{p}{\alpha}\right).$$

The dominated convergence theorem then leads, as $p \to iM$,

$$\mathbb{E}|X|^{iM} = \frac{2^{(2+iM)/2}}{\alpha\sqrt{\pi}}\Gamma\left(\frac{1}{2}+\frac{iM}{\alpha}\right),$$

and Lemma 3.4 is proved. □

Lemma 3.5. *Let (S,μ) be a measure space, $f, g \in L^\alpha(S,\mu)$ and M be a symmetric α-stable random measure on S with control measure μ. Set*

$$U = \int_S f dM,$$

$$V = \int_S g dM.$$

Assume

$$\int_S |f|^\alpha = 1,$$

$$\int_S |g|^\alpha = 1,$$

$$\int_S |fg|^{\alpha/2} \leq \eta < 1.$$

There exists a constant $C(\eta)$ such that

$$|cov(|U|^{iM}, |V|^{iM})| \leq C(\eta)\int_S |fg|^{\alpha/2}.$$

Proof. First note that, since function $x \mapsto |x|^{iM}$ does not belong to L^1, [14, Th.2.1] cannot be applied.

Let $0 < \varepsilon < \alpha/2$ and let us apply Lemma 3.3 with $p = iM + \varepsilon$.

$$|U|^{iM+\varepsilon} = \frac{1}{C_{iM+\varepsilon}}\int_{\mathbb{R}}\frac{1-e^{ixU}}{|x|^{1+\varepsilon+iM}}dx,$$

and the same for V. Fubini's Theorem then leads to

$$cov(|U|^{iM+\varepsilon}, |V|^{iM+\varepsilon}) = \frac{1}{C_{iM+\varepsilon}C_{-iM+\varepsilon}}\int_{\mathbb{R}^2}\frac{cov(e^{ixU}, e^{iyV})}{|x|^{1+\varepsilon+iM}|y|^{1+\varepsilon-iM}}dxdy$$

$$= \frac{1}{C_{iM+\varepsilon}C_{-iM+\varepsilon}}\int_{\mathbb{R}^2}\frac{\exp(-\int_S |xf+yg|^\alpha) - \exp(-\int_S |xf|^\alpha - \int_S |yg|^\alpha)}{|x|^{1+\varepsilon+iM}|y|^{1+\varepsilon-iM}}dxdy.$$

Let us use the following bound (equation (3.6) of [14, Lemma 3.1])

$$\left| \exp\left(-\int_S |xf + yg|^\alpha \right) - \exp\left(-\int_S |xf|^\alpha - \int_S |yg|^\alpha \right) \right|$$
$$\leq 2 \int_S |fg|^{\alpha/2} |xy|^{\alpha/2} \exp(-2(1-\eta)|xy|^{\alpha/2}).$$

Therefore

$$\begin{aligned}
|cov(|U|^{iM+\varepsilon}, |V|^{iM+\varepsilon})| &\leq \frac{2\int_S |fg|^{\alpha/2}}{|C_{iM+\varepsilon}C_{-iM+\varepsilon}|} I_\varepsilon \\
I_\varepsilon &= \int_{\mathbb{R}^2} \frac{|xy|^{\alpha/2}\exp(-2(1-\eta)|xy|^{\alpha/2})}{|x|^{1+\varepsilon}|y|^{1+\varepsilon}} dxdy.
\end{aligned}$$

By the dominated convergence theorem

$$\lim_{\varepsilon\to 0^+} I_\varepsilon = \int_{\mathbb{R}^2} \frac{|xy|^{\alpha/2}\exp(-2(1-\eta)|xy|^{\alpha/2})}{|x||y|} dxdy.$$

There exists therefore a constant $K_1 > 0$ such that, for all $0 < \varepsilon < \alpha/2$, $I_\varepsilon \leq K_1$. Coming back to Lemma 3.3

$$\lim_{\varepsilon\to 0^+} |C_{iM+\varepsilon}C_{-iM+\varepsilon}| = |C_{iM}C_{-iM}| \neq 0.$$

There exists therefore a constant $K_2 > 0$ such that, for all $0 < \varepsilon < \alpha/2$, $|C_{iM+\varepsilon}C_{-iM+\varepsilon}| \geq K_2$. By the dominated convergence theorem

$$\lim_{\varepsilon\to 0^+} cov(|U|^{iM+\varepsilon}, |V|^{iM+\varepsilon}) = cov(|U|^{iM}, |V|^{iM}),$$

and Lemma 3.5 is proved. $\qquad\square$

Lemma 3.6. *Set*

$$f(t) = \sum_{i=0}^K a_i |i - t|^{H-1/\alpha}.$$

There exists a constant $C > 0$ such that

$$\left(\int_{\mathbb{R}} |f(t+p)f(t)|^{\alpha/2} dt \right)^2 \leq C p^{\alpha H - \alpha(L+1)}.$$

Proof. Using the change of variable $x = t - p/2$

$$\int_{\mathbb{R}} |f(t+p)f(t)|^{\alpha/2} dt = \int_{\mathbb{R}} |f(x+p/2)f(x-p/2)|^{\alpha/2} dx.$$

A Taylor expansion proves that there exists $C > 0$ such that, for $|t| \geq 2K$

$$|f(t)| \leq C|t|^{H-1/\alpha-L-1}. \tag{3.7}$$

By Cauchy-Schwarz inequality

$$\left(\int_{|x| \geq 2K} |f(x + p/2)f(x - p/2)|^{\alpha/2} dx \right)^2 \leq \int_{|x| \geq 2K} |f(x + p/2)|^\alpha dx$$
$$\times \int_{|x| \geq 2K} |f(x - p/2)|^\alpha dx \quad (3.8)$$

(3.8) together with (3.7) lead to

$$\left(\int_{|x| \geq 2K} |f(x + p/2)f(x - p/2)|^{\alpha/2} dx \right)^2 \leq Cp^{\alpha H - \alpha(L+1)}$$

We use again (3.7) for large p. We then show that, for any p

$$\int_{|x| \leq 2K} |f(x + p/2)f(x - p/2)|^{\alpha/2} dx \leq Cp^{\alpha H - 1 - \alpha(L+1)}$$

$\square$

3.4. Takenaka's processes

Let us recall the construction of the one dimensional Takenaka's process [16]. Set, for $t \in \mathbb{R}$

$$C_t = \{(x, r) \in \mathbb{R} \times \mathbb{R}^+, |x - t| \leq r\},$$
$$S_t = C_t \Delta C_0,$$

where Δ denotes the symmetric difference between two sets. Let M be a symmetric α-stable random measure $(0 < \alpha < 2)$ with control measure

$$m(dx, dr) = r^{\beta - 2} dx dr \quad (0 < \beta < 1).$$

Takenaka process X is defined by

$$X(t) = \int_{\mathbb{R} \times \mathbb{R}^+} \mathbf{1}_{S_t}(x, r) M(dx, dr).$$

Process X is β/α-sssi.

Let us estimate $cov(|\Delta_{k,1}X|^{iM}, |\Delta_{0,1}X|^{iM})$. We apply Lemma 3.5. Set

$$f_t = \sum_{i=0}^{K} a_i \mathbf{1}_{S_{t+i}}.$$

Using $\sum_{i=0}^{K} a_i = 0$ and $\mathbf{1}_{A\Delta B} = (\mathbf{1}_A - \mathbf{1}_B)^2$

$$f_t = (1 - 2\mathbf{1}_{C_0}) \sum_{i=0}^{K} a_i \mathbf{1}_{C_{t+i}},$$

$$|f_t| = \left| \sum_{i=0}^{K} a_i \mathbf{1}_{C_{t+i}} \right|$$

$$= \left| \sum_{i=0}^{K} a_i (\mathbf{1}_{C_{t+i}} - \mathbf{1}_{C_t}) \right|$$

Therefore, we have to estimate, as $|k| \to +\infty$

$$\int_0^{+\infty} r^{\beta-2} \int_{\mathbb{R}} |f_k(x,r) f_0(x,r)|^{\alpha/2} dx dr.$$

When $k > 2r+K$, then $|f_k(x,r) f_0(x,r)| = 0$. When $k < 2r+K$, $|f_k(x,r) f_0(x,r)|$ can only be bounded by a constant. In particular, increasing the number of vanishing moments of the sequence (a_i) has no effect. Indeed, we have proved that there exists $C > 0$ such that

$$|cov(|\Delta_{k,1} X|^{iM}, |\Delta_{0,1} X|^{iM})| \leq C \int_{(k-K)/2}^{+\infty} r^{\beta-2} dr$$

$$\leq C' k^{\beta-1}.$$

Take M such that $0 < \beta/\alpha < 2\pi/(M \log(2))$. Then

$$\lim_{n \to +\infty} \widehat{H}_n = H \ (\mathbb{P}),$$

$$\widehat{H}_n - H = O_{\mathbb{P}}(n^{(\beta-1)/2}).$$

4. Identification of the distribution of $X(1)$

We just have seen that the self-similarity index H is identifiable. The aim of this section is to prove that the distribution of $X(1)$ is identifiable. Let us assume in this section the following

- The distribution of $X(1)$ is symmetric.
- Estimate $\widehat{H}_n$, with $L = 0$, and a given M, satisfies the following

$$\lim_{n \to +\infty} \log(n)(\widehat{H}_n - H) = 0 \ (\mathbb{P}). \qquad (4.1)$$

An inspection of the proofs shows that all the examples of the previous section satisfy condition (4.1). It follows that

$$n^{i\widehat{H}_n M} V_n(X) = e^{i(\widehat{H}_n - H)M \log(n)} n^{iHM} V_n(X)$$

$$\lim_{n \to +\infty} n^{i\widehat{H}_n M} V_n(X) = \mathbb{E}(\exp(iM \log |\Delta_{0,1} X|)) \ (\mathbb{P}).$$

Take the sequence $a_0 = -1, a_1 = 1$. The characteristic function of $\log|\Delta_{0,1}X| = \log|X(1)|$ is identifiable. Since $X(1)$'s distribution is symmetric, the distribution of $X(1)$ is identifiable.

5. Remarks

The following questions then arise. Are there other ways of estimating self-similarity than complex variations? Is one way optimal? There are clearly other estimators. For instance, one can use logarithmic variations (Please note that, in the following equation, logarithmic variations are used in another way that in [7, 9])

$$\widehat{H}_{\log,n} = \frac{V_{n/2}(\log, X) - V_n(\log, X)}{\log 2}.$$

An inspection of the proof indicates that Theorem 2.1 holds assuming the existence of logarithmic moments for X. Let now f be an arbitrary function from $\mathbb{R}^2$ onto $\mathbb{R}$ and g from $\mathbb{R}$ onto $\mathbb{C}$. Define the following estimate

$$\widehat{H}_{f,g,n} = f(V_{n/2}(g, X), V_n(g, X)).$$

What is the class $\mathcal{F}$ of admissible functions f and g? That is functions f and g leading to a consistent estimate under weak assumptions on X. Among this class $\mathcal{F}$, what are the best functions, say with respect to the mean square error? These questions turn out to be actually open and this paper offers no answer.

Acknowledgements

References

[1] BARDET, J.-M. AND TUDOR, C. (2010). A wavelet analysis of the Rosenblatt process: chaos expansion and estimation of the self-similarity parameter. *Stochastic Analysis and Applications* 1–25. MR2728168

[2] BREUER, P. AND MAJOR, P. (1983). Central limit theorems for non-linear functionals of Gaussian fields. *J. Mult. Anal.* 13, 425-441. MR0716933

[3] BROUSTE, A., LAMBERT-LACROIX, S. AND ISTAS, J. (2007). On Gaussian multifractal random fields simulations. *Journal of Statistical Software* **1 (23)**, 1–23.

[4] CHRONOPOULOU, A., TUDOR, C. AND VIENS, F. (2009). Variations and Hurst index estimation for a Rosenblatt process using longer filters. *Electronic Journal of Statistics* **3** 1393–1435. MR2578831

[5] Chronopoulou, A., Tudor, C. and Viens, F. (2011). Self-similarity parameter estimation and reproduction property for non-Gaussian Hermite processes. *Communications on Stochastic Analysis*, (to appear). MR2808541

[6] Coeurjolly, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statistical Inference for Stochastic Processes* **4** (2), 199–227. MR1856174

[7] Cohen, S. and Istas, J. (2005). An universal estimator of local self-similarity, unpublished.

[8] Gradshteyn, I. and Ryzhik, I. (2007). *Table of Integrals, Series, and Products*, Alan Jeffrey and Daniel Zwillinger (eds.), Seventh edition.

[9] Istas, J. (2011). Manifold indexed fractional fields. *ESAIM P & S* (to appear).

[10] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the Hölder index of a Gaussian process. *Ann. Inst. Poincaré* **33,4** 407–436. MR1465796

[11] Lacaux, C. and Loubes, J.-M. (2007). Hurst exponent estimation of fractional Lévy Motions. *Alea* **3** 143–164. MR2349806

[12] Nourdin, I., Nualart, D. and Tudor, C. (2010). Central and non-central limit theorems for weighted power variations of fractional Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist.* **46,4** 1055–1079. MR2744886

[13] Nourdin, I. and Peccati, G. (2011). Normal approximations with Malliavin calculus. From Stein's method to universality. *Cambridge University Press* (to appear).

[14] Pipiras, V., Taqqu, M. and Abry, P. (2007). Bounds for the covariance of functions of infinite variance stable random variables with applications to central limit theorems and wavelet-based estimation. *Bernoulli* **13,4** 1091–1123. MR2364228

[15] Samorodnitsky, G. and Taqqu, M. (1994). *Stable non-Gaussian random processes: stochastic models with infinite variance*, Chapman & Hall, New York. MR1280932

[16] Takenaka, S. (1991). Integral-geometric construction of self-similar stable processes. *Nagoya Math. J.* **123** 1–12. MR1126180

[17] Tudor, C. and Viens, F. (2009). Variations and estimators for the self-similarity order through Malliavin calculus. *The Annals of Probability* **37,6** 2093–2134. MR2573552