HAL open science

Expectiles for subordinated Gaussian processes with applications

Jean-François Coeurjolly, Hedi Kortas

To cite this version:

Jean-François Coeurjolly, Hedi Kortas. Expectiles for subordinated Gaussian processes with applications. Electronic Journal of Statistics , 2012, 6, pp.303-322. 10.1214/12-EJS674 . hal-00605467

HAL Id: hal-00605467

https://hal.science/hal-00605467

Submitted on 1 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Expectiles for subordinated Gaussian processes with applications

Jean-François Coeurjolly[1] and Hedi Kortas[2]

[1] *Corresponding author:* Jean-Francois.Coeurjolly@upmf-grenoble.fr
Laboratory Jean Kuntzmann, Department of Statistics, Grenoble University,
121 Avenue Centrale, 38040 Grenoble Cedex 9, France.
[2] Higher Institute of Management, Department of Quantitative Methods,
Sousse University, Tunisia.

July 4, 2011

Abstract

In this paper, we introduce a new class of estimators of the Hurst exponent of the fractional Brownian motion (fBm) process. These estimators are based on sample expectiles of discrete variations of a sample path of the fBm process. In order to derive the statistical properties of the proposed estimators, we establish asymptotic results for sample expectiles of subordinated stationary Gaussian processes with unit variance and correlation function satisfying $\rho(i) \sim \kappa|i|^{-\alpha}$ ($\kappa \in \mathbb{R}$) with $\alpha > 0$. Via a simulation study, we demonstrate the relevance of the expectile-based estimation method and show that the suggested estimators are more robust to data rounding than their sample quantile-based counterparts.

Keywords: expectiles; robustness; local shift sensitivity; subordinated Gaussian process; fractional Brownian motion.

1 Introduction

In the statistic literature, there has been a tremendous interest in analysis, estimation and simulation issues pertaining to the fractional Brownian motion (fBm) (Mandelbrot and Ness, 1968). This is due to the fact that the fBm process offers an adequate modeling framework for nonstationary self-similar stochastic processes with stationary increments and can be used to model stochastic phenomena relating to various fields of research. A fractional Brownian motion (fBm), denoted $\{B_H(t), t \in \mathbb{R}\}$ with Hurst exponent $0 < H < 1$, is a zero-mean continuous-time Gaussian stochastic process whose correlation function satisfies $\mathbb{E}[B_H(t)B_H(s)] = \frac{\sigma^2}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$ for all pairs $(t,s) \in \mathbb{R} \times \mathbb{R}$ and $\sigma^2 = \mathbb{E}(B_H(1)^2)$. The fBm is H-self-similar i.e., for all $\alpha > 0, B_H(\alpha t) \stackrel{d}{=} \alpha^H B_H(t)$, where $\stackrel{d}{=}$ means the equality of all its finite-dimensional probability distributions. The process corresponding to the first-order increments of the fBm is known as the fractional Gaussian noise (fGn) whose correlation function $\rho_H(i)$ is asymptotically of the order of $|i|^{2H-2}$ for large lag lengths i. In particular, for $1/2 < H < 1$, the correlations are not summable, i.e. $\sum_{i=-\infty}^{+\infty} \rho_H(i) = \infty$. This property is referred to as long-range dependence or long-memory whereas the case $0 < H < 1/2$ corresponds to short memory.

Several methods aimed at estimating the Hurst characteristic exponent or long-memory exponent have been developed. Among these statistical methods figure the Fourier-based methods such as the Whittle maximum likelihood estimator (see e.g. Beran (1994); Robinson (1995)) or the spectral regression based estimator (Beran, 1994). The wavelet estimators have been also extensively investigated either with an ordinary least squares (Flandrin, 1992), a weighted least

squares (see e.g. Abry et al. (2000, 2003); Bardet et al. (2000); Soltani et al. (2004)) or a maximum likelihood (see e.g. (Wornell and Oppenheim, 1992; Percival and Walden, 2000)) estimation schemes. Faÿ et al. (2009) present a deep analysis of Fourier and wavelet methods. Recently, the so-called discrete variations techniques (see e.g.Kent and Wood (1997); Istas and Lang (1997); Coeurjolly (2001)) have been introduced. Within this class of estimators, Coeurjolly (2008) proposed a new method based on sample quantiles to estimate the Hurst exponent in the more general setting of locally self-similar gaussian processes. This estimator has been proven robust when dealing with outliers (Achard and Coeurjolly, 2010). The latter, often encountered in real world applications, can induce a significant estimation bias. Actually, the advantage of quantiles is that they have a bounded gross-error-sensitivity (Hampel et al., 1986; Huber, 1981) allowing them to cope efficiently with the problem of outliers. Nevertheless, this is not the only problem faced when dealing with estimation issues. Indeed, data rounding is also a serious impediment. Data rounding is common in finance (Bijwaard and Franses, 2009; Rosenbaum, 2009), economics (Williams, 2006), computer science (Matthieu, 2006; Bois and Vignes, 1982) and computational physics (Vilmart, 2008) and can lead to several misinterpretations. Quantiles are unfortunately not robust against data rounding since their local shift sensitivity is unbounded, see Hampel et al. (1986); Huber (1981). Newey and Powell (1987) have introduced the so-called expectile which, although similar to quantile, has a bounded local shift-sensitivity and thus can handle the rounding issue.

In this paper, we derive a Bahadur-type representation for sample expectiles of a subordinated Gaussian process with unit variance and correlation function with hyperbolic decay. This allows us to investigate the statistical properties of a new discrete variations estimator of the Hurst exponent of the fBm process. In constructing this estimator, we rely mainly on the scale and location equivariance properties of expectiles (Newey and Powell, 1987). We will show via a simulation study the robustness of the proposed estimator against data rounding.

The remainder of this paper is structured as follows: Section 2 deals with asymptotic properties of sample expectiles for a class of subordinated stationary Gaussian processes with unit variance and correlation function satisfying $\rho(i) \sim \kappa|i|^{-\alpha}$ ($\kappa \in \mathbb{R}$) with $\alpha > 0$. A short simulation study is conducted to corroborate our theoretical findings. In Section 3, we discuss a sample expectile-based estimator of the Hurst exponent and derive its statistical properties. We then perform a simulation study in order to confirm the effectiveness of the suggested estimation method.

2 Expectiles for subordinated Gaussian processes

2.1 A few notation

Given some random variable Z with mean μ, F_Z is referred to the cumulative distribution function of Z and $\xi_Z(p)$ for $p \in (0,1)$ to its pth quantile. It is well-known that the pth quantile of a random variable Z can be obtained by minimizing asymmetrically the weighted mean absolute deviation

$$\xi_Z(p) := argmin_\theta \ \mathbf{E}\big[|p - \mathbf{1}_{Z\leq\theta}|.|Z-\theta|\big].$$

In order to limit the local shift sensitivity of the pth quantile, Newey and Powell (1987) defined the notion of expectile denoted by $E_Z(p)$ for some $p \in (0,1)$. Rather than an absolute deviation (function), a quadratic loss function is considered:

$$E_Z(p) := argmin_\theta \ \mathbf{E}\big[|p - \mathbf{1}_{Z\leq\theta}|.(Z-\theta)^2\big]. \tag{1}$$

We may note that the 50%-expectile if nothing else than the expectation of Z. Newey and Powell (1987) argued that providing $\mathbf{E}[Z] < +\infty$, then for every $p \in (0,1)$ the solution of (1) is unique on the set $I_{F_Z} := \{x \in \mathbb{R} : F_Z(x) \in (0,1)\}$. The expectile can also be defined as the solution of the equation $\mathbf{E}\big[|p - \mathbf{1}_{Z\leq\theta}|.(Z-\theta)\big] = 0$.

A key property of the expectile is that it is scale and location equivariant (Newey and Powell, 1987). The scale equivariance property means that for $Y = aZ$ where $a > 0$, the pth expectile of Y satisfies:

$$E_Y(p) = aE_Z(p) \tag{2}$$

The pth expectile is location equivariant in the sense that for $Y = Z + b$ where $b \in \mathbb{R}$, the pth expectile of Y is such that:

$$E_Y(p) = E_Z(p) + b \tag{3}$$

Now, let $\mathbf{Z} = (Z_1, \ldots, Z_n)$ be a sample of identically distributed random variables with common distribution F_Z, the sample expectile of order p is defined as:

$$\widehat{E}\,(p;\mathbf{Z}) := argmin_\theta \ \frac{1}{n}\sum_{i=1}^{n} |p - \mathbf{1}_{Z_i\leq\theta}|\,.\,(Z_i-\theta)^2\,.$$

2.2 Main result

In order to derive asymptotic results for Hurst exponent estimates based on expectiles, we have to provide asymptotic results for sample expectiles of nonlinear functions of (centered) subordinated stationary Gaussian processes with variance 1 and with correlation function decreasing

hyperbolically. This will be the setting of the rest of this section. Let $\{Y_i\}_{i=1}^{+\infty}$ be such a Gaussian process with correlation function $\rho(\cdot)$ satisfying $\rho(i) \sim \kappa|i|^{-\alpha}$ for $\kappa \in \mathbb{R}$ and $\alpha > 0$. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ a sample of n observations and $\mathbf{h}(\mathbf{Y}) = (h(Y_1), \ldots, h(Y_n))$ its subordinated version for some measurable function h. We wish to provide asymptotic results for the sample pth expectile defined by

$$\widehat{E}\left(p; \mathbf{h}(\mathbf{Y})\right) := argmin_\theta \ \frac{1}{n}\sum_{i=1}^{n} \left|p - \mathbf{1}_{h(Y_i)\leq\theta}\right| . \left(h(Y_i) - \theta\right)^2 . \tag{4}$$

Since the criterion is differentiable in θ, the sample pth expectile also satisfies the following estimating equation $\psi_n\left(\widehat{E}\left(p; \mathbf{h}(\mathbf{Y})\right); \mathbf{h}(\mathbf{Y})\right) = 0$ with

$$\psi_n\left(\theta; \mathbf{h}(\mathbf{Y})\right) := \frac{1}{n}\sum_{i=1}^{n} \left|p - \mathbf{1}_{h(Y_i)\leq\theta}\right| . \left(h(Y_i) - \theta\right). \tag{5}$$

In the following, we need the two following additional notation for $Y \sim \mathcal{N}(0,1)$

$$\begin{aligned}
\psi_{h(Y)}(\theta; p) &:= \mathbf{E}\left[\left|p - \mathbf{1}_{h(Y)\leq\theta}\right| .(h(Y) - \theta)\right] \\
\psi'_{h(Y)}(\theta; p) &:= -\mathbf{E}\left[\left|p - \mathbf{1}_{h(Y)\leq\theta}\right|\right] = -p(1 - F_{h(Y)}(\theta)) - (1-p)F_{h(Y)}(\theta),
\end{aligned}$$

the latter quantity corresponding to the derivative of $\psi_{h(Y)}(\cdot, p)$ if it is well-defined. Let us note that the pth expectile of $h(Y)$ satisfies $\psi_{h(Y)}(E_{h(Y)}; p) = 0$. We now present the assumption on the function h considered in our asymptotic result.

[A(h,p)] $h(\cdot)$ is a measurable function such that $\mathbf{E}h(Y)^2 < +\infty$ and such that the function $\psi_{h(Y)}(\cdot, p)$ is continuously differentiable in a neighborhood of $E_{h(Y)}(p)$ with negative derivative at this point.

Such an assumption is in particular satisfied under the following one:
[A′(h)] $h(\cdot)$ is a measurable function such that $\mathbf{E}h(Y)^2 < +\infty$, h is not "flat", i.e. for all $\theta \in \mathbb{R}$ the set $\{y \in \mathbb{R} : h(y) = \theta\}$ has null Lebesgue measure.

Indeed, if h satisfies **[A′(h)]** then $\psi(\cdot, p)$ is differentiable in θ. And since, $E_{h(Y)}(p)$ belongs to the set $I_{h(Y)} = \{x \in \mathbb{R} : F_{h(Y)}(x) \in (0,1)\}$, $\psi'(E_{h(Y)}(p); p)$ is necessarily negative. For the purpose of this paper, our main result will be applied with $h(\cdot) = |\cdot|^\beta$ (with $\beta > 0$) or $h(\cdot) = \log|\cdot|$ which obviously satisfy **[A′(h)]**.

The nature of the asymptotic result will depend on the correlation structure of the Gaussian process and on the Hermite rank, $\tau(p,\theta)$ of the function

$$\widetilde{\psi}(t;p,\theta) := \left|p - \mathbf{1}_{h(t)\leq\theta}\right|.(h(t)-\theta) - \psi_{h(Y)}(\theta;p).$$

We recall that the Hermite rank (see *e.g.* Taqqu (1977)) corresponds to the smallest integer such that the coefficient in the Hermite expansion of the considered function is not zero. For the sake of simplicity, assume that the Hermite rank of this function depends neither on θ nor p and denote it simply by τ. Again, this could be weakened since we believe that the next result could be proved with the following Hermit rank: $\inf_{\theta\in\mathcal{V}(E_{h(Y);p})}\tau(p,\theta)$. As an example, the Hermite rank of $\widetilde{\psi}(\cdot,p,\theta)$ is 1 for $h(\cdot)=\cdot$ and $(p,\theta)\in(0,1)\times\mathbb{R}$ and 2 for $h(\cdot)=|\cdot|^{\beta}$ $(\beta>0)$ or $\log|\cdot|$ for $(p,\theta)\in(0,1)\times\mathbb{R}^{+}\setminus\{0\}$. We now present our main result stating a Bahadur type representation for the sample pth expectile of a subordinated Gaussian process.

Theorem 1 *Let $\{Y_i\}_{i=1}^{+\infty}$ a (centered) stationary Gaussian process with variance 1 and correlation function satisfying $\rho(i)\sim\kappa|i|^{-\alpha}$ $(\kappa\in\mathbb{R})$, as $|i|\to+\infty$ with $\alpha>0$ and with a function h satisfying **[A(h,p)]**. Let $\mathbf{h}(\mathbf{Y})=(h(Y_1),\ldots,h(Y_n))$ a sample of n observations of the subordinated process, then, for all $p\in(0,1)$*

$$\widehat{E}\left(p;\mathbf{h}(\mathbf{Y})\right) - E_{h(Y)}(p) = -\frac{\psi_n\left(E_{h(Y)}(p);\mathbf{h}(\mathbf{Y})\right)}{\psi'\left(E_{h(Y)}(p);p\right)} \quad + \quad o_P(r_n), \tag{6}$$

where the sequence $r_n = r_n(\alpha,\tau)$ is defined by

$$r_n = \begin{cases} n^{-1/2} & \text{if} \quad \alpha\tau>1 \\ n^{-1/2}\log(n) & \text{if} \quad \alpha\tau=1 \\ n^{-\alpha\tau/2} & \text{if} \quad \alpha\tau>1. \end{cases}$$

Proof. Let us simplify the notation for sake of conciseness: let $\widehat{E}=\widehat{E}(p;\mathbf{h}(\mathbf{Y}))$, $E=E_{h(Y)}(p)$, $\psi_n(E)=\psi_n(E_{h(Y)}(p);\mathbf{h}(\mathbf{Y}))$ and $\psi'(E)=\psi'_{h(Y)}(E_{h(Y)}(p);p)$. The first thing to note is that the sequence r_n corresponds to the short-range or long-range dependence characteristic of the sequence $\widetilde{\psi}(h(Y_1);p,\theta),\ldots,\widetilde{\psi}(h(Y_n);p,\theta)$. More precisely r_n^2 corresponds to the asymptotic behavior of $\mathbf{E}\psi_n(E)^2$. Indeed, if $(c_j)_{j\geq0}$ denotes the sequence of the Hermite coefficients of the expansion of $\widetilde{\psi}(\cdot;p,E)$ in Hermite polynomials (denoted by $(H_j(t))_{j\geq0}$ and normalized in such a way that $E[H_j(Y)H_k(Y)]=j!\delta_{jk}$), we may have using standard developments on Hermite

polynomials (see *e.g.* Taqqu (1977))

$$\begin{aligned}
\mathbf{E}\psi_n(E)^2 &= \frac{1}{n^2}\sum_{i,j=1}^{n}\mathbf{E}\left[\widetilde{\psi}(Y_i;p,E)\widetilde{\psi}(Y_j;p,E)\right] \\
&= \frac{1}{n^2}\sum_{i,j=1}^{n}\sum_{k_1,k_2\geq 0}\frac{c_{k_1}c_{k_2}}{k_1!k_2!}\mathbf{E}\left[H_{k_1}(Y_i)H_{k_2}(Y_j)\right] \\
&= \frac{1}{n^2}\sum_{i,j=1}^{n}\sum_{k\geq\tau}\frac{c_k^2}{k!}\rho(j-i)^k \\
&= \mathcal{O}\Big(\underbrace{\frac{1}{n}\sum_{|i|\leq n}|\rho(i)|^\tau}_{=:\rho_n}\Big) = \mathcal{O}(r_n^2). \qquad (7)
\end{aligned}$$

Let us define $V_n := r_n^{-1}(\widehat{E}-E)$ and $W_n(E) := -r_n^{-1}\psi_n(E)/\psi'(E)$. We just have to prove that $V_n - W_n(E)$ converges in probability to 0 as $n\to+\infty$. The proof is based on the application of Lemma 1 of Ghosh (1971) which consists in satisfying the two following conditions:

(a) for all $\delta>0$, there exists $\varepsilon=\varepsilon(\delta)$ such that $P(|W_n(E)|>\varepsilon)<\delta$.

(b) for all $y\in\mathbb{R}$ and for all $\varepsilon>0$

$$\lim_{n\to+\infty}P(V_n\leq y, W_n(E)\geq y+\varepsilon) = \lim_{n\to+\infty}P(V_n\geq y+\varepsilon, W_n(E)\leq y) = 0.$$

(a) is in particular fulfilled if we prove that $\mathbf{E}W_n(E)^2=\mathcal{O}(1)$ which follows from (7) since $\mathbf{E}W_n(E)^2 = \psi'(E)^{-2}r_n^{-2}\mathbf{E}\psi_n(E)^2 = r_n^{-2}\times\mathcal{O}(\rho_n)=\mathcal{O}(1)$.

(b) We consider only the first limit. The second one follows similar developments. We first state that the map $\psi_n(\cdot)$ is decreasing. Indeed, let $\theta\leq\theta'$ and denote by $Z_i(\theta)$ the variable $|p-\mathbf{1}_{h(y_i)\leq\theta)}|.(h(Y_i)-\theta)$. If $h(Y_i)\leq\theta$ or $h(Y_i)>\theta$, we obviously get $Z_i(\theta)>Z_i(\theta')$ a.s. leading to the decreasing of $\psi_n(\cdot)$. And in the in between case, $Z_i(\theta)-Z_i(\theta')=p(\theta'-\theta)+\theta'-h(Y_i)\geq 0$ which leads to the same conclusion. Let $y\in\mathbb{R}$, then also using the fact that $\psi_n(\widehat{E})=\psi(E)=0$ and $\psi'(E)<0$, we derive

$$\begin{aligned}
\{V_n\leq y\} &= \{\widehat{E}\leq y\times r_n+E\} \\
&= \{\psi_n(\widehat{E})\geq\psi_n(y\times r_n+E)\} \\
&= \{\psi(y\times r_n+E)-\psi_n(y\times r_n+E)\geq\psi(y\times r_n+E)-\psi(E)\} \\
&= \{W_n(y\times r_n+E)\leq y_n\},
\end{aligned}$$

where $y_n = r_n^{-1}\psi'(E)^{-1}(\psi(y\times r_n+E)-\psi(E))$. Under the assumption **[A(h,p)]**, $y_n\to y$ as $n\to+\infty$. Now, let $U_n := \psi'(E)\left(W_n(E)-W_n(y\times r_n+E)\right)$, explicitly given by

$$U_n = \frac{1}{nr_n}\sum_{i=1}^{n}\Big(\widetilde{\psi}(Y_i;p,E+y\times r_n)-\widetilde{\psi}(Y_i;p,E)\Big).$$

Let $c_{j,n}$ the jth Hermite coefficient of the function $r_n^{-1}\left(\widetilde{\psi}(t;p,E+y\times r_n)-\widetilde{\psi}(t;p,E)\right)$, then under the assumption **[A(h,p)]** and from the dominated convergence theorem we can prove that

$$c_{j,n} \overset{n\to+\infty}{\longrightarrow} y\,\mathbf{E}\left[|p-\mathbf{1}_{h(Y)\leq E}|H_j(Y)\right] =: \widetilde{c}_j.$$

Therefore for n large enough,

$$\begin{aligned}
\mathbf{E}[U_n^2] &= \frac{1}{n^2}\sum_{i,j=1}^{n}\sum_{k\geq\tau}\frac{c_{k,n}^2}{k!}\rho(j-i)^k \\
&\leq \frac{2}{n}\sum_{|i|\leq n}\sum_{k\geq\tau}\frac{\widetilde{c}_k^2}{k!}\rho(i)^\tau \\
&= \mathcal{O}(\rho_n)=\mathcal{O}(r_n^2)
\end{aligned}$$

which leads to the convergence of U_n to 0 in probability. For all $\varepsilon>0$, there exists $n_0(\varepsilon)$ such that for all $n\geq n_0(\varepsilon)$, $y_n\leq y+\varepsilon/2$. Therefore for $n\geq n_0(\varepsilon)$

$$\begin{aligned}
P(V_n\leq y\ ,\ W_n\geq y+\varepsilon) &= P(W_n(y\times r_n+E)\leq y_n\ ,\ W_n\geq y+\varepsilon) \\
&\leq P(W_n(y\times r_n+E)\leq y+\varepsilon/2\ ,\ W_n(E)\geq y+\varepsilon) \\
&\leq P\left(|W_n(y\times r_n+E)-W_n(E)|\geq\varepsilon/2\right) \\
&\overset{n\to+\infty}{\to} 0,
\end{aligned}$$

which ends the proof. ■

In the case of short-range dependence, i.e. $\alpha\tau>1$ then, using the Bahadur type representation of expectiles, we derive immediately the following asymptotic normality for the sample expectile and some generalisations. This result is based on standard central limit theorem for means of subordinated Gaussian stationary processes (Taqqu, 1977; Arcones, 1994).Therefore the proof is omitted.

Corollary 2

(i) Under the assumptions of Theorem 1 with $p\in(0,1)$ and $\alpha\tau>1$, then as $n\to+\infty$

$$\sqrt{n}\left(\widehat{E}\left(p;\mathbf{h}(\mathbf{Y})\right)-E_{h(Y)}(p)\right)\xrightarrow{d}\mathcal{N}(0,\sigma^2(p)),$$

where

$$\sigma^2(p)=\frac{1}{\psi'\left(E_{h(Y)}(p);p\right)^2}\sum_{i\in\mathbb{Z}}\sum_{k\geq\tau}\frac{c_k(p)^2}{k!}\rho(i)^k$$

and where $c_k(p)$ is the kth Hermite coefficient of the expansion of the function $\psi(h(\cdot);E_{h(Y)}(p);p)$ in Hermite polynomials.

(ii) Let $\{Y_i^1\}_{i=1}^{+\infty}$ *and* $\{Y_i^2\}_{i=1}^{+\infty}$ *two (centered) stationary Gaussian processes with variances 1 and correlation functions (resp. cross-correlation functions)* ρ^1, ρ^2 *(resp.* ρ^{12}, ρ^{21}*) decreasing hyperbolically with exponents* α^1, α^2 *(resp.* α^{12}, α^{21}*). Let* $p \in (0,1)$*,* h *a function satisfying* ***[A(h,p)]*** *and let* $\mathbf{h}(\mathbf{Y^1})$ *and* $\mathbf{h}(\mathbf{Y^2})$ *be the samples of* n *observations of the two subordinated samples. If* $\min(\alpha^1, \alpha^2, \alpha^{12}, \alpha^{21}) \times \tau > 1$*, then as* $n \to +\infty$

$$\sqrt{n}\left(\widehat{E}\left(p; \mathbf{h}(\mathbf{Y^1})\right) - E_{h(Y)}(p), \widehat{E}\left(p; \mathbf{h}(\mathbf{Y^2})\right) - E_{h(Y)}(p)\right)^T \xrightarrow{d} \mathcal{N}(0, \underline{\mathbf{\Sigma}}).$$

where $\underline{\mathbf{\Sigma}}$ *is the* $(2,2)$ *matrix with entries* Σ_{ab} *for* $a, b = 1, 2$ *given by*

$$\Sigma_{ab} = \frac{1}{\psi'\left(E_{h(Y)}(p); p\right)^2} \sum_{i \in \mathbb{Z}} \sum_{k \geq \tau} \frac{c_k(p)^2}{k!} \rho^{ab}(i)^k. \tag{8}$$

As it was established for sample quantiles (Coeurjolly, 2008), a non standard limit towards a Rosenblatt process is expected in the other cases ($\alpha\tau \leq 1$). This case is not considered here.

2.3 Simulations

To illustrate a part of the previous results, we propose a short simulation study in this section. The latent stationary Gaussian process we consider here is the fractional Gaussian noise with variance 1, which is obtained by taking the discretized increments from a fractional Brownian motion. The correlation function of the fractional Gaussian noise with Hurst parameter (or self-similarity parameter) $H \in (0,1)$ satisfies the hyperbolic decreasing property required in Theorem 1 with $\alpha = 2 - 2H$. Discretized sample paths of fractional Brownian motion can be generated exactly using the embedding circulant matrix method popularized by Wood and Chan (1994) (see also Coeurjolly (2000)) which is implemented in the `R` package `dvfBm`.

Figures 1 and 2 illustrate the convergence of the sample expectiles. Three h functions are considered: $h(\cdot) = (\cdot)$, $(\cdot)^2$ and $\log|\cdot|$. The related Hermite rank of the function $\widetilde{\psi}$ is respectively 1,2 and 2 for these three h functions. The sample size of the simulation is fixed to $n = 500$. We can claim the convergence of the sample expectile $\widehat{E}(p; h(Y))$ towards $E_{h(Y)}(p)$ for all the values of α (or H), p and for the three functions h considered. If we focus on $h(\cdot) = (\cdot)$, we can also remark a higher variance of the sample estimates for $\alpha = 0.6$ compared to $\alpha = 1.4$. This is in agreement with the theory since for $\alpha = 0.6$, $\alpha\tau = 0.6 < 1$ and the rate of convergence is lower than $n^{-1/2}$ which means an increasing of the variance. For the two other functions considered, then $\alpha\tau$ is always greater than 1 (it equals either 2.8 or 1.2 in our simulations) and we do not observe such an increasing of the variance.

To put emphasis on this last point, Figure 3 shows in log-scale the average (over the 9 order of expectiles considered in the simulation, i.e. $p = 0.1, \ldots, 0.9$) of the empirical variances in terms of n for the three h functions and for the two values of $\alpha = 0.6$ and $\alpha = 1.4$. We clearly observe that as soon as $\alpha\tau > 1$, the slope of the curves is close to -1 which agrees with the result presented in Corollary 2 for example. When $h(\cdot) = 1$ and $\alpha = 0.6$, we observe that the slope is about -0.6 which seems to agree with the convergence in $n^{-\alpha\tau}$ which is expected from Theorem 1.

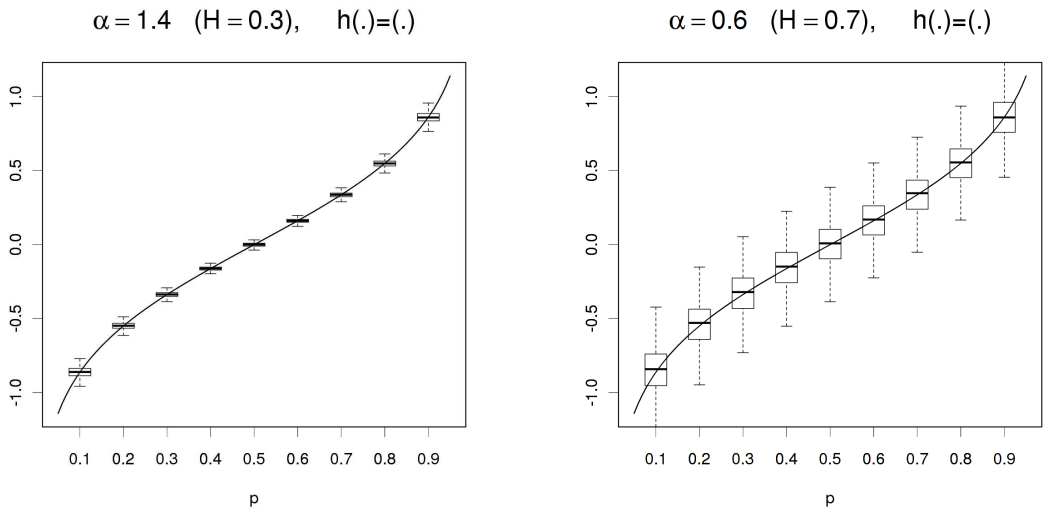

Figure 1: Boxplots of sample expectiles for expectiles of order $p = 0.1, \ldots, 0.9$ based on $m = 500$ replications of fractional Gaussian noise with length $n = 500$ and with Hurst parmeter $H = 0.3$ (left, $\alpha = 1.4$) and $H = 0.7$ (right, $\alpha = 0.6$). The h functions considered here is the identity function (with Hermite rank 1). The curves correspond to the theoretical expectile functions for $Y \sim \mathcal{N}(0, 1)$.

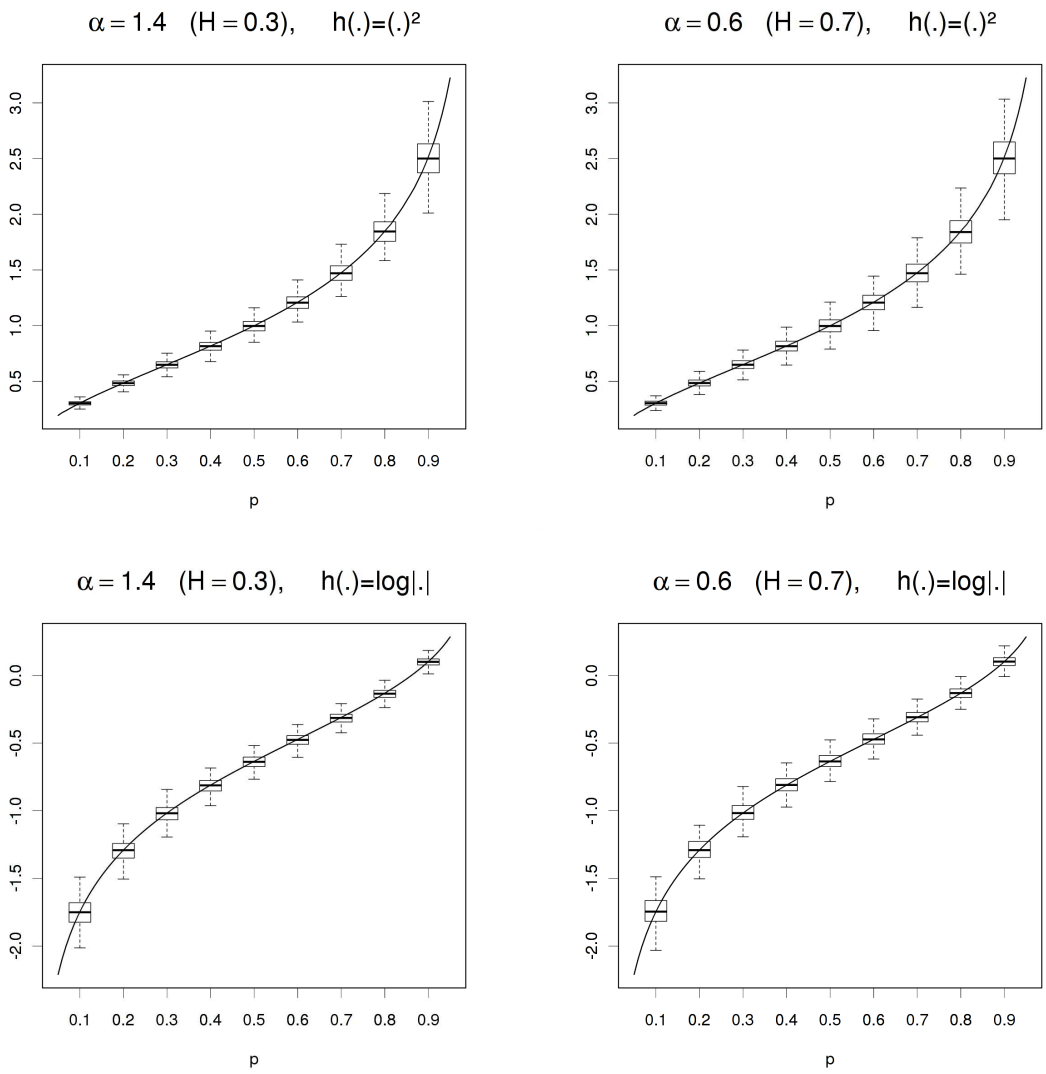

Figure 2: Boxplots of sample expectiles for expectiles of order $p = 0.1, \ldots, 0.9$ based on $m = 500$ replications of fractional Gaussian noise with length $n = 500$ and with Hurst parmeter $H = 0.3$ (left, $\alpha = 1.4$) and $H = 0.7$ (right, $\alpha = 0.6$). Two h functions with Hermite rank 2 have been considered here: $h(\cdot) = (\cdot)^2$ (top) and $h(\cdot) = \log|\cdot|$ (bottom). The curves correspond to the theoretical expectile functions for Y^2 (middle) and $\log|Y|$ (bottom) where $Y \sim \mathcal{N}(0, 1)$.

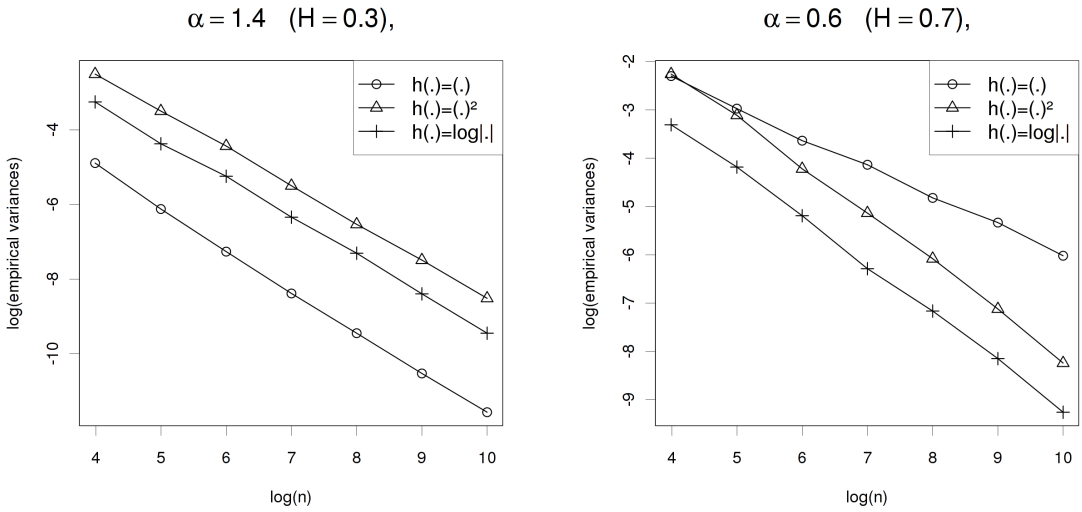

Figure 3: Means of empirical variances of sample expectiles in terms of n in log-scale based on $m = 500$ replications of fractional Gaussian noise with parameters $H = 0.3$ (left, $\alpha = 1.4$) and $H = 0.7$ (right, $\alpha = 0.6$). More precisely, we consider the vector of probability $(0.1, \ldots, 0.9)$ for the orders of the expectiles and we compute $\widehat{\sigma}_n^2 = 1/9 \times \sum_{i=1}^{9} \widehat{\sigma}_{i,n}^2$ where $\widehat{\sigma}_{i,n}^2$ is the empirical variance for the expectile with order $i/10$ for the sample size n. Three choices of h functions have been considered: $h(\cdot) = (\cdot), (\cdot)^2$ and $\log|\cdot|$.

3 Estimation of the Hurst exponent using sample expectiles and discrete variations

3.1 Estimation method and asymptotic results

Let $\mathbf{X} = (X(i))_{i=1,\ldots,n}$ be a discretized version of a fractional Brownian motion process and let $\mathbf{a}$ be a filter of length $\ell+1$ and of order $\nu \geq 1$ with real components i.e.:

$$\sum_{q=0}^{\ell} q^j a_q = 0, \text{ for } j = 0, \ldots, \nu-1 \qquad \text{and} \qquad \sum_{q=0}^{\ell} q^\nu a_q \neq 0.$$

Define also $\mathbf{X^a}$ to be the series obtained by filtering $\mathbf{X}$ with $\mathbf{a}$, then:

$$X^{\mathbf{a}}(i) = \sum_{q=0}^{\ell} a_q X(i-q), \qquad \text{for } i \geq \ell+1$$

and $\tilde{\mathbf{X}}^{\mathbf{a}}$ as the normalized vector of $\mathbf{X^a}$, i.e.:

$$\tilde{\mathbf{X}}^{\mathbf{a}} = \frac{\mathbf{X^a}}{\mathbf{E}((X^{\mathbf{a}}(1))^2)^{1/2}}.$$

It should be noticed here that the filtering operation allows to decorrelate the increments of the discretized version of the fractional Brownian motion process. Indeed, it may be proved (see *e.g.* Coeurjolly (2001)) that: $\rho_H^a(i) \sim k_H |i|^{2H-2\nu}$ as $|i| \to +\infty$.

Consider the sequence $(\mathbf{a}^m)_{m\geq 1}$ defined by:

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

which is the filter $\mathbf{a}$ dilated m times. It has been shown in Coeurjolly (2001, 2008) that:

$$\tilde{\mathbf{X}}^{\mathbf{a^m}} = \frac{\mathbf{X^{a^m}}}{\sigma_m}$$

where $\sigma_m^2 = m^{2H}\sigma^2\kappa_H^{\mathbf{a}}$ and $\kappa_H^{\mathbf{a}} = \frac{-1}{2}\sum_{q,q'=0}^{\ell} a_q a_{q'} |q-q'|^{2H}$.

The following proposition allows us to construct an ordinary least squares (OLS) estimator of the Hurst exponent H of a fBm process based on sample expectiles.

Proposition 3 *Let $\widehat{E}\left(p; \mathbf{h}(\mathbf{X^{a^m}})\right)$ and $\widehat{E}\left(p; \mathbf{h}(\tilde{\mathbf{X}}^{\mathbf{a^m}})\right)$ be the pth order sample expectiles for the filtered series $\mathbf{h}(\mathbf{X^{a^m}})$ and $\mathbf{h}(\tilde{\mathbf{X}}^{\mathbf{a^m}})$ respectively. Here two positive functions $h(\cdot)$ are considered, namely: $h(\cdot) = |\cdot|^\beta$ for $\beta > 0$ and $h(\cdot) = \log|\cdot|$. We have:*

$$\widehat{E}\left(p; |\mathbf{X}^{\mathbf{a}^m}|^\beta\right) = \sigma_m^\beta \widehat{E}\left(p; |\tilde{\mathbf{X}}^{\mathbf{a}^m}|^\beta)\right) \tag{9}$$

and

$$\widehat{E}\left(p;\log|\mathbf{X}^{\mathbf{a}^m}|\right)=\frac{1}{2}\log(\sigma_m^2)+\widehat{E}\left(p;\log|\tilde{\mathbf{X}}^{\mathbf{a}^m}|\right). \tag{10}$$

Proof.

We have:

$$\begin{aligned}\widehat{E}\left(p;|\mathbf{X}^{\mathbf{a}^m}|^\beta\right) &= argmin_\theta\frac{1}{n-m\ell}\sum_{i=m\ell}^{n-1}|p-\mathbf{1}_{\{|X^{\mathbf{a}^m}(i)|^\beta\leq\theta\}}|.(|X^{\mathbf{a}^m}(i)|^\beta-\theta)^2\\ &= argmin_\theta\frac{1}{n-ml}\sum_{i=ml}^{n-1}|p-\mathbf{1}_{\{|\tilde{X}^{\mathbf{a}^m}(i)|^\beta\leq\frac{\theta}{\sigma_m^\beta}\}}|.(|\tilde{X}^{\mathbf{a}^m}(i)|^\beta-\frac{\theta}{\sigma_m^\beta})^2.\end{aligned}$$

Setting $\theta'=\frac{\theta}{\sigma_m^\beta}$, the proof of the first relation (9) follows easily. Using the same methodology, we can demonstrate the result given by equation (10). ■

Remark 1 *It should be stressed here that the scaling relationship relating the* theoretical pth expectiles *for the series* $\mathbf{h}(\mathbf{X}^{\mathbf{a}^m})$ *and* $\mathbf{h}(\tilde{\mathbf{X}}^{\mathbf{a}^m})$ *can be obtained directly using the scale equivariance property (2) for* $h(\cdot)=|\cdot|^\beta$ *and the location equivariance property (3) for* $h(\cdot)=\log|\cdot|$.

Now applying the logarithmic transformation to both sides of (9), we get:

$$\log\widehat{E}\left(p;|\mathbf{X}^{\mathbf{a}^m}|^\beta\right)=\beta H\log(m)+\log\left(\sigma^\beta(\kappa_H^{\mathbf{a}})^{\beta/2}E_{|Y|^\beta}(p)\right)+\log\left(\frac{\widehat{E}\left(p;|\tilde{\mathbf{X}}^{\mathbf{a}^m}|^\beta\right)}{E_{|Y|^\beta}(p)}\right). \tag{11}$$

On the other hand, (10) can be reformulated in the following way:

$$\widehat{E}\left(p;\log|\mathbf{X}^{\mathbf{a}^m}|\right)=H\log(m)+\frac{1}{2}\log(\sigma^2\kappa_H^{\mathbf{a}})+E_{\log|Y|}(p)+\left(\widehat{E}\left(p;\log(|\tilde{\mathbf{X}}^{\mathbf{a}^m}|)\right)-E_{\log|Y|}(p))\right). \tag{12}$$

It is noteworthy here that we expect that $\log\widehat{E}\left(p;|\tilde{\mathbf{X}}^{\mathbf{a}^m}|^\beta\right)/E_{|Y|^\beta}(p)$ and $\widehat{E}\left(p;\log(|\tilde{\mathbf{X}}^{\mathbf{a}^m}|^\beta)\right)-E_{\log|Y|}(p)$ to converge towards 0 as $n\to\infty$. Hence, based on equations (11) and (12), we opt for an OLS regression scheme. This allows to derive the two following estimators of the hurst index defined by:

$$\widehat{H}^\beta=\frac{\mathbf{A}^T}{\beta||\mathbf{A}||^2}\left(\log\widehat{E}\left(p;|\mathbf{X}^{\mathbf{a}^m}|^\beta\right)\right)_{m=1,\ldots,M}, \tag{13}$$

and

$$\widehat{H}^{\log}=\frac{\mathbf{A}^T}{||\mathbf{A}||^2}\left(\widehat{E}\left(p;\log|\mathbf{X}^{\mathbf{a}^m}|\right)\right)_{m=1,\ldots,M}, \tag{14}$$

where $\mathbf{A}$ is the vector of length M with components $A_m=\log m-\frac{1}{M}\sum_{m=1}^M\log(m)$, $m=1,\ldots,M$ for some $M\geq 2$ whereas $||\mathbf{z}||$ for some vector $\mathbf{z}$ of length d designates the norm defined by $\left(\sum_{i=1}^d z_i^2\right)^{1/2}$. Notice here that $\widehat{H}^\beta$ and $\widehat{H}^{\log}$ do not depend on σ^2.

We would like to put the stress on the fact that (13) and (14) are really similar to the ones developed in Coeurjolly (2001, 2008). Indeed, the standard procedure developed in Coeurjolly (2001) simply consists in replacing the sample expectile by the sample variance (this method will be denoted by ST in Section 3.2). To deal with outliers, the procedure developed in Coeurjolly (2008) consists in replacing the sample expectile by either the sample median of $(\mathbf{X}^{\mathbf{a}^m})^2$ or the trimmed-means of $(\mathbf{X}^{\mathbf{a}^m})^2$. These two last methods are denoted by MED and TM in Section 3.2.

Now, we state the asymptotic results for these new estimates based on expectiles.

Proposition 4 *Let* $\mathbf{a}$ *a filter with order* $\nu \geq 2$, $p \in (0,1)$, $\beta > 0$ *then as* $n \to +\infty$, $\widehat{H}^\beta$ *and* $\widehat{H}^{\log}$ *converge in probability to* H. *Moreover, the following convergences in distribution hold*

$$\sqrt{n}\left(\widehat{H}^\beta - H\right) \xrightarrow{d} \mathcal{N}(0, \sigma_\beta^2) \quad \text{and} \quad \sqrt{n}\left(\widehat{H}^{\log} - H\right) \xrightarrow{d} \mathcal{N}(0, \sigma_{\log}^2),$$

where

$$\sigma_\beta^2 = \frac{1}{E_{|Y|^\beta}(p)^2} \times \frac{\mathbf{A}^T \underline{\mathbf{\Sigma}}^\beta \mathbf{A}}{\beta^2 \|\mathbf{A}\|^4} \quad \text{and} \quad \sigma_{\log}^2 = \frac{\mathbf{A}^T \underline{\mathbf{\Sigma}}^{\log} \mathbf{A}}{\|\mathbf{A}\|^4}$$

and where the (M, M) *matrices* $\underline{\mathbf{\Sigma}}^\beta$ *and* $\underline{\mathbf{\Sigma}}^{\log}$ *are defined by (8).*

Proof. We only provide a sketch of the proof. We claim that once Theorem 1 and Corollary 2 are established, the obtention of convergences stated in Proposition 4 are semi-routine. First of all, let us notice that

$$\widehat{H}^\beta - H = \frac{\mathbf{A}^T}{\beta \|\mathbf{A}\|^2} \left(\log \left(\frac{\widehat{E}\left(p; |\widetilde{\mathbf{X}}^{\mathbf{a}^m}|^\beta\right)}{E_{|Y|^\beta}} \right) \right)_{m=1,\ldots,M} \tag{15}$$

and

$$\widehat{H}^{\log} - H = \frac{\mathbf{A}^T}{\beta \|\mathbf{A}\|^2} \left(\widehat{E}\left(p; \log |\widetilde{\mathbf{X}}^{\mathbf{a}^m}|\right) - E_{\log|Y|}(p) \right)_{m=1,\ldots,M}. \tag{16}$$

Since the functions $|\cdot|^\beta$ and $\log|\cdot|$ have Hermite rank 2 and since the correlation function of the stationary sequence $\widetilde{\mathbf{X}}^{\mathbf{a}^m}$ decreases hyperbolically with an exponent $\alpha = 2\nu - 2H$ then for any $m \in \{1, \ldots, M\}$, Theorem 1 holds with $r_n = n^{-1/2}$ (since $\alpha\tau > 1$ for all $H \in (0,1)$). This ensures the convergence in probability of the new estimates.

The cross-correlation between $\widetilde{X}^{\mathbf{a}^{m_1}}$ and $\widetilde{X}^{\mathbf{a}^{m_2}}$ is defined by

$$\rho_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_2}}(j) = \frac{\pi_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_2}}(j)}{\pi_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_1}}(0)^{1/2} \pi_H^{\mathbf{a}^{m_2},\mathbf{a}^{m_2}}(0)^{1/2}} \quad \text{with} \quad \pi_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_2}}(j) = \sum_{q,r=0}^{\ell} a_q a_r |m_1 q - m_2 r + j|^{2H}.$$

Lemma 1 in Coeurjolly (2008) states that for all m_1, m_2 the correlation function $\rho_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_2}}$ is also decreasing hyperbolically with an exponent $\alpha = 2\nu - 2H$, then Corollary 2 may be applied to prove that

$$\left(\widehat{E}\left(p; |\widetilde{\mathbf{X}}^{\mathbf{a}^m}|^\beta\right) - E_{|Y|^\beta}\right)_{m=1,\ldots,M} \xrightarrow{d} \mathcal{N}(0, \underline{\mathbf{\Sigma}}^\beta) \tag{17}$$

and

$$\left(\widehat{E}\left(p; \log|\widetilde{\mathbf{X}}^{\mathbf{a}^m}|\right) - E_{\log|Y|}\right)_{m=1,\ldots,M} \xrightarrow{d} \mathcal{N}(0, \underline{\mathbf{\Sigma}}^{\log}), \tag{18}$$

where according to (8), the (M, M) matrices $\underline{\mathbf{\Sigma}}^\beta$ and $\underline{\mathbf{\Sigma}}^{\log}$ are respectively defined by

$$\Sigma^\beta_{m_1 m_2} = \frac{1}{\psi'\left(E_{|Y|^\beta}(p); p\right)^2} \sum_{i\in\mathbb{Z}} \sum_{k\geq 2} \frac{c_k^\beta(p)^2}{k!} \rho_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_2}}(i)^k \tag{19}$$

$$\Sigma^{\log}_{m_1 m_2} = \frac{1}{\psi'\left(E_{\log|Y|}(p); p\right)^2} \sum_{i\in\mathbb{Z}} \sum_{k\geq 2} \frac{c_k^{\log}(p)^2}{k!} \rho_H^{\mathbf{a}^{m_1},\mathbf{a}^{m_2}}(i)^k, \tag{20}$$

where $(c_k^\beta)_{k\geq 2}$ and $(c_k^{\log})_{k\geq 2}$ are respectively the Hermite coefficients of the functions $|\cdot|^\beta$ and $\log|\cdot|$. The convergences (17) and (18) combined with (15) and (16) and the use of the delta-method (for the convergence of $\widehat{H}^\beta$) end the proof. ■

3.2 A short simulation study

In this section, we investigate the interest of the new estimators based on expectiles. We consider three different models in our simulations.

(a) *standard fBm*: non-contaminated fractional Brownian motion.

(b) *fBm with additive outliers*: we contaminate 5% of the observations of the increments of the fractional Brownian motion with an independent Gaussian noise such that the SNR of the considered components equals $-20Db$.

(c) *rounded fBm*: we assume the data are given by the integer part of a discretized sample path of an original fBm.

To fix ideas, Figure 4 provides some examples of discretized sample paths of standard and contaminated fBm. The simulation results are presented in Tables 1 and 2. For these simulations, as suggested in Coeurjolly (2001), we chose the filter $a = d4$ corresponding to the wavelet Daubechies filter with order 4 (see Daubechies (1992)) and the maximum number of dilated filters $M = 5$. Also, in other simulations not presented here, we have observed that the estimates $\widehat{H}^\beta$ perform better than $\widehat{H}^{\log}$ and, among all possible choices of β, the value $\beta = 2$ seems to be

a good compromise. Therefore, we present only the result for this latter estimator, that is $\widehat{H}^{\beta}$ with $\beta = 2$.

In a first step, we had observed a quite large sensitivity to the value of the probability p defining the expectile. In order to have an efficient data-driven procedure, we propose to choose the probability parameter p via a Monte-Carlo approach as follows:

1. Estimate the parameters H and σ^2 using the standard method (the estimation of σ^2 is not described here but it may be found for example in Coeurjolly (2001)). Denote these estimates $\widehat{H}_0$ and $\widehat{\sigma}_0^2$.

2. Generate $B = 100$ contaminated fBm with Hurst parameter $\widehat{H}_0$ and scaling coefficient $\widehat{\sigma}_0^2$, define a grid of probabilities $(p_1, \ldots, p_P)$. For each new replication, we estimate $\widehat{H}_0$ with expectiles for all the p_i. The optimal p, denoted in the tables by p^{opt}, is then defined as the one achieving the smallest mean squared error (based on the $B = 100$ replications).

The procedure based on expectiles, denoted E(p) in the results, is compared to the standard method (ST) and to methods which efficiently deal with outliers, that is methods MED and TM (the last one is calculated by discarding 5% of the lowest and the highest values of $(\mathbf{X}^{\mathbf{a}^m})^2$ at each scale m).

The *standard fBm* model is used as a control to show that all methods perform well. As seen in the first two columns of Tables 1 and 2, this is indeed true. All the methods seem to be asymptotically unbiased and have a variance converging to zero. We can also remark that in this situation whatever the value of H, estimates based on expectiles exhibit a performance which is very close to the one of the standard method (wich can also be viewed as the method based on expectile with $p = 0.5$). Several types of expectiles are investigated. When the discretized sample path of the fBm is contaminated by outliers, we recover the results already shown in Coeurjolly (2008), Achard and Coeurjolly (2010) or Kouamo et al. (2010): methods based on medians or trimmed-means are very efficient which is in agreement with the fact that quantiles have a finite gross error sensitivity. The inefficiency of expectiles for such a contamination is also coherent since expectiles have infinite gross error sensitivity. Finally, the interest of the expectile-based method can be seen with the *rounded fBm* corresponding to the last two columns of each table. In this situation, expectiles are shown to be more efficient in terms of bias and its variance seems to be not too much affected by this type of strong contamination. We also put the stress on the interest and efficiency to choose the p value based on a Monte-Carlo approach.

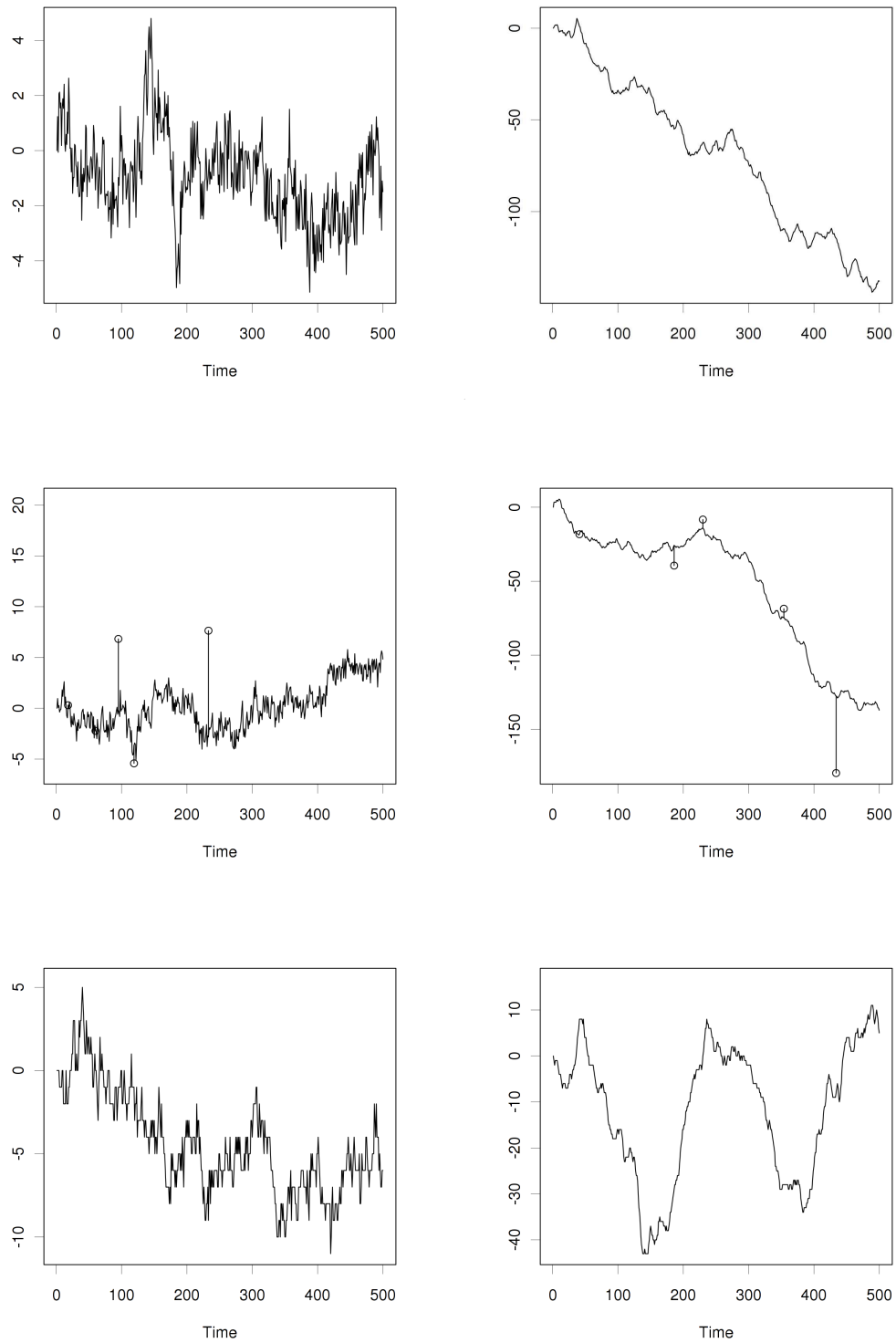

Figure 4: Examples of discretized sample path of standard fBm (top), fBm with additive outliers (middle) and rounded fBm (bottom) for $n = 500$ and with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right).

	Standard fBm		fBm with additive outliers		Rounded fBm	
	$n = 500$	$n = 5000$	$n = 500$	$n = 5000$	$n = 500$	$n = 5000$
$E(p = 0.2)$	0.198 (0.033)	0.200 (0.011)	0.280 (0.062)	0.298 (0.024)	0.334 (0.036)	0.337 (0.011)
$E(p = 0.4)$	0.198 (0.032)	0.200 (0.010)	0.288 (0.068)	0.309 (0.026)	0.298 (0.034)	0.300 (0.011)
$E(p = 0.6)$	0.199 (0.032)	0.200 (0.010)	0.298 (0.076)	0.323 (0.029)	0.284 (0.035)	0.287 (0.011)
$E(p = 0.8)$	0.199 (0.033)	0.200 (0.010)	0.311 (0.086)	0.349 (0.034)	0.275 (0.037)	0.277 (0.011)
$E(p = p^{opt})$	0.199 (0.035)	0.200 (0.011)	0.314 (0.085)	0.368 (0.033)	**0.249 (0.040)**	**0.240 (0.012)**
MED	0.197 (0.048)	0.200 (0.016)	0.227 (0.050)	0.227 (0.016)	0.451 (0.158)	0.361 (0.119)
TM	0.206 (0.034)	0.201 (0.011)	**0.222 (0.052)**	**0.225 (0.016)**	0.294 (0.038)	0.289 (0.012)
ST	**0.199 (0.032)**	**0.200 (0.010)**	0.293 (0.072)	0.315 (0.027)	0.290 (0.034)	0.292 (0.011)

Table 1: Empirical means and standard deviations of H estimates based on $m = 500$ replications of non-contaminated and contaminated fractional Brownian motions with scale parameter $\sigma = .5$, Hurst parameter $H = 0.2$ and sample size $n = 500, 5000$ are given between brackets. Methods based on expectiles, quantiles and trimmed-means as well as the standard method are considered. The filter a correspond to the Daubechies wavelet filter with order 4 (two zero moments) and we set $M_1 = 1, M_2 = 5$. According to a sample size and a model, the method achieving the lowest mean squared error is printed in bold.

References

P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. *Wavelets for the analysis, estimation, and synthesis of scaling data*, pages 39–88. In Self-similar Network Traffic and Performance Evaluation. by K. Park and W. Willinger, Wiley, New York, 2000.

P. Abry, P. Flandrin, M.S. Taqqu, and D. Veitch. *Self-similarity and long-range dependence through the wavelet lens*, pages 527–556. Theory and applications of long-range dependence. Birkhäuser, 2003.

S. Achard and J-F. Coeurjolly. Discrete variations of the fractional brownian motion in the presence of outliers and an additive noise. *Statistics Surveys*, 4:117–147, 2010.

M.A. Arcones. Limit theorems for nonlinear functionals of stationary gaussian field of vectors. *Ann. Probab.*, 22: 2242–2274, 1994.

J.M. Bardet, G. Lang, E. Moulines, and P. Soulier. Wavelet estimator of long-range dependent processes. *Statistical Inference for Stochastic Processes*, 3:85–99, 2000.

J. Beran. *Statistics for long memory processes.* Monogr. Stat. Appl. Probab. 61. Chapman and Hall, London, 1994.

G.E. Bijwaard and P. Hans Franses. The effect of rounding on payment efficiency. *Computational statistics and data analysis*, 53(4):1449–1461, 2009.

	Standard fBm		fBm with additive outliers		Rounded fBm	
	$n = 500$	$n = 5000$	$n = 500$	$n = 5000$	$n = 500$	$n = 5000$
$E(p = 0.2)$	0.796 (0.044)	0.800 (0.014)	0.725 (0.065)	0.715 (0.024)	0.775 (0.049)	0.777 (0.014)
$E(p = 0.4)$	0.795 (0.043)	0.800 (0.013)	0.712 (0.072)	0.700 (0.027)	0.725 (0.048)	0.728 (0.014)
$E(p = 0.6)$	0.795 (0.042)	0.800 (0.013)	0.692 (0.083)	0.679 (0.032)	0.702 (0.048)	0.706 (0.013)
$E(p = 0.8)$	0.794 (0.043)	0.800 (0.014)	0.653 (0.105)	0.632 (0.042)	0.702 (0.049)	0.707 (0.014)
$E(p = p^{opt})$	0.796 (0.045)	0.800 (0.014)	0.722 (0.073)	0.713 (0.027)	**0.786 (0.058)**	**0.786 (0.018)**
MED	0.799 (0.064)	0.800 (0.019)	0.817 (0.062)	0.816 (0.021)	1.242 (2.487)	0.903 (0.017)
TM	0.803 (0.045)	0.801 (0.014)	**0.815 (0.052)**	**0.814 (0.017)**	0.696 (0.048)	0.689 (0.014)
ST	**0.798 (0.042)**	**0.800 (0.013)**	0.703 (0.077)	0.691 (0.029)	0.710 (0.048)	0.714 (0.013)

Table 2: Empirical means and standard deviations of H estimates based on $m = 500$ replications of non-contaminated and contaminated fractional Brownian motions with scale parameter $\sigma = .5$, Hurst parameter $H = 0.8$ and sample size $n = 500, 5000$ are given between brackets. Methods based on expectiles, quantiles and trimmed-means as well as the standard method are considered. The filter a correspond to the Daubechies wavelet filter with order 4 (two zero moments) and we set $M_1 = 1, M_2 = 5$. According to a sample size and a model, the method achieving the lowest mean squared error is printed in bold.

P. Bois and J. Vignes. An algorithm for automatic round-off error analysis in discrete linear transforms. *International Journal of Computer Mathematics*, 12(2):161–171, 1982.

J.-F. Coeurjolly. Simulation and identification of the fractional brownian motion: a bibliographical and comparative study. *J. Stat. Softw.*, 5(7):1–53, 2000.

J.-F. Coeurjolly. Estimating the parameters of a fractional brownian motion by discrete variations of its sample paths. *Stat. Infer. Stoch. Process.*, 4(2):199–227, 2001.

J.-F. Coeurjolly. Hurst exponent estimation of locally self-similar gaussian processes using sample quantiles. *Annals of Statistics*, 36(3):1404–1434, 2008.

I. Daubechies. *Ten lectures on wavelets.* CBMS-NSF Regional Conference Series on Applied Mathematics, 61. SIAM, Philadelphia, 1992.

G. Faÿ, E. Moulines, F. Roueff, and M.S. Taqqu. Estimators of long-memory: Fourier versus wavelets. *Journal of econometrics*, 151(2):159–177, 2009.

P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Trans. Inform. Theory.*, 38(2, part 2):910–917, 1992.

J.K. Ghosh. A new proof of the bahadur representation of quantiles and an application. *The Annals of Mathematical Statistics*, 42(6):1957–1961, 1971.

F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York, 1986.

P. J. Huber. *Robust statitics*. Wiley series in probability and mathematical statistics. Wiley, 1981.

J. Istas and G. Lang. Quadratic variations and estimation of the holder index of a gaussian process. *Ann. Inst. H. Poincaré Probab. Statist.*, 33:407–436, 1997.

J.T. Kent and A.T.A. Wood. Estimating the fractal dimension of a locally self-similar gaussian process using increments. *J. Roy. Statist. Soc. Ser. B*, 59:679–700, 1997.

O. Kouamo, C. Lévy-Leduc, and E. Moulines. Central limit theorem for the robust log-regression wavelet estimation of the memory parameter in the gaussian semi-parametric context, 2010. arXiv:1011.4370.

B. B. Mandelbrot and J. W. Van Ness. Siam review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:422–437, 1968.

M. Matthieu. Semantics of roundoff error propagation in finite precision calculations. *Higher-order and symbolic computation*, 19(1):7–30, 2006.

W.K. Newey and J.L. Powell. Asymmetric least squares estimation and testing. *Econometrica*, 55:819–847, 1987.

D. B. Percival and A. T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2000.

P. Robinson. Gaussian semiparametric estimation of long range dependence. *Ann. Stat.*, 23:1630–1661, 1995.

M. Rosenbaum. Integrated volatility and round-off error. *Bernoulli*, 15(3):687–720, 2009.

S. Soltani, P. Simard, and D. Boichu. Estimation of the self-similarity parameter using the wavelet transform. *Signal Process.*, 84:117–123, 2004.

M.S. Taqqu. Law of the iterated logarithm for sums of non-linear functions of gaussian variables that exhibit a long range dependence. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 40:203–238, 1977.

G. Vilmart. Reducing round-off errors in rigid body dynamics. *Journal of Computational Physics*, 227(15): 7083–7088, 2008.

E. Williams. The effects of rounding on the consumer price index. *Monthly labor review*, 129(10):80–89, 2006.

A.T.A. Wood and G. Chan. Simulation of stationary gaussian processes in $[0,1]^d$. *J. Comput. Graph. Statist.*, 3: 409–432, 1994.

G. Wornell and A. Oppenheim. Estimation of fractal signals from noise measurements using wavelets. *IEEE Transactions on Signal Processing*, 40:611–623, 1992.