HAL open science

Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise

Sophie Achard, Jean-François Coeurjolly

▶ To cite this version:

Sophie Achard, Jean-François Coeurjolly. Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise. Statistics Surveys, 2010, 4, pp.117-147. 10.1214/09-SS059 . hal-00442760

HAL Id: hal-00442760

https://hal.science/hal-00442760v1

Submitted on 22 Dec 2009

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.

Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise

Sophie Achard[1] and Jean-François Coeurjolly[1,2]
[1] Gipsa-lab, CNRS, Grenoble, France
[2] Laboratory Jean Kuntzmann, Grenoble, France.

Abstract

This paper gives an overview of the problem of estimating the Hurst parameter of a fractional Brownian motion when the data are observed with outliers and/or with an additive noise by using methods based on discrete variations. We show that the classical estimation procedure based on the log-linearity of the variogram of dilated series is made more robust to outliers and/or an additive noise by considering sample quantiles and trimmed means of the squared series or differences of empirical variances. These different procedures are compared and discussed through a large simulation study and are implemented in the `R` package `dvfBm`.

Keywords: fractional Brownian motion, Hurst exponent estimation, discrete variations, robustness, outliers

Contents

1 Introduction **2**

2 Discrete variations of the fractional Brownian motion **3**
2.1 Some general notation . . . 3
2.2 Applications to the fractional Brownian motion . . . 4

3 Discrete variations of contaminated sample paths of the fractional Brownian motion **5**
3.1 Robustness to outliers . . . 5
3.1.1 Using sample quantiles . . . 5
3.1.2 Using trimmed means . . . 6
3.2 Robustness to an additive noise . . . 6
3.2.1 Model B0: $X(t) = B_H(t) + \sigma B^{(0)}(t)$ . . . 7
3.2.2 Model B1: $X(t) = B_H(t) + \sigma B^{(1)}(t)$ . . . 7

4 Summary and general result **8**

5 Simulation study and discussion **9**
5.1 Choice of filters and their parameters . . . 9
5.2 Robustness of the estimators to contaminated models . . . 10

A Consistency of the different procedures **25**

1 Introduction

Since the pioneer work of Mandelbrot and Ness (1968), the fractional Brownian motion (fBm) has become widely popular in a theoretical context as well as in a practical one for modelling self-similar phenomena. Fractional Brownian motion can be defined as the only centered Gaussian process, denoted by $(B_H(t))_{t\in\mathbb{R}}$, with stationary increments and with variance function $v(\cdot)$, given by $v(t) = C^2|t|^{2H}$ for all $t \in \mathbb{R}$. The parameter $H \in (0,1)$ (resp. $C > 0$) is referred to as the Hurst parameter (resp. the scaling coefficient). In particular, the case $H = 1/2$ corresponds to the standard Brownian motion. In general, the fractional Brownian motion is an H-self-similar process, that is for all $\delta > 0$, $(B_H(\delta t))_{t\in\mathbb{R}} \stackrel{d}{=} \delta^H (B_H(t))_{t\in\mathbb{R}}$ (where $\stackrel{d}{=}$ means equal in finite-dimensional distributions) with autocovariance function behaving like $O(|k|^{2H-2})$ as $|k| \to +\infty$. Thus, the discretized increments of the fractional Brownian motion (called the fractional Gaussian noise) constitute a short-range dependent process, when $H < 1/2$, and a long-range dependent process, when $H > 1/2$. The index H characterizes also the path regularity since the fractal dimension of the fractional Brownian motion is equal to $D = 2 - H$. General references on self-similar processes and long-memory processes are given in Beran (1994) or Doukhan et al. (2003).

Due to the different characteristics of the fractional Brownian motion which are of practical interest, the estimation of the Hurst parameter H is a very important (and quite difficult) task which has led to a very vast literature. We refer the interested reader to Coeurjolly (2000a), to the book of Doukhan et al. (2003) and the references therein and to the excellent paper of Fäy et al. (2009) which focuses on long-memory processes.
This paper hightlights one class of these methods, namely the method based on discrete variations, which has known great developments these last years. This method originates simultaneously from works of Kent and Wood (1997) and Istas and Lang (1997) in the context of locally self-similar Gaussian processes and more deeply in Coeurjolly (2001) in the case of the fractional Brownian motion. These ideas have then been used/extended in many other situations: *e.g.* Cohen and Istas (2002) for more general local self-similar processes, Coeurjolly (2005) for the multifractional Brownian motion, Coeurjolly (2008) for a more robust estimate, Richard and Biermé (2008) for anisotropic Gaussian random fields, Istas (2007) in the context of spherical Brownian motion, Brouste et al. (2007) for more general Gaussian random fields ...

This paper focuses on fBm-type processes by using discrete variations type procedures. Consider first, $\mathbf{B}_H$ a sample path of a fractional Brownian motion discretized at times $i = 1, \ldots, n$ and with parameters H, C. Let a be a vector with real components representing a filter and $\mathbf{B}_H^a$ the filtered series. For example, when $a = (1, -1)$ (resp. $(1, -2, 1)$), $\mathbf{B}_H^a$ corresponds to the increments (resp. the increments of the increments) of $\mathbf{B}_H$ (this is presented in Section 2). Moreover, let a^m be the filter a dilated m times (for example $(1, -2, 1)^2 = (1, 0, -2, 0, 1)$), the classical estimation procedure is based on the following property

$$Var\left(B_H^{a^m}(i)\right) = m^{2H} \times \gamma_{H,C} \quad \Leftrightarrow \quad \log\left(Var\left(B_H^{a^m}(i)\right)\right) = 2H\log(m) + \log(\gamma_{H,C}),$$

where $\gamma_{H,C}$ is a constant independent of m. It is now sufficient for different values of m to estimate the variance by its empirical version and to estimate H (actually $2H$) through a simple log-linear regression. This procedure has many advantages: it is extremely simple to implement, computationnally fast (it does not need a large number of dilated filters). In addition, the definition of the estimate is independent of the scaling coefficient and invariant of the discretization step. From a theoretical point of view (see *e.g.* Coeurjolly (2001)), this estimate is consistent and follows a central limit theorem if $p = 1$ and $H < 3/4$ and for any H if $p \geq 2$ where p is the

order of the filter (1 for $a = (1, -1)$ and 2 for $a = (1, -2, 1), \ldots$). This is proved by the fact that the correlation function of the filtered series decays as $|k|^{2H-2p}$.

The aim of this paper is to show that, when the data are contaminated by outliers and/or by an additive noise, it is still possible to adapt the previous method in order to take into account the possible contaminations and to keep its principal properties: estimation of H without estimating any other parameters, simple and computationnally fast. The Sections 2 and 3 give a survey on this topic: we show how when replacing the empirical variance by sample quantiles or trimmed means of the squared series, it is possible to define an estimate more robust to outliers. This problem has been already considered by Coeurjolly (2008). We also demonstrate that if the data are composed of a fractional Brownian motion plus a standard Brownian motion or standard Gaussian variables, it is still possible to define an estimate of H by considering differences of empirical variances. Finally, we also show that it is possible to combine these different procedures. We propose consistency results (depending on the model) proved in appendix. In Section 5, we have conducted a large simulation study where pure and contaminated sample paths of fractional Brownian motions are considered. The different estimation procedures and parameters are compared and discussed. Finally, this paper is accompanied with a simple `R` package named `dvfBm` available on the `R CRAN` (`http://cran.r-project.org/`)

2 Discrete variations of the fractional Brownian motion

2.1 Some general notation

Let $\mathbf{X} = (X(1), \ldots, X(n))$ be a sample of a stochastic process (with stationary increments and finite variance) at times $i = 1, \ldots, n$. Define a as a filter of length $\ell + 1$ with order $p \geq 1$, that is a vector with $\ell + 1$ real components satisfying

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

For instance, we shall consider the following filters: Increments 1 ($a = i1 = (-1, 1)$), Increments 2 ($a = i2 = (1, -2, 1)$), Daublets 4 ($a = d4 = (-0.09150635, -0.15849365, 0.59150635, -0.34150635)$), ... We refer the reader to Percival and Walden (2000) or Daubechies (2006) for details on Daubechies wavelet filters and extensions. We define also the vector $\mathbf{X}^a$ as the vector $\mathbf{X}$ filtered with a and given for $i = \ell + 1, \ldots, n$ by

$$X^a(i) := \sum_{q=0}^{\ell} a_q X(i-q).$$

$\widetilde{\mathbf{X}}^a$ is the normalized vector of $\mathbf{X}^a$ defined by

$$\widetilde{X}^a(i) = \frac{X^a(i)}{\mathbf{E}\left(X^a(i)^2\right)^{1/2}} = \frac{X^a(i)}{\mathbf{E}\left(X^a(1)^2\right)^{1/2}},$$

due to the stationarity of the increments of $\mathbf{X}$. Let us also denote for a function $g(\cdot)$ the vector $\mathbf{g}(\mathbf{X}) = (g(X(1)), \ldots, g(X(n)))$. Moreover, $\overline{\mathbf{X}}$, $\widehat{\xi}(p, \mathbf{X})$ (for some $0 < p < 1$) and $\overline{\mathbf{X}}^{(\beta)}$ (for some vector $\beta = (\beta_1, \beta_2)$ satisfying $0 < \beta_1, \beta_2 < 1/2$) will respectively denote the empirical mean of $\mathbf{X}$, the sample quantile of $\mathbf{X}$ and the $\beta-$trimmed mean of $\mathbf{X}$ defined by

$$\overline{\mathbf{X}}^{(\beta)} = \frac{1}{n - [n\beta_2] - [n\beta_1]} \sum_{[n\beta_1]+1}^{n-[n\beta_2]} \left(\mathbf{X}^a\right)_{(i),n},$$

where $[\cdot]$ denotes the integer part and where $(\mathbf{X})_{(1),n} \leq (\mathbf{X})_{(2),n} \leq \ldots \leq (\mathbf{X})_{(n),n}$ are the order statistics of $X(1), \ldots, X(n)$. It is well-known that $(\mathbf{X})_{(i),n} = \widehat{\xi}(i/n, \mathbf{X})$. Finally, in the sequel Z will denote a random variable following a standard Gaussian distribution.

2.2 Applications to the fractional Brownian motion

Let $\mathbf{B}_H$ be a discretized sample path of a fractional Brownian motion at times $i = 1, \ldots, n$ with Hurst parameter $H \in (0,1)$ and with scaling coefficient $C > 0$. Denote by $\mathbf{B}_H^a$ and $\widetilde{\mathbf{B}}_H^a$ its filtered version and normalized filtered version. The covariance and correlation functions of $\mathbf{B}_H^a$ are given for $i \in \mathbb{Z}$ by (see Coeurjolly (2001))

$$\mathbf{E}\left(B_H^a(j)B_H^a(i+j)\right) = C^2 \times \pi_H^a(i) \text{ with } \pi_H^a(i) = -\frac{1}{2}\sum_{q,r=0}^{\ell} a_q a_r |q-r+i|^{2H}$$

and

$$\mathbf{E}\left(\widetilde{B}_H^a(j)\widetilde{B}_H^a(i+j)\right) = \frac{\pi_H^a(i)}{\pi_H^a(1)},$$

which is clearly independent of C. The interest of filtering a discretized sample of a fractional Brownian motion is revealed by the fact that the action of filtering destroys the correlation of the increments. Indeed, it may be proved (see *e.g.* Coeurjolly (2001)) that $\rho_H^a(i) \sim k_H |i|^{2H-2p}$.

Let us now explain how one can estimate the parameter H (independently of C). First, consider the collection of dilated filters $(a^m)_{m\geq 1}$ of a filter a. Recall that a^m is the filter of length $m\ell + 1$ with order p and defined for $i = 0, \ldots, m\ell$ by

$$a_i^m = \begin{cases} a_{i/m} & \text{if } i/m \text{ is an integer} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

As a typical example, if $a := a^1 = (1, -2, 1)$, then $a^2 := (1, 0, -2, 0, 1)$. It is shown that

$$\pi_H^{a^m}(0) = -\frac{1}{2}\sum_{q,r=0}^{m\ell} a_q^m a_r^m |q-r|^{2H} = -\frac{1}{2}\sum_{q,r=0}^{\ell} a_q a_r |mq - mr|^{2H} = m^{2H}\pi_H^a(0).$$

Now, consider the empirical mean of the squared filtered coefficients denoted by $\overline{(\mathbf{B}_H^{a^m})^2}$. Since,

$$\mathbf{E}\left(\overline{(\mathbf{B}_H^{a^m})^2}\right) = C^2\pi_H^{a^m}(0) = m^{2H}C^2\pi_H^a(0),$$

one may obtain, by denoting, from now on, $\gamma = \gamma_{H,C} := C^2\pi_H^a(0)$ (which is independent of m), the following simple linear regression model

$$\log\left(\overline{(\mathbf{B}_H^{a^m})^2}\right) = 2H\log(m) + \log(\gamma) + \underbrace{\log\left(\frac{\overline{(\mathbf{B}_H^{a^m})^2}}{\mathbf{E}\left(B_H^{a^m}(1)^2\right)}\right)}_{:=\varepsilon_m^{ST}}. \tag{3}$$

The ordinary least squares estimate associated to the regression model (3) is then given by:

$$\widehat{H}^{ST} = \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2}\left(\log\left(\overline{(\mathbf{B}_H^{a^m})^2}\right)\right)_{m=M_1,\ldots,M_2},$$

where $A_m = \log(m) - \frac{1}{M_2 - M_1 + 1}\sum_{m=M_1}^{M_2}\log(m)$ and where $1 \leq M_1 \leq M_2$.

3 Discrete variations of contaminated sample paths of the fractional Brownian motion

3.1 Robustness to outliers

As noted by Coeurjolly (2008) and Shen et al. (2007), the previous procedure which is very close from a wavelet procedure may be particularly affected by outliers. The aim of this section is to propose some alternative procedures to the standard procedure based on sample quantiles or trimmed means.

3.1.1 Using sample quantiles

Let us denote by $(\mathbf{p},\mathbf{c}) = (p_k, c_k)_{k=1,\dots,K} \in ((0,1)\times\mathbb{R}^+)^K$ for an integer $1 \le K < +\infty$. Define the following statistics based on a convex combination of sample quantiles:

$$\widehat{\xi}(\mathbf{p},\mathbf{c},\mathbf{B}_H^a) = \sum_{k=1}^{K} c_k\, \widehat{\xi}(p_k,\mathbf{B}_H^a), \tag{4}$$

where $c_k, k = 1,\dots,K$ are positive real numbers such that $\sum_{k=1}^{K} c_k = 1$. For example, this corresponds to the sample median when $K = 1, \mathbf{p} = 1/2, \mathbf{c} = 1$, to a mean of quartiles when $K = 2, \mathbf{p} = (1/4, 3/4), \mathbf{c} = (1/2, 1/2)$. The estimation procedure is based on the following remark

$$\widehat{\xi}(\mathbf{p},\mathbf{c},(\mathbf{B}_H^a)^2) = \mathbf{E}\left(B_H^a(1)^2\right) \times \widehat{\xi}(\mathbf{p},\mathbf{c},(\widetilde{\mathbf{B}}_H^a)^2).$$

It may be expected (see Proposition 2) that, as $n \to +\infty$, $\widehat{\xi}(\mathbf{p},\mathbf{c},(\widetilde{\mathbf{B}}_H^a)^2)$ converges almost surely to $\xi_{Z^2}(\mathbf{p},\mathbf{c})$ where $\xi_{Z^2}(\mathbf{p},\mathbf{c}) = \sum_{k=1}^{K} c_k \xi_{Z^2}(p_k)$ and where $\xi_{Z^2}(p)$ denotes the theoretical quantile of order p of a $\chi^2(1)$ distribution. Therefore, by using the collection of dilated filters we may write

$$\log\left(\widehat{\xi}\left(\mathbf{p},\mathbf{c},\overline{(\mathbf{B}_H^{a^m})^2}\right)\right) = 2H\log(m) + \log(\gamma \times \xi_{Z^2}(\mathbf{p},\mathbf{c})) + \underbrace{\log\left(\frac{\widehat{\xi}\left(\mathbf{p},\mathbf{c},\left(\widetilde{\mathbf{B}}_H^{a^m}\right)^2\right)}{\xi_{Z^2}(\mathbf{p},\mathbf{c})}\right)}_{:=\varepsilon_m^Q} \tag{5}$$

Again, the regression model (5) allows us to define a simple estimator of H as the ordinary least squares estimator defined by

$$\widehat{H}^Q = \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2}\ \left(\log\left(\widehat{\xi}\left(\mathbf{p},\mathbf{c},\overline{(\mathbf{B}_H^{a^m})^2}\right)\right)\right)_{m=M_1,\dots,M_2}.$$

3.1.2 Using trimmed means

Let us replace the convex combination of sample quantiles by $\beta-$trimmed means. By using the notation presented in Section 2.1 and the previous ideas, we have

$$\overline{(\mathbf{B}_H^{a^m})^2}^{(\beta)} = m^{2H}\ \gamma\ \times \overline{\left(\widetilde{\mathbf{B}}_H^{a^m}\right)^2}^{(\beta)}.$$

Then, we may write the following simple linear regression model

$$\log\left(\overline{(\mathbf{B}_H^{a^m})^2}^{(\beta)}\right) = 2H\log(m) + \log(\gamma \times \overline{Z^2}^{(\beta)}) + \underbrace{\log\left(\frac{\overline{\left(\widetilde{\mathbf{B}}_H^{a^m}\right)^2}^{(\beta)}}{\overline{Z^2}^{(\beta)}}\right)}_{:=\varepsilon_m^{TM}} \tag{6}$$

where

$$\overline{Z^2}^{(\beta)} = \frac{1}{1-\beta_2-\beta_1}\int_{\beta_1}^{1-\beta_2} \xi_{Z^2}(p)dp.$$

Since it is again expected that ε_m^{TM} converges almost surely towards 0 as $n \to +\infty$, we can define the following estimator

$$\widehat{H}^{TM} = \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2}\left(\log\left(\overline{(\mathbf{B}_H^{a^m})^2}^{(\beta)}\right)\right)_{m=M_1,\ldots,M_2}.$$

3.2 Robustness to an additive noise

This section is aimed at defining alternatives to the standard procedure when the discretized sample path of the fractional Brownian motion is corrupted by an additive noise. One may distinguish two types of models:

- the fractional Gaussian noise is contaminated by an additive Gaussian white noise which means that the fractional Brownian motion is contaminated by an additive Brownian motion. The following equation summarizes this model denoted in the sequel by $B0$: one assumes observing
$$X(i) = B_H(i) \ + \ \sigma\ B^{(0)}(i),$$
where $H \neq 1/2$, $\sigma > 0$ and where $B^{(0)}(i)$ for $i = 1, \ldots, n$ is a standard Brownian motion.
- the fractional Brownian motion is contaminated by an additive Gaussian white noise. The following equation summarizes this model denoted in the sequel by $B1$: one assumes observing
$$X(i) = B_H(i) \ + \ \sigma\ B^{(1)}(i),$$
where $\sigma > 0$ and where $B^{(1)}(i)$ for $i = 1, \ldots, n$ are i.i.d. standard Gaussian variables.

The aim of this section is to propose an estimator of H that would be independent of C and σ, easily and quickly computable. This problem (in particular the model $B0$) has already been undertaken by several authors: Shen et al. (2007), Baykut et al. (2007) in a wavelet context, and Coeurjolly (2000b).

3.2.1 Model B0: $X(t) = B_H(t) + \sigma B^{(0)}(t)$

Let us see how the standard procedure is affected by this contamination: since $B^{(0)}$ is a fractional Brownian motion with Hurst parameter $H = 1/2$, the variance of the filtered series of $\mathbf{X}$ is

$$\mathbf{E}\left(X^a(i)^2\right) = C^2\ \pi_H^a(0) \ + \ \sigma^2\ \pi_{1/2}^a(0),$$

which leads to

$$\mathbf{E}\left(\overline{(\mathbf{X}^{a^m})^2}\right) = m^{2H}\gamma + m\sigma^2\pi_{1/2}^a(0).$$

Let us define $Y^{a^m}(i) = \frac{X^{a^m}(i)}{\sqrt{m}}$, then the estimation procedure is based on the following idea which is valid as soons as $H \neq 1/2$

$$\begin{aligned}\mathbf{E}\left(\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2}\right) &= ((2m)^{2H-1} - m^{2H-1})\gamma + \frac{2m}{2m}\sigma^2\pi^a_{1/2}(0) - \frac{m}{m}\sigma^2\pi^a_{1/2}(0) \\ &= m^{2H-1}\left(2^{2H-1}-1\right)\gamma.\end{aligned}$$

Now, let us consider the following regression model:

$$\log\left(\left|\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2}\right|\right) = (2H-1)\log(m) + \log(\left(\left|2^{2H-1}-1\right|\right)\gamma) + \varepsilon_m^{B0-ST} \tag{7}$$

with

$$\varepsilon_m^{B0-ST} = \log\left(\left|\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2}\right| \;/\; \left|\mathbf{E}\left(\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2}\right)\right|\right),$$

which is aimed at converging towards 0. The corresponding ordinary least squares estimate is denoted by $\widehat{H}^{B0}$ and is defined by

$$\widehat{H}^{B0-ST} = \frac{1}{2} + \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2}\left(\log\left(\left|\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2}\right|\right)\right)_{m=M_1,\ldots,M_2}.$$

This method will be denoted in the following by B0-ST. Similarly to Section 3.1, one may define two new methods denoted by B0-Q and B0-TM by replacing the sample variance by either a convex combination of sample quantiles of the squared filtered series or by a $\beta-$trimmed mean. The two new estimators are naturally denoted by $\widehat{H}^{B0-Q}$ and $\widehat{H}^{B0-TM}$.

3.2.2 Model B1: $X(t) = B_H(t) + \sigma B^{(1)}(t)$

Let us see how the standard procedure is affected by this contamination:

$$\begin{aligned}\mathbf{E}\left(X^a(i)^2\right) &= C^2\,\pi^a_H(0) \;+\; \sigma^2\sum_{q,r=0}^{\ell} a_q a_r \mathbf{E}\left(B^{(1)}(j-q)B^{(1)}(i+j-r)\right), \\ &= \gamma + \sigma^2\sum_{q,r=0}^{\ell} a_q a_r \delta_{q,r} \\ &= \gamma + \sigma^2|a|^2,\end{aligned}$$

with $|a|^2 = \sum_{q=0}^{\ell} a_q^2$. Since $|a^m|^2 = |a|^2$, this leads to

$$\mathbf{E}\left(\overline{\left(\mathbf{X}^{a^m}\right)^2}\right) = m^{2H}\gamma + \sigma^2|a|^2.$$

Therefore by using the same idea as the previous section, one may obtain the following regression model

$$\log\left(\left|\overline{\left(\mathbf{X}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{X}^{a^m}\right)^2}\right|\right) = 2H\log(m) + \log(\left(2^{2H}-1\right)\gamma) + \varepsilon_m^{B1-ST} \tag{8}$$

with

$$\varepsilon_m^{B1} = \log\left(\left|\overline{\left(\mathbf{X}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{X}^{a^m}\right)^2}\right| \;/\; \mathbf{E}\left(\overline{\left(\mathbf{X}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{X}^{a^m}\right)^2}\right)\right),$$

which is aimed at converging towards 0. The corresponding ordinary least squares estimate is denoted by $\widehat{H}^{B1-ST}$ and is defined by

$$\widehat{H}^{B1-ST} = \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2} \left(\log \left(\left| \overline{\left(\mathbf{X}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{X}^{a^m}\right)^2} \right| \right) \right)_{m=M_1,\ldots,M_2}.$$

This method will be denoted in the following by B1-ST. Similarly, one may define two new methods denoted by B1-Q and B1-TM leading to two other estimators denoted by $\widehat{H}^{B1-Q}$ and $\widehat{H}^{B1-TM}$.

4 Summary and general result

In Sections 2 and 3, we have defined estimators of the self-similarity index based on different ideas. They all exploit the property of self-similarity of the dilated-filtered initial series. They have several common points: quickly computable, definition of an estimator which is independent of the scaling coefficient and of σ^2 (in the case of an additive noise). All these estimators are obtained by an ordinary least squares procedure and may be summarized by the following equation:

$$\widehat{H}^{\bullet} = \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2} U^{\bullet}_{M_1,M_2} + \theta^{\bullet}, \text{ with } \theta^{\bullet} = \begin{cases} 1/2 & \text{if } \bullet = \text{B0-ST, B0-Q, B0-TM,} \\ 0 & \text{otherwise.} \end{cases},$$

and where $U^{\bullet}_{M_1,M_2}$ is summarized in the following table, where the vector $\mathbf{X}$ is the vector of initial data:

Method	$(U^{\bullet}_{M_1,M_2})_m$
ST	$\log\left(\overline{\left(\mathbf{X}^{a^m}\right)^2}\right)$
Q	$\log\left(\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^m}\right)^2\right)\right)$
TM	$\log\left(\overline{\left(\mathbf{X}^{a^m}\right)^2}^{(\beta)}\right)$
B0-ST	$\log\left(\left\|\overline{\left(\mathbf{X}^{a^{2m}}\right)^2}/(2m) - \overline{\left(\mathbf{X}^{a^m}\right)^2}/m\right\|\right)$
B0-Q	$\log\left(\left\|\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^{2m}}\right)^2\right)/(2m) - \widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^m}\right)^2\right)/(m)\right\|\right)$
B0-TM	$\log\left(\left\|\overline{\left(\mathbf{X}^{a^{2m}}\right)^2}^{(\beta)}/(2m) - \overline{\left(\mathbf{X}^{a^m}\right)^2}^{(\beta)}/m\right\|\right)$
B1-ST	$\log\left(\left\|\overline{\left(\mathbf{X}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{X}^{a^m}\right)^2}\right\|\right)$
B1-Q	$\log\left(\left\|\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^{2m}}\right)^2\right) - \widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^m}\right)^2\right)\right\|\right)$
B1-TM	$\log\left(\left\|\overline{\left(\mathbf{X}^{a^{2m}}\right)^2}^{(\beta)} - \overline{\left(\mathbf{X}^{a^m}\right)^2}^{(\beta)}\right\|\right)$

The next result is proved in Section A.

Proposition 1 *The following convergences hold almost surely as* $n \to +\infty$

$$\widehat{H}^{\bullet} \longrightarrow H \quad \text{with } \bullet = \begin{cases} \textit{ST, Q, TM, B0-ST, B0-Q, B0-TM,} & \\ \textit{B1-ST, B1-Q, B1-TM} & \textit{when } \mathbf{X} = \mathbf{B}_H \\ \textit{B0-ST, B0-Q, B0-TM} & \textit{when } \mathbf{X} = \mathbf{B}_H + \sigma \mathbf{B}^{(0)} \\ \textit{B1-ST, B1-Q, B1-TM} & \textit{when } \mathbf{X} = \mathbf{B}_H + \sigma \mathbf{B}^{(1)} \end{cases} \tag{9}$$

5 Simulation study and discussion

In order to study the performance of the different estimators $\widehat{H}^{\bullet}$ for $\bullet$ =ST, Q, TM, B0-ST, B0-Q, B0-TM, B1-ST, B1-Q, B1-TM, we first simulate sample paths of pure fractional Browninan motions to control the choice of the filters and their parameters. Secondly, we use three different types of contamination to test the robustness of the chosen estimators.

In the sequel, we will denote the estimators in three different classes, the classic one which corresponds to $\widehat{H}^{\bullet}$ for $\bullet$ =ST, Q, TM, the B0-class which corresponds to $\widehat{H}^{\bullet}$ for $\bullet$ =B0-ST, B0-Q, B0-TM, and finally the B1-class which corresponds to $\widehat{H}^{\bullet}$ for $\bullet$ =B1-ST, B1-Q, B1-TM.

In the following, $\mathbf{B_H} = (B_H(1), \ldots, B_H(n))$ is a sample path of a fractional Brownian motion with Hurst parameter H and with scaling coefficient C fixed to 1. Let us note that the variance of the increments that is the variance of the fGn is thus equal to 1 ($Var(B_H(i+1) - B_H(i)) = 1$).

For each simulation, we run 500 replications with time series of length 100, 1000 and 10000. We use specific values for $\widehat{H}^{Q}$, $\mathbf{p} = 1/2$ and $\mathbf{c} = 1$, this corresponds to the sample median. We specify $\beta_1 = \beta_2 = 10\%$ for $\widehat{H}^{TM}$.

5.1 Choice of filters and their parameters

Using simulations of sample paths of pure fractional Brownian motions, we test the convergence of the proposed estimators with five different filters and parameters. The filters $i1$, $i2$ and $i3$ correspond to the increments of order 1,2 and 3 respectively. The filters $d4$, $d6$ are the Daubechies wavelet filters of order 4 and 6 respectively. For each filter, M_1 was chosen equal to 1 and M_2 was chosen equal to 2 or 5.

The tables 1, 2 and 3 present the results of simulations using the estimators defined in this paper. As previously shown in Coeurjolly (2008) and by exploring the columns corresponding to a length of 10000 points in the time series, the three classic estimators $\widehat{H}^{\bullet}$ for $\bullet$ = ST, Q, TM are asymptotically without bias and converge in mean square for all the choices of filters, and for all possible values of H. The same conclusions can be written for the B0-class and B1-class of estimators.

The choice of the filters is crucial in order to minimize the variance of the estimators. Based on the tables 1, 2 and 3, we decide to choose the filters $i1, i2$ and $d4$, with $M_1 = 1$ and $M_2 = 5$ for the other simulations.

For the specific choice of the filters, the variance of the estimators can be different : there are differences within one class and between the three classes. Inside the three classes, the estimators based on the standard scheme ($\widehat{H}^{\bullet}$ for $\bullet$ = ST, B0-ST, B1-ST), have lower variance than the other estimators based on the quantiles and trimmed means. Between the three classes, the classic estimators have less variance than the estimators based on model B0 and B1.

5.2 Robustness of the estimators to contaminated models

In this section, we explore the robustness of the estimators when the simulations are not simply pure fractional Browninan motions. We consider here three different models of contamination :

- Model AO (additive outliers): the fGn is contaminated by an additive outlier model (*e.g.* Beran (1994), p. 130)

$$X(i+1) - X(i) = Z(i)(B_H(i+1) - B_H(i)) + \sigma(1 - Z(i))U(i),$$

where $Z(i)$ are independent Bernoulli random variables with parameter $p = 0.01$ and where $U(i)$ are i.i.d. standard Gaussian random variables. This means that the expectation of the number of contaminated observations is $n \times 1\%$. The parameter σ is chosen such that the contaminated observation achieves a given Signal Noise Rate (SNR), that is such that

$$SNR = 10\log_{10}\left(\frac{Var(B_H(i+1) - B_H(i))}{Var(\sigma U(i))}\right) = 10\log_{10}\left(\frac{1}{\sigma^2}\right) \Longleftrightarrow \sigma^2 = 10^{-SNR/10}.$$

The tables 4, 5, 6 present the results for the estimators using SNR equal to 0, -10 and -20 respectively. We observe that there are no major differences when the SNR is equal to 0, but when the SNR is equal to -10 or -20, the bias is reduced for the estimators based on the quantiles and trimmed means. Especially, the bias of the standard estimators is increasing when the SNR is decreasing. In contrast, the estimators based on the quantiles and trimmed means are less affected (in terms of bias and variances) by the noise.

On the figure 7, we show the mean squared errors (in short MSE) on a log-log plot. This clearly illustrates that the estimators based on the quantiles and trimmed means have the lowest values of mean square error.

- Model $B0$: the fGn is assumed to be contaminated by an additive Gaussian white noise, that is

$$X(i+1) - X(i) = B_H(i+1) - B_H(i) + \sigma B^{(1)}(i) \Longleftrightarrow X(i) = B_H(i) + \sigma B^{(0)}(i),$$

where $B^{(0)} = B(\cdot)$ is a standard Brownian motion and where $B^{(1)}(i) = B^{(0)}(i+1) - B^{(0)}(i)$. Hence, $B^{(1)}(i)$ are i.i.d. standard Gaussian variables. The parameter σ is chosen such that the increments of $X(\cdot)$ achieve a given Signal Noise Rate (SNR), that is such that

$$SNR = 10\log_{10}\left(\frac{Var(B_H(i+1) - B_H(i))}{Var(\sigma B^{(1)}(i))}\right) = 10\log_{10}\left(\frac{1}{\sigma^2}\right) \Longleftrightarrow \sigma^2 = 10^{-SNR/10}.$$

The tables 7, 8, present the results for the estimators using SNR equal to 0, 10 respectively. Under a contamination by model B0, the bias is increasing for all the estimators. When looking at the MSE, figure 10, only the MSE of the estimators $\widehat{H}^{\bullet}$ for $\bullet$ = B0-ST, B0-Q, B0-TM) seems to converge to 0. For the other class of estimators, the MSE does not seem to converge to 0. We always remark that the estimator $\widehat{H}^{B0-ST}$ is better than $\widehat{H}^{B0-TM}$ which is better than $\widehat{H}^{B0-Q}$.

- Model $B1$: the sample path of a fBm is assumed to be contaminated by an additive Gaussian white noise, that is

$$X(i) = B_H(i) + \sigma B^{(1)}(i) \Longleftrightarrow X(i+1) - X(i) = B_H(i+1) - B_H(i) + \sigma(B^{(1)}(i+1) - B^{(1)}(i),$$

where $B^{(1)}(i)$ are i.i.d. standard Gaussian random variables. Again, the parameter σ is chosen such that the increments of $X(\cdot)$ achieve a given Signal Noise Rate (SNR), that is such that

$$SNR = 10\log_{10}\left(\frac{Var(B_H(i+1) - B_H(i))}{Var(\sigma(B^{(1)}(i+1) - B^{(1)}(i))}\right) = 10\log_{10}\left(\frac{1}{2\sigma^2}\right) \Longleftrightarrow \sigma^2 = \frac{1}{2}\times 10^{-SNR/10}.$$

The tables 9, 10, present the results for the estimators using SNR equal to 0, 10 respectively. Under a contamination by model B1, the bias is increasing for all the estimators. When looking at the MSE, figure 13, only the MSE of the estimators $\widehat{H}^{\bullet}$ for $\bullet$ = B1-ST, B1-Q, B1-TM) seems to converge to 0. For the other class of estimators, the MSE does not seem to converge to 0. We always remark that the estimator $\widehat{H}^{B1-ST}$ is better than $\widehat{H}^{B1-TM}$ which is better than $\widehat{H}^{B1-Q}$.

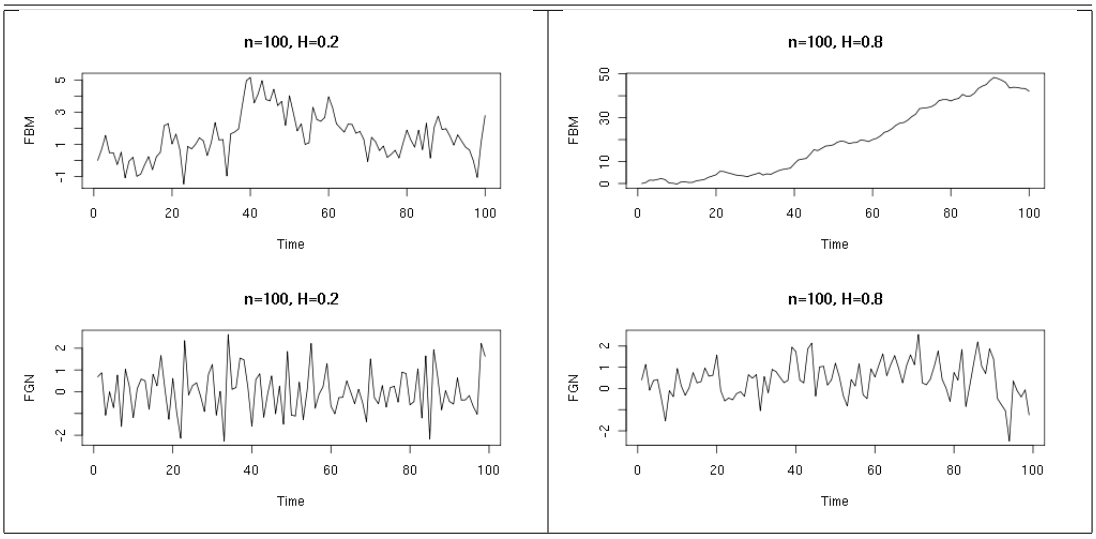

Figure 1: Example of (pure) fractional Brownian motions with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 100$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,2)	0.195	0.095	0.199	0.029	0.200	0.009	0.779	0.062	0.796	0.024	0.800	0.010
ST(i2,2)	0.189	0.155	0.199	0.048	0.201	0.015	0.789	0.127	0.799	0.037	0.799	0.012
ST(d4,2)	0.195	0.120	0.199	0.036	0.201	0.011	0.786	0.109	0.799	0.033	0.800	0.011
ST(i3,2)	0.184	0.195	0.200	0.059	0.201	0.018	0.796	0.173	0.799	0.053	0.799	0.017
ST(d6,2)	0.195	0.132	0.198	0.039	0.201	0.013	0.788	0.139	0.800	0.040	0.799	0.013
ST(i1,5)	0.201	0.059	0.199	0.019	0.200	0.006	0.774	0.072	0.795	0.028	0.800	0.012
ST(i2,5)	0.202	0.078	0.199	0.025	0.200	0.008	0.784	0.103	0.799	0.031	0.800	0.010
ST(d4,5)	0.201	0.072	0.199	0.023	0.200	0.007	0.783	0.106	0.799	0.031	0.800	0.010
ST(i3,5)	0.201	0.090	0.199	0.028	0.200	0.009	0.787	0.119	0.800	0.037	0.800	0.011
ST(d6,5)	0.202	0.079	0.199	0.024	0.200	0.007	0.783	0.122	0.799	0.035	0.800	0.011
Q(i1,2)	0.187	0.213	0.201	0.064	0.202	0.021	0.809	0.151	0.800	0.050	0.801	0.017
Q(i2,2)	0.181	0.238	0.204	0.075	0.201	0.023	0.802	0.221	0.799	0.065	0.800	0.021
Q(d4,2)	0.197	0.218	0.205	0.068	0.200	0.021	0.799	0.210	0.798	0.062	0.799	0.020
Q(i3,2)	0.160	0.287	0.200	0.088	0.200	0.028	0.813	0.266	0.795	0.079	0.798	0.026
Q(d6,2)	0.182	0.229	0.202	0.070	0.201	0.021	0.821	0.229	0.799	0.072	0.801	0.023
Q(i1,5)	0.204	0.099	0.201	0.031	0.200	0.010	0.805	0.122	0.801	0.044	0.801	0.016
Q(i2,5)	0.208	0.111	0.201	0.037	0.200	0.011	0.801	0.141	0.801	0.041	0.799	0.014
Q(d4,5)	0.208	0.111	0.201	0.034	0.200	0.010	0.799	0.149	0.800	0.041	0.800	0.014
Q(i3,5)	0.212	0.122	0.201	0.039	0.200	0.012	0.804	0.159	0.801	0.047	0.800	0.015
Q(d6,5)	0.210	0.115	0.201	0.037	0.200	0.011	0.802	0.160	0.800	0.045	0.800	0.015
TM(i1,2)	0.191	0.116	0.199	0.037	0.201	0.011	0.791	0.073	0.799	0.029	0.800	0.011
TM(i2,2)	0.181	0.174	0.199	0.053	0.201	0.017	0.794	0.142	0.799	0.044	0.799	0.014
TM(d4,2)	0.185	0.141	0.199	0.042	0.200	0.013	0.791	0.123	0.799	0.038	0.799	0.013
TM(i3,2)	0.173	0.217	0.200	0.065	0.201	0.020	0.796	0.193	0.797	0.059	0.799	0.019
TM(d6,2)	0.182	0.152	0.199	0.046	0.201	0.014	0.786	0.156	0.799	0.047	0.800	0.015
TM(i1,5)	0.234	0.065	0.202	0.022	0.200	0.007	0.820	0.080	0.802	0.032	0.801	0.013
TM(i2,5)	0.234	0.085	0.202	0.028	0.200	0.008	0.824	0.111	0.802	0.034	0.800	0.011
TM(d4,5)	0.242	0.079	0.202	0.025	0.200	0.008	0.834	0.114	0.803	0.033	0.800	0.011
TM(i3,5)	0.241	0.096	0.202	0.031	0.201	0.009	0.833	0.129	0.803	0.039	0.800	0.012
TM(d6,5)	0.248	0.087	0.203	0.027	0.200	0.008	0.836	0.130	0.803	0.037	0.800	0.012

Table 1: $m = 500$ replications of a (pure) fractional Brownian motion for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$ for the methods ST, Q and TM.

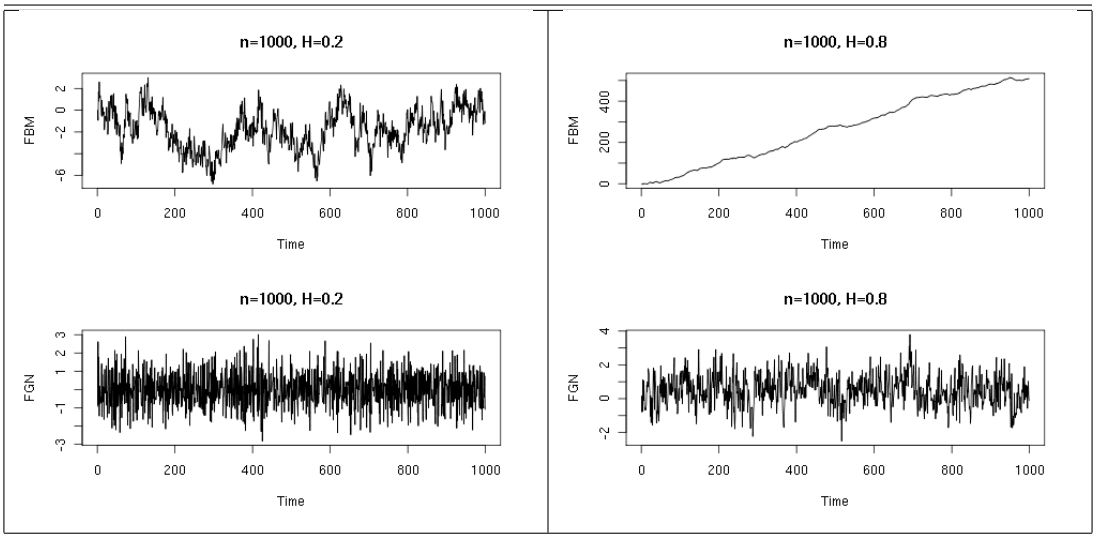

Figure 2: Example of a (pure) fractional Brownian motions with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
B0-ST(i1,2)	0.163	0.396	0.199	0.109	0.202	0.034	0.725	0.291	0.791	0.052	0.799	0.021
B0-ST(i2,2)	0.114	0.986	0.200	0.197	0.203	0.061	0.750	0.802	0.798	0.163	0.801	0.049
B0-ST(d4,2)	0.146	0.636	0.196	0.140	0.203	0.044	0.733	0.671	0.795	0.126	0.801	0.040
B0-ST(i3,2)	0.137	1.092	0.203	0.259	0.203	0.079	0.717	1.033	0.800	0.257	0.800	0.074
B0-ST(d6,2)	0.154	0.750	0.195	0.154	0.203	0.049	0.689	0.903	0.792	0.176	0.801	0.056
B0-ST(i1,5)	0.186	0.216	0.197	0.051	0.200	0.016	0.690	0.275	0.786	0.060	0.799	0.023
B0-ST(i2,5)	0.187	0.390	0.194	0.083	0.201	0.026	0.763	0.417	0.786	0.125	0.801	0.035
B0-ST(d4,5)	0.171	0.317	0.196	0.068	0.200	0.021	0.733	0.422	0.788	0.122	0.800	0.033
B0-ST(i3,5)	0.207	0.401	0.191	0.103	0.201	0.031	0.797	0.443	0.773	0.173	0.801	0.045
B0-ST(d6,5)	0.194	0.367	0.195	0.076	0.200	0.023	0.781	0.428	0.781	0.158	0.800	0.041
B0-Q(i1,2)	0.116	1.073	0.205	0.271	0.205	0.088	0.774	0.894	0.804	0.184	0.800	0.059
B0-Q(i2,2)	0.081	1.089	0.211	0.343	0.202	0.098	0.863	1.041	0.803	0.303	0.795	0.088
B0-Q(d4,2)	0.134	1.169	0.219	0.283	0.202	0.087	0.798	1.009	0.803	0.286	0.802	0.082
B0-Q(i3,2)	0.051	1.264	0.196	0.432	0.199	0.123	0.790	1.177	0.814	0.390	0.804	0.113
B0-Q(d6,2)	0.097	1.046	0.203	0.294	0.203	0.087	0.717	1.078	0.797	0.349	0.793	0.097
B0-Q(i1,5)	0.193	0.409	0.199	0.102	0.202	0.030	0.810	0.395	0.794	0.128	0.803	0.041
B0-Q(i2,5)	0.217	0.410	0.199	0.137	0.199	0.037	0.863	0.469	0.780	0.193	0.801	0.050
B0-Q(d4,5)	0.223	0.433	0.201	0.119	0.199	0.033	0.851	0.499	0.777	0.206	0.801	0.051
B0-Q(i3,5)	0.251	0.405	0.194	0.161	0.200	0.043	0.906	0.487	0.771	0.239	0.802	0.059
B0-Q(d6,5)	0.225	0.419	0.196	0.129	0.201	0.034	0.899	0.482	0.768	0.249	0.799	0.056
B0-TM(i1,2)	0.142	0.584	0.199	0.142	0.203	0.044	0.723	0.359	0.794	0.072	0.801	0.026
B0-TM(i2,2)	0.077	1.041	0.201	0.223	0.201	0.069	0.714	0.821	0.798	0.189	0.800	0.057
B0-TM(d4,2)	0.088	0.772	0.198	0.162	0.201	0.052	0.783	0.819	0.798	0.160	0.801	0.049
B0-TM(i3,2)	0.038	1.194	0.204	0.288	0.201	0.088	0.786	1.036	0.811	0.283	0.800	0.082
B0-TM(d6,2)	-0.005	0.893	0.193	0.180	0.202	0.057	0.825	0.888	0.802	0.203	0.799	0.065
B0-TM(i1,5)	0.163	0.282	0.195	0.064	0.200	0.019	0.724	0.324	0.794	0.074	0.801	0.028
B0-TM(i2,5)	0.173	0.385	0.192	0.096	0.201	0.029	0.809	0.423	0.789	0.141	0.801	0.040
B0-TM(d4,5)	0.164	0.368	0.196	0.079	0.200	0.024	0.777	0.474	0.788	0.141	0.800	0.038
B0-TM(i3,5)	0.211	0.381	0.191	0.117	0.201	0.034	0.850	0.464	0.772	0.203	0.801	0.049
B0-TM(d6,5)	0.175	0.405	0.195	0.088	0.200	0.026	0.853	0.458	0.783	0.170	0.800	0.045

Table 2: $m = 500$ replications of a (pure) fractional Brownian motion for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$ for the methods B0-ST, B0-Q and B0-TM.

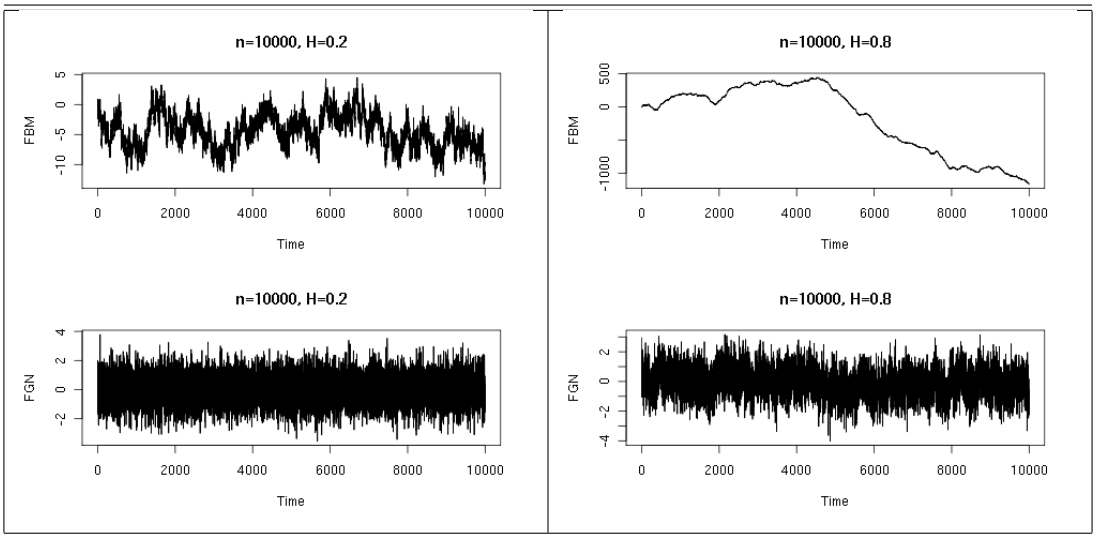

Figure 3: Example of (pure) fractional Brownian motions with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 10000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
B1-ST(i1,2)	0.248	0.794	0.198	0.173	0.197	0.052	0.765	0.110	0.794	0.035	0.800	0.015
B1-ST(i2,2)	0.230	1.118	0.196	0.318	0.196	0.092	0.772	0.252	0.800	0.075	0.800	0.023
B1-ST(d4,2)	0.202	1.004	0.202	0.223	0.195	0.066	0.772	0.229	0.799	0.062	0.800	0.020
B1-ST(i3,2)	0.241	1.221	0.185	0.451	0.196	0.120	0.763	0.375	0.802	0.110	0.800	0.033
B1-ST(d6,2)	0.204	1.045	0.204	0.244	0.195	0.073	0.766	0.310	0.798	0.082	0.800	0.026
B1-ST(i1,5)	0.144	0.402	0.199	0.086	0.199	0.028	0.740	0.155	0.792	0.041	0.799	0.018
B1-ST(i2,5)	0.192	0.444	0.199	0.139	0.197	0.042	0.753	0.239	0.796	0.060	0.801	0.018
B1-ST(d4,5)	0.163	0.450	0.199	0.113	0.198	0.035	0.743	0.249	0.797	0.061	0.800	0.018
B1-ST(i3,5)	0.221	0.440	0.194	0.183	0.195	0.050	0.721	0.321	0.794	0.071	0.801	0.022
B1-ST(d6,5)	0.196	0.472	0.198	0.127	0.198	0.038	0.702	0.351	0.796	0.070	0.801	0.021
B1-Q(i1,2)	0.249	1.170	0.199	0.451	0.194	0.133	0.801	0.297	0.803	0.085	0.800	0.029
B1-Q(i2,2)	0.311	1.237	0.175	0.644	0.197	0.150	0.786	0.466	0.804	0.127	0.798	0.038
B1-Q(d4,2)	0.314	1.194	0.167	0.492	0.196	0.133	0.783	0.440	0.803	0.121	0.801	0.037
B1-Q(i3,2)	0.362	1.257	0.201	0.758	0.201	0.189	0.767	0.648	0.808	0.155	0.801	0.049
B1-Q(d6,2)	0.364	1.131	0.199	0.534	0.196	0.132	0.754	0.506	0.801	0.138	0.797	0.043
B1-Q(i1,5)	0.231	0.437	0.191	0.190	0.196	0.048	0.768	0.277	0.802	0.070	0.802	0.026
B1-Q(i2,5)	0.233	0.480	0.188	0.235	0.198	0.059	0.783	0.307	0.798	0.081	0.801	0.025
B1-Q(d4,5)	0.259	0.466	0.189	0.200	0.200	0.052	0.779	0.342	0.797	0.087	0.801	0.026
B1-Q(i3,5)	0.243	0.484	0.191	0.289	0.197	0.067	0.784	0.363	0.798	0.091	0.802	0.028
B1-Q(d6,5)	0.307	0.463	0.188	0.233	0.198	0.053	0.770	0.389	0.797	0.092	0.800	0.028
B1-TM(i1,2)	0.238	0.918	0.196	0.224	0.195	0.066	0.776	0.139	0.798	0.043	0.801	0.017
B1-TM(i2,2)	0.231	1.264	0.191	0.370	0.199	0.104	0.767	0.286	0.800	0.085	0.800	0.026
B1-TM(d4,2)	0.276	1.105	0.199	0.259	0.199	0.078	0.793	0.267	0.800	0.074	0.800	0.023
B1-TM(i3,2)	0.297	1.193	0.177	0.571	0.199	0.134	0.788	0.407	0.807	0.119	0.800	0.036
B1-TM(d6,2)	0.362	1.091	0.209	0.284	0.198	0.085	0.808	0.343	0.803	0.093	0.800	0.030
B1-TM(i1,5)	0.179	0.390	0.200	0.104	0.199	0.032	0.769	0.167	0.798	0.048	0.801	0.020
B1-TM(i2,5)	0.232	0.455	0.197	0.165	0.197	0.046	0.783	0.249	0.799	0.065	0.801	0.020
B1-TM(d4,5)	0.218	0.466	0.197	0.129	0.198	0.040	0.770	0.269	0.798	0.068	0.800	0.020
B1-TM(i3,5)	0.222	0.460	0.189	0.229	0.196	0.055	0.752	0.331	0.796	0.077	0.801	0.024
B1-TM(d6,5)	0.239	0.495	0.196	0.148	0.198	0.042	0.755	0.349	0.798	0.076	0.800	0.023

Table 3: $m = 500$ replications of a (pure) fractional Brownian motion for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$ for the methods B1-ST, B1-Q and B1-TM.

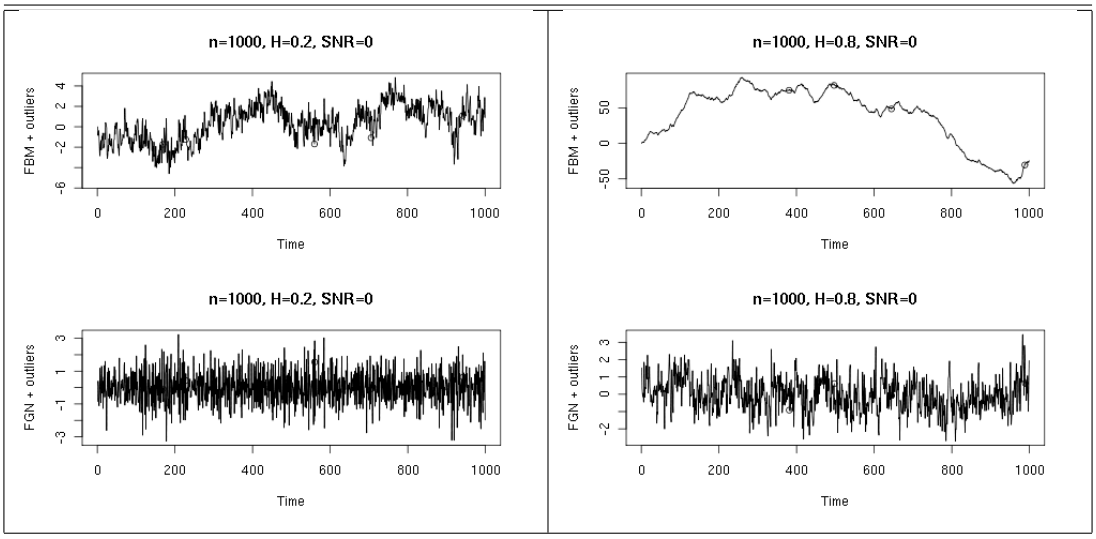

Figure 4: Examples of contaminated fractional Brownian motions (model AO with a $SNR = 0$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.201	0.063	0.202	0.019	0.203	0.006	0.771	0.072	0.793	0.027	0.798	0.012
ST(i2,5)	0.200	0.081	0.201	0.025	0.202	0.008	0.786	0.098	0.796	0.031	0.798	0.010
ST(d4,5)	0.200	0.077	0.201	0.023	0.202	0.007	0.787	0.099	0.796	0.031	0.798	0.010
Q(i1,5)	0.209	0.102	0.203	0.031	0.202	0.010	0.792	0.123	0.798	0.042	0.800	0.017
Q(i2,5)	0.205	0.119	0.200	0.037	0.202	0.011	0.802	0.134	0.799	0.043	0.799	0.013
Q(d4,5)	0.211	0.110	0.201	0.035	0.202	0.011	0.808	0.142	0.799	0.044	0.799	0.014
TM(i1,5)	0.235	0.069	0.205	0.022	0.203	0.007	0.816	0.079	0.799	0.030	0.799	0.014
TM(i2,5)	0.232	0.086	0.203	0.028	0.202	0.009	0.825	0.104	0.800	0.033	0.799	0.010
TM(d4,5)	0.240	0.081	0.204	0.025	0.202	0.008	0.836	0.106	0.801	0.033	0.799	0.011
B0-ST(i1,5)	0.174	0.216	0.197	0.054	0.200	0.016	0.677	0.319	0.784	0.058	0.797	0.023
B0-ST(i2,5)	0.175	0.369	0.195	0.085	0.201	0.025	0.757	0.434	0.790	0.119	0.799	0.035
B0-ST(d4,5)	0.160	0.305	0.196	0.071	0.200	0.021	0.727	0.412	0.790	0.113	0.798	0.033
B0-Q(i1,5)	0.196	0.403	0.197	0.104	0.200	0.030	0.803	0.409	0.794	0.121	0.799	0.041
B0-Q(i2,5)	0.194	0.412	0.190	0.137	0.200	0.037	0.878	0.475	0.788	0.188	0.799	0.049
B0-Q(d4,5)	0.209	0.440	0.190	0.116	0.201	0.034	0.848	0.496	0.785	0.182	0.799	0.049
B0-TM(i1,5)	0.143	0.282	0.196	0.065	0.200	0.019	0.720	0.347	0.788	0.070	0.798	0.027
B0-TM(i2,5)	0.164	0.364	0.193	0.095	0.200	0.028	0.796	0.448	0.793	0.131	0.799	0.037
B0-TM(d4,5)	0.147	0.321	0.195	0.080	0.200	0.024	0.772	0.437	0.790	0.127	0.799	0.036
B1-ST(i1,5)	0.172	0.358	0.205	0.090	0.206	0.027	0.746	0.143	0.789	0.040	0.798	0.017
B1-ST(i2,5)	0.196	0.470	0.198	0.144	0.204	0.041	0.762	0.230	0.796	0.058	0.799	0.018
B1-ST(d4,5)	0.196	0.434	0.201	0.124	0.205	0.034	0.746	0.252	0.796	0.057	0.798	0.018
B1-Q(i1,5)	0.237	0.452	0.202	0.184	0.205	0.045	0.767	0.255	0.799	0.068	0.800	0.026
B1-Q(i2,5)	0.255	0.495	0.192	0.256	0.205	0.057	0.775	0.341	0.799	0.081	0.800	0.025
B1-Q(d4,5)	0.281	0.467	0.216	0.214	0.203	0.052	0.754	0.358	0.798	0.082	0.800	0.025
B1-TM(i1,5)	0.212	0.386	0.206	0.106	0.205	0.030	0.770	0.160	0.793	0.047	0.799	0.020
B1-TM(i2,5)	0.254	0.456	0.200	0.164	0.204	0.045	0.783	0.249	0.798	0.063	0.799	0.019
B1-TM(d4,5)	0.243	0.444	0.204	0.133	0.205	0.038	0.762	0.287	0.797	0.063	0.799	0.019

Table 4: $m = 500$ replications of a contaminated fractional Brownian motion (model AO with a $SNR = 0$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

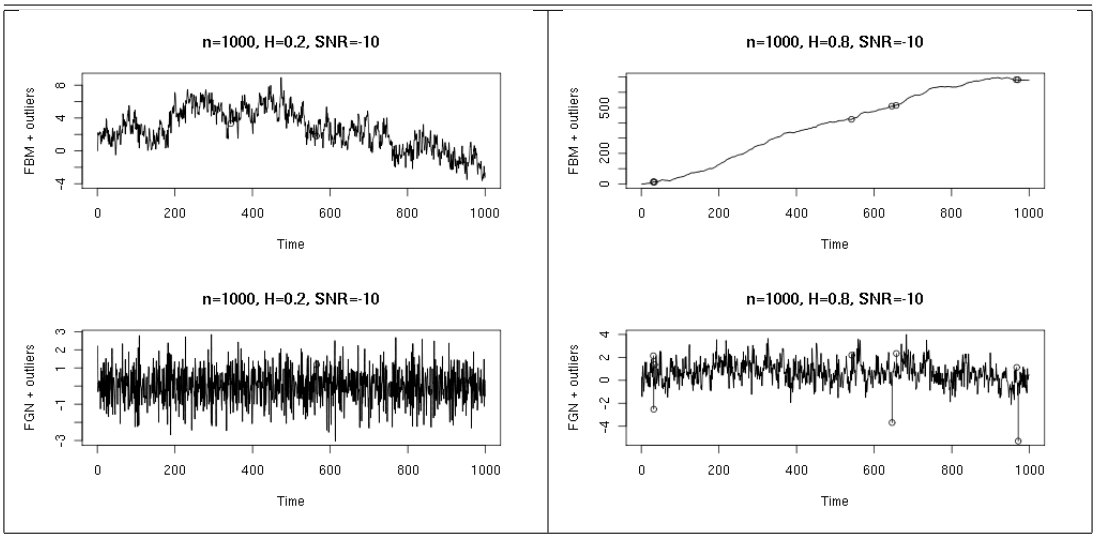

Figure 5: Example of contaminated fractional Brownian motions (model AO with a $SNR = -10$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.237	0.076	0.224	0.025	0.223	0.008	0.747	0.081	0.786	0.031	0.790	0.013
ST(i2,5)	0.228	0.089	0.219	0.028	0.218	0.009	0.756	0.109	0.783	0.033	0.781	0.011
ST(d4,5)	0.232	0.087	0.220	0.026	0.219	0.008	0.757	0.110	0.784	0.033	0.783	0.010
Q(i1,5)	0.218	0.102	0.211	0.031	0.208	0.009	0.785	0.131	0.798	0.043	0.803	0.016
Q(i2,5)	0.222	0.117	0.212	0.037	0.211	0.011	0.812	0.136	0.805	0.040	0.804	0.013
Q(d4,5)	0.227	0.114	0.213	0.035	0.211	0.010	0.806	0.148	0.804	0.042	0.803	0.013
TM(i1,5)	0.259	0.074	0.215	0.023	0.210	0.007	0.811	0.084	0.802	0.033	0.802	0.013
TM(i2,5)	0.263	0.096	0.217	0.028	0.212	0.009	0.831	0.109	0.807	0.032	0.803	0.011
TM(d4,5)	0.275	0.094	0.218	0.027	0.213	0.008	0.837	0.115	0.806	0.033	0.802	0.010
B0-ST(i1,5)	0.153	0.290	0.201	0.059	0.201	0.018	0.678	0.298	0.782	0.065	0.798	0.024
B0-ST(i2,5)	0.201	0.394	0.199	0.093	0.202	0.028	0.759	0.435	0.791	0.123	0.801	0.037
B0-ST(d4,5)	0.199	0.348	0.198	0.079	0.202	0.023	0.718	0.430	0.788	0.118	0.801	0.036
B0-Q(i1,5)	0.172	0.410	0.197	0.105	0.196	0.031	0.784	0.415	0.792	0.129	0.797	0.039
B0-Q(i2,5)	0.215	0.432	0.195	0.131	0.198	0.039	0.857	0.453	0.782	0.181	0.794	0.051
B0-Q(d4,5)	0.202	0.439	0.192	0.122	0.197	0.035	0.824	0.470	0.774	0.181	0.794	0.053
B0-TM(i1,5)	0.103	0.334	0.194	0.065	0.195	0.020	0.708	0.323	0.785	0.077	0.795	0.028
B0-TM(i2,5)	0.185	0.394	0.195	0.098	0.198	0.030	0.786	0.447	0.783	0.132	0.791	0.038
B0-TM(d4,5)	0.185	0.353	0.193	0.083	0.197	0.025	0.732	0.462	0.781	0.129	0.791	0.039
B1-ST(i1,5)	0.255	0.335	0.253	0.086	0.258	0.027	0.720	0.168	0.785	0.045	0.794	0.018
B1-ST(i2,5)	0.258	0.424	0.237	0.136	0.244	0.040	0.725	0.259	0.790	0.059	0.791	0.018
B1-ST(d4,5)	0.207	0.429	0.244	0.110	0.247	0.034	0.718	0.255	0.789	0.059	0.792	0.019
B1-Q(i1,5)	0.296	0.443	0.228	0.154	0.229	0.045	0.761	0.252	0.798	0.071	0.800	0.025
B1-Q(i2,5)	0.335	0.454	0.225	0.215	0.233	0.055	0.759	0.341	0.801	0.079	0.799	0.025
B1-Q(d4,5)	0.322	0.435	0.233	0.184	0.237	0.051	0.760	0.346	0.796	0.080	0.799	0.026
B1-TM(i1,5)	0.308	0.389	0.238	0.095	0.237	0.030	0.751	0.176	0.793	0.050	0.798	0.020
B1-TM(i2,5)	0.305	0.434	0.236	0.150	0.238	0.044	0.758	0.272	0.798	0.062	0.797	0.019
B1-TM(d4,5)	0.273	0.443	0.242	0.120	0.241	0.037	0.755	0.268	0.796	0.063	0.797	0.020

Table 5: $m = 500$ replications of a contaminated fractional Brownian motion (model AO with a $SNR = -10$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

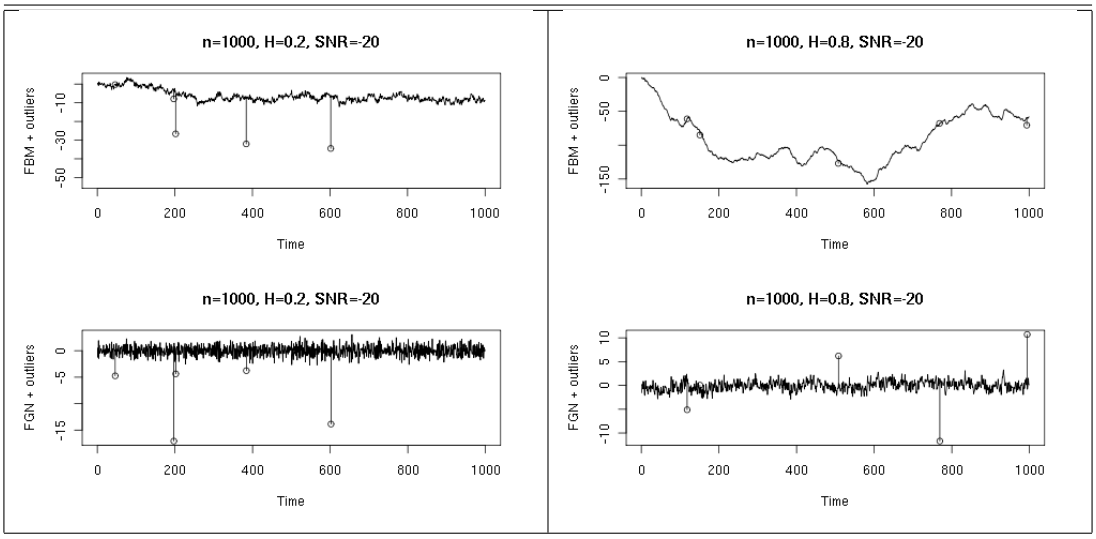

Figure 6: Example of contaminated fractional Brownian motions (model AO with a $SNR = -20$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.332	0.122	0.314	0.055	0.332	0.019	0.680	0.114	0.731	0.049	0.729	0.018
ST(i2,5)	0.318	0.130	0.297	0.054	0.311	0.019	0.663	0.140	0.693	0.059	0.684	0.019
ST(d4,5)	0.315	0.144	0.301	0.054	0.316	0.019	0.665	0.147	0.699	0.059	0.690	0.019
Q(i1,5)	0.231	0.108	0.213	0.029	0.213	0.010	0.805	0.125	0.808	0.046	0.809	0.016
Q(i2,5)	0.252	0.121	0.221	0.036	0.223	0.011	0.850	0.139	0.817	0.044	0.816	0.013
Q(d4,5)	0.268	0.126	0.227	0.036	0.227	0.011	0.849	0.143	0.818	0.045	0.816	0.014
TM(i1,5)	0.288	0.090	0.221	0.022	0.218	0.007	0.843	0.083	0.813	0.035	0.811	0.013
TM(i2,5)	0.392	0.178	0.234	0.030	0.231	0.009	0.922	0.145	0.825	0.035	0.821	0.010
TM(d4,5)	0.418	0.196	0.240	0.030	0.237	0.009	0.922	0.150	0.825	0.036	0.820	0.011
B0-ST(i1,5)	0.262	0.425	0.187	0.125	0.199	0.038	0.698	0.342	0.790	0.068	0.799	0.024
B0-ST(i2,5)	0.349	0.460	0.174	0.194	0.198	0.058	0.770	0.498	0.794	0.192	0.800	0.052
B0-ST(d4,5)	0.399	0.491	0.173	0.183	0.199	0.051	0.745	0.486	0.794	0.166	0.801	0.046
B0-Q(i1,5)	0.176	0.421	0.182	0.105	0.191	0.032	0.786	0.417	0.803	0.132	0.817	0.039
B0-Q(i2,5)	0.260	0.456	0.159	0.166	0.180	0.041	0.874	0.491	0.821	0.163	0.825	0.049
B0-Q(d4,5)	0.354	0.466	0.155	0.171	0.175	0.038	0.822	0.507	0.805	0.176	0.818	0.049
B0-TM(i1,5)	0.239	0.375	0.175	0.073	0.181	0.021	0.759	0.314	0.806	0.079	0.814	0.025
B0-TM(i2,5)	0.469	0.550	0.129	0.167	0.161	0.036	0.841	0.467	0.823	0.121	0.824	0.037
B0-TM(d4,5)	0.481	0.573	0.119	0.160	0.153	0.033	0.757	0.476	0.808	0.119	0.813	0.037
B1-ST(i1,5)	0.348	0.314	0.379	0.098	0.410	0.024	0.668	0.180	0.756	0.051	0.759	0.020
B1-ST(i2,5)	0.322	0.408	0.361	0.123	0.391	0.035	0.627	0.326	0.732	0.073	0.728	0.023
B1-ST(d4,5)	0.308	0.390	0.364	0.124	0.396	0.032	0.613	0.343	0.736	0.072	0.733	0.022
B1-Q(i1,5)	0.317	0.445	0.255	0.150	0.254	0.042	0.774	0.253	0.809	0.075	0.814	0.025
B1-Q(i2,5)	0.398	0.486	0.293	0.215	0.289	0.049	0.806	0.325	0.824	0.078	0.821	0.026
B1-Q(d4,5)	0.455	0.477	0.304	0.158	0.305	0.043	0.769	0.369	0.819	0.082	0.817	0.026
B1-TM(i1,5)	0.530	0.451	0.276	0.097	0.278	0.028	0.793	0.179	0.809	0.052	0.812	0.018
B1-TM(i2,5)	0.627	0.584	0.326	0.141	0.327	0.040	0.830	0.321	0.825	0.064	0.823	0.020
B1-TM(d4,5)	0.577	0.572	0.341	0.127	0.343	0.036	0.787	0.311	0.817	0.064	0.817	0.021

Table 6: $m = 500$ replications of a contaminated fractional Brownian motion (model AO with a $SNR = -20$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

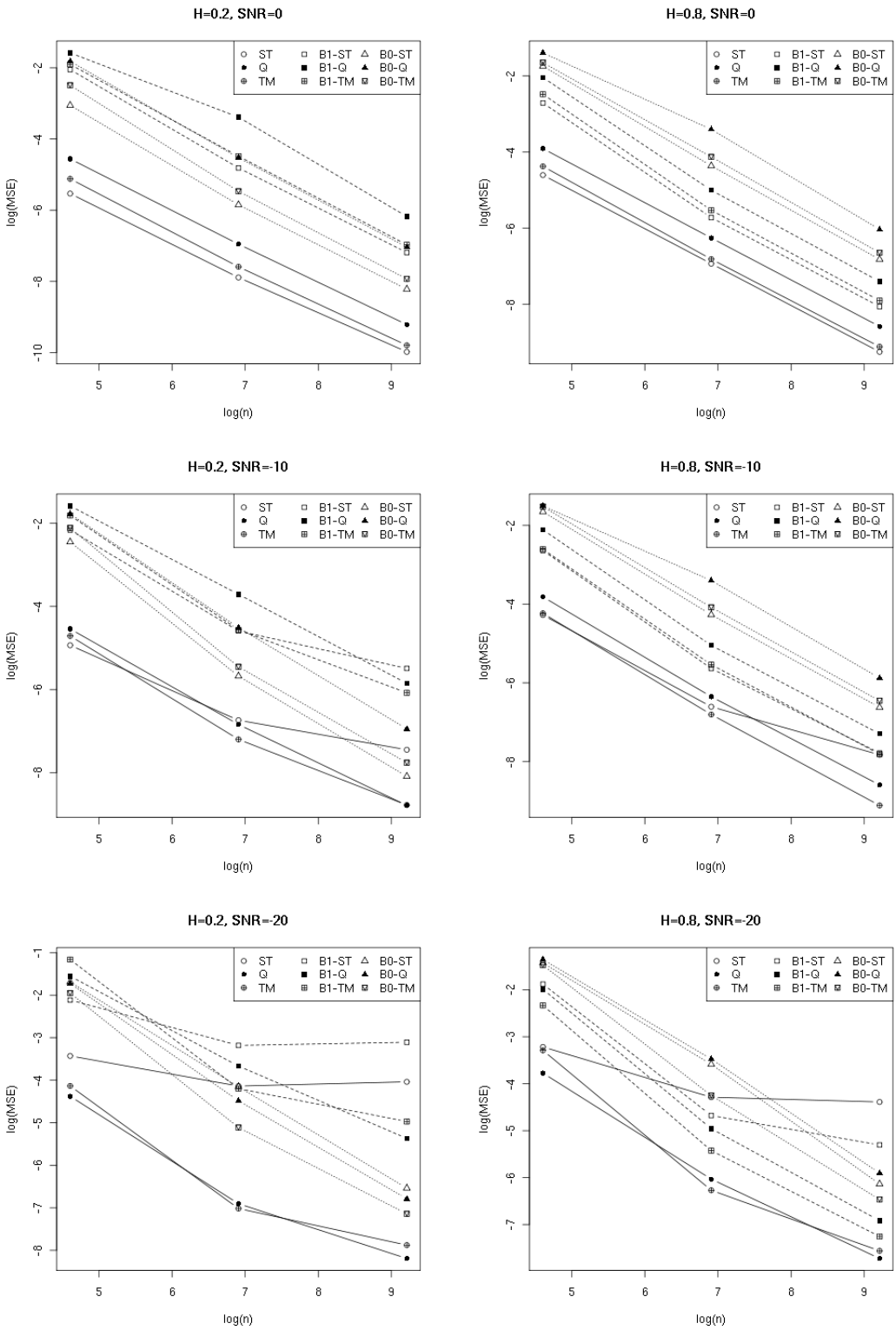

Figure 7: Empirical MSE in terms of n (for $n = 100, 1000, 10000$ in log-log scales) for the AO model for the nine methods in the previous tables for $SNR = 0, -10, -20$ and with the optimal filter $a_{opt} = i1$ for $H = 0.2$ and $a_{opt} = d4$ for $H = 0.8$, with $M_1 = 1$ and $M_2 = 5$.

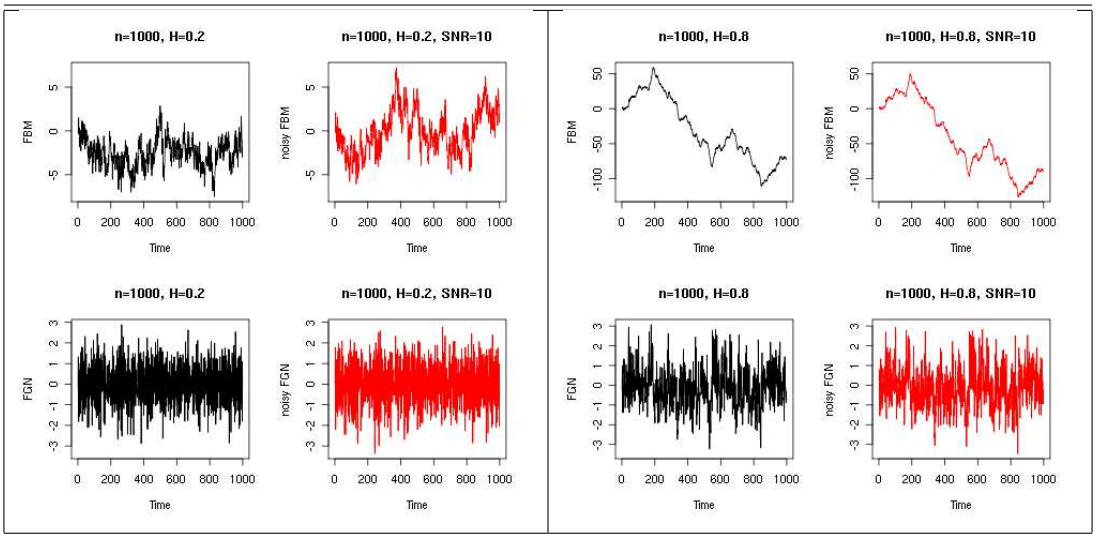

Figure 8: Example of a contaminated fractional Brownian motion (model B0 with a $SNR = 10$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.235	0.060	0.243	0.019	0.242	0.006	0.751	0.082	0.778	0.029	0.782	0.012
ST(i2,5)	0.226	0.080	0.232	0.023	0.233	0.008	0.752	0.102	0.767	0.031	0.766	0.009
ST(d4,5)	0.229	0.076	0.235	0.022	0.235	0.007	0.753	0.105	0.770	0.030	0.769	0.009
Q(i1,5)	0.241	0.099	0.242	0.032	0.242	0.009	0.773	0.131	0.782	0.044	0.783	0.015
Q(i2,5)	0.227	0.118	0.234	0.036	0.233	0.011	0.765	0.133	0.766	0.043	0.766	0.013
Q(d4,5)	0.229	0.110	0.235	0.035	0.235	0.010	0.771	0.142	0.769	0.043	0.769	0.013
TM(i1,5)	0.269	0.067	0.245	0.021	0.242	0.007	0.796	0.089	0.783	0.033	0.783	0.012
TM(i2,5)	0.258	0.086	0.235	0.026	0.233	0.009	0.789	0.107	0.770	0.034	0.766	0.010
TM(d4,5)	0.269	0.082	0.238	0.025	0.235	0.008	0.800	0.110	0.773	0.034	0.769	0.011
B0-ST(i1,5)	0.155	0.275	0.193	0.061	0.200	0.020	0.682	0.315	0.782	0.064	0.798	0.023
B0-ST(i2,5)	0.200	0.377	0.187	0.100	0.201	0.031	0.784	0.416	0.787	0.133	0.801	0.036
B0-ST(d4,5)	0.174	0.332	0.191	0.082	0.201	0.027	0.755	0.438	0.790	0.125	0.801	0.034
B0-Q(i1,5)	0.224	0.404	0.188	0.139	0.202	0.035	0.805	0.420	0.796	0.139	0.799	0.041
B0-Q(i2,5)	0.213	0.432	0.188	0.166	0.201	0.044	0.884	0.460	0.785	0.198	0.803	0.052
B0-Q(d4,5)	0.229	0.425	0.189	0.152	0.202	0.039	0.853	0.536	0.786	0.205	0.805	0.051
B0-TM(i1,5)	0.134	0.330	0.193	0.076	0.201	0.024	0.728	0.316	0.791	0.075	0.799	0.026
B0-TM(i2,5)	0.182	0.414	0.189	0.113	0.201	0.034	0.829	0.437	0.787	0.147	0.803	0.039
B0-TM(d4,5)	0.152	0.401	0.191	0.095	0.201	0.030	0.802	0.452	0.789	0.141	0.803	0.038
B1-ST(i1,5)	0.256	0.334	0.305	0.070	0.300	0.022	0.732	0.152	0.781	0.044	0.790	0.017
B1-ST(i2,5)	0.242	0.444	0.285	0.119	0.281	0.034	0.735	0.264	0.780	0.060	0.783	0.018
B1-ST(d4,5)	0.225	0.435	0.289	0.097	0.285	0.030	0.720	0.278	0.782	0.060	0.784	0.018
B1-Q(i1,5)	0.287	0.439	0.302	0.142	0.297	0.038	0.760	0.247	0.793	0.072	0.791	0.025
B1-Q(i2,5)	0.301	0.473	0.280	0.183	0.279	0.049	0.765	0.352	0.782	0.080	0.784	0.025
B1-Q(d4,5)	0.325	0.461	0.284	0.168	0.283	0.043	0.745	0.362	0.785	0.082	0.786	0.025
B1-TM(i1,5)	0.271	0.373	0.301	0.085	0.299	0.025	0.757	0.167	0.787	0.050	0.791	0.019
B1-TM(i2,5)	0.283	0.481	0.282	0.134	0.280	0.037	0.760	0.265	0.781	0.065	0.784	0.019
B1-TM(d4,5)	0.279	0.453	0.287	0.108	0.284	0.033	0.742	0.298	0.783	0.064	0.785	0.019

Table 7: $m = 500$ replications of a contaminated fractional Brownian motion (model B0 with a $SNR = 10$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

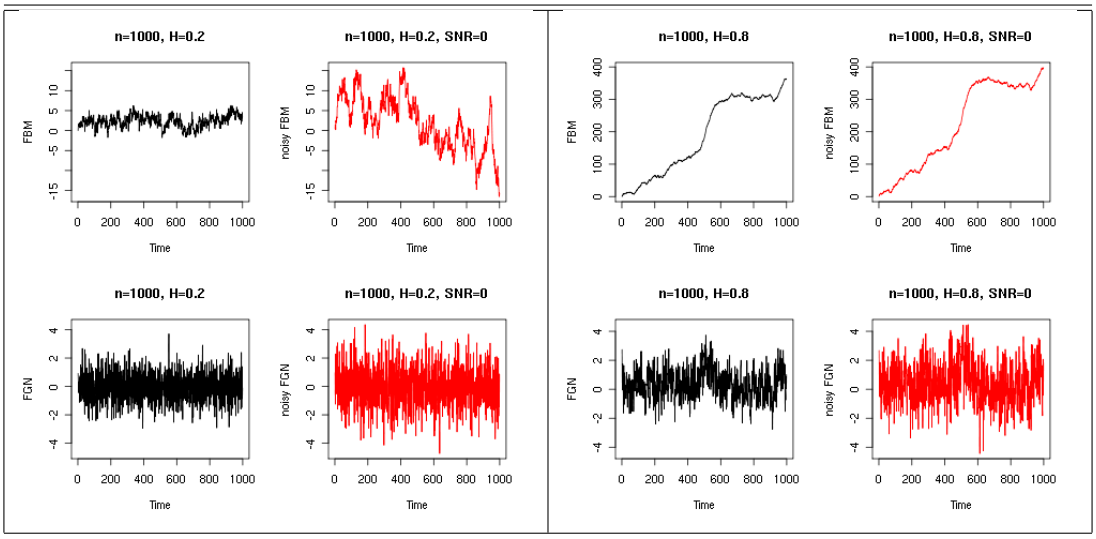

Figure 9: Example of a contaminated fractional Brownian motion (model B0 with a $SNR = 0$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.378	0.072	0.385	0.020	0.385	0.007	0.683	0.080	0.684	0.028	0.684	0.011
ST(i2,5)	0.358	0.094	0.365	0.027	0.363	0.009	0.634	0.101	0.636	0.030	0.633	0.010
ST(d4,5)	0.365	0.091	0.370	0.026	0.369	0.009	0.639	0.097	0.643	0.029	0.640	0.009
Q(i1,5)	0.387	0.109	0.385	0.033	0.385	0.011	0.699	0.132	0.686	0.044	0.685	0.015
Q(i2,5)	0.361	0.128	0.366	0.036	0.363	0.012	0.636	0.138	0.636	0.041	0.632	0.013
Q(d4,5)	0.366	0.127	0.371	0.037	0.369	0.012	0.653	0.139	0.644	0.042	0.639	0.013
TM(i1,5)	0.413	0.080	0.388	0.023	0.385	0.007	0.723	0.090	0.689	0.032	0.685	0.012
TM(i2,5)	0.391	0.102	0.368	0.029	0.363	0.010	0.668	0.108	0.639	0.032	0.633	0.011
TM(d4,5)	0.407	0.097	0.374	0.029	0.369	0.009	0.682	0.105	0.646	0.032	0.640	0.011
B0-ST(i1,5)	0.319	0.423	0.161	0.233	0.191	0.062	0.727	0.378	0.779	0.082	0.796	0.026
B0-ST(i2,5)	0.365	0.451	0.159	0.288	0.185	0.093	0.794	0.505	0.787	0.236	0.799	0.063
B0-ST(d4,5)	0.401	0.456	0.143	0.332	0.187	0.083	0.809	0.477	0.780	0.207	0.800	0.056
B0-Q(i1,5)	0.405	0.463	0.194	0.356	0.178	0.128	0.845	0.479	0.795	0.209	0.798	0.051
B0-Q(i2,5)	0.429	0.457	0.204	0.361	0.172	0.140	0.841	0.513	0.809	0.393	0.799	0.093
B0-Q(d4,5)	0.443	0.462	0.192	0.368	0.172	0.140	0.857	0.518	0.810	0.379	0.799	0.086
B0-TM(i1,5)	0.345	0.437	0.142	0.288	0.190	0.077	0.811	0.377	0.789	0.101	0.798	0.032
B0-TM(i2,5)	0.373	0.469	0.150	0.328	0.182	0.098	0.821	0.515	0.799	0.286	0.799	0.070
B0-TM(d4,5)	0.422	0.435	0.142	0.348	0.183	0.092	0.881	0.539	0.789	0.236	0.800	0.063
B1-ST(i1,5)	0.412	0.212	0.444	0.051	0.450	0.017	0.692	0.167	0.721	0.046	0.727	0.017
B1-ST(i2,5)	0.402	0.358	0.429	0.089	0.437	0.028	0.633	0.284	0.685	0.061	0.684	0.020
B1-ST(d4,5)	0.388	0.340	0.431	0.079	0.441	0.024	0.645	0.258	0.691	0.059	0.691	0.018
B1-Q(i1,5)	0.434	0.386	0.445	0.091	0.451	0.029	0.725	0.287	0.730	0.075	0.729	0.025
B1-Q(i2,5)	0.469	0.457	0.431	0.128	0.439	0.036	0.670	0.350	0.691	0.090	0.683	0.027
B1-Q(d4,5)	0.459	0.443	0.437	0.116	0.442	0.033	0.682	0.337	0.699	0.087	0.690	0.026
B1-TM(i1,5)	0.428	0.244	0.445	0.062	0.450	0.020	0.721	0.182	0.727	0.053	0.728	0.019
B1-TM(i2,5)	0.442	0.388	0.430	0.099	0.438	0.029	0.663	0.274	0.689	0.068	0.684	0.021
B1-TM(d4,5)	0.411	0.379	0.432	0.089	0.442	0.026	0.667	0.282	0.694	0.066	0.691	0.021

Table 8: $m = 500$ replications of a contaminated fractional Brownian motion (model B0 with a $SNR = 0$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

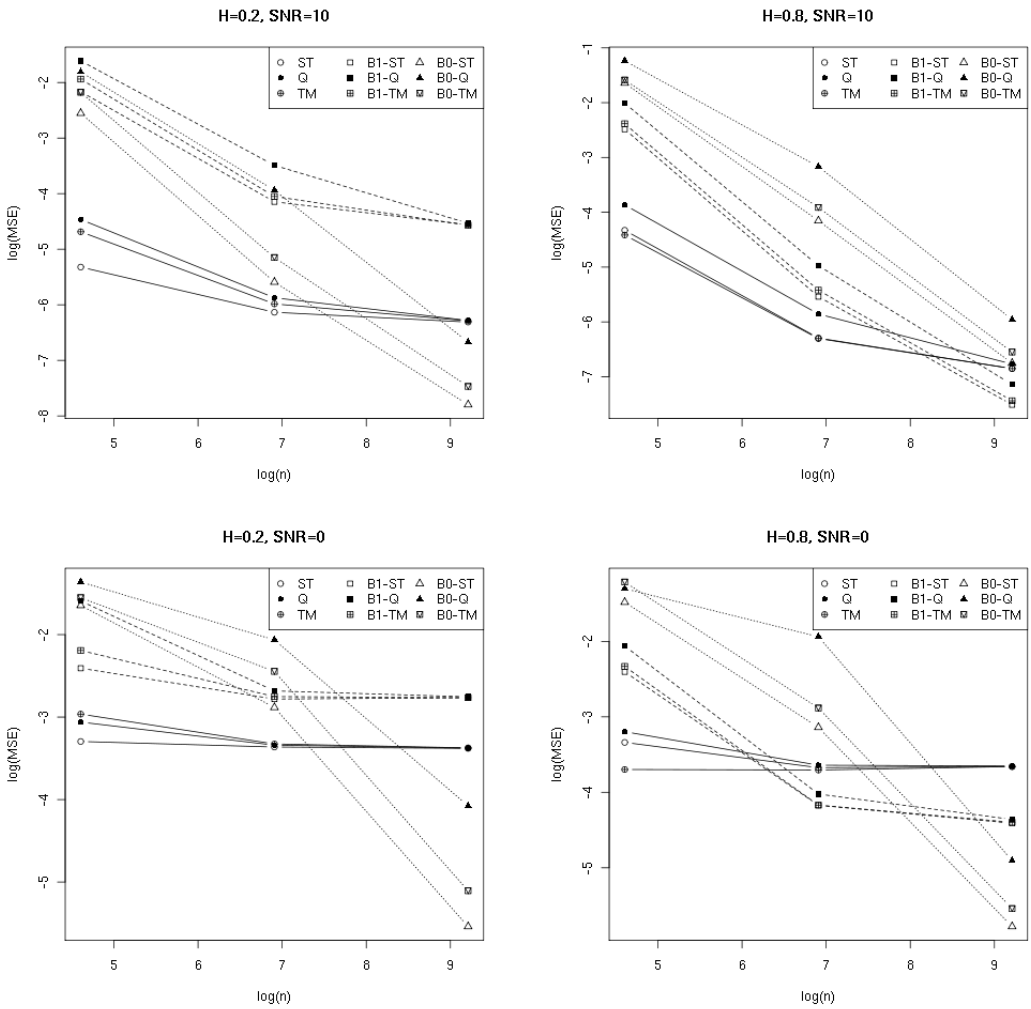

Figure 10: Empirical MSE in terms of n (for $n = 100, 1000, 10000$ in log-log scales) for the model B0 for the nine methods in the previous tables, for $SNR = 10, 0$ and with the optimal filter $a_{opt} = i1$ for $H = 0.2$ and $a_{opt} = d4$ for $H = 0.8$ with $M_1 = 1$ and $M_2 = 5$.

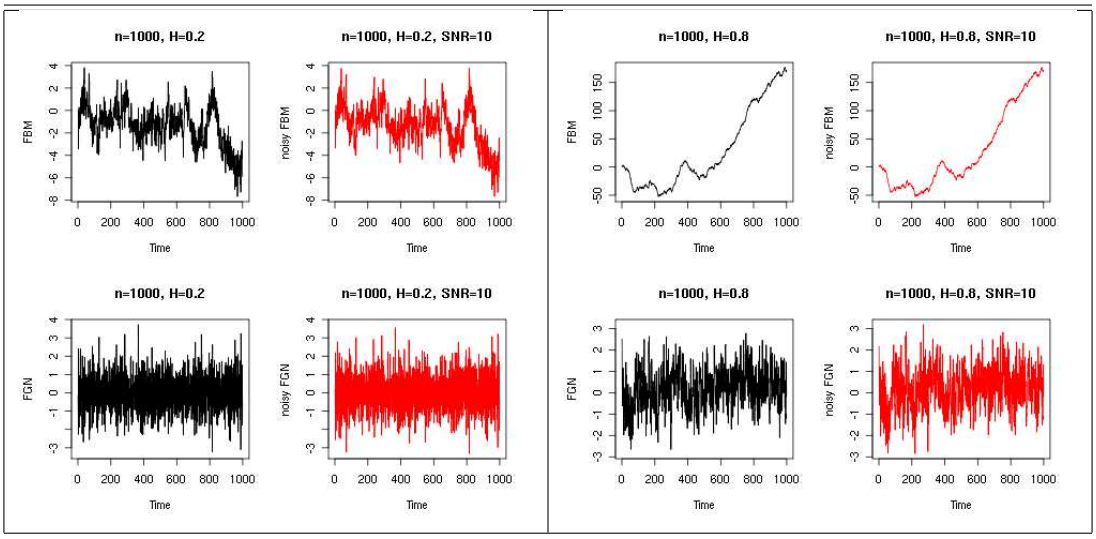

Figure 11: Example of a contaminated fractional Brownian motion (model B1 with a $SNR = 10$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.178	0.059	0.186	0.019	0.186	0.006	0.749	0.076	0.767	0.032	0.772	0.013
ST(i2,5)	0.180	0.076	0.184	0.024	0.185	0.008	0.711	0.100	0.720	0.031	0.723	0.010
ST(d4,5)	0.178	0.071	0.185	0.022	0.186	0.007	0.722	0.102	0.733	0.031	0.736	0.010
Q(i1,5)	0.180	0.103	0.185	0.031	0.186	0.009	0.778	0.129	0.772	0.049	0.773	0.017
Q(i2,5)	0.183	0.111	0.185	0.037	0.185	0.011	0.720	0.143	0.723	0.043	0.723	0.014
Q(d4,5)	0.182	0.106	0.185	0.035	0.185	0.010	0.732	0.147	0.735	0.041	0.736	0.013
TM(i1,5)	0.212	0.067	0.188	0.022	0.187	0.007	0.796	0.083	0.773	0.036	0.773	0.014
TM(i2,5)	0.210	0.084	0.187	0.027	0.185	0.008	0.746	0.109	0.724	0.034	0.723	0.011
TM(d4,5)	0.218	0.078	0.188	0.025	0.186	0.008	0.768	0.111	0.738	0.034	0.737	0.011
B0-ST(i1,5)	0.169	0.193	0.180	0.049	0.181	0.015	0.730	0.276	0.814	0.062	0.827	0.021
B0-ST(i2,5)	0.184	0.337	0.175	0.078	0.179	0.023	0.893	0.463	0.899	0.138	0.903	0.041
B0-ST(d4,5)	0.164	0.323	0.177	0.065	0.180	0.019	0.837	0.463	0.873	0.121	0.879	0.038
B0-Q(i1,5)	0.180	0.397	0.179	0.101	0.180	0.029	0.820	0.434	0.813	0.136	0.828	0.041
B0-Q(i2,5)	0.211	0.407	0.174	0.133	0.178	0.035	0.912	0.505	0.912	0.233	0.904	0.059
B0-Q(d4,5)	0.199	0.386	0.173	0.114	0.178	0.030	0.917	0.509	0.881	0.220	0.881	0.059
B0-TM(i1,5)	0.149	0.247	0.178	0.060	0.180	0.018	0.770	0.304	0.822	0.073	0.828	0.026
B0-TM(i2,5)	0.172	0.352	0.175	0.089	0.178	0.026	0.938	0.460	0.902	0.162	0.903	0.044
B0-TM(d4,5)	0.152	0.325	0.176	0.074	0.179	0.022	0.909	0.488	0.874	0.140	0.880	0.043
B1-ST(i1,5)	0.121	0.394	0.194	0.092	0.198	0.027	0.745	0.153	0.790	0.045	0.798	0.018
B1-ST(i2,5)	0.160	0.430	0.193	0.149	0.199	0.041	0.752	0.249	0.796	0.059	0.800	0.018
B1-ST(d4,5)	0.127	0.449	0.194	0.121	0.199	0.034	0.740	0.249	0.795	0.055	0.800	0.018
B1-Q(i1,5)	0.183	0.442	0.199	0.207	0.201	0.050	0.770	0.263	0.794	0.073	0.800	0.026
B1-Q(i2,5)	0.187	0.439	0.197	0.267	0.202	0.062	0.776	0.335	0.800	0.079	0.800	0.025
B1-Q(d4,5)	0.195	0.474	0.205	0.223	0.202	0.053	0.784	0.339	0.799	0.081	0.800	0.027
B1-TM(i1,5)	0.177	0.391	0.196	0.109	0.200	0.033	0.767	0.181	0.795	0.051	0.799	0.020
B1-TM(i2,5)	0.179	0.437	0.188	0.180	0.201	0.045	0.785	0.270	0.798	0.063	0.800	0.020
B1-TM(d4,5)	0.169	0.452	0.193	0.139	0.200	0.039	0.776	0.258	0.797	0.061	0.800	0.021

Table 9: $m = 500$ replications of a contaminated fractional Brownian motion (model B1 with a $SNR = 10$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

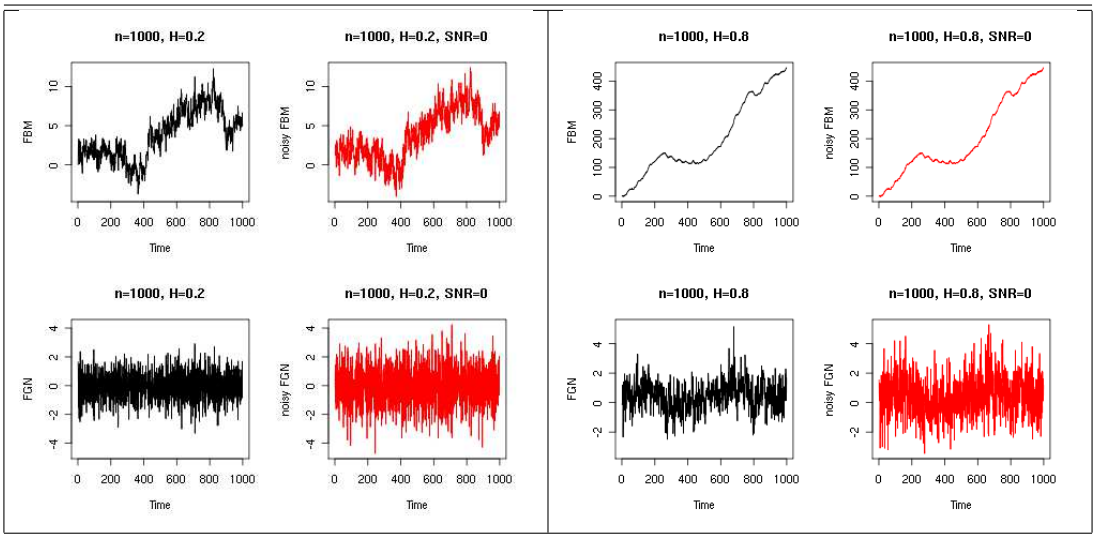

Figure 12: Example of a contaminated fractional Brownian motion (model B1 with a $SNR = 0$) with Hurst parameters $H = 0.2$ (left) and $H = 0.8$ (right) and length $n = 1000$.

	$H = 0.2$						$H = 0.8$					
	$n = 100$		$n = 1000$		$n = 10000$		$n = 100$		$n = 1000$		$n = 10000$	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ST(i1,5)	0.112	0.053	0.116	0.018	0.116	0.005	0.575	0.096	0.603	0.041	0.607	0.017
ST(i2,5)	0.105	0.067	0.111	0.023	0.110	0.007	0.422	0.093	0.424	0.031	0.424	0.009
ST(d4,5)	0.108	0.062	0.113	0.021	0.112	0.006	0.456	0.096	0.462	0.031	0.462	0.009
Q(i1,5)	0.113	0.097	0.115	0.029	0.116	0.009	0.601	0.143	0.609	0.059	0.607	0.021
Q(i2,5)	0.112	0.106	0.112	0.033	0.111	0.011	0.426	0.123	0.423	0.041	0.424	0.012
Q(d4,5)	0.111	0.102	0.112	0.031	0.112	0.010	0.463	0.132	0.463	0.041	0.462	0.013
TM(i1,5)	0.146	0.061	0.119	0.020	0.116	0.006	0.619	0.106	0.609	0.047	0.607	0.019
TM(i2,5)	0.137	0.074	0.114	0.025	0.111	0.008	0.453	0.097	0.426	0.033	0.424	0.010
TM(d4,5)	0.149	0.068	0.116	0.023	0.113	0.007	0.496	0.102	0.466	0.033	0.462	0.010
B0-ST(i1,5)	0.080	0.139	0.099	0.042	0.097	0.013	0.996	0.487	1.413	0.336	1.728	0.327
B0-ST(i2,5)	0.059	0.260	0.094	0.062	0.092	0.019	0.535	0.452	0.572	0.231	0.714	0.098
B0-ST(d4,5)	0.064	0.219	0.096	0.051	0.093	0.016	0.643	0.483	0.711	0.173	0.728	0.046
B0-Q(i1,5)	0.068	0.328	0.096	0.077	0.096	0.025	0.884	0.504	1.271	0.344	1.588	0.347
B0-Q(i2,5)	0.080	0.363	0.093	0.095	0.092	0.029	0.612	0.510	0.527	0.327	0.670	0.113
B0-Q(d4,5)	0.081	0.339	0.094	0.082	0.093	0.026	0.703	0.518	0.669	0.311	0.736	0.093
B0-TM(i1,5)	0.066	0.179	0.098	0.049	0.097	0.015	0.967	0.533	1.399	0.368	1.698	0.340
B0-TM(i2,5)	0.048	0.306	0.092	0.069	0.092	0.021	0.578	0.476	0.568	0.237	0.702	0.103
B0-TM(d4,5)	0.052	0.238	0.095	0.058	0.093	0.018	0.677	0.490	0.706	0.194	0.731	0.057
B1-ST(i1,5)	0.191	0.418	0.196	0.153	0.200	0.040	0.736	0.146	0.793	0.044	0.800	0.017
B1-ST(i2,5)	0.191	0.463	0.201	0.283	0.199	0.062	0.778	0.329	0.805	0.071	0.800	0.022
B1-ST(d4,5)	0.215	0.471	0.194	0.219	0.199	0.050	0.747	0.304	0.801	0.063	0.800	0.019
B1-Q(i1,5)	0.194	0.482	0.213	0.341	0.205	0.078	0.805	0.279	0.805	0.076	0.803	0.026
B1-Q(i2,5)	0.176	0.458	0.187	0.378	0.202	0.099	0.822	0.420	0.810	0.111	0.802	0.032
B1-Q(d4,5)	0.177	0.410	0.203	0.349	0.204	0.086	0.826	0.448	0.805	0.099	0.801	0.029
B1-TM(i1,5)	0.235	0.448	0.199	0.195	0.201	0.048	0.769	0.172	0.799	0.054	0.801	0.020
B1-TM(i2,5)	0.215	0.462	0.207	0.303	0.200	0.071	0.826	0.352	0.806	0.080	0.801	0.024
B1-TM(d4,5)	0.215	0.453	0.195	0.241	0.200	0.058	0.786	0.339	0.801	0.071	0.801	0.021

Table 10: $m = 500$ replications of a contaminated fractional Brownian motion (model B1 with a $SNR = 0$) for $n = 100, 1000, 10000$ and $H = 0.2, 0.8$ for different filters and different values of $M2$.

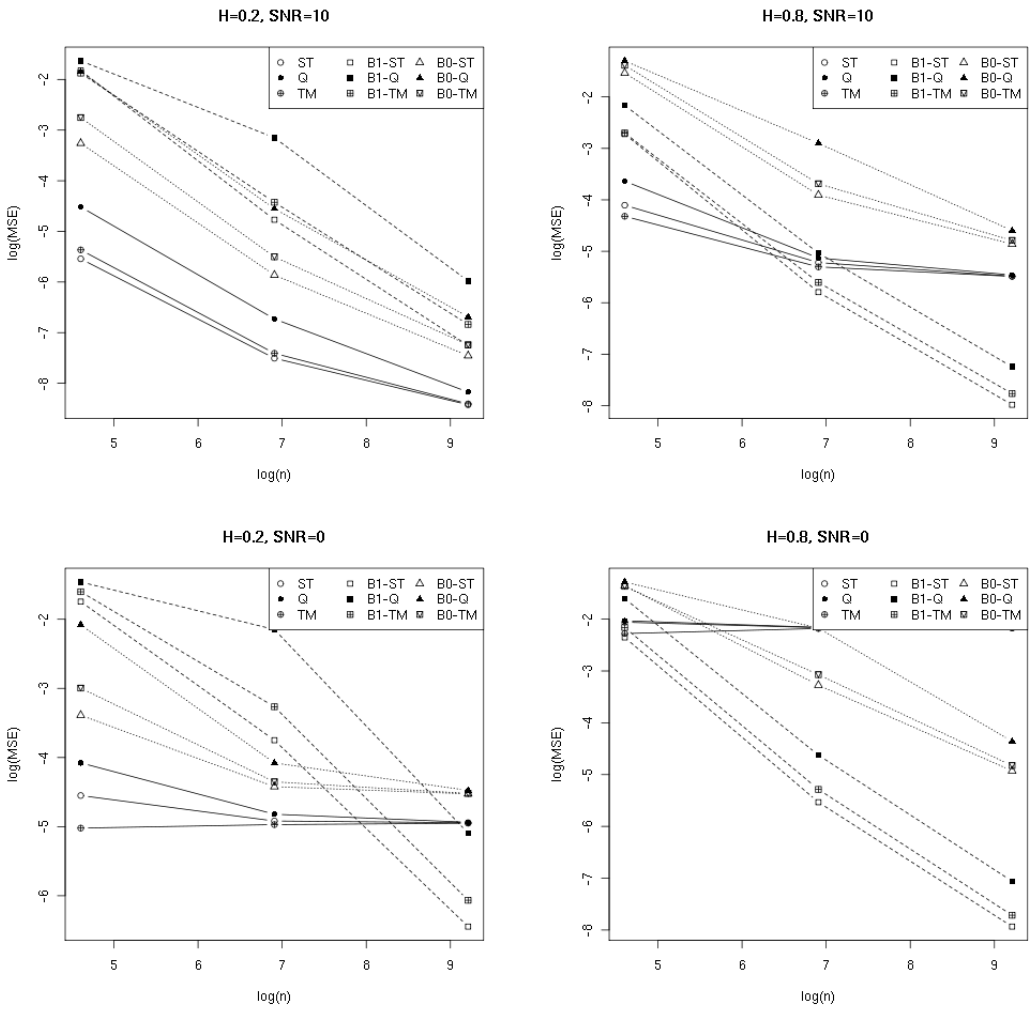

Figure 13: Empirical MSE in terms of n (for $n = 100, 1000, 10000$ in log-log scales) for the model B1 for the nine methods in the previous tables, for $SNR = 10, 0$ and with the optimal filter $a_{opt} = i1$ for $H = 0.2$ and $a_{opt} = d4$ for $H = 0.8$ with $M_1 = 1$ and $M_2 = 5$.

A Consistency of the different procedures

First of all, we leave the reader to verify that for all the methods presented in Sections 2 and 3, the variables $\varepsilon_m^\bullet$ have been defined in such a way that:

$$\widehat{H}^\bullet - H = \frac{\mathbf{A}^T}{2\|\mathbf{A}\|^2}\,(\varepsilon_m^\bullet)_{m=M_1,\ldots,M_2}$$

Now, let us present some general result

Proposition 2 *Under the previous notation, let* $\mathbf{X}$ *denote either* $\mathbf{B}_H$, $\mathbf{B}_H+\sigma\mathbf{B}^{(0)}$ *or* $\mathbf{B}_H+\sigma\mathbf{B}^{(1)}$, *then for any filter* a *of order* $p \geq 1$ *and for any* $H \in (0,1)$, *we have the following almost sure convergences as* $n \to +\infty$

$$\overline{(\mathbf{X}^a)^2} \longrightarrow \mathbf{E}\left(X^a(1)^2\right) \tag{10}$$

$$\widehat{\xi}\left(\mathbf{p},\mathbf{c},(\mathbf{X}^a)^2\right) \longrightarrow \mathbf{E}\left(X^a(1)^2\right) \times \xi_{Z^2}(\mathbf{p},\mathbf{c}) \tag{11}$$

$$\overline{(\mathbf{X}^a)^2}^{(\beta)} \longrightarrow \mathbf{E}\left(X^a(1)^2\right) \times \overline{Z^2}^{(\beta)} \tag{12}$$

Proof.

- Model $\mathbf{X} = \mathbf{B}_H$: the proof of (10) can be found in Coeurjolly (2001) (Proposition 1) while the proofs of (11) and (12) can be found in Coeurjolly (2008) (Theorem 4 and 5).
- Model $\mathbf{X} = \mathbf{B}_H + \sigma\mathbf{B}^{(0)}$: for such a model the covariance function of $\mathbf{X}^a$ is given by

  $$\mathbf{E}\left(X^a(j)X^a(i+j)\right) = C^2 \times \pi_H^a(i) \;+\; \sigma^2 \times \pi_{1/2}^a(i) \sim |i|^{2H-2p},$$

  since $\pi_{1/2}^a(i)$, which is nothing else that the covariance function of a filtered Brownian motion, vanishes for $|i| > \ell$. We may have, by following the proof of Proposition 1 of Coeurjolly (2001) to prove that for any filter and any $H \in (0,1)$,

  $$\mathbf{E}\left(\overline{(\mathbf{X}^a)^2}\right) = o(1),$$

  as $n \to +\infty$. The result (10) is then ensured by Theorem 6.2 of Doob (1953) (p. 492).

  Now, since $\widetilde{\mathbf{X}}^a$ is a Gaussian sequence with correlation function decreasing hyperbolically, Theorem 2 of Coeurjolly (2008) (resp. Theorem 3 (with $g(\cdot) = (\cdot)^2$) ensures that a Bahadur representation (resp. an uniform Bahadur representation) holds for the sample quantile $\widehat{\xi}\left(p, \left(\widetilde{\mathbf{X}}^a\right)^2\right)$ for some $p \in (0,1)$ (resp. $\sup_{p_0 \leq p \leq p_1} \widehat{\xi}\left(p, \left(\widetilde{\mathbf{X}}^a\right)^2\right)$ for $0 < p_0 < p_1 < 1$. Then, following the proofs of Theorems 4 and 5 of Coeurjolly (2008) (devoted to the case $\mathbf{X} = \mathbf{B}_H$, we can obtain the results (11) and (12).
- Model $\mathbf{X} = \mathbf{B}_H + \sigma\mathbf{B}^{(1)}$: the proof is quite similar to the previous one. The covariance function of $\mathbf{X}^a$ is given by

  $$\mathbf{E}\left(X^a(j)X^a(i+j)\right) = C^2 \times \pi_H^a(i) \;+\; \sigma^2 \times \sum_{q,r=0}^{\ell} a_q a_r \delta_{q,r-i} \sim |i|^{2H-2p},$$

  since the second term vanishes when $|i| > \ell$.

■

Now, let us prove Proposition 1.

Proof. The cases $\widehat{H}^{\bullet}$ for $\bullet =$ ST, Q, TM when $\mathbf{X} = \mathbf{B}_H$ have already obtained in Coeurjolly (2001) and Coeurjolly (2008). Since a pure fractional Brownian is a fractional Brownian motion contaminated by an additive noise ($B^{(0)}$ or $(B^{(1)})$ with $\sigma = 0$, we just have to consider the two last cases of (9).

- <u>Model $\mathbf{X} = \mathbf{B}_H + \sigma \mathbf{B}^{(0)}$:</u> for any filter a of order $p \geq 1$, for any $H \in 0, 1)$ and any $m \geq 1$ we have, from Proposition 2, as $n \to +\infty$

$$\begin{aligned}
\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2} &= \frac{\overline{\left(\mathbf{X}^{a^{2m}}\right)^2}}{2m} - \frac{\overline{\left(\mathbf{X}^{a^m}\right)^2}}{m} \\
&\xrightarrow{a.s.} \frac{\mathbf{E}\left(\overline{\left(\mathbf{X}^{a^{2m}}\right)^2}\right)}{2m} - \frac{\mathbf{E}\left(\overline{\left(\mathbf{X}^{a^m}\right)^2}\right)}{m} = m^{2H-1}\left(2^{2H-1} - 1\right)\gamma.
\end{aligned}$$

Therefore, as $n \to +\infty$

$$\left(\mathbf{U}^{B0-ST}_{M1,M2}\right)_m := \log\left(\left|\overline{\left(\mathbf{Y}^{a^{2m}}\right)^2} - \overline{\left(\mathbf{Y}^{a^m}\right)^2}\right|\right) \xrightarrow{a.s.} (2H-1)\log(m) + \log\left(\left|2^{2H-1} - 1\right|\gamma\right)$$

and so

$$\frac{\mathbf{A}^T \mathbf{U}^{B0-ST}_{m1,M_2}}{2\|\mathbf{A}\|^2} + \frac{1}{2} \xrightarrow{a.s.} (2H-1)\frac{\mathbf{A}^T\mathbf{A}}{2\|\mathbf{A}\|^2} + \frac{1}{2} = H,$$

since for $\mathbf{1} = (1, \ldots, 1)^T$, $\mathbf{A}^T\mathbf{1} = 0$. The estimators $\widehat{H}^{B0-Q}$ and $\widehat{H}^{B0-TM}$ follow the same ideas. Consider the first one for example, we have from Proposition 2

$$\begin{aligned}
\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{Y}^{a^{2m}}\right)^2\right) - \widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{Y}^{a^m}\right)^2\right) &= \widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^{2m}}\right)^2/(2m)\right) - \widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^m}\right)^2/m\right) \\
&= \frac{\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{X}^{a^{2m}}\right)^2\right)}{2m} - \frac{\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{Y}^{a^m}\right)^2\right)}{m} \\
&\xrightarrow{a.s.} \left(m^{2H-1}\left(2^{2H-1} - 1\right)\gamma\right)\xi_{Z^2}(\mathbf{p}, \mathbf{c}).
\end{aligned}$$

Therefore, as $n \to +\infty$

$$\begin{aligned}
\left(\mathbf{U}^{B0-Q}_{M1,M2}\right)_m &:= \log\left(\left|\widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{Y}^{a^{2m}}\right)^2\right) - \widehat{\xi}\left(\mathbf{p}, \mathbf{c}, \left(\mathbf{Y}^{a^m}\right)^2\right)\right|\right) \\
&\xrightarrow{a.s.} (2H-1)\log(m) + \log\left(\left|2^{2H-1} - 1\right|\gamma \times \xi_{Z^2}(\mathbf{p}, \mathbf{c})\right)
\end{aligned}$$

and so

$$\frac{\mathbf{A}^T \mathbf{U}^{B0-Q}_{m1,M_2}}{2\|\mathbf{A}\|^2} + \frac{1}{2} \xrightarrow{a.s.} (2H-1)\frac{\mathbf{A}^T\mathbf{A}}{2\|\mathbf{A}\|^2} + \frac{1}{2} = H.$$

- <u>Model $\mathbf{X} = \mathbf{B}_H + \sigma \mathbf{B}^{(1)}$:</u> the proof is omitted since it follows exactly the same ideas as the previous model.

■

Remark 1 *It is the belief of the authors that each estimate follows a central limit theorem. It is true for the methods ST, Q, TM (see Coeurjolly (2001) and Coeurjolly (2008)). Other methods need more work. However, it is not the objective of the present paper.*

References

S. Baykut, T. Akgül, and S. Ergintav. Estimation of spectral exponent parameter of 1/f process in additive white background noise. *EURASIP Journal on Advances in Signal Processing*, 2007 (15):1–7, 2007.

J. Beran. *Statistics for long-memory processes.* Chapman & Hall/CRC, 1994.

A. Brouste, J. Istas, and S. Lambert-Lacroix. On Fractional Gaussian Random Fields Simulations. *Journal of Statistical Software*, 23(1):1–23, 2007.

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

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, January 2001.

J.-F. Coeurjolly. Identification of multifractional Brownian motion. *Bernoulli*, 11(6):987–1008, 2005.

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

J.-F. Coeurjolly. *Inférence statistique pour les mouvements Browniens fractionnaire et multifractinonaire.* PhD thesis, Université Joseph Fourier, 2000b.

S. Cohen and J. Istas. An universal estimator of local self-similarity. *ESAIM: Probability and Statistics*, 2002.

I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and, Applied Mathematics*, 41(7):909–996, 2006.

J.L. Doob. *Stochastic Processes.* Wiley Classics Library, USA, 1953.

P. Doukhan, G. Oppenheim, and M.S. Taqqu. *Theory and applications of long-range dependence.* Birkhauser, 2003.

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

J. Istas. Quadratic variations of spherical fractional Brownian motions. *Stochastic Processes and their Applications*, 117(4):476–486, 2007.

J. Istas and G. Lang. Quadratic variations and estimation of the hölder 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.

B. Mandelbrot and J. Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

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

F. Richard and H. Biermé. Estimation of anisotropic gaussian fields through radon transform. *ESAIM Probab. Stat.*, 12(1):30–50, 2008.

Haipeng Shen, Zhengyuan Zhu, and Thomas C. M. Lee. Robust estimation of the self-similarity parameter in network traffic using wavelet transform. *Signal Process.*, 87(9):2111–2124, 2007.