Adaptive wavelet based estimator of long-range dependence parameter for stationary Gaussian processes

Jean-Marc Bardet[a], Hatem Bibi[a], Abdellatif Jouini[b]

bardet@univ-paris1.fr, hatem.bibi@malix.univ-paris1.fr, Abdellatif.jouini@fst.rnu.tn,

[a] Samos-Matisse-CES, Université Paris 1, CNRS UMR 8174, 90 rue de Tolbiac, 75013 Paris, FRANCE.

[b] Département de Mathématiques, Faculté des Sciences de Tunis, 1060 Tunis, TUNISIE.

Running title: Adaptive wavelet based estimator of long-memory parameter.

keyword: Adaptive estimator, Long-memory processes, Semi-parametric estimation, Wavelet analysis.

Abstract

The aim of this contribution is to provide an adaptive estimation of the long-memory parameter in the classical semi-parametric framework for Gaussian stationary processes using a wavelet method. In particular, the choice of a data-driven optimal band of scales is introduced and developed. Moreover, a central limit theorem for the estimator of the long-memory parameter reaching the minimax rate of convergence (up to a logarithm factor) is established. Simulations confirm the quality of this estimator.

1 Introduction

Let $X = (X_t)_{t\in\mathbb{R}}$ be a second order zero mean stationary process and define its covariogram

$$r(t) = \mathbb{E}(X_0 \cdot X_t), \quad \text{for } t \in \mathbb{R}.$$

We will assume that X is a long-memory process, that is,

$$\lim_{|t|\to\infty} r(t) = 0 \quad \text{and} \quad \sum_{t\in\mathbb{Z}} |r(t)| = \infty.$$

A particular case of such a property is the following form for the covariogram,

$$r(t) \sim C \cdot |t|^{-D} \quad \text{when } |t| \to \infty, \tag{1}$$

with $D \in]0, 1[$ (so-called the long-memory parameter) and $C > 0$. This property can also be translated in the spectral domain. Indeed, assume that the spectral density of X, with

$$f(\lambda) = \frac{1}{2\pi} \cdot \sum_{k\in\mathbb{Z}} r(k) \cdot e^{-ik},$$

exists for $\lambda \in [-\pi, 0)[\cup]0, \pi]$. Then, from an Abelian theorem, the asymptotic behavior (1) can be written in spectral terms, *i.e.*,

$$f(\lambda) \sim C' \cdot \lambda^{D-1} \quad \text{when } \lambda \to 0,$$

with $C' > 0$ (see Doukhan *et al.*, 2003, for more details on this part). In this paper, two following semi-parametric frameworks will be considered,

- **Assumption A1**, $X = (X_t)_{t\in\mathbb{R}}$ is a zero mean stationary Gaussian process with spectral density $f(\lambda) = |\lambda|^{D-1} \cdot f^*(\lambda)$ with $f^*(0) > 0$, and $f^* \in \mathcal{H}(D', C_{D'})$ where $0 < D' \le 2$, $C_{D'} > 0$ and

$$\mathcal{H}(D', C_{D'}) = \Big\{ g : [-\pi, \pi] \to \mathbb{R}^+, \ |g(\lambda) - g(0)| \le C_{D'} \cdot |\lambda|^{D'} \text{ for all } \lambda \in [-\pi, \pi] \Big\}.$$

- **Assumption A2**, $X = (X_t)_{t\in\mathbb{R}}$ is a zero mean stationary Gaussian process with spectral density $f(\lambda) = |\lambda|^{D-1} \cdot f^*(\lambda)$ with $f^*(0) > 0$, and $f^* \in \mathcal{H}'(D', C_{D'}, D'', C_{D''})$ where $0 < D' \le 2$, $C_{D'} > 0$, $D'' > 0$, $C_{D''} > 0$ and

$$\mathcal{H}'(D', C_{D'}, D'', C_{D''}) = \Big\{ g : [-\pi,\pi] \to \mathbb{R}^+,\ |g(\lambda) - (g(0) + C_{D'} \cdot |\lambda|^{D'})| \le C_{D''} \cdot |\lambda|^{D'+D''} \text{ for all } \lambda \in [-\pi,\pi] \Big\}.$$

Remark 1 *In numerous previous works concerning the estimation of the long range parameter in a semi-parametric framework (see for instance Robinson, 1995, Taqqu and Teverovsky, 1996, Giraitis* et al.*, 1997, Moulines and Soulier, 2003, Moulines et al., 2006), the assumption on the dependence structure is:* $f(\lambda) = |1 - e^{i\lambda}|^{-2d} \cdot f^*(\lambda)$ *with* f^* *a function such that* $|f^*(\lambda) - f^*(0)| \le f^*(0) \cdot \lambda^{\beta}$ *and* $0 < \beta \le 2$*, that is equivalent to Assumption A1, with* $2d = 1 - D$*. The Assumption A2 is a necessary condition for studying the following adaptive estimator of* D*.*

The aim of this article is the semi-parametric estimation of the parameter D using a wavelet analysis. This method has been introduced by Flandrin (1989) and numerically developed by Abry *et al.* (1998, 2001) and Veitch *et al.* (2003). Asymptotic results are provided in Bardet *et al.* (2000) and recently in Moulines *et al.* (2006). Compared with these papers, two points of our work can be highlighted : first, a central limit theorem is provided under weaker conditions than in Bardet *et al.* (2000). Secondly, an auto-driven estimator $\tilde{D}_n$ of D is defined (with a different definition than in Veitch *et al.*, 2003). A central limit theorem followed by $\tilde{D}_n$ is established and this estimator is proved to be rate optimal up to a logarithm factor (see above). Now, more details on this estimation method are provided.

Let ψ be a wavelet satisfying the following assumption, *i.e.* ψ has m first vanishing moments :

Assumption $W(m)$ **:** $\psi:\ [0,1] \mapsto \mathbb{R}$ *is a continuously differentiable function satisfying* $\psi(0) = \psi(1) = 0$ *and such that,*

1. *it exists* $m \in \mathbb{N} \setminus \{0,1\}$ *satisfying,* $\displaystyle\int_0^1 t^p \psi(t)\, dt = 0$ *for all* $p \in \{0, 1, \ldots, m-1\}$ *and* $\displaystyle\int_0^1 |\psi(t)|\, dt > 0$.
2. *it exists* $m' > 3/2$ *such that* $\sup_{\lambda\in\mathbb{R}} |\widehat{\psi}(\lambda)|(1 + |\lambda|)^{m'} < \infty$*, where* $\widehat{\psi}(u) = \int_0^1 \psi(t)\, e^{-iut} dt$ *is the Fourier transform of* ψ*.*

Remark 2 *The function* ψ *is a compactly supported function (the interval* $[0,1]$ *is just for ease of writing but the following results be easily extended to another interval) with its* m *first vanishing moments. For instance,* ψ *can be a dilated Daubeshies "mother" wavelet of order* d *with* $d \ge 5$ *to ensure the regularity of the function* ψ*. The following theory could also be extended for "essentially" compactly supported "mother" wavelet like Lemarié-Meyer wavelet. One can remark that it is not necessary to choose* ψ *being a "mother" wavelet associated to a multiresolution analysis of* $\mathbb{L}^2(\mathbb{R})$*. The whole theory can be developed without resorting to this assumption. The choice of* ψ *is then very large. However, the recent paper of Moulines* et al. *(2006) is developed under weaker conditions on* ψ*.*

For $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}$, define the wavelet coefficient $d(a, b)$ of the process X for the scale a and the shift b, *i.e.*

$$d(a,b) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \psi(\frac{t}{a} - b) X_t\, dt. \tag{2}$$

Remark 3 *Let us underline that we consider a continuous wavelet transform. However, the case of discrete wavelet transform where* $a = 2^j$*, that is numerically very interesting (using Mallat's cascade algorithm) is just a particular case. The main interest of considering continuous transform will be to offer a larger number of "scales" for computing the data-driven optimal band of scales (see above).*

If X satisfies one of the previous assumptions, the asymptotic behavior of the variance of $d(a, b)$ is a power law in the scale a. Indeed, under Assumption $W(1)$ on ψ and Assumption A1 or A2 on the process X, $(d(a, b))_{b\in\mathbb{R}}$ is a Gaussian stationary process and the following expansion can be established (see Section 2) :

$$\mathbb{E}(d^2(a,0)) \sim K_{(\psi,1-D)} \cdot a^{1-D} \quad \text{when } a \to \infty,$$

with a constant $K_{(\psi,1-D)}$ such that,

$$K_{(\psi,\alpha)} = \int_{\infty}^{\infty} |\widehat{\psi}(u)|^2 \cdot |u|^{-\alpha} du > 0 \quad \text{for all } \alpha \in]-2, 1[, \tag{3}$$

where $\widehat{\psi}$ is the Fourier's transform of ψ (the existence of $K_{(\psi,\alpha)}$ is proved in the Section 5). The principle of the wavelet-based estimation of D is linked to this power law in D of a. Indeed, let $(X_1,\ldots,X_N)$ be a sampled path of X and define $\widehat{S}_N(a)$ a sample variance of $d(a,.)$ obtained from an appropriate choice of shifts b. Then, when $a=a(N)\to\infty$ satisfies $a(N)=o(N^{1/(2D'+1)})$, a central limit theorem for $\log(\widehat{S}_N(a(N)))$ can be proved. More precisely one obtains:

$$\log(\widehat{S}_N(a(N)))=(1-D)\log(a(N))+\log(f^*(0)K_{(\psi,1-D)})+\sqrt{\frac{a(N)}{N}}\cdot\varepsilon_N,$$

with $\varepsilon_N \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2_{(\psi,D)})$ and $\sigma^2_{(\psi,D)}>0$. As a consequence, using different scales $(a_1(N),\ldots,a_\ell(N))$ a linear regression of $(\log(\widehat{S}_N(a_i(N)))_i$ on $(\log(a_i(N)))_i$ provides an estimator $\widehat{D}(a_N)$ that satisfies a central limit theorem (we suppose that it exists $m_i\in\mathbb{N}^*$ such that $a_i(N)=m_i\cdot a(N)$ $i=1,\ldots,\ell$). At this point, our result is close to Bardet *et al.* (2000) or Moulines *et al.* (2006) results, with only different conditions on the process X and the function ψ.

The main problem is now: how to select the reference scale $a(N)$ considering the fact that the smaller is $a(N)$ the faster the convergence rate of $\widehat{D}(a_N)$. An optimal choice would be to chose $a(N)$ larger but closer to $N^{1/(2D'+1)}$, but the parameter D' is supposed to be unknown. In Veitch *et al.* (2003), an automatic selection procedure is proposed using a Khi-squared goodness of fit statistic. This procedure is applied successfully on numerous numerical examples but no theoretical proofs are provided. The method we develop here is close to this one. Roughly speaking, the "optimal" choice of scale $(a(N))$ is obtained from the "best" linear regression among all the possible linear regressions of ℓ consecutive points $(a,\log(\widehat{S}_N(a)))$, where ℓ is a fixed integer number. More formally, a contrast is minimized and the chosen scale $\tilde{a}(N)$ satisfies:

$$\frac{\log(\tilde{a}(N))}{\log N}\xrightarrow[N\to\infty]{\mathcal{P}}\frac{1}{2D'+1}.$$

By this way, the adaptive estimator $\tilde{D}_N$ of D for this scale $\tilde{a}(N)$ is such that :

$$\sqrt{\frac{N}{\tilde{a}(N)}}(\tilde{D}_N-D)\xrightarrow[N\to\infty]{\mathcal{L}}\mathcal{N}(0,\sigma^2_D),$$

with $\sigma^2_D>0$. As a consequence, the minimax rate of convergence $N^{D'/(1+2D')}$, up to a logarithm factor, for the estimation of the long-memory parameter D in this semi-parametric setting (see Giraitis *et al.*, 1997) is obtained by $\tilde{D}_N$.

The paper is organized as follows. In section 2, a central limit theorem for sample variance of wavelet coefficients is established. In section 3, the automatic selection of the scale is described, the asymptotic behavior of $\tilde{D}_N$ is studied. Simulations are proposed in section 4 and proofs are given in section 5.

2 A central limit theorem for the sample variance of wavelet coefficients

The first point that explains all that follows is the following asymptotic behavior of the variance of wavelet coefficients,

Property 1 *Under assumption A1 or A2 on X and assumption $W(m)$ on ψ, for $a>0$, $(d(a,b))_{b\in\mathbb{R}}$ is a zero mean Gaussian stationary process and it exists $M>0$ not depending on a such that*

- *Under Assumption A1, for all $a\geq 1$,*

$$\Big|\mathbb{E}(d^2(a,0))-f^*(0)K_{(\psi,1-D)}\cdot a^{1-D}\Big|\leq M\cdot a^{1-(D+D')}. \quad (4)$$

- *Under Assumption A2, for all $a>0$,*

$$\Big|\mathbb{E}(d^2(a,0))-f^*(0)K_{(\psi,1-D)}\cdot a^{1-D}-C_{D'}K_{(\psi,1-(D+D'))}\cdot a^{1-(D+D')}\Big|\leq M\big(a^{-1-D}+a^{1-(D+D'+D'')}\big). \quad (5)$$

The proof of this property, like all the other proofs, is in the last section of this paper. The paper of Moulines *et al.* (2006) provides the same results under weaker assumptions but for multiresolution wavelet analysis. As it was said in the introduction, this property allows an estimation of D from a log-log regression, as soon as a consistant estimator of $\mathbb{E}(d^2(a,0))$ is provided from a sample $(X_0, X_1, \dots, X_N)$ of the time series X. Define then the normalized wavelet coefficient such that

$$\tilde{d}(a,b) = \frac{d(a,b)}{\big(f^*(0)K_{(\psi,1-D)} \cdot a^{1-D}\big)^{1/2}} \quad \text{for } a>0 \text{ and } b \in \mathbb{R}. \tag{6}$$

From property 1, it is obvious that under Assumptions A1 or A2, it exists $M'>0$ satisfying for all $a>0$,

$$\Big|\mathbb{E}(\tilde{d}^2(a,0)) - 1\Big| \leq M' \cdot \frac{1}{a^{D'}}.$$

In view of using this formula for estimating D by a log-log regression, an estimator of the variance of $d(a,0)$ should be considered. Hence, in the sequel, a sample $(X_1, \dots, X_N)$ of the process X is supposed to be known, but the different parameters $(D,\, D',\, C_D, \dots)$ are unknown. Consider the sample variance and the normalized sample variance of the wavelet coefficient, for $0<a<N$,

$$S_N(a) = \frac{1}{[\frac{N}{a}]} \sum_{k=1}^{[\frac{N}{a}]} d^2(a,k-1) \quad \text{and} \quad \tilde{S}_N(a) = \frac{1}{[\frac{N}{a}]} \sum_{k=1}^{[\frac{N}{a}]} \tilde{d}^2(a,k-1). \tag{7}$$

The following proposition specifies a central limit theorem satisfying by $\log \tilde{S}_N(a)$, which provides a first step for obtaining the asymptotic properties of the estimator by log-log regression. More generally, the following multidimensional central limit theorem for a vector $(\log \tilde{S}_N(a_i))_i$ can be established,

Proposition 1 *Let X satisfy Assumption A1 or A2 and ψ the assumption $W(m)$. Define $\ell \in \mathbb{N}\setminus\{0,1\}$ and $(r_1, \cdots, r_\ell) \in (\mathbb{N}^*)^\ell$. Let $(a_n)_{n\in\mathbb{N}}$ be such that $N/a_N \underset{N\to\infty}{\longrightarrow} \infty$ and $o(a_N) = N^{(1/1+2D')}$. Then,*

$$\sqrt{\frac{N}{a_N}}\Big(\log \tilde{S}_N(r_i a_N)\Big)_{1\leq i\leq \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,D)\big), \tag{8}$$

with $\Gamma(r_1,\cdots,r_\ell,\psi,D) = (\gamma_{ij})_{1\leq i,j\leq \ell}$ the covariance matrix such that

$$\gamma_{ij} = \frac{2(r_ir_j)^{1+D}\Big(\int_{-\infty}^{\infty} \frac{\cos\lambda}{|\lambda|^{1-D}}\,d\lambda\Big)^2}{K^2_{(\psi,1-D)} d_{ij}} \sum_{m=-\infty}^{\infty} \Big(\int_0^1\int_0^1 \frac{\psi(t)\psi(t')}{|r_it - r_jt' + d_{ij}m|^D} dt\,dt'\Big)^2. \tag{9}$$

Remark 4 *One remarks that the asymptotic covariance matrix does not depend on the used scales since theses scales are large enough compared to the length of the sample. It could be interesting to find a function ψ minimizing Γ for all D. Finally, the expression of Γ allows its estimation using an estimation of D. This can be used for constructing a test of long range dependence (see a forthcoming paper).*

The CLT (8) implies the following CLT for the vector $(\log S_N(r_i a_N))_i$,

$$\sqrt{\frac{N}{a_N}}\Big(\log S_N(r_ia_N) - (1-D)\log(r_ia_N) - \log(f^*(0)K_{(\psi,1-D)})\Big)_{1\leq i\leq\ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell(0\,;\,\Gamma).$$

There is still a problem by using wavelet coefficients $d(a,b)$ when X is not a continuous process but a time series. Indeed, this coefficients can not be exactly computed from a path $(X_1,\dots,X_N)$ of the process X. However, they can be approximated by replacing integrals by Riemann sums (roughly speaking, the Mallat's cascade algorithm proceeds with such an approximation for a discrete wavelet). This problem was also studied in Bardet (2002), Bardet and Bertrand (2006) and Moulines *et al.* (2006) in similar frameworks. Thus, consider the following approximations of wavelet coefficients and their sample variance, with $a \in \mathbb{N}^*$,

$$e(a,b) = \frac{1}{\sqrt{a}} \sum_{k=1}^{a} \psi(\frac{k}{a}) X_{k+ab} \tag{10}$$

$$\text{and} \quad T_N(a) = \frac{1}{[N/a]} \sum_{k=1}^{[N/a]} e^2(a,k-1). \tag{11}$$

Proposition 2 *Under the assumptions of the Proposition 1,*

$$\sqrt{\frac{N}{a_N}}\Big(\log T_N(r_ia_N) - (1-D)\log(r_ia_N) - \log(f^*(0)K_{(\psi,1-D)})\Big)_{1\leq i\leq\ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,D)\big).$$

3 Adaptive estimator of long range dependent parameter using data driven optimal scales

In this section, we consider the approximated wavelet coefficients $e(a,b)$ and their empirical variance T_N (for a time series). However, all the sequel is still valuable using the coefficients $d(a,b)$ and their empirical variance S_N (for instance for a continuous process).

The previous central limit theorem can also be written as,

$$\big(\log T_N(r_i a_N)\big)_{1\le i\le \ell} - \big(\log(r_i a_N)\big)_{1\le i\le \ell} = A_N \cdot \begin{pmatrix} D \\ K \end{pmatrix} + \frac{1}{\sqrt{N/a_N}}\,\big(\varepsilon_i\big)_{1\le i\le \ell},$$

with $A_N = \begin{pmatrix} -\log(r_1 a_N) & -1 \\ : & : \\ -\log(r_\ell a_N) & -1 \end{pmatrix}$, $K = \log(f^*(0)\cdot K_{(\psi,1-D)})$ and $\big(\varepsilon_i\big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}_\ell\big(0\,;\,\Gamma(r_1,\cdots,r_\ell,\psi,D)\big)$.

Therefore, a log-log regression of $\big(T_N(r_i a_N)\big)_{1\le i\le \ell}$ on scales $\big(r_i a_N\big)_{1\le i\le \ell}$ provides an estimator $\begin{pmatrix} \widehat{D}(a_N) \\ \widehat{K}(a_N) \end{pmatrix}$ of $\begin{pmatrix} D \\ K \end{pmatrix}$ such that

$$\begin{pmatrix} \widehat{D}(a_N) \\ \widehat{K}(a_N) \end{pmatrix} = (A'_N \cdot A_N)^{-1}\cdot A'_N \cdot Y_{a_N}^{(r_1,\ldots,r_\ell)} \quad \text{with} \quad Y_{a_N}^{(r_1,\ldots,r_\ell)} = \big(\log T_N(r_i a_N) - \log(r_i a_N)\big)_{1\le i\le \ell}, \tag{12}$$

which satisfies the following central limit theorem,

Proposition 3 *Under the assumptions of the Proposition 1,*

$$\sqrt{\frac{N}{a_N}}\Big(\begin{pmatrix} \widehat{D}(a_N) \\ \widehat{K}(a_N) \end{pmatrix} - \begin{pmatrix} D \\ K \end{pmatrix}\Big) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}_2(0\,;\,(A'\cdot A)^{-1}\cdot A'\cdot \Gamma(r_1,\cdots,r_\ell,\psi,D)\cdot A\cdot (A'\cdot A)^{-1}), \tag{13}$$

with $A = \begin{pmatrix} \log(r_1) & 1 \\ : & : \\ \log(r_\ell) & 1 \end{pmatrix}$ *and* $\Gamma(r_1,\cdots,r_\ell,\psi,D)$ *given by (9).*

Hence, $\widehat{D}(a_N)$ is a semi-parametric estimator of D and its asymptotic mean square error can be minimized with an appropriate scales sequence (a_N) reaching the well-known minimax rate of convergence for the long-range dependence parameter in this semi-parametric setting (see for instance Giraitis *et al.*, 1997). Indeed,

Proposition 4 *Let X satisfy Assumption A1 and ψ the assumption $W(m)$. Let (a_N) be a sequence such that $a_N = N^{1/(1+2D')}$. Then, the estimator $\widehat{D}(a_N)$ is rate optimal in the minimax sense,* i.e.

$$\limsup_{N\to\infty} \sup_{D\in]0,1[} \sup_{f^*\in\mathcal{H}(D',C_{D'})} N^{\frac{2D'}{1+2D'}}\cdot \mathbb{E}[\widehat{D}(a_N)-D)^2] < +\infty.$$

In the previous Propositions 1, 3 and 2, the rate of convergence of scale a_N is given by the following condition,

$$o(a_N) = N^{1/(1+2D')} \quad \text{for} \quad D' \le 2.$$

Now, for ease of writing, consider that $a_N = N^\alpha$. Then, the previous conditions can be written as,

$$a_N = N^\alpha \quad \text{with} \quad \alpha > \alpha^* \quad \text{and} \quad \alpha^* = \frac{1}{1+2D'} \quad \big(\ge \frac{1}{5}\big). \tag{14}$$

Thus an optimal choice, which means the faster convergence rate of the estimator, is obtained when α is larger but closer to α^*. However α^* depends on D and D' that are unknown. For solving this problem, Veitch *et al.* (2003) proposed a procedure based on a Khi-squared test (constructed from a distance between the regression line and the different points $(\log T_N(r_i a_N), \log(r_i a_N))$. It seems to be an efficient and interesting numerical way of estimating D, but no theoretical proofs are provided (instead of the log-periodogram procedure which is proven to reach the minimax convergence rate, see Moulines and Soulier, 2003).

We propose a new procedure for data-driven selection of optimal scales, *i.e.* optimal α. Let $\ell \in \mathbb{N}\backslash\{0,1,2\}$ and for $\alpha \in (0,1)$, define the vector $Y_N(\alpha) = \big(\log T_N(i\cdot N^\alpha) - \log(i\cdot N^\alpha)\big)_{1\le i\le \ell}$ (by this way, $(r_1,\dots,r_\ell)=(1,\dots,\ell)$) and the matrix $A_N(\alpha) = \begin{pmatrix} -\log(N^\alpha) & -1 \\ : & : \\ -\log(\ell\cdot N^\alpha) & -1 \end{pmatrix}$. Define the contrast,

$$Q_N(\alpha,D,K) = \Big(Y_N(\alpha) - A_N(\alpha)\cdot\Big(\begin{array}{c} D \\ K\end{array}\Big)\Big)'\cdot\Big(Y_N(\alpha) - A_N(\alpha)\cdot\Big(\begin{array}{c} D \\ K\end{array}\Big)\Big),$$

which corresponds to a squared distance between the ℓ points $\big(\log(i\cdot N^\alpha)\,,\log T_N(i\cdot N^\alpha)-\log(i\cdot N^\alpha)\big)$ and a line. The aim is to minimize this contrast for the three different parameters. It is obvious that for a fixed $\alpha\in(0,1)$, Q is minimized from the previous least square regression and therefore,

$$Q_N(\widehat{\alpha}_N,\widehat{D}(a_N),\widehat{K}(a_N)) = \min_{(\alpha,D)\in(0,1)^2,K\in\mathbb{R}} Q_N(\alpha,D,K).$$

with $(\widehat{D}(a_N),\widehat{K}(a_N))$ obtained as in relation (12). However, like $\widehat{\alpha}_N$ has to be obtained from numerical computations, the interval $(0,1)$ can be discretized as follows,

$$\widehat{\alpha}_N \in \mathcal{A}_N = \Big\{\frac{2}{\log N}\,,\frac{3}{\log N}\,,\dots,\frac{\log[N/\ell]}{\log N}\Big\}.$$

Hence, if $\alpha\in\mathcal{A}_N$, it exists $k\in\{2,3,\dots,\log[N/\ell]\}$ such that $k=\alpha\cdot\log N$.

Remark 5 *This choice of discretization is implied by the following proof of the consistence of $\widehat{\alpha}_N$. If the interval $(0,1)$ is stepped in N^β points, with $\beta>0$, the used proof is not sufficient for proving this consistence. Finally, this is the same framework than the case of discrete wavelet transform usually considered (see for instance Veitch et al., 2003) but less restricted because $\log N$ can be replaced in the previous expression of $\mathcal{A}_N$ by any function of N such is negligible compared to all functions N^β with $\beta>0$ (for instance, $(\log N)^d$ or $d\cdot\log N$ can be used).*

As a consequence, define

$$\widehat{Q}_N(\alpha) = Q_N(\alpha,\widehat{D}(a_N),\widehat{K}(a_N)) \ \text{ for } \alpha\in\mathcal{A}_N,$$

and minimizing Q_N for variables (α,D,K) is implied by minimizing $\widehat{Q}_N$ for variable $\alpha\in\mathcal{A}_N$, that is

$$\widehat{Q}_N(\widehat{\alpha}_N) = \min_{\alpha\in\mathcal{A}_N}\widehat{Q}_N(\alpha).$$

From the previous central limit theorem, it can be shown that,

Proposition 5 *Let X satisfy Assumption A2 and ψ the assumption $W(m)$. Then,*

$$\widehat{\alpha}_N = \frac{\log\widehat{a}_N}{\log N} \ \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}}\ \alpha^* = \frac{1}{1+2D'}.$$

By this way, an estimator $\widehat{D'}_N$ of the parameter D' can be also proved to be consistent,

Corollary 1 *Let X satisfy Assumption A2 and ψ the assumption $W(m)$. Then,*

$$\widehat{D'}_N = \frac{1-\widehat{\alpha}_N}{2\widehat{\alpha}_N}\ \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}}\ D'.$$

The estimator $\widehat{\alpha}_N$ defines the selected scale $\widehat{a}_N$ such that $\widehat{a}_N = N^{\widehat{\alpha}_N}$. From a direct application of the proof of the proposition 5 (see the details in the proof of theorem 1), one obtains a precise information on the asymptotic behavior of $\widehat{a}_N$, that is,

$$\Pr\Big(\frac{N^{\alpha^*}}{(\log N)^\lambda}\le N^{\widehat{\alpha}_N}\le N^{\alpha^*}\cdot(\log N)^\mu\Big)\ \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}}\ 1, \tag{15}$$

for all positive real numbers λ and μ such that $\lambda > \frac{2}{(\ell-2)D'}$ and $\mu > \frac{12}{\ell-2}$. As a consequence, the selected scale is equal to N^{α^*} up to a logarithm factor.

Finally, the proposition 5 can be used to define an adaptive estimator of D. First, define the straightforward estimator $\widehat{\widehat{D}}_N = \widehat{D}(\widehat{a}_N)$, that should minimize the mean square error using $\widehat{a}_N$. However, providing asymptotic results for the estimator $\widehat{\widehat{D}}_N$ is problematic because $\Pr(\widehat{\alpha}_N \leq \alpha^*) > 0$ and we do not know to prove if $\mathbb{E}(\sqrt{N/\widehat{a}_N}(\widehat{\widehat{D}}_N - D)) = 0$ or not. For establishing a central limit theorem satisfied by an adaptive estimator $\tilde{D}_N$ of D, an adaptive scale sequence $(\tilde{a}_N) = (N^{\tilde{\alpha}_N})$ has to be defined to ensure $\Pr(\tilde{\alpha}_N \leq \alpha^*) \underset{N\to\infty}{\longrightarrow} 0$. The following theorem provides the asymptotic behavior of such estimator,

Theorem 1 *Let X satisfy Assumption A2. Define,*

$$\tilde{\alpha}_N = \widehat{\alpha}_N + \frac{3}{(\ell-2)\widehat{D'}_N} \cdot \frac{\log\log N}{\log N}, \quad \tilde{a}_N = N^{\tilde{\alpha}_N} = N^{\widehat{\alpha}_N} \cdot \big(\log N\big)^{\frac{3}{(\ell-2)\widehat{D'}_N}} \quad \textit{and} \quad \tilde{D}_N = \widehat{D}(\tilde{a}_N).$$

Then, with $\sigma_D^2 = (1\ 0) \cdot (A' \cdot A)^{-1} \cdot A' \cdot \Gamma \cdot A \cdot (A' \cdot A)^{-1} \cdot (1\ 0)'$,

$$\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big(\tilde{D}_N - D\big) \underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}(0\,;\sigma_D^2) \quad \textit{and} \quad \forall \rho > \frac{2(1+3D')}{(\ell-2)D'}, \quad \frac{N^{\frac{D'}{1+2D'}}}{(\log N)^{\rho}} \cdot \big|\tilde{D}_N - D\big| \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} 0. \tag{16}$$

Remark 6 *Both the adaptive estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ converge to D with a rate of convergence rate equal to the minimax rate of convergence $N^{\frac{D'}{1+2D'}}$ up to a logarithm factor (this result is a classical result in this semi-parametric framework). Unfortunately, the used method does not allow to prove that the mean square errors of both those estimators reach the optimal rate and therefore to know if there are oracles.*

4 Simulations

The different previous estimators are computed from sample of different processes satisfying Assumption A2. The results are obtained from 100 generated independent samples of each process. The concrete procedure of generation of these Gaussian processes are obtained from circulant matrix method and is detailed in Doukhan *et al.* (2003).

The simulations are realized for five different values of D, $D = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$, for three different values of N, $N = 10^3,\ 10^4,\ 10^5$ of four different Gaussian processes processes which satisfy Assumption A2 and therefore Assumption A1 (the article of Moulines *et al.*, 2006, gives a lot of details on this point):

1. the fractional Gaussian noise (fGn) of parameter $H = 2 - 2D$ and $\sigma^2 = 1$. The spectral density f_{fGn} of a fGn is included in $\mathcal{H}(2, C_{D'})$ (thus $D' = 2$);
2. the FARIMA[0,d,0] with Gaussian innovations and parameter d such that $d = (1-D)/2 \in]0, 0.5[$. The spectral density f_{FARIMA} of such a process is included in the set $\mathcal{H}(2, C_{D'})$ (thus $D' = 2$);
3. the FARIMA[1,d,1] with Gaussian innovations and parameter d such that $d = (1-D)/2 \in]0, 0.5[$ and $\phi = -0.3$, $\theta = -0.7$ where ϕ denotes the AR coefficient and θ the MA coefficient. The spectral density f_{FARIMA} of such a process is included in the set $\mathcal{H}(2, C_{D'})$ (thus $D' = 2$);
4. the Gaussian stationary processes $X^{(D')}$, such that its spectral density is
$$f_4(\lambda) = \frac{1}{\lambda^{1-D}}(1 + \lambda^{D'}) \quad \text{for } \lambda \in [-\pi, \pi], \tag{17}$$
with $D \in]0, 1[$ and $D' \in]0, 2]$. The following simulations are realized for $D' = 1$ a therefore $f_4 \in \mathcal{H}(1, C_{D'})$.

The different parameters of the method of estimation are:

- A mother wavelet ψ such that $\psi(t) = 100 \cdot t^2(t-1)^2(t^2 - t + 3/14)\mathbb{I}_{0\leq t\leq 1}$. Hence, $m = 1$ and the different conditions of assumption $W(1)$ are satisfied. Choose $m > 1$ or different kind of function ψ (like dilated Daubeshies mother wavelet) leads to "similar" results (except for the bias of $\widehat{D}_N$ and $\widetilde{D}_N$). However, if $m = 0$ or if $\psi(0) \neq 0$ or if $\widehat{\psi}$ has not a fast decreasing rate to ∞, then the MSE of the estimation is larger (in particular, $\widehat{\alpha}_N > \alpha^*$...).
- A number ℓ of points used for the log-log-regression such that $m = 15$. Several simulations realized with $\ell = 5$, $\ell = 10$ or $\ell = 20$ lead to few different results (the larger N and ℓ or the smaller N and ℓ the smaller the MSE of the estimator). A better choice could be to consider ℓ as an increasing function of N, but the theoretic expression of such a function requires new technical computations... An adaptive choice of ℓ could be also developed as it is in Giraitis *et al.* (2000).

- As it was specified in Remark 5, $\log N$ is replaced by $5 \cdot \log N$ in the expression of $\mathcal{A}_N$.

For each process, the cases $D = 0.1,\ 0.3,\ 0.5,\ 0.7$ and 0.9 are investigated for 100 independent replications of samples with length varying from 1000 to 100000. The results are presented in the following table 1:

			$D=0.1$	$D=0.3$	$D=0.5$	$D=0.7$	$D=0.9$
$N = 10^3 \longrightarrow$	fGn ($H = 2-2D$)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.04, 0.08	0.29, 0.37	0.50, 0.55	0.72, 0.76	0.94, 0.95
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.14, 0.17	0.12, 0.21	0.12, 0.15	0.12, 0.18	0.12, 0.14
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.20, 0.24	0.23, 0.28	0.23, 0.28	0.23, 0.28	0.24, 0.29
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.04, 0.04	0.04, 0.04	0.05, 0.04	0.05, 0.05	0.05, 0.06
	FARIMA(0, $\frac{1-D}{2}$, 0)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.10, 0.11	0.34, 0.36	0.57, 0.57	0.76, 0.79	0.96, 0.98
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.16, 0.17	0.14, 0.16	0.13, 0.15	0.12, 0.17	0.13, 0.14
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.21, 0.24	0.22, 0.27	0.23, 0.28	0.23, 0.28	0.23, 0.28
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.05, 0.07	0.04, 0.04	0.05, 0.04	0.05, 0.04	0.04, 0.03
	FARIMA(1, $\frac{1-D}{2}$, 1)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	-0.31, -0.30	-0.23, -0.22	-0.10, -0.08	0.21, 0.26	0.52, 0.65
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.12, 0.15	0.18, 0.21	0.24, 0.30	0.31, 0.36	0.28, 0.38
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.19, 0.21	0.19, 0.20	0.20, 0.21	0.24, 0.27	0.27, 0.31
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.05, 0.07	0.05, 0.07	0.06, 0.08	0.07, 0.09	0.08, 0.09
	$X^{(D')}$, $D'=1$	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.59, 0.54	0.79, 0.74	0.98, 0.93	1.19, 1.14	1.41, 1.38
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.16, 0.21	0.16, 0.17	0.17, 0.19	0.17, 0.18	0.19, 0.18
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.23, 0.28	0.24, 0.28	0.25, 0.29	0.24, 0.29	0.24, 0.29
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.04, 0.04	0.04, 0.04	0.05, 0.05	0.05, 0.05	0.05, 0.04

			$D=0.1$	$D=0.3$	$D=0.5$	$D=0.7$	$D=0.9$
$N = 10^4 \longrightarrow$	fGn ($H = 2-2D$)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.07, 0.08	0.26, 0.27	0.48, 0.48	0.69, 0.69	0.91, 0.91
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.04, 0.04	0.04, 0.04	0.04, 0.03	0.04, 0.04	0.04, 0.04
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.18, 0.18	0.18, 0.18	0.17, 0.18	0.18, 0.19	0.18, 0.18
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.04, 0.05	0.04, 0.05	0.03, 0.05	0.04, 0.05	0.04, 0.05
	FARIMA(0, $\frac{1-D}{2}$, 0)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.14, 0.14	0.34, 0.34	0.53, 0.53	0.74, 0.73	0.93, 0.93
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.06, 0.06	0.04, 0.05	0.05, 0.05	0.05, 0.05	0.04, 0.04
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.19, 0.20	0.18, 0.19	0.18, 0.20	0.18, 0.19	0.18, 0.20
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.04, 0.05	0.03, 0.04	0.03, 0.04	0.04, 0.05	0.04, 0.05
	FARIMA(1, $\frac{1-D}{2}$, 1)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	-0.02, 0.02	0.01, 0.05	0.36, 0.39	0.57, 0.63	0.76, 0.81
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.12, 0.14	0.24, 0.27	0.20, 0.20	0.12, 0.15	0.13, 0.14
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.30, 0.33	0.24, 0.26	0.30, 0.34	0.30, 0.34	0.31, 0.34
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.04, 0.05	0.09, 0.11	0.06, 0.08	0.04, 0.05	0.04, 0.05
	$X^{(D')}$, $D'=1$	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.49, 0.45	0.66, 0.63	0.87, 0.83	1.08, 1.05	1.29, 1.26
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.14, 0.13	0.12, 0.12	0.13, 0.11	0.13, 0.11	0.15, 0.13
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.22, 0.24	0.23, 0.25	0.23, 0.25	0.22, 0.24	0.22, 0.24
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.05, 0.06	0.05, 0.06	0.05, 0.06	0.05, 0.06	0.05, 0.06

			$D=0.1$	$D=0.3$	$D=0.5$	$D=0.7$	$D=0.9$
$N = 10^5 \longrightarrow$	fGn ($H = 2-2D$)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.07, 0.07	0.27, 0.27	0.47, 0.48	0.68, 0.69	0.93, 0.90
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.01, 0.02	0.01, 0.01	0.01, 0.01	0.01, 0.02	0.01, 0.01
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.14, 0.15	0.14, 0.15	0.14, 0.15	0.15, 0.16	0.14, 0.15
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.04, 0.04	0.03, 0.04	0.03, 0.04	0.03, 0.04	0.03, 0.04
	FARIMA(0, $\frac{1-D}{2}$, 0)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.12, 0.11	0.31, 0.31	0.51, 0.51	0.71, 0.71	0.91, 0.91
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.02, 0.02	0.02, 0.02	0.02, 0.02	0.02, 0.02	0.02, 0.02
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.18, 0.20	0.18, 0.20	0.18, 0.19	0.18, 0.20	0.17, 0.18
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.02, 0.03	0.03, 0.03	0.03, 0.03	0.03, 0.04	0.03, 0.03
	FARIMA(1, $\frac{1-D}{2}$, 1)	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.06, 0.08	0.26, 0.28	0.45, 0.48	0.65, 0.67	0.86, 0.89
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.04, 0.04	0.03, 0.04	0.03, 0.03	0.04, 0.04	0.04, 0.04
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.29, 0.32	0.30, 0.33	0.29, 0.32	0.29, 0.32	0.30, 0.33
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.03, 0.04	0.03, 0.04	0.03, 0.04	0.03, 0.04	0.03, 0.03
	$X^{(D')}$, $D'=1$	mean $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.28, 0.25	0.48, 0.45	0.67, 0.64	0.87, 0.83	1.07, 1.03
		std $\widehat{\widehat{D}}_N$, $\tilde{D}_N$	0.10, 0.20	0.10, 0.24	0.13, 0.30	0.14, 0.42	0.13, 0.51
		mean $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.27, 0.30	0.27, 0.30	0.28, 0.31	0.28, 0.31	0.29, 0.32
		std $\widehat{\alpha}_N$, $\tilde{\alpha}_N$	0.05, 0.05	0.04, 0.05	0.05, 0.06	0.04, 0.05	0.05, 0.06

Table 1: Estimation of the different parameters from 100 independent samples of the different processes.

Comments on the estimation of D: These numerical results nearly follow the theoretic results. An important point is that the numerical convergence rate of the estimation of D does not depend on the value of parameter D (even if D is close to 0). Another point is that there is still an important bias for $N = 10^3$ or $N = 10^4$ for the FARIMA(1,d,1) or the process $X^{(D')}$ (but for $N = 10^5$ the bias is quite reasonable). Another choice of mother wavelet ψ (Daubeshies 5) with $m > 1$ leads to better results in this particular case.

Automatic estimation of the onset of scaling: The previous numerical results show that $\widehat{\alpha}_N$ and $\tilde{\alpha}_N$ converge (very slowly) to the optimal rate α^*, that is 0.2 for the three first processes and 1/3 for the fourth one. Two remarks: first, for fGn and FARIMA[0,d,0] the estimators $\widehat{\alpha}_N$ and $\tilde{\alpha}_N$ decrease with N and finally overestimate 0.2 when $N = 10^5$. The choice of $\ell = 15$, for which the length of the scale band nearly corresponds to $N^{1/4}$ (for $N = 10^5$) may explain this behavior. Moreover, other simulations show that $\widehat{\alpha}_N$ and $\tilde{\alpha}_N$ are larger for $N = 10^6$ than for $N = 10^5$ and are close to 0.17. A second point the slow convergence rate of this estimation for FARIMA(1,d,1) and for $X^{(D')}$ with $D' = 1$. The following graphs exhibit the evolution with N

of the log-log plotting and the choice of the onset of scaling:

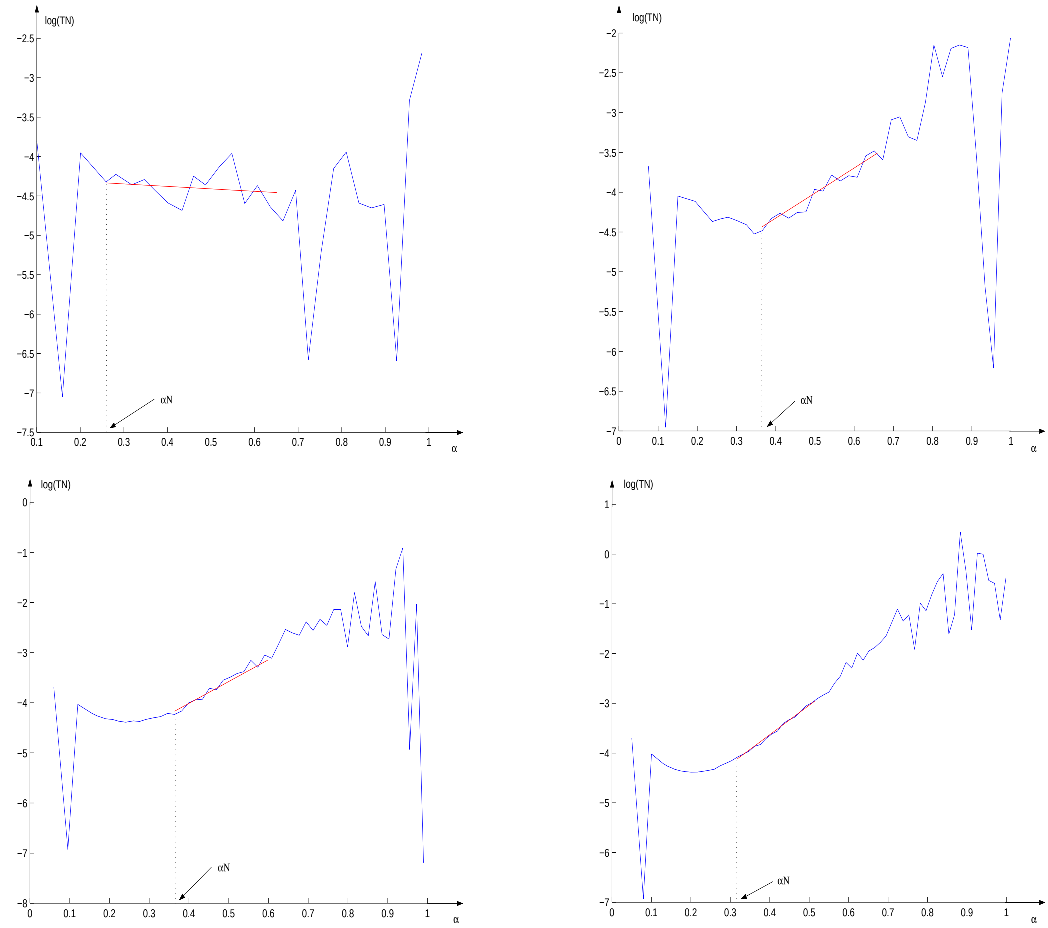

Figure 1: Log-log graphs for different samples of $X^{(D')}$ with $D = 0.5$ and $D' = 1$ when $N = 10^3$ (up and left, $\widehat{\widehat{D}}_N = 1.04$), $N = 10^4$ (up and right, $\widehat{\widehat{D}}_N = 0.66$), $N = 10^5$ (down and left, $\widehat{\widehat{D}}_N = 0.62$) and $N = 10^6$ (down and right, $\widehat{\widehat{D}}_N = 0.54$).

In Figure 1, it can be seen that that $\log T_N(i \cdot N^\alpha)$ is not a linear function of the logarithm of the scale $\log(i \cdot N^\alpha)$ when N increases and $\alpha < \alpha^*$ (it is exactly a consequence of Property 1: a bias exists in such a case). Moreover, if $\alpha > \alpha^*$ and α increases, a linear model appears with an error variance that increases.

Distribution of the estimators $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$: the following Figure 2 exhibits the histograms of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ for 100 independent samples of the process FARIMA(1,d,1) with $D = 0.5$:

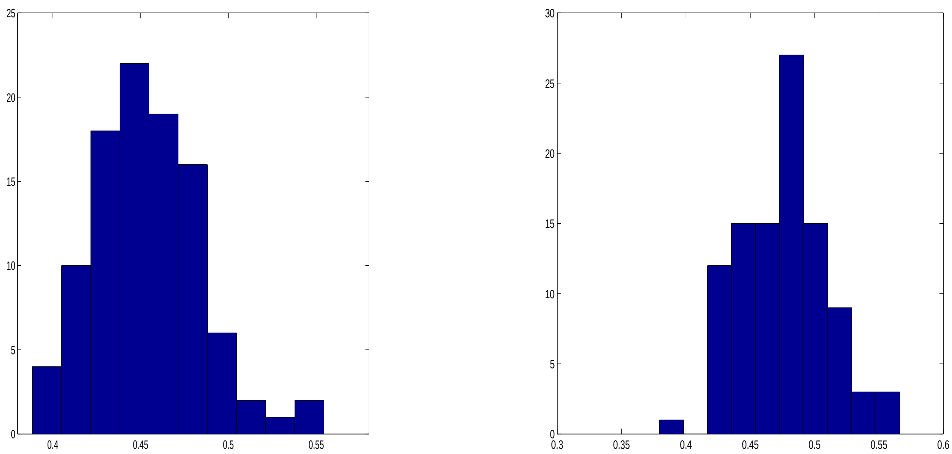

Figure 2: Histograms of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ for 100 samples of the process FARIMA(1,d,1) with $D = 0.5$ for $N = 10^5$.

Both these histograms are similar to Gaussian law histograms. It is not surprising for $\tilde{D}_N$ because Theorem 1 shows that the asymptotic law of $\tilde{D}_N$ is a Gaussian law with D mean. The asymptotic law of $\widehat{\widehat{D}}_N$ seems to be also close to a Gaussian law. A Cramer-von Mises test of normality indicates that both the distributions of $\widehat{\widehat{D}}_N$ and $\tilde{D}_N$ can be considered like Gaussian law (respectively $W \simeq 0.07$, $p-value \simeq 0.24$ and $W \simeq 0.05$, $p-value \simeq 0.54$).

5 Proofs

Proof [Property 1] First, it is obvious from the same arguments than in Abry *et al.* (1998) or Moulines *et al.* (2006) that,

$$\begin{aligned}
\mathbb{E}(d^2(a,0)) &= a \cdot \iint_{[0,1]^2} \psi(t)\psi(t')\mathbb{E}(X_{a\,t}X_{a\,t'})\,dt\,dt' \\
&= a \cdot \iint_{[0,1]^2} \psi(t)\psi(t')\,r(a(t-t'))\,dt\,dt' \\
&= a \iint_{[0,1]^2} \int_{-\pi}^{\pi} \psi(t)\psi(t')\,f(\lambda)e^{i\lambda a(t-t')}dt\,dt'\,d\lambda \\
&= \int_{-a\pi}^{a\pi} |\widehat{\psi}(u)|^2 f\big(\frac{u}{a}\big)\,du.
\end{aligned}$$

But for $1<c<3$,

$$K_{(\psi,c)} = \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(u)|^2}{|u|^c}\,du < \infty,$$

because when $u \to 0$, the hypothesis 1. of Assumption $W(m)$ implies that $|\widehat{\psi}(u)| = O(|u|^m)$ and when $u \to \infty$, the hypothesis 2. of Assumption $W(m)$ implies that $|\widehat{\psi}(u)|^2|u|^{-c} \leq |u|^{-2m'-c}$ and $2m'+c>3$. Moreover,

$$\begin{aligned}
\Big|\int_{-a\pi}^{a\pi} \frac{|\widehat{\psi}(u)|^2}{|u|^c}\,du - K_{(\psi,c)}\Big| &= 2\int_{a\pi}^{\infty} \frac{|\widehat{\psi}(u)|^2}{u^c}\,du \\
&\leq C \cdot \int_{a\pi}^{\infty} \frac{|\widehat{\psi}(u)|^2}{u^{2m'+c}}\,du \\
&\leq C' \cdot \frac{1}{a^{2m'+c-1}},
\end{aligned}$$

with $C>0$ and $C'>0$ not depending on a. As a consequence, under Assumption A1,

$$\begin{aligned}
\Big|\mathbb{E}(d^2(a,0)) - f^*(0)\cdot\int_{-\infty}^{\infty} \frac{|\widehat{\psi}(u)|^2}{|u/a|^{1-D}}\,du\Big| &\leq 2f^*(0)\cdot\int_{a\pi}^{\infty} \frac{|\widehat{\psi}(u)|^2}{(u/a)^{1-D}}\,du + C_{D'}\cdot\int_{-a\pi}^{a\pi} \frac{|\widehat{\psi}(u)|^2}{|u|^{1-(D+D')}}\,du \\
\Longrightarrow \Big|\mathbb{E}(d^2(a,0)) - f^*(0)K_{(\psi,1-D)}\cdot a^{1-D}\Big| &\leq \leq C' f^*(0)\cdot a^{1-2m'} + C_{D'}K_{(\psi,1-(D+D'))}\cdot a^{1-(D+D')}.
\end{aligned}$$

Under Assumption A2, from the same kind of inequalities, one obtains,

$$\Big|\mathbb{E}(d^2(a,0)) - f^*(0)K_{(\psi,1-D)}\cdot a^{1-D} - C_{D'}K_{(\psi,1-(D+D'))}\cdot a^{1-(D+D')}\Big|$$
$$\leq C'(f^*(0)+C_{D'})\cdot a^{1-2m'} + C_{D''}K_{(\psi,1-(D+D'+D''))}\cdot a^{1-(D+D'+D'')},$$

and this finishes the proof of this property. □

Proof [Proposition 1] This proof can be decomposed in three steps **Step 1**, **Step 2** and **Step 3**.

Step 1. In this part, $\sqrt{\dfrac{N}{a_N}}\cdot\mathrm{Cov}(\tilde{S}_N(r_i),\tilde{S}_N(r_j))_{1\leq i,j\leq \ell}$ is proved to converge to a covariance matrix Γ. First, for all $(i,j)\in\{1,\ldots,\ell\}^2$,

$$\mathrm{Cov}(\tilde{S}_N(r_i),\tilde{S}_N(r_j)) = 2\frac{1}{[N/r_ia_N]}\frac{1}{[N/r_ja_N]}\sum_{p=1}^{[N/r_ia_N]}\sum_{q=1}^{[N/r_ja_N]}\Big(\mathrm{Cov}(\tilde{d}(r_ia_N,p),\tilde{d}(r_ja_N,q)\Big)^2, \tag{18}$$

because X is a Gaussian process. Therefore, by considering only $i=j$ and $p=q$, for N and a_N large enough,

$$\mathrm{Cov}(\tilde S_N(r_i),\tilde S_N(r_j)) \geq \frac{1}{r_i}\frac{N}{a_N}. \tag{19}$$

Now, for $(p,q)\in\{1,\dots,[N/r_ia_N]\}\times\{1,\dots,[N/r_ia_N]\}$ such that $|r_ip-r_jq|\geq 2(r_i+r_j)$,

$$\mathrm{Cov}\big(\tilde d(r_ia_N,p),\tilde d(r_ja_N,q)\big) = \frac{a_N^D(r_ir_j)^{D/2}}{f^*(0)K_{(\psi,1-D)}}\int_0^1\int_0^1 \psi(t)\psi(t')r(a_N(r_it-r_jt')+a_N(r_ip-r_jq))\,dt\,dt'.$$

Now, the covariogram can be bounded from the expression of the spectral density under Assumption A1. Indeed, for $a>0$,

$$\begin{aligned} r(a) &= \frac{1}{a}\cdot\int_{-a\pi}^{a\pi} e^{-iu}f(u/a)\,du \\ &\Longrightarrow \Big|r(a)-2f^*(0)\Big(\int_0^\infty \frac{\cos\lambda}{\lambda^{1-D}}\,d\lambda-\int_{a\pi}^\infty\frac{\cos\lambda}{\lambda^{1-D}}\,d\lambda\Big)\frac{1}{a^D}\Big| \leq 2C_{D'}\Big|\int_0^{a\pi}\frac{\cos\lambda}{\lambda^{1-D-D'}}\,d\lambda\Big|\cdot\frac{1}{a^{D+D'}} \\ &\qquad\qquad \Longrightarrow \Big|r(a)-\frac{2f^*(0)}{a^D}\int_0^\infty\frac{\cos\lambda}{\lambda^{1-D}}\,d\lambda\Big|\leq M\cdot\big(\frac{1}{a}+\frac{1}{a^{D+D'}}\big), \end{aligned}$$

with a constant real number $M>0$ not depending on a, and because $\Big|\int_{a\pi}^\infty\frac{\cos\lambda}{\lambda^{1-D}}\,d\lambda\Big|\leq\frac{2}{(a\pi)^{1-D}}$ and $\Big|\int_0^{a\pi}\frac{\cos\lambda}{\lambda^{1-D-D'}}\,d\lambda\Big|\leq M'\cdot\max(1\,;\,a^{D+D'-1})$ for all $a>0$ (a constant real number $M'>0$ not depending on a).

Using this inequality, if (p,q) are such that $|r_ip-r_jq|\geq 2(r_i+r_j)$, on can write,

$$\begin{aligned} &\Big|\mathrm{Cov}\big(\tilde d(r_ia_N,p),\tilde d(r_ja_N,q)\big)-\frac{a_N^D(r_ir_j)^{D/2}}{f^*(0)K_{(\psi,1-D)}}\int_0^1\int_0^1\psi(t)\psi(t')\Big|1+a_N(r_ip-r_jq)\big(1+\frac{r_it-r_jt'}{r_ip-r_jq}\big)\Big|^{-D}dt\,dt'\Big| \\ &\qquad\leq 2M\frac{a_N^D(r_ir_j)^{D/2}}{f^*(0)K_{(\psi,1-D)}}\int_0^1\int_0^1|\psi(t)\psi(t')|\Big|a_N(r_ip-r_jq)\big(1+\frac{r_it-r_jt'}{r_ip-r_jq}\big)\Big|^{-\min(D+D',1)}dt\,dt' \\ &\qquad\leq\Big(\frac{2M\cdot\|\psi\|_\infty^2(r_ir_j)^{D/2}}{f^*(0)K_{(\psi,1-D)}}\Big(\frac{1}{2}\Big)^{-\min(D+D',1)}\Big)\frac{1}{|r_ip-r_jq|^{\min(D+D',1)}}\cdot\frac{1}{a_N^{\min(D',1-D)}}. \end{aligned}$$

In another hand, from a Taylor expansion,

$$\begin{aligned} &\Big|\int_0^1\int_0^1\psi(t)\psi(t')\left(\Big|(r_ip-r_jq)\big(1+\frac{r_it-r_jt'}{r_ip-r_jq}\big)\Big|^{-D}-|r_ip-r_jq|^{-D}\sum_{s=0}^{2m-1}b(s,D)\Big(\frac{r_it-r_jt'}{r_ip-r_jq}\Big)^s\right)dt\,dt'\Big| \\ &\qquad\leq|r_ip-r_jq|^{-D-2m}\|\psi\|_\infty^2(r_i+r_j)^{2m}b(2m,D), \end{aligned}$$

with $b(0,D)=1$ and $b(s+1,D)=(s-D)b(s,D)$ for $s\in\mathbb{N}$. Thanks to Assumption $W(m)$ satisfied by ψ, for all $s\in\{0,1,\dots,2m-1\}$, $\displaystyle\int_0^1\int_0^1\psi(t)\psi(t')\big(\frac{r_it-r_jt'}{r_ip-r_jq}\big)^s dt\,dt'=0$. As a consequence,

$$\Big|\int_0^1\int_0^1\psi(t)\psi(t')\Big|(r_ip-r_jq)\big(1+\frac{r_it-r_jt'}{r_ip-r_jq}\big)\Big|^{-D}dt\,dt'\Big|\leq|r_ip-r_jq|^{-D-2m}\|\psi\|_\infty^2(r_i+r_j)^{2m}b(2m,D),$$

and therefore, it exists $K_1>0$, such that for all $(i,j)\in\{1,\dots,\ell\}^2$, and for $(p,q)\in\{1,\dots,[N/r_ia_N]\}\times\{1,\dots,[N/r_ia_N]\}$ satisfying $|r_ip-r_jq|\geq 2(r_i+r_j)$,

$$\Big|\mathrm{Cov}\big(\tilde d(r_ia_N,p),\tilde d(r_ja_N,q)\big)\Big|\leq K_1\left(\frac{1}{|r_ip-r_jq|^{\min(D+D',1)}}\cdot\frac{1}{a_N^{\min(D',1-D)}}+\frac{1}{|r_ip-r_jq|^{D+2m}}\right). \tag{20}$$

Moreover, from Property 1, for a_N large enough and $(p,q)\in\{1,\dots,[N/r_ia_N]\}\times\{1,\dots,[N/r_ia_N]\}$ satisfying $|r_ip-r_jq|<2(r_i+r_j)$,

$$\Big|\mathrm{Cov}\big(\tilde d(r_ia_N,p),\tilde d(r_ja_N,q)\big)\Big|\leq\Big(\mathbb{E}(\tilde d^2(r_ia_N,p))\cdot\mathbb{E}(\tilde d^2(r_ia_N,p))\Big)^{1/2}\leq 2. \tag{21}$$

Thanks to (20) and (21), it exists a constant $K_2 > 0$ (not depending on N or a_N) such that the relation (18) becomes the inequality,

$$\mathrm{Cov}(\tilde S_N(r_i),\tilde S_N(r_j)) \quad\leq\quad K_2\cdot\frac{r_ir_ja_N^2}{N^2}\sum_{p=1}^{[N/r_ia_N]}\sum_{q=1}^{[N/r_ja_N]}\frac{a_N^{-2\min(D',1-D)}}{(1+|r_ip-r_jq|)^{2\min(D+D',1)}}+\frac{1}{(1+|r_ip-r_jq|)^{2D+4m}}.$$

But, from the theorem of comparison between sums and integrals, for $0<\alpha\leq 1$, N and a_N large enough,

$$\begin{aligned}\sum_{p=1}^{[N/r_ia_N]}\sum_{q=1}^{[N/r_ja_N]}(1+|r_ip-r_jq|)^{-\alpha} &\sim \int_0^{N/r_ia_N}\int_0^{N/r_ja_N}\frac{dx\,dy}{(1+|r_ix-r_jy|)^\alpha}\\ &\sim \frac{1}{r_ir_j}\int_0^{N/a_N}\int_0^{N/a_N}\frac{du\,dv}{(1+|u-v|)^\alpha}\\ &\sim \frac{2}{r_ir_j}\int_0^{N/a_N}\frac{(N/a_N-w)\,dw}{(1+w)^\alpha}\\ &\sim \begin{cases}\dfrac{2}{r_ir_j}\dfrac{1}{(1-\alpha)(2-\alpha)}\cdot\big(\dfrac{N}{a_N}\big)^{2-\alpha} & \text{if } 0<\alpha<1\\ \dfrac{2}{r_ir_j}\big(\dfrac{N}{a_N}\big)\log\big(\dfrac{N}{a_N}\big) & \text{if } \alpha=1\end{cases}\end{aligned}$$

In the same way, for $\alpha>1$,

$$\begin{aligned}\sum_{p=1}^{[N/r_ia_N]}\sum_{q=1}^{[N/r_ja_N]}(1+|r_ip-r_jq|)^{-\alpha} &\leq \frac{1}{r_ir_j}\int_0^{N/a_N}\int_0^{N/a_N}\frac{du\,dv}{(1+|u-v|)^\alpha}\\ &\leq \frac{2}{r_ir_j}\int_0^{N/a_N}\frac{N/a_N\,dw}{(1+w)^\alpha}\\ &\leq \frac{2}{(\alpha-1)r_ir_j}\cdot\frac{N}{a_N}.\end{aligned}$$

As a consequence, with $m\geq 1$, one obtains the following bound for N and a_N large enough,

$$\mathrm{Cov}(\tilde S_N(r_ia_N),\tilde S_N(r_ja_N)) \quad\leq\quad K_3\Big(\big(\frac{N}{a_N}\big)^{-1}+\log\big(\frac{N}{a_N}\big)\times\big(\frac{N}{a_N}\big)^{-2(D+D')}\frac{1}{a_N^{2D'}}+\big(\frac{N}{a_N}\big)^{-1}\frac{1}{a_N^{2(1-D)}}\Big), \quad (22)$$

with $K_3>0$ and not depending on N or a_N. Finally, from (19) and (22), if the sequence (a_n) is such that $o(a_N)=\log(N)N^{(1-2(D+D'))/(1-2D)}\mathbb{I}_{2(D+D')\leq 1}+\mathbb{I}_{2(D+D')>1}$, then,

$$\lim_{N\to\infty}\frac{N}{a_N}\Big(\mathrm{Cov}(\tilde S_N(r_ia_N),\tilde S_N(r_ja_N))\Big)_{1\leq i,j\leq \ell}=\Gamma, \quad (23)$$

with $\Gamma=(\gamma_{ij})_{1\leq i,j\leq\ell}$ a non null symmetric matrix that can be specified. Indeed, from the previous inequalities,

$$\begin{aligned}\gamma_{ij} &= \lim_{N\to\infty}\frac{2r_ir_ja_N}{N}\sum_{p=1}^{[N/r_ia_N]}\sum_{q=1}^{[N/r_ja_N]}\Big(\frac{(r_ir_j)^{D/2}}{K_{(\psi,1-D)}}\int_0^1\int_0^1\frac{\psi(t)\psi(t')}{|r_it-r_jt'+r_ip-r_jp'|^D}dt\,dt'\int_{-\infty}^{\infty}\frac{\cos\lambda}{|\lambda|^{1-D}}\,d\lambda\Big)^2\\ &= \lim_{N\to\infty}\frac{2(r_ir_j)^{1+D}\Big(\int_{-\infty}^\infty\frac{\cos\lambda}{|\lambda|^{1-D}}\,d\lambda\Big)^2a_N}{K^2_{(\psi,1-D)}N}\sum_{m=-[N/d_{ij}a_N]+1}^{[N/d_{ij}a_N]-1}\big(\frac{N}{d_{ij}a_N}-|m|\big)\Big(\int_0^1\int_0^1\frac{\psi(t)\psi(t')}{|r_it-r_jt'+d_{ij}m|^D}dt\,dt'\Big)^2\\ &= \frac{2(r_ir_j)^{1+D}\Big(\int_{-\infty}^\infty\frac{\cos\lambda}{|\lambda|^{1-D}}\,d\lambda\Big)^2}{K^2_{(\psi,1-D)}d_{ij}}\sum_{m=-\infty}^{\infty}\Big(\int_0^1\int_0^1\frac{\psi(t)\psi(t')}{|r_it-r_jt'+d_{ij}m|^D}dt\,dt'\Big)^2,\end{aligned}$$

with $d_{ij}=GCD(r_i\,;\,r_j)$. Therefore, the matrix Γ is only depending on $r_1,\cdots,r_\ell,\psi,D$.

Step 2. In a general frame, the previous result is not sufficient for obtaining the central limit theorem,

$$\sqrt{\frac{N}{a_N}}\Big(\tilde S_N(r_ia_N)-\mathbb{E}(\tilde d^2(r_ia_N,0)\Big)_{1\leq i\leq\ell}\ \xrightarrow[N\to\infty]{\mathcal{L}}\ \mathcal{N}_\ell(0,\Gamma(r_1,\cdots,r_\ell,\psi,D)). \quad (24)$$

However, each $\tilde{S}_N(r_i a_N)$ is a quadratic form of Gaussian process. *Mutatis mutandis*, it is exactly the same framework (*i.e.* a Lindeberg central limit theorem) as that of the Proposition 2.1 in Bardet (2000), and (24) is checked. Moreover, if $(a_n)_n$ is such that $o(a_N) = N^{1/1+2D'}$ when $D + D' < 1$ (which implies $o(a_N) = \log(N) N^{(1-2(D+D'))/(1-2D)}$ in the particular case $2(D + D') < 1$), then,

$$\sqrt{\frac{N}{a_N}}\Big(a_N^{-2} + a_N^{-D'}\Big) \underset{N\to\infty}{\longrightarrow} 0.$$

As a consequence, under those assumptions,

$$\sqrt{\frac{N}{a_N}}\Big(\tilde{S}_N(r_i a_N) - 1\Big)_{1\le i\le \ell} \underset{N\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}_\ell(0, \Gamma(r_1, \cdots, r_\ell, \psi, D)). \tag{25}$$

Step 3. The logarithm function $(x_1, .., x_\ell) \in]0, +\infty[^\ell \mapsto (\log x_1, .., \log x_m)$ is $\mathcal{C}^2$ on $]0, +\infty[^\ell$. As a consequence, using the Delta-method, the central limit theorem (8) for the vector $\Big(\log \tilde{S}_N(r_i a_N)\Big)_{1\le i\le\ell}$ follows with the same asymptotical covariance matrix $\Gamma(r_1, \cdots, r_\ell, \psi, D)$ (because the Jacobian matrix of the function in $(1, .., 1)$ is the identity matrix). □

Proof [Proposition 2] Since the variables $d = d(a, k)$ and $e = e(a, k)$ are Gaussian, the variables $d^2 - e^2$ have finite second order moment and Jensen's inequality implies

$$\begin{aligned}\mathbb{E}\,|S_N(a) - T_N(a)| &= \mathbb{E}\left|\frac{1}{[N/a]}\sum_{k=1}^{[N/a]}\big(d^2(a,k) - e^2(a,k)\big)\right| \\ &\le \frac{1}{[N/a]}\sum_{k=1}^{[N/a]}\sqrt{\mathbb{E}\,(d^2(a,k) - e^2(a,k))^2}\end{aligned}$$

From the same computations than in Bardet and Bertrand (2006), one obtains,

$$\mathbb{E}\,|S_N(a) - T_N(a)| \le \frac{\sqrt{6}}{[N/a]}\sum_{k=1}^{[N/a]}\sqrt{\mathbb{E}\varepsilon^2(a,k)} \times \sqrt{\mathbb{E}\,[4\,d^2(a,k) + \varepsilon^2(a,k)]}. \tag{26}$$

But

$$\begin{aligned}\mathbb{E}\varepsilon^2(a,b) &\le \frac{1}{a}\left(\sum_{k=0}^{a-1}\int_{k/a}^{(k+1)/a} dt\,\sqrt{\mathbb{E}\Big(\psi(t)X_{a(t+b)} - \psi(\frac{k}{a})X_{k+ab}\Big)^2}\right)^2 \\ &\le \frac{1}{a}\left(\sum_{k=0}^{a-1}\int_{k/a}^{(k+1)/a} dt\,\sqrt{r(0)\big(\psi(t) - \psi(\frac{k}{a})\big)^2}\right)^2,\end{aligned}$$

with $r(t) = r(0)$ for all $t \in [0, 1[$. The function ψ is supposed to be a $\mathcal{C}^1$ function. As a consequence, a Taylor expansion implies,

$$\begin{aligned}\mathbb{E}\varepsilon^2(a,b) &\le \frac{1}{a}\cdot r(0)\cdot\|\psi'\|_\infty\left(\sum_{k=0}^{a-1}\int_{k/a}^{(k+1)/a}(t - \frac{k}{a})\,dt\right)^2 \\ &\le r(0)\cdot\|\psi'\|_\infty\cdot\frac{1}{4a^3}.\end{aligned}$$

Thus, for a large enough, it exists $K > 0$ such that,

$$\mathbb{E}\,|S_N(a) - T_N(a)| \le \Big(\frac{\sqrt{6}}{2}\sqrt{r(0)\cdot\|\psi'\|_\infty}\Big)\,a^{-1-D/2}.$$

Therefore, if $o(a_N) = N^{1/1+2D'}$, then,

$$\lim_{N\to\infty}\sqrt{\frac{N}{a_N}} \times \mathbb{E}\,|S_N(a_N) - T_N(a_N)| = 0.$$

From central limit theorem (2), this implies the convergence of the finite-dimensional distribution,

$$\sqrt{\frac{N}{a_N}}\Big(T_N(r_i a_N) + C \cdot C_{(\psi,d)} \cdot (r_i a_N)^{1-D}\Big)_{1\leq i\leq \ell} \xrightarrow[N\to\infty]{D} \mathcal{N}_\ell(0\,;\Gamma).$$

Now the Delta-method can be applied and this finishes the proof. $\square$

Proof [Proposition 5] Let $\varepsilon > 0$ be a fixed positive real number, such that $\alpha^* + \varepsilon < 1$.

I. First, a bound of $\Pr(\widehat{\alpha}_N \leq \alpha^* + \varepsilon)$ is provided. Indeed,

$$\begin{aligned}
\Pr\left(\widehat{\alpha}_N \leq \alpha^* + \varepsilon\right) &\geq \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) \leq \min_{\alpha\geq\alpha^*+\varepsilon \text{ and } \alpha\in\mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \\
&\geq 1 - \Pr\Big(\bigcup_{\alpha\geq\alpha^*+\varepsilon \text{ and } \alpha\in\mathcal{A}_N} \widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \\
&\geq 1 - \sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/\ell]} \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N\big(\frac{k}{\log N}\big)\Big). \qquad (27)
\end{aligned}$$

But, for $\alpha \geq \alpha^* + 1$,

$$\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) = \Pr\Big(\Big\|P_N(\alpha^*+\varepsilon/2)\cdot Y_N(\alpha^*+\varepsilon/2)\Big\|^2 > \Big\|P_N(\alpha)\cdot Y_N(\alpha)\Big\|^2\Big)$$

with $P_N(\alpha) = I_\ell - A_N(\alpha)\cdot\big(A_N'(\alpha)\cdot A_N(\alpha)\big)^{-1}\cdot A_N(\alpha)$ for all $\alpha\in(0,1)$, *i.e.* $P_N(\alpha)$ is the matrix of orthogonal projection on the orthogonal subspace (in $\mathbb{R}^\ell$) generated by $A_N(\alpha)$ (and I_ℓ is the identity matrix in $\mathbb{R}^\ell$). From the expression of $A_N(\alpha)$, it is obvious that for all $\alpha\in(0,1)$,

$$P_N(\alpha) = P = I_\ell - A\cdot\big(A'\cdot A\big)^{-1}\cdot A,$$

with the matrix $A = \begin{pmatrix} \log(r_1) & 1 \\ \vdots & \vdots \\ \log(r_\ell) & 1\end{pmatrix}$ like in Proposition 3. By this way,

$$\begin{aligned}
\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) &= \Pr\Big(\Big\|P\cdot Y_N(\alpha^*+\varepsilon/2)\Big\|^2 > \Big\|P\cdot Y_N(\alpha)\Big\|^2\Big) \\
&= \Pr\left(\Big\|P\cdot\sqrt{\frac{N}{N^{\alpha^*+\varepsilon/2}}}\,Y_N(\alpha^*+\varepsilon/2)\Big\|^2 > N^{\alpha-(\alpha^*+\varepsilon/2)}\Big\|P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)\Big\|^2\right) \\
&\leq \Pr\Big(V_N(\alpha^*+\varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) + \Pr\Big(V_N(\alpha) \leq N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big)
\end{aligned}$$

with $V_N(\alpha) = \Big\|P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)\Big\|^2$ for all $\alpha\in(0,1)$. From Proposition 2, for all $\alpha > \alpha^*$, the asymptotic law of $P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)$ is a Gaussian law with covariance matrix $P\cdot\Gamma\cdot P'$. Moreover, the rank of the matrix is $P\cdot\Gamma\cdot P'$ is $\ell-2$ (this is the rank of P) and it exists $0<\lambda_-$, not depending on N) such that $P\cdot\Gamma\cdot P' - \lambda_- P\cdot P'$ is a non-negative matrix ($0<\lambda_-<\min\{\lambda\in\mathrm{Sp}(\Gamma)\}$). As a consequence, for N large enough,

$$\begin{aligned}
\Pr\Big(V_N(\alpha) \leq N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) &\leq 2\cdot\Pr\Big(V_- \leq N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) \\
&\leq \frac{1}{2^{\ell/2-2}\Gamma(\ell/2)}\cdot\Big(\frac{N}{\lambda_-}\Big)^{-(\frac{\ell}{2}-1)\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}},
\end{aligned}$$

with $V_- \sim \lambda_-\cdot\chi^2(\ell-2)$. Moreover, from Markov inequality,

$$\begin{aligned}
\Pr\Big(V_N(\alpha^*+\varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) &\leq 2\cdot\Pr\Big(\exp(\sqrt{V_+}) > \exp\big(N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big)\Big) \\
&\leq 2\cdot\mathbb{E}(\exp(\sqrt{V_+}))\cdot\exp\big(-N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big)
\end{aligned}$$

with $V_+ \sim \lambda_+ \cdot \chi^2(\ell-2)$ and $\lambda_+ > \max\{\lambda \in \mathrm{Sp}(\Gamma)\} > 0$. Like $\mathbb{E}(\exp(\sqrt{V_+})) < \infty$ is not depending on N, one obtains that it exists $M_1 > 0$ not depending on N, such that for N large enough,

$$\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \le M_1 \cdot N^{-(\frac{\ell}{2}-1)\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}},$$

and therefore, the inequality (27) becomes, for N large enough,

$$\begin{aligned}\Pr\big(\widehat{\alpha}_N \le \alpha^*+\varepsilon\big) &\ge 1 - M_1 \cdot \sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/\ell]} N^{-\frac{(\ell-2)}{4}\Big(\big(\frac{k}{\log N}\big)-(\alpha^*+\varepsilon/2)\Big)} \\ &\ge 1 - M_1 \cdot \log N \cdot N^{-\frac{(\ell-2)}{12}\varepsilon}. \qquad (28)\end{aligned}$$

II. Secondly, a bound of $\Pr(\widehat{\alpha}_N \ge \alpha^*-\varepsilon)$ is provided. From the same arguments and notations than previously,

$$\begin{aligned}\Pr\big(\widehat{\alpha}_N \ge \alpha^*-\varepsilon\big) &\ge \Pr\Big(\widehat{Q}_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon) \le \min_{\alpha\le\alpha^*-\varepsilon \text{ and } \alpha\in\mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \\ &\ge 1 - \sum_{k=2}^{[(\alpha^*-\varepsilon)\log N]+1} \Pr\Big(\widehat{Q}_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\frac{k}{\log N})\Big), \qquad (29)\end{aligned}$$

and like previously,

$$\begin{aligned}&\Pr\Big(\widehat{Q}_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\Big) \\ &\quad = \Pr\Bigg(\Big\|P\cdot\sqrt{\frac{N}{N^{\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon}}}\,Y_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)\Big\|^2 > N^{\alpha-(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)}\Big\|P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)\Big\|^2\Bigg). \qquad (30)\end{aligned}$$

Now, in the case $a_N = N^\alpha$ with $\alpha \le \alpha^*$, the sample variance of wavelet coefficients is biased. In such a case, from the relation of Property 1 under Assumption A2,

$$\Big(Y_N(\alpha)\Big)_{1\le i\le\ell} = \Big(\frac{C_{D'}K_{(\psi,1-(D+D'))}}{f^*(0)K_{(\psi,1-D)}}(iN^\alpha)^{-D'}(1+o_i(1))\Big)_{1\le i\le\ell} + \Big(\sqrt{\frac{N^\alpha}{N}}\cdot\varepsilon_N(\alpha)\Big)_{1\le i\le\ell},$$

with $o_i(1) \to 0$ when $N\to\infty$ for all i and $\mathbb{E}(Z_N(\alpha)) = 0$. As a consequence, for N large enough,

$$\begin{aligned}\Big\|P\cdot\sqrt{\frac{N}{N^\alpha}}\,Y_N(\alpha)\Big\|^2 &= \Big\|P\cdot\varepsilon_N(\alpha)\Big\|^2 + N^{\frac{\alpha^*-\alpha}{\alpha^*}}\Big\|P\cdot\Big(\frac{C_{D'}K_{(\psi,1-(D+D'))}}{f^*(0)K_{(\psi,1-D)}}i^{-D'}(1+o_i(1))\Big)_{1\le i\le\ell}\Big\|^2 \\ &\ge D\cdot N^{\frac{\alpha^*-\alpha}{\alpha^*}},\end{aligned}$$

with $D>0$, because the vector $(i^{-D'})_{1\le i\le\ell}$ is not in the orthogonal subspace of the subspace generated by the matrix A. Then, the relation (30) becomes,

$$\begin{aligned}\Pr\Big(\widehat{Q}_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\Big) &\le \Pr\Big(\Big\|P\cdot\sqrt{\frac{N}{N^{\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon}}}\,Y_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)\Big\|^2 \ge D\cdot N^{\alpha-(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)}\cdot N^{\frac{\alpha^*-\alpha}{\alpha^*}}\Big) \\ &\le \Pr\Big(V_+ \ge D\cdot N^{\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon)}\Big) \\ &\le M_2\cdot N^{-(\frac{\ell}{2}-1)\frac{1-\alpha^*}{2\alpha^*}\varepsilon},\end{aligned}$$

with $M_2>0$, because $V_+ \sim \lambda_+\cdot\chi^2(\ell-2)$ and $\dfrac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon) \ge \dfrac{1-\alpha^*}{2\alpha^*}\varepsilon$ for all $\alpha\le\alpha^*-\varepsilon$. Hence, from the inequality (29), for N large enough,

$$\Pr\big(\widehat{\alpha}_N \ge \alpha^*-\varepsilon\big) \ge 1 - M_2\cdot\log N\cdot N^{-(\frac{\ell}{2}-1)\frac{1-\alpha^*}{2\alpha^*}\varepsilon}. \qquad (31)$$

The inequalities (28) and (31) imply that $\Pr\big(|\widehat{\alpha}_N-\alpha|\ge\varepsilon\big) \underset{N\to\infty}{\longrightarrow} 0$. □

Proof [Theorem 1] The central limit theorem of (16) can be established from the following arguments. First, $\Pr(\tilde{\alpha}_N > \alpha^*) \underset{N\to\infty}{\longrightarrow} 1$. Indeed, thanks to the previous proof, for all $\varepsilon>0$,

$$\Pr\big(\widehat{\alpha}_N \ge \alpha^*-\varepsilon\big) \ge 1 - M_2\cdot\log N\cdot N^{-(\frac{\ell}{2}-1)\frac{1-\alpha^*}{2\alpha^*}\varepsilon}.$$

Consequently, if $\varepsilon_N = \lambda \cdot \dfrac{\log\log N}{\log N}$ with $\lambda > \dfrac{2}{(\ell-2)D'}$ then,

$$\begin{aligned} \Pr\big(\widehat{\alpha}_N \geq \alpha^* - \varepsilon_N\big) &\geq 1 - M_2 \cdot \log N \cdot N^{-\lambda \frac{(\ell-2)D'}{2} \cdot \frac{\log\log N}{\log N}} \\ &\geq 1 - M_2 \cdot \big(\log N\big)^{1-\lambda \frac{(\ell-2)D'}{2}} \\ \Longrightarrow \quad & \Pr\big(\widehat{\alpha}_N + \varepsilon_N \geq \alpha^*\big) \underset{N\to\infty}{\longrightarrow} 1. \end{aligned}$$

Now, from Corollary 1, $\widehat{D'}_N \overset{\mathcal{P}}{\underset{N\to\infty}{\longrightarrow}} D'$. Therefore, $\Pr\big(\widehat{D'}_N \leq \frac{4}{3}D'\big) \underset{N\to\infty}{\longrightarrow} 1$. Thus, with $\lambda \geq \dfrac{9}{4(\ell-2)D'}$, $\Pr\big(\tilde{\alpha}_N + \big(\varepsilon_N - \dfrac{3}{(\ell-2)\widehat{D'}_N} \cdot \dfrac{\log\log N}{\log N}\big) \geq \alpha^*\big) \underset{N\to\infty}{\longrightarrow} 1$ which implies $\Pr(\tilde{\alpha}_N > \alpha^*) \underset{N\to\infty}{\longrightarrow} 1$.
Secondly, for $x \in_R$,

$$\begin{aligned} \lim_{N\to\infty} \Pr\Big(\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big(\tilde{D}_N - D\big) \leq x\Big) &= \lim_{N\to\infty} \Pr\Big(\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big(\tilde{D}_N - D\big) \leq x \bigcap \tilde{\alpha}_N > \alpha^*\Big) \\ &\quad + \lim_{N\to\infty} \Pr\Big(\sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big(\tilde{D}_N - D\big) \leq x \bigcap \tilde{\alpha}_N \leq \alpha^*\Big) \\ &= \lim_{N\to\infty} \int_{\alpha^*}^1 \Pr\Big(\sqrt{\frac{N}{N^{\alpha}}}\big(\tilde{D}_N - D\big) \leq x\Big) f_{\widehat{\alpha}_N}(\alpha)\, d\alpha \\ &= \lim_{N\to\infty} \Pr\Big(Z_\Gamma \leq x\Big) \cdot \int_{\alpha^*}^1 f_{\widehat{\alpha}_N}(\alpha)\, d\alpha \\ &= \Pr\Big(Z_\Gamma \leq x\Big), \end{aligned}$$

with $f_{\widehat{\alpha}_N}(\alpha)$ the probability density function of $\widehat{\alpha}_N$ and $Z_\Gamma \sim \mathcal{N}(0\,;\,(A' \cdot A)^{-1} \cdot A' \cdot \Gamma \cdot A \cdot (A' \cdot A)^{-1})$.

For proving the second part of (16), one first deduces from above that

$$\Pr\Big(\alpha^* < \tilde{\alpha}_N < \alpha^* + \frac{3}{(\ell-2)\widehat{D'}_N} \cdot \frac{\log\log N}{\log N} + \mu \cdot \frac{\log\log N}{\log N}\Big) \underset{N\to\infty}{\longrightarrow} 1,$$

with $\mu > \frac{12}{\ell-2}$. Therefore, it exists $\nu < \frac{4}{(\ell-2)D'} + \frac{12}{\ell-2}$,

$$\Pr\Big(N^{\alpha^*} < N^{\tilde{\alpha}_N} < N^{\alpha^*} \cdot (\log N)^{\nu}\Big) \underset{N\to\infty}{\longrightarrow} 1.$$

Now, this inequality and the previous central limit theorem provides that for all $\rho > \nu/2$, and $\varepsilon > 0$,

$$\begin{aligned} \Pr\Big(\frac{N^{\frac{D'}{1+2D'}}}{(\log N)^{\rho}} \cdot \big|\tilde{D}_N - D\big| > \varepsilon\Big) &= \Pr\Big(\frac{N^{\frac{1}{2}(\widehat{\alpha}_N - \alpha^*)}}{(\log N)^{\rho}} \cdot \sqrt{\frac{N}{N^{\tilde{\alpha}_N}}}\big|\tilde{D}_N - D\big| > \varepsilon\Big) \\ &\underset{N\to\infty}{\longrightarrow} 0. \ \Box \end{aligned}$$

References

[1] P. Abry, P. Flandrin, M.S. Taqqu and D. Veitch, Self-similarity and long-range dependence through the wavelet lens, In P. Doukhan, G. Oppenheim and M.S. Taqqu editors, *Long-range Dependence: Theory and Applications*, Birkhäuser, 2003.

[2] P. Abry and D. Veitch, Wavelet analysis of long-range-dependent traffic, *IEEE Trans. on Info. Theory*, 44(1): 2-15, 1998.

[3] P. Abry and D. Veitch, A wavelet based joint estimator of the parameter of long-range dependence traffic, *IEEE Trans. on Info. Theory, Special Issue on "Multiscale Statistical Signal Analysis and its Application"*, 45: 878-897, 1999.

[4] P. Abry, D. Veitch and P. Flandrin, Long-range dependent: revisiting aggregation with wavelets, *J. Time Ser. Anal.* , Vol. 19, 253-266, 1998.

[5] J. M. Bardet, Testing for the presence of Self-similarity of Gaussian process with stationnary increments, *J. Time Ser. Anal.* , Vol. 21, 497-516, 2000.

[6] J. M. Bardet, Statistical Study of the Wavelet Analysis of Fractional Brownian Motion, *IEEE Trans. Inf. Theory*, Vol. 48, No. 4, 991-999, 2002.

[7] J. M. Bardet and P. Bertrand, Identification of the multiscale fractional Brownian motion with biomechanical applications, To appear in *J. Time Ser. Anal.*, 2006.

[8] J. M. Bardet, G. Lang, E. Moulines and P. Soulier, Wavelet estimator of long range-dependant processes, *Statist. Infer. Stochast. Processes*, Vol. 3, 85-99, 2000.

[9] P. Doukhan, G. Oppenheim and M.S. Taqqu (Editors), Theory and applications of long-range dependence, Birkhäuser, Boston, 2003.

[10] L. Giraitis, P. Robinson and A. Samarov, Rate optimal semi-parametric estimation of the memory parameter of the Gaussian time series with long range dependence, *J. Time Ser. Anal.*, 18:49-61, 1997.

[11] L. Giraitis, P. Robinson and A. Samarov, Adaptive semiparametric estimation of the memory parameter, *J. Multivariate Anal.*, 72:183-207, 2000.

[12] E. Moulines, F. Roueff and M.S. Taqqu, On the spectral density of the wavelet coefficients of long memory time series with application to the log-regression estimation of the memory parameter, To appear in *J. Time Ser. Anal.*, available in `http://arxiv.org/abs/math.ST/0512635`, 2006.

[13] E. Moulines and P. Soulier, Semiparametric spectral estimation for fractionnal processes, In P. Doukhan, G. Openheim and M.S. Taqqu editors, *Theory and applications of long-range dependence*, 251-301, Birkhäuser, Boston, 2003.

[14] P. M. Robinson, Gaussian semiparametric estimation of long range dependence, *The Annals of statistics*, 23:1630-1661, 1995.

[15] M. S. Taqqu and V. Teverovsky Semiparametric graphical estimation technics for long memeory data, *Lectures Notes in Statistics*, 115, 420-432, 1996.

[16] D. Veitch, P. Abry, and M.S. Taqqu. On the automatic selection of the onset of scaling, *Fractals*, 11, 377-390, 2003.